# Ankit Choudhary

ankit choudhary | Ankit | ankittarar703@gmail.com | +91.8447797845

## Summary

Data Science  Machine Learning Enthusiast with hands-on experience in Python, SQL, and predictive modeling. Skilled in developing data-driven applications using Pandas, Scikit-learn, and Power BI. Experienced in building ML pipelines, deploying models via REST APIs, and deriving insights that improve business performance.. For details, click here.

## Skills

| | |
|---|---|
| **Programming Languages** | Python, SQL, Java, JavaScript |
| **Data Science & ML Libraries** | NumPy, Pandas, Scikit-learn, TensorFlow, Matplotlib, Seaborn, NLTK |
| **Databases** | MySQL, PostgreSQL, MongoDB |
| **Data Analysis & Visualization Tools** | Power BI, Excel, Jupyter Notebook |
| **Frameworks & Web Technologies** | Django, Flask, RESTful APIs, React, MERN Stack, Tailwind CSS |
| **Machine Learning Techniques** | Regression, Classification, Feature Engineering, Model Evaluation, Time Series Forecasting |
| **Tools & Version Control** | Git, GitHub, VS Code, Postman, Selenium, PyTest, Docker (basic) |
| **Certifications** | Natural Language Processing (NPTEL, IIT Kharagpur),DSA (Apna College) |

## Experience

**Automation Engineer — BugHunters pvt. ltd., Faridabad** *Feb 2025 – Present*

- Designed and implemented **automated test scripts** using Python, Selenium,Playwright and PyTest to validate data pipelines, dashboards, and predictive models.
- Collaborated with the Data Science team to **verify ML model outputs and data accuracy**, ensuring reliability of analytical insights and production APIs.
- Developed and maintained **ETL test automation frameworks** to check data consistency across PostgreSQL and REST API layers.
- Created **Python-based scripts for data profiling and anomaly detection**, improving defect identification rate by 35%.
- Utilized **Pandas, NumPy, and SQL** for exploratory data analysis (EDA) during model validation and test data generation.
- Documented test metrics and results using **Power BI dashboards**, enabling data-driven QA reporting and improved release decisions.

**Software Trainee — PySpiders, Gurugram** *Oct 2024 – Dec 2024*

- Trained extensively in **Python programming, data analysis, and machine learning fundamentals**, applying them to real-world datasets.
- Performed **Exploratory Data Analysis (EDA)** using Pandas, NumPy, and Matplotlib to uncover trends and clean raw datasets for model training.
- Built and optimized **supervised ML models** (Linear Regression, Decision Trees) achieving up to 85% accuracy in predicting outcomes.
- Designed Python scripts to **automate data preprocessing and feature engineering**, improving model reproducibility and reducing manual work by 40%.

- Collaborated with a cross-functional team to **deploy ML models via Flask APIs**, integrating model predictions into web-based dashboards.
- Documented project workflows, model evaluation metrics, and testing procedures, contributing to improved version control and reproducibility.

## Projects

**SORTMEDIC AI**

*Tools: Python, Scikit-learn, MERN Stack, Tailwind CSS, REST APIs*

- Built an AI-powered hospital management system with an integrated **machine learning model** (Random Forest, KNN) to predict **cardiac arrest risk** from patient vitals and historical data.
- Designed a complete ML workflow including **data preprocessing, feature selection, and model evaluation**, achieving high accuracy in real-time clinical prediction.
- Deployed the model via REST API into a MERN-based admin dashboard, ensuring **secure access** through **JWT authentication** and **role-based authorization** for healthcare personnel.

**STOCK PRICE PREDICTION**

*Tools: Python, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn*

- Developed a stock price prediction model using **LSTM neural networks** and **time series forecasting techniques**, leveraging historical stock data from **Kaggle**.
- Engineered features including **moving averages, lag values, and price volatility**, and applied **MinMaxScaler** for data normalization to enhance model accuracy.
- Achieved improved performance over baseline models (**Linear Regression, ARIMA**), and evaluated predictions using **RMSE** and **MAE**, visualized with **Matplotlib**.

## Education

**J.C. Bose University of Science and Technology (YMCA)** *Aug 2023 – June 2025*
*Master of Computer Applications (MCA)* CGPA: 7.8/10

**J.C. Bose University of Science and Technology (YMCA)** *Aug 2019 – June 2022*
*Bachelor of Science (Mathematics Hons)* CGPA: 7.9/10