Fred Mubang

Email: fmubang@usf.edu
Machine Learning Engineer / Data Scientist
Linkedin: Fred Mubang
Website: fmubang.github.io

Summary: Computer Science MSc., current Ph.D candidate, and research assistant at University of South Florida. Specialties lie in building Machine Learning pipelines and data analysis.

Education

Ph.D in Computer Science, University of South Florida, GPA: 3.87/4.00 — **Aug 2018 — 2022 (Expected)**

Master of Science, Computer Science, University of South Florida, GPA:3.87/4.00 — **Aug 2018 — May 2021**

Post Bachelor Studies, Computer Science, University of South Florida, GPA:3.7/4.00 — **Aug 2017 — July 2018**

Bachelor of Arts, Music Business, Berklee College of Music — **Aug 2010 — May 2014**

Relevant Courses: Data Mining, Machine Learning, Neural Networks, Advanced Neural Networks, Social Media Mining, Network Science, Natural Language Processing, Intro to AI, Calculus 1-3, Linear Algebra, Probability and Statistics

Skills and Technologies

- **ML Skills:** Classification, Regression, Clustering, NLP, Feature engineering, Dimension reduction techniques
- **Data Analytics:** Cleaning, Manipulation, Scraping, Visualization
- **Statistics:** Inferential and Descriptive Statistics, Bayesian Methods
- **Technologies:** Python, C, C++, Linux, Scikit-learn, Pandas, Tensorflow, Keras, XGBoost, Numpy, Seaborn, Matplotlib, Networkx, Jupyter Notebooks, Excel

Experience

Research Assistant for Department of Defense/DARPA-Funded, Social Simulation Project - Link — **Oct 2017 — Dec 2021**

- Objective: High-fidelity computational simulation and of user activity across different social media platforms such as Twitter and YouTube, among others.
- Built neural network and XGBoost machine learning models to perform time series forecasting and network simulation of user activity in various social media platforms with millions of users. Achieved over 20% improvement over historical baselines in both time series and network simulation tasks.
- Participated in various hackathons against other universities such as USC, Duke, UCF, and UIUC. Model placed in 3rd place out of 60 submissions.
- Performed various data engineering tasks such as cleaning, manipulating, scraping, feature engineering, and visualization of data.
- Performed detailed social network time series analysis of various datasets and created weekly Powerpoint presentations containing data analysis and insights. Used various Python libraries to prepare results such as Pandas, Matplotlib, and Numpy.

AI Bootcamp Instructor — **May 2020 — Aug 2020**

- Taught a course on AutoEncoder Neural Networks at University of South Florida's Summer 2020 AI Bootcamp
- Prepared course materials and lectured.

Data Mining Teaching Assistant at University of South Florida, Computer Science — **Aug 2019 — Dec 2019**

- Taught students course material during office hours; Proctored and graded exams
- Gave lectures on various machine learning concepts such as Decision Trees and Association Rules

Academic Projects (With Links)

Temporal Link Prediction in Twitter (Network Science Final Project - Link) — **Dec 2020**

- Created Logistic Regression models to predict future user to user interactions in a Twitter dataset.
- Used Pandas, Networkx, and Numpy Libraries to engineer multiple features for these models

Predicting Cooking States with ResNet Neural Networks (Deep Learning Midterm Project - Link) — **March 2019**

- Trained Resnet neural networks to predict the cooking states of various food items shown in photographs (e.g. *sliced, peeled, etc*)
- Used Keras and Tensorflow libraries for feature processing and creating the networks

Publications (With Links)

- **Mubang, F.**, Hall, L.O. *VAM: An End-to-End Simulator for Time Series Regression and Temporal Link Prediction in Social Media Networks. IEEE Transactions on Social Computing* (2021 - Accepted with Revisions) - Link
- Liu, R., **Mubang, F.** *Simulating Temporal User Activity on Social Networks with Sequence to Sequence Neural Models. IEEE SMC International Conference* (2020) - Link
- Liu, R., **Mubang, F.**, et al. *Predicting Longitudinal User Activity at Fine Time Granularity in Online Collaborative Platforms. IEEE SMC International Conference* (2019) - Link