Priya Yadav

+91 8707845748 | priyayadav67426742@gmail.com | www.linkedin.com/in/priya-yadav-3a1806208/

SUMMARY

Skilled Machine Learning Engineer with expertise in developing high-performance models. Proficient in Python, TensorFlow, and data science tools, with a proven track record of improving model accuracy and driving impactful results.

EDUCATION

Madan Mohan Malaviya University of Technology, *B.Tech in CHE* | Gorakhpur   CGPA: 7.9 / 10.0   May 2023
Blossom Sr Sec School, *Intermediate* | Gorakhpur   Percentage: 85.4%   April 2019
Blossom Sr Sec School, *High School* | Gorakhpur   CGPA: 10 / 10   April 2017

EXPERIENCE

ZS Associate, *Artificial Intelligence Engineer* | Gurugram   Feb 2025- Present

- **Developed an end-to-end ML system for Stock Out (SO) & Stock Risk (SR) prediction,** automating supply chain decision-making through data preprocessing, feature engineering (anchor dates, PIR, lag features), and class balancing.
- Trained and optimized XGBoost classifiers using **Hyperopt**; evaluated with AUC, F1, and Confusion Matrix; applied **SHAP** for interpretability and Evidently AI for data drift monitoring and imbalanced class handling via **Tomek Links and Random Under-Sampling**, Amazon Web Services **AWS** for feature storing and result.
- Automated monthly predictions, report generation (CSV), and model serialization using **Pickle** for seamless deployment.
- **Built a dual-component framework:** .
  - Training Model: Automates retraining with new data to continuously improve the recommendation accuracy by processing problem-cause-solution triplets from raw operational comments.
  - User-facing Solution Model: Provides real-time solution recommendations by querying the knowledge base with material, country, date, and relevant features (PIR, INV, DOS rates).
- Preprocessed raw comment data and grouped comments into consistent conversations using **LLM (GPT-4.0), MultiThreading** for problem, cause, and solution extraction.
- Classified and categorized causes using **cosine similarity** and **Faiss**, improving solution accuracy through efficient cause matching and clustering and stored processed data (PIR, INV, DOS, triplets) in a **SQL database, creating a knowledge base** for fast, query-driven solution extraction.
- Developed a **Flask-based controller** managing both training and inference APIs; tested endpoints with **Postman** for robust performance.
- Created a **GPT-3.5-turbo powered medical chatbot** enhanced with **NER fine-tuning** ,**Spacy** for better query understanding
- Deployed the chatbot using **Flask** and integrated it with a SQL database for efficient data management.
- Applied **Object-Oriented Programming (OOP) principles and Python** best practices to develop modular and maintainable code.

Tata Consultancy Service, *Machine Learning Engineer* | Gandhinagar   Nov 2023 - Feb 2025

- Spearheaded the development of a machine learning model that stratified customers into **5** distinct segments, leveraging over **100,000+** purchasing records, which catalyzed a **30%** surge in customer engagement.
- Executed comprehensive data cleansing and normalization processes, addressing **5%** missing values and **3%** outliers, and engineered **10+** key features, resulting in a **15%** enhancement in model accuracy.
- Conducted in-depth exploratory data analysis (EDA) on a dataset comprising **1 million** data points, uncovering **3** major trends and **5** significant correlations using advanced visualization techniques such as histograms, box plots, and scatter plots.
- Applied a suite of clustering algorithms, including K-Means, Hierarchical Clustering, and DBSCAN, on a dataset of **100,000+** customer records, achieving a **25%** enhancement in clustering precision.
- Amplified conversion rates by **25%** and elevated the average order value by **20%** through meticulously targeted marketing strategies derived from model-driven insights.
- Reduced customer churn by **15%** and enhanced customer lifetime value (CLV) by **20%** through the deployment of data-driven retention methodologies.
- Streamlined inventory management across **50+ SKUs**, reducing stockouts by **20%** and mitigating overstock scenarios by **15%**.

SKILLS

Technical Skills : Python, C++, Java, SQL, HTML, JavaScript ,Machine Learning, Data Science, Deep Learning,Artificial Intelligence, NLP, Algorithms, Data Structure , YOLO, LLM, NLTK,Streamlit,OpenCV, GenAI, HuggingFace,AWS

Soft Skills Time Management, Problem-solving, Documentation, Engaging Presentation, Leadership, On-site coordination.

PROJECTS

Real-Time Object Detection Web App using YOLOv8, Flask, and Socket.IO Jan 2024 - April 2024

- Developed a real-time object detection web application using **YOLOv8, Flask, and Socket.IO**, enabling live video processing and frame-by-frame object detection.
- Trained YOLOv8 model for **35 epochs**, achieving **Precision: 83.4%, Recall: 79.3%, and mAP@0.5: 84.4%**, optimizing model generalization and accuracy
- Designed a Flask API to handle video file uploads, process frames using Ultralytics YOLOv8, and return real-time detection results.
- Implemented asynchronous, event-driven communication with WebSockets (Flask-SocketIO) for low-latency frame streaming from server to client
- Utilized **OpenCV (cv2)** for frame extraction, processing, and annotation, improving visualization of detected objects.
- Fine-tuned YOLOv8 with custom-trained weights on a **7-class dataset (anger, disgust, fear, happy, neutral, null, sad)** for specialized object classification
- Deployed the system in a **local development environment**, with potential scalability for cloud-based deployment using Docker and FastAPI.

Virtual Interview System using LLMs and Streamlit Jan 2025-

- Designed and developed an **AI-powered Virtual Interview** Application using **Streamlit LangChain**, and **GroqLLM** to conduct automated technical interviews.
- Integrated **GPT-based Large Language Models** (LLMs) to dynamically generate context-aware, domain-specific interview questions using **PyPDF2** for candidate resume parsing and **FAISS** for vector embeddings
- Implemented speech synthesis using **gTTS (Google Text-to-Speech)** to convert **interview questions into audio**, enhancing accessibility.
- Utilized **Whisper (OpenAI)** for high-accuracy **Speech-to-Text (STT)** conversion, enabling precise voice-to-text transcription of candidate responses
- Implemented **timer-based interview workflow**, **session state management** in Streamlit to track interview progress, question playback, and response submission.
- Integrated **LangChain Memory** to maintain context in **multi-turn interviews**, allowing follow-up questions based on previous responses
- Incorporated **JSON-based response storage** for structured data collection and efficient processing of candidate answers.

Tomato Leaf Disease Classification Using Convolutional Neural Networks Nov 2023 - Dec 2023

- Designed a Convolutional Neural Network (CNN) model to classify 10 distinct tomato leaf diseases, significantly aiding in early detection and management, with a final model accuracy of **94%**.
- **Leveraged Transfer Learning:** by fine-tuning a **pretrained ResNet50** model, enhancing classification accuracy by **10%** compared to traditional training methods.
- **Data Preprocessing and Augmentation:** Performed extensive preprocessing and augmentation on the dataset.
- **Evaluated model** performance using metrics such as accuracy, precision, recall, and F1-score to ensure effective disease classification, Achieved a classification accuracy of **94%** with the final model.

ACHIEVEMENTS

Tech Finalists PhonePe Women's Tech Scholar(2022-23) ,TCS Codevita Finalist(2022)