

Ehsan Barkhordar

✉ ebarkhordar23@ku.edu.tr

🌐 [ebarkhordar](https://www.linkedin.com/in/ebarkhordar)

📄 [Stack Overflow](https://stackoverflow.com/users/10000000/ebarkhordar)

🔄 [ebarkhordar](https://github.com/ebarkhordar)

🎓 [Google Scholar](https://scholar.google.com/citations?user=ebarkhordar)

Education

Koç University (Ranked #1 in Turkey, QS 2026)

M.Sc. in Computer Science, GPA: 3.8/4.0

Feb 2023 – Feb 2026

Istanbul, Turkey

- Advisors: Prof. Gözde Gül Şahin (primary), Assoc. Prof. Erdem Yörük (co-advisor).

Amirkabir University of Technology (Ranked #3 in Iran, QS 2026)

M.Sc. in Computer Science, GPA: 15.15/20 (3.1/4.0)

Sep 2018 – Mar 2021

Tehran, Iran

Amirkabir University of Technology (Ranked #3 in Iran, QS 2026)

B.Sc. in Computer Science, GPA: 14.79/20 (3.0/4.0)

Sep 2013 – Apr 2018

Tehran, Iran

Publications

Are Non-English Papers Reviewed Fairly? Language-of-Study Bias in NLP Peer Reviews

(under review)

- Ehsan Barkhordar, Abdulfattah Safa, Verena Blaschke, Erika Lombart, Marie-Catherine de Marneffe, Gözde Gül Şahin. *arXiv preprint arXiv:2604.07119*, 2026.

Assessing the Predictive Power of Social Media Data-Fed Large Language Models on Vote Preference

- Ehsan Barkhordar, Sukru Atsizelti. *Companion Publication of the 16th ACM Web Science Conference, 2024*, Stuttgart, Germany.

Why the Unexpected? Dissecting the Political and Economic Bias in Persian Small and Large Language Models

- Ehsan Barkhordar, Surendrabikram Thapa, Ashwarya Maratha, Usman Naseem. *Proceedings of the 3rd Annual Meeting of SIGUL @ LREC-COLING, 2024*, Torino, Italy.

Team Curie at HSD-2Lang 2024: Hate Speech Detection in Turkish and Arabic Tweets using BERT-based Models

- Ehsan Barkhordar, Işık Topçu, Ali Hürriyetoglu. *Proceedings of the 7th Workshop on CASE, 2024*, St. Julians, Malta.

Which Side Are You On? Investigating Politico-Economic Bias in Nepali Language Models

- Surendrabikram Thapa, Kritesh Rauniyar, Ehsan Barkhordar, Hariram Veeramani, Usman Naseem. *Proceedings of ALTA, 2024*, Canberra, Australia.

Clustering of Bank Customers using LSTM-based Encoder-Decoder and Dynamic Time Warping

- Ehsan Barkhordar, Mohammad Hassan Shiralı-Shahreza, Hamid Reza Sadeghi. *arXiv preprint arXiv:2110.11769*, 2021.

Teaching Experience

Koç University

Teaching Assistant – COMP 302: Software Engineering (2 consecutive terms)

Fall 2024 – Spring 2025

Istanbul, Turkey

- Facilitated lectures and labs on software engineering practices including testing, CI/CD, and agile methods.
- Designed assignment rubrics, evaluated semester projects, and organized discussion sections.
- Provided technical mentorship to 60+ students on system design, version control (Git), and collaborative workflows.

Koç University

Teaching Assistant – COMP 100: Introduction to CS and Programming

Spring 2024

Istanbul, Turkey

- Supported an introductory Python programming course with 200+ students enrolled.
- Graded weekly assignments, organized lab sessions, and assisted students in debugging.

Koç University

Teaching Assistant – COMP 491: Senior Design Project

Fall 2023

Istanbul, Turkey

- Advised senior students on project design, implementation, and research methodology.
- Evaluated technical documentation and presentations, and provided structured feedback.

Koç University

Teaching Assistant – COMP 202: Data Structures and Algorithms in Java

Spring 2023

Istanbul, Turkey

- Led weekly lab/tutorial sessions for 50+ undergraduate students on data structures and Java programming.
- Held office hours, graded assignments/projects, and provided one-on-one academic support.

Research & Work Experience

Koç University

Feb 2023 – Feb 2026

Researcher in NLP & Computational Social Science

Istanbul, Turkey

- Led a multimodal hate speech detection project integrating textual and visual modalities.
- Authored three first-author papers in computational social science.
- Reviewed 100+ papers for ARR, EMNLP, ICLR, and related venues.
- Produced 65+ weekly progress reports for collaborative research tracking.
- Trained and fine-tuned transformer-based models (BERT, RoBERTa, GPT) for NER and related tasks.
- Supervised two research interns, providing mentorship and task delegation.

Pishro Net Energy

Aug 2022 – Jan 2023

Machine Learning Engineer

Tehran, Iran

- Designed and deployed end-to-end ML pipelines for real-time energy monitoring across 1,000+ sensor streams (10 buildings), reducing forecasting error by 21%.
- Created data-driven dashboards for 150+ KPIs to support facility managers in decision-making.

