Indirectly Supervised Natural Language Processing

Wenpeng Yin[†], Muhao Chen[‡], Ben Zhou[◇], Qiang Ning[⋆], Kai-Wei Chang[♯], Dan Roth[◇⋆]
[†]Penn State; [‡]USC; [◇]UPenn; [⋆]AWS AI Labs; [♯]UCLA
wenpeng@psu.edu; muhaoche@usc.edu
{xyzhou,danroth}@seas.upenn.edu
qning@amazon.com; kwchang@cs.ucla.edu

Abstract

This tutorial targets researchers and practitioners who are interested in ML technologies for NLP from indirect supervision. In particular, we will present a diverse thread of indirect supervision studies that try to answer the following questions: (i) when and how can we provide supervision for a target task T, if all we have is data that corresponds to a "related" task T'? (ii) humans do not use *exhaustive* supervision; they rely on occasional feedback, and learn from incidental signals from various sources; how can we effectively incorporate such supervision in machine learning? (iii) how can we leverage multi-modal supervision to help NLP? To the end, we will discuss several lines of research that address those challenges, including (i) indirect supervision from T' that handles T with outputs spanning from a moderate size to an open space, (ii) the use of sparsely occurring and incidental signals, such as partial labels, noisy labels, knowledge-based constraints, and cross-domain or cross-task annotations—all having statistical associations with the task, (iii) principled ways to measure and understand why these incidental signals can contribute to our target tasks, and (iv) indirect supervision from vision-language signals. We will conclude the tutorial by outlining directions for further investigation.

1 Introduction

Conventional approaches to NLP rely on task-specific labeled examples of a large volume. This does not apply to scenarios where tasks may be too complicated or costly to annotate, or the system is required to handle a new task immediately. Many people increasingly perceive that pretrained language models (PLMs) use self-supervision, and therefore there is no need for supervision anymore. While this is probably true for Encoder-only models (e.g., BERT (Devlin et al., 2019)), this does not hold for Decoder models, where people nowadays use vast amounts of supervision and reinforcement learning signals. Therefore, it is still desirable to gather *supervision that has already existed in related tasks or is pretty cheap*, which is termed "indirect supervision" in this tutorial.

Recently, there have been increasing works that study indirect supervision for a wide range of NLP tasks. For example, Yin et al. (2019) and Lu et al. (2022a) respectively leveraged the rich annotation of a source task (natural language inference or summarization) to address the poorly-annotated target tasks. To make better use of the natural texts, some literature (Roth, 2017; Chen et al., 2021; He et al., 2021) proposed to explore incidental supervision, e.g., phonetic similarity and similar temporal distribution for named entity transliteration, to help downstream tasks. That sort of incidental supervision is often weak signals that exist in the data and the environment independently of the tasks at hand, and is hard to be encoded by PLMs. Furthermore, when accessing supervision from pure text is challenging, researchers turned to other modalities for indirect supervision (Li et al., 2022b).

This tutorial presents a comprehensive introduction of those lines of frontier research on indirectly supervised NLP. In particular, it tries to answer the following questions: (i) Which source task is easier to be adapted to solve various target tasks and any constraints there? (ii) What are the limitations of pretrained language models in discovering supervision from natural texts, and how can we alleviate them with incidental signals? (iii) Are there any theoretical measures that can indicate the benefits of the incidental signals to a given downstream task? (iv) How to mitigate the gap between different modalities if we want to utilize image/video knowledge to guide NLP? By addressing those critical questions, we believe it is necessary to present a timely tutorial to comprehensively summarize the new frontiers in indirectly supervised NLP research and point out the emerging challenges that deserve further investigation. Participants will learn about recent trends and emerging challenges in this topic, representative tools and learning resources to obtain ready-to-use models, and how related technologies benefit end-user NLP applications.

2 Outline of Tutorial Content

This **half-day** tutorial presents a systematic overview of recent advancements in indirect supervision methods for NLP. The detailed contents are outlined below.

