**Susmitha Arikatla**

+1(281)-750-6048| [susmitha.usk7@gmail.com |](mailto:susmitha.usk7@gmail.com%20|) [GitHub](https://github.com/Susmithareddy1994?tab=repositories) |[LinkedIn](https://www.linkedin.com/in/smithareddya/)| Houston, Texas

**EDUCATION**

**University of Houston** **Houston, Texas**

Masters in Data Science GPA: 3.6 *May 2023*

* Awards & Scholarships: Dean’s Honors List, Engineering Dean’s Master Scholarship, Master’s Competitive Scholarship
* **Relevant Coursework**: Machine Learning, Statistics, Predictive Analytics, Database Management, Big Data, Time series Forecasting, Machine Diagnostics, Hypothesis Testing, Artificial Intelligence

**SKILLS**

* **Languages:** Python, R, SQL
* **Frameworks:** Scikit, NLTK, TensorFlow, Keras, Django, Numpy, Pandas, Plotly, Jupyter Lab
* **Analysis Techniques:** Logistic Regressions, Decision Trees, Neural Networks, Random Forests, Statistical Analysis
* **Tools:** Data science pipeline (cleansing, wrangling, visualization, modeling, interpretation), Statistics, ETL, A/B testing, Hypothesis testing**,** Microsoft Outlook, Power Point, Word, MacOS, Microsoft Azure, AWS, Excel, Power BI

**PROFESSIONAL EXPERIENCE**

**Cullen College of Engineering, University of Houston Houston, Texas**

Data Visualization Teaching Assistant *Aug 2022 – May 2023*

* Utilized advanced data science techniques and machine learning models using Excel, SAS, R, and Tableau to analyze student performance data and identify areas for improvement resulting in a 12% increase in overall grades.
* Designed interactive visualizations for course material using Tableau resulting in a 17% increase in student comprehension within the first 4 months of implementation.

**Freelance Remote**

Data Analyst *June 2020 – Dec 2021*

* Developed 10 custom data visualizations using Tableau to illustrate user acquisition trends and inform cross-functional teams, resulting in a 15% increase in conversion rates.
* Collaborated with marketing team to identify and target high-value customer segments, increasing revenue by $200K within 3 months.
* Streamlined data management processes using SQL queries in MYSQL, reducing dashboard creation time by 50% and increasing accuracy of reporting.

**Tata Communications Chennai, India**

Project Planning & Implementation Engineer*Apr 2016 – May 2017*

* Orchestrated the development of project plans for 20+ projects, resulting in a 95% on-time completion rate and an average cost savings of $50K per project.
* Managed risk through continuous monitoring and adjustment of project progress, resulting in a decrease in major issues by over 60% compared to previous years.
* Developed and presented proposals for process improvements resulting in a 25% increase in efficiency and cost savings of $50k.

**PROJECT EXPERIENCE**

**Construction Safety Analysis using OSHA Dataset |**Machine Learning, Tableau| Data Analysis **|** [**Kaggle** *Apr 2023*](https://www.kaggle.com/code/susmithareddya/construction-safety-analysis-using-osha-data/edit)

* Assisted in-depth Construction Safety Analysis by leveraging a **web-scraped** dataset of 26,000 OSHA records.
* Preprocessed the data using advanced **NLP** techniques, including stemming and lemmatization, to enhance data quality and trends.
* Applied Principal Component Analysis (**PCA**) to reduce dataset dimensionality and implemented **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|** Machine Learning, Tableau| Data Analysis |[Kaggle](https://www.kaggle.com/code/susmithareddya/seoul-bike-sharing/edit) *Aug 2022*

* Curated 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.
* Implemented 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.
* Generated the performance of the model’s using metrics like mean absolute error (**MAE**), root mean squared error (**RMSE**), and R-squared. Devised visually appealing **dashboards** and **reports** to aid in data visualization and facilitate decision-making.

**CERTIFICATIONS**

* Machine Learning – Harvard University – EDX
* Google Data Analytics Professional Certificate
* Tableau Desktop Specialist (Tableau Software, LLC)