Thomas Hur

thur1@binghamton.edu - https://github.com/hurhurhurt - https://www.linkedin.com/in/thomas-hur/

Education

State University of New York, Binghamton **Expected May 2021**
Thomas J. Watson School of Engineering and Applied Sciences GPA: 3.19
Bachelor of Science in Computer Science and Economics

Technical Skills

Languages: Python, Spark, Scala, SQL, Java, Javascript, R, C, C++
Frameworks: PySpark, TensorFlow, Pandas, Numpy, Scikit-Learn, Hadoop, Matplotlib, BeautifulSoup
Coursework: Computer Systems I & II, OOP, Data Structures, Calculus, Linear Algebra, Number Systems

Professional Experience

Xaltius Tech Pte Ltd **Kent Ridge, Singapore**
Data Science Intern June 2019 - August 2019

- Built several financial use cases with PySpark and Scala to market to customers including Kickstarter Success and Financial Auditing
- Constructed multiple models of algorithms such as Support Vector Machines and utilized methods like Gradient Boosting, ML Flow, and Cross Validation to achieve highly accurate models
- Collaborated with marketing interns to design multiple presentations using Canva to showcase machine learning projects to consumers and businesses

TakenMind Organization **Remote**
Data Analytics Intern October 2018 - December 2018

- Discovered key parameters that led to high employee turnover by implementing multiple machine learning models like SVM, Decision Trees, Random Forest, and Naive Bayes
- Utilized multiple classification algorithms on the popular iris flower dataset to label each individual flower as either Iris setosa, Iris virginica or Iris versicolor

Project Experience

Asian Recipe Data Analysis **https://git.io/JeAvs**

- Established a dataset of over 1,400 recipes and 9 features from the Woks of Life, an Asian recipe website, using a Python script developed using Requests, BeautifulSoup4, and JSON.
- Conducted standard data cleaning procedures using Regex and Boolean indexing
- Visualized trends between features with the Matplotlib, Seaborn, and WordCloud libraries

Credit Card Fraud Analysis **https://git.io/JeAvY**

- Administered in-depth analysis of a Kaggle Credit Card Fraud dataset containing over 284,000 transactions and 28 features using PySpark, SQL, and IPython.
- Composed deep learning models with Keras/TensorFlow as well as SVM, LR, DTC, GBT to develop models that correctly identified fraud with 98.12% accuracy
- Optimized an SVM model with K-Fold Validation and Pipelining to boost accuracy further to 98.87%

Competition Experience

Google Coding Competition **Binghamton, New York**
4th place / 27 teams September 2019

- Collaborated with team to find solutions to complicated math puzzles in topics like linear algebra, number theory, and calculus