Bhavya Soni

bhavyasoni9587@gmail.com | linkedin.com/in/bhavyasoni01 | github.com/Bhavyasoni01

Summary

Computer Science undergraduate with hands-on experience building end-to-end machine learning and MLOps projects, including model training, explainability, experiment tracking, containerized deployment, and CI/CD on Google Cloud Platform. Seeking an internship in MLOps or ML Engineering.

Projects

Loan Default Prediction System | *XGBoost, SHAP, Optuna, MLflow, Pydantic, FastAPI, Docker, Kubernetes, GCP*

- Trained XGBoost on 2.26M Lending Club records with Optuna hyperparameter tuning; detected and removed post-loan **data leakage**, reporting an honest 0.71 ROC-AUC on application-time features only
- Engineered domain features: debt-to-installment ratio, payment-to-income ratio, and credit intensity score
- Built a SHAP explainability layer; a dedicated /explain endpoint returns the top factors driving each default prediction
- Deployed a dual-endpoint API (predict + explain) on GCP via Docker
- Automated CI/CD with GitHub Actions (metric validation on every push) and tracked experiments with an MLflow model registry for reproducible, staged model promotion
- Orchestrated deployment on Kubernetes (Minikube) with liveness/readiness probes, resource limits, and ConfigMap-based configuration; validated self-healing and horizontal replica scaling

Credit Card Fraud Detection API | *Scikit-learn, MLflow, FastAPI, Docker, GitHub Actions, GCP*

- Handled extreme class imbalance (fraud ∼0.16% of 94K transactions), achieving **0.98 ROC-AUC** with 0.85 recall and 0.82 precision on the fraud class — flagging 127 of 149 frauds at only 35 false positives
- Served the model as a low-latency REST API with FastAPI, containerized via Docker
- Managed model versioning and artifact tracking with MLflow for a reproducible train-to-deploy lifecycle
- Deployed on GCP with an automated GitHub Actions CI/CD pipeline enforcing testing quality gates

Telco Customer Churn Prediction | *Scikit-learn, MLflow*

- Built a churn classifier optimized for recall (0.84 on churners, 0.84 ROC-AUC), prioritizing catching at-risk customers over raw accuracy
- Tracked and compared model and hyperparameter runs with MLflow to select the best performer

Education

Manipal University Jaipur — Jaipur, Rajasthan
Bachelor of Technology in Information Technology — *Aug. 2024 – Aug 2028*

- CGPA: 7.49 / 10 (Most recent semester: 8.38 – upward trend)
- Relevant Coursework: Data Structures & Algorithms, DBMS, Operating Systems, Object-Oriented Programming, Computer Networks

Technical Skills

Languages: Python, Java, C
ML / Data: scikit-learn, XGBoost, pandas, NumPy, Matplotlib, SHAP, Optuna
MLOps & Tools: FastAPI, MLflow, Docker, Kubernetes, GitHub Actions, Git, Pydantic
Cloud: Google Cloud Platform (Compute Engine)