

Susmitha Arikatla

+1(281)-750-6048|susmitha.usk7@gmail.com | [GitHub](#) | [LinkedIn](#) | Houston, Texas

EDUCATION

University of Houston

Houston, Texas

Masters in Data Science (CGPA: 3.6/4.0)

May 2023

- Awards & Scholarships: Dean's Honors List, Engineering Dean's Master Scholarship, Master's Competitive Scholarship
- **Relevant Coursework:** Machine Learning, Statistics, Data Analysis, Database Management, Big Data, Time series Forecasting, Artificial Intelligence, Data Science for Security, Data Visualization, Time Series Analysis
- **Certifications:** Google Data Analytics Professional Certificate, Tableau Desktop Specialist (Tableau Software, LLC) Machine Learning-HarvardX, Python & Machine Learning for Financial Analysis

SKILLS

- **Data Visualization/Engineering:** Tableau, Looker, Power BI, Qlik Sense, Excel, Python (Seaborn, Plotly) Rshiny
- **Techniques:** Hypothesis Testing, Recommendation Engines, Customer Segmentation Analysis.
- **Tools/Frameworks:** Elastic Search, Python (Keras, scikit-learn), Hadoop, Microsoft Office, Apache Spark, Apache Kafka, Data Pipelines, SSRS, Agile, SAP, ETL, Alteryx, DAX, Informatica, Data Mining
- **Database:** Oracle, MongoDB, AWS S3, MYSQL, Snowflake
- **Data Modeling:** Logistic Regressions, Decision Trees, Multivariate Regression, Neural Networks, Random Forests

PROFESSIONAL EXPERIENCE

Cullen College of Engineering, University of Houston

Houston, Texas

Data Visualization Teaching Assistant

Aug 2022 – May 2023

- Showcased the utilization of various software tools including Excel, SAS, R, Tableau to demonstrate the implementation of data science techniques and machine learning models.
- Generated visualizations for professors to use in class, increasing student comprehension by 17% in 4 months.

Freelance

Remote

Data Analyst

June 2020 – Apr 2022

- Implemented data modeling techniques and statistical analysis to optimize business processes, resulting in a 30% increase in team efficiency and cost savings of \$50,000 per year.
- Partnered with Cross-functional teams and identified opportunities for improvement through analysis of key metrics and KPIs related to user acquisition.
- Automated several Excel Data tools (pivot tales, VBA, and Excel Formula), reports into Tableau dashboards, saving 117 hours a month.

ACADEMIC PROJECTS

Construction Safety Analysis using OSHA Dataset

[Kaggle Link](#)

Tech Stack: Numpy, Google Collab, PowerBI, Excel

- Applied web-scraping to collect the data and preprocessed using advanced NLP techniques, including stemming and lemmatization, to enhance data quality and trends.
- Applied Principal Component Analysis (PCA) to reduce dataset dimensionality and Conducted TF-IDF Vectorization for feature representation.
- Employed the K-means Clustering technique to identify the optimal number of clusters (K) and centroids.

Seoul Bike Sharing Demand

[Kaggle Link.](#)

Tech Stack: Regression, Google Collab, MySQL, Python, Tableau

- Collected and aggregated data from various sources to create a comprehensive dataset for analysis.
- Cleaned, normalized, and engineered features in the dataset to ensure data integrity and suitability for modeling.
- Developed 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.

Comparative study of supervised learning algorithms for Intrusion Detection

[Kaggle Link](#)

Tech Stack: Numpy, Algorithms, Python, Data Analysis

- Gathered Data from external sources and processed it using techniques PCA, Label encoding, and Normalization.
- Trained and evaluated Decision Tree, Random Forest, Support Vector Machines, Boosting, Bootstrap Aggregation, Naive Bayes, and Advanced Neural Networks algorithms.
- Assessed the best algorithm based on performance metrics like Accuracy Score, Execution Time, and F1 score.

VOLUNTEERING & INTERESTS

- UH-ISSO organization | Student Mentor