Masood Dastan

Data Scientist, Economist
(470) 775.6115 · masooddastan@gmail.com · portfolio ·in/ · github.com/

EXPERIENCE

Assistant Professor, The University of Texas at El Paso | TX, USA | Sep 2020 - present

- Instruct and develop undergraduate and graduate courses in machine learning, business analytics, and economics.
- Designed and delivered a comprehensive curriculum that covered a wide array of topics, including supervised learning, unsupervised learning, hypothesis testing, causal inference, data cleaning, exploratory data analysis (EDA), best practices in data visualization, and data querying using SQL.
- Authored and published a peer-reviewed paper employing advanced statistical techniques, emphasizing the importance of addressing missing data when assessing corruption's impact on firm investment in developing countries. | Link

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

- 12-week, 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.

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

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

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 comprehensively.

Reddit Post Classification | Python, NLTK, Scikit-Learn

- Employed Reddit API to gather 10,000+ posts from "Personal Finance" and "Investing" subreddits.
- 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 team to create a predictive model for credit card delinquency.
- Spearheaded data cleaning, EDA, and feature engineering efforts.
- 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%.

EDUCATION

Certificate - 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

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

SKILLS

Language Proficiencies: Python, SQL, R, STATA, Tableau
Other Misc. Technologies: Git, AWS, Spark, LaTeX
Data Analysis Skills: Machine Learning, Econometrics, Time Series Analysis, Causal Analysis, Deep Learning, NLP