

- 1, TITLE: Generative Knowledge Graph Construction: A Review  
<https://aclanthology.org/2022.emnlp-main.1>  
AUTHORS: Hongbin Ye, Ningyu Zhang, Hui Chen, Huajun Chen  
HIGHLIGHT: In this study, we summarize the recent compelling progress in generative knowledge graph construction.
  
- 2, TITLE: CDConv: A Benchmark for Contradiction Detection in Chinese Conversations  
<https://aclanthology.org/2022.emnlp-main.2>  
AUTHORS: Chujie Zheng, Jinfeng Zhou, Yinhe Zheng, Libiao Peng, Zhen Guo, Wenquan Wu, Zheng-Yu Niu, Hua Wu, Minlie Huang  
HIGHLIGHT: In this work, we propose a benchmark for Contradiction Detection in Chinese Conversations, namely CDConv.
  
- 3, TITLE: Transformer Feed-Forward Layers Build Predictions by Promoting Concepts in the Vocabulary Space  
<https://aclanthology.org/2022.emnlp-main.3>  
AUTHORS: Mor Geva, Avi Caciularu, Kevin Wang, Yoav Goldberg  
HIGHLIGHT: Transformer-based language models (LMs) are at the core of modern NLP, but their internal prediction construction process is opaque and largely not understood. In this work, we make a substantial step towards unveiling this underlying prediction process, by reverse-engineering the operation of the feed-forward network (FFN) layers, one of the building blocks of transformer models.
  
- 4, TITLE: Learning to Generate Question by Asking Question: A Primal-Dual Approach with Uncommon Word Generation  
<https://aclanthology.org/2022.emnlp-main.4>  
AUTHORS: Qifan Wang, Li Yang, Xiaojun Quan, Fuli Feng, Dongfang Liu, Zenglin Xu, Sinong Wang, Hao Ma  
HIGHLIGHT: Moreover, unseen or rare word generation has not been studied in previous works. In this paper, we propose a novel approach which incorporates question generation with its dual problem, question answering, into a unified primal-dual framework.
  
- 5, TITLE: Graph-based Model Generation for Few-Shot Relation Extraction  
<https://aclanthology.org/2022.emnlp-main.5>  
AUTHORS: Wanli Li, Tiejun Qian  
HIGHLIGHT: Existing models follow a "one-for-all" scheme where one general large model performs all individual N-way-K-shot tasks in FSRE, which prevents the model from achieving the optimal point on each task. In view of this, we propose a model generation framework that consists of one general model for all tasks and many tiny task-specific models for each individual task.
  
- 6, TITLE: Backdoor Attacks in Federated Learning by Rare Embeddings and Gradient Ensembling  
<https://aclanthology.org/2022.emnlp-main.6>  
AUTHORS: Ki Yoon Yoo, Nojun Kwak  
HIGHLIGHT: This paper investigates the feasibility of model poisoning for backdoor attacks through rare word embeddings of NLP models.
  
- 7, TITLE: Generating Natural Language Proofs with Verifier-Guided Search  
<https://aclanthology.org/2022.emnlp-main.7>  
AUTHORS: Kaiyu Yang, Jia Deng, Danqi Chen  
HIGHLIGHT: In this work, we focus on proof generation: Given a hypothesis and a set of supporting facts, the model generates a proof tree indicating how to derive the hypothesis from supporting facts.
  
- 8, TITLE: Toward Unifying Text Segmentation and Long Document Summarization  
<https://aclanthology.org/2022.emnlp-main.8>  
AUTHORS: Sangwoo Cho, Kaiqiang Song, Xiaoyang Wang, Fei Liu, Dong Yu  
HIGHLIGHT: In this paper, we explore the role that section segmentation plays in extractive summarization of written and spoken documents.
  
- 9, TITLE: The Geometry of Multilingual Language Model Representations  
<https://aclanthology.org/2022.emnlp-main.9>  
AUTHORS: Tyler Chang, Zhuowen Tu, Benjamin Bergen  
HIGHLIGHT: We assess how multilingual language models maintain a shared multilingual representation space while still encoding language-sensitive information in each language.
  
- 10, TITLE: Improving Complex Knowledge Base Question Answering via Question-to-Action and Question-to-Question Alignment  
<https://aclanthology.org/2022.emnlp-main.10>  
AUTHORS: Yechun Tang, Xiaoxia Cheng, Weiming Lu  
HIGHLIGHT: However, there is a significant semantic and structural gap between natural language and action sequences, which makes this conversion difficult. In this paper, we introduce an alignment-enhanced complex question answering framework, called ALCQA, which mitigates this gap through question-to-action alignment and question-to-question alignment.
  
- 11, TITLE: PAIR: Prompt-Aware margin Ranking for Counselor Reflection Scoring in Motivational Interviewing  
<https://aclanthology.org/2022.emnlp-main.11>  
AUTHORS: Do June Min, Verónica Pérez-Rosas, Kenneth Resnicow, Rada Mihalcea  
HIGHLIGHT: In this paper, we propose a system for the analysis of counselor reflections.

- 12, TITLE: Co-guiding Net: Achieving Mutual Guidances between Multiple Intent Detection and Slot Filling via Heterogeneous Semantics-Label Graphs  
<https://aclanthology.org/2022.emnlp-main.12>  
AUTHORS: Bowen Xing, Ivor Tsang  
HIGHLIGHT: In this paper, we propose a novel model termed Co-guiding Net, which implements a two-stage framework achieving the mutual guidances between the two tasks.
- 13, TITLE: The Importance of Being Parameters: An Intra-Distillation Method for Serious Gains  
<https://aclanthology.org/2022.emnlp-main.13>  
AUTHORS: Haoran Xu, Philipp Koehn, Kenton Murray  
HIGHLIGHT: We propose a general task-agnostic method, namely intra-distillation, appended to the regular training loss to balance parameter sensitivity.
- 14, TITLE: Interpreting Language Models with Contrastive Explanations  
<https://aclanthology.org/2022.emnlp-main.14>  
AUTHORS: Kayo Yin, Graham Neubig  
HIGHLIGHT: To disentangle the different decisions in language modeling, we focus on explaining language models contrastively: we look for salient input tokens that explain why the model predicted one token instead of another.
- 15, TITLE: RankGen: Improving Text Generation with Large Ranking Models  
<https://aclanthology.org/2022.emnlp-main.15>  
AUTHORS: Kalpesh Krishna, Yapei Chang, John Wieting, Mohit Iyyer  
HIGHLIGHT: Given an input sequence (or prefix), modern language models often assign high probabilities to output sequences that are repetitive, incoherent, or irrelevant to the prefix; as such, model-generated text also contains such artifacts. To address these issues we present RankGen, a 1.2B parameter encoder model for English that scores model generations given a prefix.
- 16, TITLE: Learning a Grammar Inducer from Massive Uncurated Instructional Videos  
<https://aclanthology.org/2022.emnlp-main.16>  
AUTHORS: Songyang Zhang, Linfeng Song, Lifeng Jin, Haitao Mi, Kun Xu, Dong Yu, Jiebo Luo  
HIGHLIGHT: While previous work focuses on building systems for inducing grammars on text that are well-aligned with video content, we investigate the scenario, in which text and video are only in loose correspondence.
- 17, TITLE: Normalized Contrastive Learning for Text-Video Retrieval  
<https://aclanthology.org/2022.emnlp-main.17>  
AUTHORS: Yookoon Park, Mahmoud Azab, Seungwhan Moon, Bo Xiong, Florian Metze, Gourab Kundu, Kirmani Ahmed  
HIGHLIGHT: Specifically, we show that many test instances are either over- or under-represented during retrieval, significantly hurting the retrieval performance. To address this problem, we propose Normalized Contrastive Learning (NCL) which utilizes the Sinkhorn-Knopp algorithm to compute the instance-wise biases that properly normalize the sum retrieval probabilities of each instance so that every text and video instance is fairly represented during cross-modal retrieval.
- 18, TITLE: Estimating Soft Labels for Out-of-Domain Intent Detection  
<https://aclanthology.org/2022.emnlp-main.18>  
AUTHORS: Hao Lang, Yinhe Zheng, Jian Sun, Fei Huang, Luo Si, Yongbin Li  
HIGHLIGHT: In this paper, we propose an adaptive soft pseudo labeling (ASoul) method that can estimate soft labels for pseudo OOD samples when training OOD detectors.
- 19, TITLE: Multi-VQG: Generating Engaging Questions for Multiple Images  
<https://aclanthology.org/2022.emnlp-main.19>  
AUTHORS: Min-Hsuan Yeh, Vincent Chen, Ting-Hao Huang, Lun-Wei Ku  
HIGHLIGHT: In this paper, we propose generating engaging questions from multiple images.
- 20, TITLE: Tomayto, Tomahto. Beyond Token-level Answer Equivalence for Question Answering Evaluation  
<https://aclanthology.org/2022.emnlp-main.20>  
AUTHORS: Jannis Bulian, Christian Buck, Wojciech Gajewski, Benjamin B?rschinger, Tal Schuster  
HIGHLIGHT: In this paper, we present the first systematic conceptual and data-driven analysis to examine the shortcomings of token-level equivalence measures.
- 21, TITLE: Non-Parametric Domain Adaptation for End-to-End Speech Translation  
<https://aclanthology.org/2022.emnlp-main.21>  
AUTHORS: Yichao Du, Weizhi Wang, Zhirui Zhang, Boxing Chen, Tong Xu, Jun Xie, Enhong Chen  
HIGHLIGHT: In this paper, we propose a novel non-parametric method that leverages in-domain text translation corpus to achieve domain adaptation for E2E-ST systems.
- 22, TITLE: Prompting for Multimodal Hateful Meme Classification  
<https://aclanthology.org/2022.emnlp-main.22>  
AUTHORS: Rui Cao, Roy Ka-Wei Lee, Wen-Haw Chong, Jing Jiang  
HIGHLIGHT: However, there is no known explicit external knowledge base that could provide such hate speech contextual information. To address this gap, we propose PromptHate, a simple yet effective prompt-based model that prompts pre-trained language models (PLMs) for hateful meme classification.

- 23, TITLE: Certified Error Control of Candidate Set Pruning for Two-Stage Relevance Ranking  
<https://aclanthology.org/2022.emnlp-main.23>  
AUTHORS: Minghan Li, Xinyu Zhang, Ji Xin, Hongyang Zhang, Jimmy Lin  
HIGHLIGHT: In this paper, we propose the concept of certified error control of candidate set pruning for relevance ranking, which means that the test error after pruning is guaranteed to be controlled under a user-specified threshold with high probability.
- 24, TITLE: Linearizing Transformer with Key-Value Memory  
<https://aclanthology.org/2022.emnlp-main.24>  
AUTHORS: Yizhe Zhang, Deng Cai  
HIGHLIGHT: We propose Memsizer, an approach towards closing the performance gap while improving the efficiency even with short generation.
- 25, TITLE: Robustness of Fusion-based Multimodal Classifiers to Cross-Modal Content Dilutions  
<https://aclanthology.org/2022.emnlp-main.25>  
AUTHORS: Gaurav Verma, Vishwa Vinay, Ryan Rossi, Srijan Kumar  
HIGHLIGHT: In this work, we investigate the robustness of multimodal classifiers to cross-modal dilutions ? a plausible variation.
- 26, TITLE: Translation between Molecules and Natural Language  
<https://aclanthology.org/2022.emnlp-main.26>  
AUTHORS: Carl Edwards, Tuan Lai, Kevin Ros, Garrett Honke, Kyunghyun Cho, Heng Ji  
HIGHLIGHT: We present MolT5 - a self-supervised learning framework for pretraining models on a vast amount of unlabeled natural language text and molecule strings.
- 27, TITLE: What Makes Instruction Learning Hard? An Investigation and a New Challenge in a Synthetic Environment  
<https://aclanthology.org/2022.emnlp-main.27>  
AUTHORS: Matthew Finlayson, Kyle Richardson, Ashish Sabharwal, Peter Clark  
HIGHLIGHT: We thus propose Hard RegSet as a challenging instruction learning dataset, and a controlled environment for studying instruction learning.
- 28, TITLE: Sentence-Incremental Neural Coreference Resolution  
<https://aclanthology.org/2022.emnlp-main.28>  
AUTHORS: Matt Grenander, Shay B. Cohen, Mark Steedman  
HIGHLIGHT: We propose a sentence-incremental neural coreference resolution system which incrementally builds clusters after marking mention boundaries in a shift-reduce method.
- 29, TITLE: SNaC: Coherence Error Detection for Narrative Summarization  
<https://aclanthology.org/2022.emnlp-main.29>  
AUTHORS: Tanya Goyal, Junyi Jessy Li, Greg Durrett  
HIGHLIGHT: In this work, we introduce SNaC, a narrative coherence evaluation framework for fine-grained annotations of long summaries.
- 30, TITLE: HydraSum: Disentangling Style Features in Text Summarization with Multi-Decoder Models  
<https://aclanthology.org/2022.emnlp-main.30>  
AUTHORS: Tanya Goyal, Nazneen Rajani, Wenhao Liu, Wojciech Kryscinski  
HIGHLIGHT: However, these are implicitly encoded within model parameters and specific styles cannot be enforced. To address this, we introduce HydraSum, a new summarization architecture that extends the single decoder framework of current models to a mixture-of-experts version with multiple decoders.
- 31, TITLE: A Good Neighbor, A Found Treasure: Mining Treasured Neighbors for Knowledge Graph Entity Typing  
<https://aclanthology.org/2022.emnlp-main.31>  
AUTHORS: Zhuoran Jin, Pengfei Cao, Yubo Chen, Kang Liu, Jun Zhao  
HIGHLIGHT: Besides, we also observe that there are co-occurrence relations between types, which is very helpful to alleviate false-negative problem. In this paper, we propose a novel method called Mining Treasured Neighbors (MiNer) to make use of these two characteristics.
- 32, TITLE: Guiding Neural Entity Alignment with Compatibility  
<https://aclanthology.org/2022.emnlp-main.32>  
AUTHORS: Bing Liu, Harrison Scells, Wen Hua, Guido Zuccon, Genghong Zhao, Xia Zhang  
HIGHLIGHT: In this work, we argue that different entities within one KG should have compatible counterparts in the other KG due to the potential dependencies among the entities.
- 33, TITLE: InstructDial: Improving Zero and Few-shot Generalization in Dialogue through Instruction Tuning  
<https://aclanthology.org/2022.emnlp-main.33>  
AUTHORS: Prakhar Gupta, Cathy Jiao, Yi-Ting Yeh, Shikib Mehri, Maxine Eskenazi, Jeffrey Bigham  
HIGHLIGHT: We introduce InstructDial, an instruction tuning framework for dialogue, which consists of a repository of 48 diverse dialogue tasks in a unified text-to-text format created from 59 openly available dialogue datasets.
- 34, TITLE: Unsupervised Boundary-Aware Language Model Pretraining for Chinese Sequence Labeling  
<https://aclanthology.org/2022.emnlp-main.34>

**AUTHORS:** Peijie Jiang, Dingkun Long, Yanzhao Zhang, Pengjun Xie, Meishan Zhang, Min Zhang  
**HIGHLIGHT:** In this work, we suggest unsupervised statistical boundary information instead, and propose an architecture to encode the information directly into pre-trained language models, resulting in Boundary-Aware BERT (BABERT).

35, **TITLE:** RetroMAE: Pre-Training Retrieval-oriented Language Models Via Masked Auto-Encoder  
<https://aclanthology.org/2022.emnlp-main.35>  
**AUTHORS:** Shitao Xiao, Zheng Liu, Yingxia Shao, Zhao Cao  
**HIGHLIGHT:** In this paper, we propose RetroMAE, a new retrieval oriented pre-training paradigm based on Masked Auto-Encoder (MAE).

36, **TITLE:** Aligning Recommendation and Conversation via Dual Imitation  
<https://aclanthology.org/2022.emnlp-main.36>  
**AUTHORS:** Jinfeng Zhou, Bo Wang, Minlie Huang, Dongming Zhao, Kun Huang, Ruifang He, Yuexian Hou  
**HIGHLIGHT:** However, existing conversational recommendation systems (CRS) ignore the advantage of user interest shift in connecting recommendation and conversation, which leads to an ineffective loose coupling structure of CRS. To address this issue, by modeling the recommendation actions as recommendation paths in a knowledge graph (KG), we propose DICR (Dual Imitation for Conversational Recommendation), which designs a dual imitation to explicitly align the recommendation paths and user interest shift paths in a recommendation module and a conversation module, respectively.

37, **TITLE:** QRelScore: Better Evaluating Generated Questions with Deeper Understanding of Context-aware Relevance  
<https://aclanthology.org/2022.emnlp-main.37>  
**AUTHORS:** Xiaoqiang Wang, Bang Liu, Siliang Tang, Lingfei Wu  
**HIGHLIGHT:** In this paper, we propose QRelScore, a context-aware Relevance evaluation metric for Question Generation.

38, **TITLE:** Abstract Visual Reasoning with Tangram Shapes  
<https://aclanthology.org/2022.emnlp-main.38>  
**AUTHORS:** Anya Ji, Noriyuki Kojima, Noah Rush, Alane Suhr, Wai Keen Vong, Robert Hawkins, Yoav Artzi  
**HIGHLIGHT:** We introduce KiloGram, a resource for studying abstract visual reasoning in humans and machines.

39, **TITLE:** UnifiedSKG: Unifying and Multi-Tasking Structured Knowledge Grounding with Text-to-Text Language Models  
<https://aclanthology.org/2022.emnlp-main.39>  
**AUTHORS:** Tianbao Xie, Chen Henry Wu, Peng Shi, Ruiqi Zhong, Torsten Scholak, Michihiro Yasunaga, Chien-Sheng Wu, Ming Zhong, Pengcheng Yin, Sida I. Wang, Victor Zhong, Bailin Wang, Chengzu Li, Connor Boyle, Ansong Ni, Ziyu Yao, Dragomir Radev, Caiming Xiong, Lingpeng Kong, Rui Zhang, Noah A. Smith, Luke Zettlemoyer, Tao Yu  
**HIGHLIGHT:** Since the inputs and outputs of SKG tasks are heterogeneous, they have been studied separately by different communities, which limits systematic and compatible research on SKG. In this paper, we overcome this limitation by proposing the UnifiedSKG framework, which unifies 21 SKG tasks into a text-to-text format, aiming to promote systematic SKG research, instead of being exclusive to a single task, domain, or dataset.

40, **TITLE:** Balanced Adversarial Training: Balancing Tradeoffs between Fickleness and Obstnacy in NLP Models  
<https://aclanthology.org/2022.emnlp-main.40>  
**AUTHORS:** Hannah Chen, Yangfeng Ji, David Evans  
**HIGHLIGHT:** We show that standard adversarial training methods focused on reducing vulnerability to fickle adversarial examples may make a model more vulnerable to obstinate adversarial examples, with experiments for both natural language inference and paraphrase identification tasks. To counter this phenomenon, we introduce Balanced Adversarial Training, which incorporates contrastive learning to increase robustness against both fickle and obstinate adversarial examples.

41, **TITLE:** When Can Transformers Ground and Compose: Insights from Compositional Generalization Benchmarks  
<https://aclanthology.org/2022.emnlp-main.41>  
**AUTHORS:** Ankur Sikarwar, Arkil Patel, Navin Goyal  
**HIGHLIGHT:** In this work, we present a simple transformer-based model that outperforms specialized architectures on ReaSCAN and a modified version (Qiu et al., 2021) of gSCAN (Ruis et al., 2020).

42, **TITLE:** Generative Language Models for Paragraph-Level Question Generation  
<https://aclanthology.org/2022.emnlp-main.42>  
**AUTHORS:** Asahi Ushio, Fernando Alva-Manchego, Jose Camacho-Collados  
**HIGHLIGHT:** However, it is difficult to measure advances in QG research since there are no standardized resources that allow a uniform comparison among approaches. In this paper, we introduce QG-Bench, a multilingual and multidomain benchmark for QG that unifies existing question answering datasets by converting them to a standard QG setting.

43, **TITLE:** A Unified Encoder-Decoder Framework with Entity Memory  
<https://aclanthology.org/2022.emnlp-main.43>  
**AUTHORS:** Zhihan Zhang, Wenhao Yu, Chenguang Zhu, Meng Jiang  
**HIGHLIGHT:** In this work, we propose an encoder-decoder framework with an entity memory, namely EDMem.

44, **TITLE:** Segmenting Numerical Substitution Ciphers  
<https://aclanthology.org/2022.emnlp-main.44>  
**AUTHORS:** Nada Aldarrab, Jonathan May

- HIGHLIGHT:** In this work, we propose the first automatic methods to segment those ciphers using Byte Pair Encoding (BPE) and unigram language models.
- 45, **TITLE:** Crossmodal-3600: A Massively Multilingual Multimodal Evaluation Dataset  
<https://aclanthology.org/2022.emnlp-main.45>  
**AUTHORS:** Ashish V. Thapliyal, Jordi Pont Tuset, Xi Chen, Radu Soricut  
**HIGHLIGHT:** In this paper we present the Crossmodal-3600 dataset (XM3600 in short), a geographically diverse set of 3600 images annotated with human-generated reference captions in 36 languages.
- 46, **TITLE:** ReSel: N-ary Relation Extraction from Scientific Text and Tables by Learning to Retrieve and Select  
<https://aclanthology.org/2022.emnlp-main.46>  
**AUTHORS:** Yuchen Zhuang, Yinghao Li, Junyang Zhang, Yue Yu, Yingjun Mou, Xiang Chen, Le Song, Chao Zhang  
**HIGHLIGHT:** We study the problem of extracting N-ary relation tuples from scientific articles.
- 47, **TITLE:** GammaE: Gamma Embeddings for Logical Queries on Knowledge Graphs  
<https://aclanthology.org/2022.emnlp-main.47>  
**AUTHORS:** Dong Yang, Peijun Qing, Yang Li, Haonan Lu, Xiaodong Ling  
**HIGHLIGHT:** An additional limitation is that the union operator is non-closure, which undermines the model to handle a series of union operators. To address these problems, we propose a novel probabilistic embedding model, namely Gamma Embeddings (GammaE), for encoding entities and queries to answer different types of FOL queries on KGs.
- 48, **TITLE:** Reasoning Like Program Executors  
<https://aclanthology.org/2022.emnlp-main.48>  
**AUTHORS:** Xinyu Pi, Qian Liu, Bei Chen, Morteza Ziyadi, Zeqi Lin, Qiang Fu, Yan Gao, Jian-Guang Lou, Weizhu Chen  
**HIGHLIGHT:** In this paper, we showcase two simple instances POET-Math and POET-Logic, in addition to a complex instance, POET-SQL.
- 49, **TITLE:** SEM-F1: an Automatic Way for Semantic Evaluation of Multi-Narrative Overlap Summaries at Scale  
<https://aclanthology.org/2022.emnlp-main.49>  
**AUTHORS:** Naman Bansal, Mousumi Akter, Shubhra Kanti Karmaker Santu  
**HIGHLIGHT:** In this paper, we exclusively focus on the automated evaluation of the SOS task using the benchmark dataset.
- 50, **TITLE:** Inducer-tuning: Connecting Prefix-tuning and Adapter-tuning  
<https://aclanthology.org/2022.emnlp-main.50>  
**AUTHORS:** Yifan Chen, Devamanyu Hazarika, Mahdi Namazifar, Yang Liu, Di Jin, Dilek Hakkani-Tur  
**HIGHLIGHT:** In this paper, we propose to understand and further develop prefix-tuning through the kernel lens.
- 51, **TITLE:** DocInfer: Document-level Natural Language Inference using Optimal Evidence Selection  
<https://aclanthology.org/2022.emnlp-main.51>  
**AUTHORS:** Puneet Mathur, Gautam Kunapuli, Riyaz Bhat, Manish Shrivastava, Dinesh Manocha, Maneesh Singh  
**HIGHLIGHT:** We present DocInfer - a novel, end-to-end Document-level Natural Language Inference model that builds a hierarchical document graph enriched through inter-sentence relations (topical, entity-based, concept-based), performs paragraph pruning using the novel SubGraph Pooling layer, followed by optimal evidence selection based on REINFORCE algorithm to identify the most important context sentences for a given hypothesis.
- 52, **TITLE:** LightEA: A Scalable, Robust, and Interpretable Entity Alignment Framework via Three-view Label Propagation  
<https://aclanthology.org/2022.emnlp-main.52>  
**AUTHORS:** Xin Mao, Wenting Wang, Yuanbin Wu, Man Lan  
**HIGHLIGHT:** In this paper, we argue that existing complex EA methods inevitably inherit the inborn defects from their neural network lineage: poor interpretability and weak scalability.
- 53, **TITLE:** Metric-guided Distillation: Distilling Knowledge from the Metric to Ranker and Retriever for Generative Commonsense Reasoning  
<https://aclanthology.org/2022.emnlp-main.53>  
**AUTHORS:** Xingwei He, Yeyun Gong, A-Long Jin, Weizhen Qi, Hang Zhang, Jian Jiao, Bartuer Zhou, Biao Cheng, Sm Yiu, Nan Duan  
**HIGHLIGHT:** Another problem is that re-ranking is very expensive, but only using retrievers will seriously degrade the performance of their generation models. To solve these problems, we propose the metric distillation rule to distill knowledge from the metric (e.g., BLEU) to the ranker.
- 54, **TITLE:** Efficient Document Retrieval by End-to-End Refining and Quantizing BERT Embedding with Contrastive Product Quantization  
<https://aclanthology.org/2022.emnlp-main.54>  
**AUTHORS:** Zexuan Qiu, Qinliang Su, Jianxing Yu, Shijing Si  
**HIGHLIGHT:** Furthermore, the Hamming distance can only be equal to one of several integer values, significantly limiting its representational ability for document distances. To address these issues, in this paper, we propose to leverage BERT embeddings to perform efficient retrieval based on the product quantization technique, which will assign for every document a real-valued codeword from the codebook, instead of a binary code as in semantic hashing.

- 55, TITLE: Curriculum Knowledge Distillation for Emoji-supervised Cross-lingual Sentiment Analysis  
<https://aclanthology.org/2022.emnlp-main.55>  
AUTHORS: Jianyang Zhang, Tao Liang, Mingyang Wan, Guowu Yang, Fengmao Lv  
HIGHLIGHT: In this work, based on the intuitive assumption that the relationships between emojis and sentiments are consistent across different languages, we investigate transferring sentiment knowledge across languages with the help of emojis.
- 56, TITLE: Correctable-DST: Mitigating Historical Context Mismatch between Training and Inference for Improved Dialogue State Tracking  
<https://aclanthology.org/2022.emnlp-main.56>  
AUTHORS: Hongyan Xie, Haoxiang Su, Shuangyong Song, Hao Huang, Bo Zou, Kun Deng, Jianghua Lin, Zhihui Zhang, Xiaodong He  
HIGHLIGHT: However, only the previously predicted dialogue state can be used in inference. This discrepancy might lead to error propagation, i.e., mistakes made by the model in the current turn are likely to be carried over to the following turns. To solve this problem, we propose Correctable Dialogue State Tracking (Correctable-DST).
- 57, TITLE: DropMix: A Textual Data Augmentation Combining Dropout with Mixup  
<https://aclanthology.org/2022.emnlp-main.57>  
AUTHORS: Fanshuang Kong, Richong Zhang, Xiaohui Guo, Samuel Mensah, Yongyi Mao  
HIGHLIGHT: In this paper, we argue that the property is essential to overcome overfitting in text learning.
- 58, TITLE: Cross-document Event Coreference Search: Task, Dataset and Modeling  
<https://aclanthology.org/2022.emnlp-main.58>  
AUTHORS: Alon Eirew, Avi Caciularu, Ido Dagan  
HIGHLIGHT: We propose an appealing, and often more applicable, complementary set up for the task ? Cross-document Coreference Search, focusing in this paper on event coreference.
- 59, TITLE: VIRT: Improving Representation-based Text Matching via Virtual Interaction  
<https://aclanthology.org/2022.emnlp-main.59>  
AUTHORS: Dan Li, Yang Yang, Hongyin Tang, Jiahao Liu, Qifan Wang, Jingang Wang, Tong Xu, Wei Wu, Enhong Chen  
HIGHLIGHT: However, these models suffer from severe performance degradation due to the lack of interactions between the pair of texts. To remedy this, we propose a Virtual InteRacTion mechanism (VIRT) for improving representation-based text matching while maintaining its efficiency.
- 60, TITLE: MAVEN-ERE: A Unified Large-scale Dataset for Event Coreference, Temporal, Causal, and Subevent Relation Extraction  
<https://aclanthology.org/2022.emnlp-main.60>  
AUTHORS: Xiaozhi Wang, Yulin Chen, Ning Ding, Hao Peng, Zimu Wang, Yankai Lin, Xu Han, Lei Hou, Juanzi Li, Zhiyuan Liu, Peng Li, Jie Zhou  
HIGHLIGHT: Different types of event relations naturally interact with each other, but existing datasets only cover limited relation types at once, which prevents models from taking full advantage of relation interactions. To address these issues, we construct a unified large-scale human-annotated ERE dataset MAVEN-ERE with improved annotation schemes.
- 61, TITLE: Entity Extraction in Low Resource Domains with Selective Pre-training of Large Language Models  
<https://aclanthology.org/2022.emnlp-main.61>  
AUTHORS: Aniruddha Mahapatra, Sharmila Reddy Nangi, Aparna Garimella, Anandhavelu N  
HIGHLIGHT: In this paper, we introduce effective ways to select data from unlabeled corpora of target domains for language model pretraining to improve the performances in target entity extraction tasks.
- 62, TITLE: How Large Language Models are Transforming Machine-Paraphrase Plagiarism  
<https://aclanthology.org/2022.emnlp-main.62>  
AUTHORS: Jan Philip Wahle, Terry Ruas, Frederic Kirstein, Bela Gipp  
HIGHLIGHT: This work explores T5 and GPT3 for machine-paraphrase generation on scientific articles from arXiv, student theses, and Wikipedia.
- 63, TITLE: M2D2: A Massively Multi-Domain Language Modeling Dataset  
<https://aclanthology.org/2022.emnlp-main.63>  
AUTHORS: Machel Reid, Victor Zhong, Suchin Gururangan, Luke Zettlemoyer  
HIGHLIGHT: We present M2D2, a fine-grained, massively multi-domain corpus for studying domain adaptation in language models (LMs).
- 64, TITLE: ?Will You Find These Shortcuts?? A Protocol for Evaluating the Faithfulness of Input Salience Methods for Text Classification  
<https://aclanthology.org/2022.emnlp-main.64>  
AUTHORS: Jasmijn Bastings, Sebastian Ebert, Polina Zablotskaia, Anders Sandholm, Katja Filippova  
HIGHLIGHT: Existing work on faithfulness evaluation is not conclusive and does not provide a clear answer as to how different methods are to be compared. Focusing on text classification and the model debugging scenario, our main contribution is a protocol for faithfulness evaluation that makes use of partially synthetic data to obtain ground truth for feature importance ranking.
- 65, TITLE: Information-Transport-based Policy for Simultaneous Translation  
<https://aclanthology.org/2022.emnlp-main.65>

AUTHORS: Shaolei Zhang, Yang Feng  
HIGHLIGHT: In this paper, we treat the translation as information transport from source to target and accordingly propose an Information-Transport-based Simultaneous Translation (ITST).

66, TITLE: Learning to Adapt to Low-Resource Paraphrase Generation

<https://aclanthology.org/2022.emnlp-main.66>

AUTHORS: Zhigen Li, Yammeng Wang, Rizhao Fan, Ye Wang, Jianfeng Li, Shaojun Wang

HIGHLIGHT: However, transferring a paraphrasing model to another domain encounters the problem of domain shifting especially when the data is sparse. At the same time, widely using large pre-trained language models (PLMs) faces the overfitting problem when training on scarce labeled data. To mitigate these two issues, we propose, LAPA, an effective adapter for PLMs optimized by meta-learning.

67, TITLE: A Distributional Lens for Multi-Aspect Controllable Text Generation

<https://aclanthology.org/2022.emnlp-main.67>

AUTHORS: Yuxuan Gu, Xiaocheng Feng, Sicheng Ma, Lingyuan Zhang, Heng Gong, Bing Qin

HIGHLIGHT: Existing methods achieve complex multi-aspect control by fusing multiple controllers learned from single-aspect, but suffer from attribute degeneration caused by the mutual interference of these controllers. To address this, we provide observations on attribute fusion from a distributional perspective and propose to directly search for the intersection areas of multiple attribute distributions as their combination for generation.

68, TITLE: ELMER: A Non-Autoregressive Pre-trained Language Model for Efficient and Effective Text Generation

<https://aclanthology.org/2022.emnlp-main.68>

AUTHORS: Junyi Li, Tianyi Tang, Wayne Xin Zhao, Jian-Yun Nie, Ji-Rong Wen

HIGHLIGHT: In this paper, we propose ELMER: an efficient and effective PLM for NAR text generation to explicitly model the token dependency during NAR generation.

69, TITLE: Multilingual Relation Classification via Efficient and Effective Prompting

<https://aclanthology.org/2022.emnlp-main.69>

AUTHORS: Yuxuan Chen, David Harbecke, Leonhard Hennig

HIGHLIGHT: In this paper, we present the first work on prompt-based multilingual relation classification (RC), by introducing an efficient and effective method that constructs prompts from relation triples and involves only minimal translation for the class labels.

70, TITLE: Topic-Regularized Authorship Representation Learning

<https://aclanthology.org/2022.emnlp-main.70>

AUTHORS: Jitkapat Sawatphol, Nonthakit Chaiwong, Can Udomcharoenchaikit, Sarana Nutanong

HIGHLIGHT: To handle a large number of unseen authors and topics, we propose Authorship Representation Regularization (ARR), a distillation framework that creates authorship representation with reduced reliance on topic-specific information.

71, TITLE: Fine-grained Contrastive Learning for Relation Extraction

<https://aclanthology.org/2022.emnlp-main.71>

AUTHORS: William Hogan, Jiacheng Li, Jingbo Shang

HIGHLIGHT: In this paper, we propose fine-grained contrastive learning (FineCL) for RE, which leverages fine-grained information about which silver labels are and are not noisy to improve the quality of learned relationship representations for RE.

72, TITLE: Curriculum Prompt Learning with Self-Training for Abstractive Dialogue Summarization

<https://aclanthology.org/2022.emnlp-main.72>

AUTHORS: Changqun Li, Linlin Wang, Xin Lin, Gerard de Melo, Liang He

HIGHLIGHT: Succinctly summarizing dialogue is a task of growing interest, but inherent challenges, such as insufficient training data and low information density impede our ability to train abstractive models. In this work, we propose a novel curriculum-based prompt learning method with self-training to address these problems.

73, TITLE: Zero-Shot Text Classification with Self-Training

<https://aclanthology.org/2022.emnlp-main.73>

AUTHORS: Ariel Gera, Alon Halfon, Eyal Shnarch, Yotam Perlitz, Liat Ein-Dor, Noam Slonim

HIGHLIGHT: However, the fact that such models are unfamiliar with the target task can lead to instability and performance issues. We propose a plug-and-play method to bridge this gap using a simple self-training approach, requiring only the class names along with an unlabeled dataset, and without the need for domain expertise or trial and error.

74, TITLE: Deconfounding Legal Judgment Prediction for European Court of Human Rights Cases Towards Better Alignment with Experts

<https://aclanthology.org/2022.emnlp-main.74>

AUTHORS: T.y.s.s Santosh, Shanshan Xu, Oana Ichim, Matthias Grabmair

HIGHLIGHT: This work demonstrates that Legal Judgement Prediction systems without expert-informed adjustments can be vulnerable to shallow, distracting surface signals that arise from corpus construction, case distribution, and confounding factors.

75, TITLE: SQuALITY: Building a Long-Document Summarization Dataset the Hard Way

<https://aclanthology.org/2022.emnlp-main.75>

AUTHORS: Alex Wang, Richard Yuanzhe Pang, Angelica Chen, Jason Phang, Samuel R. Bowman

**HIGHLIGHT:** In this work, we turn to a slower but more straightforward approach to developing summarization benchmark data: We hire highly-qualified contractors to read stories and write original summaries from scratch.

76, **TITLE:** MetaASSIST: Robust Dialogue State Tracking with Meta Learning

<https://aclanthology.org/2022.emnlp-main.76>

**AUTHORS:** Fanghua Ye, Xi Wang, Jie Huang, Shenghui Li, Samuel Stern, Emine Yilmaz

**HIGHLIGHT:** Specifically, we propose three schemes with varying degrees of flexibility, ranging from slot-wise to both slot-wise and instance-wise, to convert the weighting parameter into learnable functions.

77, **TITLE:** Multilingual Machine Translation with Hyper-Adapters

<https://aclanthology.org/2022.emnlp-main.77>

**AUTHORS:** Christos Baziotis, Mikel Artetxe, James Cross, Shruti Bhosale

**HIGHLIGHT:** However, adapters of related languages are unable to transfer information, and their total number of parameters becomes prohibitively expensive as the number of languages grows. In this work, we overcome these drawbacks using hyper-adapters ? hyper-networks that generate adapters from language and layer embeddings.

78, **TITLE:** Z-LaVI: Zero-Shot Language Solver Fueled by Visual Imagination

<https://aclanthology.org/2022.emnlp-main.78>

**AUTHORS:** Yue Yang, Wenlin Yao, Hongming Zhang, Xiaoyang Wang, Dong Yu, Jianshu Chen

**HIGHLIGHT:** However, they generally suffer from reporting bias, the phenomenon describing the lack of explicit commonsense knowledge in written text, e.g., ?an orange is orange?. To overcome this limitation, we develop a novel approach, Z-LaVI, to endow language models with visual imagination capabilities.

79, **TITLE:** Using Commonsense Knowledge to Answer Why-Questions

<https://aclanthology.org/2022.emnlp-main.79>

**AUTHORS:** Yash Kumar Lal, Niket Tandon, Tanvi Aggarwal, Horace Liu, Nathanael Chambers, Raymond Mooney, Niranjan Balasubramanian

**HIGHLIGHT:** What aspects can be made accessible via external commonsense resources? We study these questions in the context of answering questions in the TellMeWhy dataset using COMET as a source of relevant commonsense relations.

80, **TITLE:** Affective Idiosyncratic Responses to Music

<https://aclanthology.org/2022.emnlp-main.80>

**AUTHORS:** Sky CH-Wang, Evan Li, Oliver Li, Smaranda Muresan, Zhou Yu

**HIGHLIGHT:** Despite consensus that idiosyncratic factors play a key role in regulating how listeners emotionally respond to music, precisely measuring the marginal effects of these variables has proved challenging. To address this gap, we develop computational methods to measure affective responses to music from over 403M listener comments on a Chinese social music platform.

81, **TITLE:** Successive Prompting for Decomposing Complex Questions

<https://aclanthology.org/2022.emnlp-main.81>

**AUTHORS:** Dheeru Dua, Shivanshu Gupta, Sameer Singh, Matt Gardner

**HIGHLIGHT:** We introduce a way to generate synthetic dataset which can be used to bootstrap model?s ability to decompose and answer intermediate questions.

82, **TITLE:** Maieutic Prompting: Logically Consistent Reasoning with Recursive Explanations

<https://aclanthology.org/2022.emnlp-main.82>

**AUTHORS:** Jaehun Jung, Lianhui Qin, Sean Welleck, Faeze Brahman, Chandra Bhagavatula, Ronan Le Bras, Yejin Choi

**HIGHLIGHT:** In this work, we develop Maieutic Prompting, which aims to infer a correct answer to a question even from the unreliable generations of LM.

83, **TITLE:** DANLI: Deliberative Agent for Following Natural Language Instructions

<https://aclanthology.org/2022.emnlp-main.83>

**AUTHORS:** Yichi Zhang, Jianing Yang, Jiayi Pan, Shane Storks, Nikhil Devraj, Ziqiao Ma, Keunwoo Yu, Yuwei Bao, Joyce Chai

**HIGHLIGHT:** These reactive agents are insufficient for long-horizon complex tasks. To address this limitation, we propose a neuro-symbolic deliberative agent that, while following language instructions, proactively applies reasoning and planning based on its neural and symbolic representations acquired from past experience (e.g., natural language and egocentric vision).

84, **TITLE:** Tracing Semantic Variation in Slang

<https://aclanthology.org/2022.emnlp-main.84>

**AUTHORS:** Zhewei Sun, Yang Xu

**HIGHLIGHT:** We explore these theories using computational models and test them against historical slang dictionary entries, with a focus on characterizing regularity in the geographical variation of slang usages attested in the US and the UK over the past two centuries.

85, **TITLE:** Fine-grained Category Discovery under Coarse-grained supervision with Hierarchical Weighted Self-contrastive Learning

<https://aclanthology.org/2022.emnlp-main.85>

**AUTHORS:** Wenbin An, Feng Tian, Ping Chen, Siliang Tang, Qinghua Zheng, QianYing Wang



**HIGHLIGHT:** Considering most current methods cannot transfer knowledge from coarse-grained level to fine-grained level, we propose a hierarchical weighted self-contrastive network by building a novel weighted self-contrastive module and combining it with supervised learning in a hierarchical manner.

86, **TITLE:** PLM-based World Models for Text-based Games

<https://aclanthology.org/2022.emnlp-main.86>

**AUTHORS:** Minsoo Kim, Yeonjoon Jung, Dohyeon Lee, Seung-won Hwang

**HIGHLIGHT:** As the core tasks of world models are future prediction and commonsense understanding, our claim is that pre-trained language models (PLMs) already provide a strong base upon which to build world models.

87, **TITLE:** Prompt-Based Meta-Learning For Few-shot Text Classification

<https://aclanthology.org/2022.emnlp-main.87>

**AUTHORS:** Haoxing Zhang, Xiaofeng Zhang, Haibo Huang, Lei Yu

**HIGHLIGHT:** Prompt-tuning has recently proved to be another effective few-shot learner by bridging the gap between pre-train and downstream tasks. In this work, we closely combine the two promising few-shot learning methodologies in structure and propose a Prompt-Based Meta-Learning (PBML) model to overcome the above meta-learning problem by adding the prompting mechanism.

88, **TITLE:** How well can Text-to-Image Generative Models understand Ethical Natural Language Interventions?

<https://aclanthology.org/2022.emnlp-main.88>

**AUTHORS:** Hritik Bansal, Da Yin, Masoud Monajatipoor, Kai-Wei Chang

**HIGHLIGHT:** To this end, we introduce an Ethical Natural Language Interventions in Text-to-Image GENERation (ENTIGEN) benchmark dataset to evaluate the change in image generations conditional on ethical interventions across three social axes ? gender, skin color, and culture.

89, **TITLE:** Geographic Citation Gaps in NLP Research

<https://aclanthology.org/2022.emnlp-main.89>

**AUTHORS:** Mukund Rungta, Janvijay Singh, Saif M. Mohammad, Diyi Yang

**HIGHLIGHT:** In the spirit of “what we do not measure, we cannot improve”, this work asks a series of questions on the relationship between geographical location and publication success (acceptance in top NLP venues and citation impact).

90, **TITLE:** Language Models of Code are Few-Shot Commonsense Learners

<https://aclanthology.org/2022.emnlp-main.90>

**AUTHORS:** Aman Madaan, Shuyan Zhou, Uri Alon, Yiming Yang, Graham Neubig

**HIGHLIGHT:** In this paper, we show that when we instead frame structured commonsense reasoning tasks as code generation tasks, pre-trained LMs of code are better structured commonsense reasoners than LMs of natural language, even when the downstream task does not involve source code at all. We demonstrate our approach across three diverse structured commonsense reasoning tasks.

91, **TITLE:** Numerical Optimizations for Weighted Low-rank Estimation on Language Models

<https://aclanthology.org/2022.emnlp-main.91>

**AUTHORS:** Ting Hua, Yen-Chang Hsu, Felicity Wang, Qian Lou, Yilin Shen, Hongxia Jin

**HIGHLIGHT:** Unlike standard SVD, weighed value decomposition is a non-convex optimization problem that lacks a closed-form solution. We systematically investigated multiple optimization strategies to tackle the problem and examined our method by compressing Transformer-based language models.

92, **TITLE:** Generative Multi-hop Retrieval

<https://aclanthology.org/2022.emnlp-main.92>

**AUTHORS:** Hyunji Lee, Sohee Yang, Hanseok Oh, Minjoon Seo

**HIGHLIGHT:** However, such a bi-encoder approach has limitations in multi-hop settings; (1) the reformulated query gets longer as the number of hops increases, which further tightens the embedding bottleneck of the query vector, and (2) it is prone to error propagation. In this paper, we focus on alleviating these limitations in multi-hop settings by formulating the problem in a fully generative way.

93, **TITLE:** Visual Spatial Description: Controlled Spatial-Oriented Image-to-Text Generation

<https://aclanthology.org/2022.emnlp-main.93>

**AUTHORS:** Yu Zhao, Jianguo Wei, ZhiChao Lin, Yueheng Sun, Meishan Zhang, Min Zhang

**HIGHLIGHT:** Accordingly, we annotate a dataset manually to facilitate the investigation of the newly-introduced task, and then build several benchmark encoder-decoder models by using VL-BART and VL-T5 as backbones.

94, **TITLE:** M3: A Multi-View Fusion and Multi-Decoding Network for Multi-Document Reading Comprehension

<https://aclanthology.org/2022.emnlp-main.94>

**AUTHORS:** Liang Wen, Houfeng Wang, Yingwei Luo, Xiaolin Wang

**HIGHLIGHT:** In this work, we propose a novel method that tries to employ a multi-view fusion and multi-decoding mechanism to achieve it.

95, **TITLE:** COCO-DR: Combating the Distribution Shift in Zero-Shot Dense Retrieval with Contrastive and Distributionally Robust Learning

<https://aclanthology.org/2022.emnlp-main.95>

**AUTHORS:** Yue Yu, Chenyan Xiong, Si Sun, Chao Zhang, Arnold Overwijk

**HIGHLIGHT:** We present a new zero-shot dense retrieval (ZeroDR) method, COCO-DR, to improve the generalization ability of dense retrieval by combating the distribution shifts between source training tasks and target scenarios.

96, **TITLE:** Language Model Pre-Training with Sparse Latent Typing  
<https://aclanthology.org/2022.emnlp-main.96>

**AUTHORS:** Liliang Ren, Zixuan Zhang, Han Wang, Clare Voss, ChengXiang Zhai, Heng Ji

**HIGHLIGHT:** In this paper, we manage to push the language models to obtain a deeper understanding of sentences by proposing a new pre-training objective, Sparse Latent Typing, which enables the model to sparsely extract sentence-level keywords with diverse latent types.

97, **TITLE:** On the Transformation of Latent Space in Fine-Tuned NLP Models  
<https://aclanthology.org/2022.emnlp-main.97>

**AUTHORS:** Nadir Durrani, Hassan Sajjad, Fahim Dalvi, Firoj Alam

**HIGHLIGHT:** We study the evolution of latent space in fine-tuned NLP models.

98, **TITLE:** Watch the Neighbors: A Unified K-Nearest Neighbor Contrastive Learning Framework for OOD Intent Discovery  
<https://aclanthology.org/2022.emnlp-main.98>

**AUTHORS:** Yutao Mou, Keqing He, Pei Wang, Yanan Wu, Jingang Wang, Wei Wu, Weiran Xu

**HIGHLIGHT:** In this paper, we propose a unified K-nearest neighbor contrastive learning framework to discover OOD intents.

99, **TITLE:** Extracted BERT Model Leaks More Information than You Think!

<https://aclanthology.org/2022.emnlp-main.99>

**AUTHORS:** Xuanli He, Lingjuan Lyu, Chen Chen, Qiongzai Xu

**HIGHLIGHT:** This work bridges this gap by launching an attribute inference attack against the extracted BERT model.

100, **TITLE:** Do Vision-and-Language Transformers Learn Grounded Predicate-Noun Dependencies?

<https://aclanthology.org/2022.emnlp-main.100>

**AUTHORS:** Mitja Nikolaus, Emmanuelle Salin, Stephane Ayache, Abdellah Fourtassi, Benoit Favre

**HIGHLIGHT:** We take a first step in closing this gap by creating a new multimodal task targeted at evaluating understanding of predicate-noun dependencies in a controlled setup.

101, **TITLE:** A Multilingual Perspective Towards the Evaluation of Attribution Methods in Natural Language Inference

<https://aclanthology.org/2022.emnlp-main.101>

**AUTHORS:** Kerem Zaman, Yonatan Belinkov

**HIGHLIGHT:** In this work, we present a multilingual approach for evaluating attribution methods for the Natural Language Inference (NLI) task in terms of faithfulness and plausibility.

102, **TITLE:** Graph-Based Multilingual Label Propagation for Low-Resource Part-of-Speech Tagging

<https://aclanthology.org/2022.emnlp-main.102>

**AUTHORS:** Ayyoob ImaniGooghari, Silvia Severini, Masoud Jalili Sabet, Francois Yvon, Hinrich Sch?tze

**HIGHLIGHT:** In this paper, we propose a novel method for transferring labels from multiple high-resource source to low-resource target languages.

103, **TITLE:** SubeventWriter: Iterative Sub-event Sequence Generation with Coherence Controller

<https://aclanthology.org/2022.emnlp-main.103>

**AUTHORS:** Zhaowei Wang, Hongming Zhang, Tianqing Fang, Yangqiu Song, Ginny Wong, Simon See

**HIGHLIGHT:** In this paper, we propose a new task of sub-event generation for an unseen process to evaluate the understanding of the coherence of sub-event actions and objects.

104, **TITLE:** Infinite SCAN: An Infinite Model of Diachronic Semantic Change

<https://aclanthology.org/2022.emnlp-main.104>

**AUTHORS:** Seiichi Inoue, Mamoru Komachi, Toshinobu Ogiso, Hiroya Takamura, Daichi Mochihashi

**HIGHLIGHT:** In this study, we propose a Bayesian model that can jointly estimate the number of senses of words and their changes through time. The model combines a dynamic topic model on Gaussian Markov random fields with a logistic stick-breaking process that realizes Dirichlet process.

