

SUSMITHA ARIKATLA

+1(281) -750-6048|susmitha.usk7@gmail.com |<https://www.linkedin.com/in/susmithaa/>|Houston,Texas

EDUCATION

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

- Awards & Scholarships: Dean's Honors List, Engineering Dean's Master Scholarship, Masters Competitive Scholarship
- Relevant Coursework:** Machine Learning, Statistics, Data Analytics, Database Management, Big Data, Time Series Forecasting, Artificial Intelligence, Data Visualization, Time Series Analysis, Data Analysis

WORK EXPERIENCE

Data Science – Teaching Assistant | University of Houston | Houston, Texas

- 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

- 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

Visualization Tools : Tableau, Looker, Power BI, Qlik Sense, Microsoft Excel, Python (Seaborn, Matplotlib), R,Java
Tools : Microsoft Office, DAX, PowerPoint, SharePoint, Hypothesis Testing, SQL, SAS, ETL, Matillion
Database : MYSQL, SQL Server, PostgreSQL, Azure ML Studio
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

PROJECT EXPERIENCE

Construction Safety Analysis using OSHA Dataset

Apr 2023

Tech Stack: Numpy, Google Collab, Power BI, Excel

- 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

Tech Stack: Snowflake, S3Bucket, Power BI, SQL

- 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

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

- 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.

Comparative study of supervised learning algorithms for Intrusion Detection

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, F1 score.

CERTIFICATIONS

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