Pratik Sonawale

Karvenagar, Pune | pratiksonawale07@gmail.com | 9823423859Portfolio |
https://linkedin.com/in/pratik-sonawale-4309b2188 |

Summary

Dynamic and results-driven Data Scientist / ML Engineer with 1.8 years of experience in data analysis, machine learning, and geospatial analytics (GIS). Skilled in building end-to-end ML/DL pipelines, ETL workflows, and web-based AI applications using Python. Hands-on with Generative AI (LLMs, prompt engineering), computer vision, and model deployment using Flask and basic MLOps / CI-CD practices. Passionate about converting raw data into scalable, production-ready solutions that drive business value

Education

Shivaji University, Bachelors of Computer Application June 2019 – June 2022

- GPA: 9.5

  Key Subjects Studied- Software Development Life Cycle (SDLC), Project Management, Descriptive and Inferential Statistics, Hypothesis Testing, Networking Fundamentals, Data Transmission

Experience

PE-Fintech, Cognizant Technology Solutions Nov 2023 – July 2025

- Worked on data-driven projects involving large-scale datasets, responsible for data collection, ETL, data cleaning, and preprocessing for analytical and reporting use.
- Performed Exploratory Data Analysis (EDA), statistical analysis, and dashboard reporting using Excel, SQL, and Python to support decision-making.
- Wrote Python scripts and SQL queries to automate manual tasks, implement data quality checks, and improve operational efficiency.
- Contributed to basic CI/CD-style workflows by ensuring reproducible scripts, version-controlled analytics code (Git), and structured reporting processes.

Projects

Gemini RealEstate Assistant (GenAI, ML, Flask) `Source Code Link` | `Live Demo Link`

- Built an end-to-end AI Real Estate Advisor web application using Python, Flask, and Google's Gemini 2.5 Flash model.
- Utilized machine learning and deep learning techniques for financial pattern analysis and scenario modeling.
- Implemented prompt engineering, LLM-driven recommendation logic, data processing, and API integration to generate property insights.
- Designed a clean, responsive HTML interface for real-time interaction, offering market trends, risk evaluation, and investment suggestions.

Crop Recommendation `Source Code Link` | `Live Demo Link`

- Built an ML-powered web application that predicts the most suitable crop using soil nutrients (N, P, K), temperature, humidity, pH, and rainfall.
- Implemented data preprocessing and feature normalization, trained a Random Forest model, and deployed it using Flask.
- Achieved high-accuracy predictions through hyperparameter tuning, model evaluation, and real-time input validation.
- Integrated an end-to-end pipeline covering dataset handling, model training, serialization (pickle), UI design, and production-ready inference.
- Incorporated PyTorch and TensorFlow workflows for experimentation with neural-based recommendation enhancement.

Stress Score Prediction `Source Code Link | Live Demo Link`

- Built a lifestyle-based ML model that predicts stress levels using behavioural features such as sleep hours, diet score, work hours, screen time, caffeine/alcohol intake, and social activity.
- Performed data preprocessing, feature engineering, EDA, and feature scaling to improve model accuracy and stability.
- Implemented an end-to-end ML pipeline including data cleaning → model training → evaluation → model serialization (pickle) → API deployment..
- Demonstrated strong skills in ML deployment, API integration, Python programming, problem-solving, and full-stack ML development.
- Designed an interactive and responsive UI for a seamless prediction workflow and improved user experience.

Skills

Programming and Scripting:

- Python, R, SQL, basic Bash/Shell

Machine Learning and Deep Learning:

- Supervised and Unsupervised Learning (Regression, Classification, Clustering)
- Predictive Modeling, Feature Engineering, Model Evaluation, Cross-Validation
- Deep Learning (Neural Networks, basic CNN/RNN concepts)
- Generative AI, LLMs, Prompt Engineering, Reinforcement Learning

Frameworks and Libraries:

- Scikit-Learn, PyTorch, TensorFlow, Keras
- Pandas, NumPy, Matplotlib, Seaborn
- Basic XGBoost / Gradient Boosting experience
- OpenCV for image preprocessing and basic CV tasks
- Familiarity with NLP libraries (SpaCy, NLTK, Hugging Face ecosystem concepts

Data Engineering and MLOps:

- Data Cleaning, Data Wrangling, ETL Pipelines (Python + SQL-based)
- Model Deployment using Flask (ready for FastAPI migration)
- Familiarity with CI/CD concepts (Git, GitHub, automated testing ideas)

Visualization and Reporting:

- Excel, Power BI
- Matplotlib, Seaborn (Python-based visualizations)
- Dashboarding and storytelling with data

Business and Soft Skills:

- Cross-Functional Collaboration
- Requirement Gathering and Problem Solving
- Project Documentation and Reporting
- Stakeholder and Client Communication

Certifications

Python for Data Science - IBM

Time Series Analysis Stock Market Prediction Python - Great Learning

SQL and Relational Databases - IBM

Fine-Tuning Large Language Models with DPO and Hugging Face - IBM