Samet Tenekeci

Applied AI Scientist · Computational Biology · Software Engineering

İzmir, Türkiye samettenekeci@gmail.com +90 232 750 78 76 smtnkc.github.io

Summary

I hold a PhD in Computer Engineering with a research focus on natural language processing, integrative data analysis, and computational modeling for biological and software systems. My work spans domain-adapted language models, graph neural networks, and multimodal representation learning, with publications in IEEE Transactions on Computational Biology and Bioinformatics, Computational Biology and Chemistry, Automated Software Engineering, and the Journal of Systems and Software. I have taught core computer engineering courses, supervised students, organized workshops, contributed to national and international research projects, and earned multiple grants supporting applied AI research.

Education

PhD, Computer Engineering *09/2019 – 06/2025*
Izmir Institute of Technology
GPA: 4.00 / 4.00
Thesis: *"Modeling Viral Evolution with Natural Language Processing"*

MSc, Computer Engineering *09/2016 – 08/2019*
Dokuz Eylül University
GPA: 3.86 / 4.00
Thesis: *"Discovering Disease-Causing Genes by Network Analysis"*

BSc, Computer Engineering *09/2009 – 07/2014*
Izmir University
GPA: 3.23 / 4.00
Graduated with honor degree.

Experience

Research Assistant → Lecturer *07/2018 – Present*
Izmir Institute of Technology

- Teaching undergraduate and graduate courses since September 2025
- Conducting research in Software Engineering & AI Research Group
- Conducting research in Data Analytics Research Group
- Assisted in teaching more than 10 computer science courses
- Led administrative tasks within the department as the RA coordinator
- Contributed to the organization of academic conferences and workshops

Software Engineer *02/2017 – 05/2018*
Airties Wireless Networks

- Shipped client-specific features for access points
- Resolved large-scale device overheating issues in production

Teaching Assistant *11/2014 – 08/2016*
Izmir University

- Assisted in teaching and grading core computer science courses
- Conducted laboratory sessions for undergraduate courses
- Contributed to administrative tasks within the department

PUBLICATIONS

1. Tenekeci S, Sezgin E, Tekir S. **A contrastive learning framework for efficient viral escape prediction**. *IEEE Transactions on Computational Biology and Bioinformatics* (Preprint).

2. Tenekeci S, Ünlü H, Gül BA, Keleş D, Küçük M, Demirörs O. **Automating software size measurement from Python code using language models**. *Automated Software Engineering*, 2025.

3. Tenekeci S, Ünlü H. **Peer review and peer assessment as collaborative practices in computer engineering education**. *Computer Supported Cooperative Work (CSCW)* (Preprint).

4. Ünlü H, Tenekeci S, Kennouche DE, Demirörs O. **Automating software size measurement with language models: Insights from industrial case studies**. *Journal of Systems and Software*, 2025.

5. Tenekeci S, Erciyes K. **Distributed approximation algorithms for sorting unsigned genomes by reversals**. *Journal of Global Optimization* (Preprint).

6. Tenekeci S, Ünlü H, Keçeci B, İncir ME, Demirörs O. **Automated software size measurement using multilingual domain-adapted language models**. *Turkish Journal of Electrical Engineering and Computer Sciences* (Preprint).

7. Tenekeci S, Tekir S. **Identifying promoter and enhancer sequences by graph convolutional networks**. *Computational Biology and Chemistry*, 2024.

8. Ünlü H, Tenekeci S, Çiftçi C, Oral İB, Atalay T, Hacaloğlu T, Musaoğlu B, Demirörs O. **Predicting software functional size using natural language processing: An exploratory case study**. *50th Euromicro Conference on Software Engineering and Advanced Applications*, Paris, France, 2024.

9. Tenekeci S, Ünlü H, Dikenelli E, Selçuk U, Kılınç Soylu G, Demirörs O. **Predicting software size and effort from code using natural language processing**. *33rd International Workshop on Software Measurement (IWSM) & 18th International Conference on Software Process and Product Measurement (Mensura)*, Montréal, Canada, 2024.

10. Tekir S, Güzel A, Tenekeci S, Haman BU. **Quote detection: A new task and dataset for NLP**. *7th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, Dubrovnik, Croatia, 2023.

11. Sezerer E, Tenekeci S, Acar A, Baloğlu B, Tekir S. **Author reputation measurement on question and answer sites by the classification of author-generated content**. *International Journal on Artificial Intelligence Tools*, 2021.