2.1 Background and Motivation [15min]

We will begin motivating this topic with a selection of real-world applications and emerging challenges of NLP with limited end-task annotations.

2.2 Indirect Supervision from NLU Tasks [30min]

We start with indirect supervision from a source task that is efficient to handle a moderate size of outputs in the target task. For example, in most zero/few-shot text classification tasks, such as topic classification, entity typing, relation identification, etc., the main obstacle is letting systems understand the semantics of labels. In contrast to conventional supervised classifiers, which converted labels into indices, we introduce NLI (natural language inference)-based approaches that take into account the input semantics as well as label semantics. In specific, we will introduce typical work that treats different topics (Yin et al., 2019), stances (Xu et al., 2022), entity types (Li et al., 2022a; Du et al., 2023), event types (Lyu et al., 2021), entity relations (Xia et al., 2021; Sainz et al., 2021, 2022), and question-answer (Yin et al., 2021) as hypotheses and the inputs as premises, then makes use of pretrained NLI system to handle a variety of classification tasks with a given set of labels.

In addition, we will present extractive question answering (Ex-QA) based supervision that is utilized for downstream tasks (McCann et al., 2018; Keskar et al., 2019; He et al., 2020; Wu et al., 2020; Li et al., 2020). The advantage of Ex-QA based indirect supervision over the NLI-based one lies in that Ex-QA can handle sequence tagging and span detection tasks while NLI-based approaches primarily work for classification.

2.3 Indirect Supervision from NLG and IR [30min]

We will introduce methodologies that acquire indirect supervision signals from natural language generation (NLG) and information retrieval tasks to solve more low-resource discriminative tasks. Formulating discriminative tasks as generation tasks can be an efficient way to guide PLMs to leverage the semantics of decision labels (Huang et al., 2021; Lu et al., 2022a; Hsu et al., 2022; Yuan et al., 2022). A method of this kind typically leads to a sequence-to-sequence generation process that emits a verbalization of the decision label given the input sequence (Zeng et al., 2018, 2020; Ye et al., 2021; Cao and Ananiadou, 2021). Instead of predicting classification logits, these models represent the class as a concise structure and employ controlled decoding for the generation. In this way, the model allows cross-task signal transfer from high-resource NLG tasks, and captures a semantically rich representation of the discriminative task's original decision space. A representative example is SuRE (Lu et al., 2022a), which reformulates the more expensive relation extraction task into summarization with constrained decoding, leading to more precise and label-efficient sentence-level relation extraction. We will also introduce methods that reformulate as a retrieval task (Zhang et al., 2021a,b; Huang et al., 2022; Chen et al., 2020). This technique allows using the inductive bias of a dense retrieval model to handle a discriminative task with a large decision space, such as entity linking (Zhang et al., 2021a) and fine-grained typing (Huang et al., 2022).

2.4 Incidental Supervision from Natural Text [30min]

Both the indirect supervision introduced in the above sections (§2.2-§2.3) relies on transferred supervision signals from some source task annotations. Natural texts are structured to contain a large number of incidental signals that can be subsequently utilized by downstream tasks with minimal human effort. Despite the fact that the community has found that PLMs are capable of providing incidental supervision signals for a wide range of tasks, they do not provide controls over what kinds of knowledge exist. To the end, we introduce incidental relations found in natural text spans. For example, certain keywords and linguistic patterns can provide incidental supervision to downstream tasks such as relation extraction (Zhou et al., 2022b), temporal reasoning (Zhou et al., 2020, 2021), and affordance reasoning (Qasemi et al., 2022). Moreover, textual snippets can often be viewed in a structure by their global information, such as publication dates, titles, and authors, which establish relations that helps with complex tasks (Zhou et al., 2022a). Designing and collecting such linguistic patterns often require human knowledge; this process of injecting human knowledge provides signals that PLMs cannot find and produces diverse automatic supervision for many tasks.

