

Monisha R (She/her)

CONTACT: +91 8610831486 | monishavenkatesh7@gmail.com | Chennai, Tamil Nadu | [LinkedIn](#) | [GitHub](#) |

EDUCATION

Degree	Institute	CGPA or percentage
B.Tech (Civil Engineering) – 2021	SASTRA Deemed University, Thanjavur	79.40(7.94/10)

SKILLS

Software Tools: Power BI, Advanced Excel, DAX Formulation, Power Pivot, Pivot Table, SQL Queries, BeautifulSoup.

Coding Languages: Python, MySQL, Selenium.

Data Analysis: Data Cleaning, Data Visualization, Statistical Analysis, Predictive Modeling, Trend Analysis.

Cloud Technologies: Amazon Web Services(AWS), RDS, EC2, S3.

INTERNSHIPS

Project Research & Data Analyst intern	Indian Institute of Management Mumbai	January 2024 – August 2024
---	--	-----------------------------------

Project: Indian Logistics Leveraging Recent Tech in Tourism Sectors **Guide: Dr. Vijay, Asst Prof, IIM Mumbai**

- Boosted analytical efficiency by 50% by using DAX formulations and power query in Power BI.
- Increased extraction efficiency by 80% by reducing manual effort using BeautifulSoup and selenium for web scraping.
- Designed ETL processes, optimizing GDP/GSDP data analysis for 36 states and UTs, expediting decision-making.
- Delivered actionable insights on footfall patterns, driving a 70% enhancement in trade forecasting accuracy.

Stocks Data Analysis intern	Reyanna Technologies Pvt Ltd	April 2024 – June 2024
------------------------------------	-------------------------------------	-------------------------------

Project: Understanding and Identifying Price Trends of US Stocks **Guide: Mr. Divakar, Director, Strategic Projects**

- Uncovered profitable stock trading strategies through in-depth analysis of 10 years of historical data.
- Reduced man-hours by 80% by streamlining workflow through Python automation for trend visualizations.

PROJECTS

Cybersecurity Threat Classification [Big Data Feature Engineering Machine Learning] Link	September 2024
---	-----------------------

- Reduced big data size by 75% through advanced Data cleaning, optimizing processing speed and memory usage.
- Enhanced computational efficiency by 50% using parquet files and 100,000-row data chunking for large datasets.
- Boosted model accuracy by balancing TP, FP, and BP class distribution across 5.87M rows with stratified sampling.
- Reduced the Data volume by 30% by finding Feature Importance and highly correlated features by correlation matrix
- Accomplished feature engineering to combine, delete and fine-tune 77 features avoid the overfitting of the model.
- Attained 89% accuracy with a Decision Tree, validated through cross-validation with a mean CV score of 0.89.
- Improved Random Forest model performance by achieving 83% accuracy and 82% Macro F1 score through Grid Search.
- Applied stratified sampling to handle class imbalances in datasets improving model performance by 15%.

YouTube Video Recommender - [TF-IDF Hierarchical Clustering Streamlit AWS RDS] Link	August 2024
--	--------------------

- Deployed global app on EC2 using AWS S3 and RDS for efficient data handling of 4500 YouTube video data.
- Diminished computational time by 10% using NLP techniques, including lemmatization and stop word removal.
- Optimized clustering model accuracy by applying statistical methods (Elbow method) improving precision by 15%.
- Fine-tuned AWS cloud resources by reducing computational efforts by 30% by using data chunking and pickle files.
- Utilized TF-IDF vectorisation and Hierarchical clustering to a silhouette score of 0.34 for best keyword matches.
- Refined user experience by providing actionable video recommendations with a user-friendly Streamlit interface.

Customer, sales and Portfolio analysis - [SQL PowerBI KPI DAX Formulation] Link	August 2024
--	--------------------

- Managed over 700,000 data points, 5 datasets, integrating MySQL and advanced DAX formulation in Power BI.
- Designed an SQL database using SQLAlchemy in Python, ensuring seamless data transfer and error-free analysis.
- Established SQL connection and imported data into 5 SQL tables, optimizing data storage and retrieval.
- Condensed 500,000+ data points into 5 buckets generating precise and actionable insights for decision-makers.
- Facilitated Data-driven decision making with 20+ interactive visualizations with multiple KPIs.

CERTIFICATIONS / AWARDS:

- Completed Certificate Course in Pharmacovigilance [Link](#).
- Complete Generative AI Course With Langchain and Huggingface [Link](#).
- IIT-M Advanced Programming Professional & Master Data Science (GUVI) [Link](#).
- Google Data Analytics Certification (Coursera) [Link](#).
- Get Job Ready: Power BI Data Analytics for All Levels 2.0 (Codebasics) [Link](#).
- The Complete Python Bootcamp from Zero to Hero in Python (Udemy) [Link](#).
- 2nd Prize in Debate Competition at IIT Madras (Civil Engineering Association).

POSITION OF RESPONSIBILITIES

- Directed a three-day event for 500 participants as Head Organizer of Hinges 2019.
- Led the Daksh Arts Team at SASTRA University, organizing intra-university events and improving team coordination.