

Tahereh Fahi

✉ tahereh.fahi1@gmail.com
Toronto, Canada

⌚ github.com/tahereh-fahi
㏌ linkedin.com/in/tahereh-fahi

Education

Massachusetts Institute of Technology (MIT)

MicroMaster
Statistics and Data Science
GPA: 4/4

Sep 2022 – Apr 2024

Tarbiat Modares University, Tehran, Iran

Master of Science
Information Technology Engineering- IT Systems
Ranked 1st
GPA: 3.92/4

Aug 2015 – May 2017

K.N. Toosi University of Technology, Tehran, Iran

Bachelor of Science
Industrial Engineering- System Analysis
Ranked 1st
GPA: 3.35/4

Sep 2009 – Oct 2013

Research Interests

- Business Data Analytics
- ML & Optimization
- Human–AI Interaction
- Data-Driven Decision-Making
- Applications in Finance, Bioinformatics & Network Science
- Statistical & Evolutionary Models
- Operations Management
- Engineering Statistics
- Visualization

Experiences

Research Experience.....

Tarbiat Modares University Tehran, Iran
Graduate Research Assistant Aug 2016 – Aug 2017

“Integrating a Local Feature Selection and a Modified PLR Method for Stock Trading Points Prediction” ([GitHub](#))

- Addressed the challenge of non-stationary, noisy financial time series where global feature selection often fails to capture localized market dynamics.
- Developed a hybrid model integrating Piecewise Linear Representation (PLR) for time-series segmentation with Local Feature Selection (LFS) to identify the most relevant technical indicators within each market regime.
- Validated the approach on real-world stock market datasets, achieving more accurate turning-point (peak/trough) detection and improved trading signal precision with lower computational cost compared to global selection methods.

Industry Experience.....

Roboself.ai Toronto, Canada

Data Scientist Nov 2018 – Present

- Developed algorithms to detect turning points in financial time series, improving early trend reversal prediction.
- Designed hybrid AI models combining Neural Networks, GA, PSO, and NSGA-II for multi-objective optimization.
- Applied time series analysis and signal processing for feature extraction, noise reduction, and model reliability

Skillset

Languages: Python, R, C++, C#, MATLAB, MQL (a programming language for developing automated trading systems on MetaTrader 4 and 5)

Machine Learning Algorithms: Decision Tree, SVM, Linear Regression, Linear Programming, Clustering, Bayesian, Deep Learning, Gaussian Process, Convolutional Neural Network, RL (Q- Learning)

Machine Learning Technologies: TensorFlow, PyTorch, Keras, Pandas, Numpy, SciKit Learn, Matplotlib, Seaborn

Software Development Technologies: SQL, NoSQL

Big Data and Cloud Technologies: Microsoft Azure, Google Cloud

Computer Engineering: Data Structures, Algorithms, Object Oriented Programming

Projects

Deep Learning Specialization Projects.....

Generative Adversarial Networks for Realistic Image Synthesis

- Tackled the challenge of generating realistic synthetic imagery without labeled datasets.
- Built and trained PyTorch GAN models to improve adversarial stability and latent-space representation.
- Achieved visually consistent outputs and enhanced feature diversity across generated samples.

Convolutional Neural Networks for Object Detection and Image Generation

- Addressed limitations of traditional classifiers in handling image spatial hierarchies.
- Designed CNN architectures with transfer learning and residual blocks for detection and segmentation tasks.
- Improved recognition accuracy and generalization through YOLO-based detection and style transfer.

Sequence Modeling with RNNs and Attention Mechanisms for Natural Language Processing

- Confronted with the difficulty of modeling sequential linguistic dependencies in text.
- Trained RNN, LSTM, and attention-based models for translation, sentiment, and trigger-word tasks.
- Demonstrated higher accuracy and contextual retention over baseline n-gram and bag-of-words models.

Deep Learning Approaches for Emotion Recognition in Social Media Text

- Sought to detect emotional tone and polarity in informal, high-noise tweet data.

- Engineered tokenization and embedding layers with deep classifiers for multi-label emotion prediction.
- Produced robust classification results that outperformed classical sentiment baselines.

Optimization and Regularization Strategies for Neural Network Training

- Faced issues of slow convergence and overfitting in deep model training.
- Applied initialization schemes, dropout, L2 regularization, and batch normalization in TensorFlow.
- Enhanced model stability and reduced training error by >25% compared to untuned networks.

Bias–Variance Trade-Off and Transfer Learning Strategies in Machine Learning

- Identified challenges of model under/overfitting and data mismatch in ML pipelines.
- Diagnosed errors and optimized architectures using ML-strategy frameworks and transfer learning.
- Reduced generalization error and accelerated deployment by improving model selection decisions.

Implementation of Deep Neural Networks for Structured and Image Data

- Needed to build neural networks from scratch to understand backpropagation mechanics.
- Coded multilayer perceptrons and logistic regression models in NumPy.
- Verified correctness through gradient checks and achieved high accuracy on 2D planar datasets.

MIT MicroMaster Projects

Q-Learning with Function Approximation for Sparse-Reward Text-Based Games ([GitHub](#))

- Aimed to teach an agent to learn optimal policies under sparse reward conditions.
- Implemented Q-learning with linear approximation and adaptive exploration.
- Achieved stable convergence and measurable improvement in cumulative reward rates.

