

Ahmed Ikram | ML & AI Engineering

Dundee, Scotland | +44 7421 996048 | ahmedikram30@gmail.com | [LinkedIn](#) | [GitHub](#) | [Online Portfolio](#)

PROFILE

3rd-year Computer Science student at the University of Dundee on track for a First-Class degree, seeking an ML or AI Engineering placement or internship. Focused on building and evaluating end-to-end ML pipelines and LLM-powered applications - from feature engineering and model selection through to RAG architecture, local inference, and retrieval quality evaluation. Experienced in applying ML rigorously: leakage prevention, cross-validated benchmarking, and LLM-as-a-judge evaluation pipelines.

EDUCATION

High School of Dundee

Aug 2017 - June 2023

- National 5 Computer Science, Maths, English, Economics, Physics, Chemistry, Biology (A)
- Higher Computer Science, Maths and English (A), Higher Chemistry (B) and Higher Physics (C)
- Advanced Higher Computer Science (C)

University of Dundee - BSc (Hons) Computer Science (Data Science and AI)

Sep 2023 - Present

- Expected Graduation: June 2027 (with year-long placement, June 2028)

PROJECTS

ATM Log Aggregation, Analysis, Diagnostics Platform | Python, FastAPI, SQLite, LangChain, ChromaDB, Ollama

- Led a 7-person Agile team as backend/data engineering lead, delivering a production system for NCR Atleos
- Built a fully local RAG-based diagnostic assistant (LangChain, ChromaDB, Ollama) enabling natural language querying of ATM log history - architected as an air-gapped system ensuring no log data leaves the local network, directly aligned with banking-grade data privacy requirements
- Implemented semantic chunking via LangChain's SemanticChunker for context-aware log segmentation, with nomic-embed-text embeddings and ChromaDB as the vector store; inference served locally by llama3.1:8b
- Evaluated retrieval and generation quality using an LLM-as-judge pipeline - scoring relevance, faithfulness, and answer completeness across a held-out set of diagnostic queries
- Built a synthetic data generator producing statistically consistent anomaly distributions across 7 ATM anomaly/fault types, used for end-to-end pipeline validation and anomaly detector stress testing
- Worked from a formal requirements specification - translating 12 functional and 8 non-functional requirements and user stories with acceptance criteria into technical design decisions; 48 automated tests mapped directly to formal acceptance criteria ensuring full requirements traceability

Haggis Data Mining Project | Python, Scikit-learn, XGBoost, Matplotlib, Jupyter Notebook

- Engineered and validated 2 novel ratio-based features that emerged as top-3 predictors across all models; prevented data leakage using sklearn Pipelines with ColumnTransformer and RobustScaler throughout
- Trained and compared 7 classifiers including XGBoost, Random Forest, and Logistic Regression achieving ~90% accuracy; tuned hyperparameters via GridSearchCV with 5-fold cross-validation
- Applied K-Means and DBSCAN clustering to validate natural groupings against ground-truth labels; built a Linear Regression model predicting body mass ($R^2=0.756$), validating assumptions via residual diagnostics

DevSync Project Tracker with GitHub Integration | React, Flask, Socket.io, GitHub API, Pytest, AWS, Docker

- Deployed on AWS using ECS Fargate behind an Application Load Balancer in a custom VPC; RDS isolated in a private subnet with least-privilege security groups; frontend on S3 + CloudFront with HTTPS via ACM; CI/CD via GitHub Actions with OIDC federation - no static credentials
- Implemented a complete GitHub OAuth 2.0 flow and API integration, allowing users to securely connect repositories and bi-directionally link internal platform tasks to live GitHub Issues and Pull Requests
- Enforced via GitHub Actions CI/CD: 541 total tests (271 Pytest backend at 80% line coverage gate, 270 Jest frontend at 90% line coverage, 5 Cypress E2E suites) gate every PR - deployment aborts on any failure across either track

StockLens FinTech Spending & Investment App | React Native Expo, TypeScript, Firebase, Node.js, Jest

- Engineered an ML layer combining ARIMA-based time-series forecasting for per-category spending trends and Linear Regression for forward-looking portfolio projections, fed by historical CAGR computed from Alpha Vantage data - surfaced directly in the app UI alongside the OCR receipt pipeline
- Full-stack mobile app in React Native/TypeScript with secure biometric authentication, real-time Firebase synchronisation, and custom AES encryption for sensitive data at rest

PROJECTS

CineMatch AI / ML Movie Recommendation System | Python, Scikit-learn, Flask

- Implemented three recommendation approaches - collaborative filtering, content-based filtering, and a hybrid model - on the MovieLens dataset; collaborative filtering achieved ~78% hit rate and ~0.22 Precision@10
- Systematically compared all three approaches across hit rate, Precision@10, and cold-start handling — hybrid model outperformed standalone methods by combining user similarity signals with item content features, with each strategy independently swappable via dependency injection for controlled benchmarking

TECHNICAL SKILLS

Languages	Python, TypeScript, JavaScript, SQL, Bash
ML & AI	Scikit-learn, XGBoost, PyTorch, Pandas, NumPy
LLM Engineering	LangChain, ChromaDB, Ollama, Vector Search, Evaluation Pipelines
Data & Backend	Apache Airflow, ETL/ELT, SQLite, PostgreSQL, NoSQL, MySQL, REST APIs, FastAPI
Cloud & Tools	AWS (EC2, S3, RDS, IAM, SageMaker), Docker, Azure, GitHub Actions, Linux
Visualisation	Jupyter Notebook, Matplotlib

CERTIFICATES

<u>AWS Academy Graduate - Machine Learning Foundations</u>	Feb 2025
<u>AWS Academy Graduate - Cloud Foundations</u>	Mar 2024
<u>Foundations of Azure AI: Concepts, Capabilities, and Implementation</u>	Oct 2024

EXPERIENCE

Beaulieu Mini Market, Dundee, Scotland

Retail Assistant

September 2023 - Present

- Diagnosed and resolved Point-of-Sale and stock system failures rapidly to minimise service disruption

Zulus Peri Peri Grill, Dundee, Scotland

Front of House Worker

June 2021 - September 2023

- Trained 5 new staff members; managed competing priorities across dine-in, takeaway, and delivery orders under sustained pressure

TRANSFERRABLE SKILLS

Model Evaluation & Benchmarking

Benchmark models rigorously - cross-validation, leakage prevention, held-out evaluation sets, and LLM-as-judge pipelines - not just headline accuracy scores

Experimental & Hypothesis-Driven Thinking

Frame ML problems as falsifiable hypotheses; validate with held-out sets and residual diagnostics before drawing conclusions

LLM Integration & RAG Architecture

Design and evaluate end-to-end LLM-powered systems - chunking strategy, embedding selection, vector retrieval, and generation quality as first-class architectural concerns

Data Pipeline Engineering for ML

Build production-grade data pipelines that feed ML systems reliably - ingestion, transformation, synthetic data generation, and anomaly injection

Self-directed Learning

Independently acquired AWS ML and Azure AI certifications alongside full-time study; consistently extends technical depth beyond coursework

Research Methodology

Apply structured research practices, controlled evaluation, feature validation, and statistical testing, to ensure findings generalise beyond training data

Technical Communication

Present model results, retrieval quality metrics, and system trade-offs clearly to both technical peers and non-technical stakeholders

Cross-functional Collaboration

Delivered ML and AI components within a multi-disciplinary Agile team; experienced coordinating across data engineering, frontend, and analysis workstreams