

Abhiram M V

abve5411@colorado.edu | 303-668-0982 | Boulder, CO | [LinkedIn](#) | [Portfolio](#) | [GitHub](#)

EDUCATION

University of Colorado Boulder

Master of Science in Data Science (GPA: 3.63/4)

Boulder, CO

August 2024 - May 2026 (Expected)

Dayananda Sagar University

Bachelor of Technology in Electronics and Communication Engineering (GPA: 7.91/10)

Bengaluru, India

August 2018 - May 2022

SKILLS

Languages/Libraries: Python (Pandas, Numpy, NLTK, Statsmodels, Scikit-learn, PyTorch, TensorFlow), SQL (PostgreSQL)

Frameworks/Technologies: AWS (SageMaker, S3, Lambda, Step Functions, Glue), Apache Spark, Airflow, Docker, Tableau, Git

Competencies: Data Visualization, Statistical Modeling, Predictive Modeling, Machine Learning, Time Series Forecasting, NLP, RAG, Deep Learning, Causal Inference

PROFESSIONAL EXPERIENCE

ConverSight

Product Research Intern

Indianapolis, USA

May 2025 – August 2025

- Developed an AI voice agent (FastAPI, Twilio, OpenAI, LangChain) with dynamic tool invocation to automate supplier assurance calls, eliminating **20** hours/week of manual effort by extracting and storing insights in PostgreSQL to accelerate procurement cycles
- 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

Data Scientist

Coimbatore, India

August 2022 – July 2024

- Deployed a CLV estimation model using LightGBM for beauty products, scoring **8%** in MAPE and **0.93** in AUC for churn prediction, reducing churn by **9%** over **6** months and **13%** profit uplift over **12** months through targeted retention and personalized campaigns
- 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

Cognizant Technology Solutions

Programmer Analyst

Bengaluru, India

January 2022 – July 2022

- Engineered a pipeline 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%**

PROJECTS

Twitter Sentiment Analysis [\[Link\]](#)

University of Colorado Boulder | January 2025 – April 2025

- Engineered a distributed machine learning pipeline in PySpark to process over **1.6 million** tweets, leveraging advanced feature engineering (TF-IDF, N-grams, ChiSqSelector) and cross-validation to achieve an **81%** F1-score with a Logistic Regression model
- Designed and built a real-time ETL pipeline using Docker for containerization, Kafka for tweet ingestion, and Spark Structured Streaming for processing and classification, storing predictions in MongoDB and Delta Lake for downstream analytics

LLaMA Squad [\[Link\]](#)

University of Colorado Boulder | January 2025 – April 2025

- Developed a QA model by fine-tuning LLaMA 2 and 3 in PyTorch with QLoRA, achieving **~75%** Exact Match on SQuAD 2.0, a **~300%** improvement over the baseline while boosting correct abstention on unanswerable questions from **4%** to **82%**, reducing model hallucination

Walmart Forecasting [\[Link\]](#)

University of Colorado Boulder | September 2024 - December 2024

- Engineered a high-performance forecasting system for **30K+** Walmart SKUs by creating hundreds of predictive features (lags, rolling stats, price momentum) and ensembling LightGBM with an LSTM, achieving a top-**10%** ranked **0.48** WRMSSE score on Kaggle
- Architected a fully automated, serverless pipeline on AWS using Step Functions to orchestrate SageMaker Batch Transform jobs, reducing weekly forecast generation time from days to one hour and delivering actionable insights via QuickSight dashboards

Feedback Prize - Evaluating Student Writing [\[Link\]](#)

May 2024 – July 2024

- Devised a multi-stage NLP model by ensembling Longformer/BigBird transformers with an XGBoost built on learned token embeddings to identify elements in student essays, boosting the final F1 score to **0.74** with a custom Weighted Box Fusion post-processing pipeline to merge ensemble predictions (Matched the SOTA performance of the competition's 2nd place solution)

Causal Uplift Modeling for Marketing Campaign Optimization [\[Link\]](#)

October 2023 – December 2023

- 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 **80%** of the total potential campaign lift