# **Masood** Dastan

Data Scientist | Economist

## Experience

### Data Science Fellow | General Assembly

**May 2023 – Aug 2023 | Remote**

- 400+ hour training program focusing on developing skills to analyze, interpret, and effectively communicate data-driven insights from massive data sets, and to predict what happens next through predictive modeling and pattern recognition.
- Executed 20 labs and 6 projects, employing SQL for data cleaning, manipulation, and extraction and Python, R, and Tableau for data visualization and statistical modeling.
- Mastered machine learning with methods such as cross-validation, decision trees, ensemble models, SVMs, clustering, and NLP, alongside deep learning expertise in DNNs, RNNs, and CNNs.

### Assistant Professor | The University of Texas at El Paso

**Sep 2020 – current | El Paso, TX**

- Published peer-reviewed papers that utilize advanced statistical techniques to analyze cross-sectional firm data from 69 developing countries.
- Developed business analytics curricula encompassing classification, regression, hypothesis testing, causal inference, time series analysis, data cleaning, EDA, visualization best practices, and SQL querying.

### Visiting Assistant Professor | Texas A&M International University

**Aug 2019 – Aug 2020 | Laredo, TX**

- Published peer-reviewed papers on the interrelationships between economic growth, financing, and energy consumption using advanced time series methods.
- Delivered lessons on econometric modeling, encompassing topics such as ordinary least squares, instrumental variable estimation, limited dependent variable models, time series models, and panel data methods.

## Projects

### Investigating the Impact of Federal Open Market Committee Meetings on Stock Market Performance | SpaCy, NLTK, Scikit-Learn, Transfer Learning, Deep Learning

- Leveraged advanced NLP techniques with SpaCy, NLTK, Scikit-Learn, and deep learning models to analyze and extract invaluable insights from FOMC meeting minutes.
- Employed topic modeling (LDA) to discern significant US economic concerns deliberated during FOMC meetings.
- Executed sentiment analysis using zero-shot text classification to assess FOMC meetings.

### Reddit Post Classification | Python, NLTK, Scikit-Learn

- Employed Python Reddit API Wrapper to collect posts from two similar Subreddits, Personal Finance and Investing.
- Implemented hyperparameter tuning via randomized search, encompassing a spectrum of classification methods like Naïve Bayes, Logistic Regression, Random Forest, Support Vector Machine, and Gradient Boosting.

### Predicting Credit Card Delinquency | Python, Scikit-Learn, Deep Learning, Gradient Boosting

- Spearheaded data cleaning, exploratory data analysis (EDA), and feature engineering to optimize data quality.
- Implemented diverse machine learning models for credit delinquency prediction, culminating in a finely tuned Adaboost model with an 80% accuracy in the test data—signifying a remarkable 27 percentage point improvement over the benchmark.

## Contact Info


masooddastan@gmail.com

(470) 775-6115

linkedin.com/in/masooddastan

github.com/MasoodDastan


### PORTFOLIO

masooddastan.github.io

## Skills

### INTERESTS/EXPERTISE

- Machine Learning
- Deep Learning
- Causal Inference
- Econometrics
- Natural Language Processing
- Time Series Analysis

### PROGRAMMING

**Proficient:**

Python • SQL • STATA

**Experienced:**

Tableau • R • SPSS

### DATA VISUALIZATION

Matplotlib • Seaborn • GGplot2
Tableau • Plotly

### PLATFORMS/TOOLS

Git • GitHub • AWS • Spark •
LaTeX • Google Cloud (GCP)

## Education

### GENERAL ASSEMBLY

Data Science Immersive | Certificate

### GEORGIA STATE UNIVERSITY

Ph.D. in Economics

### UNIVERSITY OF TEHRAN

M.A. in Energy Economics

### UNIVERSITY OF TEHRAN

B.A. in Economics

## Certificates

### GOOGLE DATA ANALYTICS

Coursera

### PYTHON FOR DATA SCIENCE

Udemy

### MYSQL BOOTCAMP

Udemy