2.5 Theoretical Analysis of Incidental Supervision [30min]

§2.4 presents several real-world applications of incidental signals. In this part, we pose the challenge to define a principled way to measure the benefits of these signals to a given downstream task, and the challenge to further understand why and how these signals can help reduce the complexity of the learning problem in theory. We will introduce existing efforts along these two lines, mainly He et al. (2021) and Wang et al. (2020). Specifically, we introduce (i) a unified theoretical framework (Wang et al., 2020) for multi-class classification when the supervision is provided by a variable that contains nonzero mutual information with the gold label; the nature of this problem is determined by the transition probability from the gold labels to the indirect supervision variables (van Rooyen and Williamson, 2018) and the learner's prior knowledge about the transition; and (ii) a unified PAC-Bayesian motivated informativeness measure, PABI (He et al., 2021), that characterizes the uncertainty reduction provided by incidental supervision signals. We share studies in Qasemi et al. (2022) and Ning et al. (2019) that demonstrate PABI's effectiveness by quantifying the value added by various types of incidental signals to sequence tagging tasks. Finally, we will highlight the gaps that are yet to be closed in these lines, and point out future research directions on this topic.

2.6 Indirect Supervision from Multi-modalities [30min]

In the previous section, we discuss how to leverage indirect supervision from text data. Next, we will extend our discussion to introduce methods that leverage indirect supervision in multimodal data for cross-modality tasks. We will take vision-language tasks, such as answering complex high-level question about images (Zellers et al., 2019), as an example. We will first introduce methods that learn to align visual tokens and text tokens based on image caption data (Tan and Bansal, 2019; Li et al., 2019; Tan and Bansal, 2020). The cross-modality knowledge learned from indirect supervision can be used to solve various text, image, and mixed modality tasks. We will then introduce approaches that use only indirect supervision from object recognition models to learn text-image alignment from unaligned language and vision Data (Li et al., 2021). Finally, we will discuss methods for learning to ground elements of language to image regions without explicit supervision (Li et al., 2022b; Zhang et al., 2022).

2.7 Future Research Directions [15min]

Indirect supervision is the key to coping with a variety of NLP tasks that are not equipped with enough labeled data. We will conclude the tutorial by presenting further challenges and potential research topics, such as (i) explaining the model predictions when the supervision is indirect (Rajani et al., 2020; Lu et al., 2022b), (ii) injecting incidental signals that express human knowledge but cannot be learned by pretrained language models from natural texts (Yu et al., 2022), and (iii) task instructions as supervision (Wang et al., 2022).

3 Specification of the Tutorial

The proposed tutorial is considered a **cutting-edge** tutorial that introduces new frontiers in indirectly supervised NLP. The presented topic has not been covered by any *CL tutorials in the past 4 years.

Audience and Prerequisites Based on the level of interest in this topic, we expect around 150 participants. While no specific background knowledge is assumed of the audience, it would be best for the attendees to know about basic deep learning technologies, pre-trained language models (e.g. BERT). A reading list that could help provide background knowledge to the audience before attending this tutorial is given in Appx. §A.2.

Breadth We estimate that at least 60% of the work covered in this tutorial is from researchers other than the instructors of the tutorial.

Diversity Considerations This tutorial will cover indirect supervision from beyond text. We will also cover content around how indirect supervision can be applicable to a variety of low-resourced tasks. Our presenter team has a diverse background from both academia (including assistant, associate, distinguished professors, and a senior Ph.D. student) and industry (a senior scientist at AWS AI).

Our instructor team will promote our tutorial on social media to diversify our audience participation.

Material Access Online Open Access All the materials are openly available at https://cogcomp.seas.upenn.edu/page/tutorial.202307

4 Tutorial Instructors

The following are biographies of the speakers. Past tutorials given by us are listed in Appx. §A.1.

