SUSMITHA ARIKATLA

+1(281) -750-6048|susmitha.usk7@gmail.com |LinkedIn |Git Hub | 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, Master's Competitive Scholarship
- **Relevant Courses**: Neural Networks, Machine Learning, Statistics, Predictive Analytics, Database Management, Big Data, Time series Forecasting, Machine Diagnostics, Hypothesis Testing

SKILLS

Visualization Tools : Tableau, Looker, Power BI, Qlik Sense, Microsoft Excel, Python (Seaborn, Matplotlib)
Tools : Microsoft Office, DAX, PowerPoint, SharePoint, Hypothesis Testing, SQL
Database : MYSQL, SQL Server, PostgreSQL, Azure ML Studio, Oracle
Python Libraries : Pandas, Numpy, TensorFlow, Keras, Scikit-learn, PyTorch.
Cloud Technologies : AWS, GCP, Snowflake, Microsoft Azure
Environments : GitHub, Google colab, PyCharm, VSCode, Jupyter, Rstudio
Data Modeling : Logistic Regressions, Decision Trees, Multivariate Regression, Neural Networks, Random Forests

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.

Data Analyst | Freelance **|** Remote *Jan 2019 – Apr 2022*

- Implemented data visualization techniques for creating Dashboards and reports using Python and Tableau to present complex data insights to stakeholders, resulting in improved decision-making and a 20% reduction in time spent on data analysis.
- Collaborated with cross-functional teams to develop and implement machine learning algorithms for maintenance and collect and report key metrics.
- Combined advanced statistical techniques and predictive modeling to analyze a large dataset of customer behavior, resulting in a 15% increase in customer retention.

PROJECT EXPERIENCE

Visualization of Data Scientist Job Salaries *May 2023*

- 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 **S3**Bucket 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.

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.

Online Optimization of SVM Classifier using Kernel and Ensemble Techniques *Aug 2022*

- Designed custom kernel functions for **SVM** using kernel tricks to train unbalanced data.
- Used **Boosting** and **Bagging** methods on custom kernel SVM models to improve accuracy.
- Increased accuracy of the models by more than 15% when compared with inbuilt SVM functions.

Online class monitoring tool using facial recognition and emotion analysis *Jun 2022*

- Used Pandas, SciPy, Scikit-learn, PyTorch, and other libraries to process images from video frames.
- Developed modules using state-of-the-art technologies such as **FaceNet** and **DeepFace** for facial recognition and emotionanalysis. Documented the emotions of students throughout a video lecture and presented a detailed report.

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.

CERTIFICATIONS

- Machine Learning (HarvardX)
- Google Data Analytics Professional Certificate
- Tableau Desktop Specialist (TableauSoftware, LLC)