

ANKUSH KUMAR MALLICK

☎ +91 9142623599 ✉ ankushmallick750@gmail.com 🔗 linkedin.com/in/ankush-mallick750

Professional Summary

Final-year B.Tech (CSE) student specializing in Machine Learning and Data Science with hands-on experience in building end-to-end ML pipelines. Skilled in data cleaning, feature engineering, hyperparameter tuning, and model evaluation using Python, SQL, and Scikit-learn. Strong foundation in core computer science fundamentals including Data Structures and Algorithms and Database Management Systems, with experience developing deployment-ready ML applications.

Education

University of Engineering & Management, Jaipur
Bachelor of Technology in Computer Science Engineering

Aug 2022 – Present
Rajasthan, India

Technical Skills

Languages: Python, C++, C, SQL

ML Frameworks: Scikit-learn, TensorFlow, Keras, XGBoost, Streamlit

Data Science Lifecycle: Data Cleaning, Feature Engineering, Hyperparameter Tuning, Cross-Validation

Data Libraries: Pandas, NumPy, Matplotlib, Seaborn

CS Fundamentals: Data Structures & Algorithms (DSA), OOPS, DBMS

ML Topics: Regression, Classification, Random Forest, Neural Networks, Transformers, NLP

Projects

Heart Disease Risk Prediction System | *XGBoost, Scikit-learn, SHAP, Streamlit*

- Built an end-to-end ML pipeline using the Cleveland dataset: performed rigorous **Data Cleaning**, imputation, and categorical encoding while preventing data leakage.
- Optimized model performance via **Hyperparameter Tuning** and advanced models (XGBoost) with **Stratified Cross-Validation**, achieving a **0.92 ROC-AUC**.
- Implemented **SHAP-based interpretability** to identify key clinical drivers and validated influence through model behavior analysis.
- Deployed a real-time risk prediction system with Streamlit, integrating preprocessing into a single joblib pipeline.

Travel Budget Prediction System | *Python, XGBoost, Scikit-learn, Streamlit*

- Developed an ML regression model to predict travel budgets using **Feature Engineering** on user inputs such as travel style and trip duration.
- Applied numeric feature preprocessing and **Hyperparameter Tuning** using Scikit-learn, achieving a **97% R2 score** on test data.
- Built an interactive Streamlit web application for real-time budget forecasting and saved models using joblib.
- Maintained a deployment-ready structure with requirements.txt and ensured model robustness through feature selection.

Certifications

Python for Data Science – NPTEL	2024
Fundamental Algorithms: Design and Analysis – NPTEL	2024
The Joy of Computing using Python – NPTEL	2023
Problem Solving Through Programming in C – NPTEL	2023

Relevant Coursework

Mathematics: Probability & Statistics, Linear Algebra, Calculus, Discrete Mathematics

Computer Science: Data Structures & Algorithms, Object Oriented Programming, DBMS, Operating Systems

Data Science/AI: Machine Learning, Design & Analysis of Algorithms, Artificial Intelligence