Wenpeng Yin is an Assistant Professor in the Department of Computer Science and Engineering at Penn State University. Prior to joining Penn State, he was a tenure-track faculty member at Temple University (1/2022-12/2022), Senior Research Scientist at Salesforce Research (8/2019-12/2021), a postdoctoral researcher at UPenn (10/2017-7/2019), and got his Ph.D. degree from the Ludwig Maximilian University of Munich, Germany, in 2017. Dr. Yin's research focuses on natural language processing with three sub-areas: (i) learning from task instructions; (ii) information extraction; (iii) learning with limited supervision. Additional information is available at www.wenpengyin.org.

Muhao Chen is an Assistant Research Professor of Computer Science at USC, where he directs the Language Understanding and Knowledge Acquisition (LUKA) Group. His research focuses on data-driven machine learning approaches for natural language understanding and knowledge acquisition. His work has been recognized with an NSF CRII Award, a Cisco Faculty Research Award, an ACM SIGBio Best Student Paper Award, and a Best Paper Nomination at CoNLL. Muhao obtained his PhD degree from UCLA Department of Computer Science in 2019, and was a postdoctoral researcher at UPenn prior to joining USC. Additional information is available at http://luka-group.github.io.

Ben Zhou is a fourth-year Ph.D. student at the Department of Computer and Information Science, University of Pennsylvania. Ben's research interests are distant supervision extraction and experiential knowledge reasoning, and he has more than 5 recent papers on related topics. He is a recipient of the ENIAC fellowship from the University of Pennsylvania, and a finalist of the CRA outstanding undergraduate researcher award. Additional information is available at http://xuanyu.me/.

Qiang Ning is currently a senior applied scientist at AWS AI (2022-). Prior to that, Qiang was an applied scientist at Alexa AI (2020-2022) and a research scientist at the Allen Institute for AI (2019-2020). Qiang received his Ph.D. from the University of Illinois at Urbana-Champaign in 2019 in Electrical and Computer Engineering. Qiang's research interests span in information extraction, question answering, and the application of weak supervision methods in these NLP problems in both theoretical and practical aspects. Additional information is available at https://www.qiangning.info/.

Kai-Wei Chang is an associate professor in the Department of Computer Science at the University of California Los Angeles. His research interests include designing robust, fair, and accountable machine learning methods for building reliable NLP systems. His awards include the EMNLP Best Long Paper Award (2017), the KDD Best Paper Award (2010), and the Sloan Resaerch Fellowship (2021). Kai-Wei has given tutorials at NAACL 15, AAAI 16, FAccT18, EMNLP 19, AAAI 20, EMNLP 21, MLSS 21 on different research topics. Additional information is available at http://kwchang.net.

Dan Roth is the Eduardo D. Glandt Distinguished Professor at the Department of Computer and Information Science, UPenn, the NLP Lead at AWS AI Labs, and a Fellow of the AAAS, ACM, AAAI, and ACL. In 2017 Roth was awarded the John McCarthy Award, the highest award the AI community gives to mid-career AI researchers. Roth was recognized "for major conceptual and theoretical advances in the modeling of natural language understanding, machine learning, and reasoning." Roth has published broadly in machine learning, NLP, KRR, and learning theory, and has given keynote talks and tutorials in all ACL and AAAI major conferences. Roth was the Editor-in-Chief of JAIR until 2017, and was the program chair of AAAI'11, ACL'03 and CoNLL'02; he serves regularly as an area chair and senior program committee member in the major conferences in his research areas. Additional information is available at www.cis.upenn.edu/~danroth.

Acknowledgement

This presenters' research is supported in part by Contract FA8750-19-2-1004 with the US Defense Advanced Research Projects Agency (DARPA), the DARPA MCS program under Contract No. N66001-19-2-4033 with the United States Office Of Naval Research, Intelligence Advanced Research Projects Activity (IARPA) Contract No. 2019-19051600006 under the BETTER Program, the National Science Foundation (NSF) of United States Grant IIS 2105329, a subaward from NSF Cloudbank 1925001 through UCSD, an Amazon Research Award and a Cisco Research Award. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of DARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein,

Ethical Considerations

We do not anticipate any ethical issues particularly to the topics of the tutorial. Nevertheless, some work presented in this tutorial extensively uses large-scale pretrained models with self-attention, which may lead to substantial financial and environmental costs.

