Rıza Özçelik

Ph.D. candidate, Eindhoven University of Technology
Last updated: January 19, 2024

Eindhoven, The Netherlands
r.ozcelik@tue.nl
rizaozcelik.github.io

Short Bio

"To improve human lives." That motto makes me leave my bed every morning to work on my Ph.D. It's also the reason that I picked computer science major for my bachelor's and have been conducting interdisciplinary machine learning research for over 7 years. My research led to 12 publications, 12 presentations, and collaborations ranging from IBM to DeepMind and Stanford. I will stay a scientist for the rest of my life, either in academia or industry. That's why I want to experience groundbreaking research in the industry and seek my second research internship. I hope working on state-space models for generative AI at NVIDIA will be this experience.

Education

- 2022–2026: **Ph.D.**, *Biomedical Engineering, Eindhoven University of Technology, The Netherlands.*
  Thesis: Advancing Chemical Language Models for *De Novo* Drug Design
- 2018–2021: **M.Sc.**, *Computer Engineering, Boğaziçi University, Turkey.*
  GPA: 3.94/4.00 (High Honor). Thesis: Biomolecular Language Processing for Drug - Target Affinity Prediction
- 2014–2018: **B.Sc.**, *Computer Engineering, Boğaziçi University, Turkey.*
  GPA: 3.57/4.00 (High Honor). Thesis: Disease Classification based on Genomic Data with Machine Learning

Employment

- 06.2021 – 12.2021: **Great Minds Research Intern**, *IBM Research, Zurich.*
  Translated natural language prompts to application pipelines. Implemented from scratch.
- 02.2019 – 12.2021: **Teaching Assistant**, *Dept. of Computer Engineering, Boğaziçi University, Istanbul.*
  Taught lectures for more than 200 hours. More than 15% of written feedback called me "the best TA ever."
- 12.2018 – 06.2019: **Research Assistant**, *Scientific Council of Turkey (TUBITAK), Istanbul.*
  Added modules to the open-source Turkish language processing library, `nlptoolkit`, as a core developer.
- 07.2018 – 12.2018: **Data Scientist**, *ING Turkey, Istanbul.*
  Formulated cash flow management as a machine learning problem. Implemented solutions.

Publications

- **Özçelik, Rıza**, Sarah de Ruiter, Emanuele Crusciolo, and Francesca Grisoni. "Chemical Language Modeling with Structured State Spaces" *ChemRxiv preprint doi: 10.26434/chemrxiv-2023-jwmf3*, 2024.
- Budur, Emrah, **Rıza Özçelik**, Dilara Soylu, Omar Khattab, Tunga Güngör, and Christopher Potts. "Building Efficient and Effective OpenQA Systems for Low-Resource Languages." *arXiv preprint arXiv:2401.03590*, 2024.
- **Özçelik, Rıza**, Derek van Tilborg, Jose Jiménez-Luna, and Francesca Grisoni. "Structure-based Drug Discovery with Deep Learning." *ChemBioChem*, 2022.
- Büşra Temizer, Asu, Gökçe Uludoğan, **Rıza Özçelik**, Taha Koulani, Elif Ozkirimli, Kutlu Ulgen, Nilgun Karalı, and Arzucan Özgür. "Exploring Data-Driven Chemical SMILES Tokenization Approaches to Identify Key Protein-Ligand Binding Moieties." *arXiv preprints, arXiv:2210.14642*, 2022. *Accepted for publication in Molecular Informatics.*
- **Özçelik, Rıza**, Alperen Bağ, Berk Atıl, Melih Barsbey, Arzucan Özgür, and Elif Özkırımlı. "DebiasedDTA: A Framework for Improving the Generalizability of Drug-Target Affinity Prediction Models." *Journal of Computational Biology*, 2021.

- Barsbey, Melih, **Rıza Özçelik**, Alperen Bağ, Berk Atıl, Arzucan Özgür, and Elif Özkırımlı. "A Computational Software for Training Robust Drug-Target Affinity Prediction Models: pydebiaseddta." *Journal of Computational Biology*, 2022.
- **Özçelik, Rıza**, Hakime Öztürk, Arzucan Özgür, and Elif Ozkirimli. "ChemBoost: A chemical language based approach for the prediction of protein - ligand binding affinity." *Molecular Informatics*, 2021.
- Budur, Emrah, **Rıza Özçelik**, Tunga Güngör, and Christopher Potts. "Data and Representation for Turkish Language Inference." *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2020.
- Köksal, Abdullatif, Hilal Dönmez, **Rıza Özçelik**, Elif Ozkirimli, and Arzucan Özgür. "Vapur: A Search Engine to Find Related Protein–Compound Pairs in COVID-19 Literature." *Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP*, 2020.
- **Özçelik, Rıza**, Gökçe Uludogan, Selen Parlar, Özge Bakay, Özlem Ergelen and Olcay Taner Yıldız. "User Interface for Turkish Word Network KeNet." *Signal Processing Applications*, 2019.
- Uludoğan, Gökçe, **Rıza Özçelik**, Selen Parlar, Gökhan Ercan and Olcay Taner Yıldız. "User Interfaces for Turkish Natural Language Processing." *Signal Processing Applications*, 2019.
- Ercan, Gökhan, Orçun Erkek, Onur Açıkgöz, **Rıza Özçelik**, Selen Parlar and Olcay Taner Yıldız. "Data Set Generation for Analysing of Turkish Semantic and Sentence Similarity." *3rd International Conference on Computer Science and Engineering (UBMK)*, 2018.

Talks & Presentations

- Selected talk (15min) at ELLIS ML4Molecules Workshop 2023;
- Selected talk (15min) at World Chemistry Congress (IUPAC-CHAINS) 2023;
- Presentation (30min) at Chemical Biology cluster meeting at TUe Fall 2023;
- Selected talk (5min) at Instute of Complex Molecular Systems Symposium 2023;
- Guest lecturer at TUe Advanced Programming Course 2023;
- Invited talk at Oncode Cancer Institute (2023);
- Selected talk at RECOMB2023 conference (15min);
- Poster at ICLR - Machine Learning for Drug Discovery Workshop 2022;
- Two selected talks (15min each) at ISMB 2021;
- Two posters at ISMB 2021;
- Poster at UKQSAR September 2021 meeting;
- Selected talk (15min) and poster in HIBIT 2020.

Honors & Awards

- Google EMEA Academic Scholarship;
- Winner at scientific communication contest;
- Two best bachelor thesis awards (as a thesis supervisor);
- Scientific Council of Turkey (TUBITAK) graduate student scholarship;
- Multiple undergraduate scholarships for 5 years.

References

Available upon request.