

ROSHAN MADHAV

+91 9544795385 • roshanmadhav928@gmail.com • [LinkedIn](#) • [GitHub](#)

ABOUT ME

Aspiring Data Scientist pursuing an MSc in Data Science at VIT Vellore, with a strong analytical foundation built through a Physics background. Skilled in Machine Learning, Statistical Modeling, Deep Learning, Predictive Analytics, and Big Data Analytics. Experienced in developing end-to-end data science solutions involving data preprocessing, exploratory data analysis (EDA), feature engineering, predictive modeling, and data visualization using Python, R, SQL, Tableau, Power BI, and Hadoop. Passionate about developing intelligent, data-driven solutions and transforming complex data into actionable business insights.

EDUCATION

MSc Data Science , Vellore Institute of Technology Vellore, India	2025 – Present CGPA: 8.73
---	-------------------------------------

BSc Physics , Payyanur College, Kannur University Kannur, India	2022 – 2025 Percentage: 78.9%
---	---

Higher Secondary Education , AKASGVHSS Payannur, Kerala HSE Kannur, India	2020 – 2022 Percentage: 94.9%
---	---

SSLC , Kendriya Vidyalaya Payyanur, Central Board of Secondary Education Kannur, India	2020 Percentage: 87.6%
--	----------------------------------

TECHNICAL EXPERTISE

Programming Languages : Python · R · SQL · Java (Basic) · C++ (Basic)

Machine Learning : Regression Analysis · Classification · Clustering · Decision Trees · Random Forest · Support Vector Machines (SVM) · XGBoost · Feature Engineering · Model Evaluation · Hyperparameter Tuning

Deep Learning : TensorFlow · Keras · Convolutional Neural Networks (CNN) · Transfer Learning · Image Classification

Data Analytics : Exploratory Data Analysis (EDA) · Statistical Inference · Hypothesis Testing · Predictive Analytics · Data Cleaning · Data Wrangling · Feature Selection

Big Data : Hadoop · HDFS · MapReduce

Data Visualization : Tableau · Power BI · Matplotlib · Seaborn

Libraries : Pandas · NumPy · Scikit-learn · Librosa · NetworkX

Tools & Databases : Git · Jupyter Notebook · VS Code · MySQL

PROJECTS

Workforce Skill Intelligence using Network Analysis & Machine Learning	2026
---	------

Designed and developed a data-driven framework for workforce skill prioritization using network analysis and machine learning to identify emerging workforce trends.

- Integrated workforce datasets from O*NET, OECD and World Economic Forum(WEF) to analyze emerging skill demand and workforce trends.
- Performed Exploratory Data Analysis (EDA) to uncover hidden relationships between workforce skills.
- Built a weighted skill network using NetworkX and computed Degree, Betweenness, PageRank, and Closeness centrality measures.

- Designed a custom Skill Influence Score (SIS) for ranking workforce skills based on network importance.
- Applied K-Means clustering to identify AI-ready, emerging, foundational, and declining skill groups.
- Developed Random Forest and XGBoost models for predictive workforce demand analysis.
- Visualized insights using correlation heatmaps, feature importance plots, and network graphs.
- **Technologies:** Python | Pandas | NumPy | Scikit-learn | NetworkX | Random Forest | XGBoost | Matplotlib

Intelligent Voice Emotion Analysis using Audio Signal Processing

2025

Developed an intelligent speech emotion recognition system using audio signal processing and **Machine Learning** techniques for emotion classification.

- Extracted MFCC, Chroma, Spectral Contrast, Zero-Crossing Rate, Pitch, and Energy features from speech signals using Librosa.
- Built an end-to-end speech emotion recognition pipeline for classifying Happy, Sad, and Neutral emotions from voice recordings.
- Performed feature visualization and analysis to understand emotional patterns within speech signals.
- Generated personalized playlist recommendations and supportive responses based on detected emotional states.
- **Technologies:** Python | Librosa | NumPy | Pandas | Matplotlib | Scikit-learn

Intelligent Monsoon Onset Forecasting using Climate Analytics & Machine Learning

2025–Present

Developing an intelligent machine learning framework for forecasting monsoon onset using meteorological datasets and climate analytics in collaboration with CUSAT(Cochin University of Science and Technology).

- Integrating real-world meteorological datasets collected in collaboration with **CUSAT** for monsoon onset forecasting.
- Performing **data preprocessing**, **EDA**, and **feature engineering** to identify key climatic variables influencing monsoon onset.
- Developing regression-based machine learning models for forecasting monsoon onset using historical climate observations.
- Evaluating predictive performance using statistical metrics and visualization techniques to improve model reliability.
- Extending the framework with **Anomaly Detection** techniques to identify abnormal weather patterns and enhance forecasting accuracy.
- **Technologies:** Python | Pandas | NumPy | Scikit-learn | Matplotlib | Seaborn

LANGUAGES

- English: Fluent
- Malayalam: Native
- Hindi: Fluent

CERTIFICATIONS

Cloud Computing (NPTEL Course).