105, **TITLE:** Learning Instructions with Unlabeled Data for Zero-Shot Cross-Task Generalization

<https://aclanthology.org/2022.emnlp-main.105>

**AUTHORS:** Yuxian Gu, Pei Ke, Xiaoyan Zhu, Minlie Huang

**HIGHLIGHT:** In this work, we study how IT can be improved with unlabeled data.

106, **TITLE:** Counterfactual Data Augmentation via Perspective Transition for Open-Domain Dialogues

<https://aclanthology.org/2022.emnlp-main.106>

**AUTHORS:** Jiao Ou, Jinchao Zhang, Yang Feng, Jie Zhou

**HIGHLIGHT:** In this paper, we propose a data augmentation method to automatically augment high-quality responses with different semantics by counterfactual inference.

107, **TITLE:** SQUIRE: A Sequence-to-sequence Framework for Multi-hop Knowledge Graph Reasoning

<https://aclanthology.org/2022.emnlp-main.107>

AUTHORS: Yushi Bai, Xin Lv, Juanzi Li, Lei Hou, Yincen Qu, Zelin Dai, Feiyu Xiong  
HIGHLIGHT: Here we present SQUIRE, the first Sequence-to-sequence based multi-hop reasoning framework, which utilizes an encoder-decoder Transformer structure to translate the query to a path.

108, TITLE: SpeechUT: Bridging Speech and Text with Hidden-Unit for Encoder-Decoder Based Speech-Text Pre-training  
<https://aclanthology.org/2022.emnlp-main.108>

AUTHORS: Ziqiang Zhang, Long Zhou, Junyi Ao, Shujie Liu, Lirong Dai, Jinyu Li, Furu Wei  
HIGHLIGHT: In this paper, we propose a unified-modal speech-unit-text pre-training model, SpeechUT, to connect the representations of a speech encoder and a text decoder with a shared unit encoder.

109, TITLE: Learning Label Modular Prompts for Text Classification in the Wild  
<https://aclanthology.org/2022.emnlp-main.109>

AUTHORS: Hailin Chen, Amrita Saha, Shafiq Joty, Steven C.H. Hoi  
HIGHLIGHT: However, current modular approaches in NLP do not take advantage of recent advances in parameter efficient tuning of pretrained language models. To close this gap, we propose ModularPrompt, a label-modular prompt tuning framework for text classification tasks.

110, TITLE: Unbiased and Efficient Sampling of Dependency Trees  
<https://aclanthology.org/2022.emnlp-main.110>

AUTHORS: Milo? Stanojevic  
HIGHLIGHT: In this paper we show that their fastest algorithm for sampling with replacement, Wilson-RC, is in fact producing biased samples and we provide two alternatives that are unbiased.

111, TITLE: Continual Learning of Neural Machine Translation within Low Forgetting Risk Regions  
<https://aclanthology.org/2022.emnlp-main.111>

AUTHORS: Shuhao Gu, Bojie Hu, Yang Feng  
HIGHLIGHT: To solve the problem, we propose a two-stage training method based on the local features of the real loss.

112, TITLE: COST-EFF: Collaborative Optimization of Spatial and Temporal Efficiency with Slenderized Multi-exit Language Models  
<https://aclanthology.org/2022.emnlp-main.112>

AUTHORS: Bowen Shen, Zheng Lin, Yuanxin Liu, Zhengxiao Liu, Lei Wang, Weiping Wang  
HIGHLIGHT: To address the trade-off of early exiting, we propose a joint training approach that calibrates slenderization and preserves contributive structures to each exit instead of only the final layer.

113, TITLE: Rescue Implicit and Long-tail Cases: Nearest Neighbor Relation Extraction  
<https://aclanthology.org/2022.emnlp-main.113>

AUTHORS: Zhen Wan, Qianying Liu, Zhuoyuan Mao, Fei Cheng, Sadao Kurohashi, Jiwei Li  
HIGHLIGHT: In this paper, we introduce a simple enhancement of RE using k nearest neighbors (kNN-RE).

114, TITLE: StoryER: Automatic Story Evaluation via Ranking, Rating and Reasoning  
<https://aclanthology.org/2022.emnlp-main.114>

AUTHORS: Hong Chen, Duc Vo, Hiroya Takamura, Yusuke Miyao, Hideki Nakayama  
HIGHLIGHT: Existing automatic story evaluation methods place a premium on story lexical level coherence, deviating from human preference. We go beyond this limitation by considering a novel Story Evaluation method that mimics human preference when judging a story, namely StoryER, which consists of three sub-tasks: Ranking, Rating and Reasoning.

115, TITLE: Enhancing Self-Consistency and Performance of Pre-Trained Language Models through Natural Language Inference  
<https://aclanthology.org/2022.emnlp-main.115>

AUTHORS: Eric Mitchell, Joseph Noh, Siyan Li, Will Armstrong, Ananth Agarwal, Patrick Liu, Chelsea Finn, Christopher Manning  
HIGHLIGHT:  $\langle i \rangle$ . To address this failure mode, we propose a framework, Consistency Correction through Relation Detection, or  $\langle b \rangle$ ConCoRD $\langle b \rangle$ , for boosting the consistency and accuracy of pre-trained NLP models using pre-trained natural language inference (NLI) models without fine-tuning or re-training.

116, TITLE: Robustness of Demonstration-based Learning Under Limited Data Scenario  
<https://aclanthology.org/2022.emnlp-main.116>

AUTHORS: Hongxin Zhang, Yanzhe Zhang, Ruiyi Zhang, Diyi Yang  
HIGHLIGHT: In this paper, we design pathological demonstrations by gradually removing intuitively useful information from the standard ones to take a deep dive of the robustness of demonstration-based sequence labeling and show that (1) demonstrations composed of random tokens still make the model a better few-shot learner; (2) the length of random demonstrations and the relevance of random tokens are the main factors affecting the performance; (3) demonstrations increase the confidence of model predictions on captured superficial patterns.

117, TITLE: Modeling Information Change in Science Communication with Semantically Matched Paraphrases  
<https://aclanthology.org/2022.emnlp-main.117>

AUTHORS: Dustin Wright, Jiaxin Pei, David Jurgens, Isabelle Augenstein  
HIGHLIGHT: To this end, we present the SCIENTIFIC PARAPHRASE AND INFORMATION CHANGE DATASET (SPICED), the first paraphrase dataset of scientific findings annotated for degree of information change.

- 118, TITLE: Word Order Matters When You Increase Masking  
<https://aclanthology.org/2022.emnlp-main.118>  
AUTHORS: Karim Lasri, Alessandro Lenci, Thierry Poibeau  
HIGHLIGHT: However, recent experiments have shown that explicit position encoding is not always useful, since some models without such feature managed to achieve state-of-the art performance on some tasks. To understand better this phenomenon, we examine the effect of removing position encodings on the pre-training objective itself (i.e., masked language modelling), to test whether models can reconstruct position information from co-occurrences alone.
- 119, TITLE: An Empirical Analysis of Memorization in Fine-tuned Autoregressive Language Models  
<https://aclanthology.org/2022.emnlp-main.119>  
AUTHORS: Fatemehsadat Mireshghallah, Archit Uniyal, Tianhao Wang, David Evans, Taylor Berg-Kirkpatrick  
HIGHLIGHT: In this paper, we empirically study memorization of fine-tuning methods using membership inference and extraction attacks, and show that their susceptibility to attacks is very different.
- 120, TITLE: Style Transfer as Data Augmentation: A Case Study on Named Entity Recognition  
<https://aclanthology.org/2022.emnlp-main.120>  
AUTHORS: Shuguang Chen, Leonardo Neves, Tamar Solorio  
HIGHLIGHT: In this work, we take the named entity recognition task in the English language as a case study and explore style transfer as a data augmentation method to increase the size and diversity of training data in low-resource scenarios.
- 121, TITLE: Linguistic Corpus Annotation for Automatic Text Simplification Evaluation  
<https://aclanthology.org/2022.emnlp-main.121>  
AUTHORS: R?mi Cardon, Adrien Bibal, Rodrigo Wilkens, David Alfter, Magali Norr?, Adeline M?ller, Watrin Patrick, Thomas Fran?ois  
HIGHLIGHT: In this paper, we propose annotations of the ASSET corpus that can be used to shed more light on ATS evaluation.
- 122, TITLE: Semantic Framework based Query Generation for Temporal Question Answering over Knowledge Graphs  
<https://aclanthology.org/2022.emnlp-main.122>  
AUTHORS: Wentao Ding, Hao Chen, Huayu Li, Yuzhong Qu  
HIGHLIGHT: Based on the semantic framework, we propose a temporal question answering method, SF-TQA, which generates query graphs by exploring the relevant facts of mentioned entities, where the exploring process is restricted by SF-TCons.
- 123, TITLE: There Is No Standard Answer: Knowledge-Grounded Dialogue Generation with Adversarial Activated Multi-Reference Learning  
<https://aclanthology.org/2022.emnlp-main.123>  
AUTHORS: Xueliang Zhao, Tingchen Fu, Chongyang Tao, Rui Yan  
HIGHLIGHT: To this end, we establish a multi-reference KGC dataset and propose a series of metrics to systematically assess the one-to-many efficacy of existing KGC models.
- 124, TITLE: Stop Measuring Calibration When Humans Disagree  
<https://aclanthology.org/2022.emnlp-main.124>  
AUTHORS: Joris Baan, Wilker Aziz, Barbara Plank, Raquel Fernandez  
HIGHLIGHT: Recently, calibration to human majority has been measured on tasks where humans inherently disagree about which class applies. We show that measuring calibration to human majority given inherent disagreements is theoretically problematic, demonstrate this empirically on the ChaosNLI dataset, and derive several instance-level measures of calibration that capture key statistical properties of human judgements - including class frequency, ranking and entropy.
- 125, TITLE: Improving compositional generalization for multi-step quantitative reasoning in question answering  
<https://aclanthology.org/2022.emnlp-main.125>  
AUTHORS: Armineh Nourbakhsh, Cathy Jiao, Sameena Shah, Carolyn Ros?  
HIGHLIGHT: Quantitative reasoning is an important aspect of question answering, especially when numeric and verbal cues interact to indicate sophisticated, multi-step programs. In this paper, we demonstrate how modeling the compositional nature of quantitative text can enhance the performance and robustness of QA models, allowing them to capture arithmetic logic that is expressed verbally.
- 126, TITLE: A Comprehensive Comparison of Neural Networks as Cognitive Models of Inflection  
<https://aclanthology.org/2022.emnlp-main.126>  
AUTHORS: Adam Wiemerslage, Shiran Dudy, Katharina Kann  
HIGHLIGHT: This debate has gravitated into NLP by way of the question: Are neural networks a feasible account for human behavior in morphological inflection?We address that question by measuring the correlation between human judgments and neural network probabilities for unknown word inflections.
- 127, TITLE: Can Visual Context Improve Automatic Speech Recognition for an Embodied Agent?  
<https://aclanthology.org/2022.emnlp-main.127>  
AUTHORS: Pradip Pramanick, Chayan Sarkar  
HIGHLIGHT: In this work, we present a method to incorporate a robot?s visual information into an ASR system and improve the recognition of a spoken utterance containing a visible entity.

- 128, TITLE: AfroLID: A Neural Language Identification Tool for African Languages  
<https://aclanthology.org/2022.emnlp-main.128>  
AUTHORS: Ife Adebara, AbdelRahim Elmadany, Muhammad Abdul-Mageed, Alcides Inciarte  
HIGHLIGHT: Problematically, most of the world's 7000+ languages today are not covered by LID technologies. We address this pressing issue for Africa by introducing AfroLID, a neural LID toolkit for 517 African languages and varieties.
- 129, TITLE: EvEntS ReaLM: Event Reasoning of Entity States via Language Models  
<https://aclanthology.org/2022.emnlp-main.129>  
AUTHORS: Evangelia Spiliopoulou, Artidoro Pagnoni, Yonatan Bisk, Eduard Hovy  
HIGHLIGHT: This paper investigates models of event implications.
- 130, TITLE: Large language models are few-shot clinical information extractors  
<https://aclanthology.org/2022.emnlp-main.130>  
AUTHORS: Monica Agrawal, Stefan Hegselmann, Hunter Lang, Yoon Kim, David Sontag  
HIGHLIGHT: In this work, we show that large language models, such as InstructGPT (Ouyang et al., 2022), perform well at zero- and few-shot information extraction from clinical text despite not being trained specifically for the clinical domain.
- 131, TITLE: Towards a Unified Multi-Dimensional Evaluator for Text Generation  
<https://aclanthology.org/2022.emnlp-main.131>  
AUTHORS: Ming Zhong, Yang Liu, Da Yin, Yuning Mao, Yizhu Jiao, Pengfei Liu, Chenguang Zhu, Heng Ji, Jiawei Han  
HIGHLIGHT: In this paper, we propose a unified multi-dimensional evaluator UniEval for NLG.
- 132, TITLE: GeoMLAMA: Geo-Diverse Commonsense Probing on Multilingual Pre-Trained Language Models  
<https://aclanthology.org/2022.emnlp-main.132>  
AUTHORS: Da Yin, Hritik Bansal, Masoud Monajatipoor, Liunian Harold Li, Kai-Wei Chang  
HIGHLIGHT: In this paper, we introduce a benchmark dataset, Geo-diverse Commonsense Multilingual Language Models Analysis (GeoMLAMA), for probing the diversity of the relational knowledge in multilingual PLMs.
- 133, TITLE: The (Undesired) Attenuation of Human Biases by Multilinguality  
<https://aclanthology.org/2022.emnlp-main.133>  
AUTHORS: Cristina Espa?a-Bonet, Alberto Barr?n-Cede?o  
HIGHLIGHT: We introduce and release CA-WEAT, multilingual cultural aware tests to quantify biases, and compare them to previous English-centric tests.
- 134, TITLE: Entailer: Answering Questions with Faithful and Truthful Chains of Reasoning  
<https://aclanthology.org/2022.emnlp-main.134>  
AUTHORS: Oyvind Tafjord, Bhavana Dalvi Mishra, Peter Clark  
HIGHLIGHT: Our goal is a question-answering (QA) system that can show how its answers are implied by its own internal beliefs via a systematic chain of reasoning.
- 135, TITLE: Near-Negative Distinction: Giving a Second Life to Human Evaluation Datasets  
<https://aclanthology.org/2022.emnlp-main.135>  
AUTHORS: Philippe Laban, Chien-Sheng Wu, Wenhao Liu, Caiming Xiong  
HIGHLIGHT: In this paper, we propose a new and simple automatic evaluation method for NLG called Near-Negative Distinction (NND) that repurposes prior human annotations into NND tests. In an NND test, an NLG model must place a higher likelihood on a high-quality output candidate than on a near-negative candidate with a known error.
- 136, TITLE: ToKen: Task Decomposition and Knowledge Infusion for Few-Shot Hate Speech Detection  
<https://aclanthology.org/2022.emnlp-main.136>  
AUTHORS: Badr AlKhamissi, Faisal Ladhak, Srinivasan Iyer, Veselin Stoyanov, Zornitsa Kozareva, Xian Li, Pascale Fung, Lambert Mathias, Asli Celikyilmaz, Mona Diab  
HIGHLIGHT: It is also difficult to collect a large-scale hate speech annotated dataset. In this work, we frame this problem as a few-shot learning task, and show significant gains with decomposing the task into its ?constituent? parts.
- 137, TITLE: Are Hard Examples also Harder to Explain? A Study with Human and Model-Generated Explanations  
<https://aclanthology.org/2022.emnlp-main.137>  
AUTHORS: Swarnadeep Saha, Peter Hase, Nazneen Rajani, Mohit Bansal  
HIGHLIGHT: In this work, we study the connection between explainability and sample hardness by investigating the following research question ? Are LLMs and humans equally good at explaining data labels for both easy and hard samples?
- 138, TITLE: Stanceosaurus: Classifying Stance Towards Multicultural Misinformation  
<https://aclanthology.org/2022.emnlp-main.138>  
AUTHORS: Jonathan Zheng, Ashutosh Baheti, Tarek Naous, Wei Xu, Alan Ritter  
HIGHLIGHT: We present Stanceosaurus, a new corpus of 28,033 tweets in English, Hindi and Arabic annotated with stance towards 250 misinformation claims.
- 139, TITLE: Gendered Mental Health Stigma in Masked Language Models  
<https://aclanthology.org/2022.emnlp-main.139>  
AUTHORS: Inna Lin, Lucille Njoo, Anjalie Field, Ashish Sharma, Katharina Reinecke, Tim Althoff, Yulia Tsvetkov  
HIGHLIGHT: In this work, we investigate gendered mental health stigma in masked language models.

- 140, TITLE: Efficient Nearest Neighbor Search for Cross-Encoder Models using Matrix Factorization  
<https://aclanthology.org/2022.emnlp-main.140>  
AUTHORS: Nishant Yadav, Nicholas Monath, Rico Angell, Manzil Zaheer, Andrew McCallum  
HIGHLIGHT: In this paper, we present an approach that avoids the use of a dual-encoder for retrieval, relying solely on the cross-encoder.
- 141, TITLE: Prompt-and-Rerank: A Method for Zero-Shot and Few-Shot Arbitrary Textual Style Transfer with Small Language Models  
<https://aclanthology.org/2022.emnlp-main.141>  
AUTHORS: Mirac Suzgun, Luke Melas-Kyriazi, Dan Jurafsky  
HIGHLIGHT: We propose a method for arbitrary textual style transfer (TST)?the task of transforming a text into any given style?utilizing general-purpose pre-trained language models.
- 142, TITLE: Learning to Decompose: Hypothetical Question Decomposition Based on Comparable Texts  
<https://aclanthology.org/2022.emnlp-main.142>  
AUTHORS: Ben Zhou, Kyle Richardson, Xiaodong Yu, Dan Roth  
HIGHLIGHT: In this paper, we look at large-scale intermediate pre-training of decomposition-based transformers using distant supervision from comparable texts, particularly large-scale parallel news.
- 143, TITLE: Why is Winoground Hard? Investigating Failures in Visuolinguistic Compositionality  
<https://aclanthology.org/2022.emnlp-main.143>  
AUTHORS: Anuj Diwan, Layne Berry, Eunsol Choi, David Harwath, Kyle Mahowald  
HIGHLIGHT: In this paper, we identify the dataset?s main challenges through a suite of experiments on related tasks (probing task, image retrieval task), data augmentation, and manual inspection of the dataset.
- 144, TITLE: Gradient-based Constrained Sampling from Language Models  
<https://aclanthology.org/2022.emnlp-main.144>  
AUTHORS: Sachin Kumar, Biswajit Paria, Yulia Tsvetkov  
HIGHLIGHT: Large pretrained language models are successful at generating fluent text but are notoriously hard to controllably sample from. In this work, we study constrained sampling from such language models, i.e., generating text that satisfies user-defined constraints, while maintaining fluency and model?s performance in a downstream task.
- 145, TITLE: TaCube: Pre-computing Data Cubes for Answering Numerical-Reasoning Questions over Tabular Data  
<https://aclanthology.org/2022.emnlp-main.145>  
AUTHORS: Fan Zhou, Mengkang Hu, Haoyu Dong, Zhoujun Cheng, Fan Cheng, Shi Han, Dongmei Zhang  
HIGHLIGHT: However, auto-regressive PLMs are challenged by recent emerging numerical reasoning datasets, such as TAT-QA, due to the error-prone implicit calculation. In this paper, we present TaCube, to pre-compute aggregation/arithmetic results for the table in advance, so that they are handy and readily available for PLMs to answer numerical reasoning questions.
- 146, TITLE: Rich Knowledge Sources Bring Complex Knowledge Conflicts: Recalibrating Models to Reflect Conflicting Evidence  
<https://aclanthology.org/2022.emnlp-main.146>  
AUTHORS: Hung-Ting Chen, Michael Zhang, Eunsol Choi  
HIGHLIGHT: In this paper, we simulate knowledge conflicts (i.e., where parametric knowledge suggests one answer and different passages suggest different answers) and examine model behaviors.
- 147, TITLE: QA Domain Adaptation using Hidden Space Augmentation and Self-Supervised Contrastive Adaptation  
<https://aclanthology.org/2022.emnlp-main.147>  
AUTHORS: Zhenrui Yue, Huimin Zeng, Bernhard Kratzwald, Stefan Feuerriegel, Dong Wang  
HIGHLIGHT: In this paper, we propose a novel self-supervised framework called QADA for QA domain adaptation.
- 148, TITLE: When FLUE Meets FLANG: Benchmarks and Large Pretrained Language Model for Financial Domain  
<https://aclanthology.org/2022.emnlp-main.148>  
AUTHORS: Raj Shah, Kunal Chawla, Dheeraj Eidnani, Agam Shah, Wendi Du, Sudheer Chava, Natraj Raman, Charese Smiley, Jiaao Chen, Diyi Yang  
HIGHLIGHT: We propose a novel domain specific Financial LANGuage model (FLANG) which uses financial keywords and phrases for better masking, together with span boundary objective and in-filing objective.
- 149, TITLE: Retrieval as Attention: End-to-end Learning of Retrieval and Reading within a Single Transformer  
<https://aclanthology.org/2022.emnlp-main.149>  
AUTHORS: Zhengbao Jiang, Luyu Gao, Zhiruo Wang, Jun Araki, Haibo Ding, Jamie Callan, Graham Neubig  
HIGHLIGHT: These two components are usually modeled separately, which necessitates a cumbersome implementation and is awkward to optimize in an end-to-end fashion. In this paper, we revisit this design and eschew the separate architecture and training in favor of a single Transformer that performs retrieval as attention (RAA), and end-to-end training solely based on supervision from the end QA task.
- 150, TITLE: Reproducibility in Computational Linguistics: Is Source Code Enough?  
<https://aclanthology.org/2022.emnlp-main.150>  
AUTHORS: Mohammad Arvan, Lu?s Pina, Natalie Parde

**HIGHLIGHT:** This work studies trends in source code availability at major computational linguistics conferences, namely, ACL, EMNLP, LREC, NAACL, and COLING.

151, **TITLE:** Generating Information-Seeking Conversations from Unlabeled Documents

<https://aclanthology.org/2022.emnlp-main.151>

**AUTHORS:** Gangwoo Kim, Sungdong Kim, Kang Min Yoo, Jaewoo Kang

**HIGHLIGHT:** In this paper, we introduce a novel framework, **SimSeek**, (**Simulating information-Seeking** conversation from unlabeled documents), and compare its two variants.

152, **TITLE:** Distill The Image to Nowhere: Inversion Knowledge Distillation for Multimodal Machine Translation

<https://aclanthology.org/2022.emnlp-main.152>

**AUTHORS:** Ru Peng, Yawen Zeng, Jake Zhao

**HIGHLIGHT:** Thus, in this work, we introduce IKD-MMT, a novel MMT framework to support the image-free inference phase via an inversion knowledge distillation scheme.

153, **TITLE:** A Multifaceted Framework to Evaluate Evasion, Content Preservation, and Misattribution in Authorship Obfuscation Techniques

<https://aclanthology.org/2022.emnlp-main.153>

**AUTHORS:** Malik Altrkri, Thomas Scialom, Benjamin C. M. Fung, Jackie Chi Kit Cheung

**HIGHLIGHT:** Motivated by recent work on cross-topic authorship identification and content preservation in summarization, we re-evaluate different authorship obfuscation techniques on detection evasion and content preservation.

154, **TITLE:** SafeText: A Benchmark for Exploring Physical Safety in Language Models

<https://aclanthology.org/2022.emnlp-main.154>

**AUTHORS:** Sharon Levy, Emily Allaway, Melanie Subbiah, Lydia Chilton, Desmond Patton, Kathleen McKeown, William Yang Wang

**HIGHLIGHT:** We create the first benchmark dataset, SafeText, comprising real-life scenarios with paired safe and physically unsafe pieces of advice.

155, **TITLE:** Ground-Truth Labels Matter: A Deeper Look into Input-Label Demonstrations

<https://aclanthology.org/2022.emnlp-main.155>

**AUTHORS:** Kang Min Yoo, Junyeob Kim, Huhng Joon Kim, Hyunsoo Cho, Hwiyeol Jo, Sang-Woo Lee, Sang-goo Lee, Taek Kim

**HIGHLIGHT:** Intrigued by this counter-intuitive observation, we re-examine the importance of ground-truth labels in in-context learning.

156, **TITLE:** D4: a Chinese Dialogue Dataset for Depression-Diagnosis-Oriented Chat

<https://aclanthology.org/2022.emnlp-main.156>

**AUTHORS:** Binwei Yao, Chao Shi, Likai Zou, Lingfeng Dai, Mengyue Wu, Lu Chen, Zhen Wang, Kai Yu

**HIGHLIGHT:** Based on clinical depression diagnostic criteria ICD-11 and DSM-5, we designed a 3-phase procedure to construct D4: a Chinese Dialogue Dataset for Depression-Diagnosis-Oriented Chat, which simulates the dialogue between doctors and patients during the diagnosis of depression, including diagnosis results and symptom summary given by professional psychiatrists for each conversation.

157, **TITLE:** Exploiting domain-slot related keywords description for Few-Shot Cross-Domain Dialogue State Tracking

<https://aclanthology.org/2022.emnlp-main.157>

**AUTHORS:** Gao Qixiang, Guanting Dong, Yutao Mou, Liwen Wang, Chen Zeng, Daichi Guo, Mingyang Sun, Weiran Xu

**HIGHLIGHT:** In this paper, we propose a novel framework based on domain-slot related description to tackle the challenge of few-shot cross-domain DST.

158, **TITLE:** CoCoo: An Encoder-Decoder Model for Controllable Code-switched Generation

<https://aclanthology.org/2022.emnlp-main.158>

**AUTHORS:** Sneha Mondal, Ritika ., Shreya Pathak, Preethi Jyothi, Aravindan Raghuvver

**HIGHLIGHT:** We present CoCoo, an encoder-decoder translation model that converts monolingual Hindi text to Hindi-English code-switched text with both encoder-side and decoder-side interventions to achieve fine-grained controllable generation.

159, **TITLE:** Towards Climate Awareness in NLP Research

<https://aclanthology.org/2022.emnlp-main.159>

**AUTHORS:** Daniel Hershcovich, Nicolas Webersinke, Mathias Kraus, Julia Bingler, Markus Leippold

**HIGHLIGHT:** As a remedy, we propose a climate performance model card with the primary purpose of being practically usable with only limited information about experiments and the underlying computer hardware.

160, **TITLE:** Navigating Connected Memories with a Task-oriented Dialog System

<https://aclanthology.org/2022.emnlp-main.160>

**AUTHORS:** Satwik Kottur, Seungwhan Moon, Alborz Geramifard, Babak Damavandi

**HIGHLIGHT:** In this work, we propose dialogs for connected memories as a powerful tool to empower users to search their media collection through a multi-turn, interactive conversation.

161, **TITLE:** Language Model Decomposition: Quantifying the Dependency and Correlation of Language Models

<https://aclanthology.org/2022.emnlp-main.161>

- AUTHORS: Hao Zhang  
HIGHLIGHT: However, a theoretical framework for studying their relationships is still missing. In this paper, we fill this gap by investigating the linear dependency between pre-trained LMs.
- 162, TITLE: SynGEC: Syntax-Enhanced Grammatical Error Correction with a Tailored GEC-Oriented Parser  
<https://aclanthology.org/2022.emnlp-main.162>  
AUTHORS: Yue Zhang, Bo Zhang, Zhenghua Li, Zuyi Bao, Chen Li, Min Zhang  
HIGHLIGHT: This work proposes a syntax-enhanced grammatical error correction (GEC) approach named SynGEC that effectively incorporates dependency syntactic information into the encoder part of GEC models.
- 163, TITLE: Varifocal Question Generation for Fact-checking  
<https://aclanthology.org/2022.emnlp-main.163>  
AUTHORS: Nedjma Ousidhoum, Zhangdie Yuan, Andreas Vlachos  
HIGHLIGHT: In this paper, we present Varifocal, a method that generates questions based on different focal points within a given claim, i.e. different spans of the claim and its metadata, such as its source and date.
- 164, TITLE: Bilingual Lexicon Induction for Low-Resource Languages using Graph Matching via Optimal Transport  
<https://aclanthology.org/2022.emnlp-main.164>  
AUTHORS: Kelly Marchisio, Ali Saad-Eldin, Kevin Duh, Carey Priebe, Philipp Koehn  
HIGHLIGHT: In this work, we improve bilingual lexicon induction performance across 40 language pairs with a graph-matching method based on optimal transport.
- 165, TITLE: Whose Language Counts as High Quality? Measuring Language Ideologies in Text Data Selection  
<https://aclanthology.org/2022.emnlp-main.165>  
AUTHORS: Suchin Gururangan, Dallas Card, Sarah Dreier, Emily Gade, Leroy Wang, Zeyu Wang, Luke Zettlemoyer, Noah A. Smith  
HIGHLIGHT: Using a new dataset of U.S. high school newspaper articles written by students from across the country we investigate whose language is preferred by the quality filter used for GPT-3.
- 166, TITLE: ConReader: Exploring Implicit Relations in Contracts for Contract Clause Extraction  
<https://aclanthology.org/2022.emnlp-main.166>  
AUTHORS: Weiwen Xu, Yang Deng, Wenqiang Lei, Wenlong Zhao, Tat-Seng Chua, Wai Lam  
HIGHLIGHT: We study automatic Contract Clause Extraction (CCE) by modeling implicit relations in legal contracts.
- 167, TITLE: Training Dynamics for Curriculum Learning: A Study on Monolingual and Cross-lingual NLU  
<https://aclanthology.org/2022.emnlp-main.167>  
AUTHORS: Fenia Christopoulou, Gerasimos Lampouras, Ignacio Iacobacci  
HIGHLIGHT: Current approaches for Natural Language Understanding (NLU) tasks use CL to improve in-distribution data performance often via heuristic-oriented or task-agnostic difficulties. In this work, instead, we employ CL for NLU by taking advantage of training dynamics as difficulty metrics, i.e., statistics that measure the behavior of the model at hand on specific task-data instances during training and propose modifications of existing CL schedulers based on these statistics.
- 168, TITLE: Revisiting Parameter-Efficient Tuning: Are We Really There Yet?  
<https://aclanthology.org/2022.emnlp-main.168>  
AUTHORS: Guanzheng Chen, Fangyu Liu, Zaiqiao Meng, Shangsong Liang  
HIGHLIGHT: By tuning just a fraction amount of parameters comparing to full model finetuning, PETuning methods claim to have achieved performance on par with or even better than finetuning. In this work, we take a step back and re-examine these PETuning methods by conducting the first comprehensive investigation into the training and evaluation of them.
- 169, TITLE: Transfer Learning from Semantic Role Labeling to Event Argument Extraction with Template-based Slot Querying  
<https://aclanthology.org/2022.emnlp-main.169>  
AUTHORS: Zhisong Zhang, Emma Strubell, Eduard Hovy  
HIGHLIGHT: In this work, we investigate transfer learning from semantic role labeling (SRL) to event argument extraction (EAE), considering their similar argument structures.
- 170, TITLE: Calibrating Zero-shot Cross-lingual (Un-)structured Predictions  
<https://aclanthology.org/2022.emnlp-main.170>  
AUTHORS: Zhengping Jiang, Anqi Liu, Benjamin Van Durme  
HIGHLIGHT: We study different post-training calibration methods in structured and unstructured prediction tasks.
- 171, TITLE: PRINCE: Prefix-Masked Decoding for Knowledge Enhanced Sequence-to-Sequence Pre-Training  
<https://aclanthology.org/2022.emnlp-main.171>  
AUTHORS: Song Xu, Haoran Li, Peng Yuan, Youzheng Wu, Xiaodong He  
HIGHLIGHT: This paper introduces a simple yet effective pre-training paradigm, equipped with a knowledge-enhanced decoder that predicts the next entity token with noises in the prefix, explicitly strengthening the representation learning of entities that span over multiple input tokens.
- 172, TITLE: How Far are We from Robust Long Abstractive Summarization?  
<https://aclanthology.org/2022.emnlp-main.172>



- AUTHORS:** Huan Yee Koh, Jiaxin Ju, He Zhang, Ming Liu, Shirui Pan  
**HIGHLIGHT:** Abstractive summarization has made tremendous progress in recent years. In this work, we perform fine-grained human annotations to evaluate long document abstractive summarization systems (i.e., models and metrics) with the aim of implementing them to generate reliable summaries.
- 173, **TITLE:** Measuring Context-Word Biases in Lexical Semantic Datasets  
<https://aclanthology.org/2022.emnlp-main.173>  
**AUTHORS:** Qianchu Liu, Diana McCarthy, Anna Korhonen  
**HIGHLIGHT:** We question this assumption by presenting the first quantitative analysis on the context-word interaction being tested in major contextual lexical semantic tasks. To achieve this, we run probing baselines on masked input, and propose measures to calculate and visualize the degree of context or word biases in existing datasets.
- 174, **TITLE:** Iteratively Prompt Pre-trained Language Models for Chain of Thought  
<https://aclanthology.org/2022.emnlp-main.174>  
**AUTHORS:** Boshi Wang, Xiang Deng, Huan Sun  
**HIGHLIGHT:** In this work, we explore an iterative prompting framework, a new prompting paradigm which progressively elicits relevant knowledge from PLMs for multi-step inference.
- 175, **TITLE:** Unobserved Local Structures Make Compositional Generalization Hard  
<https://aclanthology.org/2022.emnlp-main.175>  
**AUTHORS:** Ben Bogin, Shivanshu Gupta, Jonathan Berant  
**HIGHLIGHT:** In this work, we investigate the factors that make generalization to certain test instances challenging.
- 176, **TITLE:** Mitigating Data Sparsity for Short Text Topic Modeling by Topic-Semantic Contrastive Learning  
<https://aclanthology.org/2022.emnlp-main.176>  
**AUTHORS:** Xiaobao Wu, Anh Tuan Luu, Xinshuai Dong  
**HIGHLIGHT:** To better address data sparsity, in this paper we propose a novel short text topic modeling framework, Topic-Semantic Contrastive Topic Model (TSCTM).
- 177, **TITLE:** Back to the Future: Bidirectional Information Decoupling Network for Multi-turn Dialogue Modeling  
<https://aclanthology.org/2022.emnlp-main.177>  
**AUTHORS:** Yiyang Li, Hai Zhao, Zhuosheng Zhang  
**HIGHLIGHT:** Recent studies of dialogue modeling commonly employ pre-trained language models (PrLMs) to encode the dialogue history as successive tokens, which is insufficient in capturing the temporal characteristics of dialogues. Therefore, we propose Bidirectional Information Decoupling Network (BiDeN) as a universal dialogue encoder, which explicitly incorporates both the past and future contexts and can be generalized to a wide range of dialogue-related tasks.
- 178, **TITLE:** Calibration Meets Explanation: A Simple and Effective Approach for Model Confidence Estimates  
<https://aclanthology.org/2022.emnlp-main.178>  
**AUTHORS:** Dongfang Li, Baotian Hu, Qingcai Chen  
**HIGHLIGHT:** Similarly, the explanations may tell us when the model might know and when it does not. Inspired by this, we propose a method named CME that leverages model explanations to make the model less confident with non-inductive attributions.
- 179, **TITLE:** Non-Autoregressive Neural Machine Translation: A Call for Clarity  
<https://aclanthology.org/2022.emnlp-main.179>  
**AUTHORS:** Robin Schmidt, Telmo Pires, Stephan Peitz, Jonas L??f  
**HIGHLIGHT:** In this work, we take a step back and revisit several techniques that have been proposed for improving non-autoregressive translation models and compare their combined translation quality and speed implications under third-party testing environments.
- 180, **TITLE:** RED-ACE: Robust Error Detection for ASR using Confidence Embeddings  
<https://aclanthology.org/2022.emnlp-main.180>  
**AUTHORS:** Zorik Gekhman, Dina Zverinski, Jonathan Mallinson, Genady Beryozkin  
**HIGHLIGHT:** Specifically, we add an ASR Confidence Embedding (ACE) layer to the AED modelâ€™s encoder, allowing us to jointly encode the confidence scores and the transcribed text into a contextualized representation.
- 181, **TITLE:** Fast-R2D2: A Pretrained Recursive Neural Network based on Pruned CKY for Grammar Induction and Text Representation  
<https://aclanthology.org/2022.emnlp-main.181>  
**AUTHORS:** Xiang Hu, Haitao Mi, Liang Li, Gerard de Melo  
**HIGHLIGHT:** However, its rule-based pruning process suffers from local optima and slow inference. In this paper, we propose a unified R2D2 method that overcomes these issues.
- 182, **TITLE:** A Localized Geometric Method to Match Knowledge in Low-dimensional Hyperbolic Space  
<https://aclanthology.org/2022.emnlp-main.182>  
**AUTHORS:** Bo Hui, Tian Xia, Wei-Shinn Ku  
**HIGHLIGHT:** In this paper, we propose a localized geometric method to find equivalent entities in hyperbolic space.
- 183, **TITLE:** Memory-assisted prompt editing to improve GPT-3 after deployment  
<https://aclanthology.org/2022.emnlp-main.183>

- AUTHORS: Aman Madaan, Niket Tandon, Peter Clark, Yiming Yang  
HIGHLIGHT: Large LMs such as GPT-3 are powerful, but can commit mistakes that are obvious to humans. For example, GPT-3 would mistakenly interpret "What word is similar to good?" to mean a homophone, while the user intended a synonym. Our goal is to effectively correct such errors via user interactions with the system but without retraining, which will be prohibitively costly.
- 184, TITLE: LVP-M3: Language-aware Visual Prompt for Multilingual Multimodal Machine Translation  
<https://aclanthology.org/2022.emnlp-main.184>  
AUTHORS: Hongcheng Guo, Jiaheng Liu, Haoyang Huang, Jian Yang, Zhoujun Li, Dongdong Zhang, Zheng Cui  
HIGHLIGHT: To this end, we first propose the Multilingual MMT task by establishing two new Multilingual MMT benchmark datasets covering seven languages. Then, an effective baseline LVP-M3 using visual prompts is proposed to support translations between different languages, which includes three stages (token encoding, language-aware visual prompt generation, and language translation).
- 185, TITLE: PromptEHR: Conditional Electronic Healthcare Records Generation with Prompt Learning  
<https://aclanthology.org/2022.emnlp-main.185>  
AUTHORS: Zifeng Wang, Jimeng Sun  
HIGHLIGHT: In this work, we propose to formulate EHRs generation as a text-to-text translation task by language models (LMs), which suffices to highly flexible event imputation during generation.
- 186, TITLE: ROSE: Robust Selective Fine-tuning for Pre-trained Language Models  
<https://aclanthology.org/2022.emnlp-main.186>  
AUTHORS: Lan Jiang, Hao Zhou, Yankai Lin, Peng Li, Jie Zhou, Rui Jiang  
HIGHLIGHT: However, they are still limited due to redundant attack search spaces and the inability to defend against various types of attacks. In this work, we present a novel fine-tuning approach called ROBust SElective fine-tuning (ROSE) to address this issue.
- 187, TITLE: CodeRetriever: A Large Scale Contrastive Pre-Training Method for Code Search  
<https://aclanthology.org/2022.emnlp-main.187>  
AUTHORS: Xiaonan Li, Yeyun Gong, Yelong Shen, Xipeng Qiu, Hang Zhang, Bolun Yao, Weizhen Qi, Daxin Jiang, Weizhu Chen, Nan Duan  
HIGHLIGHT: In this paper, we propose the CodeRetriever model, which learns the function-level code semantic representations through large-scale code-text contrastive pre-training.
- 188, TITLE: Open-Topic False Information Detection on Social Networks with Contrastive Adversarial Learning  
<https://aclanthology.org/2022.emnlp-main.188>  
AUTHORS: Guanghui Ma, Chunming Hu, Ling Ge, Hong Zhang  
HIGHLIGHT: In this open-topic scenario, we empirically find that the existing models suffer from impairment in the detection performance for seen or unseen topic data, resulting in poor overall model performance. To address this issue, we propose a novel Contrastive Adversarial Learning Network, CALN, that employs an unsupervised topic clustering method to capture topic-specific features to enhance the model's performance for seen topics and an unsupervised adversarial learning method to align data representation distributions to enhance the model's generalisation to unseen topics.
- 189, TITLE: Mitigating Inconsistencies in Multimodal Sentiment Analysis under Uncertain Missing Modalities  
<https://aclanthology.org/2022.emnlp-main.189>  
AUTHORS: Jiandian Zeng, Jiantao Zhou, Tianyi Liu  
HIGHLIGHT: To tackle the issue, we propose an Ensemble-based Missing Modality Reconstruction (EMMR) network to detect and recover semantic features of the key missing modality.
- 190, TITLE: ConvTrans: Transforming Web Search Sessions for Conversational Dense Retrieval  
<https://aclanthology.org/2022.emnlp-main.190>  
AUTHORS: Kelong Mao, Zhicheng Dou, Hongjin Qian, Fengran Mo, Xiaohua Cheng, Zhao Cao  
HIGHLIGHT: In this paper, we present ConvTrans, a data augmentation method that can automatically transform easily-accessible web search sessions into conversational search sessions to fundamentally alleviate the data scarcity problem for conversational dense retrieval.
- 191, TITLE: MUSIED: A Benchmark for Event Detection from Multi-Source Heterogeneous Informal Texts  
<https://aclanthology.org/2022.emnlp-main.191>  
AUTHORS: Xiangyu Xi, Jianwei Lv, Shuai Peng Liu, Wei Ye, Fan Yang, Guanglu Wan  
HIGHLIGHT: As a pioneering exploration that expands event detection to the scenarios involving informal and heterogeneous texts, we propose a new large-scale Chinese event detection dataset based on user reviews, text conversations, and phone conversations in a leading e-commerce platform for food service.
- 192, TITLE: Reproducibility Issues for BERT-based Evaluation Metrics  
<https://aclanthology.org/2022.emnlp-main.192>  
AUTHORS: Yanran Chen, Jonas Belouadi, Steffen Eger  
HIGHLIGHT: In this paper, we ask whether results and claims from four recent BERT-based metrics can be reproduced.
- 193, TITLE: Improving Multi-task Stance Detection with Multi-task Interaction Network  
<https://aclanthology.org/2022.emnlp-main.193>  
AUTHORS: Heyan Chai, Siyu Tang, Jinhao Cui, Ye Ding, Binxing Fang, Qing Liao

**HIGHLIGHT:** However, they neglect to explore capturing the fine-grained task-specific interaction between stance detection and sentiment tasks, thus degrading performance. To address this issue, this paper proposes a novel multi-task interaction network (MTIN) for improving the performance of stance detection and sentiment analysis tasks simultaneously.

194, **TITLE:** Neural-based Mixture Probabilistic Query Embedding for Answering FOL queries on Knowledge Graphs  
<https://aclanthology.org/2022.emnlp-main.194>

**AUTHORS:** Xiao Long, Liansheng Zhuang, Li Aodi, Shafei Wang, Houqiang Li

**HIGHLIGHT:** In this paper, we propose a Neural-based Mixture Probabilistic Query Embedding Model (NMP-QEM) that encodes the answer set of each mini-query as a mixed Gaussian distribution with multiple means and covariance parameters, which can approximate any random distribution arbitrarily well in real KGs.

195, **TITLE:** Improving Multi-turn Emotional Support Dialogue Generation with Lookahead Strategy Planning  
<https://aclanthology.org/2022.emnlp-main.195>

**AUTHORS:** Yi Cheng, Wenge Liu, Wenjie Li, Jiashuo Wang, Ruihui Zhao, Bang Liu, Xiaodan Liang, Yefeng Zheng

**HIGHLIGHT:** In comparison, multi-turn ES conversation systems can provide ES more effectively, but face several new technical challenges, including: (1) how to adopt appropriate support strategies to achieve the long-term dialogue goal of comforting the user's emotion; (2) how to dynamically model the user's state. In this paper, we propose a novel system MultiESC to address these issues.

196, **TITLE:** Conformal Predictor for Improving Zero-Shot Text Classification Efficiency

<https://aclanthology.org/2022.emnlp-main.196>

**AUTHORS:** Prafulla Kumar Choubey, Yu Bai, Chien-Sheng Wu, Wenhao Liu, Nazneen Rajani

**HIGHLIGHT:** In this work, we improve the efficiency of such cross-encoder-based 0shot models by restricting the number of likely labels using another fast base classifier-based conformal predictor (CP) calibrated on samples labeled by the 0shot model.

197, **TITLE:** Effective and Efficient Query-aware Snippet Extraction for Web Search

<https://aclanthology.org/2022.emnlp-main.197>

**AUTHORS:** Jingwei Yi, Fangzhao Wu, Chuhan Wu, Xiaolong Huang, Binxing Jiao, Guangzhong Sun, Xing Xie

**HIGHLIGHT:** In this paper, we propose an effective query-aware webpage snippet extraction method named DeepQSE.

198, **TITLE:** You Only Need One Model for Open-domain Question Answering

<https://aclanthology.org/2022.emnlp-main.198>

**AUTHORS:** Haejun Lee, Akhil Kedia, Jongwon Lee, Ashwin Paranjape, Christopher Manning, Kyoung-Gu Woo

**HIGHLIGHT:** This allows us to use a single question answering model trained end-to-end, which is a more efficient use of model capacity and also leads to better gradient flow. We present a pre-training method to effectively train this architecture and evaluate our model on the Natural Questions and TriviaQA open datasets.

199, **TITLE:** Generative Entity Typing with Curriculum Learning

<https://aclanthology.org/2022.emnlp-main.199>

**AUTHORS:** Siyu Yuan, Deqing Yang, Jiaqing Liang, Zhixu Li, Jinxi Liu, Jingyue Huang, Yanghua Xiao

**HIGHLIGHT:** The traditional classification-based entity typing paradigm has two unignorable drawbacks: 1) it fails to assign an entity to the types beyond the predefined type set, and 2) it can hardly handle few-shot and zero-shot situations where many long-tail types only have few or even no training instances. To overcome these drawbacks, we propose a novel generative entity typing (GET) paradigm: given a text with an entity mention, the multiple types for the role that the entity plays in the text are generated with a pre-trained language model (PLM).

200, **TITLE:** SetGNER: General Named Entity Recognition as Entity Set Generation

<https://aclanthology.org/2022.emnlp-main.200>

**AUTHORS:** Yuxin He, Buzhou Tang

**HIGHLIGHT:** Motivated by the observation that the target output of NER is essentially a set of sequences, we propose a novel entity set generation framework for general NER scenes in this paper.

201, **TITLE:** Opinion Summarization by Weak-Supervision from Mix-structured Data

<https://aclanthology.org/2022.emnlp-main.201>

**AUTHORS:** Yizhu Liu, Qi Jia, Kenny Zhu

**HIGHLIGHT:** In this paper, we convert each review into a mix of structured and unstructured data, which we call opinion-aspect pairs (OAs) and implicit sentences (ISs).

202, **TITLE:** Multi-level Distillation of Semantic Knowledge for Pre-training Multilingual Language Model

<https://aclanthology.org/2022.emnlp-main.202>

**AUTHORS:** Mingqi Li, Fei Ding, Dan Zhang, Long Cheng, Hongxin Hu, Feng Luo

**HIGHLIGHT:** In this paper, we propose Multi-level Multilingual Knowledge Distillation (MMKD), a novel method for improving multilingual language models.

203, **TITLE:** Empowering Dual-Encoder with Query Generator for Cross-Lingual Dense Retrieval

<https://aclanthology.org/2022.emnlp-main.203>

**AUTHORS:** Houxing Ren, Linjun Shou, Ning Wu, Ming Gong, Daxin Jiang

**HIGHLIGHT:** In this paper, we propose to use a query generator as the teacher in the cross-lingual setting, which is less dependent on enough training samples and high-quality negative samples.

