SUSMITHA ARIKATLA

Data Analyst

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

**Data Science Teaching Assistant** *Aug 2022 – May 2023*

* Showcased the utilization of various software tools including Excel, Power BI, R, Tableau to demonstrate the implementation of data science techniques and machine learning models.

# WORK EXPERIENCE

**Data Analyst |** Freelance | Remote, India *June 2020 – Apr 2022*

# Data analysis and data cleansing using Python and data visualization using Tableau.

# Developed interactive dashboards on Power BI to visualize key performance indicators (KPIs) of maintenance processes and presented findings to higher management.

# Created training materials and process documents using Microsoft Office, facilitating the implementation of predictive analytics for improved decision-making.

# SKILLS

|  |  |
| --- | --- |
| **Programming Language:**  Python (NumPy, Scikit-learn, Pandas, TensorFlow, Keras, PyTorch), R | |
| **Database Tools** | : MS Access, MySQL, Oracle DB, Hadoop, Azure ML Studio, Snowflake, ETL | |
| **Visualization Tools** | : Tableau, Power BI, Microsoft Office, Python, VBA, DAX, Excel, Looker, Qlik Sense, PowerPoint | |
| **Environments** | : GitHub, Google colab, PyCharm, VSCode, Jupyter, RStudio | |
| **Key Concepts** | : Machine Learning, Statistics, Adobe Analytics, Database Mgmt., Big Data Analytics | |

**PROJECTS**

**Construction Safety Analysis using OSHA Dataset** *Jun 2023*

* Collected the data using web scrapping. Preprocessed the data using advancedNLP 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. Employed the K-means Clustering Technique to identify the optimal K and Centroids

# Optimization of SVM Classifier using Kernel and Ensemble Techniques *Dec 2022*

* Designed 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. Increased accuracy of the models by more than 15% when compared with inbuilt SVM functions.

# Online class monitoring tool using facial recognition and emotion analysis *Aug 2022*

* Used Pandas, SciPy, Scikit-learn, PyTorch, and other libraries to process images from video frames. Developed modules using state-of-the-art technologies such as FaceNet and DeepFace for facial recognition and emotion analysis. Documented the emotions of students throughout a video lecture and presented a detailed report.

# Comparative study of supervised learning algorithms for Intrusion Detection *Apr 2022*

* Gathered Data from sources and processed it using techniques PCA, Label encoding, and Normalization. Trained the Decision Tree, Random Forest, SVM, XGBoost, Naive Bayes, and Advanced Neural Networks algorithms. Validated and presented the best algorithm based on Accuracy Score, Execution Time, and F1 score metrics.

# CERTIFICATIONS

* + Google Data Analytics Professional Certificate
  + Machine Learning-HarvardX
  + Python & Machine Learning for Financial Analysis

# VOLUNTEERING & INTERESTS

* + UHISSO organization | Student Mentor
* Data Visualization with Tableau
* Machine Learning Pipelines with Azure ML Studio
  + Advancing Community Engagement Services Institute, UH | Academic Impact Cougar Tutor
  + Interests: Photography, Hiking, Volleyball