References

Jiarun Cao and Sophia Ananiadou. 2021. GenerativeRE: Incorporating a novel copy mechanism and pretrained model for joint entity and relation extraction. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 2119–2126, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Muhao Chen, Weijia Shi, Ben Zhou, and Dan Roth. 2021. Cross-lingual entity alignment with incidental supervision. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, EACL 2021, Online, April 19 - 23, 2021*, pages 645–658. Association for Computational Linguistics.

Muhao Chen, Hongming Zhang, Haoyu Wang, and Dan Roth. 2020. What are you trying to do? semantic typing of event processes. In *Proceedings of the 24th Conference on Computational Natural Language Learning*, pages 531–542, Online. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 4171–4186.

Jiangshu Du, Wenpeng Yin, Congying Xia, and Philip S. Yu. 2023. Learning to select from multiple options. In *AAAI*.

Hangfeng He, Qiang Ning, and Dan Roth. 2020. Quase: Question-answer driven sentence encoding. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 8743–8758.

Hangfeng He, Mingyuan Zhang, Qiang Ning, and Dan Roth. 2021. Foreseeing the Benefits of Incidental Supervision. In *Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

I-Hung Hsu, Kuan-Hao Huang, Elizabeth Boschee, Scott Miller, Prem Natarajan, Kai-Wei Chang, and Nanyun Peng. 2022. DEGREE: A data-efficient generation-based event extraction model. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1890–1908, Seattle, United States. Association for Computational Linguistics.

James Y. Huang, Bangzheng Li, Jiashu Xu, and Muhao Chen. 2022. Unified semantic typing with meaningful label inference. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2642–2654, Seattle, United States. Association for Computational Linguistics.

Kung-Hsiang Huang, Sam Tang, and Nanyun Peng. 2021. Document-level entity-based extraction as template generation. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 5257–5269, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Nitish Shirish Keskar, Bryan McCann, Caiming Xiong, and Richard Socher. 2019. Unifying question answering and text classification via span extraction. *CoRR*, abs/1904.09286.

Bangzheng Li, Wenpeng Yin, and Muhao Chen. 2022a. Ultra-fine entity typing with indirect supervision from natural language inference. *Transactions of the Association for Computational Linguistics*, 10:607–622.

Liunian Harold Li, Mark Yatskar, Da Yin, Cho-Jui Hsieh, and Kai-Wei Chang. 2019. Visualbert: A simple and performant baseline for vision and language. In *Arxiv*.

Liunian Harold Li, Haoxuan You, Zhecan Wang, Alireza Zareian, Shih-Fu Chang, and Kai-Wei Chang. 2021. Unsupervised vision-and-language pre-training without parallel images and captions. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5339–5350, Online. Association for Computational Linguistics.

Liunian Harold Li, Pengchuan Zhang, Haotian Zhang, Jianwei Yang, Chunyuan Li, Yiwu Zhong, Lijuan Wang, Lu Yuan, Lei Zhang, Jenq-Neng Hwang, Kai-Wei Chang, and Jianfeng Gao. 2022b. Grounded language-image pre-training. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, June 18-24, 2022*, pages 10955–10965. IEEE.

Xiaoya Li, Jingrong Feng, Yuxian Meng, Qinghong Han, Fei Wu, and Jiwei Li. 2020. A unified MRC framework for named entity recognition. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5849–5859, Online. Association for Computational Linguistics.

Keming Lu, I-Hung Hsu, Mingyu Derek Ma, Wenxuan Zhou, and Muhao Chen. 2022a. Summarization as indirect supervision for relation extraction. In *EMNLP - Findings*.

Pan Lu, Swaroop Mishra, Tony Xia, Liang Qiu, Kai-Wei Chang, Song-Chun Zhu, Oyvind Tafjord, Peter Clark, and Ashwin Kalyan. 2022b. Learn to explain: Multimodal reasoning via thought chains for science question answering. In *NeurIPS*.

