Furkan Şahinuç

Doctoral Researcher, Ubiquitous Knowledge Processing (UKP) Lab
FB 20 / Department of Computer Science
Technical University of Darmstadt
Hochschulstr. 10, 64289 Darmstadt, Germany
Office: S2|02|B106
E-mail: furkan.sahinuc@tu-darmstadt.de
Web page: furkansahinuc.github.io

 furkansahinuc | furkan-sahinuc | Google Scholar | Semantic Scholar

Professional Experience

Ubiquitous Knowledge Processing Lab, Darmstadt, Germany

    Research Intern   Nov 2022 - Oct 2023

- Research on automatic scientific writing via large language models (particularly related work generation) in InterText Group.

ASELSAN Research Center

    Research Engineer in NLP Team   June 2019 - Sep 2022

- Research on hate speech detection, misinformation detection in social media.
- Other research topics are interpretability of word embeddings, signal processing tools in NLP, gender bias.

Education

Technical University of Darmstadt, Darmstadt, Germany

    Ph.D., Computer Science, Ubiquitous Knowledge Processing Lab   2023 -

        Advisor: Prof. Iryna Gurevych

Bilkent University, Ankara, Türkiye

    M.Sc., Electrical and Electronics Engineering (GPA: 3.68/4.00)   2019 - 2021

        Advisor: Asst. Prof. Aykut Koç, **Co-advisor:** Prof. Haldun Ozaktas

        Thesis title: *Disclosing Zipfian Regularities in Semantic Breadth of Words via Multimodal Gaussian Embeddings.*

    B.S., Electrical and Electronics Engineering (GPA: 3.71/4.00)   2014 - 2019

Publications

Conference Proceedings

[C10]   **F. Şahinuç**, T. T. Tran, Y. Grishina, Y. Hou, B. Chen, and I. Gurevych. (2024). Efficient Performance Tracking: Leveraging Large Language Models for Automated Construction of Scientific Leaderboards. *In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, (pp. 7963–7977). doi: 10.18653/v1/2024.emnlp-main.453

[C9]   **F. Şahinuç**, I. Kuznetsov, Y. Hou, and I. Gurevych. (2024). Systematic Task Exploration with LLMs: A Study in Citation Text Generation. *In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, (pp. 4832–4855). doi: 10.18653/v1/2024.acl-long.265

[C8]   C. Toraman, O. Ozcelik, **F. Şahinuç**, and F. Can. (2024). MiDe22: An Annotated Multi-Event Tweet Dataset for Misinformation Detection. *In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, (pp. 11283–11295). URL: https://aclanthology.org/2024.lrec-main.986/

[C7]   C. Toraman, **F. Şahinuç**, and E. H. Yilmaz. (2022). Large-scale Hate Speech Detection with Cross-domain Transfer. *In Proceedings of The Language Resources and Evaluation Conference (LREC)*, (pp. 2215–2225). URL: https://aclanthology.org/2022.lrec-1.238/

[C6]   C. Toraman, **F. Şahinuç**, and E. H. Yilmaz. (2022). BlackLivesMatter2020: An Analysis of Deleted and Suspended Users in Twitter. *In 14th ACM Web Science Conference*, (pp. 290–295). doi: 10.1145/3501247.3531539

[C5]   C. Toraman, O. Ozcelik, **F. Şahinuç**, and U. Sahin. (2022). ARC-NLP at CheckThat! 2022: Contradiction for Harmful Tweet Detection. *In Working notes of CLEF 2022 - Conference and Labs of the Evaluation Forum*, (pp. 722-739). URL: https://ceur-ws.org/Vol-3180/paper-59.pdf

[C4]   O. Ozcelik, E. H. Yilmaz, **F. Şahinuç**, and C. Toraman. (2022). Slot Filling for Voice Assistants. *In 30th Signal Processing and Communications Applications Conference (SIU)*, (pp. 1–4). doi: 10.1109/SIU55565.2022.9864737

[C3]   **F. Şahinuç** and C. Toraman. (2021). Tweet Length Matters: A Comparative Analysis on Topic Detection in Microblogs. *In European Conference on Information Retrieval (ECIR)*, (pp. 471–478). doi: 10.1007/978-3-030-72240-1_50

[C2]   **F. Şahinuç**, C. Toraman, and A. Koç. (2021). Topic Detection Based on Deep Learning Language Model in Turkish Microblogs. *In 29th Signal Processing and Communications Applications Conference (SIU)*, (pp. 1–4). doi: 10.1109/SIU53274.2021.9477781

