Abhiram M V

Boulder, CO | LinkedIn | 303-668-0982 | abve5411@colorado.edu | 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/Tools: Python, SQL, R, Tableau, AWS (IAM, S3, Lambda, EC2, SageMaker), PySpark, Streamlit, Docker, Git
Frameworks/Libraries: Pandas, NumPy, Matplotlib, Seaborn, Statsmodels, Scikit-learn, Keras, TensorFlow
Competencies: Data Analysis, Data Visualization, Statistical Analysis, Predictive Modeling, Machine Learning, Deep Learning

RELEVANT EXPERIENCE

ConverSight — **Coimbatore, India**
Data Scientist — August 2022 – July 2024

- Developed an automated demand forecasting pipeline using Prophet, ARIMA, and Holt-Winters with dynamic model selection and top-down disaggregation, improving forecast accuracy by **20%** and reducing inventory turnover costs by **$200K** annually
- Built a document Q&A system using Retrieval-Augmented Generation (RAG) with Python, Go, AWS S3, and OpenAI, enhancing answer relevance via hybrid semantic-sparse search and enabling traceability of source content for validation
- Fine-tuned LLaMA-2, Mistral-7B, and GPT-3.5-Turbo using LoRA/QLoRA and Chain-of-Thought (CoT) prompting on domain-specific queries, increasing AI assistant's text-to-SQL accuracy from **85%** to **97%**

Cognizant Technology Solutions — **Bengaluru, India**
Programmer Analyst — January 2022 – July 2022

- Engineered a real-time fraud detection pipeline on millions of UPI transactions using SQL, PySpark, and CatBoost within Apache Zeppelin, achieving **96%** fraud recall and **89%** chargeback recall, reducing financial risk for banks and merchants

PROJECTS

Instacart Market Basket Analysis [Link]

- Segmented **200K**+ Instacart users using PCA + K-Means clustering on aisle-level purchase behavior, identifying **5** distinct profiles (e.g., fruit-heavy, water-focused), enabling targeted marketing strategies, promotions and discounts
- Built an XGBoost classifier with **30**+ features (e.g., reorder ratio, cart position) to predict whether a product will be reordered, achieving an AUC score of **0.835**, supporting personalized product recommendations and improving retention

Fire in Focus: Analytical Approach to Wildfire Analysis [Link]

- Engineered predictive models by merging NASA FIRMS and Open-Meteo data (2020–2024), transforming satellite and weather variables into a labeled dataset of **113K** rows via spatial joins and class balancing
- Developed and tuned ensemble classifiers (LightGBM, XGBoost) with Optuna and BayesSearchCV, achieving AUC-ROC of **0.94** and identifying key predictors (temperature, humidity, soil features) to enable proactive fire risk alerts across California

Customer Loyalty Prediction Using ML

- Predicted loyalty scores for **50%** of untagged customers using a Random Forest model (Adj. R^2 = **0.955**), recovering critical loyalty insights to improve accurate customer tracking, targeting, and comms
- Deployed an ensemble model via Amazon SageMaker endpoint after AutoML training with SageMaker Canvas, improving Mean Squared Error (MSE) and streamlining model delivery

Starbucks Nutritional Analysis [Link]

- Constructed a MLR model in R to predict the calorie content of **240**+ Starbucks beverages based on key nutrients (Fat, Sugars, Carbs). Achieved an MSPE of ~**44** on the reduced model, optimizing predictors through VIF, adjusted R^2, and stepwise AIC

Solar Panel Dust Detection [Link]

- Fine-tuned a MobileNet CNN using transfer learning to classify dusty solar panels from aerial imagery, achieving **87%** accuracy and **0.856** AUC. Deployed the model as a Flask API for real-time inference by allowing users to upload images