

EvoNLP 2022

The First Workshop on Ever Evolving NLP

Proceedings of the Workshop

December 7, 2022

The EvoNLP organizers gratefully acknowledge the support from the following sponsors.

Sponsored by

Snap Inc.

©2022 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-959429-09-8

Preface by the Workshop Organizers

We are excited to welcome you to EvoNLP 2022, the 1st Workshop on Ever Evolving NLP. This workshop follows a hybrid format, and is being held on December 7, 2022, co-located with EMNLP 2022, which will also follow a hybrid format (Abu Dhabi and remote).

EvoNLP is a forum to discuss the challenges posed by the dynamic nature of language in the specific context of the current NLP paradigm, dominated by language models. This year, the program includes a regular session, a session dedicated to the time-aware Word-in-Context classification shared task, as well as non-archival presentations and Findings of EMNLP papers. Finally, are delighted to have the following renowned invited speakers: Eunsol Choi, Jacob Eisenstein, Adam Jatowt, Ozan Sener and Nazneen Rajani.

15 papers will be presented at EvoNLP. In particular, 3 research papers, 5 TempoWiC system description papers, 4 Findings of EMNLP papers, and 3 non-archival submissions. These papers cover a variety of topics, for example few shot learning using incremental approaches, clustering techniques for dynamic topic discovery, construction of temporal benchmarks, or leveraging time-dependent features for offensive language detection. Regarding the TempoWiC shared task, our participants leveraged techniques such as multitask learning, mixture-of-experts or generative approaches.

We would like to thank the Program Committee members for their support of this event in form of reviewing and feedback, without whom we would not be able to ensure the overall quality of the workshop.

Organizing Committee

Organizers

Francesco Barbieri, Snap Inc.
Jose Camacho-Collados, Cardiff University
Bhuwan Dhingra, Duke University
Luis Espinosa-Anke, Cardiff University / AMPLYFI
Elena Gribovskaya, DeepMind
Angeliki Lazaridou, DeepMind
Daniel Loureiro, Cardiff University
Leonardo Neves, Snap Inc.

Program Committee

Workshop Reviewers

Alessandro Raganato, University of Milano-Bicocca
Dimosthenis Antypas, Cardiff University
David Vilares, University of A Coruña
Ian Stewart, University of Michigan
Indira Sen, GESIS, Leibniz Institute
David Jurgens, University of Michigan
Sebastian Ruder, Google
Sanjaya Wijeratne, TrueNorth
Steven Schockaert, Cardiff University

Shared Task Reviewers

Daniel Loureiro, Cardiff University
Jose Camacho-Collados, Cardiff University
Chenyang Lyu, Dublin City University
Elizaveta Tukhtina, HSE University
Ze Chen, NetEase, Inc.

Keynote Talk: Knowledge-rich NLP models in a dynamic real world

Eunsol Choi
UT-Austin

Abstract: To address knowledge-rich tasks such as question answering, NLP models should combine knowledge from multiple sources – memorized knowledge in the language model (LM), and passages retrieved from an evidence corpus. Prior work assumed information stored in various sources is consistent with each other. Yet, a mismatch in knowledge sources is common in the real world: some sources are updated while others remain stale, and different sources can interpret the same question differently or propose differing opinions. How should we resolve such complex knowledge conflicts? In this talk, I will describe our recent work on (1) an evaluation framework for updating knowledge in LMs and (2) how state-of-the-art models behave under knowledge conflicts. I will conclude my talk with paths for handling real-world scenarios, continual learning of models and calibrating them to avoid answering when provided with incomplete or conflicting information.

Bio: Eunsol Choi is an assistant professor in the Computer Science department at the University of Texas at Austin and a visiting researcher at Google AI. Her research area spans natural language processing and machine learning. She is particularly interested in interpreting and reasoning about text in a rich real-world context. She received a Ph.D. from University of Washington and B.A from Cornell University. She is a recipient of Facebook research fellowship, Google faculty research award, and outstanding paper award at EMNLP 2021.

Keynote Talk: What can we learn from language change?