12. Ünlü H, Tenekeci S, Yıldız A, Demirörs O. **Event oriented vs object oriented analysis for microservice architecture: An exploratory case study**. *47th Euromicro Conference on Software Engineering and Advanced Applications*, Palermo, Italy, 2021.

13. Çiftçi O, Tenekeci S, Ülgentürk C. **Artist recommendation based on association rule mining and community detection**. *13th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, Valletta, Malta, 2021.

14. Tenekeci S, Işık Z. **Integrative biological network analysis to identify shared genes in metabolic disorders**. *IEEE Transactions on Computational Biology and Bioinformatics*, 2020.

PROJECTS

MorphoMark: Deep learning models for facial morphology analysis *10/2025 – Present*

- Built a deep learning model for 3D facial landmark detection on mesh and point cloud data. Enabled automated anatomical point localization and geometric measurement for skeletal classification in orthodontic and craniofacial analysis.
- Techstack: Python, PyTorch, Git, Blender, MeshLab

OptiMine: AI agent for process optimization in mining and mineral processing *10/2025 – Present*

- Built an intelligent agent that integrates multimodal sensor data (vibration, pressure, temperature, throughput) to monitor operations, predict equipment failures, optimize process parameters, and enhance operational efficiency in mining plants.
- Techstack: Python, LangChain, XGBoost, PostgreSQL, MLflow, Streamlit

AI-Estimator: Automated size measurement for software projects *01/2023 – 01/2025*

- Designed and deployed task-specific BERT models for automated software analytics. Conducted industrial case studies. Funded by the Scientific and Technological Research Council of Türkiye and published in *Journal of Systems and Software*.
- Techstack: Python, PyTorch, Google Vertex AI, HTML, CSS, JS, Git, Docker

CoV-SNN: An efficient framework for viral escape analysis *06/2023 – 06/2025*

- Developed a lightweight protein language model and a contrastive learning framework for viral escape prediction. Achieved 97% accuracy and 125× speedup. Funded by the Council of Higher Education of Türkiye and published in *IEEE TCBB*.
- Techstack: Python, PyTorch, Streamlit, HTML, CSS, JS, Git

GCN4EPI: Graph neural networks to identify gene-regulatory elements *02/2023 – 06/2024*

- Designed a multimodal graph convolutional network model to identify gene regulatory elements. Integrated DNA sequence and Enhancer-Promoter interaction data. Achieved +10% accuracy and published in *Computational Biology and Chemistry*.
- Techstack: Python, PyTorch, TensorFlow, Bash, Git

FastSbR: Distributed approximation algorithms for sorting by reversals *01/2019 – 01/2022*

- Developed global optimization algorithms for an NP-hard problem: Sorting unsigned genomes by reversals. Achieved 5.6-fold speedup compared to baselines. Supported by the Scientific and Technological Research Council of Türkiye.
- Techstack: C, MPI, OpenMP, Bash, SLURM, Git

GO-cluster: Multimodal networks to identify shared disease genes *06/2017 – 06/2019*

- Designed weighted gene co-expression networks integrating protein-protein interaction, gene expression, and gene ontology data. Identified 22 shared genes in three metabolic disorders. Published the method and results in *IEEE TCBB*.
- Techstack: R, Python, GOSemSim, STRINGdb

Skills

Areas: Natural language processing, Bioinformatics, Computational biology, Machine learning, Data science, Graph neural networks, Complex networks, High-performance computing, Parallel algorithms, Software sizing, Software engineering, Sequence analysis, Interdisciplinary research

Tech: Python, PyTorch, TensorFlow, NumPy, Pandas, R, C, HTML, CSS, JavaScript, Git, Docker, Anaconda, Google Vertex AI, Hugging Face, Streamlit, Slurm, Bash, Linux

Organizations

Data Analytics Research Group darg.iyte.edu.tr

- Contributed to project proposals and grant applications
- Received grants from TÜBİTAK, EuroHPC, and YÖK
- Led the development efforts in various projects including CoV-SNN
- Built and managed GPU workstations

Software Engineering & Artificial Intelligence Research Group softw-ai.com

- Contributed to Horizon Europe and British Council grant proposals
- Led AI development efforts in various projects including AI-Estimator
- Contributed to several peer-reviewed publications and conference papers

GitHub Education github.blog

- Organized workshops as the first Campus Advisor in Türkiye.