UncertaiNLP 2024

Workshop on Uncertainty-Aware NLP (UncertaiNLP 2024)

Proceedings of the Workshop

March 22, 2024

©2024 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 979-8-89176-075-2

Introduction

Human languages are inherently ambiguous and understanding language input is subject to interpretation and complex contextual dependencies. Nevertheless, the main body of research in NLP is still based on the assumption that ambiguities and other types of underspecification can and have to be resolved. The UncertaiNLP workshop (workshop on uncertainty-aware NLP) aims to provide a platform for research that embraces variability in human language and aims to represent and evaluate the uncertainty that arises from it, and from modeling tools themselves.

This volume contains the proceedings of the first edition of the UncertaiNLP workshop hosted on March 22nd 2024, co-located with the 18th Conference of the European Chapter of the European Chapter of the Association for Computational Linguistics in the Radisson Blu and Corinthia St George's Bay hotel in St Julian's, Malta. We invited paper submissions on a wide variety of topics, including representing, documenting or modeling uncertainty, parameter estimation, probabilistic inference, decision making and evaluation. We received a total of 28 submissions, of which we accepted 8 long and 7 short papers, amounting to an acceptance rate of 65.22

We are also grateful to our invited keynote speakers: Kristin Lennox (Exponent) discussed TODO, Mohit Bansal (UNC Chapel Hill) contributed a talk on confidence-based rephrasing, while Clara Meister (ETH Zürich) presented TODO and Chrysoula Zerva (Instituto Superior Tecnico) provided insights on TODO. Lastly, we want to express our gratitude to Raul Vasquez for helping to compile these proceedings.

We would also like to thank the Research Council of Finland for their support of the workshop through the project Uncertainty-Aware Neural Language Models and the EU's Horizon Europe research and innovation program for support through the Unified Transcription and Translation (UTTER) project.

The UncertaiNLP organizers,

Wilker Aziz, Joris Baan, Hande Celikkanat, Marie-Catherine de Marneffe, Barbara Plank, Swabha Swayamdipta, Jörg Tiedemann, Dennis Ulmer, Raúl Vázquez

Program Committee

Program Chairs

Wilker Aziz, University of Amsterdam
Joris Baan, University of Amsterdam
Hande Celikkanat, University of Helsinki
Marie-Catherine de Marneffe, UCLouvain
Barbara Plank, Ludwig-Maximilians-Universität München and IT University of Copenhagen
Swabha Swayamdipta, University of Southern California
Jörg Tiedemann, University of Helsinki
Dennis Ulmer, IT University of Copenhagen and Pioneer Centre for Artificial Intelligence
Raúl Vázquez, University of Helsinki

Publication Chair

Raúl Vázquez, University of Helsinki

Reviewers

Luigi Acerbi, University of Helsinki
Mathias Creutz, University of Helsinki
Nico Daheim, Technische Universität Darmstadt
Stella Frank, University of Copenhagen
Markus Heinonen, Aalto University
Evgenia Ilia, University of Amsterdam
Robin Jia, University of Southern California
Gabriella Lapesa, University of Stuttgart
Putra Manggala, University of Amsterdam
Timothee Mickus, University of Helsinki
Eric Nalisnick, University of Amsterdam
Natalie Schluter, Technical University of Denmark, Apple and IT University of Copenhagen
Philip Schulz, Amazon
Sebastian Schuster, Universität des Saarlandes
Rico Sennrich, University of Zurich and University of Edinburgh
Edwin Simpson, University of Bristol
Arno Solin, Aalto University
Dharmesh Tailor, University of Amsterdam
Aarne Talman, University of Helsinki
Ivan Titov, University of Edinburgh and University of Amsterdam
Teemu Vahtola, University of Helsinki
Sami Virpioja, University of Helsinki
Raúl Vázquez, University of Helsinki
Xinpeng Wang, Ludwig-Maximilians-Universität München
Leon Weber-Genzel, Ludwig-Maximilians-Universität München
Roman Yangarber, University of Helsinki

Keynote Talk: Uncertainty in NLP: Quantification, interpretation and evaluation

Chrysoula Zerva
Instituto Superior Tecnico, Portugal
2024-03-22 09:10:00 – Room: **Corinthia, Bastion 2**

Abstract: As the availability (and size) of language models keeps increasing, so do their applications to different tasks, rendering them ubiquitous in modern society. This in turn, brings forward the question of reliability. We know models don't always know what they don't knowand hence being able to quantify the uncertainty over their predictions is a key step in the path towards reliability. But how can we estimate uncertainty when we have multiple sources of it, and frequently no or limited access to the parameters of the models? And how do we know if we can trust our uncertainty estimations? In this talk I will discuss uncertainty quantification in NLP, emphasising its interpretation and evaluation. I will focus on generation and evaluation tasks, using machine translation as the main paradigm.

