Abhiram M V

abve5411@colorado.edu | 303-668-0982 | Boulder, CO | LinkedIn | Portfolio | GitHub

EDUCATION

University of Colorado Boulder   **Boulder, CO**
Master of Science in Data Science (GPA: 3.63/4)   August 2024 - May 2026 (Expected)

Dayananda Sagar University   **Bengaluru, India**
Bachelor of Technology in Electronics and Communication Engineering (GPA: 7.91/10)   August 2018 - May 2022

SKILLS

Languages/Libraries: Python (Pandas, Numpy, NLTK, Statsmodels, Scikit-learn, PyTorch, TensorFlow), SQL (PostgreSQL), Go
Frameworks/Technologies: AWS (S3, SageMaker, EC2), Apache Spark, Airflow, Tableau, Git
Competencies: Data Visualization, Statistical Modeling, Predictive Modeling, Machine Learning, Time Series Forecasting, Deep Learning, NLP, CV, RAG, Causal Inference

RELEVANT EXPERIENCE

ConverSight   **Indianapolis, USA**
Product Research Intern   May 2025 – August 2025
- Developed an AI voice application using FastAPI to conduct supplier assurance calls. Integrated OpenAI and LangChain for dynamic tool invocation, extracting and storing insights in PostgreSQL, saving up to **5 hrs/day** and reducing procurement cycle times by **40%**
- Built a RAG-based document Q&A system as a POC for a manufacturing client using Python, Qdrant, AWS S3, and OpenAI. Implemented hybrid retrieval (BM25 + vector similarity) for more accurate, reliable answers and to reduce answer refusal rates

ConverSight   **Coimbatore, India**
Data Scientist   August 2022 – July 2024
- Deployed a CLV estimation model using LightGBM, scoring **8%** in MAPE for CLV estimation and **0.93** in AUC for churn prediction, resulting in a **25%** increase in retention of high-value customers for beauty products
- Led pipeline development to detect fraudulent UPI transactions from a database of millions using PySpark with CatBoost, achieving **96%** recall in fraud detection and **89%** recall in chargeback detection, reducing financial exposure for banks and merchants by **20%**
- Architected an automated demand-forecasting system for **5K**+ SKUs using Prophet, ARIMA, and Holt-Winters, improving forecast accuracy by **20%** (**11%** MAPE) and preventing over **$200K** in annual stock-out and excess inventory costs
- Fine-tuned LLaMA-2, Mistral-7B, and GPT-3.5-Turbo models on domain-specific data using LoRA/QLoRA and Chain-of-Thought prompting to increase the AI assistant's text-to-SQL accuracy from **85%** to **97%** on a custom evaluation dataset

PROJECTS

Walmart Forecasting [Link]
- Engineered a forecasting system for **30K**+ Walmart SKUs by deploying a fully automated AWS batch prediction pipeline (SageMaker, Lambda, S3) with an ensemble model (LightGBM, LSTM) that ranked in the top **10%** on Kaggle for accuracy (**0.48** WRMSSE)

Feedback Prize - Evaluating Student Writing [Link]
- Architected a multi-stage NLP model by ensembling Longformer and BigBird transformers with an XGBoost to identify elements in student essays, boosting the final F1 score to **0.74** via a custom Weighted Box Fusion post-processing pipeline

Causal Uplift Modeling for Marketing Campaign Optimization [Link]
- Optimized marketing campaign ROI by developing a causal uplift model (Python, CausalML, XGBoost) to identify and target persuadable customers, demonstrating that targeting the top **30%** of users could capture over **80%** of the total potential campaign lift

BUFF-RoSTOREing [Link]
- Engineered an automated shift assignment system leveraging Python, Airflow, and an event-driven AWS architecture (S3, Lambda, SQS, SES) to slash weekly scheduling time by **95%** (from 1 hour to 3 minutes) and ensure **100%** coverage of critical shifts

Fire in Focus: Analytical Approach to Wildfire Analysis [Link]
- Experimented with machine learning models (XGBoost, Decision Tree, Random Forest, SVM) to predict wildfires in Southern California using NASA FIRMS and Open-Meteo data, achieving **83.5%** accuracy (Random Forest) and a **0.94** AUC-ROC (XGBoost)

Solar Panel Dust Detection [Link]
- Built and deployed a dust detection system for solar panels using a fine-tuned MobileNetV2 model exposed through a Flask REST API, achieving **90%** classification accuracy