Pratik Kujur

6264902455 | pratikkujur121@gmail.com| PratikKujur | PratikKujur

Experience

PORTER | Machine Learning Intern Jan 2025 – Present

Skills: Python, Sklearn, Snowflake, Mlflow, Chalk AI, GCP, BigQuery, Streamlit *Bangalore, Karnataka*

- Developed and deployed production-grade ML models (Logistic Regression, XGBoost) on GCP , achieving **Precision: 0.79 and Recall: 0.76** with **100K+** daily predictions and ensure **99.9% uptime.**
- Orchestrated a cross-platform **ETL** pipeline using **Airflow DAG** integrating **BigQuery** and **Snowflake** for feature retrieval, processing **5M+** records daily for model training .
- Engineered **15+** features using **Chalk AI (Feature Store)** to strengthen training pipelines for Driver Ranking models and monitor city-wise model performance.
- Automated **real-time feature skewness monitoring** to ensure data stability and feature correctness.
- Contributed to **MLOPS** dashboard to observe **Concept drift** and **Feature drift** in Driver-Ranking models.

Education

National Institute of Technology Karnataka Surathkal, Karnataka

Master of Technology, Machine Learning *Aug. 2023 – July 2025*

Government Engineering College Raipur Raipur, CG

Bachelor of Technology, Computer Science and Engineering *Aug. 2019 – July 2023*

Projects

AI/ML Based Personalized Fitness Recommender. | Github| Project Link

Skills: Python, Scikit-learn, Google-Gemini, Streamlit

- Developed a personalized fitness recommendation/ranking system using an XGBoost model, leveraging **LDA-based feature selection** to maximize class separation and reduce noise, with ranking performance of **NDCG@10 of %81** and **Precision@10 of 84%**
- Integrated **Gemini 2.5 Flash** to generate dynamic personalized exercise + meal plans to increasing user engagement and enabling fully automated, real-time wellness recommendations.

Melanoma Skin Cancer Detection Using VGG16 | Github| Project Link

Skills: Python,TensorFlow, Keras, MLflow, DVC, AWS (S3, EC2), Flask

- Developed end-to-end deep learning system using **transfer learning** with **VGG16** architecture, achieving **87% accuracy** in binary classification of **10,000+** dermatoscopic images into benign and melanoma categories
- Implemented comprehensive **MLOps workflow** with MLflow for experiment tracking (15+ experiments) and DVC for dataset versioning, ensuring reproducibility and enabling rollback capabilities for model iterations

NLP-Based Hate Sentiment Classifier. | Github| Project Link

Skills:Python, TensorFlow, Keras, NLTK, FastAPI

- Architected LSTM-based sentiment analysis model for binary classification of 50K+ tweets, achieving **85% accuracy** and **F1-score of 0.83**.
- Preprocessed textual data using advanced NLP techniques including tokenization, lemmatization, and custom stopword removal, reducing vocabulary size by **60%** while preserving semantic information

Technical Skills

Languages: Python, C/C++, SQL
ML/DL Frameworks: TensorFlow, Keras, PyTorch, scikit-learn, XGBoost, NLTK, Langchain
MLOps & Tools: MLflow, DVC, Docker, Git, GitHub. Snowflake, Chalk AI (Feature Store)
Cloud & Deployment: AWS/GCP (EC2, S3), Flask, FastAPI, Streamlit
Data Science: pandas, NumPy, Matplotlib, Seaborn, Statistical Analysis, A/B Testing, Feature Engineering
Specializations: Computer Vision (YOLO, VGG16), NLP (LSTM, BERT, Transformers),LLm, RAG Systems

Achivements

Face Counting Challenge | Github

Skills: Python, YOLOv11.

- Analytics Vidhya Face Counting Challenge – **Ranked 24th** out of **500+** participants with **RMSE of 0.96** by **fine-tuning YOLOv11** object detection model on custom annotated dataset with 2000+ images