[C1]   **F. Şahinuç**, V. Yücesoy, and A. Koç. (2020). Intent Classification and Slot Filling for Turkish Dialogue Systems. *In 28th Signal Processing and Communications Applications Conference (SIU)*, (pp. 1–4). doi: 10.1109/SIU49456.2020.9302308

Journal Articles

[J9]   N. Sevim, E. O. Ozyedek, **F. Şahinuç**, T. Ozates, and A. Koç. (2025). Fast-FNet: Accelerating Transformer Encoder Models via Efficient Fourier Layers. *Signal, Image and Video Processing*, 19(12). doi: 10.1007/s11760-025-04552-3

[J8]   C. Toraman, **F. Şahinuç**, E. H. Yilmaz, and O. Ozcelik. (2023). Impact of Tokenization on Language Models: An Analysis for Turkish. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 22(4), (pp. 1-21). doi: 10.1145/3578707

[J7]   **F. Şahinuç** and A. Koç. (2022). Fractional Fourier Transform Meets Transformer Encoder. *IEEE Signal Processing Letters*, 29, (pp. 2258-2262). doi: 10.1109/LSP.2022.3217975

[J6]   **F. Şahinuç**, E. H. Yilmaz, C. Toraman, and A. Koç. (2022). The Effect of Gender Bias on Hate Speech Detection. *Signal, Image and Video Processing*, 17(4), (pp. 1591–1597). doi: 10.1007/s11760-022-02368-z

[J5]   L. K. Şenel, **F. Şahinuç**, V. Yücesoy, H. Schütze, T. Çukur and A. Koç. (2022). Learning Interpretable Word Embeddings via Bidirectional Alignment of Embedding Dimensions with Semantic Concepts. *Information Processing & Management*, 59(3), (pp. 102925). doi: 10.1016/j.ipm.2022.102925

[J4]   N. Sevim, **F. Şahinuç**, and A. Koç. (2022). Gender Bias in Legal Corpora and Debiasing it. *Natural Language Engineering*, 29(2), (pp. 449-482). doi: 10.1017/S1351324922000122

[J3] C. Toraman, **F. Şahinuç**, E. H. Yilmaz, and I. B. Akkaya. (2022). Understanding Social Engagements: A Comparative Analysis of User and Text Features in Twitter. *Social Network Analysis and Mining*, 12(47). doi: 10.1007/s13278-022-00872-1

[J2] **F. Şahinuç** and A. Koç. (2021). Zipfian Regularities in Non-point Word Representations. *Information Processing & Management*, 58(3), (pp. 102493). doi: 10.1016/j.ipm.2021.102493

[J1] L. K. Şenel, I. Utlu, **F. Şahinuç**, H. M. Ozaktas, and A. Koç. (2020). Imparting Interpretability to Word Embeddings While Preserving Semantic Structure. *Natural Language Engineering*, 27(6), (pp. 721–746). doi: 10.1017/S1351324920000315

Preprints

[P1] **F. Şahinuç**, S. Dutta, and I. Gurevych. (2025). Expert Preference-based Evaluation of Automated Related Work Generation. *arXiv:2508.07955*. URL: https://arxiv.org/abs/2508.07955

Student Supervision

- Shuqi Zhang, M.Sc. Thesis, *Enhancing Synthetic Related Work Paragraphs through LLM Revision*, Technical University of Darmstadt, 2024.

- Jakob Wölfle, B.S. Thesis, *Evaluating the Effectiveness of LLMs as Virtual Research Assistants*, Technical University of Darmstadt, 2024.

Skills

Languages: Turkish (Native), English (Full professional)

Coding: Python (PyTorch, Huggingface, vLLM), C, C++, Java, MATLAB, HTML LaTeX

Awards - Honors

- Best Paper Award: *The Effect of Gender Bias on Hate Speech Detection.* 30th IEEE Signal Processing and Communications Applications Conference (SIU), 2022.

- Ranked 1st in Turkish Harmful Tweet Detection Competition at 2022 CheckThat! Lab on Fighting the COVID-19 Infodemic and Fake News Detection.

- Full Scholarship, Bilkent University (B.S., M.Sc.)

  - High Honor List ($\times$6), Semester GPA > 3.5/4.0.
  - Honor List ($\times$2), 3.0/4.0 < Semester GPA < 3.5/4.0.

- Top 0.003% (59th of 2,000,000) in Turkey's nationwide university entrance exam, 2014.

- Top Scoring Student (96.38/100), Suleyman Demirel Science High School, Afyonkarahisar, Turkey, 2014