Mihir Parmar

+91 9875014540 | mihirparmar0913@gmail.com | LinkedIn | GitHub

Ahmedabad, Gujarat- 380015, India

Education

- **Lok Jagruti Kendra University,** *Nov 2021 - June 2025*
  Bachelor of Engineering (BE) | *Computer Science & Engineering* Ahmedabad, India
  - CGPA: 8.67

Experience

- **Data Science Intern** *Feb 2025 - Present*
  Petpooja Ahmedabad, India
  - Validated, annotated, and cleaned OCR-extracted invoice & receipt data (image + JSON) across multiple formats to ensure accuracy and reliability for downstream LLM models.
  - Working on a payroll face-attendance system that uses facial recognition to accurately verify employee check-ins and check-outs. The system identifies, classifies, and records false positives (FP) and false negatives (FN) to ensure high accuracy and reliability. It also captures detailed logs, including timestamps, confidence scores, and comparison data, which are later used for performance tuning and error analysis.
  - Designed and implemented web scraping scripts to collect structured data for product insights. Built an AI agent performance dashboard for business stakeholders and developers to track errors, identify improvement areas, and optimize workflows.Collaborated with AI/ML teams to align data pipelines with product goals, ensuring high-quality datasets for production use.

Technical Skills

- **Programming Languages:** Python, C, Java
- **Mathematics:** Probability, Statistics, Calculus, Linear Algebra
- **Data Analysis and Visualization:** Pandas, NumPy, Matplotlib, Seaborn
- **Machine Learning Libraries:** Scikit-learn, LightGBM, CatBoost, XGBoost
- **Deep Learning & NLP:** PyTorch, Keras, ANN, CNN, RNN, LSTM, GRU, Transformer, BERT
- **Generative AI:** RAG, LangChain,HuggingFace
- **Web Technologies:** HTML, CSS, JavaScript
- **Database Management:** SQL, Vector Databases
- **Version Control:** Git, GitHub
- **Other Tools:** Jupyter Notebook, Google Colab, Kaggle

Projects

- **AI-Powered Recipe Recommendation System (RAG-Based)**
  Tools: [FastAPI, ChromaDB, FastEmbed, Gemini API, Flask, Python]
  - Built a Retrieval-Augmented Generation (RAG) system that recommends Indian recipes based on user-provided ingredients using semantic search and vector embeddings.
  - Implemented ingredient-level fuzzy matching, synonym normalization, similarity scoring, and strict-mode filtering to improve recommendation accuracy.
  - Developed a full-stack application with a FastAPI backend (RAG + embeddings pipeline) and Flask frontend for recipe search, ingredient-based recommendations, and detailed recipe view.
- **Schemafin**
  Tools: [Python, PyQt5, JSON, shutil]
  - Developed a dual-pane desktop application using Python and PyQt5 to synchronize large-scale JSON datasets with invoice imagery, streamlining visual verification for document processing pipelines.
  - Designed a bidirectional synchronization engine supporting both a dynamic Form View for tabular entry and a RawJSONView,featuring a custom draggable mapping table for reordering Bank Statement (BSR) records.

- **Job Fraud Detection System (NLP + Machine Learning)**
  Tools: [Python, Scikit-learn, TF-IDF, Logistic Regression, Streamlit] GitHub[]
  - Built an NLP-based classification system to detect fraudulent job postings using text features from titles, descriptions, and requirements.
  - Addressed severe class imbalance (95:5) using class weighting and decision-threshold tuning.
  - Evaluated model performance using Precision–Recall curves instead of accuracy to handle imbalanced data.
  - Deployed the model using Streamlit to enable real-time fraud prediction from raw job descriptions.

CERTIFICATIONS

- **Machine Learning With Python**
- **Building Generative AI-Powered Applications with Python**
- **Generative AI: Introduction and Applications**

ACHIEVEMENTS

- **HackerRank: 4-Star Coder**
- **LeetCode: 50+ Problems Solved**