

# UNIVERSITY OF SOUTH FLORIDA

## Defense of a Doctoral Dissertation

Social Media Time Series Forecasting and User-Level Activity Prediction with Gradient Boosting, Deep Learning, and Data Augmentation

by

**Fred Mubang**

For the Ph.D. degree in Computer Science and Engineering

In the overall history of technological innovations, social media has only existed for a brief time, however its influence is undeniable. Researchers have found that it can be used to influence elections, spread health misinformation, and aid with financial pump-and-dump schemes. Keeping all this in mind, it is clear that more research is needed to predict the spread of information on social media in order to combat its malicious use. To that end, in this dissertation, we explore the use of gradient boosting, deep learning, and data augmentation algorithms to perform time series forecasting and user-level activity prediction in social media. We address the myriad of challenges that come with these tasks such as: (1) the inherent uncertainty of predicting the future, (2) the lack of historical data, (3) the selection of the appropriate model for prediction, (4) the large number of users on a given social media platform, (5) the infrequent activity of most users, (6) the selection of appropriate metrics, and (7) the modeling of the arrival and activity of new users.

### Examining Committee

Chris Tsokos, Ph.D., Chairperson

Lawrence O. Hall, Ph.D., Major Professor

Adriana Iamnitchi, Ph.D.

John Skvoretz, Ph.D.

Yu Sun, Ph.D.

Mingyang Li, Ph.D.

Wednesday, October 12<sup>th</sup>, 2022

1:00 PM

Hybrid (ENB 313 and Online)

Please email [fmubang@usf.edu](mailto:fmubang@usf.edu) for more information

THE PUBLIC IS INVITED

### Publications

- 1) **Fred Mubang** and Lawrence O. Hall. 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).
- 2) **Fred Mubang** and Lawrence O. Hall. Simulating New and Old Twitter User Activity with XGBoost and Probabilistic Hybrid Models. 21st International Conference on Machine Learning and Applications, 2022 (Accepted).
- 3) Renhao Liu, **Fred Mubang**, and Lawrence O. Hall. Simulating Temporal User Activity on Social Networks with Sequence to Sequence Neural Models. IEEE SMC International Conference, 2020.
- 4) Renhao Liu, **Fred Mubang**, Lawrence O. Hall, Sameera Horawalavithana, Adriana Iamnitchi, and John Skvoretz.