

# Gökçe Uludoğan

gokce.uludogan@bogazici.edu.tr  
gokceuludogan.github.io

---

- EXPERIENCE**
- Researcher**, Understanding the Language of Life: Identifying and Characterizing the Language Units in Protein Sequences (ERC-funded) Nov 2023–Present
- Developing protein function predictors by leveraging state-of-the-art protein language models and graph neural networks.
- Researcher**, Utilizing Digital Technology for Social Cohesion (EuropeAid/170389/DD/ACT/Multi) Jun 2023–Present
- Compiling datasets for Turkish hate speech in social and print media, benchmarking and developing new state-of-the-art Turkish hate speech detection models.
- Teaching Assistant**, Department of Computer Engineering, Bogazici University Feb 2019–Present
- Providing academic support, developing teaching materials, and conducting evaluations across a diverse range of courses, including database systems, information retrieval, systems programming, computer organization, and bioinformatics.
- Researcher**, *Chemical Language Processing for Target-based Drug Design* Apr 2020–Oct 2023
- Conducted research on molecular representation learning and developed novel methods for targeted drug design, funded by The Scientific and Technological Research Council of Turkey (TÜBİTAK) with grant number 119E133.

- EDUCATION**
- Ph.D. in Computer Engineering**, Bogazici University (exp.) 2026
- M.S. in Computer Engineering**, Bogazici University 2021
- B.S. in Computer Engineering**, Bogazici University 2018

- SELECTED PUBLICATIONS**
- [1] G. Uludogan, Z. Yirmibesoglu Balal, F. Akkurt, M. Turker, O. Gungor and S. Uskurdarli. TURNA: A Turkish Encoder-Decoder Language Model for Enhanced Understanding and Generation. In: *arXiv preprint*, arXiv:2401.14373, 2024.
- [2] G. Uludogan, S. Dehghan, Inanc Arin, E. Erol, B. Yanikoglu, A. Özgür. Overview of the Hate Speech Detection in Turkish and Arabic Tweets (HSD-2Lang) Shared Task at CASE 2024. In: *Proceedings of the 7th Workshop on Challenges and Applications of Automated Extraction of Sociopolitical Events from Text (CASE 2024)*, p229-233, 2024
- [3] A. B. Temizer, G. Uludogan, R. Ozcelik, T. Koulani, E. Ozkirimli, K. O. Ulgen, N. Karali and A. Ozgur. Exploring Data-Driven Chemical SMILES Tokenization Approaches to Identify Key Protein-Ligand Binding Moieties. In: *Molecular Informatics*, 2024. doi: 10.1002/minf.202300249
- [4] G. Uludogan, E. Ozkirimli, K. O. Ulgen, N. Karali and A. Ozgur. Exploiting pretrained biochemical language models for targeted drug design. In: *Bioinformatics*, 38(Supplement\_2), ii155-ii161, 2022. doi: 10.1093/bioinformatics/btac482

- SERVICES**
- Voluntary Experience**
- The 18th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2024)
  - The 31st Annual Intelligent Systems For Molecular Biology and the 22nd Annual European Conference on Computational Biology (ISMB/ECCB 2023)
  - 27th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2023)
  - 17th Women in Machine Learning Workshop (WiML 2022)