- 204, TITLE: R2F: A General Retrieval, Reading and Fusion Framework for Document-level Natural Language Inference  
<https://aclanthology.org/2022.emnlp-main.204>  
AUTHORS: Hao Wang, Yixin Cao, Yangguang Li, Zhen Huang, Kun Wang, Jing Shao  
HIGHLIGHT: In this paper, we establish a general solution, named Retrieval, Reading and Fusion (R2F) framework, and a new setting, by analyzing the main challenges of DOCNLI: interpretability, long-range dependency, and cross-sentence inference.
- 205, TITLE: Revisiting Pre-trained Language Models and their Evaluation for Arabic Natural Language Processing  
<https://aclanthology.org/2022.emnlp-main.205>  
AUTHORS: Abbas Ghaddar, Yimeng Wu, Sunyam Bagga, Ahmad Rashid, Khalil Bibi, Mehdi Rezagholizadeh, Chao Xing, Yasheng Wang, Xinyu Duan, Zhefeng Wang, Baoxing Huai, Xin Jiang, Qun Liu, Phillippe Langlais  
HIGHLIGHT: However, they are still limited due to redundant attack search spaces and the inability to defend against various types of attacks. In this work, we present a novel fine-tuning approach called ROBust SElective fine-tuning (ROSE) to address this issue.
- 206, TITLE: KECP: Knowledge Enhanced Contrastive Prompting for Few-shot Extractive Question Answering  
<https://aclanthology.org/2022.emnlp-main.206>  
AUTHORS: Jianing Wang, Chengyu Wang, Minghui Qiu, Qiuhui Shi, Hongbin Wang, Jun Huang, Ming Gao  
HIGHLIGHT: However, most existing approaches for MRC may perform poorly in the few-shot learning scenario. To solve this issue, we propose a novel framework named Knowledge Enhanced Contrastive Prompt-tuning (KECP).
- 207, TITLE: Knowledge Prompting in Pre-trained Language Model for Natural Language Understanding  
<https://aclanthology.org/2022.emnlp-main.207>  
AUTHORS: Jianing Wang, Wenkang Huang, Minghui Qiu, Qiuhui Shi, Hongbin Wang, Xiang Li, Ming Gao  
HIGHLIGHT: Then we design multiple continuous prompts rules and transform the knowledge sub-graph into natural language prompts. To further leverage the factual knowledge from these prompts, we propose two novel knowledge-aware self-supervised tasks including prompt relevance inspection and masked prompt modeling.
- 208, TITLE: On the Evaluation Metrics for Paraphrase Generation  
<https://aclanthology.org/2022.emnlp-main.208>  
AUTHORS: Lingfeng Shen, Lema Liu, Haiyun Jiang, Shuming Shi  
HIGHLIGHT: In this paper we revisit automatic metrics for paraphrase evaluation and obtain two findings that disobey conventional wisdom: (1) Reference-free metrics achieve better performance than their reference-based counterparts.
- 209, TITLE: Curriculum Learning Meets Weakly Supervised Multimodal Correlation Learning  
<https://aclanthology.org/2022.emnlp-main.209>  
AUTHORS: Sijie Mai, Ya Sun, Haifeng Hu  
HIGHLIGHT: In this paper, we inject curriculum learning into weakly supervised multimodal correlation learning.
- 210, TITLE: Rethinking Positional Encoding in Tree Transformer for Code Representation  
<https://aclanthology.org/2022.emnlp-main.210>  
AUTHORS: Han Peng, Ge Li, Yunfei Zhao, Zhi Jin  
HIGHLIGHT: In this work, we propose a novel tree Transformer encoding node positions based on our new description method for tree structures.
- 211, TITLE: RASAT: Integrating Relational Structures into Pretrained Seq2Seq Model for Text-to-SQL  
<https://aclanthology.org/2022.emnlp-main.211>  
AUTHORS: Jixing Qi, Jingyao Tang, Ziwei He, Xiangpeng Wan, Yu Cheng, Chenghu Zhou, Xinbing Wang, Quanshi Zhang, Zhouhan Lin  
HIGHLIGHT: However, introducing these structural relations comes with prices: they often result in a specialized model structure, which largely prohibits using large pretrained models in text-to-SQL. To address this problem, we propose RASAT: a Transformer seq2seq architecture augmented with relation-aware self-attention that could leverage a variety of relational structures while inheriting the pretrained parameters from the T5 model effectively.
- 212, TITLE: COM-MRC: A COntext-Masked Machine Reading Comprehension Framework for Aspect Sentiment Triplet Extraction  
<https://aclanthology.org/2022.emnlp-main.212>  
AUTHORS: Zepeng Zhai, Hao Chen, Fangxiang Feng, Ruifan Li, Xiaojie Wang  
HIGHLIGHT: In this paper, we propose a novel COntext-Masked MRC (COM-MRC) framework for ASTE.
- 213, TITLE: CEM: Machine-Human Chatting Handoff via Causal-Enhance Module  
<https://aclanthology.org/2022.emnlp-main.213>  
AUTHORS: Shanshan Zhong, Jinghui Qin, Zhongzhan Huang, Daifeng Li  
HIGHLIGHT: These variables are significantly associated with handoff decisions, resulting in prediction bias and cost increase. Therefore, we propose Causal-Enhance Module (CEM) by establishing the causal graph of MHCH based on these two variables, which is a simple yet effective module and can be easy to plug into the existing MHCH methods.
- 214, TITLE: Nearest Neighbor Zero-Shot Inference  
<https://aclanthology.org/2022.emnlp-main.214>  
AUTHORS: Weijia Shi, Julian Michael, Suchin Gururangan, Luke Zettlemoyer

**HIGHLIGHT:** Retrieval-augmented language models (LMs) use non-parametric memory to substantially outperform their non-retrieval counterparts on perplexity-based evaluations, but it is an open question whether they achieve similar gains in few- and zero-shot end-task accuracy. We extensively study one such model, the k-nearest neighbor LM (kNN-LM), showing that the gains marginally transfer.

215, **TITLE:** Robots-Dont-Cry: Understanding Falsely Anthropomorphic Utterances in Dialog Systems

<https://aclanthology.org/2022.emnlp-main.215>

**AUTHORS:** David Gros, Yu Li, Zhou Yu

**HIGHLIGHT:** We collect human ratings on the feasibility of approximately 900 two-turn dialogs sampled from 9 diverse data sources.

216, **TITLE:** A Joint Learning Framework for Restaurant Survival Prediction and Explanation

<https://aclanthology.org/2022.emnlp-main.216>

**AUTHORS:** Xin Li, Xiaojie Zhang, Peng JiaHao, Rui Mao, Mingyang Zhou, Xing Xie, Hao Liao

**HIGHLIGHT:** In this paper, we tackle the practical problem of restaurant survival prediction.

217, **TITLE:** Making Pretrained Language Models Good Long-tailed Learners

<https://aclanthology.org/2022.emnlp-main.217>

**AUTHORS:** Chen Zhang, Lei Ren, Jingang Wang, Wei Wu, Dawei Song

**HIGHLIGHT:** This motivates us to check the hypothesis that prompt-tuning is also a promising choice for long-tailed classification, since the tail classes are intuitively few-shot ones. To achieve this aim, we conduct empirical studies to examine the hypothesis.

218, **TITLE:** UniGeo: Unifying Geometry Logical Reasoning via Reformulating Mathematical Expression

<https://aclanthology.org/2022.emnlp-main.218>

**AUTHORS:** Jiaqi Chen, Tong Li, Jinghui Qin, Pan Lu, Liang Lin, Chongyu Chen, Xiaodan Liang

**HIGHLIGHT:** However, in essence, these two tasks have similar problem representations and overlapped math knowledge which can improve the understanding and reasoning ability of a deep model on both two tasks. Therefore, we construct a large-scale Unified Geometry problem benchmark, UniGeo, which contains 4,998 calculation problems and 9,543 proving problems.

219, **TITLE:** Face-Sensitive Image-to-Emotional-Text Cross-modal Translation for Multimodal Aspect-based Sentiment Analysis

<https://aclanthology.org/2022.emnlp-main.219>

**AUTHORS:** Hao Yang, Yanyan Zhao, Bing Qin

**HIGHLIGHT:** In this work, we introduce a face-sensitive image-to-emotional-text translation (FITE) method, which focuses on capturing visual sentiment cues through facial expressions and selectively matching and fusing with the target aspect in textual modality.

220, **TITLE:** FineD-Eval: Fine-grained Automatic Dialogue-Level Evaluation

<https://aclanthology.org/2022.emnlp-main.220>

**AUTHORS:** Chen Zhang, Luis Fernando D'Haro, Qiquan Zhang, Thomas Friedrichs, Haizhou Li

**HIGHLIGHT:** To this end, we are motivated to propose a multi-dimensional dialogue-level metric, which consists of three sub-metrics with each targeting a specific dimension.

221, **TITLE:** Sentence Representation Learning with Generative Objective rather than Contrastive Objective

<https://aclanthology.org/2022.emnlp-main.221>

**AUTHORS:** Bohong Wu, Hai Zhao

**HIGHLIGHT:** We instead propose a novel generative self-supervised learning objective based on phrase reconstruction.

222, **TITLE:** RLPrompt: Optimizing Discrete Text Prompts with Reinforcement Learning

<https://aclanthology.org/2022.emnlp-main.222>

**AUTHORS:** Mingkai Deng, Jianyu Wang, Cheng-Ping Hsieh, Yihan Wang, Han Guo, Tianmin Shu, Meng Song, Eric Xing, Zhiting Hu

**HIGHLIGHT:** This paper proposes RLPrompt, an efficient discrete prompt optimization approach with reinforcement learning (RL).

223, **TITLE:** DisCup: Discriminator Cooperative Unlikelihood Prompt-tuning for Controllable Text Generation

<https://aclanthology.org/2022.emnlp-main.223>

**AUTHORS:** Hanqing Zhang, Dawei Song

**HIGHLIGHT:** In this paper, we propose a new CTG approach, namely DisCup, which incorporates the attribute knowledge of discriminator to optimize the control-prompts, steering a frozen CLM to produce attribute-specific texts.

224, **TITLE:** CPL: Counterfactual Prompt Learning for Vision and Language Models

<https://aclanthology.org/2022.emnlp-main.224>

**AUTHORS:** Xuehai He, Diji Yang, Weixi Feng, Tsu-Jui Fu, Arjun Akula, Varun Jampani, Pradyumna Narayana, Sugato Basu, William Yang Wang, Xin Wang

**HIGHLIGHT:** Towards non-spurious and efficient prompt learning from limited examples, this paper presents a novel Counterfactual Prompt Learning (CPL) method for vision and language models, which simultaneously employs counterfactual generation and contrastive learning in a joint optimization framework.

- 225, TITLE: Red Teaming Language Models with Language Models  
<https://aclanthology.org/2022.emnlp-main.225>  
AUTHORS: Ethan Perez, Saffron Huang, Francis Song, Trevor Cai, Roman Ring, John Aslanides, Amelia Glaese, Nat McAleese, Geoffrey Irving  
HIGHLIGHT: In this work, we automatically find cases where a target LM behaves in a harmful way, by generating test cases (?red teaming?) using another LM.
- 226, TITLE: CapOnImage: Context-driven Dense-Captioning on Image  
<https://aclanthology.org/2022.emnlp-main.226>  
AUTHORS: Yiqi Gao, Xinglin Hou, Yuanmeng Zhang, Tiezheng Ge, Yuning Jiang, Peng Wang  
HIGHLIGHT: In this work, we introduce a new task called captioning on image (CapOnImage), which aims to generate dense captions at different locations of the image based on contextual information.
- 227, TITLE: SpanProto: A Two-stage Span-based Prototypical Network for Few-shot Named Entity Recognition  
<https://aclanthology.org/2022.emnlp-main.227>  
AUTHORS: Jianing Wang, Chengyu Wang, Chuanqi Tan, Minghui Qiu, Songfang Huang, Jun Huang, Ming Gao  
HIGHLIGHT: To this end, we propose a seminal span-based prototypical network (SpanProto) that tackles few-shot NER via a two-stage approach, including span extraction and mention classification.
- 228, TITLE: Discovering Differences in the Representation of People using Contextualized Semantic Axes  
<https://aclanthology.org/2022.emnlp-main.228>  
AUTHORS: Li Lucy, Divya Tadimeti, David Bamman  
HIGHLIGHT: In particular, past work has compared embeddings against “semantic axes” that represent two opposing concepts. We extend this paradigm to BERT embeddings, and construct contextualized axes that mitigate the pitfall where antonyms have neighboring representations.
- 229, TITLE: Generating Literal and Implied Subquestions to Fact-check Complex Claims  
<https://aclanthology.org/2022.emnlp-main.229>  
AUTHORS: Jifan Chen, Aniruddh Sriram, Eunsol Choi, Greg Durrett  
HIGHLIGHT: In this work, we focus on decomposing a complex claim into a comprehensive set of yes-no subquestions whose answers influence the veracity of the claim.
- 230, TITLE: Machine Translation Robustness to Natural Asemantic Variation  
<https://aclanthology.org/2022.emnlp-main.230>  
AUTHORS: Jacob Bremerman, Xiang Ren, Jonathan May  
HIGHLIGHT: An important yet under-studied category involves minor variations in nuance (non-typos) that preserve meaning w.r.t. the target language. We introduce and formalize this category as Natural Asemantic Variation (NAV) and investigate it in the context of MT robustness.
- 231, TITLE: Natural Language to Code Translation with Execution  
<https://aclanthology.org/2022.emnlp-main.231>  
AUTHORS: Freda Shi, Daniel Fried, Marjan Ghazvininejad, Luke Zettlemoyer, Sida I. Wang  
HIGHLIGHT: In this work, we introduce execution result-based minimum Bayes risk decoding (MBR-EXEC) for program selection and show that it improves the few-shot performance of pretrained code models on natural-language-to-code tasks.
- 232, TITLE: Life is a Circus and We are the Clowns: Automatically Finding Analogies between Situations and Processes  
<https://aclanthology.org/2022.emnlp-main.232>  
AUTHORS: Oren Sultan, Dafna Shahaf  
HIGHLIGHT: Our goal is to automatically extract entities and their relations from the text and find a mapping between the different domains based on relational similarity (e.g., blood is mapped to water).
- 233, TITLE: Language Contamination Helps Explain the Cross-lingual Capabilities of English Pretrained Models  
<https://aclanthology.org/2022.emnlp-main.233>  
AUTHORS: Terra Blevins, Luke Zettlemoyer  
HIGHLIGHT: These models are generally presented as being trained only on English text but have been found to transfer surprisingly well to other languages. We investigate this phenomenon and find that common English pretraining corpora actually contain significant amounts of non-English text: even when less than 1% of data is not English (well within the error rate of strong language classifiers), this leads to hundreds of millions of foreign language tokens in large-scale datasets.
- 234, TITLE: Analyzing the Mono- and Cross-Lingual Pretraining Dynamics of Multilingual Language Models  
<https://aclanthology.org/2022.emnlp-main.234>  
AUTHORS: Terra Blevins, Hila Gonen, Luke Zettlemoyer  
HIGHLIGHT: However, because these analyses have focused on fully trained multilingual models, little is known about the dynamics of the multilingual pretraining process. We investigate when these models acquire their in-language and cross-lingual abilities by probing checkpoints taken from throughout XLM-R pretraining, using a suite of linguistic tasks.
- 235, TITLE: Neural Machine Translation with Contrastive Translation Memories  
<https://aclanthology.org/2022.emnlp-main.235>  
AUTHORS: Xin Cheng, Shen Gao, Lemao Liu, Dongyan Zhao, Rui Yan

**HIGHLIGHT:** Different from previous works that make use of mutually similar but redundant translation memories (TMs), we propose a new retrieval-augmented NMT to model contrastively retrieved translation memories that are holistically similar to the source sentence while individually contrastive to each other providing maximal information gain in three phases.

236, **TITLE:** Distilling Causal Effect from Miscellaneous Other-Class for Continual Named Entity Recognition  
<https://aclanthology.org/2022.emnlp-main.236>

**AUTHORS:** Junhao Zheng, Zhanxian Liang, Haibin Chen, Qianli Ma

**HIGHLIGHT:** Thanks to the causal inference, we identify that the forgetting is caused by the missing causal effect from the old data. To this end, we propose a unified causal framework to retrieve the causality from both new entity types and Other-Class.

237, **TITLE:** Exploring the Secrets Behind the Learning Difficulty of Meaning Representations for Semantic Parsing  
<https://aclanthology.org/2022.emnlp-main.237>

**AUTHORS:** Zhenwen Li, Jiaqi Guo, Qian Liu, Jian-Guang Lou, Tao Xie

**HIGHLIGHT:** In this paper, we propose a data-aware metric called ISS (denoting incremental structural stability) of MRs, and demonstrate that ISS is highly correlated with the final performance.

238, **TITLE:** That's the Wrong Lung! Evaluating and Improving the Interpretability of Unsupervised Multimodal Encoders for Medical Data  
<https://aclanthology.org/2022.emnlp-main.238>

**AUTHORS:** Jered McInerney, Geoffrey Young, Jan-Willem van de Meent, Byron Wallace

**HIGHLIGHT:** Our main finding is that the text has an often weak or unintuitive influence on attention; alignments do not consistently reflect basic anatomical information.

239, **TITLE:** Unsupervised Tokenization Learning  
<https://aclanthology.org/2022.emnlp-main.239>

**AUTHORS:** Anton Kolonin, Vignav Ramesh

**HIGHLIGHT:** In the presented study, we discover that the so-called 'transition freedom' metric appears superior for unsupervised tokenization purposes in comparison to statistical metrics such as mutual information and conditional probability, providing F-measure scores in range from 0.71 to 1.0 across explored multilingual corpora.

240, **TITLE:** A Template-based Method for Constrained Neural Machine Translation  
<https://aclanthology.org/2022.emnlp-main.240>

**AUTHORS:** Shuo Wang, Peng Li, Zhixing Tan, Zhaopeng Tu, Maosong Sun, Yang Liu

**HIGHLIGHT:** In this work, we propose a template-based method that can yield results with high translation quality and match accuracy and the inference speed of our method is comparable with unconstrained NMT models.

241, **TITLE:** PATS: Sensitivity-aware Noisy Learning for Pretrained Language Models  
<https://aclanthology.org/2022.emnlp-main.241>

**AUTHORS:** Yupeng Zhang, Hongzhi Zhang, Sirui Wang, Wei Wu, Zhoujun Li

**HIGHLIGHT:** In this paper, we present PATS (Perturbation According To Sensitivity), a noisy training mechanism which considers each parameter's importance in the downstream task to help fine-tune PLMs.

242, **TITLE:** Towards Reinterpreting Neural Topic Models via Composite Activations  
<https://aclanthology.org/2022.emnlp-main.242>

**AUTHORS:** Jia Peng Lim, Hady Lauw

**HIGHLIGHT:** In this paper, we present a model-free two-stage process to reinterpret NTM and derive further insights on the state of the trained model.

243, **TITLE:** Few-shot Query-Focused Summarization with Prefix-Merging  
<https://aclanthology.org/2022.emnlp-main.243>

**AUTHORS:** Ruifeng Yuan, Zili Wang, Ziqiang Cao, Wenjie Li

**HIGHLIGHT:** In this paper, we investigate the idea that whether we can integrate and transfer the knowledge of text summarization and question answering to assist the few-shot learning in query-focused summarization.

244, **TITLE:** Cross-Align: Modeling Deep Cross-lingual Interactions for Word Alignment  
<https://aclanthology.org/2022.emnlp-main.244>

**AUTHORS:** Siyu Lai, Zhen Yang, Fandong Meng, Yufeng Chen, Jinan Xu, Jie Zhou

**HIGHLIGHT:** However, we find that the existing approaches capture few interactions between the input sentence pairs, which degrades the word alignment quality severely, especially for the ambiguous words in the monolingual context. To remedy this problem, we propose Cross-Align to model deep interactions between the input sentence pairs, in which the source and target sentences are encoded separately with the shared self-attention modules in the shallow layers, while cross-lingual interactions are explicitly constructed by the cross-attention modules in the upper layers.

245, **TITLE:** BERTScore is Unfair: On Social Bias in Language Model-Based Metrics for Text Generation  
<https://aclanthology.org/2022.emnlp-main.245>

**AUTHORS:** Tianxiang Sun, Junliang He, Xipeng Qiu, Xuanjing Huang

**HIGHLIGHT:** To that end, this work presents the first systematic study on the social bias in PLM-based metrics.

246, **TITLE:** HPT: Hierarchy-aware Prompt Tuning for Hierarchical Text Classification  
<https://aclanthology.org/2022.emnlp-main.246>

- AUTHORS: Zihan Wang, Peiyi Wang, Tianyu Liu, Binghuai Lin, Yunbo Cao, Zhifang Sui, Houfeng Wang  
HIGHLIGHT: To bridge the gap, in this paper, we propose HPT, a Hierarchy-aware Prompt Tuning method to handle HTC from a multi-label MLM perspective.
- 247, TITLE: Not to Overfit or Underfit the Source Domains? An Empirical Study of Domain Generalization in Question Answering  
<https://aclanthology.org/2022.emnlp-main.247>  
AUTHORS: Md Arafat Sultan, Avi Sil, Radu Florian  
HIGHLIGHT: Here we examine the contrasting view that multi-source domain generalization (DG) is first and foremost a problem of mitigating source domain underfitting: models not adequately learning the signal already present in their multi-domain training data.
- 248, TITLE: Neural Theory-of-Mind? On the Limits of Social Intelligence in Large LMs  
<https://aclanthology.org/2022.emnlp-main.248>  
AUTHORS: Maarten Sap, Ronan Le Bras, Daniel Fried, Yejin Choi  
HIGHLIGHT: In this work, we examine the open question of social intelligence and Theory of Mind in modern NLP systems from an empirical and theorybased perspective.
- 249, TITLE: Improving Passage Retrieval with Zero-Shot Question Generation  
<https://aclanthology.org/2022.emnlp-main.249>  
AUTHORS: Devendra Sachan, Mike Lewis, Mandar Joshi, Armen Aghajanyan, Wen-tau Yih, Joelle Pineau, Luke Zettlemoyer  
HIGHLIGHT: We propose a simple and effective re-ranking method for improving passage retrieval in open question answering.
- 250, TITLE: Summarizing Community-based Question-Answer Pairs  
<https://aclanthology.org/2022.emnlp-main.250>  
AUTHORS: Ting-Yao Hsu, Yoshi Suhara, Xiaolan Wang  
HIGHLIGHT: To help users quickly digest the key information, we propose the novel CQA summarization task that aims to create a concise summary from CQA pairs.
- 251, TITLE: Logical Reasoning with Span-Level Predictions for Interpretable and Robust NLI Models  
<https://aclanthology.org/2022.emnlp-main.251>  
AUTHORS: Joe Stacey, Pasquale Minervini, Haim Dubossarsky, Marek Rei  
HIGHLIGHT: Unlike prior work, we show that improved interpretability can be achieved without decreasing the predictive accuracy.
- 252, TITLE: How to disagree well: Investigating the dispute tactics used on Wikipedia  
<https://aclanthology.org/2022.emnlp-main.252>  
AUTHORS: Christine De Kock, Andreas Vlachos  
HIGHLIGHT: Disagreements are frequently studied from the perspective of either detecting toxicity or analysing argument structure. We propose a framework of dispute tactics which unifies these two perspectives, as well as other dialogue acts which play a role in resolving disputes, such as asking questions and providing clarification.
- 253, TITLE: Chapter Ordering in Novels  
<https://aclanthology.org/2022.emnlp-main.253>  
AUTHORS: Allen Kim, Steve Skiena  
HIGHLIGHT: Understanding narrative flow and text coherence in long-form documents (novels) remains an open problem in NLP. To gain insight, we explore the task of chapter ordering, reconstructing the original order of chapters in novel given a random permutation of the text.
- 254, TITLE: Open-ended Knowledge Tracing for Computer Science Education  
<https://aclanthology.org/2022.emnlp-main.254>  
AUTHORS: Naiming Liu, Zichao Wang, Richard Baraniuk, Andrew Lan  
HIGHLIGHT: We develop an initial solution to the OKT problem, a student knowledge-guided code generation approach, that combines program synthesis methods using language models with student knowledge tracing methods.
- 255, TITLE: Logical Neural Networks for Knowledge Base Completion with Embeddings & Rules  
<https://aclanthology.org/2022.emnlp-main.255>  
AUTHORS: Prithviraj Sen, Breno William Carvalho, Ibrahim Abdelaziz, Pavan Kapanipathi, Salim Roukos, Alexander Gray  
HIGHLIGHT: To this end, we propose to utilize logical neural networks (LNN), a powerful neuro-symbolic AI framework that can express both kinds of rules and learn these end-to-end using gradient-based optimization.
- 256, TITLE: MedCLIP: Contrastive Learning from Unpaired Medical Images and Text  
<https://aclanthology.org/2022.emnlp-main.256>  
AUTHORS: Zifeng Wang, Zhenbang Wu, Dinesh Agarwal, Jimeng Sun  
HIGHLIGHT: In this paper, we decouple images and texts for multimodal contrastive learning, thus scaling the usable training data in a combinatorial magnitude with low cost.



- 257, TITLE: GA-SAM: Gradient-Strength based Adaptive Sharpness-Aware Minimization for Improved Generalization  
<https://aclanthology.org/2022.emnlp-main.257>  
AUTHORS: Zhiyuan Zhang, Ruixuan Luo, Qi Su, Xu Sun  
HIGHLIGHT: However, it has some difficulty implying SAM to some natural language tasks, especially to models with drastic gradient changes, such as RNNs. In this work, we analyze the relation between the flatness of the local minimum and its generalization ability from a novel and straightforward theoretical perspective.
- 258, TITLE: Sparse Teachers Can Be Dense with Knowledge  
<https://aclanthology.org/2022.emnlp-main.258>  
AUTHORS: Yi Yang, Chen Zhang, Dawei Song  
HIGHLIGHT: To remove the parameters that result in student-unfriendliness, we propose a sparse teacher trick under the guidance of an overall knowledgeable score for each teacher parameter.
- 259, TITLE: BBTv2: Towards a Gradient-Free Future with Large Language Models  
<https://aclanthology.org/2022.emnlp-main.259>  
AUTHORS: Tianxiang Sun, Zhengfu He, Hong Qian, Yunhua Zhou, Xuanjing Huang, Xipeng Qiu  
HIGHLIGHT: In this paper, we present BBTv2, an improved version of Black-Box Tuning, to drive PTMs for few-shot learning.
- 260, TITLE: Passage-Mask: A Learnable Regularization Strategy for Retriever-Reader Models  
<https://aclanthology.org/2022.emnlp-main.260>  
AUTHORS: Shujian Zhang, Chengyue Gong, Xingchao Liu  
HIGHLIGHT: Retriever-reader models achieve competitive performance across many different NLP tasks such as open question answering and dialogue conversations. In this work, we notice these models easily overfit the top-rank retrieval passages and standard training fails to reason over the entire retrieval passages.
- 261, TITLE: Mixed-effects transformers for hierarchical adaptation  
<https://aclanthology.org/2022.emnlp-main.261>  
AUTHORS: Julia White, Noah Goodman, Robert Hawkins  
HIGHLIGHT: In this paper, we introduce the mixed-effects transformer (MET), a novel approach for learning hierarchically-structured prefixes? lightweight modules prepended to an input sequence? to account for structured variation in language use.
- 262, TITLE: On Measuring the Intrinsic Few-Shot Hardness of Datasets  
<https://aclanthology.org/2022.emnlp-main.262>  
AUTHORS: Xinran Zhao, Shikhar Murty, Christopher Manning  
HIGHLIGHT: We consider an extensive set of recent few-shot learning methods and show that their performance across a large number of datasets is highly correlated, showing that few-shot hardness may be intrinsic to datasets, for a given pre-trained model.
- 263, TITLE: Group is better than individual: Exploiting Label Topologies and Label Relations for Joint Multiple Intent Detection and Slot Filling  
<https://aclanthology.org/2022.emnlp-main.263>  
AUTHORS: Bowen Xing, Ivor Tsang  
HIGHLIGHT: Therefore, in this paper, we first construct a Heterogeneous Label Graph (HLG) containing two kinds of topologies: (1) statistical dependencies based on labelsâ€™ co-occurrence patterns and hierarchies in slot labels; (2) rich relations among the label nodes. Then we propose a novel model termed ReLa-Net. It can capture beneficial correlations among the labels from HLG.
- 264, TITLE: An Empirical Study on Finding Spans  
<https://aclanthology.org/2022.emnlp-main.264>  
AUTHORS: Weiwei Gu, Boyuan Zheng, Yunmo Chen, Tongfei Chen, Benjamin Van Durme  
HIGHLIGHT: We present an empirical study on methods for span finding, the selection of consecutive tokens in text for some downstream tasks.
- 265, TITLE: MGDoc: Pre-training with Multi-granular Hierarchy for Document Image Understanding  
<https://aclanthology.org/2022.emnlp-main.265>  
AUTHORS: Zilong Wang, Jiuxiang Gu, Chris Tensmeyer, Nikolaos Barmpalios, Ani Nenkova, Tong Sun, Jingbo Shang, Vlad Morariu  
HIGHLIGHT: In contrast, region-level models attempt to encode regions corresponding to paragraphs or text blocks into a single embedding, but they perform worse with additional word-level features. To deal with these issues, we propose MGDoc, a new multi-modal multi-granular pre-training framework that encodes page-level, region-level, and word-level information at the same time.
- 266, TITLE: Understanding Jargon: Combining Extraction and Generation for Definition Modeling  
<https://aclanthology.org/2022.emnlp-main.266>  
AUTHORS: Jie Huang, Hanyin Shao, Kevin Chen-Chuan Chang, Jinjun Xiong, Wen-mei Hwu  
HIGHLIGHT: In this paper, we propose to combine extraction and generation for jargon definition modeling: first extract self- and correlative definitional information of target jargon from the Web and then generate the final definitions by incorporating the extracted definitional information.

- 267, TITLE: ProsocialDialog: A Prosocial Backbone for Conversational Agents  
<https://aclanthology.org/2022.emnlp-main.267>  
AUTHORS: Hyunwoo Kim, Youngjae Yu, Liwei Jiang, Ximing Lu, Daniel Khashabi, Gunhee Kim, Yejin Choi, Maarten Sap  
HIGHLIGHT: Most existing dialogue systems fail to respond properly to potentially unsafe user utterances by either ignoring or passively agreeing with them. To address this issue, we introduce ProsocialDialog, the first large-scale multi-turn dialogue dataset to teach conversational agents to respond to problematic content following social norms.
- 268, TITLE: Exploiting Global and Local Hierarchies for Hierarchical Text Classification  
<https://aclanthology.org/2022.emnlp-main.268>  
AUTHORS: Ting Jiang, Deqing Wang, Leilei Sun, Zhongzhi Chen, Fuzhen Zhuang, Qinghong Yang  
HIGHLIGHT: To exploit global and local hierarchies, we propose Hierarchy-guided BERT with Global and Local hierarchies (HBGL), which utilizes the large-scale parameters and prior language knowledge of BERT to model both global and local hierarchies.
- 269, TITLE: Semantic-aware Contrastive Learning for More Accurate Semantic Parsing  
<https://aclanthology.org/2022.emnlp-main.269>  
AUTHORS: Shan Wu, Chunlei Xin, Bo Chen, Xianpei Han, Le Sun  
HIGHLIGHT: In this paper, we propose a semantic-aware contrastive learning algorithm, which can learn to distinguish fine-grained meaning representations and take the overall sequence-level semantic into consideration.
- 270, TITLE: Scientific Paper Extractive Summarization Enhanced by Citation Graphs  
<https://aclanthology.org/2022.emnlp-main.270>  
AUTHORS: Xiuying Chen, Mingzhe Li, Shen Gao, Rui Yan, Xin Gao, Xiangliang Zhang  
HIGHLIGHT: In this work, we focus on leveraging citation graphs to improve scientific paper extractive summarization under different settings.
- 271, TITLE: Hardness-guided domain adaptation to recognise biomedical named entities under low-resource scenarios  
<https://aclanthology.org/2022.emnlp-main.271>  
AUTHORS: Ngoc Dang Nguyen, Lan Du, Wray Buntine, Changyou Chen, Richard Beare  
HIGHLIGHT: In this paper, we present a simple yet effective hardness-guided domain adaptation framework for bioNER tasks that can effectively leverage the domain hardness information to improve the adaptability of the learnt model in the low-resource scenarios.
- 272, TITLE: Syntactic Multi-view Learning for Open Information Extraction  
<https://aclanthology.org/2022.emnlp-main.272>  
AUTHORS: Kuicai Dong, Aixin Sun, Jung-Jae Kim, Xiaoli Li  
HIGHLIGHT: In this paper, we model both constituency and dependency trees into word-level graphs, and enable neural OpenIE to learn from the syntactic structures.
- 273, TITLE: TRIPS: Efficient Vision-and-Language Pre-training with Text-Relevant Image Patch Selection  
<https://aclanthology.org/2022.emnlp-main.273>  
AUTHORS: Chaoya Jiang, Haiyang Xu, Chenliang Li, Ming Yan, Wei Ye, Shikun Zhang, Bin Bi, Songfang Huang  
HIGHLIGHT: Though previous VLP works have proved the effectiveness of ViTs, they still suffer from computational efficiency brought by the long visual sequence. To tackle this problem, in this paper, we propose an efficient vision-and-language pre-training model with Text-Relevant Image Patch Selection, namely TRIPS, which reduces the visual sequence progressively with a text-guided patch-selection layer in the visual backbone for efficient training and inference.
- 274, TITLE: CGoDial: A Large-Scale Benchmark for Chinese Goal-oriented Dialog Evaluation  
<https://aclanthology.org/2022.emnlp-main.274>  
AUTHORS: Yinpei Dai, Wanwei He, Bowen Li, Yuchuan Wu, Zheng Cao, Zhongqi An, Jian Sun, Yongbin Li  
HIGHLIGHT: Practical dialog systems need to deal with various knowledge sources, noisy user expressions, and the shortage of annotated data. To better solve the above problems, we propose CGoDial, a new challenging and comprehensive Chinese benchmark for multi-domain Goal-oriented Dialog evaluation.
- 275, TITLE: Kernel-Whitening: Overcome Dataset Bias with Isotropic Sentence Embedding  
<https://aclanthology.org/2022.emnlp-main.275>  
AUTHORS: SongYang Gao, Shihan Dou, Qi Zhang, Xuanjing Huang  
HIGHLIGHT: However, such two-stage methods scale up the computational complexity of training process and obstruct valid feature information while mitigating bias. To address this issue, we utilize the representation normalization method which aims at disentangling the correlations between features of encoded sentences.
- 276, TITLE: A Unified Positive-Unlabeled Learning Framework for Document-Level Relation Extraction with Different Levels of Labeling  
<https://aclanthology.org/2022.emnlp-main.276>  
AUTHORS: Ye Wang, Xinxin Liu, Wenxin Hu, Tao Zhang  
HIGHLIGHT: To solve the common incomplete labeling problem, we propose a unified positive-unlabeled learning framework - shift and squared ranking loss positive-unlabeled (SSR-PU) learning.
- 277, TITLE: Automatic Generation of Socratic Subquestions for Teaching Math Word Problems  
<https://aclanthology.org/2022.emnlp-main.277>

- AUTHORS:** Kumar Shridhar, Jakub Macina, Mennatallah El-Assady, Tanmay Sinha, Manu Kapur, Mrinmaya Sachan  
**HIGHLIGHT:** We propose various guided question generation schemes based on input conditioning and reinforcement learning.
- 278, **TITLE:** Mixture of Attention Heads: Selecting Attention Heads Per Token  
<https://aclanthology.org/2022.emnlp-main.278>  
**AUTHORS:** Xiaofeng Zhang, Yikang Shen, Zeyu Huang, Jie Zhou, Wenge Rong, Zhang Xiong  
**HIGHLIGHT:** This paper proposes the Mixture of Attention Heads (MoA), a new architecture that combines multi-head attention with the MoE mechanism.
- 279, **TITLE:** The Optimal BERT Surgeon: Scalable and Accurate Second-Order Pruning for Large Language Models  
<https://aclanthology.org/2022.emnlp-main.279>  
**AUTHORS:** Eldar Kurtic, Daniel Campos, Tuan Nguyen, Elias Frantar, Mark Kurtz, Benjamin Fineran, Michael Goin, Dan Alistarh  
**HIGHLIGHT:** In this paper, we consider the problem of sparsifying BERT models, which are a key building block for natural language processing, in order to reduce their storage and computational cost.
- 280, **TITLE:** Information-Theoretic Text Hallucination Reduction for Video-grounded Dialogue  
<https://aclanthology.org/2022.emnlp-main.280>  
**AUTHORS:** Sunjae Yoon, Eunseop Yoon, Hee Suk Yoon, Junyeong Kim, Chang Yoo  
**HIGHLIGHT:** Hence, we design Text Hallucination Mitigating (THAM) framework, which incorporates Text Hallucination Regularization (THR) loss derived from the proposed information-theoretic text hallucination measurement approach.
- 281, **TITLE:** DSM: Question Generation over Knowledge Base via Modeling Diverse Subgraphs with Meta-learner  
<https://aclanthology.org/2022.emnlp-main.281>  
**AUTHORS:** Shasha Guo, Jing Zhang, Yanling Wang, Qianyi Zhang, Cuiping Li, Hong Chen  
**HIGHLIGHT:** In this work, we show that making use of the past experience on semantically similar subgraphs can reduce the learning difficulty and promote the performance of KBQG models. To achieve this, we propose a novel approach to model diverse subgraphs with meta-learner (DSM).
- 282, **TITLE:** ReLU-Net: Syntax-aware Graph U-Net for Relational Triple Extraction  
<https://aclanthology.org/2022.emnlp-main.282>  
**AUTHORS:** Yunqi Zhang, Yubo Chen, Yongfeng Huang  
**HIGHLIGHT:** This is due to the absence of entity locations, which is the prerequisite for pruning noisy edges from the dependency tree, when extracting relational triples. In this paper, we propose a unified framework to tackle this challenge and incorporate syntactic information for relational triple extraction.
- 283, **TITLE:** Evidence > Intuition: Transferability Estimation for Encoder Selection  
<https://aclanthology.org/2022.emnlp-main.283>  
**AUTHORS:** Elisa Bagnina, Max Müller-Eberstein, Mike Zhang, Barbara Plank  
**HIGHLIGHT:** In this paper, we propose to generate quantitative evidence to predict which LM, out of a pool of models, will perform best on a target task without having to fine-tune all candidates.
- 284, **TITLE:** Chunk-based Nearest Neighbor Machine Translation  
<https://aclanthology.org/2022.emnlp-main.284>  
**AUTHORS:** Pedro Henrique Martins, Zita Marinho, André F. T. Martins  
**HIGHLIGHT:** In this paper, we introduce a chunk-based kNN-MT model which retrieves chunks of tokens from the datastore, instead of a single token.
- 285, **TITLE:** FiE: Building a Global Probability Space by Leveraging Early Fusion in Encoder for Open-Domain Question Answering  
<https://aclanthology.org/2022.emnlp-main.285>  
**AUTHORS:** Akhil Kedia, Mohd Abbas Zaidi, Haejun Lee  
**HIGHLIGHT:** We propose to extend transformer encoders with the ability to fuse information from multiple passages, using global representation to provide cross-sample attention over all tokens across samples.
- 286, **TITLE:** Inductive Relation Prediction with Logical Reasoning Using Contrastive Representations  
<https://aclanthology.org/2022.emnlp-main.286>  
**AUTHORS:** Yudai Pan, Jun Liu, Lingling Zhang, Tianzhe Zhao, Qika Lin, Xin Hu, Qianying Wang  
**HIGHLIGHT:** To this end, we propose a novel graph convolutional network (GCN)-based model LogCo with logical reasoning by contrastive representations.
- 287, **TITLE:** Improving Chinese Spelling Check by Character Pronunciation Prediction: The Effects of Adaptivity and Granularity  
<https://aclanthology.org/2022.emnlp-main.287>  
**AUTHORS:** Jiahao Li, Quan Wang, Zhendong Mao, Junbo Guo, Yanyan Yang, Yongdong Zhang  
**HIGHLIGHT:** As most of these spelling errors are caused by phonetic similarity, effectively modeling the pronunciation of Chinese characters is a key factor for CSC. In this paper, we consider introducing an auxiliary task of Chinese pronunciation prediction (CPP) to improve CSC, and, for the first time, systematically discuss the adaptivity and granularity of this auxiliary task.

- 288, TITLE: MT-GenEval: A Counterfactual and Contextual Dataset for Evaluating Gender Accuracy in Machine Translation  
<https://aclanthology.org/2022.emnlp-main.288>  
AUTHORS: Anna Currey, Maria Nadejde, Raghavendra Reddy Pappagari, Mia Mayer, Stanislas Lauly, Xing Niu, Benjamin Hsu, Georgiana Dinu  
HIGHLIGHT: In this paper, we introduce MT-GenEval, a benchmark for evaluating gender accuracy in translation from English into eight widely-spoken languages.
- 289, TITLE: A Span-level Bidirectional Network for Aspect Sentiment Triplet Extraction  
<https://aclanthology.org/2022.emnlp-main.289>  
AUTHORS: Yuqi Chen, Chen Keming, Xian Sun, Zequn Zhang  
HIGHLIGHT: In this paper, we present a span-level bidirectional network which utilizes all possible spans as input and extracts triplets from spans bidirectionally.
- 290, TITLE: On the Calibration of Massively Multilingual Language Models  
<https://aclanthology.org/2022.emnlp-main.290>  
AUTHORS: Kabir Ahuja, Sunayana Sitaram, Sandipan Dandapat, Monojit Choudhury  
HIGHLIGHT: Overall, our work contributes towards building more reliable multilingual models by highlighting the issue of their miscalibration, understanding what language and model-specific factors influence it, and pointing out the strategies to improve the same.
- 291, TITLE: Momentum Contrastive Pre-training for Question Answering  
<https://aclanthology.org/2022.emnlp-main.291>  
AUTHORS: Minda Hu, Muzhi Li, Yasheng Wang, Irwin King  
HIGHLIGHT: Existing pre-training methods for extractive Question Answering (QA) generate cloze-like queries different from natural questions in syntax structure, which could overfit pre-trained models to simple keyword matching. In order to address this problem, we propose a novel Momentum Contrastive pRe-training fOr queStion anSwering (MCROSS) method for extractive QA.
- 292, TITLE: A Second Wave of UD Hebrew Treebanking and Cross-Domain Parsing  
<https://aclanthology.org/2022.emnlp-main.292>  
AUTHORS: Amir Zeldes, Nick Howell, Noam Ordan, Yifat Ben Moshe  
HIGHLIGHT: This paper presents a new, freely available UD treebank of Hebrew stratified from a range of topics selected from Hebrew Wikipedia.
- 293, TITLE: Finding Dataset Shortcuts with Grammar Induction  
<https://aclanthology.org/2022.emnlp-main.293>  
AUTHORS: Dan Friedman, Alexander Wettig, Danqi Chen  
HIGHLIGHT: In this work, we propose to use probabilistic grammars to characterize and discover shortcuts in NLP datasets.
- 294, TITLE: Retrieval Augmentation for Commonsense Reasoning: A Unified Approach  
<https://aclanthology.org/2022.emnlp-main.294>  
AUTHORS: Wenhao Yu, Chenguang Zhu, Zhihan Zhang, Shuohang Wang, Zhuosheng Zhang, Yuwei Fang, Meng Jiang  
HIGHLIGHT: In this paper, we systematically investigate how to leverage commonsense knowledge retrieval to improve commonsense reasoning tasks.
- 295, TITLE: Open World Classification with Adaptive Negative Samples  
<https://aclanthology.org/2022.emnlp-main.295>  
AUTHORS: Ke Bai, Guoyin Wang, Jiwei Li, Sunghyun Park, Sungjin Lee, Puyang Xu, Ricardo Henao, Lawrence Carin  
HIGHLIGHT: We propose an approach based on Adaptive Negative Samples (ANS) designed to generate effective synthetic open category samples in the training stage and without requiring any prior knowledge or external datasets.
- 296, TITLE: Re3: Generating Longer Stories With Recursive Reprompting and Revision  
<https://aclanthology.org/2022.emnlp-main.296>  
AUTHORS: Kevin Yang, Yuandong Tian, Nanyun Peng, Dan Klein  
HIGHLIGHT: Compared to prior work on shorter stories, long-range plot coherence and relevance are more central challenges here. We propose the Recursive Reprompting and Revision framework (Re3) to address these challenges by (a) prompting a general-purpose language model to construct a structured overarching plan, and (b) generating story passages by repeatedly injecting contextual information from both the plan and current story state into a language model prompt.
- 297, TITLE: Does Joint Training Really Help Cascaded Speech Translation?  
<https://aclanthology.org/2022.emnlp-main.297>  
AUTHORS: Viet Anh Khoa Tran, David Thulke, Yingbo Gao, Christian Herold, Hermann Ney  
HIGHLIGHT: In this work, we seek to answer the question of whether joint training really helps cascaded speech translation.
- 298, TITLE: MasakhaNER 2.0: Africa-centric Transfer Learning for Named Entity Recognition  
<https://aclanthology.org/2022.emnlp-main.298>  
AUTHORS: David Adelani, Graham Neubig, Sebastian Ruder, Shruti Rijhwani, Michael Beukman, Chester Palen-Michel, Constantine Lignos, Jesujoba Alabi, Shamsuddeen Muhammad, Peter Nabende, Cheikh M. Bamba Dione, Andiswa Bukula, Rooweither Mabuya, Bonaventure F. P. Dossou, Blessing Sibanda, Happy Buzaaba, Jonathan Mukiibi, Godson Kalipe, Derguene

Mbaye, Amelia Taylor, Fatoumata Kabore, Chris Chinenye Emezue, Anuoluwapo Aremu, Perez Ogayo, Catherine Gitau, Edwin Munkoh-Buabeng, Victoire Memdjokam Koagne, Allahsera Auguste Tapo, Tebogo Macucwa, Vukosi Marivate, Mboning Tchiaze Elvis, Tajuddeen Gwadabe, Tosin Adewumi, Orevaoghene Ahia, Joyce Nakatumba-Nabende, Neo Lerato Mokono, Ignatius Ezeani, Chiamaka Chukwunke, Mofetoluwa Oluwaseun Adeyemi, Gilles Quentin Hacheme, Idris Abdulmumin, Odunayo Ogundepo, Oreen Yousuf, Tatiana Moteu, Dietrich Klakow

**HIGHLIGHT:** Multiple challenges exist, including the limited availability of annotated training and evaluation datasets as well as the lack of understanding of which settings, languages, and recently proposed methods like cross-lingual transfer will be effective. In this paper, we aim to move towards solutions for these challenges, focusing on the task of named entity recognition (NER).

299, **TITLE:** Ethics consideration sections in natural language processing papers

<https://aclanthology.org/2022.emnlp-main.299>

**AUTHORS:** Luciana Benotti, Patrick Blackburn

**HIGHLIGHT:** In this paper, we present the results of a manual classification of all ethical consideration sections for ACL 2021.

300, **TITLE:** Continued Pretraining for Better Zero- and Few-Shot Promptability

<https://aclanthology.org/2022.emnlp-main.300>

**AUTHORS:** Zhaofeng Wu, Robert L Logan IV, Pete Walsh, Akshita Bhagia, Dirk Groeneveld, Sameer Singh, Iz Beltagy

**HIGHLIGHT:** In this work, we investigate if a dedicated continued pretraining stage could improve promptability?, i.e., zero-shot performance with natural language prompts or few-shot performance with prompt tuning.

301, **TITLE:** Less is More: Summary of Long Instructions is Better for Program Synthesis

<https://aclanthology.org/2022.emnlp-main.301>

**AUTHORS:** Kirby Kuznia, Swaroop Mishra, Mihir Parmar, Chitta Baral

**HIGHLIGHT:** We show that LMs benefit from the summarized version of complicated questions.

302, **TITLE:** Is a Question Decomposition Unit All We Need?

<https://aclanthology.org/2022.emnlp-main.302>

**AUTHORS:** Pruthvi Patel, Swaroop Mishra, Mihir Parmar, Chitta Baral

**HIGHLIGHT:** With the growing number of new benchmarks, we build bigger and more complex LMs.

303, **TITLE:** Discourse-Aware Soft Prompting for Text Generation

<https://aclanthology.org/2022.emnlp-main.303>

**AUTHORS:** Marjan Ghazvininejad, Vladimir Karpukhin, Vera Gor, Asli Celikyilmaz

**HIGHLIGHT:** We show that structured design of prefix parameters yields more coherent, faithful and relevant generations than the baseline prefix-tuning on all generation tasks.

304, **TITLE:** ExpUNations: Augmenting Puns with Keywords and Explanations

<https://aclanthology.org/2022.emnlp-main.304>

**AUTHORS:** Jiao Sun, Anjali Narayan-Chen, Shereen Oraby, Alessandra Cervone, Tagyoung Chung, Jing Huang, Yang Liu, Nanyun Peng

**HIGHLIGHT:** In this paper, we present the ExpUNations (ExpUN) dataset, in which we augment an existing dataset of puns with detailed crowdsourced annotations of keywords denoting the most distinctive words that make the text funny, pun explanations describing why the text is funny, and fine-grained funniness ratings.

305, **TITLE:** SLING: Sino Linguistic Evaluation of Large Language Models

<https://aclanthology.org/2022.emnlp-main.305>

**AUTHORS:** Yixiao Song, Kalpesh Krishna, Rajesh Bhatt, Mohit Iyyer

**HIGHLIGHT:** To understand what kinds of linguistic knowledge are encoded by pretrained Chinese language models (LMs), we introduce the benchmark of Sino LINGuistics (SLING), which consists of 38K minimal sentence pairs in Mandarin Chinese grouped into 9 high-level linguistic phenomena.

306, **TITLE:** Context-Situated Pun Generation

<https://aclanthology.org/2022.emnlp-main.306>

**AUTHORS:** Jiao Sun, Anjali Narayan-Chen, Shereen Oraby, Shuyang Gao, Tagyoung Chung, Jing Huang, Yang Liu, Nanyun Peng

**HIGHLIGHT:** In this work, we propose a new task, context-situated pun generation, where a specific context represented by a set of keywords is provided, and the task is to first identify suitable pun words that are appropriate for the context, then generate puns based on the context keywords and the identified pun words.

307, **TITLE:** Retrieval-Augmented Generative Question Answering for Event Argument Extraction

<https://aclanthology.org/2022.emnlp-main.307>

**AUTHORS:** Xinya Du, Heng Ji

**HIGHLIGHT:** We propose a retrieval-augmented generative QA model (R-GQA) for event argument extraction.

308, **TITLE:** Concadia: Towards Image-Based Text Generation with a Purpose

<https://aclanthology.org/2022.emnlp-main.308>

**AUTHORS:** Elisa Kreiss, Fei Fang, Noah Goodman, Christopher Potts

**HIGHLIGHT:** Descriptions focus on visual features and are meant to replace an image (often to increase accessibility), whereas captions appear alongside an image to supply additional information. To motivate this distinction and help people put it into

practice, we introduce the publicly available Wikipedia-based dataset Concadia consisting of 96,918 images with corresponding English-language descriptions, captions, and surrounding context.

