Masood Dastan

Data Scientist, Economist
(470)775.6115·masooddastan@gmail.com·Washington DC·**Portfolio:** masooddastan.github.io

Dedicated Data Scientist and Economist with a rigorous academic background and hands-on data analytics expertise. Proficient in applying advanced statistical techniques and machine learning models to interpret complex datasets and deliver actionable insights. Published researcher and university professor with a commitment to lifelong learning.

SKILLS

Data Analysis Skills: Machine Learning, Econometrics, Time Series Analysis, Causal Analysis, Deep Learning, Natural Language Processing, Classification, Regression, Unsupervised Learning
Language Proficiencies: Python, SQL, R, Tableau, STATA
Data Visualization: Matplotlib, Seaborn, Plotly, Tableau, GGplot2
Other Technologies and Tools: Git, GitHub, JupyterLab, AWS, Spark, LaTeX, Google Colab, SPSS, Slack, Microsoft Office

DATA SCIENCE PROJECTS

Investigating the Impact of Federal Open Market Committee (FOMC) 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.
- Utilized NLTK, CountVectorizer, TF-IDF, and Scikit-Learn models for data preprocessing and data modeling.
- 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.
- Achieved exceptional model performance, surpassing benchmark accuracy rates by an impressive 30 percentage points and attaining a remarkable 97% AUC score.

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

- Led a collaborative team effort to create a predictive model for credit card delinquency, ensuring effective version control and seamless collaboration using Git and GitHub.
- 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.

Housing Market Analysis | Python, Scikit-Learn

- Enhanced data quality by rectifying missing values, outliers, and incorrect data types.
- Utilized linear regression, Lasso, Ridge, and Elastic Net models. Implemented forward selection to identify impactful variables, resulting in a model surpassing the benchmark by 75%.

EXPERIENCE

Data Science Fellow, General Assembly | Remote | May 2023 – Aug 2023

- 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 | TX, USA | Sep 2020 - present

- Instruct and develop courses in machine learning, business analytics, and economics; created and taught a curriculum encompassing classification, regression, hypothesis testing, causal inference, time series analysis, data cleaning, EDA, visualization best practices, and SQL querying.
- Authored and published a peer-reviewed paper that applies advanced statistical techniques to address the problem of missing data in assessing the effect of corruption on firm investment in developing countries.

Visiting Assistant Professor, Texas A&M International University | TX, USA | Sep 2019 – Aug 2020

- Instructed courses in Econometrics, Microeconomics, and Macroeconomics.
- 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.

Publications

- **Understanding the Links Between Firm Size, Exposure to Public Officials, and Firm Corruption**. Under Review
- **The Effect of Corruption on Firm Investment in the Presence of Missing Data.** *American Journal of Economics and Sociology*, 82, 79–93.
- **Electricity Consumption and Economic Growth in OPEC Countries.** *OPEC Energy Review*, 39: 1-16.
- **Analysis of Islamic Banks' Financing and Economic Growth: A Panel Cointegration Approach**, *International Journal of Islamic and Middle Eastern Finance and Management*, Vol. 6 No. 2, pp. 156-172.

EDUCATION

Data Science Immersive | **General Assembly** | Aug 2023
Ph.D. in Economics | **Georgia State University** | Georgia, USA | Aug 2019
M.A. in Energy Economics | **University of Tehran** | Tehran, Iran | May 2013
B.A. in Economics | **University of Tehran** | Tehran, Iran | May 2010

CERTIFICATIONS

Google Data Analytics, Coursera
Python for Data Science and Machine Learning Bootcamp, Udemy
The Ultimate MySQL Bootcamp: Go from SQL Beginner to Expert, Udemy