Jacob Eisenstein

Google

Abstract: Language changes are shaped by world events and social structures. With an increasingly rich set of resources for studying text over time, this raises the possibility of reverse-engineering language change to uncover new insights about the world. This talk will survey work on diachronic corpora of social media, historical newspapers, and scientific research papers. We use these corpora to build models of cultural transmission between US cities, the effects of social media platforms' policies and norms, the leaders of the movement to abolish slavery in the United States, and which ACL papers are likely to get the most citations in the future. Collaborators include Sandeep Soni, Umashanthi Pavalanathan, Lauren F. Klein, Kristina Lerman, and David Bamman.

Bio: Jacob Eisenstein is a research scientist at Google, where he is focused on making language technology more robust and trustworthy. He was previously on the faculty of the Georgia Institute of Technology, where he supervised six successful doctoral dissertations, received the NSF CAREER Award for research on computational sociolinguistics, and wrote a textbook on natural language processing. He completed his Ph.D. at MIT, winning the George M. Sprowls award for a dissertation on computational models of speech and gesture. Thanks to his brief appearance in the documentary film *If These Knishes Could Talk*, Jacob has a Bacon number of 2.

<https://jacobeisenstein.github.io/>

Keynote Talk: Automatic Question Answering over Temporal News Collections

Adam Jatowt

University of Innsbruck

Abstract: The field of automatic question answering has been rapidly advancing recently. The existing QA approaches are however generally working on synchronic document collections such as Wikipedia, Web data or short-term news corpora. This talk is about our latest efforts in automatic question answering over diachronic news collections composed of articles published over several decades. We will first discuss an unsupervised re-ranking approach that works by utilizing temporal information embedded in the temporal document collection. Next, we will introduce a solution for finding the occurrence dates of events described in input questions based on the underlying news dataset. Finally, we will introduce ArchivalQA - a large-scale question answering dataset which has been automatically created from a two decades' long news article collection, and which contains over 500k question-answer pairs. The dataset has been processed to remove temporally ambiguous questions for which more than one correct answer exist.

Bio: Adam Jatowt is a Professor at the Department of Computer Science and Digital Science Center at the University of Innsbruck. He received his Ph.D. in Information Science and Technology from the University of Tokyo, Japan in 2005. Before moving to Austria, Adam worked at Kyoto University as an Assistant and later as an Associate Professor. His research interests include information retrieval, natural language processing, digital libraries, and digital humanities. Adam has published over 180 research papers in international conferences and journals. He is an editorial board member of IP&M, JASIST, IJDL, JIIS, and IEEE JoSC journals. Adam has received the Vannevar Bush Best Paper Award at JCDL2021, the best short paper award at ECIR2018, and the best demo award at ECIR2019.

Keynote Talk: Going from Continual Learning Algorithms to Continual Learning Systems

Ozan Sener

Apple

Abstract: We envision continual learning systems to interact with humans, with each other, and the physical world through time – and continue to learn and adapt as they do. On the other hand, we develop and benchmark continual learning algorithms, hoping they will be the future founding blocks of such systems. I will argue that going from algorithmic to system thinking requires carefully questioning various assumptions. Consider the common assumption of storage being fixed and limited through the agent’s lifetime. The economics of data storage suggests that the cost of storing data decreases over time and is negligible compared to the cost of computing. Following this economics, computing systems around us are not constrained by storage but rather by computation. This motivates us to explore online continual learning with computation constraints instead of storage constraints. Using this new setting, we set a new state of the art on the largest large-scale continual learning datasets.

Bio: Ozan Sener is interested in machine learning and its applications in computer vision and robotics. Specifically, Ozan is interested in designing machine learning algorithms which can process large-amount of multimodal information with no/weak supervision.

Ozan is a research scientist at Apple. Ozan received his PhD from Cornell University, advised by Ashutosh Saxena. Following the advisor’s move, Ozan spent three beautiful years at Stanford AI Lab as a visiting PhD student (2015-2016) and a postdoc (2017) working with Silvio Savarese. Ozan obtained his BSc and MSc degrees from METU.

