



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

- **Certifications:** Machine Learning (HarvardX), SnowPro Core

SKILLS

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

Programming Languages: Python, R, Java

Azure Technologies: Azure Synapse, Azure ADLS Gen 2, Azure Data Factory, Azure SQL, Data Lake, Databricks

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, Apache HiveQL

Cloud Platforms: AWS, Azure, GCP

Tools: Git, Docker, Microsoft Office Suite, DAX, Power Query, SSIS, Informatica

WORK EXPERIENCE

Data Engineer Intern | ATVS

Oct 2023- Present

- Prepared and processed large datasets for analysis, utilizing Excel, resulting in enhanced data quality and accuracy for further processing with Big Data technologies.
- Revamped and analyzed existing data dictionaries to include a comprehensive history, ensuring consistent development across various domains.
- Utilized acquired knowledge of Project Management (Agile) and Software Development Life Cycle (SDLC) to design tests that yielded valid solutions to business problems.

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 2019 – Dec 2021

- 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.
- 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

- Authored 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.

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.

Optimization of SVM Classifier using Kernel and Ensemble Techniques

Dec 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.
- Streamlined accuracy of the models by more than 15% when compared with inbuilt SVM functions.