Sokan Platform

Apr 2021 – Feb 2022

Machine Learning Engineer

Tehran, Iran

- Built predictive models for customer analytics, including churn prediction (82% accuracy) and cohort-based sales insights.
- Optimized large-scale data pipelines (1B+ records) using TensorFlow, Kubeflow, and Mini-batch K-means.
- Applied FP-Growth to optimize product bundling strategy, improving cross-sell opportunities and data-driven merchandising decisions.

Sokan Platform

Sep 2020 – Apr 2021

Senior Software Engineer

Tehran, Iran

- Implemented asynchronous task management in Django with Celery, reducing computational overhead by 40%.
- Accelerated analytics retrieval speed 5x via ClickHouse migration.
- Delivered 50+ GraphQL APIs using Django Graphene, reducing frontend data-fetching overhead by 35% and consolidating multiple REST endpoints.
- Architected and deployed automated CI/CD pipelines with Docker and GitLab, cutting deployment time by 60% and enabling 3x more frequent releases.

Bale Messenger

Jan 2018 – Feb 2020

Software Engineer

Tehran, Iran

- Contributed to the development of a cross-platform messaging application with 8M+ users.
- Created a Python-based bot SDK adopted by 200+ developers, powering 500+ bots and broadening accessibility for both developers and non-developers.
- Participated in migration to Golang microservices, achieving >7x performance gains.
- Implemented monitoring pipelines with Prometheus and Grafana to track 20+ key system metrics.

Skills

Programming: Python (Advanced, OOP, Async), SQL (Analytical & OLTP), Bash/Shell

Python Frameworks: PyTorch, Hugging Face (Transformers, Datasets, Accelerate, PEFT), Scikit-learn, Pandas, NumPy

GenAI & NLP: LLMs (Fine-tuning, RAG), Embeddings & Retrieval, Prompt Engineering, LLM Reasoning, LLM Evaluation, Text Classification, Multilingual NLP

Research Tools: LaTeX (ACL/IEEE/ACM styles), Overleaf, Git/GitHub, Jupyter, MLflow/W&B, Experiment Tracking, Reproducibility Pipelines

Databases & Analytics: PostgreSQL (Production, Performance Tuning), MongoDB, ClickHouse

Backend & APIs: FastAPI, REST APIs, GraphQL, Celery, Redis, Auth (JWT/OAuth2), HTML/CSS, JavaScript

Dev Tools: Git/GitHub, Docker, Jupyter, Cursor

Languages

Persian: Native

English: Fluent – TOEFL iBT 103/120 (R:26, L:28, S:25, W:24), Dec 2024

Turkish: Beginner

Selected Projects

LLM Collusion and Identifiability in Code Generation | *May 2025 - Present*

- Investigates whether large language models can recognize or attribute code authored by themselves or other LLMs.
- Implemented three attribution tasks: Self-Recognition, Target Identification, and Full Attribution.
- Evaluated across multiple benchmarks (MBPP, HumanEval, DS-1000) with code generated by GPT-5, Claude, Gemini, Grok, and DeepSeek.
- Fine-tuned lightweight judge models to identify model-specific code signatures, highlighting risks of collusion and leakage in code evaluation.

Multimodal Framework for Detecting Implicit Hate Speech on Social Media | *Sept 2024 - Apr 2025*

- Designed a hybrid framework combining text encoders (BERT, LLaMA) with vision backbones (Swin Transformer, ViT) to detect subtle hate speech.
- Improved robustness via focal loss, progressive unfreezing, and adversarial training.
- Demonstrated superior performance on multimodal hate speech datasets, addressing explainability and ethical implications.

Named Entity Recognition in Turkish Tweets (Yildirim Platform) | *Jun 2024 - Sept 2024*

- Developed an annotation platform (Yildirim) for labeling named entities in Turkish tweets.
- Annotated 5,000 tweets from a 700M+ tweet dataset to improve NER performance on low-resource, informal language.

Procedural Text Understanding via Python Code Representation | *Jul 2023 - Sept 2023*

- Explored Python code representation as an intermediate reasoning layer for LLMs.
- Code-based prompting improved accuracy from 68.29% to 86.76%.

Intent Prediction and Slot Filling with Alexa MASSIVE Dataset | *Feb 2022 - Apr 2022*

- Participated in the Massively Multilingual NLU 2022 shared task.
- Trained multilingual intent and slot-filling models on the Alexa MASSIVE dataset using PyTorch.

Shilling Attack Detection in Recommender Systems | *May 2020*

- Investigated detection of profile injection (shilling) attacks in recommender systems.
- Extracted RDMA, WDA, and TF-IDF features; evaluated multiple classifiers (KNN, Random Forest, SVM).

Honors and Awards

2nd Rank – Hate Speech Detection in Turkish and Arabic Tweets (HSD-2Lang)	2024
Invited Participant – Lisbon Machine Learning School (LxMLS 2023), Lisbon, Portugal	2023
One of 5 Selected Teams – Fusion Startup Event (sponsored by divar.ir)	2022
28th Place / 1608 Participants – Django/Python Code Cup, quera.ir	2020
2nd Team – Hackathon, Amirkabir University	2019
Winner – BOT-CUP, Sharif University of Technology	2018