

# Tahereh Fahi

📍 Toronto, Canada

✉ tahereh.fahi1@gmail.com

🌐 tahereh-fahi.github.io

🐙 [github.com/tahereh-fahi](https://github.com/tahereh-fahi)

in [linkedin.com/in/tahereh-fahi](https://www.linkedin.com/in/tahereh-fahi)

## Education

---

### Massachusetts Institute of Technology (MIT)

*MicroMaster*

Statistics and Data Science

Sep 2022 – Apr 2024

**GPA: 4/4**

### Tarbiat Modares University, Tehran, Iran

*Master of Science*

Aug 2015 – May 2017

Information Technology Engineering- IT Systems

**Ranked 1<sup>st</sup>**

**GPA: 3.92/4**

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

*Bachelor of Science*

Sep 2009 – Oct 2013

Industrial Engineering- System Analysis

**Ranked 1<sup>st</sup>**

**GPA: 3.35/4**

## Research Interests

---

- Business Data Analytics
- Engineering Statistics
- Data-Driven Decision-Making
- ML for Multimodal Sensor Data
- Human–AI Interaction
- Predictive Modeling in Health and Behavioral Data
- Modeling Biological and Physiological Data
- Applications in Finance, Bioinformatics & Networks

## Experiences

---

### 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](#))

- Tackled non-stationary, noisy financial time series by focusing on localized market behavior.
- Combined PLR-based time-series segmentation with Local Feature Selection to pick signals.

### 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 NN, GA, PSO, and NSGA-II for multi-objective optimization.

## Skillset

---

**Languages:** Python, R, C++, C#, MATLAB, MQL

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

**ML Technologies:** TensorFlow, PyTorch, Keras, Pandas, 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 challenge of generating realistic synthetic imagery without labeled datasets.
- Built and trained PyTorch GAN models to improve adversarial stability.
- Achieved visually consistent outputs and enhanced feature diversity across 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.
- Improved accuracy and generalization through YOLO-based detection.

#### ***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.

#### ***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.

#### ***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 transfer learning.

#### ***Implementation of Deep Neural Networks for Structured and Image Data***

- building neural networks from scratch to understand backpropagation mechanics.

- Verified correctness through gradient checks.

## **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 CNN***

- Investigated performance limits of linear classifiers on handwritten digits.
- Constructed a 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.
- Discovered distinct cell types 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.

### ***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.

### ***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.

### ***Forecasting CO<sub>2</sub> 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.

### ***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 interpolated flow patterns from Philippine Archipelago data.

## TensorFlow Course Projects.....

**Design and Optimization of Multilayer Feedforward Neural Architectures Using TensorFlow**

**Advanced Convolutional Frameworks for Image Recognition in TensorFlow**

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

### Related Extra Courses

- o **Machine Learning on Google Cloud**, Coursera 2020
- o **Deep Learning Specialization**, Coursera (Offered by deeplearning.ai) 2020
- o **Machine Learning**, Coursera (Stanford University) 2024

### Related Attended Courses

#### **Graduate Courses:**

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

#### **Undergraduate Courses:**

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

### Awards

- o Ranked **99<sup>th</sup>** among approximately 32,000 participants in the Information Technology Engineering National M.Sc. Entrance Examination for universities of Iran 2015
- o Ranked among top **7%** among more than 260000 participants in the 'Iranian National universities' entrance exam for B.Sc. degree 2009
- o 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

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