Melika Baghi

mbaghi3@gatech.edu | melikabaghi.github.io | Google Scholar | linkedin.com/in/melika-baghi | github.com/MelikaBaghi | Atlanta, GA

Summary

Ph.D. Candidate in Industrial & Systems Engineering at Georgia Tech with an M.S. in Statistics. Work covers experiment design, ETL and feature pipelines, predictive modeling and statistical validation, through to production-oriented pipelines. Seeking data scientist and applied machine learning roles.

Technical Skills

Programming & data: Python, SQL, PySpark, R, scikit-learn, pandas, NumPy, SciPy, Airflow, MLflow

Machine learning: scikit-learn, XGBoost, PyTorch, ensemble methods, deep learning, representation and multimodal learning

Statistics: experiment design, hypothesis testing, nonparametric methods, calibration, uncertainty quantification, causal and time-series modeling

Delivery: ETL pipelines, model tracking, dashboards (Power BI), Git, Linux, GPU/HPC, cloud computing (AWS)

Experience

Graduate Research Assistant — Georgia Institute of Technology *Aug 2023 – Present*

H. Milton Stewart School of Industrial & Systems Engineering · Advisors: K. Paynabar, X. Liu

- Develop machine learning and statistical methods for high-dimensional data, end to end from problem formulation through validation, reproducible code and publication.
- Built a constrained gradient-boosting estimator that keeps predictions valid during fitting rather than correcting them afterwards, cutting reconstruction error 5× against interpolation baselines; released as an installable open-source package.
- Built a tensor-structured surrogate that reduced interpolation cost 35.6×–564× and prediction time 4×–108× while matching baseline accuracy within 0.35 percentage points.

Teaching Assistant, then Head Teaching Assistant — ISyE 6525: High-Dimensional Data Analytics *Jan 2024 – Present*

Georgia Institute of Technology

- Lead the TA team for a graduate course of 150+ students per semester on dimensionality reduction, sparse modeling and high-dimensional inference. Named **Outstanding Online TA**, 2025.

Course Instructor — Artificial Intelligence Laboratory *Aug 2022 – Feb 2023*

Royan Institute, Tehran

- Designed and taught a project-based applied AI curriculum, supervising student projects from problem framing through evaluation.

Selected Projects

- **Hospital capacity and patient flow** — hybrid agent-based and discrete-event simulation of a hospital serving two patient populations. Cut the tourist queue wait **82%** (13.7 to 2.4 days) and the admission queue **43%** (25.9 to 14.8 days, 30 replications, $p < 0.001$), with a 256-run fractional factorial design over 16 operational factors. *AnyLogic, Python*
- **Predictive maintenance at scale** — end-to-end pipeline for sensor-based fault detection. ARIMA and LSTM forecasting with ensemble anomaly classifiers over multivariate telemetry gave over **15 hours** of advance warning; deployed on Spark for low-latency inference. *Python, PySpark*
- **End-to-end ML pipelines** — CRISP-DM workflows over simulation, sensor and transactional data. ETL in PySpark and SQL, model tracking automated with Airflow and MLflow, results delivered as dashboards in Python, R and Power BI.
- **Clinical risk from EHR** — adverse drug event prediction from longitudinal records, compared across regularized regression, tree ensembles and neural models, with calibration and fairness checks across demographic groups. *Python*
- **Credit risk modeling** — ensemble classifiers for consumer loan default, evaluated on ranking performance and probability calibration rather than accuracy alone. *Python*
- **Cross-national statistical inference** — nonparametric study of internet access against female employment and management representation across 92 countries; Mann–Whitney tests with Bonferroni correction and Anderson–Darling normality diagnostics. *Python*

Education

Ph.D., Industrial & Systems Engineering (System Informatics & Control), Georgia Institute of Technology *Aug 2023 – Present*

M.S., Statistics, Georgia Institute of Technology *Aug 2023 – May 2026*

M.S., Industrial Engineering — Healthcare Systems, Amirkabir University of Technology; GPA 4.0/4.0 *2021 – 2023*

B.S., Industrial Engineering, Amirkabir University of Technology; GPA 3.91/4.0 *2018 – 2021*

B.S., Aerospace Engineering, Amirkabir University of Technology; GPA 3.79/4.0 *2016 – 2020*

Selected Publications

Four preprints and manuscripts under review, including arXiv:2605.04130 (constrained gradient boosting), arXiv:2608.15520 (adaptive modality acquisition), arXiv:2605.19139 (hospital simulation) and arXiv:2609.01761 (delayed bandits). Full list: Google Scholar.