

# Fred Mubang

Email: fmubang@usf.edu **Machine Learning Researcher / Data Scientist**

Linkedin: Fred Mubang  
Website: fmubang.github.io

**Summary:** Machine Learning Researcher and Data Scientist with 5 years of work experience. I have spent most of that time working for the Defense Advanced Research Projects Agency (DARPA), an agency within the Department of Defense. In addition, I am a Computer Science MSc., current Ph.D candidate, and Researcher at University of South Florida.

## WORK EXPERIENCE

- Machine Learning Researcher/Data Scientist for DoD/DARPA Social Simulation Project - [Link](#)** Oct 2017 — May 2022 (5 yrs.)
- 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 Scikit, Pandas, Matplotlib, and Numpy.
- Data Structures Teaching Assistant at University of South Florida, Computer Science** May 2022 — Present
- Taught students Data Structures and C++ coding. Graded coursework and exams.
- AI Bootcamp Instructor** May 2020 — Aug 2020
- Prepared coursework and taught students various Artificial Intelligence concepts
- Data Mining Teaching Assistant at University of South Florida, Computer Science** Aug 2019 — Dec 2019
- Prepared coursework and taught students various Data Mining concepts. Graded coursework and exams.

## EDUCATION

- Ph.D in Computer Science, University of South Florida, GPA: 3.87/4.00** Aug 2018 — Oct 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

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

## 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* (2022 - In Press) - [Link](#)
- Mubang, F., Hall, L.O.** *Simulating User-Level Twitter Activity with XGBoost and Probabilistic Hybrid Models.* *arXiv preprint arXiv:2202.08964* (2022) - [Link](#)
- Mubang, F., Hall, L.O.** *A Survey of Recent Artificial-Intelligence Driven Frameworks for User-Level Activity Prediction in Github.* *Engineering Applications of Artificial Intelligence* (2022 - Under Review) - [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](#)

## 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