Progressive Digit Recognition on MNIST: From Linear Models to Convolutional Neural Networks

- Investigated performance limits of linear classifiers on handwritten digits.
- Constructed a progressive pipeline from logistic regression to CNN architectures with dropout.
- Attained >98% accuracy while improving model robustness and computational efficiency.

Dimensionality Reduction and Classification of Single-Cell RNA-Seq Data ([GitHub](#))

- Addressed the problem of analyzing noisy, high-dimensional genomic datasets.
- Applied PCA and t-SNE for dimensionality reduction and clustered cells using supervised learning.
- Discovered distinct cell types and validated grouping accuracy with cross-validation metrics.

Comparative Graph Analysis of Facebook and Twitter Networks Using Structural and Statistical Models ([GitHub](#))

- Explored assortativity and topological differences between social networks.

- Computed degree distributions, centralities, and power-law fits to evaluate network models.
- Found significant cross-platform structural variations that supported hypotheses on influence dynamics.

Graph-Based Centrality and Community Detection in Criminal and Citation Networks

- Tackled identification of key actors and clusters within complex networks.
- Applied centrality metrics and community detection algorithms on graph datasets.
- Revealed critical nodes and relationships informing intervention and recommendation strategies.

Collaborative Filtering with Gaussian Mixture Models for Sparse Movie Rating Prediction

- Confronted the problem of missing user ratings in recommendation systems.
- Modeled latent preferences using Gaussian mixture models with expectation-maximization.
- Improved prediction consistency and reduced BIC-based model penalty under sparse data.

Sentiment Classification of Amazon Product Reviews Using Logistic Regression

- Needed to extract consumer sentiment from large text corpora.
- Implemented bag-of-words features and regularized logistic regression for classification.
- Achieved high F1 scores and validated model generalization through cross-validation.

Forecasting CO₂ Concentrations and Economic Indicators Using ARIMA Models ([GitHub](#))

- Faced non-stationary patterns in environmental and economic time series.
- Built ARIMA models with residual diagnostics and forecast validation.
- Produced accurate short-term forecasts validated against observed measurements.

CPI and BER Inflation Data Analysis with External Regressors and Model Improvements ([GitHub](#))

- Investigated relationship between CPI inflation and BER expectations for forecasting.
- Integrated external regressors into ARIMAX models and tested cross-correlation effects.
- Enhanced predictive accuracy and model stability across validation horizons.

Gaussian Process Models for Spatial Prediction of Environmental Data ([GitHub](#))

- Aimed to predict spatial variables (such as ocean currents and temperature fields).
- Built Gaussian Process models with custom covariance kernels and hyperparameter optimization.
- Accurately interpolated flow patterns from Philippine Archipelago data with low prediction error.

TensorFlow Course Projects.....

Design and Optimization of Multilayer Feedforward Neural Architectures Using TensorFlow

- Addressed need for efficient model design across diverse data types.
- Constructed and optimized deep feedforward networks with TensorFlow.
- Improved classification accuracy and training efficiency through architecture tuning.

Advanced Convolutional Frameworks for Image Recognition and Feature Extraction in TensorFlow

- Tackled feature extraction limitations in raw image classification tasks.
- Implemented CNN architecture with pooling and regularization to enhance robustness.
- Increased test accuracy and reduced overfitting through parameter refinement.

Leveraging Pretrained Deep Models for Domain-Specific Applications via Transfer Learning

- Faced data scarcity in target domains requiring high-accuracy models.
- Fine-tuned pretrained CNNs to specialized datasets using transfer learning.
- Achieved significant performance gain with minimal training time and data usage.

Related Extra Courses

- | | |
|---|------|
| ○ Machine Learning on Google Cloud , Coursera | 2020 |
| Completed a 5-course program covering Vertex AI, AutoML, BigQuery ML, custom model training with Docker, feature engineering, and hyperparameter tuning with Vertex Vizier. Gained hands-on experience in distributed TensorFlow, end-to-end ML pipeline development, and model deployment and optimization on Google Cloud Platform. | |
| ○ Deep Learning Specialization , Coursera (Offered by deeplearning.ai) | 2020 |
| Neural Network architectures (CNNs, RNNs, LSTMs, Transformers) using Python and TensorFlow for real-world applications and ML project structuring) | |
| ○ Machine Learning , Coursera (Stanford University) | 2024 |

Related Attended Courses

Graduate Courses:

- Machine learning
- Data Mining & Knowledge Discovery
- Information & IT
- Web Architecture & Programming

Undergraduate Courses:

- Systems Analysis
- Computer Programming
- Engineering Statistics
- Computer Applications in IE

Awards

- Ranked **99th** among approximately 32,000 participants in the Information Technology Engineering National M.Sc. Entrance Examination for universities of Iran 2015
- Ranked among top **7%** among more than 260000 participants in the 'Iranian National universities' entrance exam for B.Sc. degree 2009
- Selected as a qualified person at the first stage of "Iranian National Computer Olympiad" 2006

Language Proficiency

English: Full professional proficiency: CELPIP–General LS: Listening 10/12, Speaking 7/12

Persian: Native or bilingual proficiency

REFERENCES

- **Dr. Seyed Kamal Chaharsooghi**, Professor at Tarbiat Modares University, Department of Industrial and Systems Engineering, Tehran, Iran.
Email: skch@modares.ac.ir
- **Dr. Milad Jasemi**, Professor at University of Montevallo, Assistant Professor of Data Analytics/Department Chair
Email: emjasemiz@montevallo.edu