SUSMITHA ARIKATLA

+1(281) -750-6048|susmitha.usk7@gmail.com | **GitHub** |Houston, Texas

EDUCATION

MSc in Data Science | University of Houston | Houston, Texas | GPA: 3.6   *May 2023*

- Awards & Scholarships: Dean's Honors List, Engineering Dean's Master Scholarship, Masters Competitive Scholarship

WORK EXPERIENCE

Data Science – Teaching Assistant | University of Houston | Houston, Texas   *Aug 2022- May 2023*

- Highlighted the utilization of various software tools including Excel, Power BI, R, Tableau to demonstrate the implementation of data science techniques and machine learning models.

Project Planning & Implementation Engineer | Tata Communication |Chennai, India

- Designed project plans that include project scope, timelines, budget, and resource requirements.
- Coordinated with vendors, contractors, and other stakeholders to ensure that projects are executed according to plan.
- Monitored project progress and adjustments as necessary to keep projects on track.
- Maintained clear and consistent communication with stakeholders throughout the project, including project sponsors and clients.

SKILLS

Data Visualization   : Tableau, Looker, Power BI, Qlik Sense, Python (Seaborn, Matplotlib), R

Database Management: MySQL, SQL Server, PostgreSQL, Azure ML Studio, Oracle, Snowflake

Python Libraries   : Pandas, Numpy, TensorFlow, Keras, Scikit-learn, PyTorch

ML Algorithms   : Logistic Regression, Decision Trees, Neural Networks, Random Forests

Data Transformation   : ETL, Power Query

Statistical Analysis   : Hypothesis Testing

Big Data Technologies   : Databricks, Synapse, Apache Spark

Tools   : Microsoft Office Suite (Word, Excel, PowerPoint, Outlook), DAX, Pivot Tables

Certifications   : Machine Learning (HarvardX), Google Data Analytics Professional Certificate

PROJECT EXPERIENCE

Construction Safety Analysis using OSHA Dataset   *Apr 2023*

- Extracted and analyzed 100k records from OSHA Website using advanced web-scraping techniques, resulting in enhanced data quality and improved understanding of safety trends within the construction industry.
- Applied Principal Component Analysis (PCA) to effectively reduce the dimensionality of the dataset, resulting in more efficient data representation and analysis.
- Employed K-means Clustering technique to determine the optimal number of clusters (K) and identify distinct safety profiles among construction companies, allowing for targeted safety interventions and improvements.

Visualization of Data Scientist Job Salaries   *Dec 2022*

- Performed extensive data analysis on job salary data for data scientist positions, identifying key industry trends and patterns, resulting in actionable insights for optimizing compensation strategies.
- Utilized advanced cloud-based technologies such as Snowflake and S3Bucket to efficiently store and retrieve large volumes of job salary data, streamlining the analytical process by 30%.
- Designed and Crafted interactive dashboards using Power BI to visually depict the distribution of Data Science job salaries based on experience levels and job titles for 2023.

Seoul Bike Sharing Demand   *Aug 2022*

- Collected and aggregated data from various sources to create a comprehensive dataset for analysis (EDA)
- Cleaned, normalized, and engineered features in the dataset to ensure data integrity and suitability for modeling.
- Applied predictive models, including regression and machine learning algorithms, to accurately forecast the demand for bikes at different stations. Achieved 80% accuracy rate, significantly improving bike sharing demand forecasting in Seoul.

Optimization of SVM Classifier using Kernel and Ensemble Techniques   *Apr 2022*

- Created custom kernel functions for SVM using Sine, Cosine, RBF, Gaussian, and Polynomial functions using kernel tricks to train unbalanced data. Used Boosting and Bagging methods on custom kernel SVM models to improve accuracy. Enhanced accuracy of the models by more than 15% when compared with inbuilt SVM functions.

Machine Learning Pipeline to predict adult income using Azure ML Studio   *Apr 2022*

- Collected adult census data from the source onto **Azure ML Studio** and analyzed Data (EDA) to find the correlation among the features. Trained a Two-Class Boosted Decision Tree Model and Directed Hyperparameter Tuning to optimize the parameters.
- Evaluated and published the model as a web service to predict income based on user input.