

Katerina Margatina

Ph.D. Student in Natural Language Processing

✉ k.margatina@sheffield.ac.uk  <https://katerinamargatina.github.io>
 Twitter  Github  LinkedIn  Google Scholar  Semantic Scholar

Research Interests

active learning, large language models, dynamic adversarial data collection, evaluation & benchmarking, language modeling, robustness, fairness, temporal drift, disagreement, cross-cultural natural language processing

Education

Oct. 2019 – Ongoing	University of Sheffield, UK <i>Ph.D. in Computer Science</i> <ul style="list-style-type: none">Natural Language Processing Group (SheffieldNLP)Supervisors: Prof. Nikolaos AletrasFunding: Amazon Alexa FellowshipThesis: “Active Learning Algorithms for Data Efficient Language Models”Expected Graduation: August 2023
Jan. 2016 – Jul. 2016	Universidad Carlos III de Madrid, Spain <i>MEng in Computer Engineering (Erasmus Mobility in Masters)</i>
Sep. 2012 – Jul. 2019	National Technical University of Athens, Greece <i>BEng & MEng in Electrical and Computer Engineering</i> <ul style="list-style-type: none">GPA: 8.06/10 (top 15%)Major: Electronics and SystemsThesis: “Transfer Learning and Attention-based Conditioning Methods for Natural Language Processing” - supervised by Prof. Alexandros Potamianos

Professional Experience

Aug. 2022 – Mar. 2023	Meta, UK <i>Research Scientist Intern at Meta AI Research (FAIR Labs) in London</i> <ul style="list-style-type: none">Active Learning Principles for In-Context Learning with Large Language Models (paper under review)
Apr. 2022 – Aug. 2022	Amazon Web Services, USA <i>Applied Scientist Intern at Amazon Comprehend in NYC</i> <ul style="list-style-type: none">Dynamic Benchmarking of Masked Language Models on Temporal Concept Drift with Multiple Views (paper published in EACL 2023)
Jun. 2021 – Feb. 2022	University of Copenhagen, Denmark <i>Research Intern / Visiting Ph.D. in Computer Science at CoAStAL</i>
Feb. 2020 - Jun. 2021	University of Sheffield, UK <i>Teaching Assistant at the Department of Computer Science</i> <ul style="list-style-type: none">COM4513/6513 Natural Language Processing (2020 & 2021)COM3110/4115 Text Processing (2020)
Nov. 2018 - Jul. 2019	DeepSea Technologies, Piraeus, Greece <i>Machine Learning Engineer at the AI team</i>

Teaching & Other Academic Experience

2022 - Ongoing	Workshop Co-Organiser <ul style="list-style-type: none">DADC: The First Workshop on Dynamic Adversarial Data Collection at NAACL 2022 (website)
2021	MSc. Dissertation Co-Supervisor <ul style="list-style-type: none">Co-advised Ard Snijders, a MSc student from the University of Amsterdam, along with Douwe Kiela on the thesis project <i>Investigating Multi-source Active Learning for Natural Language Inference</i> (published in EACL 2023).Co-advised Atsuki Yamaguchi, a MSc student from the University of Sheffield, along with George Chrysostomou and Nikolaos Aletras on the project <i>Frustratingly Simple Alternatives to Masked Language Modeling</i> (published in EMNLP 2021).
2019 - Ongoing	Teaching Assistant <ul style="list-style-type: none">Assisted students during the lab sessions of the Text Processing and Natural Language Processing modules.Marked lab reports following a pre-determined scheme set by the Course Leader.Experience in face-to-face teaching with students that needed extra guidance.
2018 - Ongoing	Program Committee/Reviewer <ul style="list-style-type: none">ACL, EMNLP, NAACL, EACL, AACL, WASSA, NLLP, DADC

Invited Talks & Presentations

Jun. 2022	Bloomberg, AI Group / in-person in NYC
Apr. 2022	NLPhD Speaker Series , University of Saarland (slides) / remote

