Farhad Saffaraval

San Francisco, CA | (925) 336-0159 | fsaffaraval@gmail.com| [LinkedIn](#) | [Scholar](#) | US Permanent Resident

Technical Skills

PROGRAMMING LANGUAGES: Python | SQL | MATLAB | Flask | git

DATA SCIENCE TOOLS: Pandas | NumPy | SpaCy | NLTK | SciPy | Scikit-Learn | PySpark | Image Processing | Data Visualization (Altair, Tableau, Matplotlib) | Feature Engineering | TensorFlow | Keras | Natural Language Processing (NLP) | Statistical Methods

MACHINE LEARNING: Linear Regression | Logistic regression | Time Series Analysis | SVM | Neural Networks (CNN | RNN) | PCA | Random Forest | K-means Clustering | Deep Learning | Transfer learning | Word2vec | XGBoost

Education

Ph.D. in Mechanical Engineering | University of Maryland | Aug 2018

MS in Mechanical Engineering | Washington State University | May 2011

BS in Mechanical Engineering | Amirkabir University | Aug 2008

Experience

The Data Incubator (TDI)   **Sep 2020– Jan 2021**
Data science scholar

- Selected as a skilled scholar to participate in TDI fellowship program. TDI is a highly selective data science cohort, designed for career transition from academia to industry

- Finished seven weekly projects involving machine learning, SQL, web scrapping, time series predictions, sentiment analysis, distributed computing using PySpark, and Image classification employing TensorFlow

- Created an NLP model for the [capstone](#) project to predict the personality type by analyzing a person's social media posts; Started with logistic regression and improved the performance by using Neural Networks and Boosting algorithms

Food and Drug Administration (FDA)   **June 2019 – Sep 2020**
ORISE Fellowship Researcher   **June 2017 – Jan 2018**

- Collected and examined the experimental data to quantify cleanliness of medical devices

- Participated in discussions with senior scientists to identify flaws in medical devices which lead to soilage

- Implemented statistical methods such as student t-test and ANOVA to determine correlations between flowrate and sample cleanness and collaborates as first-author to publish research findings

National Institute of Health (NIH)   **July 2015 – Jan 2016**
Research Intern

- Developed ensemble learning algorithm based on historical data to discover optimized monthly operating strategy for efficient conduction of chillers at NIH Central Utility Plant

University of Maryland   **Aug 2012 – Aug 2018**
Graduate Research Assistant

- Led the investigation of a bio-inspired pumping enhancement in an array of parallel plates; collected, scrubbed, analyzed, visualized, and interpreted velocity kinematic data to propose an efficient new system to pump fluid

- Published and communicated research findings in the form of journal and conference publications and presentations

- Supervised and taught graduate research topics and projects to interested undergrad students

Washington State University Vancouver   **Jan 2009 – May 2011**
Graduate Research Assistant

- Trained innovative function-learning linear regression algorithm to present>1GB of velocity data in one simple formula. I used the [empirical formula](#) in the numerical simulation to predict the collapse of the volcanic column

Selected project:

Text mining of Arabian nights   Under Review for Publication

"Text Mining by Word2vec to Find Word Similarities and Address the Issue of Obscure Origins of Arabian Nights." Generated innovative natural language processing algorithm to quantify the influence of Persian and Arabian culture on a famous tale.