Keynote Talk: Takeaways from a systematic study of 75K models on Hugging Face

Nazneen Fatema Rajani
Hugging Face

Abstract: Language models trained using transformers dominate the NLP model landscape, making Hugging Face (HF) the de facto hub for sharing, benchmarking, and evaluating NLP models. The HF hub provides a rich resource for understanding how language models evolved, opening up research questions on how factors such as model age and documentation affect their usage. We conducted a systematic study of the hub to glean insights into the ever-evolving model landscape and what factors affect model usage. We also studied model documentation for semantic drifts, and observed an evolution in the use of specific keywords (such as “train,” “evaluate,” “impact”), indicating a paradigm shift from model-centric to more data-centric ML development. In this talk, I will give a macro-level view of this evolving model landscape and discuss the results from our systematic study of 75K HF models.

Bio: Nazneen is a Research Lead at HuggingFace, a startup with a mission to democratize ML, leading data-centric ML research which involves systematically analyzing, curating, and automatically annotating data. Before HF, she worked at Salesforce Research with Richard Socher and led a team of researchers focused on building robust natural language generation systems based on LLMs. She completed her Ph.D. in CS at UT-Austin with Prof. Ray Mooney.

Nazneen has over 30 papers published at ACL, EMNLP, NAACL, NeurIPS, and ICLR and has her research covered by Quanta magazine, VentureBeat, SiliconAngle, ZDNet, and Datanami. She is also teaching a course on interpreting ML models with Corise – <http://corise.com/go/nazneen>. More details about her work can be found here <https://www.nazneenrajani.com/>.

Table of Contents

<i>MLLabs-LIG at TempoWiC 2022: A Generative Approach for Examining Temporal Meaning Shift</i> Chenyang Lyu, Yongxin Zhou and Tianbo Ji	1
<i>Using Deep Mixture-of-Experts to Detect Word Meaning Shift for TempoWiC</i> Ze Chen, Kangxu Wang, Zijian Cai, Jiewen Zheng, Jiarong He, Max Gao and Jason Zhang	7
<i>Using Two Losses and Two Datasets Simultaneously to Improve TempoWiC Accuracy</i> Mohammad Javad Pirhadi, Motahhare Mirzaei and Sauleh Eetemadi	12
<i>Class Incremental Learning for Intent Classification with Limited or No Old Data</i> Debjit Paul, Daniil Sorokin and Judith Gaspers	16
<i>CC-Top: Constrained Clustering for Dynamic Topic Discovery</i> Jann Goschenhofer, Pranav Ragupathy, Christian Heumann, Bernd Bischl and Matthias Aßenma- cher	26
<i>HSE at TempoWiC: Detecting Meaning Shift in Social Media with Diachronic Language Models</i> Elizaveta Tukhtina, Kseniia Kashleva and Svetlana Vydrina	35
<i>Leveraging time-dependent lexical features for offensive language detection</i> Barbara McGillivray, Malithi Alahapperuma, Jonathan Cook, Chiara Di Bonaventura, Albert Meroño-Peñuela, Gareth Tyson and Steven R. Wilson	39
<i>Temporal Word Meaning Disambiguation using TimeLMs</i> Mihir Godbole, Parth Dandavate and Aditya Kane	55

Program

Wednesday, December 7, 2022

- 09:00 - 09:30 *Opening remarks*
- 09:30 - 10:00 *Jacob Eisenstein - What can we learn from language change?*
- 10:00 - 10:30 *Eunsol Choi - Knowledge-rich NLP models in a dynamic real world*
- 10:30 - 11:00 *Adam Jatowt - Automatic Question Answering over Temporal News Collections*
- 11:00 - 12:30 *Workshop poster session (virtual and on-site)*
- 12:30 - 14:00 *Lunch break*
- 14:00 - 15:00 *Findings and non-archival session*
- 15:00 - 15:30 *Coffee break*
- 15:30 - 16:00 *Nazneen Rajani - Takeaways from a systematic study of 75K models on Hugging Face*
- 16:00 - 16:30 *Ozan Sener - Going from Continual Learning Algorithms to Continual Learning Systems*
- 16:30 - 17:00 *Workshop oral session*
- 17:00 - 17:30 *Shared task session*
- 17:30 - 18:00 *Best paper awards and closing*