

Thomas Hur

thur1@binghamton.edu | (917)-345-4511 | hurhurhurt.github.io | linkedin.com/in/thomas-hur

Technical Skills

Computer Languages: Python, SQL (MySQL, Spark SQL), Java, C++

Frameworks: AWS (S3 Bucket, Athena, Redshift), Azure (Databricks, DevOps), Data Science (Sci-kit Learn, TensorFlow, Pandas, Numpy), Visualization (Tableau, Matplotlib, Seaborn)

Relevant Coursework: DS & Algorithms, Probability & Statistics, Linear Algebra, OOP, Operating Systems, Automata Theory

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, Minor in Economics

Professional Experience

McAfee

Plano, Texas

Incoming Data Science Analyst Co-Op

Sep 2020 - Dec 2020

- Generating product insight from data on the Mobile App Product Team

Data Science Analyst Intern

May 2020 - Aug 2020

- Segmented customers through K-Means Clustering to identify key app features and predict customer churn, retention, and value, potentially generating thousands in additional revenue through targeted advertising to new McAfee users
- Delivered end-to-end dashboard visualization with Tableau and SQL to provide stakeholders with App Store and Google Play metric data, reducing time spent by 90% manually browsing through app store data
- Leveraged AWS Athena/S3 Bucket, Azure Databricks, Python, and SQL to analyze success of app campaign messaging

Xaltius Tech Pte Ltd

Kent Ridge, Singapore

Data Science Intern

Jun 2019 - Aug 2019

- Built several financial use cases with PySpark pipelines to market to customers including Kickstarter Success and Auditing
- Constructed multiple models of algorithms such as Support Vector Machines and utilized methods like Gradient Boosting and Cross Validation to achieve robust models in terms of measures like accuracy, misclassification rate, recall, etc
- Collaborated with marketing interns to design multiple presentations using Canva to showcase machine learning projects to consumers and businesses in a comprehensive, easy-to-understand manner

TakenMind Organization

Remote

Data Analyst Intern

Oct 2018 - Dec 2018

- Discovered key parameters that led to high employee turnover by implementing multiple machine learning models like SVM, Decision Trees, Random Forest, Naive Bayes, KNN, Logistic & Linear Regression, etc
 - Utilized multiple classification algorithms on the popular iris flower dataset to predict flower genera with 98% accuracy
-

Project Experience

Asian Recipe Recommender System

git.io/JeAvs

- Established a dataset of over 1,400 recipes and 9 features from the Woks of Life, a popular Asian recipe website, using a Python script developed with Requests, BeautifulSoup4, and JSON that conformed to site's data scraping requirements
- Developed a robust recommender system comparing different recipes utilizing a custom cosine similarity NLP algorithm that compares recipes on the basis of ingredient similarity, as well as average rating and review count
- Visualized trends between features via graphs and diagrams with the Matplotlib, Seaborn, and WordCloud libraries

Credit Card Fraud Prediction

git.io/JeAvY

- Administered in-depth analysis of a Kaggle Credit Card Fraud dataset containing over 284,000 transactions and 28 features using on the Databricks platform, applying languages like Python/PySpark and SQL
 - Applied Principal Component Analysis (PCA) for dimensionality reduction in order to reduce model's chance to overfit
 - Composed and optimized an SVM model with Cross Validation and Ensembling to identify fraud with 98.87% accuracy
-

Competition Experience

Fizzle

devpost.com/software/fizzle

7th Annual HackBU

Feb 2020

- Deployed "Fizzle", a dating website to find your archnemesi using Python, Django, HTML, CSS, and JS
- Created an algorithm relying on quiz results & feature importances to match users with users already in the database
- Leveraged a SQLite backend to create a encrypted login system with the PBKDF2 algorithm and SHA256 hash