Chaitanya Khairnar

Data Scientist | Data Analyst | Machine Learning Enthusiast

+91 9284426130 | khairnarchaitanya93@gmail.com | GitHub | LinkedIn

Professional Summary

Results-driven Data Scientist with hands-on experience in data analysis and machine learning. Strong in Python, SQL, and statistical modeling with proven ability to convert raw data into actionable insights.

Skills & Technologies

Programming Languages:	Python (OOP, Scripting)
Databases:	SQL (Joins, Subqueries, Aggregations)
Data Analysis:	Pandas, NumPy, EDA
Machine Learning:	Regression, Classification, Clustering
Visualization:	Power BI, Tableau, Matplotlib, Seaborn
Tools:	Scikit-learn, Git, GitHub, Jupyter

Project Experience

Diamond Price Prediction GitHub

- Built an end-to-end ML regression pipeline to predict diamond prices using physical and quality attributes with production-ready architecture.
- Achieved 93% R^2 accuracy by training and evaluating multiple models including Random Forest, Gradient Boosting, and XGBoost.

Phishing Website Classifier GitHub

- Developed an end-to-end ML classification system to identify phishing websites using structured URL and website features with MongoDB-backed data ingestion.
- Achieved 97% Precision and 94% Recall, ensuring high phishing detection while minimizing false positives.

Customer Categorizer GitHub

- Designed an end-to-end customer segmentation system using unsupervised learning to group customers based on behavior and transactional patterns.
- Achieved a Silhouette Score of 0.61 with K-Means clustering, enabling clearly separable and actionable customer segments for targeted marketing.

Spam Detection System GitHub

- Implemented an NLP-based spam classification system with Dockerized ML pipelines and a REST API, deployable on AWS for real-time inference.
- Achieved 96% F1-score with TF-IDF and Naive Bayes/Logistic Regression, ensuring balanced spam detection and low false positives.

Climate Visibility Prediction GitHub

- Developed a regression-based ML system to predict visibility using climate and environmental features such as temperature, humidity, wind, and air quality.
- Achieved R^2 score of 0.89 with low RMSE, enabling reliable visibility forecasting for transportation and aviation safety use cases.

Myntra Product Scraper & Analysis GitHub

- Built a modular web scraping pipeline to extract structured product data (price, ratings, discounts, availability) from Myntra for market analysis.
- Successfully scraped and cleaned 10,000+ product records with automated pagination handling, producing analysis-ready datasets for BI and ML use.

Education

D.Y. Patil International University, Pune 2021 – 2023
MCA – Data Science & Artificial Intelligence **71%**

IMRD, Shahada 2018 – 2021
BCA **CGPA: 9.21**