SUSMITHA ARIKATLA

Susmitha.usk7@gmail.com | GitHub |LinkedIn | 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
- **Relevant Coursework**: Machine Learning, Probability & Statistics, Data Analytics, Database Management, Big Data, Time Series Forecasting, Artificial Intelligence, Data Visualization, Data Science for Security
- **Certifications:** Machine Learning (HarvardX), Google Data Analytics Professional Certificate

WORK EXPERIENCE

Data Science – Teaching Assistant | University of Houston *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 *Jan 2020 – Apr 2022*

- Utilized advanced Excel functions, including data formatting, Pivot tables, V-lookup, and dashboard creation to streamline data representation and enhance decision-making processes.
- Executed SQL queries against current databases to derive meaningful insights and generated analytical reports to support data-driven decision-making.
- Applied Agile methodology throughout the project lifecycle, actively participating in weekly and daily release management activities.
- Transformed third-party spending data into user-friendly deliverables using Excel and Python libraries such as NumPy and Pandas.
- Developed Tableau data visualizations, including Scatter Plots, Geographic Maps, Pie Charts, Bar Charts, and Density Charts, enhancing data interpretation.
- Designed and implemented Tableau dashboards, incorporating calculated fields, parameters, calculations, groups, sets, and hierarchies to effectively communicate critical Key Performance Indicators (KPIs) to senior management.

Project Planning & Implementation Engineer | Tata Communication *Apr 2016 – May 2017*

- Analyzed 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 clients.
- Achieved 95% of project deliverables on time with zero budgetary overruns and within scope.

SKILLS

Data Visualization: Tableau, Looker, Power BI, Qlik Sense

Programming Languages: Python, R

Databases: MySQL, SQL Server, PostgreSQL, Azure ML Studio, Oracle, Snowflake

Data Science Libraries: Pandas, NumPy, TensorFlow, Keras, Scikit-learn, PyTorch.

Machine Learning: Logistic Regression, Decision Trees, Neural Networks, Random Forests

Big Data Technologies: Apache Spark, Apache Kafka, Apache Airflow, Hadoop

Cloud Platforms: AWS, Azure, GCP

Tools: Git, Docker, Microsoft Office Suite, DAX, Power Query, SQL, A/B testing

PROJECT EXPERIENCE

Construction Safety Analysis using OSHA Dataset *Apr 2023*

- Extracted and analyzed 100k records from OSHA Website using advanced NLP techniques.
- Applied Principal Component Analysis (PCA) and TF-IDF Vectorization for feature representation.
- Employed K-means Clustering to identify distinct safety profiles among construction companies.

Visualization of Data Scientist Job Salaries *Dec 2022*

- Conducted data analysis on job salary data for data scientist positions.
- Utilized Snowflake and S3Bucket for efficient storage and retrieval of large volumes of job salary data.
- Generated interactive Power BI dashboards to visualize salary distribution based on experience levels and job titles.

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.