- 309, TITLE: Context Matters for Image Descriptions for Accessibility: Challenges for Referenceless Evaluation Metrics  
<https://aclanthology.org/2022.emnlp-main.309>  
AUTHORS: Elisa Kreiss, Cynthia Bennett, Shayan Hooshmand, Eric Zelikman, Meredith Ringel Morris, Christopher Potts  
HIGHLIGHT: The fundamental shortcoming of these metrics is that they do not take context into account, whereas contextual information is highly valued by BLV users. To substantiate these claims, we present a study with BLV participants who rated descriptions along a variety of dimensions.
- 310, TITLE: MetaLogic: Logical Reasoning Explanations with Fine-Grained Structure  
<https://aclanthology.org/2022.emnlp-main.310>  
AUTHORS: Yinya Huang, Hongming Zhang, Ruixin Hong, Xiaodan Liang, Changshui Zhang, Dong Yu  
HIGHLIGHT: In this paper, we propose a comprehensive benchmark to investigate models' logical reasoning capabilities in complex real-life scenarios.
- 311, TITLE: Explicit Query Rewriting for Conversational Dense Retrieval  
<https://aclanthology.org/2022.emnlp-main.311>  
AUTHORS: Hongjin Qian, Zhicheng Dou  
HIGHLIGHT: In this paper, we propose a model CRDR that can perform query rewriting and context modelling in a unified framework in which the query rewriting's supervision signals further enhance the context modelling.
- 312, TITLE: Efficient Nearest Neighbor Emotion Classification with BERT-whitening  
<https://aclanthology.org/2022.emnlp-main.312>  
AUTHORS: Wenbiao Yin, Lin Shang  
HIGHLIGHT: In this paper, we propose kNN-EC, a simple and efficient non-parametric emotion classification (EC) method using nearest neighbor retrieval.
- 313, TITLE: FastClass: A Time-Efficient Approach to Weakly-Supervised Text Classification  
<https://aclanthology.org/2022.emnlp-main.313>  
AUTHORS: Tingyu Xia, Yue Wang, Yuan Tian, Yi Chang  
HIGHLIGHT: This paper proposes FastClass, an efficient weakly-supervised classification approach.
- 314, TITLE: Neural-Symbolic Inference for Robust Autoregressive Graph Parsing via Compositional Uncertainty Quantification  
<https://aclanthology.org/2022.emnlp-main.314>  
AUTHORS: Zi Lin, Jeremiah Liu, Jingbo Shang  
HIGHLIGHT: In this work, we study compositionality-aware approach to neural-symbolic inference informed by model confidence, performing fine-grained neural-symbolic reasoning at subgraph level (i.e., nodes and edges) and precisely targeting subgraph components with high uncertainty in the neural parser.
- 315, TITLE: A Speaker-Aware Co-Attention Framework for Medical Dialogue Information Extraction  
<https://aclanthology.org/2022.emnlp-main.315>  
AUTHORS: Yuan Xia, Zhenhui Shi, Jingbo Zhou, Jiayu Xu, Chao Lu, Yehui Yang, Lei Wang, Haifeng Huang, Xia Zhang, Junwei Liu  
HIGHLIGHT: To this end, in this paper, we propose a speaker-aware co-attention framework for medical dialogue information extraction.
- 316, TITLE: Towards Interactivity and Interpretability: A Rationale-based Legal Judgment Prediction Framework  
<https://aclanthology.org/2022.emnlp-main.316>  
AUTHORS: Yiquan Wu, Yifei Liu, Weiming Lu, Yating Zhang, Jun Feng, Changlong Sun, Fei Wu, Kun Kuang  
HIGHLIGHT: Following the judge's real trial logic, in this paper, we propose a novel Rationale-based Legal Judgment Prediction (RLJP) framework.
- 317, TITLE: ReICLIP: Adapting Language-Image Pretraining for Visual Relationship Detection via Relational Contrastive Learning  
<https://aclanthology.org/2022.emnlp-main.317>  
AUTHORS: Yi Zhu, Zhaoqing Zhu, Bingqian Lin, Xiaodan Liang, Feng Zhao, Jianzhuang Liu  
HIGHLIGHT: In this paper, we introduce compact language information of relation labels for regularizing the representation learning of visual relations.
- 318, TITLE: Candidate Soups: Fusing Candidate Results Improves Translation Quality for Non-Autoregressive Translation  
<https://aclanthology.org/2022.emnlp-main.318>  
AUTHORS: Huanran Zheng, Wei Zhu, Pengfei Wang, Xiaoling Wang  
HIGHLIGHT: In this paper, we propose a simple but effective method called 'Candidate Soups,' which can obtain high-quality translations while maintaining the inference speed of NAT models.
- 319, TITLE: Evaluating Parameter Efficient Learning for Generation  
<https://aclanthology.org/2022.emnlp-main.319>

- AUTHORS: Peng Xu, Mostofa Patwary, Shrimai Prabhunoye, Virginia Adams, Ryan Prenger, Wei Ping, Nayeon Lee, Mohammad Shoeybi, Bryan Catanzaro  
HIGHLIGHT: In this paper, we present comparisons between PERMs and finetuning from three new perspectives: (1) the effect of sample and model size to in-domain evaluations, (2) generalization to unseen domains and new datasets, and (3) the faithfulness of generations.
- 320, TITLE: McQueen: a Benchmark for Multimodal Conversational Query Rewrite  
<https://aclanthology.org/2022.emnlp-main.320>  
AUTHORS: Yifei Yuan, Chen Shi, Runze Wang, Liyi Chen, Feijun Jiang, Yuan You, Wai Lam  
HIGHLIGHT: In this paper, we propose the task of multimodal conversational query rewrite (McQR), which performs query rewrite under the multimodal visual conversation setting.
- 321, TITLE: Self-supervised Graph Masking Pre-training for Graph-to-Text Generation  
<https://aclanthology.org/2022.emnlp-main.321>  
AUTHORS: Jiuzhou Han, Ehsan Shareghi  
HIGHLIGHT: Additionally, PLMs are typically pre-trained on free text which introduces domain mismatch between pre-training and downstream G2T generation tasks. To address these shortcomings, we propose graph masking pre-training strategies that neither require supervision signals nor adjust the architecture of the underlying pre-trained encoder-decoder model.
- 322, TITLE: Improving Stability of Fine-Tuning Pretrained Language Models via Component-Wise Gradient Norm Clipping  
<https://aclanthology.org/2022.emnlp-main.322>  
AUTHORS: Chenghao Yang, Xuezhe Ma  
HIGHLIGHT: In this paper, we first point out that this method does not always work out due to the different convergence speeds of different layers/modules. Inspired by this observation, we propose a simple component-wise gradient norm clipping method to adjust the convergence speed for different components.
- 323, TITLE: Differentially Private Language Models for Secure Data Sharing  
<https://aclanthology.org/2022.emnlp-main.323>  
AUTHORS: Justus Mattern, Zhijing Jin, Benjamin Weggenmann, Bernhard Schoelkopf, Mrinmaya Sachan  
HIGHLIGHT: In this paper, we approach the problem at hand using global differential privacy, particularly by training a generative language model in a differentially private manner and consequently sampling data from it.
- 324, TITLE: Conditional set generation using Seq2seq models  
<https://aclanthology.org/2022.emnlp-main.324>  
AUTHORS: Aman Madaan, Dheeraj Rajagopal, Niket Tandon, Yiming Yang, Antoine Bosselut  
HIGHLIGHT: We propose a novel algorithm for effectively sampling informative orders over the combinatorial space of label orders.
- 325, TITLE: Analyzing and Evaluating Faithfulness in Dialogue Summarization  
<https://aclanthology.org/2022.emnlp-main.325>  
AUTHORS: Bin Wang, Chen Zhang, Yan Zhang, Yiming Chen, Haizhou Li  
HIGHLIGHT: In this work, we first perform the fine-grained human analysis on the faithfulness of dialogue summaries and observe that over 35% of generated summaries are faithfully inconsistent respective the source dialogues. Furthermore, we present a new model-level faithfulness evaluation method.
- 326, TITLE: Twist Decoding: Diverse Generators Guide Each Other  
<https://aclanthology.org/2022.emnlp-main.326>  
AUTHORS: Jungo Kasai, Keisuke Sakaguchi, Ronan Le Bras, Hao Peng, Ximing Lu, Dragomir Radev, Yejin Choi, Noah A. Smith  
HIGHLIGHT: We introduce Twist decoding, a simple and general text generation algorithm that benefits from diverse models at inference time.
- 327, TITLE: Exploring Representation-level Augmentation for Code Search  
<https://aclanthology.org/2022.emnlp-main.327>  
AUTHORS: Haochen Li, Chunyan Miao, Cyril Leung, Yanxian Huang, Yuan Huang, Hongyu Zhang, Yanlin Wang  
HIGHLIGHT: In this paper, we explore augmentation methods that augment data (both code and query) at representation level which does not require additional data processing and training, and based on this we propose a general format of representation-level augmentation that unifies existing methods.
- 328, TITLE: Learning Semantic Textual Similarity via Topic-informed Discrete Latent Variables  
<https://aclanthology.org/2022.emnlp-main.328>  
AUTHORS: Erxin Yu, Lan Du, Yuan Jin, Zhepei Wei, Yi Chang  
HIGHLIGHT: In this paper, we develop a topic-informed discrete latent variable model for semantic textual similarity, which learns a shared latent space for sentence-pair representation via vector quantization.
- 329, TITLE: STRUDEL: Structured Dialogue Summarization for Dialogue Comprehension  
<https://aclanthology.org/2022.emnlp-main.329>  
AUTHORS: Borui Wang, Chengcheng Feng, Arjun Nair, Madelyn Mao, Jai Desai, Asli Celikyilmaz, Haoran Li, Yashar Mehdad, Dragomir Radev

**HIGHLIGHT:** In this paper, we propose a novel type of dialogue summarization task - STRUctured DiaLoguE Summarization (STRUDEL) - that can help pre-trained language models to better understand dialogues and improve their performance on important dialogue comprehension tasks.

330, **TITLE:** Competency-Aware Neural Machine Translation: Can Machine Translation Know its Own Translation Quality?  
<https://aclanthology.org/2022.emnlp-main.330>

**AUTHORS:** Pei Zhang, Baosong Yang, Hao-Ran Wei, Dayiheng Liu, Kai Fan, Luo Si, Jun Xie

**HIGHLIGHT:** This is in sharp contrast to human translators who give feedback or conduct further investigations whenever they are in doubt about predictions. To fill this gap, we propose a novel competency-aware NMT by extending conventional NMT with a self-estimator, offering abilities to translate a source sentence and estimate its competency.

331, **TITLE:** PASTA: Table-Operations Aware Fact Verification via Sentence-Table Cloze Pre-training  
<https://aclanthology.org/2022.emnlp-main.331>

**AUTHORS:** Zihui Gu, Ju Fan, Nan Tang, Preslav Nakov, Xiaoman Zhao, Xiaoyong Du

**HIGHLIGHT:** Yet, progress has been limited due to the lack of datasets that can be used to pre-train language models (LMs) to be aware of common table operations, such as aggregating a column or comparing tuples. To bridge this gap, this paper introduces PASTA for table-based fact verification via pre-training with synthesized sentence?table cloze questions.

332, **TITLE:** Sentiment-Aware Word and Sentence Level Pre-training for Sentiment Analysis  
<https://aclanthology.org/2022.emnlp-main.332>

**AUTHORS:** Shuai Fan, Chen Lin, Haonan Li, Zhenghao Lin, Jinsong Su, Hang Zhang, Yeyun Gong, Jian Guo, Nan Duan

**HIGHLIGHT:** In this paper, we propose SentiWSP, a novel Sentiment-aware pre-trained language model with combined Word-level and Sentence-level Pre-training tasks.

333, **TITLE:** Towards Multi-Modal Sarcasm Detection via Hierarchical Congruity Modeling with Knowledge Enhancement  
<https://aclanthology.org/2022.emnlp-main.333>

**AUTHORS:** Hui Liu, Wenya Wang, Haoliang Li

**HIGHLIGHT:** In this paper, we propose a novel hierarchical framework for sarcasm detection by exploring both the atomic-level congruity based on multi-head cross attentions and the composition-level congruity based on graph neural networks, where a post with low congruity can be identified as sarcasm.

334, **TITLE:** Efficiently Tuned Parameters Are Task Embeddings  
<https://aclanthology.org/2022.emnlp-main.334>

**AUTHORS:** Wangchunshu Zhou, Canwen Xu, Julian McAuley

**HIGHLIGHT:** In this paper, we anticipate that task-specific parameters updated in parameter-efficient tuning methods are likely to encode task-specific information.

335, **TITLE:** COPEN: Probing Conceptual Knowledge in Pre-trained Language Models  
<https://aclanthology.org/2022.emnlp-main.335>

**AUTHORS:** Hao Peng, Xiaozhi Wang, Shengding Hu, Hailong Jin, Lei Hou, Juanzi Li, Zhiyuan Liu, Qun Liu

**HIGHLIGHT:** Inspired by knowledge representation schemata, we comprehensively evaluate conceptual knowledge of PLMs by designing three tasks to probe whether PLMs organize entities by conceptual similarities, learn conceptual properties, and conceptualize entities in contexts, respectively.

336, **TITLE:** Capturing Global Structural Information in Long Document Question Answering with Compressive Graph Selector Network  
<https://aclanthology.org/2022.emnlp-main.336>

**AUTHORS:** Yuxiang Nie, Heyan Huang, Wei Wei, Xian-Ling Mao

**HIGHLIGHT:** However, these methods usually ignore the global structure of the long document, which is essential for long-range understanding. To tackle this problem, we propose Compressive Graph Selector Network (CGSN) to capture the global structure in a compressive and iterative manner.

337, **TITLE:** Structural generalization is hard for sequence-to-sequence models  
<https://aclanthology.org/2022.emnlp-main.337>

**AUTHORS:** Yuekun Yao, Alexander Koller

**HIGHLIGHT:** However, recent work on compositional generalization has shown that seq2seq models achieve very low accuracy in generalizing to linguistic structures that were not seen in training. We present new evidence that this is a general limitation of seq2seq models that is present not just in semantic parsing, but also in syntactic parsing and in text-to-text tasks, and that this limitation can often be overcome by neurosymbolic models that have linguistic knowledge built in.

338, **TITLE:** Contrastive Learning enhanced Author-Style Headline Generation  
<https://aclanthology.org/2022.emnlp-main.338>

**AUTHORS:** Hui Liu, Weidong Guo, Yige Chen, Xiangyang Li

**HIGHLIGHT:** Besides, we propose two methods to use the learned stylistic features to guide both the pointer and the decoder during the generation.

339, **TITLE:** Multi-Granularity Optimization for Non-Autoregressive Translation  
<https://aclanthology.org/2022.emnlp-main.339>

**AUTHORS:** Yafu Li, Leyang Cui, Yongjing Yin, Yue Zhang



- HIGHLIGHT:** This assumption is further strengthened by cross-entropy loss, which encourages a strict match between the hypothesis and the reference token by token. To alleviate this issue, we propose multi-granularity optimization for NAT, which collects model behaviours on translation segments of various granularities and integrates feedback for backpropagation.
- 340, **TITLE:** Super-NaturalInstructions: Generalization via Declarative Instructions on 1600+ NLP Tasks  
<https://aclanthology.org/2022.emnlp-main.340>  
**AUTHORS:** Yizhong Wang, Swaroop Mishra, Pegah Alipoormolabashi, Yeganeh Kordi, Amirreza Mirzaei, Atharva Naik, Arjun Ashok, Arut Selvan Dhanasekaran, Anjana Arunkumar, David Stap, Eshaan Pathak, Giannis Karamanolakis, Haizhi Lai, Ishan Purohit, Ishani Mondal, Jacob Anderson, Kirby Kuznia, Krma Doshi, Kuntal Kumar Pal, Maitreya Patel, Mehrad Moradshahi, Mihir Parmar, Mirali Purohit, Neeraj Varshney, Phani Rohitha Kaza, Pulkit Verma, Ravsehaj Singh Puri, Rushang Karia, Savan Doshi, Shailaja Keyur Sampat, Siddhartha Mishra, Sujan Reddy A, Sumanta Patro, Tanay Dixit, Xudong Shen  
**HIGHLIGHT:** Furthermore, we build Tk-Instruct, a transformer model trained to follow a variety of in-context instructions (plain language task definitions or k-shot examples).
- 341, **TITLE:** MetaFill: Text Infilling for Meta-Path Generation on Heterogeneous Information Networks  
<https://aclanthology.org/2022.emnlp-main.341>  
**AUTHORS:** Zequn Liu, Kefei Duan, Junwei Yang, Hanwen Xu, Ming Zhang, Sheng Wang  
**HIGHLIGHT:** Existing meta-path generation approaches cannot fully exploit the rich textual information in HINs, such as node names and edge type names. To address this problem, we propose MetaFill, a text-infilling-based approach for meta-path generation.
- 342, **TITLE:** DRLK: Dynamic Hierarchical Reasoning with Language Model and Knowledge Graph for Question Answering  
<https://aclanthology.org/2022.emnlp-main.342>  
**AUTHORS:** Miao Zhang, Rufeng Dai, Ming Dong, Tingting He  
**HIGHLIGHT:** In this paper, we propose DRLK (Dynamic Hierarchical Reasoning with Language Model and Knowledge Graphs), a novel model that utilizes dynamic hierarchical interactions between the QA context and KG for reasoning.
- 343, **TITLE:** AEG: Argumentative Essay Generation via A Dual-Decoder Model with Content Planning  
<https://aclanthology.org/2022.emnlp-main.343>  
**AUTHORS:** Jianzhu Bao, Yasheng Wang, Yitong Li, Fei Mi, Ruifeng Xu  
**HIGHLIGHT:** In this paper, we propose a new task, Argumentative Essay Generation (AEG).
- 344, **TITLE:** BotsTalk: Machine-sourced Framework for Automatic Curation of Large-scale Multi-skill Dialogue Datasets  
<https://aclanthology.org/2022.emnlp-main.344>  
**AUTHORS:** Minju Kim, Chaehyeon Kim, Yong Ho Song, Seung-won Hwang, Jinyoung Yeo  
**HIGHLIGHT:** To build open-domain chatbots that are able to use diverse communicative skills, we propose a novel framework BotsTalk, where multiple agents grounded to the specific target skills participate in a conversation to automatically annotate multi-skill dialogues.
- 345, **TITLE:** Wider & Closer: Mixture of Short-channel Distillers for Zero-shot Cross-lingual Named Entity Recognition  
<https://aclanthology.org/2022.emnlp-main.345>  
**AUTHORS:** Jun-Yu Ma, Beiduo Chen, Jia-Chen Gu, Zhenhua Ling, Wu Guo, Quan Liu, Zhigang Chen, Cong Liu  
**HIGHLIGHT:** In this study, a mixture of short-channel distillers (MSD) method is proposed to fully interact the rich hierarchical information in the teacher model and to transfer knowledge to the student model sufficiently and efficiently.
- 346, **TITLE:** An Efficient Memory-Augmented Transformer for Knowledge-Intensive NLP Tasks  
<https://aclanthology.org/2022.emnlp-main.346>  
**AUTHORS:** Yuxiang Wu, Yu Zhao, Baotian Hu, Pasquale Minervini, Pontus Stenertorp, Sebastian Riedel  
**HIGHLIGHT:** To combine the strength of both approaches, we propose the Efficient Memory-Augmented Transformer (EMAT) which encodes external knowledge into a key-value memory and exploits the fast maximum inner product search for memory querying.
- 347, **TITLE:** Supervised Prototypical Contrastive Learning for Emotion Recognition in Conversation  
<https://aclanthology.org/2022.emnlp-main.347>  
**AUTHORS:** Xiaohui Song, Longtao Huang, Hui Xue, Songlin Hu  
**HIGHLIGHT:** In this paper, we propose a Supervised Prototypical Contrastive Learning (SPCL) loss for the ERC task.
- 348, **TITLE:** RuCoLA: Russian Corpus of Linguistic Acceptability  
<https://aclanthology.org/2022.emnlp-main.348>  
**AUTHORS:** Vladislav Mikhailov, Tatiana Shamardina, Max Ryabinin, Alena Pestova, Ivan Smurov, Ekaterina Artemova  
**HIGHLIGHT:** To this end, we introduce the Russian Corpus of Linguistic Acceptability (RuCoLA), built from the ground up under the well-established binary LA approach.
- 349, **TITLE:** Complex Hyperbolic Knowledge Graph Embeddings with Fast Fourier Transform  
<https://aclanthology.org/2022.emnlp-main.349>  
**AUTHORS:** Huiru Xiao, Xin Liu, Yangqiu Song, Ginny Wong, Simon See  
**HIGHLIGHT:** This paper aims to utilize the representation capacity of the complex hyperbolic geometry in multi-relational KG embeddings.
- 350, **TITLE:** Towards Knowledge-Intensive Text-to-SQL Semantic Parsing with Formulaic Knowledge

<https://aclanthology.org/2022.emnlp-main.350>

AUTHORS: Longxu Dou, Yan Gao, Xuqi Liu, Mingyang Pan, Dingzirui Wang, Wanxiang Che, Dechen Zhan, Min-Yen Kan, Jian-Guang Lou

HIGHLIGHT: In this paper, we study the problem of knowledge-intensive text-to-SQL, in which domain knowledge is necessary to parse expert questions into SQL queries over domain-specific tables.

351, TITLE: Should We Ban English NLP for a Year?

<https://aclanthology.org/2022.emnlp-main.351>

AUTHORS: Anders S?gaard

HIGHLIGHT: Many have argued that it is almost impossible to mitigate inequality amplification. I argue that, on the contrary, it is quite simple to do so, and that counter-measures would have little-to-no negative impact, except for, perhaps, in the very short term.

352, TITLE: LittleBird: Efficient Faster & Longer Transformer for Question Answering

<https://aclanthology.org/2022.emnlp-main.352>

AUTHORS: Minchul Lee, Kijong Han, Myeong Cheol Shin

HIGHLIGHT: But it has a limitation dealing with long inputs due to its attention mechanism. Longformer, ETC and BigBird addressed this issue and effectively solved the quadratic dependency problem. However we find that these models are not sufficient, and propose LittleBird, a novel model based on BigBird with improved speed and memory footprint while maintaining accuracy.

353, TITLE: WeTS: A Benchmark for Translation Suggestion

<https://aclanthology.org/2022.emnlp-main.353>

AUTHORS: Zhen Yang, Fandong Meng, Yingxue Zhang, Ernan Li, Jie Zhou

HIGHLIGHT: To break these limitations mentioned above and spur the research in TS, we create a benchmark dataset, called WeTS, which is a golden corpus annotated by expert translators on four translation directions.

354, TITLE: Discrete Cross-Modal Alignment Enables Zero-Shot Speech Translation

<https://aclanthology.org/2022.emnlp-main.354>

AUTHORS: Chen Wang, Yuchen Liu, Boxing Chen, Jiajun Zhang, Wei Luo, Zhongqiang Huang, Chengqing Zong

HIGHLIGHT: In order to enable zero-shot ST, we propose a novel Discrete Cross-Modal Alignment (DCMA) method that employs a shared discrete vocabulary space to accommodate and match both modalities of speech and text.

355, TITLE: Abstractive Summarization Guided by Latent Hierarchical Document Structure

<https://aclanthology.org/2022.emnlp-main.355>

AUTHORS: Yifu Qiu, Shay B. Cohen

HIGHLIGHT: To address this shortcoming, we propose a hierarchy-aware graph neural network (HierGNN) which captures such dependencies through three main steps: 1) learning a hierarchical document structure through a latent structure tree learned by a sparse matrix-tree computation; 2) propagating sentence information over this structure using a novel message-passing node propagation mechanism to identify salient information; 3) using graph-level attention to concentrate the decoder on salient information.

356, TITLE: Explainable Question Answering based on Semantic Graph by Global Differentiable Learning and Dynamic Adaptive Reasoning

<https://aclanthology.org/2022.emnlp-main.356>

AUTHORS: Jianguo Mao, Wenbin Jiang, Xiangdong Wang, Hong Liu, Yu Xia, Yajuan Lyu, QiaoQiao She

HIGHLIGHT: To alleviate it, we propose a simple yet effective Global Differentiable Learning strategy to explore optimal reasoning paths from the latent probability space so that the model learns to solve intermediate reasoning processes without expert annotations.

357, TITLE: DuReader-Retrieval: A Large-scale Chinese Benchmark for Passage Retrieval from Web Search Engine

<https://aclanthology.org/2022.emnlp-main.357>

AUTHORS: Yifu Qiu, Hongyu Li, Yingqi Qu, Ying Chen, QiaoQiao She, Jing Liu, Hua Wu, Haifeng Wang

HIGHLIGHT: In this paper, we present DuReader-retrieval, a large-scale Chinese dataset for passage retrieval.

358, TITLE: Pair-Based Joint Encoding with Relational Graph Convolutional Networks for Emotion-Cause Pair Extraction

<https://aclanthology.org/2022.emnlp-main.358>

AUTHORS: Junlong Liu, Xichen Shang, Qianli Ma

HIGHLIGHT: This lead to an imbalance in inter-task feature interaction where features extracted later have no direct contact with the former. To address this issue, we propose a novel \*\*P\*\*air-\*\*B\*\*ased \*\*J\*\*oint \*\*E\*\*ncoding (\*\*PBJE\*\*) network, which generates pairs and clauses features simultaneously in a joint feature encoding manner to model the causal relationship in clauses.

359, TITLE: Affective Knowledge Enhanced Multiple-Graph Fusion Networks for Aspect-based Sentiment Analysis

<https://aclanthology.org/2022.emnlp-main.359>

AUTHORS: Siyu Tang, Heyan Chai, Ziyi Yao, Ye Ding, Cuiyun Gao, Binxing Fang, Qing Liao

HIGHLIGHT: In this paper, we propose a novel multi-graph fusion network (MGFN) based on latent graph to leverage the richer syntax dependency relation label information and affective semantic information of words.

360, TITLE: IndicNLG Benchmark: Multilingual Datasets for Diverse NLG Tasks in Indic Languages

<https://aclanthology.org/2022.emnlp-main.360>

- AUTHORS: Aman Kumar, Himani Shrotriya, Prachi Sahu, Amogh Mishra, Raj Dabre, Ratish Puduppully, Anoop Kunchukuttan, Mitesh M. Khapra, Pratyush Kumar  
HIGHLIGHT: We present the IndicNLG Benchmark, a collection of datasets for benchmarking NLG for 11 Indic languages.
- 361, TITLE: Improving Machine Translation with Phrase Pair Injection and Corpus Filtering  
<https://aclanthology.org/2022.emnlp-main.361>  
AUTHORS: Akshay Batheja, Pushpak Bhattacharyya  
HIGHLIGHT: In this paper, we show that the combination of Phrase Pair Injection and Corpus Filtering boosts the performance of Neural Machine Translation (NMT) systems.
- 362, TITLE: An Anchor-based Relative Position Embedding Method for Cross-Modal Tasks  
<https://aclanthology.org/2022.emnlp-main.362>  
AUTHORS: Ya Wang, Xingwu Sun, Lian Fengzong, ZhanHui Kang, Chengzhong Xu Xu  
HIGHLIGHT: In this paper, we propose a unified position embedding method for these problems, called AnChor-basEd Relative Position Embedding (ACE-RPE), in which we first introduce an anchor locating mechanism to bridge the semantic gap and locate anchors from different modalities.
- 363, TITLE: Norm-based Noisy Corpora Filtering and Refurbishing in Neural Machine Translation  
<https://aclanthology.org/2022.emnlp-main.363>  
AUTHORS: Yu Lu, Jiajun Zhang  
HIGHLIGHT: In this paper, we propose a norm-based noisy corpora filtering and refurbishing method with no external data and costly scorers.
- 364, TITLE: TeleMelody: Lyric-to-Melody Generation with a Template-Based Two-Stage Method  
<https://aclanthology.org/2022.emnlp-main.364>  
AUTHORS: Zeqian Ju, Peiling Lu, Xu Tan, Rui Wang, Chen Zhang, Songruoyao Wu, Kejun Zhang, Xiang-Yang Li, Tao Qin, Tie-Yan Liu  
HIGHLIGHT: In this paper, we develop TeleMelody, a two-stage lyric-to-melody generation system with music template (e.g., tonality, chord progression, rhythm pattern, and cadence) to bridge the gap between lyrics and melodies (i.e., the system consists of a lyric-to-template module and a template-to-melody module).
- 365, TITLE: SEEN: Structured Event Enhancement Network for Explainable Need Detection of Information Recall Assistance  
<https://aclanthology.org/2022.emnlp-main.365>  
AUTHORS: You-En Lin, An-Zi Yen, Hen-Hsen Huang, Hsin-Hsi Chen  
HIGHLIGHT: We propose a pilot model?structured event enhancement network (SEEN) that detects life event inconsistency, additional information in life events, and forgotten events.
- 366, TITLE: Rethinking Style Transformer with Energy-based Interpretation: Adversarial Unsupervised Style Transfer using a Pretrained Model  
<https://aclanthology.org/2022.emnlp-main.366>  
AUTHORS: Hojun Cho, Dohee Kim, Seungwoo Ryu, ChaeHun Park, Hyungjong Noh, Jeong-in Hwang, Minseok Choi, Edward Choi, Jaegul Choo  
HIGHLIGHT: However, adversarial training significantly degrades fluency compared to the other two metrics. In this work, we explain this phenomenon using energy-based interpretation, and leverage a pretrained language model to improve fluency.
- 367, TITLE: Towards Robust k-Nearest-Neighbor Machine Translation  
<https://aclanthology.org/2022.emnlp-main.367>  
AUTHORS: Hui Jiang, Ziyao Lu, Fandong Meng, Chulun Zhou, Jie Zhou, Degen Huang, Jinsong Su  
HIGHLIGHT: To alleviate the impact of noise, we propose a confidence-enhanced kNN-MT model with robust training.
- 368, TITLE: Tiny-NewsRec: Effective and Efficient PLM-based News Recommendation  
<https://aclanthology.org/2022.emnlp-main.368>  
AUTHORS: Yang Yu, Fangzhao Wu, Chuhan Wu, Jingwei Yi, Qi Liu  
HIGHLIGHT: In this paper, we propose Tiny-NewsRec, which can improve both the effectiveness and the efficiency of PLM-based news recommendation.
- 369, TITLE: TABS: Efficient Textual Adversarial Attack for Pre-trained NL Code Model Using Semantic Beam Search  
<https://aclanthology.org/2022.emnlp-main.369>  
AUTHORS: YunSeok Choi, Hyojun Kim, Jee-Hyong Lee  
HIGHLIGHT: In this paper, we propose TABS, an efficient beam search black-box adversarial attack method.
- 370, TITLE: Investigating the Robustness of Natural Language Generation from Logical Forms via Counterfactual Samples  
<https://aclanthology.org/2022.emnlp-main.370>  
AUTHORS: Chengyuan Liu, Leilei Gan, Kun Kuang, Fei Wu  
HIGHLIGHT: State-of-the-art methods based on pre-trained models have achieved remarkable performance on the standard test dataset. However, we question whether these methods really learn how to perform logical reasoning, rather than just relying on the spurious correlations between the headers of the tables and operators of the logical form.
- 371, TITLE: Helping the Weak Makes You Strong: Simple Multi-Task Learning Improves Non-Autoregressive Translators

<https://aclanthology.org/2022.emnlp-main.371>

AUTHORS: Xinyou Wang, Zaixiang Zheng, Shujian Huang

HIGHLIGHT: In this paper, we propose a simple and model-agnostic multi-task learning framework to provide more informative learning signals.

372, TITLE: RACE: Retrieval-augmented Commit Message Generation

<https://aclanthology.org/2022.emnlp-main.372>

AUTHORS: Ensheng Shi, Yanlin Wang, Wei Tao, Lun Du, Hongyu Zhang, Shi Han, Dongmei Zhang, Hongbin Sun

HIGHLIGHT: In this paper, we propose RACE, a new retrieval-augmented neural commit message generation method, which treats the retrieved similar commit as an exemplar and leverages it to generate an accurate commit message.

373, TITLE: PLOG: Table-to-Logic Pretraining for Logical Table-to-Text Generation

<https://aclanthology.org/2022.emnlp-main.373>

AUTHORS: Ao Liu, Haoyu Dong, Naoaki Okazaki, Shi Han, Dongmei Zhang

HIGHLIGHT: In this work, we propose a Pretrained Logical Form Generator (PLOG) framework to improve generation fidelity.

374, TITLE: GHAN: Graph-Based Hierarchical Aggregation Network for Text-Video Retrieval

<https://aclanthology.org/2022.emnlp-main.374>

AUTHORS: Yahan Yu, Bojie Hu, Yu Li

HIGHLIGHT: However, there are structural and semantic differences between text and video, making this approach challenging for fine-grained understanding. In order to solve this, we propose an end-to-end graph-based hierarchical aggregation network for text-video retrieval according to the hierarchy possessed by text and video.

375, TITLE: MuRAG: Multimodal Retrieval-Augmented Generator for Open Question Answering over Images and Text

<https://aclanthology.org/2022.emnlp-main.375>

AUTHORS: Wenhui Chen, Hexiang Hu, Xi Chen, Pat Verga, William Cohen

HIGHLIGHT: However, these methods are restricted to retrieving only textual knowledge, neglecting the ubiquitous amount of knowledge in other modalities like images ? much of which contains information not covered by any text. To address this limitation, we propose the first Multimodal Retrieval-Augmented Transformer (MuRAG), which accesses an external non-parametric multimodal memory to augment language generation.

376, TITLE: PHEE: A Dataset for Pharmacovigilance Event Extraction from Text

<https://aclanthology.org/2022.emnlp-main.376>

AUTHORS: Zhaoyue Sun, Jiazheng Li, Gabriele Pergola, Byron Wallace, Bino John, Nigel Greene, Joseph Kim, Yulan He

HIGHLIGHT: We present PHEE, a novel dataset for pharmacovigilance comprising over 5000 annotated events from medical case reports and biomedical literature, making it the largest such public dataset to date.

377, TITLE: OTSeq2Set: An Optimal Transport Enhanced Sequence-to-Set Model for Extreme Multi-label Text Classification

<https://aclanthology.org/2022.emnlp-main.377>

AUTHORS: Jie Cao, Yin Zhang

HIGHLIGHT: However, the labels in XMTC tasks are essentially an unordered set rather than an ordered sequence, the default order of labels restrains Seq2Seq models in training. To address this limitation in Seq2Seq, we propose an autoregressive sequence-to-set model for XMTC tasks named OTSeq2Set.

378, TITLE: SimQA: Detecting Simultaneous MT Errors through Word-by-Word Question Answering

<https://aclanthology.org/2022.emnlp-main.378>

AUTHORS: HyoJung Han, Marine Carpuat, Jordan Boyd-Graber

HIGHLIGHT: Yet, evaluations of simultaneous machine translation (SimulMT) fail to capture if systems correctly translate the most salient elements of a question: people, places, and dates. To address this problem, we introduce a downstream word-by-word question answering evaluation task (SimQA): given a source language question, translate the question word by word into the target language, and answer as soon as possible.

379, TITLE: Discovering Low-rank Subspaces for Language-agnostic Multilingual Representations

<https://aclanthology.org/2022.emnlp-main.379>

AUTHORS: Zhihui Xie, Handong Zhao, Tong Yu, Shuai Li

HIGHLIGHT: In this work, we provide a novel view of projecting away language-specific factors from a multilingual embedding space.

380, TITLE: Rethinking the Authorship Verification Experimental Setups

<https://aclanthology.org/2022.emnlp-main.380>

AUTHORS: Florin Brad, Andrei Manolache, Elena Burceanu, Antonio Barbalau, Radu Tudor Ionescu, Marius Popescu

HIGHLIGHT: To this end, we improve the experimental setup by proposing five new public splits over the PAN dataset, specifically designed to isolate and identify biases related to the text topic and to the author's writing style.

381, TITLE: Borrowing Human Senses: Comment-Aware Self-Training for Social Media Multimodal Classification

<https://aclanthology.org/2022.emnlp-main.381>

AUTHORS: Chunpu Xu, Jing Li

HIGHLIGHT: To better glue the cross-modal semantics therein, we capture hinting features from user comments, which are retrieved via jointly leveraging visual and lingual similarity.

382, TITLE: Training Language Models with Memory Augmentation

<https://aclanthology.org/2022.emnlp-main.382>

AUTHORS: Zexuan Zhong, Tao Lei, Danqi Chen

HIGHLIGHT: In this work, we present TRIME, a novel yet simple training approach designed for training LMs with memory augmentation.

383, TITLE: Data-Efficient Strategies for Expanding Hate Speech Detection into Under-Resourced Languages

<https://aclanthology.org/2022.emnlp-main.383>

AUTHORS: Paul R?ttger, Debora Nozza, Federico Bianchi, Dirk Hovy

HIGHLIGHT: To mitigate these issues, we explore data-efficient strategies for expanding hate speech detection into under-resourced languages. In a series of experiments with mono- and multilingual models across five non-English languages, we find that 1) a small amount of target-language fine-tuning data is needed to achieve strong performance, 2) the benefits of using more such data decrease exponentially, and 3) initial fine-tuning on readily-available English data can partially substitute target-language data and improve model generalisability.

384, TITLE: Dimension Reduction for Efficient Dense Retrieval via Conditional Autoencoder

<https://aclanthology.org/2022.emnlp-main.384>

AUTHORS: Zhenghao Liu, Han Zhang, Chenyan Xiong, Zhiyuan Liu, Yu Gu, Xiaohua Li

HIGHLIGHT: To reduce the embedding dimensions of dense retrieval, this paper proposes a Conditional Autoencoder (ConAE) to compress the high-dimensional embeddings to maintain the same embedding distribution and better recover the ranking features.

385, TITLE: Controlled Text Reduction

<https://aclanthology.org/2022.emnlp-main.385>

AUTHORS: Aviv Slobodkin, Paul Roit, Eran Hirsch, Ori Ernst, Ido Dagan

HIGHLIGHT: Concretely, we formalize Controlled Text Reduction as a standalone task, whose input is a source text with marked spans of targeted content (?highlighting?).

386, TITLE: Questioning the Validity of Summarization Datasets and Improving Their Factual Consistency

<https://aclanthology.org/2022.emnlp-main.386>

AUTHORS: Yanzhu Guo, Chlo? Clavel, Moussa Kamal Eddine, Michalis Vazirgiannis

HIGHLIGHT: Due to this lack of well-defined formulation, a large number of popular abstractive summarization datasets are constructed in a manner that neither guarantees validity nor meets one of the most essential criteria of summarization: factual consistency. In this paper, we address this issue by combining state-of-the-art factual consistency models to identify the problematic instances present in popular summarization datasets.

387, TITLE: Invariant Language Modeling

<https://aclanthology.org/2022.emnlp-main.387>

AUTHORS: Maxime Peyrard, Sarvjeet Ghotra, Martin Josifoski, Vidhan Agarwal, Barun Patra, Dean Carignan, Emre Kiciman, Saurabh Tiwary, Robert West

HIGHLIGHT: Inspired by recent progress in causal machine learning, in particular the invariant risk minimization (IRM) paradigm, we propose invariant language modeling, a framework for learning invariant representations that generalize better across multiple environments.

388, TITLE: AdaMix: Mixture-of-Adaptations for Parameter-efficient Model Tuning

<https://aclanthology.org/2022.emnlp-main.388>

AUTHORS: Yaqing Wang, Sahaj Agarwal, Subhabrata Mukherjee, Xiaodong Liu, Jing Gao, Ahmed Hassan Awadallah, Jianfeng Gao

HIGHLIGHT: We propose AdaMix as a general PEFT method that tunes a mixture of adaptation modules ? given the underlying PEFT method of choice ? introduced in each Transformer layer while keeping most of the PLM weights frozen.

389, TITLE: How ?Multi? is Multi-Document Summarization?

<https://aclanthology.org/2022.emnlp-main.389>

AUTHORS: Ruben Wolhandler, Arie Cattan, Ori Ernst, Ido Dagan

HIGHLIGHT: Accordingly, it is expected that both reference summaries in MDS datasets, as well as system summaries, would indeed be based on such dispersed information. In this paper, we argue for quantifying and assessing this expectation.

390, TITLE: BioReader: a Retrieval-Enhanced Text-to-Text Transformer for Biomedical Literature

<https://aclanthology.org/2022.emnlp-main.390>

AUTHORS: Giacomo Frisoni, Miki Mizutani, Gianluca Moro, Lorenzo Valgimigli

HIGHLIGHT: We introduce BioReader, the first retrieval-enhanced text-to-text model for biomedical natural language processing.

391, TITLE: T-Modules: Translation Modules for Zero-Shot Cross-Modal Machine Translation

<https://aclanthology.org/2022.emnlp-main.391>

AUTHORS: Paul-Ambroise Duquenne, Hongyu Gong, Beno?t Sagot, Holger Schwenk

**HIGHLIGHT:** We present a new approach to perform zero-shot cross-modal transfer between speech and text for translation tasks.

392, **TITLE:** LILA: A Unified Benchmark for Mathematical Reasoning

<https://aclanthology.org/2022.emnlp-main.392>

**AUTHORS:** Swaroop Mishra, Matthew Finlayson, Pan Lu, Leonard Tang, Sean Welleck, Chitta Baral, Tanmay Rajpurohit, Oyvind Tafford, Ashish Sabharwal, Peter Clark, Ashwin Kalyan

**HIGHLIGHT:** Towards evaluating and improving AI systems in this domain, we propose LILA, a unified mathematical reasoning benchmark consisting of 23 diversetasks along four dimensions: (i) mathematical abilities e.g., arithmetic, calculus (ii) language format e.g., question-answering, fill-in-the-blanks (iii) language diversity e.g., no language, simple language (iv) external knowledge e.g., commonsense, physics.

393, **TITLE:** Leveraging Affirmative Interpretations from Negation Improves Natural Language Understanding

<https://aclanthology.org/2022.emnlp-main.393>

**AUTHORS:** Md Mosharaf Hossain, Eduardo Blanco

**HIGHLIGHT:** Inspired by the fact that understanding a negated statement often requires humans to infer affirmative interpretations, in this paper we show that doing so benefits models for three natural language understanding tasks.

394, **TITLE:** GraphQ IR: Unifying the Semantic Parsing of Graph Query Languages with One Intermediate Representation

<https://aclanthology.org/2022.emnlp-main.394>

**AUTHORS:** Lunyiu Nie, Shulin Cao, Jiaxin Shi, Jiuding Sun, Qi Tian, Lei Hou, Juanzi Li, Jidong Zhai

**HIGHLIGHT:** This paper proposes a unified intermediate representation for graph query languages, named GraphQ IR.

395, **TITLE:** InforMask: Unsupervised Informative Masking for Language Model Pretraining

<https://aclanthology.org/2022.emnlp-main.395>

**AUTHORS:** Nafis Sadeq, Canwen Xu, Julian McAuley

**HIGHLIGHT:** In this paper, we propose InforMask, a new unsupervised masking strategy for training masked language models.

396, **TITLE:** CTRLsum: Towards Generic Controllable Text Summarization

<https://aclanthology.org/2022.emnlp-main.396>

**AUTHORS:** Junxian He, Wojciech Kryscinski, Bryan McCann, Nazneen Rajani, Caiming Xiong

**HIGHLIGHT:** Current summarization systems yield generic summaries that are disconnected from users' preferences and expectations. To address this limitation, we present CTRLsum, a generic framework to control generated summaries through a set of keywords.

397, **TITLE:** Missing Counter-Evidence Renders NLP Fact-Checking Unrealistic for Misinformation

<https://aclanthology.org/2022.emnlp-main.397>

**AUTHORS:** Max Glockner, Yufang Hou, Iryna Gurevych

**HIGHLIGHT:** Here, we contrast and compare NLP fact-checking with how professional fact-checkers combat misinformation in the absence of counter-evidence.

398, **TITLE:** A Framework for Adapting Pre-Trained Language Models to Knowledge Graph Completion

<https://aclanthology.org/2022.emnlp-main.398>

**AUTHORS:** Justin Lovelace, Carolyn Ros?

**HIGHLIGHT:** In this work, we conduct a comprehensive exploration of how to best extract and incorporate those embeddings into knowledge graph completion models.

399, **TITLE:** Mutual Information Alleviates Hallucinations in Abstractive Summarization

<https://aclanthology.org/2022.emnlp-main.399>

**AUTHORS:** Liam van der Poel, Ryan Cotterell, Clara Meister

**HIGHLIGHT:** In this paper, we identify a simple criterion under which models are significantly more likely to assign more probability to hallucinated content during generation: high model uncertainty.

400, **TITLE:** Toward the Limitation of Code-Switching in Cross-Lingual Transfer

<https://aclanthology.org/2022.emnlp-main.400>

**AUTHORS:** Yukun Feng, Feng Li, Philipp Koehn

**HIGHLIGHT:** This paper mitigates the limitation of the code-switching method by not only making the token replacement but considering the similarity between the context and the switched tokens so that the newly substituted sentences are grammatically consistent during both training and inference.

401, **TITLE:** Syntactically Rich Discriminative Training: An Effective Method for Open Information Extraction

<https://aclanthology.org/2022.emnlp-main.401>

**AUTHORS:** Frank Mtumbuka, Thomas Lukasiewicz

**HIGHLIGHT:** We propose several new methods for training neural OIE models in this paper.

402, **TITLE:** Transformer-based Entity Typing in Knowledge Graphs

<https://aclanthology.org/2022.emnlp-main.402>

**AUTHORS:** Zhiwei Hu, Victor Gutierrez-Basulto, Zhiliang Xiang, Ru Li, Jeff Pan

- HIGHLIGHT:** In this paper, we propose a novel Transformer-based Entity Typing (TET) approach, effectively encoding the content of neighbours of an entity by means of a transformer mechanism.
- 403, **TITLE:** NewsClaims: A New Benchmark for Claim Detection from News with Attribute Knowledge  
<https://aclanthology.org/2022.emnlp-main.403>  
**AUTHORS:** Revanth Gangi Reddy, Sai Chetan Chinthakindi, Zhenhailong Wang, Yi Fung, Kathryn Conger, Ahmed ELsayed, Martha Palmer, Preslav Nakov, Eduard Hovy, Kevin Small, Heng Ji  
**HIGHLIGHT:** In this work, we present NewsClaims, a new benchmark for attribute-aware claim detection in the news domain.
- 404, **TITLE:** IsoVec: Controlling the Relative Isomorphism of Word Embedding Spaces  
<https://aclanthology.org/2022.emnlp-main.404>  
**AUTHORS:** Kelly Marchisio, Neha Verma, Kevin Duh, Philipp Koehn  
**HIGHLIGHT:** We incorporate global measures of isomorphism directly into the skipgram loss function, successfully increasing the relative isomorphism of trained word embedding spaces and improving their ability to be mapped to a shared cross-lingual space.
- 405, **TITLE:** Adversarial Concept Erasure in Kernel Space  
<https://aclanthology.org/2022.emnlp-main.405>  
**AUTHORS:** Shauli Ravfogel, Francisco Vargas, Yoav Goldberg, Ryan Cotterell  
**HIGHLIGHT:** We propose a kernelization of the recently-proposed linear concept-removal objective, and show that it is effective in guarding against the ability of certain nonlinear adversaries to recover the concept.
- 406, **TITLE:** The Authenticity Gap in Human Evaluation  
<https://aclanthology.org/2022.emnlp-main.406>  
**AUTHORS:** Kawin Ethayarajh, Dan Jurafsky  
**HIGHLIGHT:** We suggest improvements to the standard protocol to make it more theoretically sound, but even in its improved form, it cannot be used to evaluate open-ended tasks like story generation.
- 407, **TITLE:** BERT in Plutarch's Shadows  
<https://aclanthology.org/2022.emnlp-main.407>  
**AUTHORS:** Ivan Yamshchikov, Alexey Tikhonov, Yorgos Pantis, Charlotte Schubert, J?rgen Jost  
**HIGHLIGHT:** This paper presents a BERT language model for Ancient Greek.
- 408, **TITLE:** Leveraging Locality in Abstractive Text Summarization  
<https://aclanthology.org/2022.emnlp-main.408>  
**AUTHORS:** Yixin Liu, Ansong Ni, Linyong Nan, Budhaditya Deb, Chenguang Zhu, Ahmed Hassan Awadallah, Dragomir Radev  
**HIGHLIGHT:** However, the quadratic memory complexity of the self-attention module with respect to the input length hinders their applications in long text summarization. Instead of designing more efficient attention modules, we approach this problem by investigating if models with a restricted context can have competitive performance compared with the memory-efficient attention models that maintain a global context by treating the input as a single sequence.
- 409, **TITLE:** Saliency Allocation as Guidance for Abstractive Summarization  
<https://aclanthology.org/2022.emnlp-main.409>  
**AUTHORS:** Fei Wang, Kaiqiang Song, Hongming Zhang, Lifeng Jin, Sangwoo Cho, Wenlin Yao, Xiaoyang Wang, Muhao Chen, Dong Yu  
**HIGHLIGHT:** In this paper, we propose a novel summarization approach with a flexible and reliable saliency guidance, namely SEASON (Saliency Allocation as Guidance for Abstractive Summarization).
- 410, **TITLE:** Fine-tuned Language Models are Continual Learners  
<https://aclanthology.org/2022.emnlp-main.410>  
**AUTHORS:** Thomas Scialom, Tuhin Chakrabarty, Smaranda Muresan  
**HIGHLIGHT:** To address this limitation, we argue that a model should be able to keep extending its knowledge and abilities, without forgetting previous skills.
- 411, **TITLE:** Natural Logic-guided Autoregressive Multi-hop Document Retrieval for Fact Verification  
<https://aclanthology.org/2022.emnlp-main.411>  
**AUTHORS:** Rami Aly, Andreas Vlachos  
**HIGHLIGHT:** We propose a novel retrieve-and-rerank method for multi-hop retrieval, that consists of a retriever that jointly scores documents in the knowledge source and sentences from previously retrieved documents using an autoregressive formulation and is guided by a proof system based on natural logic that dynamically terminates the retrieval process if the evidence is deemed sufficient.
- 412, **TITLE:** AX-MABSA: A Framework for Extremely Weakly Supervised Multi-label Aspect Based Sentiment Analysis  
<https://aclanthology.org/2022.emnlp-main.412>  
**AUTHORS:** Sabyasachi Kamila, Walid Magdy, Sourav Dutta, MingXue Wang  
**HIGHLIGHT:** We explore unsupervised language model post-training to improve the overall performance, and propose a multi-label generator model to generate multiple aspect category-sentiment pairs per review sentence.
- 413, **TITLE:** Transfer Learning with Synthetic Corpora for Spatial Role Labeling and Reasoning

- <https://aclanthology.org/2022.emnlp-main.413>  
AUTHORS: Roshanak Mirzaee, Parisa Kordjamshidi  
HIGHLIGHT: We provide two new data resources on multiple spatial language processing tasks.
- 414, TITLE: A Survey of Active Learning for Natural Language Processing  
<https://aclanthology.org/2022.emnlp-main.414>  
AUTHORS: Zhisong Zhang, Emma Strubell, Eduard Hovy  
HIGHLIGHT: In this work, we provide a literature review of active learning (AL) for its applications in natural language processing (NLP).
- 415, TITLE: Bernice: A Multilingual Pre-trained Encoder for Twitter  
<https://aclanthology.org/2022.emnlp-main.415>  
AUTHORS: Alexandra DeLucia, Shijie Wu, Aaron Mueller, Carlos Aguirre, Philip Resnik, Mark Dredze  
HIGHLIGHT: We introduce Bernice, the first multilingual RoBERTa language model trained from scratch on 2.
- 416, TITLE: CEFR-Based Sentence Difficulty Annotation and Assessment  
<https://aclanthology.org/2022.emnlp-main.416>  
AUTHORS: Yuki Arase, Satoru Uchida, Tomoyuki Kajiwara  
HIGHLIGHT: To address this problem, we created the CEFR-based Sentence Profile (CEFR-SP) corpus, containing 17k English sentences annotated with the levels based on the Common European Framework of Reference for Languages assigned by English-education professionals.
- 417, TITLE: Simple Questions Generate Named Entity Recognition Datasets  
<https://aclanthology.org/2022.emnlp-main.417>  
AUTHORS: Hyunjae Kim, Jaehyo Yoo, Seunghyun Yoon, Jinhyuk Lee, Jaewoo Kang  
HIGHLIGHT: This work introduces an ask-to-generate approach, which automatically generates NER datasets by asking simple natural language questions to an open-domain question answering system (e. g. ,  $\tilde{A}_i \tilde{A}_j$  Which disease?  $\tilde{A}_i \tilde{A}_j$ )
- 418, TITLE: TemporalWiki: A Lifelong Benchmark for Training and Evaluating Ever-Evolving Language Models  
<https://aclanthology.org/2022.emnlp-main.418>  
AUTHORS: Joel Jang, Seonghyeon Ye, Changho Lee, Sohee Yang, Joongbo Shin, Janghoon Han, Gyeonghun Kim, Minjoon Seo  
HIGHLIGHT: To this end, we introduce TemporalWiki, a lifelong benchmark for ever-evolving LMs that utilizes the difference between consecutive snapshots of English Wikipedia and English Wikidata for training and evaluation, respectively.
- 419, TITLE: Bi-Directional Iterative Prompt-Tuning for Event Argument Extraction  
<https://aclanthology.org/2022.emnlp-main.419>  
AUTHORS: Lu Dai, Bang Wang, Wei Xiang, Yijun Mo  
HIGHLIGHT: In this paper, we propose a bi-directional iterative prompt-tuning method for EAE, where the EAE task is treated as a cloze-style task to take full advantage of entity information and pre-trained language models (PLMs).
- 420, TITLE: Learning Robust Representations for Continual Relation Extraction via Adversarial Class Augmentation  
<https://aclanthology.org/2022.emnlp-main.420>  
AUTHORS: Peiyi Wang, Yifan Song, Tianyu Liu, Binghui Lin, Yunbo Cao, Sujian Li, Zhifang Sui  
HIGHLIGHT: Most previous work attributes catastrophic forgetting to the corruption of the learned representations as new relations come, with an implicit assumption that the CRE models have adequately learned the old relations. In this paper, through empirical studies we argue that this assumption may not hold, and an important reason for catastrophic forgetting is that the learned representations do not have good robustness against the appearance of analogous relations in the subsequent learning process.
- 421, TITLE: ConvFinQA: Exploring the Chain of Numerical Reasoning in Conversational Finance Question Answering  
<https://aclanthology.org/2022.emnlp-main.421>  
AUTHORS: Zhiyu Chen, Shiyang Li, Charese Smiley, Zhiqiang Ma, Sameena Shah, William Yang Wang  
HIGHLIGHT: In this work, we investigate the application domain of finance that involves real-world, complex numerical reasoning.
- 422, TITLE: A Span-based Multimodal Variational Autoencoder for Semi-supervised Multimodal Named Entity Recognition  
<https://aclanthology.org/2022.emnlp-main.422>  
AUTHORS: Baohang Zhou, Ying Zhang, Kehui Song, Wenya Guo, Guoqing Zhao, Hongbin Wang, Xiaojie Yuan  
HIGHLIGHT: To fuse the text and image features for MNER effectively under semi-supervised setting, we propose a novel span-based multimodal variational autoencoder (SMVAE) model for semi-supervised MNER.
- 423, TITLE: R-TeaFor: Regularized Teacher-Forcing for Abstractive Summarization  
<https://aclanthology.org/2022.emnlp-main.423>  
AUTHORS: Guan-Yu Lin, Pu-Jen Cheng  
HIGHLIGHT: Nevertheless, they do not consider the pairwise relationship between the original training data and the modified ones, which provides more information during training. Hence, we propose Regularized Teacher-Forcing (R-TeaFor) to utilize this relationship for better regularization.
- 424, TITLE: Modeling Consistency Preference via Lexical Chains for Document-level Neural Machine Translation  
<https://aclanthology.org/2022.emnlp-main.424>



AUTHORS: Xinglin Lyu, Junhui Li, Shimin Tao, Hao Yang, Ying Qin, Min Zhang  
HIGHLIGHT: In this paper we aim to relieve the issue of lexical translation inconsistency for document-level neural machine translation (NMT) by modeling consistency preference for lexical chains, which consist of repeated words in a source-side document and provide a representation of the lexical consistency structure of the document.

