Mansi Kaushik

+91-8839914681 | mansikaushik0705@gmail.com | LinkedIn | Github

Professional Summary

Computer Science undergraduate with experience in data analysis, predictive modeling, and scalable data pipelines. Proficient in Python, SQL, Pandas, Scikit-learn, and Power BI, focused on extracting insights and deploying data-driven solutions..

Education

Vellore Institute of Technology, India **Sep 2023 – May 2027**

BTech in Computer Science and Engineering (CGPA: 8.8)

Technical Skills

- **Languages:** Python, Java, SQL .
- **Libraries/Frameworks:** Numpy, Pandas, Matplotlib, Scikit-learn, TensorFlow, PyTorch, Flask, OpenCV.
- **Cloud Deployment:** AWS, Docker, Streamlit .
- **Tools:** Power BI, Jupyter Notebook, Git/GitHub

Research, Patents & Publications

IEEE Xplore

(A Multimodal Narration System for Enhancing Accessibility for the Visually Impaired)

- Published in IC-CGU 2025: Developed an AI-driven OCR/TTS system for real-time visual-to-audio narration.
- Achieved 90.96% OCR accuracy and 92.77% F1-score; optimized TTS module for 4.14/5 mean user rating.

Patent (Under Review)

(A Distributed Meta-Learning Framework for Rapid Anomaly Adaptation in Non-IID Financial Networks)

- Developed a novel Federated Meta-Learning framework that replaces global model convergence with meta-initialization, enabling few-shot anomaly detection (5 samples) in highly non-IID financial environments.
- Introduced a meta-gradient exchange mechanism with bi-level optimization, achieving faster edge adaptation and significantly reducing communication overhead and latency in decentralized systems.

Projects

Explainable Credit Risk Platform

(Python, XGBoost, SHAP, Scikit-learn, Streamlit, Pandas)

- Architected an end-to-end ETL pipeline and performed extensive Exploratory Data Analysis (EDA) on 1.3M+ records to identify multi-collinearity and distribution shifts, ensuring high data integrity.
- Developed a high-performance XGBoost classification model optimized for precision-recall, achieving a significant reduction in high-cost False Negatives within the credit lending lifecycle.
- Integrated Explainable AI (XAI) using SHAP to provide granular, feature-level transparency for model decisions, ensuring alignment with Fintech regulatory requirements for model interpretability.
- Deployed a real-time interactive dashboard using Streamlit to bridge the gap between complex model outputs and stakeholder-facing risk assessment tools, enhancing decision-making efficiency.

T2-Hydro: Predictive Geospatial Analytics

(Python, PyTorch, SQL, STAC API, LSTM, SHAP)

- Engineered a high-throughput data ingestion pipeline utilizing STAC APIs to process multi-gigabyte Sentinel-2 imagery, reducing raw data preprocessing latency by 40%.
- Developed a Dual-Stream CNN-LSTM architecture for drought severity forecasting, achieving 85% predictive accuracy by fusing high-resolution satellite imagery with real-time meteorological sensor data.
- Implemented Transfer Learning to optimize model deployment in data-scarce environments, reducing required training compute by 80% compared to baseline models.
- Deployed SHAP-based explainability to deliver feature-level interpretability, enabling risk-management teams to correlate reservoir levels with macro-economic indicators.

Certifications

- Applied Machine Learning in Python- University of Michigan.
- Data Analysis with Python– freeCodeCamp.