

AFTHAB MUHAMMED H

afthab.m.h2004@gmail.com | +91 8113058003 | Trivandrum, Kerala
linkedin.com/in/afthab1001

Profile

Integrated M.Sc. student in Computational Statistics and Data Analytics with hands-on experience in Machine Learning, NLP, Data Engineering, and Real-Time Analytics. Skilled in building predictive models, distributed data pipelines, and anomaly detection systems using Python, SQL, Apache Spark, Kafka, and modern ML frameworks.

Education

Vellore Institute of Technology, Vellore

Sept 2022 – May 2027

Integrated M.Sc. in Computational Statistics and Data Analytics

CGPA: 8.73 , Class XII: 75.6% , Class X: 91.2% .

Projects

Large-Scale Market Surveillance System for Trade Abuse Detection

Dec 2025

- Built a real-time surveillance pipeline processing 200,000+ cryptocurrency trade records using Apache Spark Structured Streaming and Kafka.
- Implemented Wash Trading detection using Z-score based anomaly detection and Pump-and-Dump detection using price and volume analysis.
- Developed PostgreSQL-backed storage, automated alerting, and Streamlit dashboards for low-latency monitoring.

Heart Disease Risk Prediction System

Dec 2024 – Feb 2025

- Developed a predictive classification model using the UCI Cleveland Heart Disease dataset with 13 clinical features.
- Achieved **98.7% classification accuracy** using XGBoost after preprocessing, feature scaling and model optimization.
- Evaluated performance using ROC-AUC, Precision-Recall Curve and Confusion Matrix.

Financial Named Entity Recognition System

Jan 2026

- Built a domain-specific Named Entity Recognition model using spaCy and the Financial PhraseBank dataset.
- Trained custom extraction for COMPANY, MONEY, DATE and PERCENT entities.
- Achieved **100% Precision, Recall and F1-Score** on the evaluation dataset.

Quantum Dynamic Jellyfish Swarm Optimization for Image Segmentation

2026

- Developed a hybrid segmentation framework combining QDJFSO and Fuzzy C-Means clustering.
- Benchmarked against 11 optimization algorithms across 30 independent experimental runs.
- Achieved the lowest mean fitness score of **34.98** and standard deviation of **0.0585**.

Technical Skills

Programming: Python, R, SQL

Machine Learning: XGBoost, Random Forest, SVM, KNN, Clustering, SMOTE, Feature Engineering

Deep Learning: PyTorch, TensorFlow

Data Engineering: Apache Spark, Hadoop HDFS, Apache Kafka, PostgreSQL, BigQuery, dbt, ETL Pipelines

Visualization: Tableau, Power BI, Matplotlib, ggplot2, Streamlit

Cloud & Tools: Oracle Cloud Infrastructure (OCI), Git, Jupyter, NLP

Certifications

- Oracle Cloud Infrastructure 2025 Certified AI Foundations Associate
- Tableau Certification Training: Desktop, Prep and Cloud
- Analytics Engineering Bootcamp (dbt, BigQuery, Snowflake)
- Complete SQL Bootcamp
- Microsoft Excel – Beginner to Advanced

Languages

English (Fluent), Malayalam (Fluent), Tamil (Fluent), Hindi (Fluent)