

Megh Niles Kashilkar

Bristol BS1 5UL, UK • +44 7352 678457 • meghtreasurer@gmail.com
linkedin.com/in/megh-kashilkar13 • github.com/MeghKashilkar

Summary

MSc Data Science student at the University of Bristol with a Computer Engineering background, building end-to-end machine learning projects in Python — from imbalanced-classification and NLP through to deployed, tested services. Seeking a graduate role in the data field (Data Scientist / Data Analyst / ML Engineer).

Skills

Languages: Python (pandas, NumPy, scikit-learn, XGBoost), SQL

Machine Learning: classification, regression, imbalanced-data & cost-sensitive modelling, SHAP explainability, model selection & cross-validation, A/B testing

NLP & GenAI: spaCy, LLM applications, RAG, agentic tool-calling, prompt-injection defence (Gemini API)

Deployment & MLOps: FastAPI, Streamlit, Docker, Git, pytest, Render, Vercel

Visualisation & BI: Tableau, matplotlib

Projects

Credit-Card Fraud Detection — imbalance, cost-sensitive decisions & GenAI analyst

Python, XGBoost, SHAP, FastAPI, Gemini API, Docker

2026

- Problem: detect fraud in a highly imbalanced transaction stream (~0.6% fraud), where a missed £2,000 fraud and a false alarm carry very different costs.
- Approach: engineered interpretable features on a 200,000-transaction time-ordered split; compared Logistic Regression, Random Forest and XGBoost on PR-AUC; added a cost-sensitive approve/review/block threshold layer and a Gemini GenAI analyst (RAG with citations, PII redaction, prompt-injection defence).
- Outcome: XGBoost reached **0.895 PR-AUC** (0.93 recall); cost-optimal thresholds cut fraud losses by ~98% (£122k) vs no model on a 40,000-transaction test set. Deployed as a FastAPI + web-UI service; 37 tests passing.

Customer Churn Prediction System

Python, scikit-learn, SHAP, FastAPI, Streamlit, Docker

2026

- Problem: predict which telecom customers are about to churn so retention teams can intervene.
- Approach: cleaned and feature-engineered 7,000+ customer records; compared four classifiers with 5-fold stratified cross-validation; added global and per-prediction SHAP explainability.
- Outcome: Gradient Boosting achieved **0.839 ROC-AUC** under ~27% class imbalance; shipped a shared FastAPI API and Streamlit app (batch scoring 7,000 rows in a single vectorised call), deployed on Render and Streamlit Cloud.

UK Data & Tech Job Market Tracker — NLP skill-demand & salary analytics

Python, spaCy, XGBoost, FastAPI

2026

- Problem: measure real, current UK data/tech skill demand and salaries from live, unlabelled job postings.
- Approach: self-sourced 1,460+ postings via the Adzuna API; built a 56-skill spaCy PhraseMatcher taxonomy, rule-based role/seniority classification, and a salary regressor (four models compared).

- Outcome: XGBoost salary model reached **R^2 0.308** (MAE £15k); served through a FastAPI backend with a dependency-free HTML/JS dashboard and hand-built SVG charts; 71 tests, deployed on Vercel + Render.

Experience

Personal Shopper — Dot Com (Online Fulfilment)

Tesco

Bristol

Feb 2026 – Present

- Pick and pack online grocery orders to tight per-order time targets, maintaining accuracy while working under pressure.
- Assist in-store customers in locating products, delivering clear and helpful customer service.
- Apply close attention to detail to ensure order accuracy and correct substitutions.

Education

MSc Data Science

University of Bristol • Result pending

Bristol, UK

2025 – 2026

B.E. Computer Engineering

AISSMS IOIT • CGPA 7.95/10 — First Class with Distinction

Pune, India

2021 – 2025

Languages

English (fluent), Hindi (fluent), Marathi (native)