425, TITLE: Just Fine-tune Twice: Selective Differential Privacy for Large Language Models  
<https://aclanthology.org/2022.emnlp-main.425>

AUTHORS: Weiyang Shi, Ryan Shea, Si Chen, Chiyuan Zhang, Ruoxi Jia, Zhou Yu  
HIGHLIGHT: In this paper, we develop a novel framework, \*Just Fine-tune Twice\* (JFT), that achieves SDP for state-of-the-art large transformer-based models.

426, TITLE: Factorizing Content and Budget Decisions in Abstractive Summarization of Long Documents  
<https://aclanthology.org/2022.emnlp-main.426>

AUTHORS: Marcio Fonseca, Yftah Ziser, Shay B. Cohen  
HIGHLIGHT: We argue that disentangling content selection from the budget used to cover salient content improves the performance and applicability of abstractive summarizers.

427, TITLE: Open-Domain Sign Language Translation Learned from Online Video  
<https://aclanthology.org/2022.emnlp-main.427>

AUTHORS: Bowen Shi, Diane Brentari, Gregory Shakhnarovich, Karen Livescu  
HIGHLIGHT: In this paper, we introduce OpenASL, a large-scale American Sign Language (ASL) - English dataset collected from online video sites (e. g. , YouTube).

428, TITLE: Improving Temporal Generalization of Pre-trained Language Models with Lexical Semantic Change  
<https://aclanthology.org/2022.emnlp-main.428>

AUTHORS: Zhaochen Su, Zecheng Tang, Xinyan Guan, Lijun Wu, Min Zhang, Juntao Li  
HIGHLIGHT: In this paper, we empirically observe that temporal generalization is closely affiliated with lexical semantic change, which is one of the essential phenomena of natural languages.

429, TITLE: ULN: Towards Underspecified Vision-and-Language Navigation  
<https://aclanthology.org/2022.emnlp-main.429>

AUTHORS: Weixi Feng, Tsu-Jui Fu, Yujie Lu, William Yang Wang  
HIGHLIGHT: As a primary step toward ULN, we propose a VLN framework that consists of a classification module, a navigation agent, and an Exploitation-to-Exploration (E2E) module.

430, TITLE: Federated Model Decomposition with Private Vocabulary for Text Classification  
<https://aclanthology.org/2022.emnlp-main.430>

AUTHORS: Zhuo Zhang, Xiangjing Hu, Lizhen Qu, Qifan Wang, Zenglin Xu  
HIGHLIGHT: In this paper, we propose a federated model decomposition method that protects the privacy of vocabularies, shorted as FEDEVOCAB.

431, TITLE: ReCo: Reliable Causal Chain Reasoning via Structural Causal Recurrent Neural Networks  
<https://aclanthology.org/2022.emnlp-main.431>

AUTHORS: Kai Xiong, Xiao Ding, Zhongyang Li, Li Du, Ting Liu, Bing Qin, Yi Zheng, Baoxing Huai  
HIGHLIGHT: In other words, the causal pairs to be spliced may have a conflicting threshold boundary or scenario. To address these issues, we propose a novel Reliable Causal chain reasoning framework (ReCo), which introduces exogenous variables to represent the threshold and scene factors of each causal pair within the causal chain, and estimates the threshold and scene contradictions across exogenous variables via structural causal recurrent neural networks (SRNN).

432, TITLE: Video Question Answering: Datasets, Algorithms and Challenges  
<https://aclanthology.org/2022.emnlp-main.432>

AUTHORS: Yaoyao Zhong, Wei Ji, Junbin Xiao, Yicong Li, Weihong Deng, Tat-Seng Chua  
HIGHLIGHT: This survey aims to sort out the recent advances in video question answering (VideoQA) and point towards future directions.

433, TITLE: Retrofitting Multilingual Sentence Embeddings with Abstract Meaning Representation  
<https://aclanthology.org/2022.emnlp-main.433>

AUTHORS: Deng Cai, Xin Li, Jackie Chun-Sing Ho, Lidong Bing, Wai Lam  
HIGHLIGHT: We introduce a new method to improve existing multilingual sentence embeddings with Abstract Meaning Representation (AMR).

434, TITLE: Breaking the Representation Bottleneck of Chinese Characters: Neural Machine Translation with Stroke Sequence Modeling  
<https://aclanthology.org/2022.emnlp-main.434>

AUTHORS: Zhijun Wang, Xuebo Liu, Min Zhang  
HIGHLIGHT: In this paper, we introduce a novel representation method for Chinese characters to break the bottlenecks, namely StrokeNet, which represents a Chinese character by a Latinized stroke sequence (e. g. ,  $\tilde{A}_i^{\tilde{A}_i}$ ).

435, TITLE: Boundary-Driven Table-Filling for Aspect Sentiment Triplet Extraction  
<https://aclanthology.org/2022.emnlp-main.435>

- AUTHORS: Yice Zhang, Yifan Yang, Yihui Li, Bin Liang, Shiwei Chen, Yixue Dang, Min Yang, Ruifeng Xu  
HIGHLIGHT: To overcome these issues, we propose Boundary-Driven Table-Filling (BDTF), which represents each triplet as a relation region in the 2D table and transforms the ASTE task into detection and classification of relation regions.
- 436, TITLE: Attention and Edge-Label Guided Graph Convolutional Networks for Named Entity Recognition  
<https://aclanthology.org/2022.emnlp-main.436>  
AUTHORS: Renjie Zhou, Zhongyi Xie, Jian Wan, Jilin Zhang, Yong Liao, Qiang Liu  
HIGHLIGHT: In this paper, we propose the Attention and Edge-Label guided Graph Convolution Network (AELGCN) model.
- 437, TITLE: Title2Event: Benchmarking Open Event Extraction with a Large-scale Chinese Title Dataset  
<https://aclanthology.org/2022.emnlp-main.437>  
AUTHORS: Haolin Deng, Yanan Zhang, Yangfan Zhang, Wangyang Ying, Changlong Yu, Jun Gao, Wei Wang, Xiaoling Bai, Nan Yang, Jin Ma, Xiang Chen, Tianhua Zhou  
HIGHLIGHT: In this paper, we present Title2Event, a large-scale sentence-level dataset benchmarking Open Event Extraction without restricting event types.
- 438, TITLE: Cascading Biases: Investigating the Effect of Heuristic Annotation Strategies on Data and Models  
<https://aclanthology.org/2022.emnlp-main.438>  
AUTHORS: Chaitanya Malaviya, Sudeep Bhatia, Mark Yatskar  
HIGHLIGHT: In this work, we study cognitive heuristic use in the context of annotating multiple-choice reading comprehension datasets.
- 439, TITLE: Teaching Broad Reasoning Skills for Multi-Step QA by Generating Hard Contexts  
<https://aclanthology.org/2022.emnlp-main.439>  
AUTHORS: Harsh Trivedi, Niranjan Balasubramanian, Tushar Khot, Ashish Sabharwal  
HIGHLIGHT: We show how to use question decompositions to teach language models these broad reasoning skills in a robust fashion.
- 440, TITLE: ADDMU: Detection of Far-Boundary Adversarial Examples with Data and Model Uncertainty Estimation  
<https://aclanthology.org/2022.emnlp-main.440>  
AUTHORS: Fan Yin, Yao Li, Cho-Jui Hsieh, Kai-Wei Chang  
HIGHLIGHT: Existing methods show worse than random guess performance under this scenario. To overcome this limitation, we propose a new technique, ADDMU, adversary detection with data and model uncertainty, which combines two types of uncertainty estimation for both regular and FB adversarial example detection.
- 441, TITLE: G-MAP: General Memory-Augmented Pre-trained Language Model for Domain Tasks  
<https://aclanthology.org/2022.emnlp-main.441>  
AUTHORS: Zhongwei Wan, Yichun Yin, Wei Zhang, Jiabin Shi, Lifeng Shang, Guangyong Chen, Xin Jiang, Qun Liu  
HIGHLIGHT: However, this domain-adaptive pre-training (DAPT (CITATION)) tends to forget the previous general knowledge acquired by general PLMs, which leads to a catastrophic forgetting phenomenon and sub-optimal performance. To alleviate this problem, we propose a new framework of Memory-Augmented Pre-trained Language Model (MAP), which augments the domain-specific PLM by a memory built from the frozen general PLM without losing the general knowledge.
- 442, TITLE: Towards Unifying Reference Expression Generation and Comprehension  
<https://aclanthology.org/2022.emnlp-main.442>  
AUTHORS: Duo Zheng, Tao Kong, Ya Jing, Jiaan Wang, Xiaojie Wang  
HIGHLIGHT: To address the problems, we propose a unified model for REG and REC, named UniRef.
- 443, TITLE: Textual Manifold-based Defense Against Natural Language Adversarial Examples  
<https://aclanthology.org/2022.emnlp-main.443>  
AUTHORS: Dang Nguyen Minh, Anh Tuan Luu  
HIGHLIGHT: In this study, we find a similar phenomenon occurs in the contextualized embedding space of natural sentences induced by pretrained language models in which textual adversarial examples tend to have their embeddings diverge off the manifold of natural sentence embeddings.
- 444, TITLE: Tiny-Attention Adapter: Contexts Are More Important Than the Number of Parameters  
<https://aclanthology.org/2022.emnlp-main.444>  
AUTHORS: Hongyu Zhao, Hao Tan, Hongyuan Mei  
HIGHLIGHT: In this paper, we investigate the effectiveness of using tiny-attention  $\tilde{A}_i \hat{A}_i^{-1}$ , attention with extremely small per-head dimensionality  $\tilde{A}_i \hat{A}_i^{-1}$  as adapters.
- 445, TITLE: Reduce Catastrophic Forgetting of Dense Retrieval Training with Teleportation Negatives  
<https://aclanthology.org/2022.emnlp-main.445>  
AUTHORS: Si Sun, Chenyan Xiong, Yue Yu, Arnold Overwijk, Zhiyuan Liu, Jie Bao  
HIGHLIGHT: In this paper, we investigate the instability in the standard dense retrieval training, which iterates between model training and hard negative selection using the being-trained model.
- 446, TITLE: ATTEMPT: Parameter-Efficient Multi-task Tuning via Attentional Mixtures of Soft Prompts  
<https://aclanthology.org/2022.emnlp-main.446>  
AUTHORS: Akari Asai, Mohammadreza Salehi, Matthew Peters, Hannaneh Hajishirzi

- HIGHLIGHT:** This work introduces a new multi-task, parameter-efficient language model (LM) tuning method that learns to transfer knowledge across different tasks via a mixture of soft prompts and small prefix embedding vectors pre-trained for different tasks.
- 447, **TITLE:** Exploration of the Usage of Color Terms by Color-blind Participants in Online Discussion Platforms  
<https://aclanthology.org/2022.emnlp-main.447>  
**AUTHORS:** Ella Rabinovich, Boaz Carmeli  
**HIGHLIGHT:** We study this question by making a step forward towards a better understanding of the conceptual perception of colors by color-blind individuals, as reflected in their spontaneous linguistic productions.
- 448, **TITLE:** DEER: Descriptive Knowledge Graph for Explaining Entity Relationships  
<https://aclanthology.org/2022.emnlp-main.448>  
**AUTHORS:** Jie Huang, Kerui Zhu, Kevin Chen-Chuan Chang, Jinjun Xiong, Wen-mei Hwu  
**HIGHLIGHT:** To construct DEER, we propose a self-supervised learning method to extract relation descriptions with the analysis of dependency patterns and generate relation descriptions with a transformer-based relation description synthesizing model, where no human labeling is required.
- 449, **TITLE:** META-GUI: Towards Multi-modal Conversational Agents on Mobile GUI  
<https://aclanthology.org/2022.emnlp-main.449>  
**AUTHORS:** Liangtai Sun, Xingyu Chen, Lu Chen, Tianle Dai, Zichen Zhu, Kai Yu  
**HIGHLIGHT:** In this paper, we propose a new TOD architecture: GUI-based task-oriented dialogue system (GUI-TOD).
- 450, **TITLE:** Understanding and Improving Knowledge Distillation for Quantization Aware Training of Large Transformer Encoders  
<https://aclanthology.org/2022.emnlp-main.450>  
**AUTHORS:** Minsoo Kim, Sihwa Lee, Suk-Jin Hong, Du-Seong Chang, Jungwook Choi  
**HIGHLIGHT:** In this work, we provide an in-depth analysis of the mechanism of KD on attention recovery of quantized large Transformers.
- 451, **TITLE:** Exploring Mode Connectivity for Pre-trained Language Models  
<https://aclanthology.org/2022.emnlp-main.451>  
**AUTHORS:** Yujia Qin, Cheng Qian, Jing Yi, Weize Chen, Yankai Lin, Xu Han, Zhiyuan Liu, Maosong Sun, Jie Zhou  
**HIGHLIGHT:** In this paper, we investigate the geometric connections of different minima through the lens of mode connectivity, which measures whether two minima can be connected with a low-loss path.
- 452, **TITLE:** Synergy with Translation Artifacts for Training and Inference in Multilingual Tasks  
<https://aclanthology.org/2022.emnlp-main.452>  
**AUTHORS:** Jaehoon Oh, Jongwoo Ko, Se-Young Yun  
**HIGHLIGHT:** Translation has played a crucial role in improving the performance on multilingual tasks: (1) to generate the target language data from the source language data for training and (2) to generate the source language data from the target language data for inference. However, prior works have not considered the use of both translations simultaneously. This paper shows that combining them can synergize the results on various multilingual sentence classification tasks.
- 453, **TITLE:** Increasing Visual Awareness in Multimodal Neural Machine Translation from an Information Theoretic Perspective  
<https://aclanthology.org/2022.emnlp-main.453>  
**AUTHORS:** Baijun Ji, Tong Zhang, Yicheng Zou, Bojie Hu, Si Shen  
**HIGHLIGHT:** In this paper, we endeavor to improve MMT performance by increasing visual awareness from an information theoretic perspective.
- 454, **TITLE:** Improving Event Coreference Resolution Using Document-level and Topic-level Information  
<https://aclanthology.org/2022.emnlp-main.454>  
**AUTHORS:** Sheng Xu, Peifeng Li, Qiaoming Zhu  
**HIGHLIGHT:** They failed to capture the interactions and contextual cues among those long-distance event mentions. Besides, high-level information, such as event topics, is rarely considered to enhance representation learning for ECR. To address the above two issues, we first apply a Longformer-based encoder to obtain the document-level embeddings and an encoder with a trigger-mask mechanism to learn sentence-level embeddings based on local context. In addition, we propose an event topic generator to infer the latent topic-level representations.
- 455, **TITLE:** Vector-Quantized Input-Contextualized Soft Prompts for Natural Language Understanding  
<https://aclanthology.org/2022.emnlp-main.455>  
**AUTHORS:** Rishabh Bhardwaj, Amrita Saha, Steven C.H. Hoi, Soujanya Poria  
**HIGHLIGHT:** A fixed prompt, however, may not generalize well to the diverse kinds of inputs the task comprises. In order to address this, we propose Vector-quantized Input-contextualized Prompts (VIP) as an extension to the soft prompt tuning framework.
- 456, **TITLE:** Boosting Natural Language Generation from Instructions with Meta-Learning  
<https://aclanthology.org/2022.emnlp-main.456>  
**AUTHORS:** Budhaditya Deb, Ahmed Hassan Awadallah, Guoqing Zheng  
**HIGHLIGHT:** In this paper we investigate whether meta-learning applied to MTIL can further improve generalization to unseen tasks in a zero-shot setting.

- 457, TITLE: Topical Segmentation of Spoken Narratives: A Test Case on Holocaust Survivor Testimonies  
<https://aclanthology.org/2022.emnlp-main.457>  
AUTHORS: Eitan Wagner, Renana Keydar, Amit Pinchevski, Omri Abend  
HIGHLIGHT: We tackle the task of segmenting running (spoken) narratives, which poses hitherto unaddressed challenges.
- 458, TITLE: Unifying the Convergences in Multilingual Neural Machine Translation  
<https://aclanthology.org/2022.emnlp-main.458>  
AUTHORS: Yichong Huang, Xiaocheng Feng, Xinwei Geng, Bing Qin  
HIGHLIGHT: In this paper, we propose a novel training strategy named LSSD (LanguageSpecific Self-Distillation), which can alleviate the convergence inconsistency and help MNMT models achieve the best performance on each language pair simultaneously.
- 459, TITLE: Modeling Label Correlations for Ultra-Fine Entity Typing with Neural Pairwise Conditional Random Field  
<https://aclanthology.org/2022.emnlp-main.459>  
AUTHORS: Chengyue Jiang, Yong Jiang, Weiqi Wu, Pengjun Xie, Kewei Tu  
HIGHLIGHT: To this end, we use an undirected graphical model called pairwise conditional random field (PCRF) to formulate the UFET problem, in which the type variables are not only unarily influenced by the input but also pairwise relate to all the other type variables.
- 460, TITLE: Help me write a Poem - Instruction Tuning as a Vehicle for Collaborative Poetry Writing  
<https://aclanthology.org/2022.emnlp-main.460>  
AUTHORS: Tuhin Chakrabarty, Vishakh Padmakumar, He He  
HIGHLIGHT: Building on the prior success of large language models in the realm of computer assisted creativity, in this work, we present CoPoet, a collaborative poetry writing system, with the goal of to study if LLMs actually improve the quality of the generated content.
- 461, TITLE: Open Relation and Event Type Discovery with Type Abstraction  
<https://aclanthology.org/2022.emnlp-main.461>  
AUTHORS: Sha Li, Heng Ji, Jiawei Han  
HIGHLIGHT: This calls for systems that can automatically infer new types from given corpora, a task which we refer to as type discovery. To tackle this problem, we introduce the idea of type abstraction, where the model is prompted to generalize and name the type.
- 462, TITLE: Enhancing Multilingual Language Model with Massive Multilingual Knowledge Triples  
<https://aclanthology.org/2022.emnlp-main.462>  
AUTHORS: Linlin Liu, Xin Li, Ruidan He, Lidong Bing, Shafiq Joty, Luo Si  
HIGHLIGHT: In this work, we explore methods to make better use of the multilingual annotation and language agnostic property of KG triples, and present novel knowledge based multilingual language models (KMLMs) trained directly on the knowledge triples.
- 463, TITLE: Revisiting Grammatical Error Correction Evaluation and Beyond  
<https://aclanthology.org/2022.emnlp-main.463>  
AUTHORS: Peiyuan Gong, Xuebo Liu, Heyan Huang, Min Zhang  
HIGHLIGHT: To alleviate the limitation, we propose a novel GEC evaluation metric to achieve the best of both worlds, namely PT-M2 which only uses PT-based metrics to score those corrected parts.
- 464, TITLE: R2D2: Robust Data-to-Text with Replacement Detection  
<https://aclanthology.org/2022.emnlp-main.464>  
AUTHORS: Linyong Nan, Lorenzo Jaime Flores, Yilun Zhao, Yixin Liu, Luke Benson, Weijin Zou, Dragomir Radev  
HIGHLIGHT: We introduce R2D2, a training framework that addresses unfaithful Data-to-Text generation by training a system both as a generator and a faithfulness discriminator with additional replacement detection and unlikelihood learning tasks.
- 465, TITLE: IDK-MRC: Unanswerable Questions for Indonesian Machine Reading Comprehension  
<https://aclanthology.org/2022.emnlp-main.465>  
AUTHORS: Rifki Afina Putri, Alice Oh  
HIGHLIGHT: Existing Indonesian MRC datasets (Purwarianti et al. , 2007; Clark et al. , 2020) are still inadequate because of the small size and limited question types, i. e. , they only cover answerable questions. To fill this gap, we build a new Indonesian MRC dataset called Indonesian Know- MRC (IDK-MRC) by combining the automatic and manual unanswerable question generation to minimize the cost of manual dataset construction while maintaining the dataset quality.
- 466, TITLE: XLM-D: Decorate Cross-lingual Pre-training Model as Non-Autoregressive Neural Machine Translation  
<https://aclanthology.org/2022.emnlp-main.466>  
AUTHORS: Yong Wang, Shilin He, Guanhua Chen, Yun Chen, Daxin Jiang  
HIGHLIGHT: In this work, we establish the connection between a pre-trained masked language model (MLM) and non-autoregressive generation on machine translation.
- 467, TITLE: Cross-stitching Text and Knowledge Graph Encoders for Distantly Supervised Relation Extraction  
<https://aclanthology.org/2022.emnlp-main.467>  
AUTHORS: Qin Dai, Benjamin Heinzerling, Kentaro Inui

**HIGHLIGHT:** Here, we introduce cross-stitch bi-encoders, which allow full interaction between the text encoder and the KG encoder via a cross-stitch mechanism.

468, **TITLE:** Assist Non-native Viewers: Multimodal Cross-Lingual Summarization for How2 Videos  
<https://aclanthology.org/2022.emnlp-main.468>

**AUTHORS:** Nayu Liu, Kaiwen Wei, Xian Sun, Hongfeng Yu, Fanglong Yao, Li Jin, Guo Zhi, Guangluan Xu

**HIGHLIGHT:** However, existing works are restricted to monolingual video scenarios, ignoring the demands of non-native video viewers to understand the cross-language videos in practical applications. It stimulates us to propose a new task, named Multimodal Cross-Lingual Summarization for videos (MCLS), which aims to generate cross-lingual summaries from multimodal inputs of videos.

469, **TITLE:** PACIFIC: Towards Proactive Conversational Question Answering over Tabular and Textual Data in Finance  
<https://aclanthology.org/2022.emnlp-main.469>

**AUTHORS:** Yang Deng, Wenqiang Lei, Wenxuan Zhang, Wai Lam, Tat-Seng Chua

**HIGHLIGHT:** To facilitate conversational question answering (CQA) over hybrid contexts in finance, we present a new dataset, named PACIFIC.

470, **TITLE:** Generative Data Augmentation with Contrastive Learning for Zero-Shot Stance Detection

<https://aclanthology.org/2022.emnlp-main.470>

**AUTHORS:** Yang Li, Jiawei Yuan

**HIGHLIGHT:** Among them, one of the important challenges is to reduce the domain transfer between seen and unseen targets. To tackle this problem, we propose a generative data augmentation approach to generate training samples containing targets and stances for testing data, and map the real samples and generated synthetic samples into the same embedding space with contrastive learning, then perform the final classification based on the augmented data.

471, **TITLE:** Better Few-Shot Relation Extraction with Label Prompt Dropout

<https://aclanthology.org/2022.emnlp-main.471>

**AUTHORS:** Peiyuan Zhang, Wei Lu

**HIGHLIGHT:** Instead, we present a novel approach called label prompt dropout, which randomly removes label descriptions in the learning process.

472, **TITLE:** Break it Down into BTS: Basic, Tiniest Subword Units for Korean

<https://aclanthology.org/2022.emnlp-main.472>

**AUTHORS:** Nayeon Kim, Jun-Hyung Park, Joon-Young Choi, Eojin Jeon, Youjin Kang, SangKeun Lee

**HIGHLIGHT:** We introduce Basic, Tiniest Subword (BTS) units for the Korean language, which are inspired by the invention principle of Hangeul, the Korean writing system.

473, **TITLE:** The Devil in Linear Transformer

<https://aclanthology.org/2022.emnlp-main.473>

**AUTHORS:** Zhen Qin, Xiaodong Han, Weixuan Sun, Dongxu Li, Lingpeng Kong, Nick Barnes, Yiran Zhong

**HIGHLIGHT:** However, they usually suffer from degraded performances on various tasks and corpus. In this paper, we examine existing kernel-based linear transformers and identify two key issues that lead to such performance gaps: 1) unbounded gradients in the attention computation adversely impact the convergence of linear transformer models; 2) attention dilution which trivially distributes attention scores over long sequences while neglecting neighbouring structures.

474, **TITLE:** Zero-Shot Learners for Natural Language Understanding via a Unified Multiple Choice Perspective

<https://aclanthology.org/2022.emnlp-main.474>

**AUTHORS:** Ping Yang, Junjie Wang, Ruyi Gan, Xinyu Zhu, Lin Zhang, Ziwei Wu, Xinyu Gao, Jiaying Zhang, Tetsuya Sakai

**HIGHLIGHT:** We propose a new paradigm for zero-shot learners that is format agnostic, i. e. , it is compatible with any format and applicable to a list of language tasks, such as text classification, commonsense reasoning, coreference resolution, and sentiment analysis.

475, **TITLE:** Hypoformer: Hybrid Decomposition Transformer for Edge-friendly Neural Machine Translation

<https://aclanthology.org/2022.emnlp-main.475>

**AUTHORS:** Sunzhu Li, Peng Zhang, Guobing Gan, Xiuqing Lv, Benyou Wang, Junqiu Wei, Xin Jiang

**HIGHLIGHT:** To compress and accelerate Transformer, we propose a Hybrid Tensor-Train (HTT) decomposition, which retains full rank and meanwhile reduces operations and parameters.

476, **TITLE:** FigMemes: A Dataset for Figurative Language Identification in Politically-Opinionated Memes

<https://aclanthology.org/2022.emnlp-main.476>

**AUTHORS:** Chen Liu, Gregor Geigle, Robin Krebs, Iryna Gurevych

**HIGHLIGHT:** To enable future research into this area, we first present FigMemes, a dataset for figurative language classification in politically-opinionated memes.

477, **TITLE:** UniRel: Unified Representation and Interaction for Joint Relational Triple Extraction

<https://aclanthology.org/2022.emnlp-main.477>

**AUTHORS:** Wei Tang, Benfeng Xu, Yuyue Zhao, Zhendong Mao, Yifeng Liu, Yong Liao, Haiyong Xie

**HIGHLIGHT:** Therefore, the rich correlations are not fully exploited by existing works. In this paper, we propose UniRel to address these challenges.

- 478, TITLE: X-FACTOR: A Cross-metric Evaluation of Factual Correctness in Abstractive Summarization  
<https://aclanthology.org/2022.emnlp-main.478>  
AUTHORS: Subhajit Chaudhury, Sarathkrishna Swaminathan, Chulaka Gunasekara, Maxwell Crouse, Srinivas Ravishankar, Daiki Kimura, Keerthiram Murugesan, Ram-Fernandez Astudillo, Tahira Naseem, Pavan Kapanipathi, Alexander Gray  
HIGHLIGHT: In this paper, we present X-FACTOR, a cross-evaluation of three high-performing fact-aware abstractive summarization methods.
- 479, TITLE: ParaTag: A Dataset of Paraphrase Tagging for Fine-Grained Labels, NLG Evaluation, and Data Augmentation  
<https://aclanthology.org/2022.emnlp-main.479>  
AUTHORS: Shuohang Wang, Ruochen Xu, Yang Liu, Chenguang Zhu, Michael Zeng  
HIGHLIGHT: To this end, we propose a novel fine-grained paraphrase annotation schema that labels the minimum spans of tokens in a sentence that don't have the corresponding paraphrases in the other sentence.
- 480, TITLE: Factual Accuracy is not Enough: Planning Consistent Description Order for Radiology Report Generation  
<https://aclanthology.org/2022.emnlp-main.480>  
AUTHORS: Toru Nishino, Yasuhide Miura, Tomoki Taniguchi, Tomoko Ohkuma, Yuki Suzuki, Shoji Kido, Noriyuki Tomiyama  
HIGHLIGHT: We employ a planning-based radiology report generation system that generates the overall structure of reports as a plan prior to generating reports that are accurate and consistent in order. Additionally, we propose a novel reinforcement learning and inference method, Coordinated Planning (CoPlan), that includes a content planner and a text generator to train and infer in a coordinated manner to alleviate the cascading of errors that are often inherent in planning-based models.
- 481, TITLE: FLUTE: Figurative Language Understanding through Textual Explanations  
<https://aclanthology.org/2022.emnlp-main.481>  
AUTHORS: Tuhin Chakrabarty, Arkadiy Saakyan, Debanjan Ghosh, Smaranda Muresan  
HIGHLIGHT: Yet no such data exists for figurative language, making it harder to assess genuine understanding of such expressions. To address this issue, we release FLUTE, a dataset of 9,000 figurative NLI instances with explanations, spanning four categories: Sarcasm, Simile, Metaphor, and Idioms.
- 482, TITLE: Precisely the Point: Adversarial Augmentations for Faithful and Informative Text Generation  
<https://aclanthology.org/2022.emnlp-main.482>  
AUTHORS: Wenhao Wu, Wei Li, Jiachen Liu, Xinyan Xiao, Sujian Li, Yajuan Lyu  
HIGHLIGHT: In this paper, we conduct the first quantitative analysis on the robustness of pre-trained Seq2Seq models.
- 483, TITLE: RLET: A Reinforcement Learning Based Approach for Explainable QA with Entailment Trees  
<https://aclanthology.org/2022.emnlp-main.483>  
AUTHORS: Tengxiao Liu, Qipeng Guo, Xiangkun Hu, Yue Zhang, Xipeng Qiu, Zheng Zhang  
HIGHLIGHT: In this work, we propose RLET, a Reinforcement Learning based Entailment Tree generation framework, which is trained utilising the cumulative signals across the whole tree.
- 484, TITLE: Let the CAT out of the bag: Contrastive Attributed explanations for Text  
<https://aclanthology.org/2022.emnlp-main.484>  
AUTHORS: Saneem Chemmengath, Amar Prakash Azad, Ronny Luss, Amit Dhurandhar  
HIGHLIGHT: In this paper, we propose a method Contrastive Attributed explanations for Text (CAT) which provides contrastive explanations for natural language text data with a novel twist as we build and exploit attribute classifiers leading to more semantically meaningful explanations.
- 485, TITLE: monoQA: Multi-Task Learning of Reranking and Answer Extraction for Open-Retrieval Conversational Question Answering  
<https://aclanthology.org/2022.emnlp-main.485>  
AUTHORS: Sarawoot Kongyoung, Craig Macdonald, Iadh Ounis  
HIGHLIGHT: In this paper, we investigate the use of Multi-Task Learning (MTL) to improve performance on the ORConvQA task by sharing the reranker and reader's learned structure in a generative model.
- 486, TITLE: Composing Ci with Reinforced Non-autoregressive Text Generation  
<https://aclanthology.org/2022.emnlp-main.486>  
AUTHORS: Yan Song  
HIGHLIGHT: Moreover, consider that with the format prepared, Ci generation can be operated by an efficient synchronous process, where autoregressive models are limited in doing so since they follow the character-by-character generation protocol. Therefore, in this paper, we propose to compose Ci through a non-autoregressive approach, which not only ensure that the generation process accommodates tune patterns by controlling the rhythm and essential meaning of each sentence, but also allow the model to perform synchronous generation.
- 487, TITLE: MetaTKG: Learning Evolutionary Meta-Knowledge for Temporal Knowledge Graph Reasoning  
<https://aclanthology.org/2022.emnlp-main.487>  
AUTHORS: Yuwei Xia, Mengqi Zhang, Qiang Liu, Shu Wu, Xiao-Yu Zhang

- HIGHLIGHT:** Since existing models highly rely on historical information to learn embeddings for entities, they perform poorly on such entities with little historical information. To tackle these issues, we propose a novel Temporal Meta-learning framework for TKG reasoning, MetaTKG for brevity.
- 488, **TITLE:** mPLUG: Effective and Efficient Vision-Language Learning by Cross-modal Skip-connections  
<https://aclanthology.org/2022.emnlp-main.488>  
**AUTHORS:** Chenliang Li, Haiyang Xu, Junfeng Tian, Wei Wang, Ming Yan, Bin Bi, Jiabo Ye, He Chen, Guohai Xu, Zheng Cao, Ji Zhang, Songfang Huang, Fei Huang, Jingren Zhou, Luo Si  
**HIGHLIGHT:** This paper presents mPLUG, a new vision-language foundation model for both cross-modal understanding and generation.
- 489, **TITLE:** Q-TOD: A Query-driven Task-oriented Dialogue System  
<https://aclanthology.org/2022.emnlp-main.489>  
**AUTHORS:** Xin Tian, Yingzhan Lin, Mengfei Song, Siqi Bao, Fan Wang, Huang He, Shuqi Sun, Hua Wu  
**HIGHLIGHT:** In this paper, we introduce a novel query-driven task-oriented dialogue system, namely Q-TOD.
- 490, **TITLE:** Dial2vec: Self-Guided Contrastive Learning of Unsupervised Dialogue Embeddings  
<https://aclanthology.org/2022.emnlp-main.490>  
**AUTHORS:** Che Liu, Rui Wang, Junfeng Jiang, Yongbin Li, Fei Huang  
**HIGHLIGHT:** In this paper, we introduce the task of learning unsupervised dialogue embeddings.
- 491, **TITLE:** WR-One2Set: Towards Well-Calibrated Keyphrase Generation  
<https://aclanthology.org/2022.emnlp-main.491>  
**AUTHORS:** Binbin Xie, Xiangpeng Wei, Baosong Yang, Huan Lin, Jun Xie, Xiaoli Wang, Min Zhang, Jinsong Su  
**HIGHLIGHT:** Nevertheless, we observe serious calibration errors outputted by ONE2SET, especially in the over-estimation of  $\hat{A}_i^{\frac{1}{2}}$  token (means  $\hat{A}_i^{\frac{1}{2}}$ no corresponding keyphrase $\hat{A}_i^{\frac{1}{2}}$ ). In this paper, we deeply analyze this limitation and identify two main reasons behind: 1) the parallel generation has to introduce excessive  $\hat{A}_i^{\frac{1}{2}}$  as padding tokens into training instances; and 2) the training mechanism assigning target to each slot is unstable and further aggravates the  $\hat{A}_i^{\frac{1}{2}}$  token over-estimation.
- 492, **TITLE:** Eeny, meeny, miny, moe. How to choose data for morphological inflection.  
<https://aclanthology.org/2022.emnlp-main.492>  
**AUTHORS:** Saliha Muradoglu, Mans Hulden  
**HIGHLIGHT:** In this paper, we explore four sampling strategies for the task of morphological inflection using a Transformer model: a pair of oracle experiments where data is chosen based on correct/incorrect predictions by the model, model confidence, entropy, and random selection.
- 493, **TITLE:** An Adaptive Logical Rule Embedding Model for Inductive Reasoning over Temporal Knowledge Graphs  
<https://aclanthology.org/2022.emnlp-main.493>  
**AUTHORS:** Xin Mei, Libin Yang, Xiaoyan Cai, Zuwei Jiang  
**HIGHLIGHT:** We combine the two methods to capture deep causal logic by learning rule embeddings, and propose an interpretable model for temporal knowledge graph reasoning called adaptive logical rule embedding model for inductive reasoning (ALRE-IR).
- 494, **TITLE:** UniNL: Aligning Representation Learning with Scoring Function for OOD Detection via Unified Neighborhood Learning  
<https://aclanthology.org/2022.emnlp-main.494>  
**AUTHORS:** Yutao Mou, Pei Wang, Keqing He, Yanan Wu, Jingang Wang, Wei Wu, Weiran Xu  
**HIGHLIGHT:** In this paper, we propose a unified neighborhood learning framework (UniNL) to detect OOD intents.
- 495, **TITLE:** Open-domain Video Commentary Generation  
<https://aclanthology.org/2022.emnlp-main.495>  
**AUTHORS:** Edison Marrese-Taylor, Yumi Hamazono, Tatsuya Ishigaki, Goran Topic, Yusuke Miyao, Ichiro Kobayashi, Hiroya Takamura  
**HIGHLIGHT:** We detail the construction of a new large-scale dataset of transcribed commentary aligned with videos containing various human actions in a variety of domains, and propose approaches based on well-known neural architectures to tackle the task.
- 496, **TITLE:** One size does not fit all: Investigating strategies for differentially-private learning across NLP tasks  
<https://aclanthology.org/2022.emnlp-main.496>  
**AUTHORS:** Manuel Senge, Timour Igamberdiev, Ivan Habernal  
**HIGHLIGHT:** In this short paper, we provide an extensive analysis of different privacy preserving strategies on seven downstream datasets in five different  $\hat{A}_i^{\frac{1}{2}}$ typical $\hat{A}_i^{\frac{1}{2}}$  NLP tasks with varying complexity using modern neural models based on BERT and XtremeDistil architectures.
- 497, **TITLE:** Counterfactual Recipe Generation: Exploring Compositional Generalization in a Realistic Scenario  
<https://aclanthology.org/2022.emnlp-main.497>  
**AUTHORS:** Xiao Liu, Yansong Feng, Jizhi Tang, Chengang Hu, Dongyan Zhao  
**HIGHLIGHT:** In this paper, we investigate whether pretrained language models can perform compositional generalization in a realistic setting: recipe generation.

- 498, TITLE: Tutoring Helps Students Learn Better: Improving Knowledge Distillation for BERT with Tutor Network  
<https://aclanthology.org/2022.emnlp-main.498>  
AUTHORS: Junho Kim, Jun-Hyung Park, Mingyu Lee, Wing-Lam Mok, Joon-Young Choi, SangKeun Lee  
HIGHLIGHT: In this paper, we propose a novel KD framework, Tutor-KD, which improves the distillation effectiveness by controlling the difficulty of training examples during pre-training.
- 499, TITLE: Does Corpus Quality Really Matter for Low-Resource Languages?  
<https://aclanthology.org/2022.emnlp-main.499>  
AUTHORS: Mikel Artetxe, Itziar Aldabe, Rodrigo Agerri, Olatz Perez-de-Vi<sup>^-</sup>^<sup>^</sup>aspre, Aitor Soroa  
HIGHLIGHT: Taking representation learning in Basque as a case study, we explore tailored crawling (manually identifying and scraping websites with high-quality content) as an alternative to filtering CommonCrawl.
- 500, TITLE: Unifying Data Perspectivism and Personalization: An Application to Social Norms  
<https://aclanthology.org/2022.emnlp-main.500>  
AUTHORS: Joan Plepi, BA<sup>^-</sup>^<sup>^</sup>la Neuendorf, Lucie Flek, Charles Welch  
HIGHLIGHT: In this work, we examine a corpus of social media posts about conflict from a set of 13k annotators and 210k judgements of social norms.
- 501, TITLE: Does Self-Rationalization Improve Robustness to Spurious Correlations?  
<https://aclanthology.org/2022.emnlp-main.501>  
AUTHORS: Alexis Ross, Matthew Peters, Ana Marasovic  
HIGHLIGHT: Specifically, we evaluate how training self-rationalization models with free-text rationales affects robustness to spurious correlations in fine-tuned encoder-decoder and decoder-only models of six different sizes.
- 502, TITLE: Efficient Pre-training of Masked Language Model via Concept-based Curriculum Masking  
<https://aclanthology.org/2022.emnlp-main.502>  
AUTHORS: Mingyu Lee, Jun-Hyung Park, Junho Kim, Kang-Min Kim, SangKeun Lee  
HIGHLIGHT: In this paper, we propose a novel concept-based curriculum masking (CCM) method to efficiently pre-train a language model.
- 503, TITLE: Subword Evenness (SuE) as a Predictor of Cross-lingual Transfer to Low-resource Languages  
<https://aclanthology.org/2022.emnlp-main.503>  
AUTHORS: Olga Pelloni, Anastassia Shaitarova, Tanja Samardzic  
HIGHLIGHT: In this study, we show that languages written in non-Latin and non-alphabetic scripts (mostly Asian languages) are the best choices for improving performance on the task of Masked Language Modelling (MLM) in a diverse set of 30 low-resource languages and that the success of the transfer is well predicted by our novel measure of Subword Evenness (SuE).
- 504, TITLE: A Unified Neural Network Model for Readability Assessment with Feature Projection and Length-Balanced Loss  
<https://aclanthology.org/2022.emnlp-main.504>  
AUTHORS: Wenbiao Li, Wang Ziyang, Yunfang Wu  
HIGHLIGHT: In this paper, we propose a BERT-based model with feature projection and length-balanced loss (BERT-FP-LBL) to determine the difficulty level of a given text.
- 505, TITLE: Speaker Overlap-aware Neural Diarization for Multi-party Meeting Analysis  
<https://aclanthology.org/2022.emnlp-main.505>  
AUTHORS: Zhihao Du, ShiLiang Zhang, Siqi Zheng, Zhi-Jie Yan  
HIGHLIGHT: To overcome the disadvantages, we reformulate overlapped speaker diarization task as a single-label prediction problem via the proposed power set encoding (PSE).
- 506, TITLE: GREENER: Graph Neural Networks for News Media Profiling  
<https://aclanthology.org/2022.emnlp-main.506>  
AUTHORS: Panayot Panayotov, Utsav Shukla, Husrev Taha Sencar, Mohamed Nabeel, Preslav Nakov  
HIGHLIGHT: We study the problem of profiling news media on the Web with respect to their factuality of reporting and bias.
- 507, TITLE: Graph Hawkes Transformer for Extrapolated Reasoning on Temporal Knowledge Graphs  
<https://aclanthology.org/2022.emnlp-main.507>  
AUTHORS: Haohai Sun, Shangyi Geng, Jialun Zhong, Han Hu, Kun He  
HIGHLIGHT: To this end, we propose a Graph Hawkes Transformer (GHT) for both TKG entity prediction and time prediction tasks in the future time.
- 508, TITLE: UniRPG: Unified Discrete Reasoning over Table and Text as Program Generation  
<https://aclanthology.org/2022.emnlp-main.508>  
AUTHORS: Yongwei Zhou, Junwei Bao, Chaoqun Duan, Youzheng Wu, Xiaodong He, Tiejun Zhao  
HIGHLIGHT: In this paper, we propose UniRPG, a semantic-parsing-based approach advanced in interpretability and scalability, to perform Unified discrete Reasoning over heterogeneous knowledge resources, i. e. , table and text, as Program Generation.
- 509, TITLE: Don<sup>^-</sup>^<sup>^</sup>t Prompt, Search! Mining-based Zero-Shot Learning with Language Models  
<https://aclanthology.org/2022.emnlp-main.509>



- AUTHORS:** Mozes van de Kar, Mengzhou Xia, Danqi Chen, Mikel Artetxe  
**HIGHLIGHT:** In this paper, we propose an alternative mining-based approach for zero-shot learning.
- 510, **TITLE:** SEMGraph: Incorporating Sentiment Knowledge and Eye Movement into Graph Model for Sentiment Analysis  
<https://aclanthology.org/2022.emnlp-main.510>  
**AUTHORS:** Bingbing Wang, Bin Liang, Jiachen Du, Min Yang, Ruifeng Xu  
**HIGHLIGHT:** This paper investigates the sentiment analysis task from a novel perspective by incorporating sentiment knowledge and eye movement into a graph architecture, aiming to draw the eye movement-based sentiment relationships for learning the sentiment expression of the context.
- 511, **TITLE:** Cross-lingual neural fuzzy matching for exploiting target-language monolingual corpora in computer-aided translation  
<https://aclanthology.org/2022.emnlp-main.511>  
**AUTHORS:** Miquel Espinall-Gomis, Víctor M. Sánchez-Cartagena, Juan Antonio Pérez-Ortiz, Felipe Sánchez-Martínez  
**HIGHLIGHT:** However, the reduced availability of in-domain TMs, as compared to in-domain monolingual corpora, limits its adoption for a number of translation tasks. In this paper, we introduce a novel neural approach aimed at overcoming this limitation by exploiting not only TMs, but also in-domain target-language (TL) monolingual corpora, and still enabling a similar functionality to that offered by conventional TM-based CAT tools.
- 512, **TITLE:** Multi-Label Intent Detection via Contrastive Task Specialization of Sentence Encoders  
<https://aclanthology.org/2022.emnlp-main.512>  
**AUTHORS:** Ivan Vulic, Iñigo Casanueva, Georgios Spithourakis, Avishek Mondal, Tsung-Hsien Wen, Paweł Budzianowski  
**HIGHLIGHT:** Deploying task-oriented dialog ToD systems for new domains and tasks requires natural language understanding models that are 1) resource-efficient and work under low-data regimes; 2) adaptable, efficient, and quick-to-train; 3) expressive and can handle complex ToD scenarios with multiple user intents in a single utterance. Motivated by these requirements, we introduce a novel framework for multi-label intent detection (mID): Multi-ConvFiT (Multi-Label Intent Detection via Contrastive Conversational Fine-Tuning).
- 513, **TITLE:** Discovering Language-neutral Sub-networks in Multilingual Language Models  
<https://aclanthology.org/2022.emnlp-main.513>  
**AUTHORS:** Negar Foroutan, Mohammadreza Banaei, Rémi Lebreton, Antoine Bosselut, Karl Aberer  
**HIGHLIGHT:** However, the extent to which they learn language-neutral representations (i. e. , shared representations that encode similar phenomena across languages), and the effect of such representations on cross-lingual transfer performance, remain open questions. In this work, we conceptualize language neutrality of multilingual models as a function of the overlap between language-encoding sub-networks of these models.
- 514, **TITLE:** Parameter-Efficient Tuning Makes a Good Classification Head  
<https://aclanthology.org/2022.emnlp-main.514>  
**AUTHORS:** Zhuoyi Yang, Ming Ding, Yanhui Guo, Qingsong Lv, Jie Tang  
**HIGHLIGHT:** In this paper, we find that parameter-efficient tuning makes a good classification head, with which we can simply replace the randomly initialized heads for a stable performance gain.
- 515, **TITLE:** STGN: an Implicit Regularization Method for Learning with Noisy Labels in Natural Language Processing  
<https://aclanthology.org/2022.emnlp-main.515>  
**AUTHORS:** Tingting Wu, Xiao Ding, Minji Tang, Hao Zhang, Bing Qin, Ting Liu  
**HIGHLIGHT:** However, previous studies exert identical perturbation for all samples, which may cause overfitting on incorrect ones or optimizing correct ones inadequately. To facilitate this, we propose a novel stochastic tailor-made gradient noise (STGN), mitigating the effect of inherent label noise by introducing tailor-made benign noise for each sample.
- 516, **TITLE:** Cross-Modal Similarity-Based Curriculum Learning for Image Captioning  
<https://aclanthology.org/2022.emnlp-main.516>  
**AUTHORS:** Hongkuan Zhang, Saku Sugawara, Akiko Aizawa, Lei Zhou, Ryohei Sasano, Koichi Takeda  
**HIGHLIGHT:** In this paper, we propose a simple yet efficient difficulty measurement for image captioning using cross-modal similarity calculated by a pretrained vision-language model.
- 517, **TITLE:** Debiasing Masks: A New Framework for Shortcut Mitigation in NLU  
<https://aclanthology.org/2022.emnlp-main.517>  
**AUTHORS:** Johannes Mario Meissner, Saku Sugawara, Akiko Aizawa  
**HIGHLIGHT:** We propose a new debiasing method in which we identify debiased pruning masks that can be applied to a finetuned model.
- 518, **TITLE:** Extending Phrase Grounding with Pronouns in Visual Dialogues  
<https://aclanthology.org/2022.emnlp-main.518>  
**AUTHORS:** Panzhong Lu, Xin Zhang, Meishan Zhang, Min Zhang  
**HIGHLIGHT:** First, we construct a dataset of phrase grounding with both noun phrases and pronouns to image regions. Based on the dataset, we test the performance of phrase grounding by using a state-of-the-art literature model of this line. Then, we enhance the baseline grounding model with coreference information which should help our task potentially, modeling the coreference structures with graph convolutional networks.