Bio: Chrysoula (Chryssa) Zerva is an Assistant Professor in Artificial Intelligence at the Instituto Superior Tecnico in Lisbon, Portugal. She is also a member of LUMLIS, the Lisbon ELLIS unit. She obtained her Ph.D. in 2019 from the University of Manchester working on "Automated identification of textual uncertainty" under the supervision of Prof. Sophia Ananiadou. She was subsequently awarded the EPSRC doctoral prize fellowship to study (mis)information propagation in health and science. In 2021, she joined the Instituto de Telecomunicações in Lisbon as a post-doc for the DeepSPIN project under the supervision of Prof. André Martins and worked on a range of machine learning and NLP related topics including uncertainty quantification, machine translation and quality estimation.

Keynote Talk: A Quantification of Semantic Uncertainty in Language Models

Clara Meister
ETH Zürich, Switzerland
2024-03-22 13:15:00 – Room: **Corinthia, Bastion 2**

Abstract: In machine learning, we are often concerned with model-related (i.e., epistemic) uncertainty. In natural language tasks specifically though, there exists quite a bit of inherent data-centric uncertainty, coming from characteristics of natural language such as ambiguity in meaning and the ability to express the same idea via multiple surface forms. In a field where standard evaluation largely assumes that there is a single correct answer, knowledge of the "semantic uncertainty" of a situation can prove useful. For example, it can provide insights into when there are several interpretations of a source sentence in machine translation or when there are multiple plausible answers to a question in question answering. We propose a simple method for quantifying uncertainty in standard, embedding-based language models (LM) that does not require fine-tuning or external models. We use the LM's embedding space to approximate the underlying distribution over semantic meanings of continuations, then analyzing this distribution to get a quantification of semantic uncertainty. We show some empirical use cases for this quantification of semantic uncertainty.

Bio: Clara Meister is a PhD student in Computer Science with Prof. Ryan Cotterell at ETH Zürich, supported by a Google PhD Fellowship. She is passionate about the general application of statistics and information theory to natural language processing. A large portion of her research in the last years has been on natural language generation—specifically, on decoding methods for probabilistic models. Her additional interests within the field of natural language generation include evaluation metrics and the incorporation of uncertainty into decoding methods.

Keynote Talk: TBA

Kristin Lennox
Exponent, US
2024-03-22 16:30:00 – Room: **Corinthia, Bastion 2**

Abstract: TBA

Bio: Kristin Lennox is a consultant at Exponent with more than ten years of experience applying statistics, machine learning, and operations research techniques to scientific and engineering problems. Dr. Lennox received her Ph.D. in statistics from Texas A&M University in 2010. She then joined Lawrence Livermore National Laboratory, where she cofounded and served as the first director of their internal statistical consulting service. After leaving the laboratory she spent several years in the software industry with a focus on AI in industrial settings, and she currently serves as a consultant regarding statistics and AI implementation for applications in many areas, including environmental science, automotive and consumer product risk, and software. Her expertise includes experimental design, analysis of computer experiments, and risk assessment in high consequence environments. Dr. Lennox's recent professional experience has focused on methods to characterize safety benefits of advanced driver assistance systems (ADAS) and automated driving. Dr. Lennox is passionate about statistics and AI education and has created a series of videos for technical and lay audiences on these topics.

Table of Contents

Calibration-Tuning: Teaching Large Language Models to Know What They Don't Know
Sanyam Kapoor, Nate Gruver, Manley Roberts, Arka Pal, Samuel Dooley, Micah Goldblum and Andrew Gordon Wilson 1

Context Tuning for Retrieval Augmented Generation
Raviteja Anantha and Danil Vodianik 15

Optimizing Relation Extraction in Medical Texts through Active Learning: A Comparative Analysis of Trade-offs
Siting Liang, Pablo Valdunciel Sánchez and Daniel Sonntag 23

Linguistic Obfuscation Attacks and Large Language Model Uncertainty
Sebastian Steindl, Ulrich Schäfer, Bernd Ludwig and Patrick Levi 35

Aligning Uncertainty: Leveraging LLMs to Analyze Uncertainty Transfer in Text Summarization
Zahra Kolagar and Alessandra Zarcone 41

How Does Beam Search improve Span-Level Confidence Estimation in Generative Sequence Labeling?
Kazuma Hashimoto, Iftekhar Naim and Karthik Raman 62

Efficiently Acquiring Human Feedback with Bayesian Deep Learning
Haishuo Fang, Jeet Gor and Edwin Simpson 70

