CHETAN PAWAR

Nashik, Maharashtra • chetanpawar2002official@gmail.com • 9359208098 • GitHub • LinkedIn

SKILLS

Programming Languages: Python, SQL
Frameworks & Tools: Pandas, NumPy, Data Visualization (Matplotlib, Seaborn), Power BI, Tableau, Excel, Streamlit, Flask, Jupyter Notebook, Git & GitHub, Docker
Data Manipulation: Data Cleaning, Data Wrangling, Feature Engineering
Machine Learning, Natural Language Processing (NLP), Generative AI (GenAI), Deep Learning, EDA, MLOps, Statistics, AI Agents

PROJECTS

RepoBot AI   [GitHub Link]   [Live Link]
- **Developed** an GenAI-driven web application using **Flask** and **Groq AI** to analyze and understand **over 500 GitHub code files**, enhancing code comprehension and query handling.
- **Optimized** vector-based search with **Chroma and HuggingFace embeddings**, enabling accurate retrieval across **400+ text chunks** for better code insights.
- **Automated** GitHub repository ingestion and embedding storage pipeline, increasing data processing efficiency and enabling seamless indexing for scalable code analysis.
- **Integrated** conversational AI with memory-based models, enhancing user interactions for **complex code-related queries** through dynamic responses.

End To End Food Delivery Time Prediction   [GitHub Link]   [Live Link]
- Built and deployed an end-to-end **regression model** for **predicting** food delivery time, helping optimize **logistics** and reduce delivery delays.
- Addressed extensive **outliers** and **missing values**, resolved data scaling issues, and **handled multicollinearity** among features during **feature engineering**.
- Employed multiple models, including **Linear Regression**, **Decision Trees**, **Random Forest**, and **XGBoost**, achieving an **R-squared** score of **0.82** and reducing **RMSE** to **3.98** on the test set using **XGBoost**
- **Deployed** the solution using **Streamlit** for user interaction, enabling real-time food delivery predictions. The project has the potential to reduce average delivery time by up to **15%**, improving overall **customer satisfaction**.

Airline Passenger Referral Program Development   [GitHub Link]   [Live Link]
- Built and deployed a **classification** model to predict whether a passenger will refer the airline to others, addressing a critical business need for improving **customer satisfaction and retention**
- Handled **imbalanced** classes through **oversampling** techniques, tackled **missing values**, and implemented **feature engineering** to improve **prediction accuracy**.
- Achieved a model accuracy of **87%** using **Random Forest** and a **precision score of 85%** for the positive class.
- **Deployed** the model using **Streamlit**, allowing for interactive testing and feedback from business users, with a potential to increase referral rates by **20%** based on **prediction analysis**.

INTERNSHIP EXPERIENCE

Accenture North America Data Analytics and Visualization Job Simulation on Forage   Jan 2024 - Feb 2024
- Conducted a data analysis simulation at Accenture for a social media client, providing insights for strategic recommendations.
- Cleaned and analyzed 7 datasets to identify trends, enhancing decision-making.
- Created a PowerPoint and video presentation viewed by 20+ stakeholders, driving alignment on project priorities..

Alfido Tech (Remote) | Data Analytics intern   Sep 2023 - Oct 2023
- Analyzed diverse datasets to extract actionable insights.
- Improved forecasting accuracy through data trends analysis.
- Optimized resource allocation and customer experience using data insights.

CERTIFICATIONS
- Career Essentials in Data Analysis (**Microsoft** & **LinkedIn**)
- Introduction to Data Science (**Infosys Springboard**)
- Complete Machine Learning, NLP Bootcamp, **MLOPS** & **Deployment** (**Krish Naik, Udemy**)

EDUCATION

Sandip University, Nashik   2022 - 2025
Bachelor of Technology, Computer Science and Engineering   **CGPA : 8.3/10**
MVPS's RSM Polytechnic, Nashik   2020 - 2022
Diploma in Computer Technology   **82%**