PRATHMESH VAIDYA

Data Engineer & AI/ML Engineer

7666342411 prathmeshvaidya20@gmail.com Linkedin | Github | Portfolio Navi Mumbai, India, 400614

SUMMARY

I'm a Data Engineer & AI/ML Engineer passionate about turning complex data challenges into intelligent, scalable solutions. My expertise spans large language models, machine learning, and robust data engineering. With hands-on experience in developing AI-driven applications, I focus on ensuring that data is transformed into actionable insights, creating systems that offer both analytical and practical benefits

EDUCATION

Post Graduate Diploma in Big Data Analytics
Mumbai
C-DAC
08/2021 - 08/2025

Bachelor of Engineering in Electronics and Telecommunication
Mumbai University
Terna Engineering College
08/2021 - 08/2024

SKILLS

Microsoft Excel, Microsoft Power BI, Microsoft Power Point, NLP, ML, DL, Pandas, Pug, Python, Random Forest, Scikit-Learn, Seaborn, SQL, Tableau, Gmail

PROJECTS

THERAPY CONNECT
01/2023 - 01/1970

Therapy Connect is a mental health support system leveraging advanced AI technologies

- Designed and deployed an end-to-end therapy support system delivering personalized, empathetic analysis of spoken thoughts.
- Built a Flask (Python) backend with a JavaScript frontend, ensuring smooth session handling, logging, and Linux-based deployment.
- Designed and implemented a BiLSTM model for six-core emotion classification, and integrated a RAG pipeline (MiniLM + ChromaDB) to deliver context-aware insights from 31 psychology books.

Funding Rounds Analytics Dashboard
08/2025 - 2025

Built an **end-to-end pipeline** to capture and analyze **equity & debt funding rounds** from filings, APIs, and press releases, storing structured data in SQL.

- Developed **interactive Power BI dashboards** for comparative analysis (Equity vs Debt, sectoral & regional trends, top investors).
- Automated **data cleaning, validation, and enrichment workflows** using Python & SQL, improving accuracy and efficiency by 40%.
- Applied **LEAN methodology & governance checks** to ensure SOP compliance and enhance data quality across funding datasets.

US TRAFFIC ACCIDENT SEVERITY
01/2023 - 01/1970

This project focuses on predicting traffic accident severity using machine learning techniques

- Built a predictive model in Python and SQL to classify accident severity, achieving 82.4% accuracy and 70.9% macro-average precision.
- Developed robust data pipelines with cleaning, feature engineering, and clustering methods.
- Performed data pipeline operations (cleaning, feature engineering, clustering).
- Applied DBSCAN and K-Means to uncover high-risk zones and accident density patterns, enabling actionable insights for traffic safety.

SBIN Stock Analysis Project
Date period

Collected and cleaned historical stock data for **State Bank Of India** then engineered key technical indicators (MA, RSI, MACD) to inform trend analysis.

- **Conducted trend detection** and time-series forecasting using ARIMA, Prophet, and LSTM models, evaluating performance via RMSE, MAE, and MAPE to support data-driven decision-making.
- **Automated analytical workflows** in Jupyter Notebook using Pandas, NumPy, and Python visualization tools (Matplotlib, Seaborn) for seamless data processing and result generation.

STRENGTHS

Analytical & Research Skills
Ability to identify patterns, trends, and anomalies in large datasets to extract actionable insights.

Data Handling & Automation
Skilled at collecting, cleaning, and automating workflows to improve efficiency and reduce manual effort.

Insight Communication
Ability to translate complex analysis into clear visualizations and business-relevant insights.