- 519, TITLE: EUR-Lex-Sum: A Multi- and Cross-lingual Dataset for Long-form Summarization in the Legal Domain  
<https://aclanthology.org/2022.emnlp-main.519>  
AUTHORS: Dennis Aumiller, Ashish Chouhan, Michael Gertz  
HIGHLIGHT: In this work, we propose a novel dataset, called EUR-Lex-Sum, based on manually curated document summaries of legal acts from the European Union law platform (EUR-Lex).
- 520, TITLE: Differentiable Data Augmentation for Contrastive Sentence Representation Learning  
<https://aclanthology.org/2022.emnlp-main.520>  
AUTHORS: Tianduo Wang, Wei Lu  
HIGHLIGHT: Although the contrastive learning framework has shown its superiority on sentence representation learning over previous methods, the potential of such a framework is under-explored so far due to the simple method it used to construct positive pairs. Motivated by this, we propose a method that makes hard positives from the original training examples.
- 521, TITLE: Text Style Transferring via Adversarial Masking and Styled Filling  
<https://aclanthology.org/2022.emnlp-main.521>  
AUTHORS: Jiarui Wang, Richong Zhang, Junfan Chen, Jacin Kim, Yongyi Mao  
HIGHLIGHT: To tackle both challenges, in this study, we propose a style transfer model, with an adversarial masking approach and a styled filling technique (AMSF).
- 522, TITLE: Character-level White-Box Adversarial Attacks against Transformers via Attachable Subwords Substitution  
<https://aclanthology.org/2022.emnlp-main.522>  
AUTHORS: Aiwei Liu, Honghai Yu, Xuming Hu, Shu $\tilde{A}$  $\tilde{A}$  $\tilde{A}$ ang Li, Li Lin, Fukun Ma, Yawen Yang, Lijie Wen  
HIGHLIGHT: We propose the first character-level white-box adversarial attack method against transformer models.
- 523, TITLE: Query-based Instance Discrimination Network for Relational Triple Extraction  
<https://aclanthology.org/2022.emnlp-main.523>  
AUTHORS: Zeqi Tan, Yongliang Shen, Xuming Hu, Wenqi Zhang, Xiaoxia Cheng, Weiming Lu, Yueting Zhuang  
HIGHLIGHT: However, they still suffer from error propagation, relation redundancy and lack of high-level connections between triples. To address these issues, we propose a novel query-based approach to construct instance-level representations for relational triples.
- 524, TITLE: Learning Inter-Entity-Interaction for Few-Shot Knowledge Graph Completion  
<https://aclanthology.org/2022.emnlp-main.524>  
AUTHORS: Yuling Li, Kui Yu, Xiaoling Huang, Yuhong Zhang  
HIGHLIGHT: Such practice, however, ignores the inter-entity interaction, resulting in low-discrimination representations for entity pairs, especially when these entity pairs are associated with 1-to-N, N-to-1, and N-to-N relations. To address this issue, this paper proposes a novel FKGC model, named Cross-Interaction Attention Network (CIAN) to investigate the inter-entity interaction between head and tail entities.
- 525, TITLE: Empowering the Fact-checkers! Automatic Identification of Claim Spans on Twitter  
<https://aclanthology.org/2022.emnlp-main.525>  
AUTHORS: Megha Sundriyal, Atharva Kulkarni, Vaibhav Pulastya, Md. Shad Akhtar, Tanmoy Chakraborty  
HIGHLIGHT: In this work, we introduce the novel task of Claim Span Identification (CSI).
- 526, TITLE: ClidSum: A Benchmark Dataset for Cross-Lingual Dialogue Summarization  
<https://aclanthology.org/2022.emnlp-main.526>  
AUTHORS: Jiaan Wang, Fandong Meng, Ziyao Lu, Duo Zheng, Zhixu Li, Jianfeng Qu, Jie Zhou  
HIGHLIGHT: We present ClidSum, a benchmark dataset towards building cross-lingual summarization systems on dialogue documents.
- 527, TITLE: Spectral Probing  
<https://aclanthology.org/2022.emnlp-main.527>  
AUTHORS: Max M $\tilde{A}$  $\tilde{A}$  $\tilde{A}$ ller-Eberstein, Rob van der Goot, Barbara Plank  
HIGHLIGHT: Contextualized embeddings have analogously been found to capture these phenomena at distinctive layers and frequencies. Leveraging these findings, we develop a fully learnable frequency filter to identify spectral profiles for any given task.
- 528, TITLE: QASem Parsing: Text-to-text Modeling of QA-based Semantics  
<https://aclanthology.org/2022.emnlp-main.528>  
AUTHORS: Ayal Klein, Eran Hirsch, Ron Eliav, Valentina Pyatkin, Avi Caciularu, Ido Dagan  
HIGHLIGHT: More recently, an appealing trend introduces semi-structured natural-language structures as an intermediate meaning-capturing representation, often in the form of questions and answers. In this work, we further promote this line of research by considering three prior QA-based semantic representations.
- 529, TITLE: Keypphrase Generation via Soft and Hard Semantic Corrections  
<https://aclanthology.org/2022.emnlp-main.529>  
AUTHORS: Guangzhen Zhao, Guoshun Yin, Peng Yang, Yu Yao  
HIGHLIGHT: To tackle the above biases, we propose a novel correction model CorrKG on top of the MLE pipeline, where the biases are corrected via the optimal transport (OT) and a frequency-based filtering-and-sorting (FreqFS) strategy.

- 530, TITLE: Modal-specific Pseudo Query Generation for Video Corpus Moment Retrieval  
<https://aclanthology.org/2022.emnlp-main.530>  
AUTHORS: Minjoon Jung, SeongHo Choi, JooChan Kim, Jin-Hwa Kim, Byoung-Tak Zhang  
HIGHLIGHT: Previous works have shown promising results; however, they relied on the expensive query annotations for the VCMR, i. e. , the corresponding moment intervals. To overcome this problem, we propose a self-supervised learning framework: Modal-specific Pseudo Query Generation Network (MPGN).
- 531, TITLE: DuQM: A Chinese Dataset of Linguistically Perturbed Natural Questions for Evaluating the Robustness of Question Matching Models  
<https://aclanthology.org/2022.emnlp-main.531>  
AUTHORS: Hongyu Zhu, Yan Chen, Jing Yan, Jing Liu, Yu Hong, Ying Chen, Hua Wu, Haifeng Wang  
HIGHLIGHT: In this paper, we focus on the robustness evaluation of Chinese Question Matching (QM) models.
- 532, TITLE: DivEMT: Neural Machine Translation Post-Editing Effort Across Typologically Diverse Languages  
<https://aclanthology.org/2022.emnlp-main.532>  
AUTHORS: Gabriele Sarti, Arianna Bisazza, Ana Guerberof-Arenas, Antonio Toral  
HIGHLIGHT: We introduce DivEMT, the first publicly available post-editing study of Neural Machine Translation (NMT) over a typologically diverse set of target languages.
- 533, TITLE: Bridging Fairness and Environmental Sustainability in Natural Language Processing  
<https://aclanthology.org/2022.emnlp-main.533>  
AUTHORS: Marius Hesselthaler, Emma Strubell, Dirk Hovy, Anne Lauscher  
HIGHLIGHT: This lacuna is highly problematic, since there is increasing evidence that an exclusive focus on fairness can actually hinder environmental sustainability, and vice versa. In this work, we shed light on this crucial intersection in NLP by (1) investigating the efficiency of current fairness approaches through surveying example methods for reducing unfair stereotypical bias from the literature, and (2) evaluating a common technique to reduce energy consumption (and thus environmental impact) of English NLP models, knowledge distillation (KD), for its impact on fairness.
- 534, TITLE: UniMSE: Towards Unified Multimodal Sentiment Analysis and Emotion Recognition  
<https://aclanthology.org/2022.emnlp-main.534>  
AUTHORS: Guimin Hu, Ting-En Lin, Yi Zhao, Guangming Lu, Yuchuan Wu, Yongbin Li  
HIGHLIGHT: In this paper, we propose a multimodal sentiment knowledge-sharing framework (UniMSE) that unifies MSA and ERC tasks from features, labels, and models.
- 535, TITLE: Is the Brain Mechanism for Hierarchical Structure Building Universal Across Languages? An fMRI Study of Chinese and English  
<https://aclanthology.org/2022.emnlp-main.535>  
AUTHORS: Xiaohan Zhang, Shaonan Wang, Nan Lin, Chengqing Zong  
HIGHLIGHT: In this paper, we first analyze the differences in language structure between two diverse languages: Chinese and English. By computing the working memory requirements when applying parsing strategies to different language structures, we find that top-down parsing generates less memory load for the right-branching English and bottom-up parsing is less memory-demanding for Chinese.
- 536, TITLE: HashFormers: Towards Vocabulary-independent Pre-trained Transformers  
<https://aclanthology.org/2022.emnlp-main.536>  
AUTHORS: Huiyin Xue, Nikolaos Aletras  
HIGHLIGHT: However, these methods are not pre-trained. Inspired by this line of work, we propose HashFormers, a new family of vocabulary-independent pre-trained transformers that support an unlimited vocabulary (i. e. all possible tokens in a corpus) given a substantially smaller fixed-sized embedding matrix.
- 537, TITLE: MatchPrompt: Prompt-based Open Relation Extraction with Semantic Consistency Guided Clustering  
<https://aclanthology.org/2022.emnlp-main.537>  
AUTHORS: Jiaxin Wang, Lingling Zhang, Jun Liu, Xi Liang, Yujie Zhong, Yaqiang Wu  
HIGHLIGHT: In this work, we propose a new prompt-based framework named MatchPrompt, which can realize OpenRE with efficient knowledge transfer from only a few pre-defined relational instances as well as mine the specific meanings for cluster interpretability.
- 538, TITLE: Improving Aspect Sentiment Quad Prediction via Template-Order Data Augmentation  
<https://aclanthology.org/2022.emnlp-main.538>  
AUTHORS: Mengting Hu, Yike Wu, Hang Gao, Yin hao Bai, Shiwan Zhao  
HIGHLIGHT: Specifically, we use the pre-trained language model to select the orders with minimal entropy.
- 539, TITLE: SocioProbe: What, When, and Where Language Models Learn about Sociodemographics  
<https://aclanthology.org/2022.emnlp-main.539>  
AUTHORS: Anne Lauscher, Federico Bianchi, Samuel R. Bowman, Dirk Hovy  
HIGHLIGHT: We address this research gap by probing the sociodemographic knowledge of different single-GPU PLMs on multiple English data sets via traditional classifier probing and information-theoretic minimum description length probing.
- 540, TITLE: When does Parameter-Efficient Transfer Learning Work for Machine Translation?  
<https://aclanthology.org/2022.emnlp-main.540>

AUTHORS: Ahmet Attar, Asa Cooper Stickland  
HIGHLIGHT: We conduct a comprehensive empirical study of PEFTs for MT, considering (1) various parameter budgets, (2) a diverse set of language-pairs, and (3) different pre-trained models.

541, TITLE: Hyper-X: A Unified Hypernetwork for Multi-Task Multilingual Transfer  
<https://aclanthology.org/2022.emnlp-main.541>

AUTHORS: Ahmet Attar, Arianna Bisazza, Gosse Bouma, Gertjan van Noord, Sebastian Ruder  
HIGHLIGHT: However, existing methods are unable to fully leverage training data when it is available in different task-language combinations. To exploit such heterogeneous supervision, we propose Hyper-X, a single hypernetwork that unifies multi-task and multilingual learning with efficient adaptation.

542, TITLE: Towards Robust Numerical Question Answering: Diagnosing Numerical Capabilities of NLP Systems  
<https://aclanthology.org/2022.emnlp-main.542>

AUTHORS: Jialiang Xu, Mengyu Zhou, Xinyi He, Shi Han, Dongmei Zhang  
HIGHLIGHT: In this paper, we propose to conduct numerical capability diagnosis on a series of Numerical Question Answering systems and datasets.

543, TITLE: Enhancing Joint Multiple Intent Detection and Slot Filling with Global Intent-Slot Co-occurrence  
<https://aclanthology.org/2022.emnlp-main.543>

AUTHORS: Mengxiao Song, Bowen Yu, Li Quangang, Wang Yubin, Tingwen Liu, Hongbo Xu  
HIGHLIGHT: In this paper, we aim to make full use of the statistical co-occurrence frequency between intents and slots as prior knowledge to enhance joint multiple intent detection and slot filling.

544, TITLE: Towards Pragmatic Production Strategies for Natural Language Generation Tasks  
<https://aclanthology.org/2022.emnlp-main.544>

AUTHORS: Mario Giulianelli  
HIGHLIGHT: This position paper proposes a conceptual framework for the design of Natural Language Generation (NLG) systems that follow efficient and effective production strategies in order to achieve complex communicative goals.

545, TITLE: LiteVL: Efficient Video-Language Learning with Enhanced Spatial-Temporal Modeling  
<https://aclanthology.org/2022.emnlp-main.545>

AUTHORS: Dongsheng Chen, Chaofan Tao, Lu Hou, Lifeng Shang, Xin Jiang, Qun Liu  
HIGHLIGHT: However, the pre-training process is computationally expensive due to the requirement of millions of video-text pairs and the redundant data structure of each video. To mitigate these problems, we propose LiteVL, which adapts a pre-trained image-language model BLIP into a video-text model directly on downstream tasks, without heavy pre-training.

546, TITLE: Communication breakdown: On the low mutual intelligibility between human and neural captioning  
<https://aclanthology.org/2022.emnlp-main.546>

AUTHORS: Roberto Dessi, Eleonora Gualdoni, Francesca Franzon, Gemma Boleda, Marco Baroni  
HIGHLIGHT: We compare the 0-shot performance of a neural caption-based image retriever when given as input either human-produced captions or captions generated by a neural captioner.

547, TITLE: Normalizing Mutual Information for Robust Adaptive Training for Translation  
<https://aclanthology.org/2022.emnlp-main.547>

AUTHORS: Youngwon Lee, Changmin Lee, Hojin Lee, Seung-won Hwang  
HIGHLIGHT: The score is obtained by combining the probability from the translation model and the target language model, which is then used to assign different weights to losses from sentences and tokens. Meanwhile, we argue this metric is not properly normalized, for which we propose Normalized Pointwise Mutual Information (NPMI).

548, TITLE: Bilingual Synchronization: Restoring Translational Relationships with Editing Operations  
<https://aclanthology.org/2022.emnlp-main.548>

AUTHORS: Jitao Xu, Josep Crego, François Yvon  
HIGHLIGHT: We consider here a more general setting which assumes an initial target sequence, that must be transformed into a valid translation of the source, thereby restoring parallelism between source and target.

549, TITLE: Human-Machine Collaboration Approaches to Build a Dialogue Dataset for Hate Speech Countering  
<https://aclanthology.org/2022.emnlp-main.549>

AUTHORS: Helena Bonaldi, Sara Dellantonio, Serra Sinem Tekiroglu, Marco Guerini  
HIGHLIGHT: In this paper, we present a hybrid approach for dialogical data collection, which combines the intervention of human expert annotators over machine generated dialogues obtained using 19 different configurations.

550, TITLE: JANUS: Joint Autoregressive and Non-autoregressive Training with Auxiliary Loss for Sequence Generation  
<https://aclanthology.org/2022.emnlp-main.550>

AUTHORS: Xiaobo Liang, Lijun Wu, Juntao Li, Min Zhang  
HIGHLIGHT: In this paper, we propose JANUS, a Joint Autoregressive and Non-autoregressive training method using aUxiliary losS to enhance the model performance in both AR and NAR manner simultaneously and effectively alleviate the problem of distribution discrepancy.

551, TITLE: Entity-Focused Dense Passage Retrieval for Outside-Knowledge Visual Question Answering  
<https://aclanthology.org/2022.emnlp-main.551>

- AUTHORS: Jialin Wu, Raymond Mooney  
HIGHLIGHT: Also, the naturally available supervision (whether the passage contains the correct answer) is weak and does not guarantee question relevancy. To address these issues, we propose an Entity-Focused Retrieval (EnFoRe) model that provides stronger supervision during training and recognizes question-relevant entities to help retrieve more specific knowledge.
- 552, TITLE: Cross-Linguistic Syntactic Difference in Multilingual BERT: How Good is It and How Does It Affect Transfer?  
<https://aclanthology.org/2022.emnlp-main.552>  
AUTHORS: Ningyu Xu, Tao Gui, Ruotian Ma, Qi Zhang, Jingting Ye, Menghan Zhang, Xuanjing Huang  
HIGHLIGHT: In this work, we investigate the distributions of grammatical relations induced from mBERT in the context of 24 typologically different languages.
- 553, TITLE: It's Not Just Hate: A Multi-Dimensional Perspective on Detecting Harmful Speech Online  
<https://aclanthology.org/2022.emnlp-main.553>  
AUTHORS: Federico Bianchi, Stefanie Hills, Patricia Rossini, Dirk Hovy, Rebekah Tromble, Nava Tintarev  
HIGHLIGHT: We show that a more fine-grained multi-label approach to predicting incivility and hateful or intolerant content addresses both conceptual and performance issues.
- 554, TITLE: Long Text Generation with Topic-aware Discrete Latent Variable Model  
<https://aclanthology.org/2022.emnlp-main.554>  
AUTHORS: Erguang Yang, Mingtong Liu, Deyi Xiong, Yujie Zhang, Yufeng Chen, Jinan Xu  
HIGHLIGHT: In this work, we investigate whether discrete latent codes can learn information of topics.
- 555, TITLE: TIARA: Multi-grained Retrieval for Robust Question Answering over Large Knowledge Base  
<https://aclanthology.org/2022.emnlp-main.555>  
AUTHORS: Yiheng Shu, Zhiwei Yu, Yuhan Li, Björn Karlsson, Tingting Ma, Yuzhong Qu, Chin-Yew Lin  
HIGHLIGHT: In this paper, we present a new KBQA model, TIARA, which addresses those issues by applying multi-grained retrieval to help the PLM focus on the most relevant KB context, viz., entities, exemplary logical forms, and schema items.
- 556, TITLE: Structure-Unified M-Tree Coding Solver for Math Word Problem  
<https://aclanthology.org/2022.emnlp-main.556>  
AUTHORS: Bin Wang, Jiangzhou Ju, Yang Fan, Xinyu Dai, Shujian Huang, Jiajun Chen  
HIGHLIGHT: In this paper, we propose the Structure-Unified M-Tree Coding Solver (SUMC-Solver), which applies a tree with any M branches (M-tree) to unify the output structures.
- 557, TITLE: FormLM: Recommending Creation Ideas for Online Forms by Modelling Semantic and Structural Information  
<https://aclanthology.org/2022.emnlp-main.557>  
AUTHORS: Yijia Shao, Mengyu Zhou, Yifan Zhong, Tao Wu, Hongwei Han, Shi Han, Gideon Huang, Dongmei Zhang  
HIGHLIGHT: To assist form designers, in this work we present FormLM to model online forms (by enhancing pre-trained language model with form structural information) and recommend form creation ideas (including question / options recommendations and block type suggestion).
- 558, TITLE: Generate, Discriminate and Contrast: A Semi-Supervised Sentence Representation Learning Framework  
<https://aclanthology.org/2022.emnlp-main.558>  
AUTHORS: Yiming Chen, Yan Zhang, Bin Wang, ZuoZhu Liu, Haizhou Li  
HIGHLIGHT: In this work, we propose a semi-supervised sentence embedding framework, GenSE, that effectively leverages large-scale unlabeled data.
- 559, TITLE: GPS: Genetic Prompt Search for Efficient Few-Shot Learning  
<https://aclanthology.org/2022.emnlp-main.559>  
AUTHORS: Hanwei Xu, Yujun Chen, Yulun Du, Nan Shao, Wang Yanggang, Haiyu Li, Zhilin Yang  
HIGHLIGHT: In this paper, we introduce Genetic Prompt Search (GPS) to improve few-shot learning with prompts, which utilizes a genetic algorithm to automatically search for the best prompt.
- 560, TITLE: Multitask Instruction-based Prompting for Fallacy Recognition  
<https://aclanthology.org/2022.emnlp-main.560>  
AUTHORS: Tariq Alhindi, Tuhin Chakraborty, Elena Musi, Smaranda Muresan  
HIGHLIGHT: Moreover, a big challenge for computational models lies in the fact that fallacies are formulated differently across the datasets with differences in the input format (e.g., question-answer pair, sentence with fallacy fragment), genre (e.g., social media, dialogue, news), as well as types and number of fallacies (from 5 to 18 types per dataset). To move towards solving the fallacy recognition task, we approach these differences across datasets as multiple tasks and show how instruction-based prompting in a multitask setup based on the T5 model improves the results against approaches built for a specific dataset such as T5, BERT or GPT-3.
- 561, TITLE: Rethinking Multi-Modal Alignment in Multi-Choice VideoQA from Feature and Sample Perspectives  
<https://aclanthology.org/2022.emnlp-main.561>  
AUTHORS: Shaoning Xiao, Long Chen, Kaifeng Gao, Zhao Wang, Yi Yang, Zhimeng Zhang, Jun Xiao  
HIGHLIGHT: In this paper, we reconsider the multi-modal alignment problem in VideoQA from feature and sample perspectives to achieve better performance.

- 562, TITLE: Towards Table-to-Text Generation with Pretrained Language Model: A Table Structure Understanding and Text Deliberating Approach  
<https://aclanthology.org/2022.emnlp-main.562>  
AUTHORS: Miao Chen, Xinjiang Lu, Tong Xu, Yanyan Li, Zhou Jingbo, Dejing Dou, Hui Xiong  
HIGHLIGHT: In this paper, to implement the table-to-text generation with pretrained language model, we propose a table structure understanding and text deliberating approach, namely T ASD.
- 563, TITLE: Hierarchical Phrase-Based Sequence-to-Sequence Learning  
<https://aclanthology.org/2022.emnlp-main.563>  
AUTHORS: Bailin Wang, Ivan Titov, Jacob Andreas, Yoon Kim  
HIGHLIGHT: This paper describes a neural transducer that maintains the flexibility of standard sequence-to-sequence (seq2seq) models while incorporating hierarchical phrases as a source of inductive bias during training and as explicit constraints during inference.
- 564, TITLE: Natural Language Deduction with Incomplete Information  
<https://aclanthology.org/2022.emnlp-main.564>  
AUTHORS: Zayne Sprague, Kaj Bostrom, Swarat Chaudhuri, Greg Durrett  
HIGHLIGHT: We propose a new system that can handle the underspecified setting where not all premises are stated at the outset; that is, additional assumptions need to be materialized to prove a claim.
- 565, TITLE: Character-centric Story Visualization via Visual Planning and Token Alignment  
<https://aclanthology.org/2022.emnlp-main.565>  
AUTHORS: Hong Chen, Rujun Han, Te-Lin Wu, Hideki Nakayama, Nanyun Peng  
HIGHLIGHT: To tackle the challenge, we propose to adapt a recent work that augments VQ-VAE with a text-to-visual-token (transformer) architecture.
- 566, TITLE: ASQA: Factoid Questions Meet Long-Form Answers  
<https://aclanthology.org/2022.emnlp-main.566>  
AUTHORS: Ivan Stelmakh, Yi Luan, Bhuwan Dhingra, Ming-Wei Chang  
HIGHLIGHT: The hurdles include a lack of high-quality data and the absence of a well-defined notion of an answer quality. In this work, we address these problems by releasing a novel dataset and a task that we call ASQA (Answer Summaries for Questions which are Ambiguous); and proposing a reliable metric for measuring performance on ASQA.
- 567, TITLE: Algorithms for Acyclic Weighted Finite-State Automata with Failure Arcs  
<https://aclanthology.org/2022.emnlp-main.567>  
AUTHORS: Anej Svete, Benjamin Dayan, Ryan Cotterell, Tim Vieira, Jason Eisner  
HIGHLIGHT: In this work, we present more efficient algorithms for computing the pathsum in sparse acyclic WFSA s, i. e., WFSA s with average out symbol fraction  $s \hat{A}_i \hat{A} \frac{1}{2}$ .
- 568, TITLE: Towards Better Document-level Relation Extraction via Iterative Inference  
<https://aclanthology.org/2022.emnlp-main.568>  
AUTHORS: Liang Zhang, Jinsong Su, Yidong Chen, Zhongjian Miao, Min Zijun, Qingguo Hu, Xiaodong Shi  
HIGHLIGHT: Existing methods usually directly predict the relations of all entity pairs of input document in a one-pass manner, ignoring the fact that predictions of some entity pairs heavily depend on the predicted results of other pairs. To deal with this issue, in this paper, we propose a novel document-level RE model with iterative inference.
- 569, TITLE: Efficient Adversarial Training with Robust Early-Bird Tickets  
<https://aclanthology.org/2022.emnlp-main.569>  
AUTHORS: Zhiheng Xi, Rui Zheng, Tao Gui, Qi Zhang, Xuanjing Huang  
HIGHLIGHT: Delving into the optimization process of adversarial training, we find that robust connectivity patterns emerge in the early training phase (typically 0).
- 570, TITLE: Quantifying Privacy Risks of Masked Language Models Using Membership Inference Attacks  
<https://aclanthology.org/2022.emnlp-main.570>  
AUTHORS: Fatemehsadat Mireshghallah, Kartik Goyal, Archit Uniyal, Taylor Berg-Kirkpatrick, Reza Shokri  
HIGHLIGHT: Prior attempts at measuring leakage of MLMs via membership inference attacks have been inconclusive, implying potential robustness of MLMs to privacy attacks. In this work, we posit that prior attempts were inconclusive because they based their attack solely on the MLM  $\hat{A}_i \hat{A} \frac{1}{2}$ s model score.
- 571, TITLE: SMaLL-100: Introducing Shallow Multilingual Machine Translation Model for Low-Resource Languages  
<https://aclanthology.org/2022.emnlp-main.571>  
AUTHORS: Alireza Mohammadshahi, Vassilina Nikoulina, Alexandre Berard, Caroline Brun, James Henderson, Laurent Besacier  
HIGHLIGHT: We introduce SMaLL-100, a distilled version of the M2M-100(12B) model, a massively multilingual machine translation model covering 100 languages.
- 572, TITLE: TextFusion: Privacy-Preserving Pre-trained Model Inference via Token Fusion  
<https://aclanthology.org/2022.emnlp-main.572>  
AUTHORS: Xin Zhou, Jinzhu Lu, Tao Gui, Ruotian Ma, Zichu Fei, Yuran Wang, Yong Ding, Yibo Cheung, Qi Zhang, Xuanjing Huang

**HIGHLIGHT:** However, recent studies have shown that intermediate representations can also be recovered to plain text with reasonable accuracy, thus the risk of privacy leakage still exists. To address this issue, we propose TextFusion, a novel method for preserving inference privacy.

573, **TITLE:** Learning to Explain Selectively: A Case Study on Question Answering

<https://aclanthology.org/2022.emnlp-main.573>

**AUTHORS:** Shi Feng, Jordan Boyd-Graber

**HIGHLIGHT:** We propose learning to explain selectively: for each decision that the user makes, we use a model to choose the best explanation from a set of candidates and update this model with feedback to optimize human performance.

574, **TITLE:** ConsistTL: Modeling Consistency in Transfer Learning for Low-Resource Neural Machine Translation

<https://aclanthology.org/2022.emnlp-main.574>

**AUTHORS:** Zhaocong Li, Xuebo Liu, Derek F. Wong, Lidia S. Chao, Min Zhang

**HIGHLIGHT:** In this paper, we propose a novel transfer learning method for NMT, namely ConsistTL, which can continuously transfer knowledge from the parent model during the training of the child model.

575, **TITLE:** Better Hit the Nail on the Head than Beat around the Bush: Removing Protected Attributes with a Single Projection

<https://aclanthology.org/2022.emnlp-main.575>

**AUTHORS:** Pantea Haghighatkhah, Antske Fokkens, Pia Sommerauer, Bettina Speckmann, Kevin Verbeek

**HIGHLIGHT:** We introduce two methods that find a single targeted projection: Mean Projection (MP, more efficient) and Tukey Median Projection (TMP, with theoretical guarantees).

576, **TITLE:** IELM: An Open Information Extraction Benchmark for Pre-Trained Language Models

<https://aclanthology.org/2022.emnlp-main.576>

**AUTHORS:** Chenguang Wang, Xiao Liu, Dawn Song

**HIGHLIGHT:** We introduce a new open information extraction (OIE) benchmark for pre-trained language models (LM).

577, **TITLE:** ConNER: Consistency Training for Cross-lingual Named Entity Recognition

<https://aclanthology.org/2022.emnlp-main.577>

**AUTHORS:** Ran Zhou, Xin Li, Lidong Bing, Erik Cambria, Luo Si, Chunyan Miao

**HIGHLIGHT:** We propose ConNER as a novel consistency training framework for cross-lingual NER, which comprises of: (1) translation-based consistency training on unlabeled target-language data, and (2) dropout-based consistency training on labeled source-language data.

578, **TITLE:** A Sequential Flow Control Framework for Multi-hop Knowledge Base Question Answering

<https://aclanthology.org/2022.emnlp-main.578>

**AUTHORS:** Minghui Xie, Chuzhan Hao, Peng Zhang

**HIGHLIGHT:** Existing methods, however, (i) infer the dynamic question representation only through coarse-grained attention mechanisms, which may bring information loss, (ii) and have not effectively modeled the sequential logic, which is crucial for the multi-hop reasoning process in KBQA. To address these issues, we propose a sequential reasoning self-attention mechanism to capture the crucial reasoning information of each single hop in a more fine-grained way.

579, **TITLE:** ACENet: Attention Guided Commonsense Reasoning on Hybrid Knowledge Graph

<https://aclanthology.org/2022.emnlp-main.579>

**AUTHORS:** Chuzhan Hao, Minghui Xie, Peng Zhang

**HIGHLIGHT:** In this paper, we propose an Attention guided Commonsense Reasoning Network (ACENet) to endow the neural network with the capability of integrating hybrid knowledge.

580, **TITLE:** Revisiting DocRED - Addressing the False Negative Problem in Relation Extraction

<https://aclanthology.org/2022.emnlp-main.580>

**AUTHORS:** Qingyu Tan, Lu Xu, Lidong Bing, Hwee Tou Ng, Sharifah Mahani Aljunied

**HIGHLIGHT:** However, we find that the annotation of DocRED is incomplete, i.e., false negative samples are prevalent. We analyze the causes and effects of the overwhelming false negative problem in the DocRED dataset.

581, **TITLE:** Towards Summary Candidates Fusion

<https://aclanthology.org/2022.emnlp-main.581>

**AUTHORS:** Mathieu Ravaut, Shafiq Joty, Nancy Chen

**HIGHLIGHT:** To bypass this limitation, we propose a new paradigm in second-stage abstractive summarization called SummaFusion that fuses several summary candidates to produce a novel abstractive second-stage summary.

582, **TITLE:** Multimodal Robustness for Neural Machine Translation

<https://aclanthology.org/2022.emnlp-main.582>

**AUTHORS:** Yuting Zhao, Ioan Calapodescu

**HIGHLIGHT:** In this paper, we look at the case of a Generic text-to-text NMT model that has to deal with data coming from various modalities, like speech, images, or noisy text extracted from the web.

583, **TITLE:** TranSHER: Translating Knowledge Graph Embedding with Hyper-Ellipsoidal Restriction

<https://aclanthology.org/2022.emnlp-main.583>

**AUTHORS:** Yizhi Li, Wei Fan, Chao Liu, Chenghua Lin, Jiang Qian

**HIGHLIGHT:** However, such a method strictly restricts entities on the hyper-ellipsoid surfaces which limits the optimization of entity distribution, leading to suboptimal performance of knowledge graph completion. To address this issue, we propose a novel score function TranSHER, which leverages relation-specific translations between head and tail entities to relax the constraint of hyper-ellipsoid restrictions.

584, **TITLE:** IRRGN: An Implicit Relational Reasoning Graph Network for Multi-turn Response Selection  
<https://aclanthology.org/2022.emnlp-main.584>

**AUTHORS:** Jingcheng Deng, Hengwei Dai, Xuewei Guo, Yuanchen Ju, Wei Peng

**HIGHLIGHT:** In addition, few studies consider differences between the options before and after reasoning. In this paper, we propose an Implicit Relational Reasoning Graph Network to address these issues, which consists of the Utterance Relational Reasoner (URR) and the Option Dual Comparator (ODC).

585, **TITLE:** Predicting Prerequisite Relations for Unseen Concepts

<https://aclanthology.org/2022.emnlp-main.585>

**AUTHORS:** Yaxin Zhu, Hamed Zamani

**HIGHLIGHT:** However, many real-world scenarios deal with concepts that are left undiscovered at training time, which is relatively unexplored. This paper studies this problem and proposes a novel alternating knowledge distillation approach to take advantage of both content- and graph-based models for this task.

586, **TITLE:** Contrastive Learning with Expectation-Maximization for Weakly Supervised Phrase Grounding

<https://aclanthology.org/2022.emnlp-main.586>

**AUTHORS:** Keqin Chen, Richong Zhang, Samuel Mensah, Yongyi Mao

**HIGHLIGHT:** Specifically, we propose a novel contrastive learning framework based on the expectation-maximization algorithm that adaptively refines the target prediction.

587, **TITLE:** Beyond prompting: Making Pre-trained Language Models Better Zero-shot Learners by Clustering Representations

<https://aclanthology.org/2022.emnlp-main.587>

**AUTHORS:** Yu Fei, Zhao Meng, Ping Nie, Roger Wattenhofer, Mrinmaya Sachan

**HIGHLIGHT:** In this work, we show that zero-shot text classification can be improved simply by clustering texts in the embedding spaces of PLMs.

588, **TITLE:** Generalizing over Long Tail Concepts for Medical Term Normalization

<https://aclanthology.org/2022.emnlp-main.588>

**AUTHORS:** Beatrice Portelli, Simone Scabro, Enrico Santus, Hooman Sedghamiz, Emmanuele Chersoni, Giuseppe Serra

**HIGHLIGHT:** In this paper we introduce a simple and effective learning strategy that leverages such information to enhance the generalizability of both discriminative and generative models.

589, **TITLE:** Unsupervised Opinion Summarisation in the Wasserstein Space

<https://aclanthology.org/2022.emnlp-main.589>

**AUTHORS:** Jiayu Song, Iman Munire Bilal, Adam Tsakalidis, Rob Procter, Maria Liakata

**HIGHLIGHT:** Such posts are noisy and have unpredictable structure, posing additional challenges for the construction of the summary distribution and the preservation of meaning compared to online reviews, which has been so far the focus on opinion summarisation. To address these challenges we present WassOS, an unsupervised abstractive summarization model which makes use of the Wasserstein distance.

590, **TITLE:** Bloom Library: Multimodal Datasets in 300+ Languages for a Variety of Downstream Tasks

<https://aclanthology.org/2022.emnlp-main.590>

**AUTHORS:** Colin Leong, Joshua Nemecek, Jacob Mansdorfer, Anna Filighera, Abraham Owodunni, Daniel Whitenack

**HIGHLIGHT:** We present Bloom Library, a linguistically diverse set of multimodal and multilingual datasets for language modeling, image captioning, visual storytelling, and speech synthesis/recognition.

591, **TITLE:** Disentangling Uncertainty in Machine Translation Evaluation

<https://aclanthology.org/2022.emnlp-main.591>

**AUTHORS:** Chrysoula Zerva, Taisiya Glushkova, Ricardo Rei, André F. T. Martins

**HIGHLIGHT:** In this paper, we propose more powerful and efficient uncertainty predictors for MT evaluation, and we assess their ability to target different sources of aleatoric and epistemic uncertainty.

592, **TITLE:** Does Your Model Classify Entities Reasonably? Diagnosing and Mitigating Spurious Correlations in Entity Typing

<https://aclanthology.org/2022.emnlp-main.592>

**AUTHORS:** Nan Xu, Fei Wang, Bangzheng Li, Mingtao Dong, Muhao Chen

**HIGHLIGHT:** To comprehensively investigate the faithfulness and reliability of entity typing methods, we first systematically define distinct kinds of model biases that are reflected mainly from spurious correlations. Particularly, we identify six types of existing model biases, including mention-context bias, lexical overlapping bias, named entity bias, pronoun bias, dependency bias, and overgeneralization bias. To mitigate model biases, we then introduce a counterfactual data augmentation method.

593, **TITLE:** EDIN: An End-to-end Benchmark and Pipeline for Unknown Entity Discovery and Indexing

<https://aclanthology.org/2022.emnlp-main.593>

**AUTHORS:** Nora Kassner, Fabio Petroni, Mikhail Plekhanov, Sebastian Riedel, Nicola Cancedda



**HIGHLIGHT:** Building on dense-retrieval based entity linking, we introduce the end-to-end EDIN-pipeline that detects, clusters, and indexes mentions of unknown entities in context.

594, **TITLE:** POQue: Asking Participant-specific Outcome Questions for a Deeper Understanding of Complex Events  
<https://aclanthology.org/2022.emnlp-main.594>

**AUTHORS:** Sai Vallurupalli, Sayontan Ghosh, Katrin Erk, Niranjana Balasubramanian, Francis Ferraro

**HIGHLIGHT:** We show that by pre-identifying a participant in a complex event, crowdworkers are able to (1) infer the collective impact of salient events that make up the situation, (2) annotate the volitional engagement of participants in causing the situation, and (3) ground the outcome of the situation in state changes of the participants.

595, **TITLE:** Measuring the Mixing of Contextual Information in the Transformer  
<https://aclanthology.org/2022.emnlp-main.595>

**AUTHORS:** Javier Ferrando, Gerard I. Gállego, Marta R. Costa-jussà

**HIGHLIGHT:** In this paper, we consider the whole attention block multi-head attention, residual connection, and layer normalization and define a metric to measure token-to-token interactions within each layer.

596, **TITLE:** Dealing with Abbreviations in the Slovenian Biographical Lexicon

<https://aclanthology.org/2022.emnlp-main.596>

**AUTHORS:** Angel Daza, Antske Fokkens, Tomaž Erjavec

**HIGHLIGHT:** In this paper, we propose a new method for addressing the problems caused by a high density of domain-specific abbreviations in a text.

597, **TITLE:** AfriCLIRMatrix: Enabling Cross-Lingual Information Retrieval for African Languages

<https://aclanthology.org/2022.emnlp-main.597>

**AUTHORS:** Odunayo Ogundepo, Xinyu Zhang, Shuo Sun, Kevin Duh, Jimmy Lin

**HIGHLIGHT:** For search, most existing datasets feature few or no African languages, directly impacting researchers' ability to build and improve information access capabilities in those languages. Motivated by this, we created AfriCLIRMatrix, a test collection for cross-lingual information retrieval research in 15 diverse African languages.

598, **TITLE:** CONDAQA: A Contrastive Reading Comprehension Dataset for Reasoning about Negation

<https://aclanthology.org/2022.emnlp-main.598>

**AUTHORS:** Abhilasha Ravichander, Matt Gardner, Ana Marasovic

**HIGHLIGHT:** To facilitate the future development of models that can process negation effectively, we present CONDAQA, the first English reading comprehension dataset which requires reasoning about the implications of negated statements in paragraphs.

599, **TITLE:** Towards Opening the Black Box of Neural Machine Translation: Source and Target Interpretations of the Transformer

<https://aclanthology.org/2022.emnlp-main.599>

**AUTHORS:** Javier Ferrando, Gerard I. Gállego, Belen Alastruey, Carlos Escolano, Marta R. Costa-jussà

**HIGHLIGHT:** In this work, we propose an interpretability method that tracks input tokens' attributions for both contexts.

600, **TITLE:** ArtELingo: A Million Emotion Annotations of WikiArt with Emphasis on Diversity over Language and Culture

<https://aclanthology.org/2022.emnlp-main.600>

**AUTHORS:** Youssef Mohamed, Mohamed Abdelfattah, Shyma Alhuwaidar, Feifan Li, Xiangliang Zhang, Kenneth Church, Mohamed Elhoseiny

**HIGHLIGHT:** This paper introduces ArtELingo, a new benchmark and dataset, designed to encourage work on diversity across languages and cultures.

601, **TITLE:** Decoding a Neural Retriever's Latent Space for Query Suggestion

<https://aclanthology.org/2022.emnlp-main.601>

**AUTHORS:** Leonard Adolphs, Michelle Chen Huebscher, Christian Buck, Sertan Girgin, Olivier Bachem, Massimiliano Ciaramita, Thomas Hofmann

**HIGHLIGHT:** However, neural systems lack the interpretability of bag-of-words models; it is not trivial to connect a query change to a change in the latent space that ultimately determines the retrieval results. To shed light on this embedding space, we learn a query decoder that, given a latent representation of a neural search engine, generates the corresponding query.

602, **TITLE:** T-STAR: Truthful Style Transfer using AMR Graph as Intermediate Representation

<https://aclanthology.org/2022.emnlp-main.602>

**AUTHORS:** Anubhav Jangra, Preksha Nema, Aravindan Raghuvier

**HIGHLIGHT:** In this work, we study the usefulness of Abstract Meaning Representation (AMR) graph as the intermediate style agnostic representation.

603, **TITLE:** PromptBERT: Improving BERT Sentence Embeddings with Prompts

<https://aclanthology.org/2022.emnlp-main.603>

**AUTHORS:** Ting Jiang, Jian Jiao, Shaohan Huang, Zihan Zhang, Deqing Wang, Fuzhen Zhuang, Furu Wei, Haizhen Huang, Denvy Deng, Qi Zhang

**HIGHLIGHT:** We propose PromptBERT, a novel contrastive learning method for learning better sentence representation.

604, **TITLE:** Extending Logic Explained Networks to Text Classification

- <https://aclanthology.org/2022.emnlp-main.604>  
AUTHORS: Rishabh Jain, Gabriele Ciravegna, Pietro Barbiero, Francesco Giannini, Davide Buffelli, Pietro Lio  
HIGHLIGHT: However, these models have only been applied to vision and tabular data, and they mostly favour the generation of global explanations, while local ones tend to be noisy and verbose. For these reasons, we propose LEN<sup>p</sup>, improving local explanations by perturbing input words, and we test it on text classification.
- 605, TITLE: Uni-Parser: Unified Semantic Parser for Question Answering on Knowledge Base and Database  
<https://aclanthology.org/2022.emnlp-main.605>  
AUTHORS: Ye Liu, Semih Yavuz, Rui Meng, Dragomir Radev, Caiming Xiong, Yingbo Zhou  
HIGHLIGHT: In this work, we propose Uni-Parser, a unified semantic parser for question answering (QA) on both KB and DB.
- 606, TITLE: RAPO: An Adaptive Ranking Paradigm for Bilingual Lexicon Induction  
<https://aclanthology.org/2022.emnlp-main.606>  
AUTHORS: Zhoujin Tian, Chaozhuo Li, Shuo Ren, Zhiqiang Zuo, Zengxuan Wen, Xinyue Hu, Xiao Han, Haizhen Huang, Denvy Deng, Qi Zhang, Xing Xie  
HIGHLIGHT: In this work, we propose a novel ranking-oriented induction model RAPO to learn personalized mapping function for each word.
- 607, TITLE: On Parsing as Tagging  
<https://aclanthology.org/2022.emnlp-main.607>  
AUTHORS: Afra Amini, Ryan Cotterell  
HIGHLIGHT: There are many proposals to reduce constituency parsing to tagging. To figure out what these approaches have in common, we offer a unifying pipeline, which consists of three steps: linearization, learning, and decoding.
- 608, TITLE: Distilled Dual-Encoder Model for Vision-Language Understanding  
<https://aclanthology.org/2022.emnlp-main.608>  
AUTHORS: Zekun Wang, Wenhui Wang, Haichao Zhu, Ming Liu, Bing Qin, Furu Wei  
HIGHLIGHT: To get the best of both worlds, we propose DiDE, a framework that distills the knowledge of the fusion-encoder teacher model into the dual-encoder student model.
- 609, TITLE: Argument Mining for Review Helpfulness Prediction  
<https://aclanthology.org/2022.emnlp-main.609>  
AUTHORS: Zaiqian Chen, Daniel Verdi do Amarante, Jenna Donaldson, Yohan Jo, Joonsuk Park  
HIGHLIGHT: To this end, we present the AMazon Argument Mining (AM2) corpus<sup>1</sup> a corpus of 878 Amazon reviews on headphones annotated according to a theoretical argumentation model designed to evaluate argument quality.
- 610, TITLE: Hierarchical Multi-Label Classification of Scientific Documents  
<https://aclanthology.org/2022.emnlp-main.610>  
AUTHORS: Mobashir Sadat, Cornelia Caragea  
HIGHLIGHT: In this paper, we introduce a new dataset for hierarchical multi-label text classification (HMLTC) of scientific papers called SciHTC, which contains 186,160 papers and 1,234 categories from the ACM CCS tree.
- 611, TITLE: Rainier: Reinforced Knowledge Introspector for Commonsense Question Answering  
<https://aclanthology.org/2022.emnlp-main.611>  
AUTHORS: Jiacheng Liu, Skyler Hallinan, Ximing Lu, Pengfei He, Sean Welleck, Hannaneh Hajishirzi, Yejin Choi  
HIGHLIGHT: We present Rainier, or Reinforced Knowledge Introspector, that learns to generate contextually relevant knowledge in response to given questions.
- 612, TITLE: A Major Obstacle for NLP Research: Let's Talk about Time Allocation!  
<https://aclanthology.org/2022.emnlp-main.612>  
AUTHORS: Katharina Kann, Shiran Dudy, Arya D. McCarthy  
HIGHLIGHT: However, this paper argues that we have been less successful than we *should* have been and reflects on where and how the field fails to tap its full potential.
- 613, TITLE: Towards Inter-character Relationship-driven Story Generation  
<https://aclanthology.org/2022.emnlp-main.613>  
AUTHORS: Anvesh Rao Vijjini, Faeze Brahman, Snigdha Chaturvedi  
HIGHLIGHT: In this paper, we introduce the task of modeling interpersonal relationships for story generation.
- 614, TITLE: Incorporating Relevance Feedback for Information-Seeking Retrieval using Few-Shot Document Re-Ranking  
<https://aclanthology.org/2022.emnlp-main.614>  
AUTHORS: Tim Baumgartner, Leonardo F. R. Ribeiro, Nils Reimers, Iryna Gurevych  
HIGHLIGHT: Specifically, we introduce a kNN approach that re-ranks documents based on their similarity with the query and the documents the user considers relevant.
- 615, TITLE: ReasTAP: Injecting Table Reasoning Skills During Pre-training via Synthetic Reasoning Examples  
<https://aclanthology.org/2022.emnlp-main.615>  
AUTHORS: Yilun Zhao, Linyong Nan, Zhenting Qi, Rui Zhang, Dragomir Radev

**HIGHLIGHT:** In this work, we develop ReasTAP to show that high-level table reasoning skills can be injected into models during pre-training without a complex table-specific architecture design.

616, **TITLE:** Few-shot Learning with Multilingual Generative Language Models  
<https://aclanthology.org/2022.emnlp-main.616>

**AUTHORS:** Xi Victoria Lin, Todor Mihaylov, Mikel Artetxe, Tianlu Wang, Shuohui Chen, Daniel Simig, Myle Ott, Naman Goyal, Shruti Bhosale, Jingfei Du, Ramakanth Pasunuru, Sam Shleifer, Punit Singh Koura, Vishrav Chaudhary, Brian O’Halloran, Jeff Wang, Luke Zettlemoyer, Zornitsa Kozareva, Mona Diab, Veselin Stoyanov, Xian Li

**HIGHLIGHT:** In this work, we train multilingual generative language models on a corpus covering a diverse set of languages, and study their few- and zero-shot learning capabilities in a wide range of tasks.

617, **TITLE:** Are representations built from the ground up? An empirical examination of local composition in language models

<https://aclanthology.org/2022.emnlp-main.617>

**AUTHORS:** Emmy Liu, Graham Neubig

**HIGHLIGHT:** At the same time, many phrases are non-compositional, carrying a meaning beyond that of each part in isolation. Representing both of these types of phrases is critical for language understanding, but it is an open question whether modern language models (LMs) learn to do so; in this work we examine this question.