Order Effects in Annotation Tasks: Further Evidence of Annotation Sensitivity
Jacob Beck, Stephanie Eckman, Bolei Ma, Rob Chew and Frauke Kreuter 81

The Effect of Generalisation on the Inadequacy of the Mode
Bryan Eikema 87

Uncertainty Resolution in Misinformation Detection
Yury Orlovskiy, Camille Thibault, Anne Imouza, Jean-François Godbout, Reihaneh Rabbany and Kellin Pelrine 93

Don't Blame the Data, Blame the Model: Understanding Noise and Bias When Learning from Subjective Annotations
Abhishek Anand, Negar Mokhberian, Prathyusha Naresh Kumar, Anweasha Saha, Zihao He, Ashwin Rao, Fred Morstatter and Kristina Lerman 102

Combining Confidence Elicitation and Sample-based Methods for Uncertainty Quantification in Misinformation Mitigation
Mauricio Rivera, Jean-François Godbout, Reihaneh Rabbany and Kellin Pelrine 114

Linguistically Communicating Uncertainty in Patient-Facing Risk Prediction Models
Adarsa Sivaprasad and Ehud Reiter 127

Program

Friday, March 22, 2024

09:00 - 09:10 *Opening Remarks*

09:10 - 09:55 *Keynote Talk - Chrysoula Zerva*

09:55 - 10:15 *Presentation Session 1*

Aligning Uncertainty: Leveraging LLMs to Analyze Uncertainty Transfer in Text Summarization
Zahra Kolagar and Alessandra Zarcone

10:15 - 10:30 *Poster Spotlights*

Context Tuning for Retrieval Augmented Generation
Raviteja Anantha and Danil Vodianik

Linguistic Obfuscation Attacks and Large Language Model Uncertainty
Sebastian Steindl, Ulrich Schäfer, Bernd Ludwig and Patrick Levi

Order Effects in Annotation Tasks: Further Evidence of Annotation Sensitivity
Jacob Beck, Stephanie Eckman, Bolei Ma, Rob Chew and Frauke Kreuter

The Effect of Generalisation on the Inadequacy of the Mode
Bryan Eikema

Uncertainty Resolution in Misinformation Detection
Yury Orlovskiy, Camille Thibault, Anne Imouza, Jean-François Godbout, Reihaneh Rabbany and Kellin Pelrine

Combining Confidence Elicitation and Sample-based Methods for Uncertainty Quantification in Misinformation Mitigation
Mauricio Rivera, Jean-François Godbout, Reihaneh Rabbany and Kellin Pelrine

Linguistically Communicating Uncertainty in Patient-Facing Risk Prediction Models
Adarsa Sivaprasad and Ehud Reiter

Friday, March 22, 2024 (continued)

10:30 - 11:00 *Coffee Break*

11:00 - 12:15 *In-Person Poster Session*

12:15 - 13:15 *Lunch Break*

13:15 - 14:00 *Keynote Talk - Clara Meister*

14:00 - 15:30 *Presentation Session 2*

Optimizing Relation Extraction in Medical Texts through Active Learning: A Comparative Analysis of Trade-offs
Siting Liang, Pablo Valdunciel Sánchez and Daniel Sonntag

Efficiently Acquiring Human Feedback with Bayesian Deep Learning
Haishuo Fang, Jeet Gor and Edwin Simpson

Don't Blame the Data, Blame the Model: Understanding Noise and Bias When Learning from Subjective Annotations
Abhishek Anand, Negar Mokhberian, Prathyusha Naresh Kumar, Anweasha Saha, Zihao He, Ashwin Rao, Fred Morstatter and Kristina Lerman

Teaching Probabilistic Logical Reasoning to Transformers
Aliakbar Nafar, K. Brent Venable and Parisa Kordjamshidi

Calibration-Tuning: Teaching Large Language Models to Know What They Don't Know
Sanyam Kapoor, Nate Gruver, Manley Roberts, Arka Pal, Samuel Dooley, Micah Goldblum and Andrew Gordon Wilson

15:30 - 16:00 *Coffee Break*

16:00 - 16:30 *Presentation Session 3*

How Does Beam Search improve Span-Level Confidence Estimation in Generative Sequence Labeling?
Kazuma Hashimoto, Iftekhar Naim and Karthik Raman

Friday, March 22, 2024 (continued)

Consistent Joint Decision-Making with Heterogeneous Learning Models
Hossein Rajaby Faghihi and Parisa Kordjamshidi

16:30 - 17:15 *Keynote Talk - Kristin Lennox*

17:15 - 17:30 *Closing Remarks*