

Prasad Shantaram Sangale

 E-mail |  +91-7385484184 |  github |  kaggle |  Portfolio

Education

B.E. in Artificial Intelligence & Data Science

S.B. Patil College of Engineering, Indapur – Savitribai Phule Pune University

2022-2026

CGPA: 9.10/10

Technical Skills

- **Programming & Data:** Python, SQL, Pandas, NumPy
- **Machine Learning:** Scikit-learn, XGBoost, Feature Engineering, Model Evaluation
- **Deep Learning & NLP:** PyTorch, Hugging Face Transformers, spaCy, NLTK
- **MLOps & Engineering:** MLflow, DVC, Git, GitHub Actions, Docker
- **Model Deployment:** FastAPI, Streamlit
- **Cloud Platforms:** AWS (EC2, S3)
- **Data Visualization:** Matplotlib, Power BI
- **Familiarity:** Large Language Models (LLMs), Retrieval-Augmented Generation (RAG) , Chromadb

Experience

AWS Trainee

Dec 2024 – Jan 2025

Rajavi Technologies Pvt. Ltd., Pune, India

- Gained hands-on experience with AWS services including EC2, S3, and RDS for deploying cloud-based applications.
- Learned to monitor infrastructure performance and optimize resource utilization through guided labs and practical exercises.

Projects

CrowdPulse – Cross-Source NLP Intelligence Platform [link](#)

- Built an end-to-end NLP platform analyzing Reddit and real-time news narratives.
- Applied DistilBERT for sentiment analysis and FASTopic for scalable topic modeling across diverse text sources.
- Built a semantic embedding pipeline to align related topics between social media discourse and news coverage using similarity-based matching.
- Integrated the Gemini API to generate narrative summaries for cross-source topic interpretation.
- built an interactive Streamlit dashboard to visualize sentiment distribution, topic overlap, and narrative insights in near real time.

Customer Churn Prediction System [link](#)

- Developed an XGBoost-based classification system to predict customer churn and identify high-risk segments.
- Engineered predictive features and evaluated model performance using ROC-AUC and classification metrics.
- Deployed an interactive Streamlit application and built Power BI dashboards to analyze churn drivers.

Neural Symptom-Based Disease Prediction [link](#)

- Built a PyTorch-based neural network to predict top-K probable diseases from patient symptoms.
- Implemented confidence-calibrated predictions to improve inference reliability.
- Designed a Streamlit interface for real-time diagnostic support.