Qing Lyu, Hongming Zhang, Elior Sulem, and Dan Roth. 2021. Zero-shot event extraction via transfer learning: Challenges and insights. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 322–332, Online. Association for Computational Linguistics.

Bryan McCann, Nitish Shirish Keskar, Caiming Xiong, and Richard Socher. 2018. The natural language decathlon: Multitask learning as question answering. *CoRR*, abs/1806.08730.

Qiang Ning, Hangfeng He, Chuchu Fan, and Dan Roth. 2019. Partial or Complete, That's The Question. In *Proc. of the Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*.

Ehsan Qasemi, Piyush Khanna, Qiang Ning, and Muhao Chen. 2022. PInKS: Preconditioned commonsense inference with minimal supervision. In *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 320–336, Online only. Association for Computational Linguistics.

Nazneen Fatema Rajani, Ben Krause, Wengpeng Yin, Tong Niu, Richard Socher, and Caiming Xiong. 2020. Explaining and improving model behavior with k nearest neighbor representations. *CoRR*, abs/2010.09030.

Dan Roth. 2017. Incidental supervision: Moving beyond supervised learning. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA*, pages 4885–4890.

Oscar Sainz, Oier Lopez de Lacalle, Gorka Labaka, Ander Barrena, and Eneko Agirre. 2021. Label verbalization and entailment for effective zero and few-shot relation extraction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, pages 1199–1212.

Oscar Sainz, Itziar Gonzalez-Dios, Oier Lopez de Lacalle, Bonan Min, and Eneko Agirre. 2022. Textual entailment for event argument extraction: Zero- and few-shot with multi-source learning. In *Findings of the Association for Computational Linguistics: NAACL 2022, Seattle, WA, United States, July 10-15, 2022*, pages 2439–2455.

Hao Tan and Mohit Bansal. 2019. LXMERT: Learning cross-modality encoder representations from transformers. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 5100–5111, Hong Kong, China. Association for Computational Linguistics.

Hao Tan and Mohit Bansal. 2020. Vokenization: Improving language understanding with contextualized, visual-grounded supervision. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2066–2080, Online. Association for Computational Linguistics.

Brendan van Rooyen and Robert C. Williamson. 2018. A Theory of Learning with Corrupted Labels. *Journal of Machine Learning Research*, 18(228):1–50.

Kaifu Wang, Qiang Ning, and Dan Roth. 2020. Learnability with Indirect Supervision Signals. In *Proc. of the Conference on Neural Information Processing Systems (NeurIPS)*.

Yizhong Wang, Swaroop Mishra, Pegah Alipoormolabashi, Yeganeh Kordi, Amirreza Mirzaei, Anjana Arunkumar, Arjun Ashok, Arut Selvan Dhanasekaran, Atharva Naik, David Stap, Eshaan Pathak, Giannis Karamanolakis, Haizhi Gary Lai, Ishan Purohit, Ishani Mondal, Jacob Anderson, Kirby Kuznia, Krima Doshi, Maitreya Patel, Kuntal Kumar Pal, Mehrad Moradshahi, Mihir Parmar, Mirali Purohit, Neeraj Varshney, Phani Rohitha Kaza, Pulkit Verma, Ravsehaj Singh Puri, Rushang Karia, Shailaja Keyur Sampat, Savan Doshi, Siddhartha

Mishra, Sujan Reddy, Sumanta Patro, Tanay Dixit, Xudong Shen, Chitta Baral, Yejin Choi, Hannaneh Hajishirzi, Noah A. Smith, and Daniel Khashabi. 2022. Benchmarking generalization via in-context instructions on 1, 600+ language tasks. *CoRR*, abs/2204.07705.

Wei Wu, Fei Wang, Arianna Yuan, Fei Wu, and Jiwei Li. 2020. CorefQA: Coreference resolution as query-based span prediction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6953–6963, Online. Association for Computational Linguistics.

