Şaziye Betül Özateş

Contact Information

Institute for Data Science and AI
Boğaziçi University
Güney Kampüs, Bebek
34342 İstanbul, Turkey
saziye.ozates@bogazici.edu.tr
linkedin.com/in/şaziye-betül-özateş-a86982211

Research Interests

Natural language processing, deep learning, machine learning, computational linguistics, artificial intelligence.

About

I am a computer scientist who specializes in natural language processing and machine learning. I have received my bachelor's degree, and master's degree from the Department of Computer Engineering in Boğaziçi University. I completed my PhD studies in the same place under the guidance of Dr. Arzucan Özgür and Dr. Tunga Güngör. From 2020 to 2022, I was a researcher at the Institute of Natural Language Processing at the University of Stuttgart and worked as a post-doctoral research fellow at KUIS-AI Center between 2022 and 2023. Currently, I am an assistant professor at the Institute for Data Science and Artificial Intelligence at Boğaziçi University.

I worked on various computational linguistics, machine learning, and data mining projects throughout my career. My current research focuses on syntactic and semantic analysis of natural language data with the help of various deep learning methods.

Experience

Assistant Professor - Full Time

DSAI, Boğaziçi University, 2023-

Research Fellow - Full Time

KUIS AI Center, 2022-2023

Researcher - Full Time

University of Stuttgart, 2020-2022

Teaching Assistant - Part Time

Boğaziçi University, 2015-2016, 2020

Research Assistant - Full Time

Turkish-German University, 2016-2018

Education

Boğaziçi University, İstanbul, Turkey

Ph.D. in Computer Engineering, May 2022

- Thesis Topic: Deep Learning-based Dependency Parsing for Turkish
- Advisor: Assoc. Prof. Arzucan Özgür and Prof. Tunga Güngör
- Earned 4.0 GPA on 4.0 scale in coursework

Boğaziçi University, İstanbul, Turkey

M.S. in Computer Engineering, June 2014

- Thesis Topic: Multi-document Summarization using Dependency Grammars
- Advisor: Assoc. Prof. Arzucan Özgür
- Earned 4.0 GPA on 4.0 scale in coursework

Boğaziçi University, İstanbul, Turkey

B.S. in Computer Engineering, June 2012

- Senior Project Topic: A Comparison of Financial Time Series Forecasting Using Multilayer Perceptron, Support Vector Machines, and Bayesian Network
- Advisor: Prof. Fikret Gürgen
- Highest honors in Engineering Faculty
- Ranked 2^{nd} in graduating class with 3.64 GPA on 4.0 scale

Publications

Refereed International Journal Papers

Özateş, Ş. B., Özgür, A., Güngör, T., Öztürk, B. (2022). "A Hybrid Approach to Dependency Parsing: Combining Rules and Morphology with Deep Learning."*IEEE Access, 10, 93867-93886.*

Türk, U., Atmaca, F., Özateş, Ş. B. *et al.* (2022). "Resources for Turkish Dependency Parsing: Introducing the BOUN Treebank and the BoAT Annotation Tool." *Language Resources & Evaluation, 56(1), 259-307.*

Refereed International Conference Papers

Karagöz, F. B., Doğan, B., Özateş, Ş. B. (2024). "Towards a Clean Text Corpus for Ottoman Turkish." *Proceedings of the First Workshop on Natural Language Processing for Turkic Languages (SIGTURK 2024), Bangkok, Thailand. Association for Computational Linguistics.*

Özateş, Ş. B., Tıraş, T. E., Genç, E. E., Bilgin Taşdemir, E. F. (2024). "Dependency Annotation of Ottoman Turkish with Multilingual BERT." *Proceedings of the 18th Linguistic Annotation Workshop, St. Julians, Malta. Association for Computational Linguistics.*

Özateş, Ş. B., Özgür, A., Güngör, T., Çetinoğlu, Ö. (2022). "Improving Code-Switching Dependency Parsing with Semi-Supervised Auxiliary Tasks." *Findings of the Association for Computational Linguistics: NAACL 2022, pp. 1159-1171.*

Marşan, B., Akkurt, S. F., Şen, M., Gürbüz, M., Özateş, Ş. B., Üsküdarlı, S., Özgür, A., Güngör, T., Öztürk, B. (2022). "Enhancements to the BOUN Treebank Reflecting the

Agglutinative Nature of Turkish." *International Conference on Agglutinative Language Technologies as a challenge of Natural Language Processing (ALTNLP), Virtual Event, Koper, Slovenia.*

Özateş, Ş. B., Çetinoğlu, Ö. (2021). "A Language-aware Approach to Code-switched Morphological Tagging." *Proceedings of the Fifth Workshop on Computational Approaches to Linguistic Code-Switching, pp. 72-83. Association for Computational Linguistics.*

Türk, U., Bayar, K., Özercan, A. D., Öztürk, G. Y., Özateş, Ş. B. (2020). "First Steps towards Universal Dependencies for Laz." *In Proceedings of the Fourth Workshop on Universal Dependencies (UDW 2020), pp. 189-194. Association for Computational Linguistics.*

Türk, U., Atmaca, F., Özateş, Ş. B. *et al.* (2019). "Improving the Annotations in the Turkish Universal Dependency Treebank." *In Universal Dependencies Workshop (UDW 2017) at SyntaxFest 2019, pp. 108-115, Paris, France. Association for Computational Linguistics.*

Türk, U., Atmaca, F., Özateş, Ş. B. *et al.* (2019). "Turkish Treebanking: *Unifying and Constructing Efforts." Proceedings of the 13th Linguistic Annotation Workshop, pp. 166-177, Florence, Italy. Association for Computational Linguistics.*

