

Jacob McCabe

<https://www.linkedin.com/in/jacobmccabe023/>

jacobmccabe023@gmail.com | +1 (717) 575-4996

<https://github.com/jamccabe1/Portfolio>

Data Scientist and Analyst inspired by tough problems.

Technical Skills:

Technology: Python • R • C • C++ • Java • SQL • Git • PowerBI • NumPy • SciKit-Learn • TensorFlow

Quantitative: Machine Learning • Time Series Analysis • NLP • Mathematical Modeling • Supervised Learning • Unsupervised Learning • Optimization • SVMs • Neural Networks • Regression • LSTM

Education:

BS Computer Science, Western Washington University, June 2023 (Major GPA: 3.73)

BS Applied Mathematics, Western Washington University, June 2023 (Major GPA: 3.84)

Work Experience:

Research Assistant - PISCES International & WWU

July. 2022 – June 2023

- Trained and tested a logistic regression model in R able to predict student success with 75% accuracy.
- Performed Data Analysis and Feature Engineering to eliminate some of the random noise and reduce the dimensionality of data.
- Wrote technical reports such as a research proposal, project plan, research paper, and data documentation.
- Presented research findings and its applications to audiences > 40 people.

Portfolio:

Research on Education of Security Operations Center Analysts (*Quantitative Analysis: R, SQL, Regression*)

Purpose: Make a supervised learning model to predict the resolution time of tickets flagging suspicious network traffic to improve the education of SOC Analysts.

Outcome: Trained a logistic regression model able to predict ticket resolution time with 75.39% accuracy, 85.33% F-1 Score, 92.24% Recall, and 79.39% Precision.

Technical Stock Analysis with LSTM (*Time Series Analysis: Python, Deep Learning, LSTM*)

Purpose: Create a LSTM model to forecast stock market closing prices for General Electric, using a Simple Moving Average and Exponential Moving Average as benchmark models.

Outcome: With little training, the LSTM model was able to forecast stock prices - scoring RMSE = \$1.74 and MAPE = 1.99% on test data, outperforming benchmark models by 290%.

Bank Marketing Analysis (*Quantitative Analysis: Python, SciKit-Learn, Supervised Learning*)

Purpose: Determine a way to determine which subset of clients are more likely to subscribe to a term deposit from a banking institution's direct marketing campaign.

Outcome: Trained several supervised classification models, and chose the model with the highest F-Score on the test set. Accuracy = 89.9%, F-Score = 0.549 (beta = 0.5).

Machine Learning Algorithms (*Machine Learning: Python, NumPy, Matplotlib, Scikit-Learn, Tensorflow*)

Purpose: Demonstrate a thorough understanding and demonstrate usage of various machine learning algorithms including K-Means, K-Nearest Neighbors, Support Vector Machines, Neural Networks, Decision Trees, Random Forests, and Boosted Tree Ensembles.