Congying Xia, Wenpeng Yin, Yihao Feng, and Philip S. Yu. 2021. Incremental few-shot text classification with multi-round new classes: Formulation, dataset and system. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021*, pages 1351–1360.

Hanzi Xu, Slobodan Vucetic, and Wenpeng Yin. 2022. Openstance: Real-world zero-shot stance detection. volume Proceedings of CoNLL.

Hongbin Ye, Ningyu Zhang, Shumin Deng, Mosha Chen, Chuanqi Tan, Fei Huang, and Huajun Chen. 2021. Contrastive triple extraction with generative transformer. In *AAAI*.

Wenpeng Yin, Jamaal Hay, and Dan Roth. 2019. Benchmarking zero-shot text classification: Datasets, evaluation and entailment approach. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019, Hong Kong, China, November 3-7, 2019*, pages 3912–3921.

Wenpeng Yin, Dragomir R. Radev, and Caiming Xiong. 2021. Docnli: A large-scale dataset for document-level natural language inference. In *Findings of ACL/IJCNLP*, pages 4913–4922.

Donghan Yu, Chenguang Zhu, Yiming Yang, and Michael Zeng. 2022. Jaket: Joint pre-training of knowledge graph and language understanding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 11630–11638.

Hongyi Yuan, Zheng Yuan, and Sheng Yu. 2022. Generative biomedical entity linking via knowledge base-guided pre-training and synonyms-aware fine-tuning. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4038–4048, Seattle, United States. Association for Computational Linguistics.

Rowan Zellers, Yonatan Bisk, Ali Farhadi, and Yejin Choi. 2019. From recognition to cognition: Visual commonsense reasoning. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Daojian Zeng, Haoran Zhang, and Qianying Liu. 2020. Copymtl: Copy mechanism for joint extraction of entities and relations with multi-task learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 9507–9514.

Xiangrong Zeng, Daojian Zeng, Shizhu He, Kang Liu, and Jun Zhao. 2018. Extracting relational facts by an end-to-end neural model with copy mechanism. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 506–514, Melbourne, Australia. Association for Computational Linguistics.

Haotian* Zhang, Pengchuan* Zhang, Xiaowei Hu, Yen-Chun Chen, Liunian Harold Li, Xiyang Dai, Lijuan Wang, Lu Yuan, Jenq-Neng Hwang, and Jianfeng Gao. 2022. Glipv2: Unifying localization and vision-language understanding. *arXiv preprint arXiv:2206.05836*.

Wenzheng Zhang, Wenyue Hua, and Karl Stratos. 2021a. Entqa: Entity linking as question answering. In *International Conference on Learning Representations*.

Yue Zhang, Hongliang Fei, and Ping Li. 2021b. Readsre: Retrieval-augmented distantly supervised relation extraction. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2257–2262.

Ben Zhou, Qiang Ning, Daniel Khashabi, and Dan Roth. 2020. Temporal Common Sense Acquisition with Minimal Supervision. In *Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL)*.

Ben Zhou, Kyle Richardson, Qiang Ning, Tushar Khot, Ashish Sabharwal, and Dan Roth. 2021. Temporal reasoning on implicit events from distant supervision. *NAACL*.

Ben Zhou, Kyle Richardson, Xiaodong Yu, and Dan Roth. 2022a. Learning to decompose: Hypothetical question decomposition based on comparable texts. *EMNLP*.

Ben Zhou, Dian Yu, Dong Yu, and Dan Roth. 2022b. Cross-lingual speaker identification using distant supervision. *Arxiv*.

A Appendix

A.1 Past Tutorials by the Instructors

The presenters of this tutorial have given the following tutorials at leading international conferences in the past.

- Muhao Chen:
  - NAACL'22: New Frontiers of Information Extraction.

- ACL'21: Event-Centric Natural Language Processing.
- AAAI'21: Event-Centric Natural Language Understanding.
- KDD'21: From Tables to Knowledge: Recent Advances in Table Understanding.
- AAAI'20: Recent Advances of Transferable Representation Learning.

