Samet Tenekeci

NLP & Bioinformatics Researcher

İzmir, Türkiye samettenekeci@gmail.com +90 554 165 50 85 smtnkc.github.io

Summary

I hold a PhD in Computer Engineering with a focus on natural language processing and bioinformatics. As a member of two research groups, I have developed dual expertise in computational biology and software engineering, contributing to academic and industrial projects. My expertise spans interdisciplinary research, large-scale data analysis, multimodal learning, complex networks, graph neural networks, and automated software sizing.

Education

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

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

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

Experience

Founder *2025 – Present*
NeuronAct

- Developing custom AI models, assistants, and agents for businesses and research groups
- Consulting for businesses on AI adoption and transformation processes

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

- Teaching undergraduate and graduate level courses as a lecturer
- Conducting research in Data Analytics Research Group
- Conducting research in Software Engineering & Artificial Intelligence Research Group
- Assisted 10 different computer science courses as a research assistant
- Led departmental organizational tasks as head research assistant

Software Engineer *2017 – 2018*
Airties Wireless Networks

- Designed and implemented client-specific features for access points
- Resolved large-scale device overheating issues in production

Teaching Assistant *2014 – 2016*
Izmir University

- Assisted core computer science courses as a teaching assistant

Software Engineer *2013 – 2014*
BroadAngle

- Developed UI/UX for a health and fitness application (DDP Yoga)

Publications

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

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

3. Tenekeci S, Ünlü H. **Peer review and assessment improves software engineering education: Insights from multiple survey studies**. *ACM Transactions on Computing Education* (Under Review).

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* (Under Review).

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

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* (Under Review).

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/ACM Transactions on Computational Biology and Bioinformatics*, 2020.

Projects

DentAI: AI-based 3D modeling tool for dentistry practices *Proprietary*

- Used graph convolutional networks for 3D dental model segmentation.
- Defined requirements and led software development and management activities.
- Coordinated a team of AI engineers and dentists.
- Techstack: Python, PyTorch, HTML, CSS, JS, Git, Blender, MeshLab

AI-Estimator: Automated size measurement for software projects *Proprietary*

- Curated datasets, trained and deployed task-specific BERT models.
- Contributed to project proposal and conducted industrial case studies.
- Funded by the Scientific and Technological Research Council of Türkiye.
- Techstack: Python, PyTorch, Google Vertex AI, HTML, CSS, JS, Git, Docker

CoV-SNN: An efficient framework for viral escape analysis *Open-Source*

- Developed a task-specific transformer-based protein language model.
- Curated datasets, configured workstations, and deployed AI models.
- Achieved 97% accuracy and 125× speedup in escape prediction.
- Funded by the Council of Higher Education of Türkiye.
- Techstack: Python, PyTorch, Streamlit, HTML, CSS, JS, Git

GCN4EPI: Graph neural networks to identify gene-regulatory elements *Open-Source*

- Designed a multimodal model using graph convolutional neural networks.
- Integrated DNA sequence and Enhancer-Promoter Interaction data.
- Published in Computational Biology and Chemistry.
- Techstack: Python, PyTorch, TensorFlow, Bash, Git

FastSbR: Distributed approximation algorithms for sorting by reversals *Open-Source*

- Developed global optimization algorithms for an NP-hard problem.
- 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 *Open-Source*

- Integrated gene expression, protein-protein interaction, and gene ontology data.
- Identified 22 shared genes in three metabolic disorders.
- Published in IEEE TCBB.
- Techstack: R, Python, Git

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
Led the development team, built and managed GPU workstations.

Software Engineering & Artificial Intelligence Research Group softw-ai.com
Led the AI team and contributed to several publications.

GitHub Education github.blog
Organized workshops as the first Campus Advisor in Türkiye.