Pranav Gujjar

Machine Learning Engineer | Data Scientist
Bangalore, India | +91-9008913366 | mr.pranavgujjar@gmail.com
LinkedIn: linkedin.com/in/pranav-b-gujjar | GitHub: github.com/Gujjar-Pranav
Portfolio: Live ML Systems & Case Studies
Visa: India (Citizen) | Open to UK / EU / US roles (sponsorship as per role)

Professional Summary

Machine Learning Engineer with an MSc in Data Science (Distinction) and hands-on experience designing, evaluating, and deploying production-style ML systems. Strong background in translating structured and unstructured data into actionable insights through robust modeling, explainability, and deployment-focused engineering. Experience spans NLP, speech processing, computer vision, and time-series forecasting, with emphasis on reproducibility, interpretability, and real-world evaluation.

Core Skills

- **Languages:** Python, SQL, PySpark, R, MATLAB
- **Machine Learning:** Supervised & Unsupervised Learning, Classification, Regression, Clustering, Feature Engineering, Feature Selection, Cross-Validation, Hyperparameter Tuning (Grid Search, Random Search), ROC-AUC, Precision-Recall, F1-score, MAE, RMSE
- **Deep Learning:** Neural Networks, CNNs, RNN, LSTM, GRU, Backpropagation, Dropout, Batch Normalization, TensorFlow, PyTorch
- **Computer Vision:** Image Preprocessing, CNN-based Models, U-Net, Dense U-Net, Image Segmentation, Data Augmentation, CLAHE, Patch-Based Learning, IoU, Dice
- **NLP & Speech:** TF-IDF, Linear SVM, Rule-Based NLP, Whisper ASR
- **Time Series:** ARIMA, Exponential Smoothing, RNN-based Forecasting
- **MLOps / Deployment:** FastAPI, REST APIs, Streamlit, Docker, CI/CD (GitHub Actions), Model Persistence
- **Data & Viz:** PostgreSQL, MongoDB, Tableau, Power BI, Matplotlib

Deployed ML Systems (End-to-End)

Live demos and source code: pranav-gujjar-portfolio.vercel.app

Retina-AI – Clinical Diabetic Retinopathy Screening Platform

- Built an AI-assisted screening system using retinal fundus images with PyTorch-based DR / No-DR classification and Grad-CAM explainability.
- Implemented confidence gating, risk stratification, and automated clinical PDF report generation.
- Developed full workflow including patient registry, role-based authentication, screening, and report management via Streamlit.
- Deployed with GitHub Actions CI, Ruff linting, and Streamlit Cloud.

Strategic Intelligence Stack – Customer Segmentation & Decision Intelligence Platform

- Built a production-grade segmentation system converting raw customer data into interpretable clusters and executive-ready insights.
- Implemented deterministic ML pipelines with run-based reproducibility, artifact persistence, and REST APIs.
- Developed scenario simulation engine enabling what-if analysis without retraining models.
- Deployed FastAPI backend with Next.js frontend (Vercel).

ReviewSense AI – Sentiment Intelligence Platform

- Designed and deployed sentiment analysis on **293 real-world reviews**.
- Selected TF-IDF + Linear SVM achieving **85.55% accuracy** and **0.853 F1-score**.
- Delivered dashboards for sentiment distribution and confidence-based decision support.

Glass Identification – ML Classification System

- Built ML pipeline with winsorization, feature engineering, and SMOTE imbalance correction.
- Implemented stacking ensemble achieving **90.70% test accuracy**.
- Deployed via FastAPI with Streamlit inference UI.

Diabetes Risk Assessment – Clinical Decision Support

- Developed interpretable Logistic Regression on **768 clinical records** achieving **ROC-AUC 0.813**.
- Implemented probability-based risk stratification with feature-level explanations.

Intelligent Task Miner – Audio-to-Task AI

- Built offline Whisper ASR + rule-based NLP pipeline converting meeting audio into structured task data.
- Exported normalized JSON outputs for workflow automation.

Work Experience

Machine Learning Engineer (Freelance) **Aug 2025 – Present**

- Designed end-to-end ML systems spanning ingestion, feature engineering, model training, evaluation, and deployment.
- Deployed inference services with FastAPI and Streamlit dashboards.

Data Science Intern **Jun 2022 – Dec 2022**

Vertexblue Pvt Ltd, India

- Improved forecasting accuracy by **15%**.
- Delivered analytics contributing to **10%+** cost reduction and reduced manual processing by **30%**.

Applied ML Studies

Recommendation System

- Built a content-based recommender using cosine similarity over **12,294 items** with feature normalization and ranking logic.
- Evaluated recommendation quality through similarity distribution analysis and cold-start handling.

Neural Networks & Time Series

- Implemented feedforward neural networks from scratch and sequence models (RNN, LSTM, GRU) for temporal learning.
- Built ARIMA-based statistical models and RNN forecasting pipelines on seasonal production and FX datasets.
- Analyzed convergence behavior, residual diagnostics, and out-of-sample forecast stability.

Clustering Analysis

- Applied K-Means, Hierarchical Clustering, and Gaussian Mixture Models to discover latent structure in unlabeled data.
- Evaluated clustering quality using Silhouette Score and Davies–Bouldin Index to guide model selection.

Academic Projects

Retinal Image Segmentation (MSc Dissertation)

- Developed U-Net and Dense U-Net models on the DRIVE dataset achieving **97.34% ROC-AUC**.
- Applied CLAHE preprocessing, patch-based learning, data augmentation, and evaluated using Dice and IoU.

AI Chatbot for UK Train Ticketing

- Built an NLP-based conversational system with intent classification and entity extraction for travel queries.
- Engineered features from **400k+** time-series records and benchmarked multiple ML models.

Education

MSc in Data Science (Distinction) Sept 2023 – Sept 2024

University of East Anglia, UK

Languages

English (Advanced), Hindi (Advanced), Gujarati (Native)