Özateş, Ş. B., Özgür, A., Güngör, T., Öztürk, B. (2018). "A Morphology-based Representation Model for LSTM-based Dependency Parsing of Agglutinative Languages." *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies, pp. 238-247, Brussels, Belgium. Association for Computational Linguistics.*

Özateş, Ş. B., Özgür, A., Güngör, Radev, D. R. (2016). "Sentence Similarity based on Dependency Tree Kernels for Multi-document Summarization." *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), pp. 2833-2838, Portorož, Slovenia. European Language Resources Association (ELRA).*

Thesis

Şaziye Betül Özateş. "Deep Learning-based Dependency Parsing for Turkish", *Ph.D. Thesis, Computer Engineering, Boğaziçi University, İstanbul, Turkey, May 2022.*
Thesis Advisors: Assoc. Prof. Arzucan Özgür and Prof. Tunga Güngör.

Şaziye Betül Bilgin. "Multi-document Summarization using Dependency Grammars", *M.S. Thesis, Computer Engineering, Boğaziçi University, İstanbul, Turkey, June 2014.*
Thesis Advisor: Assoc. Prof. Arzucan Özgür.

DFG Project (German Research Foundation)

- Computational Structural Analysis of German-Turkish Code-Switching (*August 2020 - February 2022*)
  - The aim of this project is to analyse Turkish-German code-switching (CS) from a computational perspective. We developed deep learning-based computational models for morphological and syntactic analyses of CS data.

Tübitak Projects

- ARDEB 3005 - Automatic Processing and Analysis of Kazasker Ruznamçe Records with Digital Methods *(2024 -)*
- BİLGEM - Development of a Deep Learning-Based Data Expansion Tool for Predicting the Efficacy of Antibacterial Nanoparticles Used in Nanomedicine *(2024 -)*
- ARDEB 2232 - Applications of Deep Learning Architectures for Turkey: GANs for Healthcare *(2019 - 2020)*
  - The purpose of this project is investigating generative adversarial network (GAN) structures and proposing deep learning-based solutions for improving the healthcare applications in Turkish.
- ARDEB 1005 - A Deep Learning based Turkish Dependency Parser *(2018 - 2019)*
  - We developed a novel state-of-the-art deep learning-based dependency parsing system and created a manually annotated dependency treebank for Turkish. With this system and the newly introduced treebank, we reached the best performance on the dependency parsing task that is needed in many natural language processing applications using deep learning techniques for Turkish.

Attended International Conferences

- EACL, The 18th Conference of the European Chapter of the Association for Computational Linguistics, March 17-22, 2024, St. Julians, Malta.
- NAACL, 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics, June 6–11, 2021, Online
- ACL, The 57th Annual Meeting of the Association for Computational Linguistics (ACL) July 28-August 2, 2019, Florence, Italy
- EMNLP, 2018 Conference on Empirical Methods in Natural Language Processing, October 31–November 4, 2018, Brussels, Belgium
- LREC, The 10th edition of the Language Resources and Evaluation Conference, May 23-28, 2016, Portorož, Slovenia

Invited Talks

- Şaziye Betül Özateş, "A Semi-supervised Approach for Improving Code-switching Dependency Parsing", Graduate Seminar, Boğaziçi University, Istanbul, Turkey (29 Nov, 2022)
- Şaziye Betül Özateş, "A Semi-supervised Dependency Parsing Approach for Low-resource Languages", CS and AI Talks, Boğaziçi University, Istanbul, Turkey (11 Nov, 2022)
- Şaziye Betül Özateş, "Deep Learning-based Dependency Parsing For Turkish", ClingDing Talks, Indiana University, Bloomington, IN, USA (20 Oct, 2022)

- Şaziye Betül Özateş, "Natural Language Processing for Low-resource Languages using Deep Learning and Linguistic Features", KUIS AI Talks, Koç University, İstanbul, Turkey (18 Jul, 2022)

- Şaziye Betül Özateş, "Improving Dependency Parsing of Turkish: A Hybrid Neural Parser and the BOUN Treebank", Institutsversammlung, IMS, Universität Stuttgart, Stuttgart, Germany (4 Nov, 2020)

Teaching Experience

Instructor
Institute for Data Science and Artificial Intelligence, Boğaziçi University

DSA I545, Natural Language Processing

Institute for Data Science and Artificial Intelligence, Boğaziçi University

DSAI 303, Probability and Statistics for Data Science and Artificial Intelligence

Teaching Assistant
Computer Engineering Department, Boğaziçi University

CMPE 150, Introduction to Computing

Course Assistant
Computer Engineering Department, Boğaziçi University

CMPE 493, Introduction to Information Retrieval

Honors and Awards

2019-2020	Tubitak ARDEB 2232 - Scholarship from The International Fellowship for Outstanding Researchers Program
2018-2019	Tubitak ARDEB 1005 - Scholarship from National New Ideas and Products Research Support Program
2014–2019	Tubitak BIDEB 2211 - Scholarship for PhD Students
2012–2014	Tubitak BIDEB 2228 - MSc and PhD Scholarship for Senior Undergraduate Students
2012	Highest honors in Engineering Faculty, Boğaziçi University

Computer Skills

- Python, Java, C, C++
- Matlab, Latex, Perl
- Deep neural networks, Transformers, NLP applications
- OS: Linux-based systems, Windows

Relevant Skills

Languages: Turkish (*Native*), English (*Advanced*), German (*Intermediate*)