Honors & Awards

2022	Dynabench: Awarded a research grant of \$46,825 from Meta for our proposal : “Adversarial NERDs: Optimising feedback between humans-and-model-in-the-loop” (with Hannah Rose Kirk and Dr. Scott Hale).
2020	EMNLP: Awarded a scholarship to participate to the virtual Conference on Empirical Methods in Natural Language Processing (EMNLP 2020). MLSS: Selected to participate to the highly competitive Machine Learning Summer School (MLSS) organized by the Max Planck Institute for Intelligent Systems (180 out of 1300+ applications were accepted).
2019	Amazon Alexa Fellowship: Awarded an Amazon Alexa Fellowship for a 3-year Ph.D. program at the University of Sheffield under the supervision of Dr. Nikolaos Aletras. EETN Scholarship Award: Awarded a scholarship of from the Hellenic Artificial Intelligence Society (EETN) to attend the first Athens Natural Language Processing Summer School (AthNLP). ACM-W Scholarship Award: Awarded an ACM-W scholarship of \$600 to attend the Annual Meeting of the Association for Computational Linguistics in Florence (ACL 2019).
2018	WASSA IEST: Ranked 3 rd out of 30 participants at the WASSA 2018 Implicit Emotion Shared Task (IEST) conducted by the EMNLP 2018 Conference.

Publications

- | | |
|------|--|
| 2023 | 1. Margatina, K. & Aletras, N. <i>On the Limitations of Simulating Active Learning in Findings of ACL</i> (2023). |
| | 2. Margatina, K. , Schick, T., Aletras, N. & Dwivedi-Yu, J. <i>Active Learning Principles for In-Context Learning with Large Language Models</i> . <i>Arxiv</i> (2023). |
| | 3. Margatina, K. et al. <i>Dynamic Benchmarking of Masked Language Models on Temporal Concept Drift with Multiple Views</i> in <i>EACL</i> (2023). |
| | 4. Snijders, A., Kiela, D. & Margatina, K. <i>Investigating Multi-source Active Learning for Natural Language Inference</i> in <i>EACL</i> (2023). |
| 2022 | 5. Hershcovich, D. et al. <i>Challenges and Strategies in Cross-Cultural NLP</i> in <i>ACL</i> (2022). |
| | 6. Margatina, K. , Barrault, L. & Aletras, N. <i>On the Importance of Effectively Adapting Pretrained Language Models for Active Learning</i> in <i>ACL</i> (2022). |
| 2021 | 7. Margatina, K. , Vernikos, G., Barrault, L. & Aletras, N. <i>Active Learning by Acquiring Contrastive Examples</i> in <i>EMNLP</i> (2021). |
| | 8. Yamaguchi, A., Chrysostomou, G., Margatina, K. & Aletras, N. <i>Frustratingly Simple Alternatives to Masked Language Modeling</i> in <i>EMNLP</i> (2021). |
| 2020 | 9. Vernikos, G., Margatina, K. , Chronopoulou, A. & Androutsopoulos, I. <i>Domain Adversarial Fine-Tuning as an Effective Regularizer</i> in <i>Findings of EMNLP</i> (2020). |
| 2019 | 10. Margatina, K. <i>Transfer Learning and Attention-based Conditioning Methods for Natural Language Processing</i> . Thesis, NTUA (2019). |
| | 11. Margatina, K. , Baziotis, C. & Potamianos, A. <i>Attention-based Conditioning Methods for External Knowledge Integration</i> in <i>ACL</i> (2019). |
| 2018 | 12. Chronopoulou*, A., Margatina*, K., Baziotis, C. & Potamianos, A. <i>NTUA-SLP at IEST 2018: Ensemble of Neural Transfer Methods for Implicit Emotion Classification</i> in <i>Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis</i> (2018). |

Coding

Expert: Python, Bash / PyTorch, HuggingFace transformers

Familiar with: C, C++, Matlab, SQL, Java / Tensorflow, Keras

Other: Git, \LaTeX , Unix, Flask, AWS

Languages

Greek (*Native*), English (*C2*), Spanish (*B2*), French (*B1*)

Personal Interests

Tennis, Yoga, Boulderling, Hiking, Skiing, Sailing (diploma)