618, **TITLE:** Detecting Label Errors by Using Pre-Trained Language Models

<https://aclanthology.org/2022.emnlp-main.618>

**AUTHORS:** Derek Chong, Jenny Hong, Christopher Manning

**HIGHLIGHT:** To this end, we contribute a novel method for introducing realistic, human-originated label noise into existing crowdsourced datasets such as SNLI and TweetNLP.

619, **TITLE:** Intriguing Properties of Compression on Multilingual Models

<https://aclanthology.org/2022.emnlp-main.619>

**AUTHORS:** Kelechi Ogueji, Orevaoghene Ahia, Gbemileke Onilude, Sebastian Gehrmann, Sara Hooker, Julia Kreutzer

**HIGHLIGHT:** In this work, we propose an experimental framework to characterize the impact of sparsifying multilingual pre-trained language models during fine-tuning.

620, **TITLE:** Sequence Models for Document Structure Identification in an Undeciphered Script

<https://aclanthology.org/2022.emnlp-main.620>

**AUTHORS:** Logan Born, M. Monroe, Kathryn Kelley, Anoop Sarkar

**HIGHLIGHT:** This work describes the first thorough analysis of  $\bar{\text{A}}\bar{\text{I}}\bar{\text{A}}\bar{\text{I}}$  signs in proto-Elamite, an undeciphered script from 3100-2900 BCE.

621, **TITLE:** English Contrastive Learning Can Learn Universal Cross-lingual Sentence Embeddings

<https://aclanthology.org/2022.emnlp-main.621>

**AUTHORS:** Yaoshian Wang, Ashley Wu, Graham Neubig

**HIGHLIGHT:** In this work, we propose mSimCSE, which extends SimCSE to multilingual settings and reveal that contrastive learning on English data can surprisingly learn high-quality universal cross-lingual sentence embeddings without any parallel data.

622, **TITLE:** Active Example Selection for In-Context Learning

<https://aclanthology.org/2022.emnlp-main.622>

**AUTHORS:** Yiming Zhang, Shi Feng, Chenhao Tan

**HIGHLIGHT:** We formulate example selection for in-context learning as a sequential decision problem, and propose a reinforcement learning algorithm for identifying generalizable policies to select demonstration examples.

623, **TITLE:** Improving Factual Consistency in Summarization with Compression-Based Post-Editing

<https://aclanthology.org/2022.emnlp-main.623>

**AUTHORS:** Alex Fabbri, Prafulla Kumar Choubey, Jesse Vig, Chien-Sheng Wu, Caiming Xiong

**HIGHLIGHT:** We propose to use sentence-compression data to train the post-editing model to take a summary with extrinsic entity errors marked with special tokens and output a compressed, well-formed summary with those errors removed.

624, **TITLE:** Evaluating the Impact of Model Scale for Compositional Generalization in Semantic Parsing

<https://aclanthology.org/2022.emnlp-main.624>

**AUTHORS:** Linlu Qiu, Peter Shaw, Panupong Pasupat, Tianze Shi, Jonathan Herzig, Emily Pitler, Fei Sha, Kristina Toutanova

**HIGHLIGHT:** We evaluate encoder-decoder models up to 11B parameters and decoder-only models up to 540B parameters, and compare model scaling curves for three different methods for applying a pre-trained language model to a new task: fine-tuning all parameters, prompt tuning, and in-context learning.

625, **TITLE:**  $\bar{\text{A}}\bar{\text{I}}\bar{\text{A}}\bar{\text{I}}$  sorry to hear that  $\bar{\text{A}}\bar{\text{I}}\bar{\text{A}}\bar{\text{I}}$ : Finding New Biases in Language Models with a Holistic Descriptor Dataset

<https://aclanthology.org/2022.emnlp-main.625>

**AUTHORS:** Eric Michael Smith, Melissa Hall, Melanie Kambadur, Eleonora Presani, Adina Williams

**HIGHLIGHT:** In this work, we present a new, more inclusive bias measurement dataset, HolisticBias, which includes nearly 600 descriptor terms across 13 different demographic axes.

- 626, TITLE: Understanding ME? Multimodal Evaluation for Fine-grained Visual Commonsense  
<https://aclanthology.org/2022.emnlp-main.626>  
AUTHORS: Zhecan Wang, Haoxuan You, Yicheng He, Wenhao Li, Kai-Wei Chang, Shih-Fu Chang  
HIGHLIGHT: To provide an in-depth analysis, we present a Multimodal Evaluation (ME) pipeline to automatically generate question-answer pairs to test models' understanding of the visual scene, text, and related knowledge.
- 627, TITLE: Semantic Novelty Detection and Characterization in Factual Text Involving Named Entities  
<https://aclanthology.org/2022.emnlp-main.627>  
AUTHORS: Nianzu Ma, Sahisnu Mazumder, Alexander Poltowicz, Bing Liu, Eric Robertson, Scott Grigsby  
HIGHLIGHT: This paper proposes an effective model (called PAT-SND) to solve the problem, which can also characterize the novelty.
- 628, TITLE: CN-AutoMIC: Distilling Chinese Commonsense Knowledge from Pretrained Language Models  
<https://aclanthology.org/2022.emnlp-main.628>  
AUTHORS: Chenhao Wang, Jiachun Li, Yubo Chen, Kang Liu, Jun Zhao  
HIGHLIGHT: In this paper, we propose a large-scale Chinese CKG generated from multilingual PLMs, named as **CN-AutoMIC**, aiming to fill the research gap of non-English CKGs.
- 629, TITLE: Calibrating Student Models for Emotion-related Tasks  
<https://aclanthology.org/2022.emnlp-main.629>  
AUTHORS: Mahshid Hosseini, Cornelia Caragea  
HIGHLIGHT: In this paper, we study KD on the emotion-related tasks from a new perspective: calibration.
- 630, TITLE: Overcoming Catastrophic Forgetting in Zero-Shot Cross-Lingual Generation  
<https://aclanthology.org/2022.emnlp-main.630>  
AUTHORS: Tu Vu, Aditya Barua, Brian Lester, Daniel Cer, Mohit Iyyer, Noah Constant  
HIGHLIGHT: In this paper, we explore the challenging problem of performing a generative task in a target language when labeled data is only available in English, using summarization as a case study.
- 631, TITLE: Improving Large-scale Paraphrase Acquisition and Generation  
<https://aclanthology.org/2022.emnlp-main.631>  
AUTHORS: Yao Dou, Chao Jiang, Wei Xu  
HIGHLIGHT: We present a new Multi-Topic Paraphrase in Twitter (MultiPIT) corpus that consists of a total of 130k sentence pairs with crowdsourcing (MultiPIT\_crowd) and expert (MultiPIT\_expert) annotations using two different paraphrase definitions for paraphrase identification, in addition to a multi-reference test set (MultiPIT\_NMR) and a large automatically constructed training set (MultiPIT\_Auto) for paraphrase generation.
- 632, TITLE: Entropy- and Distance-Based Predictors From GPT-2 Attention Patterns Predict Reading Times Over and Above GPT-2 Surprisal  
<https://aclanthology.org/2022.emnlp-main.632>  
AUTHORS: Byung-Doh Oh, William Schuler  
HIGHLIGHT: Under this framework, this work first defines an entropy-based predictor that quantifies the diffuseness of self-attention, as well as distance-based predictors that capture the incremental change in attention patterns across timesteps.
- 633, TITLE: A Survey of Computational Framing Analysis Approaches  
<https://aclanthology.org/2022.emnlp-main.633>  
AUTHORS: Mohammad Ali, Naeemul Hassan  
HIGHLIGHT: The growing scholarship, however, lacks a comprehensive understanding and resources of computational framing analysis methods. Aiming to address the gap, this article surveys existing computational framing analysis approaches and puts them together.
- 634, TITLE: Learning Cross-Task Dependencies for Joint Extraction of Entities, Events, Event Arguments, and Relations  
<https://aclanthology.org/2022.emnlp-main.634>  
AUTHORS: Minh Van Nguyen, Bonan Min, Franck Dernoncourt, Thien Nguyen  
HIGHLIGHT: However, the cross-task dependencies in prior work are not optimal as they are only designed manually according to some task heuristics. To address this issue, we propose a novel model for JointIE that aims to learn cross-task dependencies from data.
- 635, TITLE: Don't Copy the Teacher: Data and Model Challenges in Embodied Dialogue  
<https://aclanthology.org/2022.emnlp-main.635>  
AUTHORS: So Yeon Min, Hao Zhu, Ruslan Salakhutdinov, Yonatan Bisk  
HIGHLIGHT: This paper contributes to that conversation, by arguing that imitation learning (IL) and related low-level metrics are actually misleading and do not align with the goals of embodied dialogue research and may hinder progress.
- 636, TITLE: ALFRED-L: Investigating the Role of Language for Action Learning in Interactive Visual Environments  
<https://aclanthology.org/2022.emnlp-main.636>  
AUTHORS: Arjun Akula, Spandana Gella, Aishwarya Padmakumar, Mahdi Namazifar, Mohit Bansal, Jesse Thomason, Dilek Hakkani-Tur  
HIGHLIGHT: In this work, we examine ALFRED, a challenging benchmark for embodied task completion, with the goal of gaining insight into how effectively models utilize language.

- 637, TITLE: Dungeons and Dragons as a Dialog Challenge for Artificial Intelligence  
<https://aclanthology.org/2022.emnlp-main.637>  
AUTHORS: Chris Callison-Burch, Gaurav Singh Tomar, Lara Martin, Daphne Ippolito, Suma Bailis, David Reitter  
HIGHLIGHT: In this paper, we frame D&D specifically as a dialogue system challenge, where the tasks are to both generate the next conversational turn in the game and predict the state of the game given the dialogue history.
- 638, TITLE: Unsupervised Entity Linking with Guided Summarization and Multiple-Choice Selection  
<https://aclanthology.org/2022.emnlp-main.638>  
AUTHORS: Young Min Cho, Li Zhang, Chris Callison-Burch  
HIGHLIGHT: We address two challenge in entity linking: how to leverage wider contexts surrounding a mention, and how to deal with limited training data.
- 639, TITLE: Weakly-Supervised Temporal Article Grounding  
<https://aclanthology.org/2022.emnlp-main.639>  
AUTHORS: Long Chen, Yulei Niu, Brian Chen, Xudong Lin, Guangxing Han, Christopher Thomas, Hammad Ayyubi, Heng Ji, Shih-Fu Chang  
HIGHLIGHT: To this end, we propose a new challenging grounding task: Weakly-Supervised temporal Article Grounding (WSAG). Specifically, given an article and a relevant video, WSAG aims to localize all  $\hat{A} \hat{A}_i \hat{A}_i \hat{A}_i$  sentences to the video, and these sentences are possibly at different semantic scales.
- 640, TITLE: Exploring Dual Encoder Architectures for Question Answering  
<https://aclanthology.org/2022.emnlp-main.640>  
AUTHORS: Zhe Dong, Jianmo Ni, Dan Bikel, Enrique Alfonseca, Yuan Wang, Chen Qu, Imed Zitouni  
HIGHLIGHT: In this work, we explore the dual encoder architectures for QA retrieval tasks.
- 641, TITLE: arXivEdits: Understanding the Human Revision Process in Scientific Writing  
<https://aclanthology.org/2022.emnlp-main.641>  
AUTHORS: Chao Jiang, Wei Xu, Samuel Stevens  
HIGHLIGHT: In this work, we provide a complete computational framework for studying text revision in scientific writing.
- 642, TITLE: Why Do You Feel This Way? Summarizing Triggers of Emotions in Social Media Posts  
<https://aclanthology.org/2022.emnlp-main.642>  
AUTHORS: Hongli Zhan, Tiberiu Sosea, Cornelia Caragea, Junyi Jessy Li  
HIGHLIGHT: This paper takes a novel angle, namely, emotion detection and trigger summarization, aiming to both detect perceived emotions in text, and summarize events and their appraisals that trigger each emotion. To support this goal, we introduce CovidET (Emotions and their Triggers during Covid-19), a dataset of ~1,900 English Reddit posts related to COVID-19, which contains manual annotations of perceived emotions and abstractive summaries of their triggers described in the post.
- 643, TITLE: Analogical Math Word Problems Solving with Enhanced Problem-Solution Association  
<https://aclanthology.org/2022.emnlp-main.643>  
AUTHORS: Zhenwen Liang, Jipeng Zhang, Xiangliang Zhang  
HIGHLIGHT: In this paper, we propose to build a novel MWP solver by leveraging analogical MWPs, which advance the solver  $\hat{A} \hat{A}_i \hat{A}_i$ s generalization ability across different kinds of MWPs.
- 644, TITLE: Towards Teachable Reasoning Systems: Using a Dynamic Memory of User Feedback for Continual System Improvement  
<https://aclanthology.org/2022.emnlp-main.644>  
AUTHORS: Bhavana Dalvi Mishra, Oyvind Tafjord, Peter Clark  
HIGHLIGHT: Our goal is a teachable reasoning system for question-answering (QA), where a user can interact with faithful answer explanations, and correct its errors so that the system improves over time.
- 645, TITLE: Knowledge Transfer from Answer Ranking to Answer Generation  
<https://aclanthology.org/2022.emnlp-main.645>  
AUTHORS: Matteo Gabburo, Rik Koncel-Kedziorski, Siddhant Garg, Luca Soldaini, Alessandro Moschitti  
HIGHLIGHT: In this paper, we propose to train a GenQA model by transferring knowledge from a trained AS2 model, to overcome the aforementioned issue.
- 646, TITLE: Perturbation Augmentation for Fairer NLP  
<https://aclanthology.org/2022.emnlp-main.646>  
AUTHORS: Rebecca Qian, Candace Ross, Jude Fernandes, Eric Michael Smith, Douwe Kiela, Adina Williams  
HIGHLIGHT: In this work, we ask whether training on demographically perturbed data leads to fairer language models.
- 647, TITLE: Automatic Document Selection for Efficient Encoder Pretraining  
<https://aclanthology.org/2022.emnlp-main.647>  
AUTHORS: Yukun Feng, Patrick Xia, Benjamin Van Durme, Jo  $\hat{A} \hat{A}_i \hat{A}_i$  Sedoc  
HIGHLIGHT: We propose an alternative to larger training sets by automatically identifying smaller yet domain-representative subsets.
- 648, TITLE: The Aligned Multimodal Movie Treebank: An audio, video, dependency-parse treebank

<https://aclanthology.org/2022.emnlp-main.648>

**AUTHORS:** Adam Yaari, Jan DeWitt, Henry Hu, Bennett Stankovits, Sue Felshin, Yevgeni Berzak, Helena Aparicio, Boris Katz, Ignacio Cases, Andrei Barbu  
**HIGHLIGHT:** We introduce the Aligned Multimodal Movie Treebank (AMMT), an English language treebank derived from dialog in Hollywood movies which includes transcriptions of the audio-visual streams with word-level alignment, as well as part of speech tags and dependency parses in the Universal Dependencies formalism.

649, **TITLE:** DEMETR: Diagnosing Evaluation Metrics for Translation

<https://aclanthology.org/2022.emnlp-main.649>

**AUTHORS:** Marzena Karpinska, Nishant Raj, Katherine Thai, Yixiao Song, Ankita Gupta, Mohit Iyyer  
**HIGHLIGHT:** The operations of newer learned metrics (e. g. , BLEURT, COMET), which leverage pretrained language models to achieve higher correlations with human quality judgments than BLEU, are opaque in comparison. In this paper, we shed light on the behavior of these learned metrics by creating DEMETR, a diagnostic dataset with 31K English examples (translated from 10 source languages) for evaluating the sensitivity of MT evaluation metrics to 35 different linguistic perturbations spanning semantic, syntactic, and morphological error categories.

650, **TITLE:** Empowering Language Models with Knowledge Graph Reasoning for Open-Domain Question Answering

<https://aclanthology.org/2022.emnlp-main.650>

**AUTHORS:** Ziniu Hu, Yichong Xu, Wenhao Yu, Shuohang Wang, Ziyi Yang, Chenguang Zhu, Kai-Wei Chang, Yizhou Sun

**HIGHLIGHT:** In this work, we propose knOwledge REAsOning empowered Language Model(OREO-LM), which consists of a novel Knowledge Interaction Layer that can be flexibly plugged into existing Transformer-based LMs to interact with a differentiable Knowledge Graph Reasoning module collaboratively.

651, **TITLE:** Debiasing Pretrained Text Encoders by Paying Attention to Paying Attention

<https://aclanthology.org/2022.emnlp-main.651>

**AUTHORS:** Yacine Gaci, Boualem Benatallah, Fabio Casati, Khalid Benabdeslem

**HIGHLIGHT:** In this paper, we propose a debiasing method for pre-trained text encoders that both reduces social stereotypes, and inflicts next to no semantic damage.

652, **TITLE:** MEE: A Novel Multilingual Event Extraction Dataset

<https://aclanthology.org/2022.emnlp-main.652>

**AUTHORS:** Amir Pouran Ben Veysch, Javid Ebrahimi, Franck Demoncourt, Thien Nguyen

**HIGHLIGHT:** However, one limitation of current research for EE involves the under-exploration for non-English languages in which the lack of high-quality multilingual EE datasets for model training and evaluation has been the main hindrance. To address this limitation, we propose a novel Multilingual Event Extraction dataset (MEE) that provides annotation for more than 50K event mentions in 8 typologically different languages.

653, **TITLE:** RobustLR: A Diagnostic Benchmark for Evaluating Logical Robustness of Deductive Reasoners

<https://aclanthology.org/2022.emnlp-main.653>

**AUTHORS:** Soumya Sanyal, Zeyi Liao, Xiang Ren

**HIGHLIGHT:** To this end, we present RobustLR, a diagnostic benchmark that evaluates the robustness of language models to minimal logical edits in the inputs and different logical equivalence conditions.

654, **TITLE:** Evaluating and Improving Factuality in Multimodal Abstractive Summarization

<https://aclanthology.org/2022.emnlp-main.654>

**AUTHORS:** David Wan, Mohit Bansal

**HIGHLIGHT:** We propose CLIPBERTSCORE, a simple weighted combination of CLIPScore and BERTScore to leverage the robustness and strong factuality detection performance between image-summary and document-summary, respectively.

655, **TITLE:** Referee: Reference-Free Sentence Summarization with Sharper Controllability through Symbolic Knowledge Distillation

<https://aclanthology.org/2022.emnlp-main.655>

**AUTHORS:** Melanie Sclar, Peter West, Sachin Kumar, Yulia Tsvetkov, Yejin Choi

**HIGHLIGHT:** We present Referee, a novel framework for sentence summarization that can be trained reference-free (i. e. , requiring no gold summaries for supervision), while allowing direct control for compression ratio.

656, **TITLE:** Algorithms for Weighted Pushdown Automata

<https://aclanthology.org/2022.emnlp-main.656>

**AUTHORS:** Alexandra Butoi, Brian DuSell, Tim Vieira, Ryan Cotterell, David Chiang

**HIGHLIGHT:** In this paper, we develop novel algorithms that operate directly on WPDAs.

657, **TITLE:** MABEL: Attenuating Gender Bias using Textual Entailment Data

<https://aclanthology.org/2022.emnlp-main.657>

**AUTHORS:** Jacqueline He, Mengzhou Xia, Christiane Fellbaum, Danqi Chen

**HIGHLIGHT:** To this end, we propose MABEL (a Method for Attenuating Gender Bias using Entailment Labels), an intermediate pre-training approach for mitigating gender bias in contextualized representations.

658, **TITLE:** Breakpoint Transformers for Modeling and Tracking Intermediate Beliefs

<https://aclanthology.org/2022.emnlp-main.658>

- AUTHORS: Kyle Richardson, Ronen Tamari, Oren Sultan, Dafna Shahaf, Reut Tsarfaty, Ashish Sabharwal  
HIGHLIGHT: To this end, we propose a representation learning framework called breakpoint modeling that allows for efficient and robust learning of this type.
- 659, TITLE: Late Fusion with Triplet Margin Objective for Multimodal Ideology Prediction and Analysis  
<https://aclanthology.org/2022.emnlp-main.659>  
AUTHORS: Changyuan Qiu, Winston Wu, Xinliang Frederick Zhang, Lu Wang  
HIGHLIGHT: In this work, we introduce the task of multimodal ideology prediction, where a model predicts binary or five-point scale ideological leanings, given a text-image pair with political content.
- 660, TITLE: Leveraging QA Datasets to Improve Generative Data Augmentation  
<https://aclanthology.org/2022.emnlp-main.660>  
AUTHORS: Dheeraj Mekala, Tu Vu, Timo Schick, Jingbo Shang  
HIGHLIGHT: In this work, we propose CONDA, an approach to further improve GLM's ability to generate synthetic data by reformulating data generation as context generation for a given question-answer (QA) pair and leveraging QA datasets for training context generators.
- 661, TITLE: Meta-Learning Fast Weight Language Models  
<https://aclanthology.org/2022.emnlp-main.661>  
AUTHORS: Kevin Clark, Kelvin Guu, Ming-Wei Chang, Panupong Pasupat, Geoffrey Hinton, Mohammad Norouzi  
HIGHLIGHT: We present Fast Weight Layers (FWLs), a neural component that provides the benefits of dynamic evaluation much more efficiently by expressing gradient updates as linear attention.
- 662, TITLE: CTL++: Evaluating Generalization on Never-Seen Compositional Patterns of Known Functions, and Compatibility of Neural Representations  
<https://aclanthology.org/2022.emnlp-main.662>  
AUTHORS: Rbert Csords, Kazuki Irie, Juergen Schmidhuber  
HIGHLIGHT: Here we introduce CTL++, a new diagnostic dataset based on compositions of unary symbolic functions.
- 663, TITLE: Learning with Rejection for Abstractive Text Summarization  
<https://aclanthology.org/2022.emnlp-main.663>  
AUTHORS: Meng Cao, Yue Dong, Jingyi He, Jackie Chi Kit Cheung  
HIGHLIGHT: In this work, we propose a training objective for abstractive summarization based on rejection learning, in which the model learns whether or not to reject potentially noisy tokens.
- 664, TITLE: Adaptive Label Smoothing with Self-Knowledge in Natural Language Generation  
<https://aclanthology.org/2022.emnlp-main.664>  
AUTHORS: Dongkyu Lee, Ka Chun Cheung, Nevin Zhang  
HIGHLIGHT: In other words, label smoothing does not reflect the change in probability distribution mapped by a model over the course of training. To address this issue, we propose a regularization scheme that brings dynamic nature into the smoothing parameter by taking model probability distribution into account, thereby varying the parameter per instance.
- 665, TITLE: Hard Gate Knowledge Distillation - Leverage Calibration for Robust and Reliable Language Model  
<https://aclanthology.org/2022.emnlp-main.665>  
AUTHORS: Dongkyu Lee, Zhiliang Tian, Yingxiu Zhao, Ka Chun Cheung, Nevin Zhang  
HIGHLIGHT: Knowledge of a teacher is considered a subject that holds inter-class relations which send a meaningful supervision to a student; hence, much effort has been put to find such knowledge to be distilled. In this paper, we explore a question that has been given little attention: when to distill such knowledge.
- 666, TITLE: Are All Spurious Features in Natural Language Alike? An Analysis through a Causal Lens  
<https://aclanthology.org/2022.emnlp-main.666>  
AUTHORS: Nitish Joshi, Xiang Pan, He He  
HIGHLIGHT: Therefore, a more fine-grained treatment of spurious features is needed to specify the desired model behavior. We formalize this distinction using a causal model and probabilities of necessity and sufficiency, which delineates the causal relations between a feature and a label.
- 667, TITLE: Correcting Diverse Factual Errors in Abstractive Summarization via Post-Editing and Language Model Infilling  
<https://aclanthology.org/2022.emnlp-main.667>  
AUTHORS: Vidhisha Balachandran, Hannaneh Hajishirzi, William Cohen, Yulia Tsvetkov  
HIGHLIGHT: In this work, we propose to generate hard, representative synthetic examples of non-factual summaries through infilling language models.
- 668, TITLE: Coordinated Topic Modeling  
<https://aclanthology.org/2022.emnlp-main.668>  
AUTHORS: Pritom Saha Akash, Jie Huang, Kevin Chen-Chuan Chang  
HIGHLIGHT: We propose a new problem called coordinated topic modeling that imitates human behavior while describing a text corpus.
- 669, TITLE: Large Dual Encoders Are Generalizable Retrievers  
<https://aclanthology.org/2022.emnlp-main.669>

**AUTHORS:** Jianmo Ni, Chen Qu, Jing Lu, Zhuyun Dai, Gustavo Hernandez Abrego, Ji Ma, Vincent Zhao, Yi Luan, Keith Hall, Ming-Wei Chang, Yinfei Yang

**HIGHLIGHT:** One widespread belief is that the bottleneck layer of a dual encoder, where the final score is simply a dot-product between a query vector and a passage vector, is too limited compared to models with fine-grained interactions between the query and the passage. In this paper, we challenge this belief by scaling up the size of the dual encoder model while keeping the bottleneck layer as a single dot-product with a fixed size.

670, **TITLE:** CRIPP-VQA: Counterfactual Reasoning about Implicit Physical Properties via Video Question Answering  
<https://aclanthology.org/2022.emnlp-main.670>

**AUTHORS:** Maitreya Patel, Tejas Gokhale, Chitta Baral, Yezhou Yang

**HIGHLIGHT:** In this paper, we introduce CRIPP-VQA, a new video question answering dataset for reasoning about the implicit physical properties of objects in a scene.

671, **TITLE:** Entity-centered Cross-document Relation Extraction

<https://aclanthology.org/2022.emnlp-main.671>

**AUTHORS:** Fengqi Wang, Fei Li, Hao Fei, Jingye Li, Shengqiong Wu, Fangfang Su, Wenxuan Shi, Donghong Ji, Bo Cai

**HIGHLIGHT:** Moreover, they utilize all the text paths in a document bag in a coarse-grained way, without considering the connections between these text paths. In this paper, we aim to address both of these shortages and push the state-of-the-art for cross-document RE.

672, **TITLE:** Exploring Document-Level Literary Machine Translation with Parallel Paragraphs from World Literature

<https://aclanthology.org/2022.emnlp-main.672>

**AUTHORS:** Katherine Thai, Marzena Karpinska, Kalpesh Krishna, Bill Ray, Moira Inghilleri, John Wieting, Mohit Iyyer

**HIGHLIGHT:** The experts note that MT outputs contain not only mistranslations, but also discourse-disrupting errors and stylistic inconsistencies. To address these problems, we train a post-editing model whose output is preferred over normal MT output at a rate of 69% by experts.

673, **TITLE:** Label-aware Multi-level Contrastive Learning for Cross-lingual Spoken Language Understanding

<https://aclanthology.org/2022.emnlp-main.673>

**AUTHORS:** Shining Liang, Linjun Shou, Jian Pei, Ming Gong, Wanli Zuo, Xianglin Zuo, Daxin Jiang

**HIGHLIGHT:** In this paper, we propose to model the utterance-slot-word structure by a multi-level contrastive learning framework at the utterance, slot and word levels to facilitate explicit alignment.

674, **TITLE:** Polyglot Prompt: Multilingual Multitask Prompt Training

<https://aclanthology.org/2022.emnlp-main.674>

**AUTHORS:** Jinlan Fu, See-Kiong Ng, Pengfei Liu

**HIGHLIGHT:** This paper aims for a potential architectural improvement for multilingual learning and asks: Can different tasks from different languages be modeled in a monolithic framework, i. e. without any task/language-specific module?

675, **TITLE:** VisToT: Vision-Augmented Table-to-Text Generation

<https://aclanthology.org/2022.emnlp-main.675>

**AUTHORS:** Prajwal Gatti, Anand Mishra, Manish Gupta, Mithun Das Gupta

**HIGHLIGHT:** For example, in the tourism domain, images can be used to infer knowledge such as the type of landmark (e. g. , church), its architecture (e. g. , Ancient Roman), and composition (e. g. , white marble). Therefore, in this paper, we introduce the novel task of Vision-augmented Table-To-Text Generation (VisToT, defined as follows: given a table and an associated image, produce a descriptive sentence conditioned on the multimodal input.

676, **TITLE:** Generative Entity-to-Entity Stance Detection with Knowledge Graph Augmentation

<https://aclanthology.org/2022.emnlp-main.676>

**AUTHORS:** Xinliang Frederick Zhang, Nick Beauchamp, Lu Wang

**HIGHLIGHT:** In this paper, we emphasize the imperative need for studying interactions among entities when inferring stances.

677, **TITLE:** Symptom Identification for Interpretable Detection of Multiple Mental Disorders on Social Media

<https://aclanthology.org/2022.emnlp-main.677>

**AUTHORS:** Zhiling Zhang, Siyuan Chen, Mengyue Wu, Kenny Zhu

**HIGHLIGHT:** This paper introduces PsySym, the first annotated symptom identification corpus of multiple psychiatric disorders, to facilitate further research progress.

678, **TITLE:** Improving Iterative Text Revision by Learning Where to Edit from Other Revision Tasks

<https://aclanthology.org/2022.emnlp-main.678>

**AUTHORS:** Zae Myung Kim, Wanyu Du, Vipul Raheja, Dhruv Kumar, Dongyeop Kang

**HIGHLIGHT:** In this work, we aim to build an end-to-end text revision system that can iteratively generate helpful edits by explicitly detecting editable spans (where-to-edit) with their corresponding edit intents and then instructing a revision model to revise the detected edit spans.

679, **TITLE:** CONQRR: Conversational Query Rewriting for Retrieval with Reinforcement Learning

<https://aclanthology.org/2022.emnlp-main.679>

**AUTHORS:** Zeqiu Wu, Yi Luan, Hannah Rashkin, David Reitter, Hannaneh Hajishirzi, Mari Ostendorf, Gaurav Singh Tomar



- HIGHLIGHT:** Moreover, it can be expensive to re-train well-established retrievers such as search engines that are originally developed for non-conversational queries. To facilitate their use, we develop a query rewriting model CONQRR that rewrites a conversational question in the context into a standalone question.
- 680, **TITLE:** Specializing Multi-domain NMT via Penalizing Low Mutual Information  
<https://aclanthology.org/2022.emnlp-main.680>  
**AUTHORS:** Jiyoung Lee, Hantae Kim, Hyunchang Cho, Edward Choi, Cheonbok Park  
**HIGHLIGHT:** In this paper, we investigate domain-specific information through the lens of mutual information (MI) and propose a new objective that penalizes low MI to become higher.
- 681, **TITLE:** A Simple Contrastive Learning Framework for Interactive Argument Pair Identification via Argument-Context Extraction  
<https://aclanthology.org/2022.emnlp-main.681>  
**AUTHORS:** Lida Shi, Fausto Giunchiglia, Rui Song, Daqian Shi, Tongtong Liu, Xiaolei Diao, Hao Xu  
**HIGHLIGHT:** However, current context-based methods achieve limited improvements since the entire context typically contains much irrelevant information. In this paper, we propose a simple contrastive learning framework to solve this problem by extracting valuable information from the context.
- 682, **TITLE:** Sentence-level Media Bias Analysis Informed by Discourse Structures  
<https://aclanthology.org/2022.emnlp-main.682>  
**AUTHORS:** Yuanyuan Lei, Ruihong Huang, Lu Wang, Nick Beauchamp  
**HIGHLIGHT:** In this paper, we aim to identify sentences within an article that can illuminate and explain the overall bias of the entire article.
- 683, **TITLE:** Towards Efficient Dialogue Pre-training with Transferable and Interpretable Latent Structure  
<https://aclanthology.org/2022.emnlp-main.683>  
**AUTHORS:** Xueliang Zhao, Lema Liu, Tingchen Fu, Shuming Shi, Dongyan Zhao, Rui Yan  
**HIGHLIGHT:** This paper proposes a novel dialogue generation model with a latent structure that is easily transferable from the general domain to downstream tasks in a lightweight and transparent way.
- 684, **TITLE:** An Empirical Revisiting of Linguistic Knowledge Fusion in Language Understanding Tasks  
<https://aclanthology.org/2022.emnlp-main.684>  
**AUTHORS:** Changlong Yu, Tianyi Xiao, Lingpeng Kong, Yangqiu Song, Wilfred Ng  
**HIGHLIGHT:** Hence we call for attention to using trivial graphs as necessary baselines to design advanced knowledge fusion methods in the future.
- 685, **TITLE:** Unsupervised Non-transferable Text Classification  
<https://aclanthology.org/2022.emnlp-main.685>  
**AUTHORS:** Guangtao Zeng, Wei Lu  
**HIGHLIGHT:** In this paper, we propose a novel unsupervised non-transferable learning method for the text classification task that does not require annotated target domain data.
- 686, **TITLE:** Adaptive Contrastive Learning on Multimodal Transformer for Review Helpfulness Prediction  
<https://aclanthology.org/2022.emnlp-main.686>  
**AUTHORS:** Thong Nguyen, Xiaobao Wu, Anh Tuan Luu, Zhen Hai, Lidong Bing  
**HIGHLIGHT:** To overcome the aforementioned issues, we propose Multi-modal Contrastive Learning for Multimodal Review Helpfulness Prediction (MRHP) problem, concentrating on mutual information between input modalities to explicitly elaborate cross-modal relations.
- 687, **TITLE:** Adaptive Token-level Cross-lingual Feature Mixing for Multilingual Neural Machine Translation  
<https://aclanthology.org/2022.emnlp-main.687>  
**AUTHORS:** Junpeng Liu, Kaiyu Huang, Jiuyi Li, Huan Liu, Jinsong Su, Degen Huang  
**HIGHLIGHT:** In this paper, we propose a novel token-level feature mixing method that enables the model to capture different features and dynamically determine the feature sharing across languages.
- 688, **TITLE:** A Dataset for Hyper-Relational Extraction and a Cube-Filling Approach  
<https://aclanthology.org/2022.emnlp-main.688>  
**AUTHORS:** Yew Ken Chia, Lidong Bing, Sharifah Mahani Aljunied, Luo Si, Soujanya Poria  
**HIGHLIGHT:** Hence, we propose CubeRE, a cube-filling model inspired by table-filling approaches and explicitly considers the interaction between relation triplets and qualifiers.
- 689, **TITLE:** Low-resource Neural Machine Translation with Cross-modal Alignment  
<https://aclanthology.org/2022.emnlp-main.689>  
**AUTHORS:** Zhe Yang, Qingkai Fang, Yang Feng  
**HIGHLIGHT:** In this paper, we turn to connect several low-resource languages to a particular high-resource one by additional visual modality.
- 690, **TITLE:** Prompt-based Distribution Alignment for Domain Generalization in Text Classification  
<https://aclanthology.org/2022.emnlp-main.690>  
**AUTHORS:** Chen Jia, Yue Zhang

**HIGHLIGHT:** To improve domain generalization with prompting, we learn distributional invariance across source domains via two alignment regularization loss functions.

691, **TITLE:** Two is Better than Many? Binary Classification as an Effective Approach to Multi-Choice Question Answering  
<https://aclanthology.org/2022.emnlp-main.691>  
**AUTHORS:** Deepanway Ghosal, Navonil Majumder, Rada Mihalcea, Soujanya Poria  
**HIGHLIGHT:** We propose a simple refactoring of multi-choice question answering (MCQA) tasks as a series of binary classifications.

692, **TITLE:** HEGEL: Hypergraph Transformer for Long Document Summarization  
<https://aclanthology.org/2022.emnlp-main.692>  
**AUTHORS:** Haopeng Zhang, Xiao Liu, Jiawei Zhang  
**HIGHLIGHT:** This paper proposes HEGEL, a hypergraph neural network for long document summarization by capturing high-order cross-sentence relations.

693, **TITLE:** Adapting a Language Model While Preserving its General Knowledge  
<https://aclanthology.org/2022.emnlp-main.693>  
**AUTHORS:** Zixuan Ke, Yijia Shao, Haowei Lin, Hu Xu, Lei Shu, Bing Liu  
**HIGHLIGHT:** This paper shows that the existing methods are suboptimal and proposes a novel method to perform a more informed adaptation of the knowledge in the LM by (1) soft-masking the attention heads based on their importance to best preserve the general knowledge in the LM and (2) contrasting the representations of the general and the full (both general and domain knowledge) to learn an integrated representation with both general and domain-specific knowledge.

694, **TITLE:** Human Guided Exploitation of Interpretable Attention Patterns in Summarization and Topic Segmentation  
<https://aclanthology.org/2022.emnlp-main.694>  
**AUTHORS:** Raymond Li, Wen Xiao, Linzi Xing, Lanjun Wang, Gabriel Murray, Giuseppe Carenini  
**HIGHLIGHT:** In another vein, researchers propose new attention augmentation methods to make transformers more accurate, efficient and interpretable. In this paper, we combine these two lines of research in a human-in-the-loop pipeline to first discover important task-specific attention patterns.

695, **TITLE:** Continual Training of Language Models for Few-Shot Learning  
<https://aclanthology.org/2022.emnlp-main.695>  
**AUTHORS:** Zixuan Ke, Haowei Lin, Yijia Shao, Hu Xu, Lei Shu, Bing Liu  
**HIGHLIGHT:** Adapting or posttraining an LM using an unlabeled domain corpus can produce even better performance for end-tasks in the domain. This paper proposes the problem of continually extending an LM by incrementally post-train the LM with a sequence of unlabeled domain corpora to expand its knowledge without forgetting its previous skills.

696, **TITLE:** Dictionary-Assisted Supervised Contrastive Learning  
<https://aclanthology.org/2022.emnlp-main.696>  
**AUTHORS:** Patrick Wu, Richard Bonneau, Joshua Tucker, Jonathan Nagler  
**HIGHLIGHT:** We introduce the dictionary-assisted supervised contrastive learning (DASCL) objective, allowing researchers to leverage specialized dictionaries when fine-tuning pretrained language models.

697, **TITLE:** Fine-Tuning Pre-trained Transformers into Decaying Fast Weights  
<https://aclanthology.org/2022.emnlp-main.697>  
**AUTHORS:** Huanru Henry Mao  
**HIGHLIGHT:** Recent work proposes kernel-based methods to approximate causal self-attention by replacing it with recurrent formulations with various update rules and feature maps to achieve O(1) time and memory complexity. We explore these approaches and find that they are unnecessarily complex, and propose a simple alternative - decaying fast weights - that runs fast on GPU, outperforms prior methods, and retains 99% of attention performance for GPT-2.

698, **TITLE:** PRO-CS : An Instance-Based Prompt Composition Technique for Code-Switched Tasks  
<https://aclanthology.org/2022.emnlp-main.698>  
**AUTHORS:** Srijan Bansal, Suraj Tripathi, Sumit Agarwal, Teruko Mitamura, Eric Nyberg  
**HIGHLIGHT:** In this paper, we propose a novel instance-based prompt composition technique, PRO-CS, for CS tasks that combine language and task knowledge.

699, **TITLE:** SentBS: Sentence-level Beam Search for Controllable Summarization  
<https://aclanthology.org/2022.emnlp-main.699>  
**AUTHORS:** Chenhui Shen, Liying Cheng, Lidong Bing, Yang You, Luo Si  
**HIGHLIGHT:** However, current structure-controlling methods have limited effectiveness in enforcing the desired structure. To address this limitation, we propose a sentence-level beam search generation method (SentBS), where evaluation is conducted throughout the generation process to select suitable sentences for subsequent generations.

700, **TITLE:** A Fine-grained Chinese Software Privacy Policy Dataset for Sequence Labeling and Regulation Compliant Identification  
<https://aclanthology.org/2022.emnlp-main.700>  
**AUTHORS:** Kaifa Zhao, Le Yu, Shiyao Zhou, Jing Li, Xiapu Luo, Yat Fei Aemon Chiu, Yutong Liu  
**HIGHLIGHT:** In this paper, we construct the first Chinese privacy policy dataset, namely CA4P-483, to facilitate the sequence labeling tasks and regulation compliance identification between privacy policies and software.

- 701, TITLE: Saving Dense Retriever from Shortcut Dependency in Conversational Search  
<https://aclanthology.org/2022.emnlp-main.701>  
AUTHORS: Sungdong Kim, Gangwoo Kim  
HIGHLIGHT: In this paper, we demonstrate the existence of a retrieval shortcut in CS, which causes models to retrieve passages solely relying on partial history while disregarding the latest question.
- 702, TITLE: Graph-Induced Transformers for Efficient Multi-Hop Question Answering  
<https://aclanthology.org/2022.emnlp-main.702>  
AUTHORS: Giwon Hong, Jeonghwan Kim, Junmo Kang, Sung-Hyon Myaeng  
HIGHLIGHT: Our work proposes the Graph-Induced Transformer (GIT) that applies graph-derived attention patterns directly into a PLM, without the need to employ external graph modules.
- 703, TITLE: DiscoSense: Commonsense Reasoning with Discourse Connectives  
<https://aclanthology.org/2022.emnlp-main.703>  
AUTHORS: Prajwal Bhargava, Vincent Ng  
HIGHLIGHT: We present DiscoSense, a benchmark for commonsense reasoning via understanding a wide variety of discourse connectives.
- 704, TITLE: Boosting Document-Level Relation Extraction by Mining and Injecting Logical Rules  
<https://aclanthology.org/2022.emnlp-main.704>  
AUTHORS: Shengda Fan, Shasha Mo, Jianwei Niu  
HIGHLIGHT: In this paper, we propose MILR, a logic enhanced framework that boosts DocRE by Mining and Injecting Logical Rules.
- 705, TITLE: MOCHA: A Multi-Task Training Approach for Coherent Text Generation from Cognitive Perspective  
<https://aclanthology.org/2022.emnlp-main.705>  
AUTHORS: Zhe Hu, Hou Pong Chan, Lifu Huang  
HIGHLIGHT: In this work, we propose a novel multi-task training strategy for long text generation grounded on the cognitive theory of writing, which empowers the model to learn essential subskills needed for writing including planning and reviewing besides end-to-end generation.
- 706, TITLE: Variational Autoencoder with Disentanglement Priors for Low-Resource Task-Specific Natural Language Generation  
<https://aclanthology.org/2022.emnlp-main.706>  
AUTHORS: Zhuang Li, Lizhen Qu, Qionгкаi Xu, Tongtong Wu, Tianyang Zhan, Gholamreza Haffari  
HIGHLIGHT: In this paper, we propose a variational autoencoder with disentanglement priors, VAE-Dprior, for task-specific natural language generation with none or a handful of task-specific labeled examples.
- 707, TITLE: CISLR: Corpus for Indian Sign Language Recognition  
<https://aclanthology.org/2022.emnlp-main.707>  
AUTHORS: Abhinav Joshi, Ashwani Bhat, Pradeep S, Priya Gole, Shashwat Gupta, Shreyansh Agarwal, Ashutosh Modi  
HIGHLIGHT: In recent years researchers have actively worked for sign languages like American Sign Languages, however, Indian Sign language is still far from data-driven tasks like machine translation. To address this gap, in this paper, we introduce a new dataset CISLR (Corpus for Indian Sign Language Recognition) for word-level recognition in Indian Sign Language using videos.
- 708, TITLE: Mask the Correct Tokens: An Embarrassingly Simple Approach for Error Correction  
<https://aclanthology.org/2022.emnlp-main.708>  
AUTHORS: Kai Shen, Yichong Leng, Xu Tan, Siliang Tang, Yuan Zhang, Wenjie Liu, Edward Lin  
HIGHLIGHT: Since the error rate of the incorrect sentence is usually low (e. g. , 10%), the correction model can only learn to correct on limited error tokens but trivially copy on most tokens (correct tokens), which harms the effective training of error correction. In this paper, we argue that the correct tokens should be better utilized to facilitate effective training and then propose a simple yet effective masking strategy to achieve this goal.
- 709, TITLE: AMAL: Meta Knowledge-Driven Few-Shot Adapter Learning  
<https://aclanthology.org/2022.emnlp-main.709>  
AUTHORS: S. K. Hong, Tae Young Jang  
HIGHLIGHT: In this study, we present a meta-learning-driven low-rank adapter pooling method, called AMAL, for leveraging pre-trained language models even with just a few data points.
- 710, TITLE: Discourse Context Predictability Effects in Hindi Word Order  
<https://aclanthology.org/2022.emnlp-main.710>  
AUTHORS: Sidharth Ranjan, Marten van Schijndel, Sumeet Agarwal, Rajakrishnan Rajkumar  
HIGHLIGHT: We test the hypothesis that discourse predictability influences Hindi syntactic choice.
- 711, TITLE:  $\hat{A}^{-\hat{A}_i^{\hat{A}}}$ Covid vaccine is against Covid but Oxford vaccine is made at Oxford! $\hat{A}^{-\hat{A}_i^{\hat{A}}}$  Semantic Interpretation of Proper Noun Compounds  
<https://aclanthology.org/2022.emnlp-main.711>  
AUTHORS: Keshav Kolluru, Gabriel Stanovsky, Mausam -

**HIGHLIGHT:** These are commonly used in short-form domains, such as news headlines, but are largely ignored in information-seeking applications. To address this limitation, we release a new manually annotated dataset, ProNCL, consisting of 22.5K proper noun compounds along with their free-form semantic interpretations.

712, **TITLE:** Context Limitations Make Neural Language Models More Human-Like  
<https://aclanthology.org/2022.emnlp-main.712>

**AUTHORS:** Tatsuki Kuribayashi, Yohei Oseki, Ana Brassard, Kentaro Inui

**HIGHLIGHT:** This study highlights a limitation of modern neural LMs as the model of choice for this purpose: there is a discrepancy between their context access capacities and that of humans.

713, **TITLE:** A Generative Model for End-to-End Argument Mining with Reconstructed Positional Encoding and Constrained Pointer Mechanism  
<https://aclanthology.org/2022.emnlp-main.713>

**AUTHORS:** Jianzhu Bao, Yuhang He, Yang Sun, Bin Liang, Jiachen Du, Bing Qin, Min Yang, Ruifeng Xu

**HIGHLIGHT:** In this paper, we investigate the end-to-end AM task from a novel perspective by proposing a generative framework, in which the expected outputs of AM are framed as a simple target sequence.

714, **TITLE:** Reflect, Not Reflex: Inference-Based Common Ground Improves Dialogue Response Quality  
<https://aclanthology.org/2022.emnlp-main.714>

**AUTHORS:** Pei Zhou, Hyundong Cho, Pegah Jandaghi, Dong-Ho Lee, Bill Yuchen Lin, Jay Pujara, Xiang Ren

**HIGHLIGHT:** Human communication relies on common ground (CG), the mutual knowledge and beliefs shared by participants, to produce coherent and interesting conversations. In this paper, we demonstrate that current response generation (RG) models produce generic and dull responses in dialogues because they act reflexively, failing to explicitly model CG, both due to the lack of CG in training data and the standard RG training procedure.

715, **TITLE:** FlowEval: A Consensus-Based Dialogue Evaluation Framework Using Segment Act Flows  
<https://aclanthology.org/2022.emnlp-main.715>

**AUTHORS:** Jianqiao Zhao, Yanyang Li, Wanyu Du, Yangfeng Ji, Dong Yu, Michael Lyu, Liwei Wang

**HIGHLIGHT:** Hence, we propose segment act, an extension of dialog act from utterance level to segment level, and crowdsource a large-scale dataset for it.

716, **TITLE:** FaD-VLP: Fashion Vision-and-Language Pre-training towards Unified Retrieval and Captioning  
<https://aclanthology.org/2022.emnlp-main.716>

**AUTHORS:** Suvir Mirchandani, Licheng Yu, Mengjiao Wang, Animesh Sinha, Wenwen Jiang, Tao Xiang, Ning Zhang

**HIGHLIGHT:** Additionally, these works have mainly been restricted to multimodal understanding tasks. To address these gaps, we make two key contributions. First, we propose a novel fashion-specific pre-training framework based on weakly-supervised triplets constructed from fashion image-text pairs.

717, **TITLE:** MM-Align: Learning Optimal Transport-based Alignment Dynamics for Fast and Accurate Inference on Missing Modality Sequences  
<https://aclanthology.org/2022.emnlp-main.717>

**AUTHORS:** Wei Han, Hui Chen, Min-Yen Kan, Soujanya Poria

**HIGHLIGHT:** In this paper, we present a novel approach named MM-Align to address the missing-modality inference problem.

718, **TITLE:** Evaluating the Knowledge Dependency of Questions  
<https://aclanthology.org/2022.emnlp-main.718>

**AUTHORS:** Hyeongdon Moon, Yoonseok Yang, Hangyeol Yu, Seunghyun Lee, Myeongho Jeong, Juneyoung Park, Jamin Shin, Minsam Kim, Seungtaek Choi

**HIGHLIGHT:** They fail to evaluate the MCQA's ability to assess the student's knowledge of the corresponding target fact. To tackle this issue, we propose a novel automatic evaluation metric, coined Knowledge Dependent Answerability (KDA), which measures the MCQA's answerability given knowledge of the target fact.

719, **TITLE:** MoSE: Modality Split and Ensemble for Multimodal Knowledge Graph Completion  
<https://aclanthology.org/2022.emnlp-main.719>

**AUTHORS:** Yu Zhao, Xiangrui Cai, Yike Wu, Haiwei Zhang, Ying Zhang, Guoqing Zhao, Ning Jiang

**HIGHLIGHT:** In this paper, we propose MoSE, a Modality Split representation learning and Ensemble inference framework for MKGC.

720, **TITLE:** Entropy-Based Vocabulary Substitution for Incremental Learning in Multilingual Neural Machine Translation  
<https://aclanthology.org/2022.emnlp-main.720>

**AUTHORS:** Kaiyu Huang, Peng Li, Jin Ma, Yang Liu

**HIGHLIGHT:** In this work, we propose an entropy-based vocabulary substitution (EVS) method that just needs to walk through new language pairs for incremental learning in a large-scale multilingual data updating while remaining the size of the vocabulary.

