Akhila Pagadala

Email: apagadal@gmu.edu | *Phone:* +1 5714917568 | *LinkedIn:* www.linkedin.com/in/apagadal

EDUCATION

Master of Science - Computer Engineering | GPA: 3.93/4 May 2021
George Mason University, Fairfax, Virginia
Courses: IoT, Big Data Technologies, Database Management Systems, Machine Learning, Data Structures

Bachelor of Technology - Electronics and Computer Engineering May 2018
TKR College of Engineering and Technology, JNTU, Hyderabad, India

TECHNICAL SKILLS

Programming Languages: SQL, Python, C, Shell Scripting, Core Java
Machine Learning: Regression, Classification, Association, CART, Clustering, Gradient Boosting using various algorithms
Big data: Hadoop Ecosystem, Apache Spark, Spark SQL, Zeppelin Notebook
Tools: Jupyter Notebook, Git, MS Visio, Lucid Chart, Microsoft Excel, Tableau, Kibana
Database : Oracle, MySql, Microsoft SQL Server, Nosql: Amazon Elasticsearch
AI: KNN, Multi-Layer Perceptron, CNN
Cloud Platforms: AWS, Snowflake Basics

EXPERIENCE

Graduate Teaching Assistant - *George Mason University, VA* Jan 2020 - May 2021
Instructed students, held office hours, graded and assisted students with lab experiments for ECE 332 and ECE 350

Intern - *NetElixir, Hyderabad, India* Feb 2018 - Aug 2018

- Boosted revenue by 54% for NetElixir's biggest client - Lenovo in the SEM team & handled three Adword accounts
- Tools: Google Ads, Google Analytics, Microsoft Excel

PROJECTS

Sentiment Analysis of COVID-19 Vaccine Tweets - ***PySpark, Apache Spark's MLib, Pandas, Pyplot, Jupyter Notebook***

- 1M+ tweets are converted to **parquet** format and preprocessed. Pretrained tweets are obtained using TextBlob.
- Built a **pipeline** with feature transformers tokenizer, stopwords remover, N-gram and feature extraction TF-IDF.
- Grid search with stratified 10-fold cross-validation was employed to obtain optimal hyperparameters for Linear Regression, Decision Tree & Random Forest. Evaluated the models using accuracy, precision, recall and F1-score.
- Visualized the data to get insights on three different vaccines using word clouds, bar graphs and top N grams.

Loan Prediction Problem – ***Sklearn, Pandas, Numpy, Pyplot, Jupyter Notebook***

- Solved binary classification problem of loan prediction by implementing various classification algorithms like KNN, Decision Tree, SVM, Logistic Regression after exploratory data analysis and preprocessing the data.
- Used Model evaluation parameters like Jaccard, F-1 Score and LogLoss to report the results.

NYC Taxi Trips Analysis - ***Scala API, Spark SQL, Zeppelin Notebook***

- Analyzed the timings and locations at which customers were picked-up and dropped-off, their trip distances and rate types to find the best on-duty times and locations for NYC taxi drivers.
- Visualized the data using Spark SQL in Zeppelin.

Deep Learning Model for Image Classification - ***Keras, Python, Jupyter Notebook***

- The model consists of CNN layers, Max pooling layers, Dropout layers, and Dense layers.
- Trained using CIFAR-10 dataset and obtained an accuracy of 86%.

Wardriving Rig - ***C++, REST API, Amazon ElasticSearch, Kibana***

- Designed using Arduino Nano 33 IoT, BN 220 GPS and Micro SD card modules.
- Collected GPS coordinates and WiFi attributes in JSON format and sent to the Amazon ElasticSearch database using REST API. Kibana is used for data visualization in a real-time map.

Nutrition-Analysis Dashboard - ***Tableau***

- Analyzed daily calorie consumption, calories intake per meal and calories per food source.
- Identified the top most consumed unhealthy foods and various other trends.

Database Design - ***Oracle database, Lucidchart***

- Used LucidChart to draw the EER diagram and created a database model for a house rental service using SQL