• Qiang Ning:

- ACL'21: Event-Centric Natural Language Processing.
- AAAI'21: Event-Centric Natural Language Understanding.

• Ben Zhou:

- NAACL'22: New Frontiers of Information Extraction

• Kai-Wei Chang:

- EMNLP'21: Robustness and Adversarial Examples in Natural Language Processing
- AAAI'20: Recent Advances of Transferable Representation Learning.
- EMNLP '19: A tutorial on Bias and Fairness in Natural Language Processing.
- ACM FAT*'18: A tutorial on Quantifying and Reducing Gender Stereotypes in Word Embeddings.
- TAAI'17: A tutorial on Structured Predictions: Practical Advancements and Applications in Natural Language Processing.
- AAAI'16: A tutorial on Learning and Inference in Structured Prediction Models.
- NAACL'15: A tutorial on Hands-on Learning to Search for Structured Prediction.

• Dan Roth:

- NAACL'22: New Frontiers of Information Extraction.
- ACL'21: Event-Centric Natural Language Processing.
- AAAI'21: Event-Centric Natural Language Understanding.
- ACL'20: Commonsense Reasoning for Natural Language Processing.
- AAAI'20: Recent Advances of Transferable Representation Learning.
- ACL'18: A tutorial on Multi-lingual Entity Discovery and Linking.
- EACL'17: A tutorial on Integer Linear Programming Formulations in Natural Language Processing.
- AAAI'16: A tutorial on Structured Prediction.
- ACL'14: A tutorial on Wikification and Entity Linking.
- AAAI'13: Information Trustworthiness.
- COLING'12: A Tutorial on Temporal Information Extraction and Shallow Temporal Reasoning.
- NAACL'12: A Tutorial on Constrained Conditional Models: Structured Predictions in NLP.
- NAACL'10: A Tutorial on Integer Linear Programming Methods in NLP.
- EACL'09: A Tutorial on Constrained Conditional Models.
- ACL'07: A Tutorial on Textual Entailment.

A.2 Recommended Paper List

The following is a reading list that could help provide background knowledge to the audience before attending this tutorial:

- Wenpeng Yin, Jamaal Hay, Dan Roth. Benchmarking Zero-shot Text Classification: Datasets, Evaluation and Entailment Approach. EMNLP 2019.
- Oscar Sainz, Itziar Gonzalez-Dios, Oier Lopez de Lacalle, Bonan Min, Eneko Agirre. Textual Entailment for Event Argument Extraction: Zero- and Few-Shot with Multi-Source Learning. Findings of NAACL 2022.
- Wenzheng Zhang, Wenyue Hua, Karl Stratos. EntQA: Entity Linking as Question Answering. ICLR 2022.
- Keming Lu, I-Hung Hsu, Wenxuan Zhou, Mingyu Derek Ma, Muhao Chen. Summarization as Indirect Supervision for Relation Extraction. EMNLP - Findings, 2022.
- Sarah Wiegreffe, Jack Hessel, Swabha Swayamdipta, Mark O. Riedl, Yejin Choi. Reframing human-AI collaboration for generating free-text explanations. NAACL, 2022.
- Ben Zhou, Kyle Richardson, Xiaodong Yu, Dan Roth. Learning to decompose: Hypothetical question decomposition based on comparable texts. EMNLP, 2022.
- Hangfeng He, Mingyuan Zhang, Qiang Ning, and Dan Roth. Foreseeing the Benefits of Incidental Supervision. EMNLP 2021.
- Kaifu Wang, Qiang Ning, and Dan Roth. Learnability with Indirect Supervision Signals. NeurIPS 2020.

- Rowan Zellers, Yonatan Bisk, Ali Farhadi, and Yejin Choi. From recognition to cognition: Visual commonsense reasoning. CVPR 2019.
- Hao Tan and Mohit Bansal. Vokenization: Improving language understanding with contextualized, visual-grounded supervision. EMNLP 2020.