

# Susmitha Arikatla

+1(281)-750-6048||[susmitha.usk7@gmail.com](mailto:susmitha.usk7@gmail.com)||[LinkedIn](#)||[GitHub](#)||Houston, Teas

## 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, Master's Competitive Scholarship
- Relevant Courses:** Neural Networks, Machine Learning, Statistics, Predictive Analytics, Database Management, Big Data, Time series Forecasting, Hypothesis Testing

## SKILLS

- SQL (SQL Server, MYSQL, PostgreSQL), ETL, R
- Microsoft Azure, Snowflake, Informatica, AWS
- Python (Pandas, Numpy, SciPy, Matplotlib, TensorFlow, NLTK)
- Modeling (Regression, Decision Tree, Clustering, CNN)
- Tableau, Microsoft Power BI, Qlik Sense
- Excel (VLOOKUP, Conditional Formatting, Pivot Tables), PowerPoint, Word, DAX

## ACEDAMIC PROJECTS

### Construction Safety Analysis using OSHA Dataset.

- Performed Data Extraction using web-scraping from OSHA Website. Preprocessed the data using advanced NLP techniques such as stemming and lemmatization to improve data quality.
- Utilized Principal Component Analysis (PCA) to reduce dataset dimensionality and applied TF-IDF Vectorization.
- Employed K-means Clustering technique to identify optimal K value and centroids, and named resulting clusters based on safety incident type or severity. Results were used to prioritize safety improvements and identify areas requiring additional resources.

### Online Optimization of SVM Classifier using Kernel and Ensemble Techniques

- Designed custom kernel functions for SVM 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.

### Machine Learning Pipeline to predict adult income using Azure ML Studio

- Collected adult census data from the source onto Azure ML Studio and analyzed Data (EDA) to find the correlation among the features. Trained a Two-Class Boosted Decision Tree Model and Directed Hyperparameter Tuning to optimize the parameters.
- Evaluated and published the model as a web service to predict income based on user input.

### Visualization of Data Scientist Job Salaries

- Performed data cleaning and filtering techniques in Excel to prepare dataset for analysis.
- Conducted Exploratory Data Analysis (EDA) using Python Pandas to identify patterns in the dataset.
- Utilized Tableau Desktop to create visually impactful representations of Data Science job salaries based on experience levels and job titles for 2023.
- Designed and implemented unique visual representations using Glyphs to enhance the understanding of various experience levels in the data.

## WORK EXPERIENCE

### Information Visualization – Teaching Assistant | University of Houston | Houston, Texas

- Assisted in designing/editing course content, administered student activity, and performed administrative assessments.
- Graded and provided detailed feedback on 60+ assignments per week, resulting in a 20% reduction in student inquiries and a 15% improvement in student grades.

### Project Planning & Implementation Engineer | Tata Communication | Chennai, India

- Developed 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.
- Identified and mitigated project risks and issues.
- Maintained clear and consistent communication with stakeholders throughout the project, including project sponsors, clients, and team members.

## CERTIFICATIONS

- Machine Learning – Harvard University - EDX
- Python and Machine Learning for Financial Analysis
- Google Data Analytics Professional Certificate,
- Tableau Desktop Specialist (Tableau Software, LLC)