MARIE HASEGAWA

(863) 969-7908 ◇ Auburndale, FL
mariehasegawa12@gmail.com ◇ linkedin.com/in/marie-hasegawa-590456203/ ◇ github.com/mhasegawa7045

OBJECTIVE

Magna Cum Laude Data Science major with 1 year of experience at FedEx and Real Estate Bot. Skilled in building Machine Learning models, Data Warehouses, Data Pipelines, ETL Pipelines, and Cloud Implementation Solutions. Currently seeking a full-time position in Data Analysis, Data Engineering, or Software Engineering.

EDUCATION

Bachelor of Science - Data Science: Big Data Analytics, Florida Polytechnic University 2018-2022
Accomplishments: Magna Cum Laude (3.88 GPA) and President's List
Relevant Coursework: Data Warehouse, Machine Learning, Statistical Learning, Data Mining & Text Mining, Time Series & Forecasting, Cloud Implementation
Portfolio: Collaborated in 3 group projects and built 6 personal projects.

SKILLS

Programming Languages: PROFICIENT: *Python, SQL, R Programming, HTML;* NOVICE: *Java, Bash, C/C++;*

Databases: PROFICIENT: *PostgreSQL, MySQL;* INTERMEDIATE: *T-SQL, Azure Cosmos DB, MongoDB, NoSQL, DynamoDB;*

AI/Machine Learning: Exploratory Data Analysis (EDA), Regression, Support Vector Machine (SVM), Clustering, Decision Trees, Random Forests, Ensemble Learning, Bagging, Boosting, Natural Language Processing (NLP)

Libraries: PYTHON: *numPy, pandas, matplotlib, scikit-learn;* R: *tidyr, dplyr, ggplot2, caret, rpart, cluster;*

Tools: SQL Server Management Studio (SSMS), SQL Server Integration Services (SSIS), Microsoft Visual Studio, Tableau, Power BI, Jupyter Notebook, RStudio, GitHub, Microsoft Azure, Microsoft Project, Microsoft Excel, Visual Studio Code (VS Code), Stata, Amazon Web Services (AWS)

EXPERIENCE

IT Data Science and Analysis Intern: Research and Development Jun 2021 - Aug 2021
FedEx Corporation *Harrison, AR (Remote)*

- Developed and optimized automated data pipelines that transformed and loaded millions of records of real-time raw data from the local machines to a centralized database with Azure Cosmos DB and Blob Storage, which enhanced data accuracy and reliability by 20%.
- Created Python scripts and JSON files in VSCode, and developed data trends and data visualizations with Power BI to drive informed decision making.
- Data analysis was performed based on the compiled data in the centralized database and generated possible uses and solutions from this project to stimulate future business growth.

Data Science/Analysis Capstone Contractor Sep 2020 - May 2021
Real Estate Bot *Lakeland, FL (Remote)*

- Utilized a chatbot to collect and store chat flows and messages that were held between the chatbot and the buyer or seller.
- Constructed a flexible NoSQL database that stored personal information of the buyer and the seller, and used data cleaning techniques with the databases.
- Executed data analysis on the conversations between the chatbot and the user in order to discover patterns that could be utilized for further development of the chatbot.

PROJECTS

Nobel Prize Laureate Data Warehouse: Managed a team with 3 members to collaborate in implementing a complex Data Warehouse that organizes and transforms the raw data of every known Nobel Prize laureate's personal information. SQL Server Management Studio, SQL Server Integration Services, Visual Studio, and Tableau were used to build the Data Warehouse, views, T-SQL queries, ETL pipeline, and data visualizations. Link

Movie Ratings and Profitability Statistical and Machine Learning Project: Determine the best kind of movie that will lead to the most profits and positive feedback for their future films with machine learning techniques: exploratory data analysis (EDA), K-means clustering, decision trees, bagging, boosting, random forests, and AUC-ROC. Evaluate the correlation between a movie's profitability and audience ratings. Link

Film Story Text Mining and Sentimental Analysis: Implemented to see if text mining, sentimental analysis, and natural language processing (NLP) techniques can ease better information gathering for categorizing movies with the raw text movie descriptions. Link

Course Selection Database System: Assigned as project manager of a team of 3 members to coordinate database development and determine project scope and limitations. A database system that acts as a course selection network which would ease the process of course registration for both students and teachers by using PostgreSQL and MongoDB. Link