721, **TITLE:** Eliciting Knowledge from Large Pre-Trained Models for Unsupervised Knowledge-Grounded Conversation  
<https://aclanthology.org/2022.emnlp-main.721>

**AUTHORS:** Yanyang Li, Jianqiao Zhao, Michael Lyu, Liwei Wang

**HIGHLIGHT:** In this work, we answer the aforementioned question in unsupervised knowledge-grounded conversation.

- 722, TITLE: An Unsupervised, Geometric and Syntax-aware Quantification of Polysemy  
<https://aclanthology.org/2022.emnlp-main.722>  
AUTHORS: Anmol Goel, Charu Sharma, Ponnurangam Kumaraguru  
HIGHLIGHT: With scarce attention paid to polysemy in computational linguistics, and even scarcer attention toward quantifying polysemy, in this paper, we propose a novel, unsupervised framework to compute and estimate polysemy scores for words in multiple languages.
- 723, TITLE: Reorder and then Parse, Fast and Accurate Discontinuous Constituency Parsing  
<https://aclanthology.org/2022.emnlp-main.723>  
AUTHORS: Kailai Sun, Zuchao Li, Hai Zhao  
HIGHLIGHT: Motivated by the observation that a discontinuous constituent tree can be simply transformed into a pseudo-continuous one by artificially reordering words in the sentence, we propose a novel reordering method, thereby construct fast and accurate discontinuous constituency parsing systems working in continuous way.
- 724, TITLE: Making Science Simple: Corpora for the Lay Summarisation of Scientific Literature  
<https://aclanthology.org/2022.emnlp-main.724>  
AUTHORS: Tomas Goldsack, Zhihao Zhang, Chenghua Lin, Carolina Scarton  
HIGHLIGHT: However, current corpora for this task are limited in their size and scope, hindering the development of broadly applicable data-driven approaches. Aiming to rectify these issues, we present two novel lay summarisation datasets, PLOS (large-scale) and eLife (medium-scale), each of which contains biomedical journal articles alongside expert-written lay summaries.
- 725, TITLE: Looking at the Overlooked: An Analysis on the Word-Overlap Bias in Natural Language Inference  
<https://aclanthology.org/2022.emnlp-main.725>  
AUTHORS: Sara Rajae, Yadollah Yaghoobzadeh, Mohammad Taher Pilehvar  
HIGHLIGHT: In this paper, we focus on an overlooked aspect of the overlap bias in the NLI models: the reverse word-overlap bias.
- 726, TITLE: An Empirical Study on the Transferability of Transformer Modules in Parameter-efficient Fine-tuning  
<https://aclanthology.org/2022.emnlp-main.726>  
AUTHORS: Mohammad AkbarTajari, Sara Rajae, Mohammad Taher Pilehvar  
HIGHLIGHT: Parameter-efficient fine-tuning has garnered lots of attention in recent studies. On this subject, we investigate the capability of different transformer modules in transferring knowledge from a pre-trained model to a downstream task.
- 727, TITLE: CODER: An efficient framework for improving retrieval through COntextual Document Embedding Reranking  
<https://aclanthology.org/2022.emnlp-main.727>  
AUTHORS: George Zerveas, Navid Rekabsaz, Daniel Cohen, Carsten Eickhoff  
HIGHLIGHT: In this work, we investigate the impact of ranking context - an often overlooked aspect of learning dense retrieval models.
- 728, TITLE: AdapterShare: Task Correlation Modeling with Adapter Differentiation  
<https://aclanthology.org/2022.emnlp-main.728>  
AUTHORS: Zhi Chen, Bei Chen, Lu Chen, Kai Yu, Jian-Guang Lou  
HIGHLIGHT: More deeply, In this paper, we propose AdapterShare, an adapter differentiation method to explicitly model the task correlation among multiple tasks.
- 729, TITLE: Rethinking Task-Specific Knowledge Distillation: Contextualized Corpus as Better Textbook  
<https://aclanthology.org/2022.emnlp-main.729>  
AUTHORS: Chang Liu, Chongyang Tao, Jianxin Liang, Tao Shen, Jiazhan Feng, Quzhe Huang, Dongyan Zhao  
HIGHLIGHT: To mitigate the issues in the two gapped corpora, we present a better textbook for the student to learn: contextualized corpus that contextualizes task corpus with large-scale general corpus through relevance-based text retrieval.
- 730, TITLE: Recovering Gold from Black Sand: Multilingual Dense Passage Retrieval with Hard and False Negative Samples  
<https://aclanthology.org/2022.emnlp-main.730>  
AUTHORS: Tianhao Shen, Mingtong Liu, Ming Zhou, Deyi Xiong  
HIGHLIGHT: In this paper, we propose a novel multilingual dense passage retrieval framework, mHFN, to recover and utilize hard and false negative samples.
- 731, TITLE: The  $\tilde{\Delta}_i \hat{\Delta}_i$  Problem of Human Label Variation: On Ground Truth in Data, Modeling and Evaluation  
<https://aclanthology.org/2022.emnlp-main.731>  
AUTHORS: Barbara Plank  
HIGHLIGHT: However, this conventional practice assumes that there exists a \*ground truth\*, and neglects that there exists genuine human variation in labeling due to disagreement, subjectivity in annotation or multiple plausible answers. In this position paper, we argue that this big open problem of human label variation persists and critically needs more attention to move our field forward.
- 732, TITLE: Quality Scoring of Source Words in Neural Translation Models  
<https://aclanthology.org/2022.emnlp-main.732>

- AUTHORS: Priyesh Jain, Sunita Sarawagi, Tushar Tomar  
HIGHLIGHT: We propose a simple approach based on comparing the difference of probabilities from two language models.
- 733, TITLE: Pneg: Prompt-based Negative Response Generation for Dialogue Response Selection Task  
<https://aclanthology.org/2022.emnlp-main.733>  
AUTHORS: Nyoungwoo Lee, ChaeHun Park, Ho-Jin Choi, Jaegul Choo  
HIGHLIGHT: Nevertheless, collecting human-written adversarial responses is expensive, and existing synthesizing methods often have limited scalability. To overcome these limitations, this paper proposes a simple but efficient method for generating adversarial negative responses leveraging a large-scale language model.
- 734, TITLE: Facilitating Contrastive Learning of Discourse Relational Senses by Exploiting the Hierarchy of Sense Relations  
<https://aclanthology.org/2022.emnlp-main.734>  
AUTHORS: Wanqiu Long, Bonnie Webber  
HIGHLIGHT: Here we do more “incorporating the sense hierarchy into the recognition process itself and using it to select the negative examples used in contrastive learning.
- 735, TITLE: Simplified Graph Learning for Inductive Short Text Classification  
<https://aclanthology.org/2022.emnlp-main.735>  
AUTHORS: Kaixin Zheng, Yaqing Wang, Quanming Yao, Dejing Dou  
HIGHLIGHT: In this paper, we present SimpleSTC which handles inductive STC problem but only leverages words.
- 736, TITLE: Don't Stop Fine-Tuning: On Training Regimes for Few-Shot Cross-Lingual Transfer with Multilingual Language Models  
<https://aclanthology.org/2022.emnlp-main.736>  
AUTHORS: Fabian David Schmidt, Ivan Vulic, Goran Glavač  
HIGHLIGHT: In this work, we present a systematic study focused on a spectrum of FS-XLT fine-tuning regimes, analyzing key properties such as effectiveness, (in)stability, and modularity.
- 737, TITLE: Towards Compositional Generalization in Code Search  
<https://aclanthology.org/2022.emnlp-main.737>  
AUTHORS: Hojae Han, Seung-won Hwang, Shuai Lu, Nan Duan, Seungtaek Choi  
HIGHLIGHT: Thus we propose CTBERT, or Code Template BERT, representing codes using automatically extracted templates as building blocks.
- 738, TITLE: Towards relation extraction from speech  
<https://aclanthology.org/2022.emnlp-main.738>  
AUTHORS: Tongtong Wu, Guitao Wang, Jinming Zhao, Zhaoran Liu, Guilin Qi, Yuan-Fang Li, Gholamreza Haffari  
HIGHLIGHT: In this paper, we propose a new listening information extraction task, i. e. , speech relation extraction.
- 739, TITLE: Structural Constraints and Natural Language Inference for End-to-End Flowchart Grounded Dialog Response Generation  
<https://aclanthology.org/2022.emnlp-main.739>  
AUTHORS: Dinesh Raghu, Suraj Joshi, Sachindra Joshi, Mausam -  
HIGHLIGHT: In such cases, it fails to understand the correct polarity of the answer. To overcome these issues, we propose Structure-Aware FLONET (SA-FLONET) which infuses structural constraints derived from the connectivity structure of flowcharts into the RAG framework.
- 740, TITLE: SLICER: Sliced Fine-Tuning for Low-Resource Cross-Lingual Transfer for Named Entity Recognition  
<https://aclanthology.org/2022.emnlp-main.740>  
AUTHORS: Fabian David Schmidt, Ivan Vulic, Goran Glavač  
HIGHLIGHT: In this work, we introduce a simple yet highly effective approach for improving zero-shot transfer for NER to low-resource languages.
- 741, TITLE: EdgeFormer: A Parameter-Efficient Transformer for On-Device Seq2seq Generation  
<https://aclanthology.org/2022.emnlp-main.741>  
AUTHORS: Tao Ge, Si-Qing Chen, Furu Wei  
HIGHLIGHT: We introduce EdgeFormer a parameter-efficient Transformer for on-device seq2seq generation under the strict computation and memory constraints.
- 742, TITLE: End-to-End Unsupervised Vision-and-Language Pre-training with Referring Expression Matching  
<https://aclanthology.org/2022.emnlp-main.742>  
AUTHORS: Chi Chen, Peng Li, Maosong Sun, Yang Liu  
HIGHLIGHT: In this paper, we explore end-to-end unsupervised VLP with a vision encoder to directly encode images.
- 743, TITLE: Faithful Knowledge Graph Explanations in Commonsense Question Answering  
<https://aclanthology.org/2022.emnlp-main.743>  
AUTHORS: Guy Aglionby, Simone Teufel

**HIGHLIGHT:** A common way of incorporating facts from the graph is to encode them separately from the question, and then combine the two representations to select an answer. In this paper, we argue that highly faithful graph-based explanations cannot be extracted from existing models of this type.

744, **TITLE:** KOLD: Korean Offensive Language Dataset

<https://aclanthology.org/2022.emnlp-main.744>

**AUTHORS:** Younghoon Jeong, Juhyun Oh, Jongwon Lee, Jaimeen Ahn, Jihyung Moon, Sungjoon Park, Alice Oh  
**HIGHLIGHT:** In this paper, we present the Korean Offensive Language Dataset (KOLD) comprising 40,429 comments, which are annotated hierarchically with the type and the target of offensive language, accompanied by annotations of the corresponding text spans.

745, **TITLE:** Evade the Trap of Mediocrity: Promoting Diversity and Novelty in Text Generation via Concentrating Attention  
<https://aclanthology.org/2022.emnlp-main.745>

**AUTHORS:** Wenhao Li, Xiaoyuan Yi, Jinyi Hu, Maosong Sun, Xing Xie

**HIGHLIGHT:** Nevertheless, these models tend to produce dull high-frequency phrases, severely hurting the diversity and novelty of generated text. In this work, we dig into the intrinsic mechanism of this problem and found that sparser attention values in Transformer could improve diversity.

746, **TITLE:** The better your Syntax, the better your Semantics? Probing Pretrained Language Models for the English Comparative Correlative

<https://aclanthology.org/2022.emnlp-main.746>

**AUTHORS:** Leonie Weissweiler, Valentin Hofmann, Abdullatif K $\ddot{a}$ ksal, Hinrich Sch $\ddot{a}$ tzle

**HIGHLIGHT:** As a first step towards assessing the compatibility of CxG with the syntactic and semantic knowledge demonstrated by state-of-the-art pretrained language models (PLMs), we present an investigation of their capability to classify and understand one of the most commonly studied constructions, the English comparative correlative (CC).

747, **TITLE:** ProofInfer: Generating Proof via Iterative Hierarchical Inference

<https://aclanthology.org/2022.emnlp-main.747>

**AUTHORS:** Zichu Fei, Qi Zhang, Xin Zhou, Tao Gui, Xuanjing Huang

**HIGHLIGHT:** To this end, we propose a divide-and-conquer algorithm to encode the proof tree as the plain text without losing structure information.

748, **TITLE:** ECTSum: A New Benchmark Dataset For Bullet Point Summarization of Long Earnings Call Transcripts

<https://aclanthology.org/2022.emnlp-main.748>

**AUTHORS:** Rajdeep Mukherjee, Abhinav Bohra, Akash Banerjee, Soumya Sharma, Manjunath Hegde, Afreen Shaikh, Shivani Shrivastava, Koustuv Dasgupta, Niloy Ganguly, Saptarshi Ghosh, Pawan Goyal

**HIGHLIGHT:** In this work, we present ECTSum, a new dataset with transcripts of earnings calls (ECTs), hosted by publicly traded companies, as documents, and experts-written short telegram-style bullet point summaries derived from corresponding Reuters articles.

749, **TITLE:** Cross-domain Generalization for AMR Parsing

<https://aclanthology.org/2022.emnlp-main.749>

**AUTHORS:** Xuefeng Bai, Sen Yang, Leyang Cui, Linfeng Song, Yue Zhang

**HIGHLIGHT:** Based on our observation, we investigate two approaches to reduce the domain distribution divergence of text and AMR features, respectively.

750, **TITLE:** CiteSum: Citation Text-guided Scientific Extreme Summarization and Domain Adaptation with Limited Supervision

<https://aclanthology.org/2022.emnlp-main.750>

**AUTHORS:** Yuning Mao, Ming Zhong, Jiawei Han

**HIGHLIGHT:** In this paper, we propose a simple yet effective approach to automatically extracting TLDR summaries for scientific papers from their citation texts.

751, **TITLE:** FETA: A Benchmark for Few-Sample Task Transfer in Open-Domain Dialogue

<https://aclanthology.org/2022.emnlp-main.751>

**AUTHORS:** Alon Albalak, Yi-Lin Tuan, Pegah Jandaghi, Connor Pryor, Luke Yoffe, Deepak Ramachandran, Lise Getoor, Jay Pujara, William Yang Wang

**HIGHLIGHT:** We utilize three popular language models and three learning algorithms to analyze the transferability between 132 source-target task pairs and create a baseline for future work.

752, **TITLE:** Do Children Texts Hold The Key To Commonsense Knowledge?

<https://aclanthology.org/2022.emnlp-main.752>

**AUTHORS:** Julien Romero, Simon Razniewski

**HIGHLIGHT:** This paper explores whether children's texts hold the key to commonsense knowledge compilation, based on the hypothesis that such content makes fewer assumptions on the reader's knowledge, and therefore spells out commonsense more explicitly.

753, **TITLE:** On the Limitations of Reference-Free Evaluations of Generated Text

<https://aclanthology.org/2022.emnlp-main.753>

**AUTHORS:** Daniel Deutsch, Rotem Dror, Dan Roth

- HIGHLIGHT:** Therefore, we recommend that reference-free metrics should be used as diagnostic tools for analyzing and understanding model behavior instead of measures of how well models perform a task, in which the goal is to achieve as high of a score as possible.
- 754, **TITLE:** Sampling-Based Approximations to Minimum Bayes Risk Decoding for Neural Machine Translation  
<https://aclanthology.org/2022.emnlp-main.754>  
**AUTHORS:** Bryan Eikema, Wilker Aziz  
**HIGHLIGHT:** Recently, an approximation to minimum Bayes risk (MBR) decoding has been proposed as an alternative decision rule that would likely not suffer from the same problems. We analyse this approximation and establish that it has no equivalent to the beam search curse.
- 755, **TITLE:** IndicXNLI: Evaluating Multilingual Inference for Indian Languages  
<https://aclanthology.org/2022.emnlp-main.755>  
**AUTHORS:** Divyanshu Aggarwal, Vivek Gupta, Anoop Kunchukuttan  
**HIGHLIGHT:** To this end, we introduce INDICXNLI, an NLI dataset for 11 Indic languages.
- 756, **TITLE:** Model Cascading: Towards Jointly Improving Efficiency and Accuracy of NLP Systems  
<https://aclanthology.org/2022.emnlp-main.756>  
**AUTHORS:** Neeraj Varshney, Chitta Baral  
**HIGHLIGHT:** In this work, we present an explorative study on  $\tilde{A}^{-\hat{A}_i^{\hat{A}}/2}$  model cascading  $\tilde{A}^{-\hat{A}_i^{\hat{A}}/2}$ , a simple technique that utilizes a collection of models of varying capacities to accurately yet efficiently output predictions.
- 757, **TITLE:** Semantic Simplification for Sentiment Classification  
<https://aclanthology.org/2022.emnlp-main.757>  
**AUTHORS:** Xiaotong Jiang, Zhongqing Wang, Guodong Zhou  
**HIGHLIGHT:** To this end, we enhance the original text with a sentiment-driven simplified clause to intensify its sentiment.
- 758, **TITLE:** XPrompt: Exploring the Extreme of Prompt Tuning  
<https://aclanthology.org/2022.emnlp-main.758>  
**AUTHORS:** Fang Ma, Chen Zhang, Lei Ren, Jingang Wang, Qifan Wang, Wei Wu, Xiaojun Quan, Dawei Song  
**HIGHLIGHT:** While prompt tuning has gradually reached the performance level of fine-tuning as the model scale increases, there is still a large performance gap between prompt tuning and fine-tuning for models of moderate and small scales (typically less than 11B parameters). In this paper, we empirically show that the trained prompt tokens can have a negative impact on a downstream task and thus degrade its performance.
- 759, **TITLE:** Rethinking the Role of Demonstrations: What Makes In-Context Learning Work?  
<https://aclanthology.org/2022.emnlp-main.759>  
**AUTHORS:** Sewon Min, Xinxi Lyu, Ari Holtzman, Mikel Artetxe, Mike Lewis, Hannaneh Hajishirzi, Luke Zettlemoyer  
**HIGHLIGHT:** In this paper, we show that ground truth demonstrations are in fact not required  $\tilde{A}^{-\hat{A}_i^{\hat{A}}/2}$  randomly replacing labels in the demonstrations barely hurts performance on a range of classification and multi-choice tasks, consistently over 12 different models including GPT-3.
- 760, **TITLE:** The Curious Case of Control  
<https://aclanthology.org/2022.emnlp-main.760>  
**AUTHORS:** Elias Stengel-Eskin, Benjamin Van Durme  
**HIGHLIGHT:** We find that models can be categorized by behavior into three separate groups, with broad differences between the groups.
- 761, **TITLE:** SHARE: a System for Hierarchical Assistive Recipe Editing  
<https://aclanthology.org/2022.emnlp-main.761>  
**AUTHORS:** Shuyang Li, Yufei Li, Jianmo Ni, Julian McAuley  
**HIGHLIGHT:** To help them, we propose the task of controllable recipe editing: adapt a base recipe to satisfy a user-specified dietary constraint.
- 762, **TITLE:** IM<sup>2</sup>: an Interpretable and Multi-category Integrated Metric Framework for Automatic Dialogue Evaluation  
<https://aclanthology.org/2022.emnlp-main.762>  
**AUTHORS:** Zhihua Jiang, Guanghui Ye, Dongning Rao, Di Wang, Xin Miao  
**HIGHLIGHT:** To mitigate the problem, this paper proposes an interpretable, multi-faceted, and controllable framework IM<sup>2</sup> (Interpretable and Multi-category Integrated Metric) to combine a large number of metrics which are good at measuring different qualities.
- 763, **TITLE:** PEVL: Position-enhanced Pre-training and Prompt Tuning for Vision-language Models  
<https://aclanthology.org/2022.emnlp-main.763>  
**AUTHORS:** Yuan Yao, Qianyu Chen, Ao Zhang, Wei Ji, Zhiyuan Liu, Tat-Seng Chua, Maosong Sun  
**HIGHLIGHT:** To address the challenge, we introduce PEVL that enhances the pre-training and prompt tuning of VLP models with explicit object position modeling.
- 764, **TITLE:** Pre-training Language Models with Deterministic Factual Knowledge  
<https://aclanthology.org/2022.emnlp-main.764>  
**AUTHORS:** Shaobo Li, Xiaoguang Li, Lifeng Shang, Chengjie Sun, Bingquan Liu, Zhenzhou Ji, Xin Jiang, Qun Liu



**HIGHLIGHT:** However, some analyses reveal that PLMs fail to perform it robustly, e. g. , being sensitive to the changes of prompts when extracting factual knowledge. To mitigate this issue, we propose to let PLMs learn the deterministic relationship between the remaining context and the masked content.

765, **TITLE:** Finding Skill Neurons in Pre-trained Transformer-based Language Models  
<https://aclanthology.org/2022.emnlp-main.765>

**AUTHORS:** Xiaozhi Wang, Kaiyue Wen, Zhengyan Zhang, Lei Hou, Zhiyuan Liu, Juanzi Li

**HIGHLIGHT:** In this paper, we find that after prompt tuning for specific tasks, the activations of some neurons within pre-trained Transformers are highly predictive of the task labels.

766, **TITLE:** Prompt Conditioned VAE: Enhancing Generative Replay for Lifelong Learning in Task-Oriented Dialogue  
<https://aclanthology.org/2022.emnlp-main.766>

**AUTHORS:** Yingxiu Zhao, Yinhe Zheng, Zhiliang Tian, Chang Gao, Jian Sun, Nevin L. Zhang

**HIGHLIGHT:** In this paper, we propose a novel method, prompt conditioned VAE for lifelong learning (PCLL), to enhance generative replay by incorporating tasks  $\hat{A}_i^{-1/2}$  statistics.

767, **TITLE:** PreQuEL: Quality Estimation of Machine Translation Outputs in Advance

<https://aclanthology.org/2022.emnlp-main.767>

**AUTHORS:** Shachar Don-Yehiya, Leshem Choshen, Omri Abend

**HIGHLIGHT:** We present the task of PreQuEL, Pre-(Quality-Estimation) Learning.

768, **TITLE:** Can Transformers Reason in Fragments of Natural Language?

<https://aclanthology.org/2022.emnlp-main.768>

**AUTHORS:** Viktor Schlegel, Kamen Pavlov, Ian Pratt-Hartmann

**HIGHLIGHT:** In this paper we carry out a large-scale empirical study investigating the detection of formally valid inferences in controlled fragments of natural language for which the satisfiability problem becomes increasingly complex.

769, **TITLE:** Textless Speech Emotion Conversion using Discrete & Decomposed Representations

<https://aclanthology.org/2022.emnlp-main.769>

**AUTHORS:** Felix Kreuk, Adam Polyak, Jade Copet, Eugene Kharonov, Tu Anh Nguyen, Morgan Rivi $\hat{A}_i^{-1/2}$ , Wei-Ning Hsu, Abdelrahman Mohamed, Emmanuel Dupoux, Yossi Adi

**HIGHLIGHT:** In this study, we cast the problem of emotion conversion as a spoken language translation task.

770, **TITLE:** Textual Backdoor Attacks Can Be More Harmful via Two Simple Tricks

<https://aclanthology.org/2022.emnlp-main.770>

**AUTHORS:** Yangyi Chen, Fanchao Qi, Hongcheng Gao, Zhiyuan Liu, Maosong Sun

**HIGHLIGHT:** In this paper, we find two simple tricks that can make existing textual backdoor attacks much more harmful.

771, **TITLE:** Why Should Adversarial Perturbations be Imperceptible? Rethink the Research Paradigm in Adversarial NLP

<https://aclanthology.org/2022.emnlp-main.771>

**AUTHORS:** Yangyi Chen, Hongcheng Gao, Ganqu Cui, Fanchao Qi, Longtao Huang, Zhiyuan Liu, Maosong Sun

**HIGHLIGHT:** In this paper, we rethink the research paradigm of textual adversarial samples in security scenarios.

772, **TITLE:** Retrieval Augmented Visual Question Answering with Outside Knowledge

<https://aclanthology.org/2022.emnlp-main.772>

**AUTHORS:** Weizhe Lin, Bill Byrne

**HIGHLIGHT:** Instead, we propose a joint training scheme which includes differentiable DPR integrated with answer generation so that the system can be trained in an end-to-end fashion.

773, **TITLE:** Instance Regularization for Discriminative Language Model Pre-training

<https://aclanthology.org/2022.emnlp-main.773>

**AUTHORS:** Zhuosheng Zhang, Hai Zhao, Ming Zhou

**HIGHLIGHT:** To model explicit signals of instance contribution, this work proposes to estimate the complexity of restoring the original sentences from corrupted ones in language model pre-training.

774, **TITLE:** GuoFeng: A Benchmark for Zero Pronoun Recovery and Translation

<https://aclanthology.org/2022.emnlp-main.774>

**AUTHORS:** Mingzhou Xu, Longyue Wang, Derek F. Wong, Hongye Liu, Linfeng Song, Lidia S. Chao, Shuming Shi, Zhaopeng Tu

**HIGHLIGHT:** To bridge the data and evaluation gaps, we propose a benchmark testset for target evaluation on Chinese-English ZP translation.

775, **TITLE:** ScienceWorld: Is your Agent Smarter than a 5th Grader?

<https://aclanthology.org/2022.emnlp-main.775>

**AUTHORS:** Ruoyao Wang, Peter Jansen, Marc-Alexandre C $\hat{A}_i^{-1/2}$ , Prithviraj Ammanabrolu

**HIGHLIGHT:** We present ScienceWorld, a benchmark to test agents  $\hat{A}_i^{-1/2}$  scientific reasoning abilities in a new interactive text environment at the level of a standard elementary school science curriculum.

776, **TITLE:** Improving Embeddings Representations for Comparing Higher Education Curricula: A Use Case in Computing

<https://aclanthology.org/2022.emnlp-main.776>

- AUTHORS: Jeffri Murrugarra-Llerena, Fernando Alva-Manchego, Nils Murrugarra-Llerena  
HIGHLIGHT: We propose an approach for comparing curricula of study programs in higher education.
- 777, TITLE: Mitigating Spurious Correlation in Natural Language Understanding with Counterfactual Inference  
<https://aclanthology.org/2022.emnlp-main.777>  
AUTHORS: Can Udomcharoenchaikit, Wuttikorn Ponwitayarat, Patomporn Payoungkhamdee, Kanruethai Masuk, Weerayut Buaphet, Ekapol Chuangsuwanich, Sarana Nutanong  
HIGHLIGHT: In this paper, we propose a causal analysis framework to help debias NLU models.
- 778, TITLE: End-to-End Neural Discourse Deixis Resolution in Dialogue  
<https://aclanthology.org/2022.emnlp-main.778>  
AUTHORS: Shengjie Li, Vincent Ng  
HIGHLIGHT: We adapt Lee et al. (2018) span-based entity coreference model to the task of end-to-end discourse deixis resolution in dialogue, specifically by proposing extensions to their model that exploit task-specific characteristics.
- 779, TITLE: Balancing out Bias: Achieving Fairness Through Balanced Training  
<https://aclanthology.org/2022.emnlp-main.779>  
AUTHORS: Xudong Han, Timothy Baldwin, Trevor Cohn  
HIGHLIGHT: To achieve Equal Opportunity fairness, such as equal job opportunity without regard to demographics, this paper introduces a simple, but highly effective, objective for countering bias using balanced training.
- 780, TITLE: Prompting ELECTRA: Few-Shot Learning with Discriminative Pre-Trained Models  
<https://aclanthology.org/2022.emnlp-main.780>  
AUTHORS: Mengzhou Xia, Mikel Artetxe, Jingfei Du, Danqi Chen, Veselin Stoyanov  
HIGHLIGHT: In this work, we adapt prompt-based few-shot learning to ELECTRA and show that it outperforms masked language models in a wide range of tasks.
- 781, TITLE: Identifying Physical Object Use in Sentences  
<https://aclanthology.org/2022.emnlp-main.781>  
AUTHORS: Tianyu Jiang, Ellen Riloff  
HIGHLIGHT: We define a new task called ObjectUse Classification that determines whether a physical object mentioned in a sentence was used or likely will be used. We introduce a new dataset for this task and present a classification model that exploits data augmentation methods and FrameNet when fine-tuning a pre-trained language model.
- 782, TITLE: CDialog: A Multi-turn Covid-19 Conversation Dataset for Entity-Aware Dialog Generation  
<https://aclanthology.org/2022.emnlp-main.782>  
AUTHORS: Deeksha Varshney, Aizan Zafar, Niranshu Behera, Asif Ekbal  
HIGHLIGHT: CDialog: A Multi-turn Covid-19 Conversation Dataset for Entity-Aware Dialog Generation
- 783, TITLE: Robustifying Sentiment Classification by Maximally Exploiting Few Counterfactuals  
<https://aclanthology.org/2022.emnlp-main.783>  
AUTHORS: Maarten De Raedt, Frédéric Godin, Chris Develder, Thomas Demeester  
HIGHLIGHT: Thus, we propose a novel solution that only requires annotation of a small fraction (e. g. , 1%) of the original training data, and uses automatic generation of extra counterfactuals in an encoding vector space.
- 784, TITLE: Data-Efficient Playlist Captioning With Musical and Linguistic Knowledge  
<https://aclanthology.org/2022.emnlp-main.784>  
AUTHORS: Giovanni Gabbolini, Romain Hennequin, Elena Epure  
HIGHLIGHT: In this work, we propose PlayNTell, a data-efficient multi-modal encoder-decoder model for automatic playlist captioning.
- 785, TITLE: Improved grammatical error correction by ranking elementary edits  
<https://aclanthology.org/2022.emnlp-main.785>  
AUTHORS: Alexey Sorokin  
HIGHLIGHT: We offer a two-stage reranking method for grammatical error correction: the first model serves as edit generator, while the second classifies the proposed edits as correct or false.
- 786, TITLE: Improving Tokenisation by Alternative Treatment of Spaces  
<https://aclanthology.org/2022.emnlp-main.786>  
AUTHORS: Edward Gow-Smith, Harish Tayyar Madabushi, Carolina Scarton, Aline Villavicencio  
HIGHLIGHT: Improving Tokenisation by Alternative Treatment of Spaces
- 787, TITLE: GENIE: Toward Reproducible and Standardized Human Evaluation for Text Generation  
<https://aclanthology.org/2022.emnlp-main.787>  
AUTHORS: Daniel Khashabi, Gabriel Stanovsky, Jonathan Bragg, Nicholas Lourie, Jungo Kasai, Yejin Choi, Noah A. Smith, Daniel Weld  
HIGHLIGHT: Furthermore, we develop an automated mechanism for maintaining annotator quality via a probabilistic model that detects and excludes noisy annotators. Putting these lessons together, we introduce GENIE: a system for running standardized human evaluations across different generation tasks.

- 788, TITLE: Attentional Probe: Estimating a Module's Functional Potential  
<https://aclanthology.org/2022.emnlp-main.788>  
AUTHORS: Tiago Pimentel, Josef Valvoda, Niklas Stoeck, Ryan Cotterell  
HIGHLIGHT: Attentional Probe: Estimating a Module's Functional Potential
- 789, TITLE: When More Data Hurts: A Troubling Quirk in Developing Broad-Coverage Natural Language Understanding Systems  
<https://aclanthology.org/2022.emnlp-main.789>  
AUTHORS: Elias Stengel-Eskin, Emmanouil Antonios Platanios, Adam Pauls, Sam Thomson, Hao Fang, Benjamin Van Durme, Jason Eisner, Yu Su  
HIGHLIGHT: This requires additional training data and results in ever-growing datasets. We present the first systematic investigation into this incremental symbol learning scenario.
- 790, TITLE: Zero-shot Cross-lingual Transfer of Prompt-based Tuning with a Unified Multilingual Prompt  
<https://aclanthology.org/2022.emnlp-main.790>  
AUTHORS: Lianzhe Huang, Shuming Ma, Dongdong Zhang, Furu Wei, Houfeng Wang  
HIGHLIGHT: To alleviate the effort of designing different prompts for multiple languages, we propose a novel model that uses a unified prompt for all languages, called UniPrompt.
- 791, TITLE: Three Real-World Datasets and Neural Computational Models for Classification Tasks in Patent Landscaping  
<https://aclanthology.org/2022.emnlp-main.791>  
AUTHORS: Subhash Pujari, Jannik Strötgen, Mark Giereth, Michael Gertz, Annemarie Friedrich  
HIGHLIGHT: With this paper, we release three labeled datasets for PLS-oriented classification tasks covering two diverse domains.
- 792, TITLE: Topic Modeling With Topological Data Analysis  
<https://aclanthology.org/2022.emnlp-main.792>  
AUTHORS: Ciaran Byrne, Danijela Horak, Karo Moilanen, Amandla Mabona  
HIGHLIGHT: We present an unsupervised topic modeling method which harnesses Topological Data Analysis (TDA) to extract a topological skeleton of the manifold upon which contextualised word embeddings lie.
- 793, TITLE: Predicting Fine-Tuning Performance with Probing  
<https://aclanthology.org/2022.emnlp-main.793>  
AUTHORS: Zining Zhu, Soroosh Shahtalebi, Frank Rudzicz  
HIGHLIGHT: This paper explores the utility of probing deep NLP models to extract a proxy signal widely used in model development: the fine-tuning performance.
- 794, TITLE: Diverse Parallel Data Synthesis for Cross-Database Adaptation of Text-to-SQL Parsers  
<https://aclanthology.org/2022.emnlp-main.794>  
AUTHORS: Abhijeet Awasthi, Ashutosh Sathe, Sunita Sarawagi  
HIGHLIGHT: We present ReFill, a framework for synthesizing high-quality and textually diverse parallel datasets for adapting Text-to-SQL parsers.
- 795, TITLE: Agent-Specific Deontic Modality Detection in Legal Language  
<https://aclanthology.org/2022.emnlp-main.795>  
AUTHORS: Abhilasha Sancheti, Aparna Garimella, Balaji Vasan Srinivasan, Rachel Rudinger  
HIGHLIGHT: To this end, we introduce, LEXDEMOD, a corpus of English contracts annotated with deontic modality expressed with respect to a contracting party or agent along with the modal triggers.
- 796, TITLE: COLD: A Benchmark for Chinese Offensive Language Detection  
<https://aclanthology.org/2022.emnlp-main.796>  
AUTHORS: Jiawen Deng, Jingyan Zhou, Hao Sun, Chujie Zheng, Fei Mi, Helen Meng, Minlie Huang  
HIGHLIGHT: To this end, we propose a benchmark COLD for Chinese offensive language analysis, including a Chinese Offensive Language Dataset COLDDATASET and a baseline detector COLDETECTOR which is trained on the dataset.
- 797, TITLE: Fixing Model Bugs with Natural Language Patches  
<https://aclanthology.org/2022.emnlp-main.797>  
AUTHORS: Shikhar Murty, Christopher Manning, Scott Lundberg, Marco Tulio Ribeiro  
HIGHLIGHT: We model the task of determining if a patch applies separately from the task of integrating patch information, and show that with a small amount of synthetic data, we can teach models to effectively use real patches on real data. 1 to 7 patches improve accuracy by ~14 accuracy points on different slices of a sentiment analysis dataset, and F1 by 7 points on a relation extraction dataset.
- 798, TITLE: WeDef: Weakly Supervised Backdoor Defense for Text Classification  
<https://aclanthology.org/2022.emnlp-main.798>  
AUTHORS: Lesheng Jin, Zihan Wang, Jingbo Shang  
HIGHLIGHT: To defend different trigger types at once, we start from the class-irrelevant nature of the poisoning process and propose a novel weakly supervised backdoor defense framework WeDef.

- 799, TITLE: Interventional Training for Out-Of-Distribution Natural Language Understanding  
<https://aclanthology.org/2022.emnlp-main.799>  
AUTHORS: Sicheng Yu, Jing Jiang, Hao Zhang, Yulei Niu, Qianru Sun, Lidong Bing  
HIGHLIGHT: In this paper, we propose a novel interventional training method called Bottom-up Automatic Intervention (BAI) that performs multi-granular intervention with identified multifactorial confounders.
- 800, TITLE: Pseudo-Relevance for Enhancing Document Representation  
<https://aclanthology.org/2022.emnlp-main.800>  
AUTHORS: Jihyuk Kim, Seung-won Hwang, Seoho Song, Hyeseon Ko, Young-In Song  
HIGHLIGHT: Our contribution is to reduce the size of the multi-vector representation, without compromising the effectiveness, supervised by query logs.
- 801, TITLE: ZeroGen: Efficient Zero-shot Learning via Dataset Generation  
<https://aclanthology.org/2022.emnlp-main.801>  
AUTHORS: Jiacheng Ye, Jiahui Gao, Qintong Li, Hang Xu, Jiangtao Feng, Zhiyong Wu, Tao Yu, Lingpeng Kong  
HIGHLIGHT: In this paper, we study a flexible and efficient zero-short learning method, ZeroGen.
- 802, TITLE: Neighborhood Contrastive Learning for Scientific Document Representations with Citation Embeddings  
<https://aclanthology.org/2022.emnlp-main.802>  
AUTHORS: Malte Ostendorff, Nils Rethmeier, Isabelle Augenstein, Bela Gipp, Georg Rehm  
HIGHLIGHT: Instead, we use controlled nearest neighbor sampling over citation graph embeddings for contrastive learning.
- 803, TITLE: SPE: Symmetrical Prompt Enhancement for Fact Probing  
<https://aclanthology.org/2022.emnlp-main.803>  
AUTHORS: Yiyuan Li, Tong Che, Yezhen Wang, Zhengbao Jiang, Caiming Xiong, Snigdha Chaturvedi  
HIGHLIGHT: In this work, we propose Symmetrical Prompt Enhancement (SPE), a continuous prompt-based method for factual probing in PLMs that leverages the symmetry of the task by constructing symmetrical prompts for subject and object prediction.
- 804, TITLE: Efficient Large Scale Language Modeling with Mixtures of Experts  
<https://aclanthology.org/2022.emnlp-main.804>  
AUTHORS: Mikel Artetxe, Shrutu Bhosale, Naman Goyal, Todor Mihaylov, Myle Ott, Sam Shleifer, Xi Victoria Lin, Jingfei Du, Srinivasan Iyer, Ramakanth Pasunuru, Giridharan Anantharaman, Xian Li, Shuohui Chen, Halil Akin, Mandeep Baines, Louis Martin, Xing Zhou, Punit Singh Koura, Brian O'ÅÃ½Horo, Jeffrey Wang, Luke Zettlemoyer, Mona Diab, Zornitsa Kozareva, Veselin Stoyanov  
HIGHLIGHT: This paper presents a detailed empirical study of how autoregressive MoE language models scale in comparison with dense models in a wide range of settings: in- and out-of-domain language modeling, zero- and few-shot priming, and full-shot fine-tuning.
- 805, TITLE: MedJEx: A Medical Jargon Extraction Model with WikiÅÃ½s Hyperlink Span and Contextualized Masked Language Model Score  
<https://aclanthology.org/2022.emnlp-main.805>  
AUTHORS: Sunjae Kwon, Zonghai Yao, Harmon Jordan, David Levy, Brian Corner, Hong Yu  
HIGHLIGHT: This paper proposes a new natural language processing (NLP) application for identifying medical jargon terms potentially difficult for patients to comprehend from electronic health record (EHR) notes.
- 806, TITLE: Discourse Comprehension: A Question Answering Framework to Represent Sentence Connections  
<https://aclanthology.org/2022.emnlp-main.806>  
AUTHORS: Wei-Jen Ko, Cutter Dalton, Mark Simmons, Eliza Fisher, Greg Durrett, Junyi Jessy Li  
HIGHLIGHT: A key challenge in building and evaluating models for this type of discourse comprehension is the lack of annotated data, especially since collecting answers to such questions requires high cognitive load for annotators. This paper presents a novel paradigm that enables scalable data collection targeting the comprehension of news documents, viewing these questions through the lens of discourse.
- 807, TITLE: Learning to Generate Overlap Summaries through Noisy Synthetic Data  
<https://aclanthology.org/2022.emnlp-main.807>  
AUTHORS: Naman Bansal, Mousumi Akter, Shubhra Kanti Karmaker Santu  
HIGHLIGHT: One of the major challenges for solving this task is the lack of existing datasets for supervised training. To address this challenge, we propose a novel data augmentation technique, which allows us to create large amount of synthetic data for training a seq-to-seq model that can perform the SOS task.
- 808, TITLE: Mutual Exclusivity Training and Primitive Augmentation to Induce Compositionality  
<https://aclanthology.org/2022.emnlp-main.808>  
AUTHORS: Yichen Jiang, Xiang Zhou, Mohit Bansal  
HIGHLIGHT: Recent datasets expose the lack of the systematic generalization ability in standard sequence-to-sequence models. In this work, we analyze this behavior of seq2seq models and identify two contributing factors: a lack of mutual exclusivity bias (one target sequence can only be mapped to one source sequence), and the tendency to memorize whole examples rather than separating structures from contents.
- 809, TITLE: Directions for NLP Practices Applied to Online Hate Speech Detection

<https://aclanthology.org/2022.emnlp-main.809>

**AUTHORS:** Paula Fortuna, Monica Dominguez, Leo Wanner, Zeerak Talat  
**HIGHLIGHT:** However, hate speech is a deeply complex and situated concept that eludes such static and disembodied practices. In this position paper, we critically reflect on these methodologies for hate speech detection, we argue that many conventions in NLP are poorly suited for the problem and encourage researchers to develop methods that are more appropriate for the task.

810, **TITLE:** Pre-training Transformer Models with Sentence-Level Objectives for Answer Sentence Selection

<https://aclanthology.org/2022.emnlp-main.810>

**AUTHORS:** Luca Di Liello, Siddhant Garg, Luca Soldaini, Alessandro Moschitti  
**HIGHLIGHT:** In this paper, we propose three novel sentence-level transformer pre-training objectives that incorporate paragraph-level semantics within and across documents, to improve the performance of transformers for AS2, and mitigate the requirement of large labeled datasets.

811, **TITLE:** OpenCQA: Open-ended Question Answering with Charts

<https://aclanthology.org/2022.emnlp-main.811>

**AUTHORS:** Shankar Kantharaj, Xuan Long Do, Rixie Tiffany Leong, Jia Qing Tan, Enamul Hoque, Shafiq Joty  
**HIGHLIGHT:** Answering such questions are often difficult and time-consuming as it requires a lot of cognitive and perceptual efforts. To address this challenge, we introduce a new task called OpenCQA, where the goal is to answer an open-ended question about a chart with descriptive texts.

812, **TITLE:** A Systematic Investigation of Commonsense Knowledge in Large Language Models

<https://aclanthology.org/2022.emnlp-main.812>

**AUTHORS:** Xiang Lorraine Li, Adhiguna Kuncoro, Jordan Hoffmann, Cyprien de Masson d'Aÿautume, Phil Blunsom, Aida Nematzadeh  
**HIGHLIGHT:** Language models (LMs) trained on large amounts of data have shown impressive performance on many NLP tasks under the zero-shot and few-shot setup. Here we aim to better understand the extent to which such models learn commonsense knowledge a critical component of many NLP applications.

813, **TITLE:** Transforming Sequence Tagging Into A Seq2Seq Task

<https://aclanthology.org/2022.emnlp-main.813>

**AUTHORS:** Karthik Raman, Iftekhar Naim, Jiecao Chen, Kazuma Hashimoto, Kiran Yalasangi, Krishna Srinivasan  
**HIGHLIGHT:** In this paper, we rigorously study different formats one could use for casting input text sentences and their output labels into the input and target (i. e. , output) of a Seq2Seq model.

814, **TITLE:** CycleKQR: Unsupervised Bidirectional Keyword-Question Rewriting

<https://aclanthology.org/2022.emnlp-main.814>

**AUTHORS:** Andrea Iovine, Anjie Fang, Besnik Fetahu, Jie Zhao, Oleg Rokhlenko, Shervin Malmasi  
**HIGHLIGHT:** We propose the keyword-question rewriting task to improve query understanding capabilities of NLU systems for all surface forms. To achieve this, we present CycleKQR, an unsupervised approach, enabling effective rewriting between keyword and question queries using non-parallel data.

815, **TITLE:** Model Criticism for Long-Form Text Generation

<https://aclanthology.org/2022.emnlp-main.815>

**AUTHORS:** Yuntian Deng, Volodymyr Kuleshov, Alexander Rush  
**HIGHLIGHT:** Here, we propose to apply a statistical tool, model criticism in latent space, to evaluate the high-level structure of the generated text.

816, **TITLE:** Improving Faithfulness by Augmenting Negative Summaries from Fake Documents

<https://aclanthology.org/2022.emnlp-main.816>

**AUTHORS:** Tianshu Wang, Faisal Ladhak, Esin Durmus, He He  
**HIGHLIGHT:** However, the commonly used maximum likelihood training does not disentangle factual errors from other model errors. To address this issue, we propose a back-translation-style approach to augment negative samples that mimic factual errors made by the model.

817, **TITLE:** Joint Completion and Alignment of Multilingual Knowledge Graphs

<https://aclanthology.org/2022.emnlp-main.817>

**AUTHORS:** Soumen Chakrabarti, Harkanwar Singh, Shubham Lohiya, Prachi Jain, Mausam -  
**HIGHLIGHT:** Many effective algorithms have been proposed for completion and alignment as separate tasks. Here we show that these tasks are synergistic and best solved together.

818, **TITLE:** Offer a Different Perspective: Modeling the Belief Alignment of Arguments in Multi-party Debates

<https://aclanthology.org/2022.emnlp-main.818>

**AUTHORS:** Suzanna Sia, Kokil Jaidka, Hansin Ahuja, Niyati Chhaya, Kevin Duh  
**HIGHLIGHT:** We adopt a hierarchical generative Variational Autoencoder as our model and impose structural constraints that reflect competing hypotheses about the nature of argumentation.

819, **TITLE:** A Federated Approach to Predicting Emojis in Hindi Tweets

<https://aclanthology.org/2022.emnlp-main.819>

**AUTHORS:** Deep Gandhi, Jash Mehta, Nirali Parekh, Karan Waghela, Lynette D'Aÿmello, Zeerak Talat

**HIGHLIGHT:** In this paper, we seek to address the dual concerns of emphasising high resource languages for emoji prediction and risking the privacy of people's data.

**820, TITLE:** Injecting Domain Knowledge in Language Models for Task-oriented Dialogue Systems  
<https://aclanthology.org/2022.emnlp-main.820>

**AUTHORS:** Denis Emelin, Daniele Bonadiman, Sawsan Alqahtani, Yi Zhang, Saab Mansour

**HIGHLIGHT:** In this paper, we showcase the advantages of injecting domain-specific knowledge prior to fine-tuning on TOD tasks.

**821, TITLE:** TASA: Deceiving Question Answering Models by Twin Answer Sentences Attack  
<https://aclanthology.org/2022.emnlp-main.821>

**AUTHORS:** Yu Cao, Dianqi Li, Meng Fang, Tianyi Zhou, Jun Gao, Yibing Zhan, Dacheng Tao

**HIGHLIGHT:** We present Twin Answer Sentences Attack (TASA), an adversarial attack method for question answering (QA) models that produces fluent and grammatical adversarial contexts while maintaining gold answers.

**822, TITLE:** Improving Low-Resource Languages in Pre-Trained Multilingual Language Models  
<https://aclanthology.org/2022.emnlp-main.822>

**AUTHORS:** Viktor Hangya, Hossain Shaikh Saadi, Alexander Fraser

**HIGHLIGHT:** However, languages with small available monolingual corpora are often not well-supported by these models leading to poor performance. We propose an unsupervised approach to improve the cross-lingual representations of low-resource languages by bootstrapping word translation pairs from monolingual corpora and using them to improve language alignment in pre-trained language models.

**823, TITLE:** SCROLLS: Standardized Comparison Over Long Language Sequences  
<https://aclanthology.org/2022.emnlp-main.823>

**AUTHORS:** Uri Shaham, Elad Segal, Maor Ivgi, Avia Efrat, Ori Yoran, Adi Haviv, Ankit Gupta, Wenhan Xiong, Mor Geva, Jonathan Berant, Omer Levy

**HIGHLIGHT:** We introduce SCROLLS, a suite of tasks that require reasoning over long texts.

**824, TITLE:** PAR: Political Actor Representation Learning with Social Context and Expert Knowledge  
<https://aclanthology.org/2022.emnlp-main.824>

**AUTHORS:** Shangbin Feng, Zhaoxuan Tan, Zilong Chen, Ningnan Wang, Peisheng Yu, Qinghua Zheng, Xiaojun Chang, Minnan Luo

**HIGHLIGHT:** In this paper, we propose PAR, a Political Actor Representation learning framework that jointly leverages social context and expert knowledge.

**825, TITLE:** JDDC 2.1: A Multimodal Chinese Dialogue Dataset with Joint Tasks of Query Rewriting, Response Generation, Discourse Parsing, and Summarization  
<https://aclanthology.org/2022.emnlp-main.825>

**AUTHORS:** Nan Zhao, Haoran Li, Youzheng Wu, Xiaodong He

**HIGHLIGHT:** In this paper, we construct JDDC 2.1, a large-scale multimodal multi-turn dialogue dataset collected from a mainstream Chinese E-commerce platform, containing about 246K dialogue sessions, 3M utterances, and 507K images, along with product knowledge bases and image category annotations.

**826, TITLE:** PCL: Peer-Contrastive Learning with Diverse Augmentations for Unsupervised Sentence Embeddings  
<https://aclanthology.org/2022.emnlp-main.826>

**AUTHORS:** Qiyu Wu, Chongyang Tao, Tao Shen, Can Xu, Xiubo Geng, Daxin Jiang

**HIGHLIGHT:** As one answer, we propose a novel Peer-Contrastive Learning (PCL) with diverse augmentations.

**827, TITLE:** Digging Errors in NMT: Evaluating and Understanding Model Errors from Partial Hypothesis Space  
<https://aclanthology.org/2022.emnlp-main.827>

**AUTHORS:** Jianhao Yan, Chenming Wu, Fandong Meng, Jie Zhou

**HIGHLIGHT:** To tackle the problem of exponentially large space, we propose two approximation methods, top region evaluation along with an exact top-k decoding algorithm, which finds top-ranked hypotheses in the whole hypothesis space, and Monte Carlo sampling evaluation, which simulates hypothesis space from a broader perspective.

**828, TITLE:** DialogConv: A Lightweight Fully Convolutional Network for Multi-view Response Selection  
<https://aclanthology.org/2022.emnlp-main.828>

**AUTHORS:** Yongkang Liu, Shi Feng, Wei Gao, Daling Wang, Yifei Zhang

**HIGHLIGHT:** In this paper, we propose a novel lightweight fully convolutional architecture, called DialogConv, for response selection.