

Amal S

amalsreekumar94@gmail.com — +91 7356902532 — Kollam, Kerala
linkedin.com/in/amal-s — github.com/Amalsreekumar1 — Portfolio

Summary

Data Science student specializing in Machine Learning and AI systems, with hands-on experience in building explainable models, retrieval-augmented generation (RAG) pipelines, and code intelligence systems. Skilled in Python, Scikit-learn, and deep learning fundamentals, with practical experience in GraphCodeBERT, SHAP, and LangChain. Passionate about developing scalable and interpretable AI solutions.

Experience

Data Analyst Intern, Unified Mentor 11/2024 – 12/2024

Remote

- Gathered information from various sources, performing data validation and cleansing using Python and Excel for business analysis.
- Developed interactive dashboards in Power BI to visualize trends, produce and track key performance indicators (KPIs), and present valid inferences to management.
- Analyzed complex datasets to deliver insights that identified areas to increase efficiency, optimize business processes, and automate data workflows.

Education

Bachelor of Computer Applications in Data Science 08/2023 – 05/2026

Amrita Vishwa Vidyapeetham, Amritapuri, Kollam, Kerala
Expected Graduation: May 2026 — CGPA: 8.43

12th Grade, Science 11/2021 – 03/2023

GOVT. Model Higher Secondary School, Kulasekharapuram

Percentage: 87%

10th Grade 06/2020 – 03/2021

GOVT. Model Higher Secondary School, Kulasekharapuram

Percentage: 98%

Skills

- **Programming:** Python, SQL, HTML, CSS, JavaScript.
- **Libraries & Tools:** Pandas, NumPy, Matplotlib, Power BI, Excel.
- **Machine Learning:** Scikit-learn, GraphCodeBERT, SHAP, LangChain, FAISS, BM25.
- **Concepts & Soft Skills:** Cloud computing, problem-solving, Communication, Teamwork.

Projects

Amazon Sales Data Analysis 11/2024 – 12/2024

github

- Performed end-to-end ETL and Exploratory Data Analysis (EDA) on Amazon sales datasets using Python and Excel.
- Identified monthly and yearly sales trends and derived KPIs to establish relationships between business attributes.
- **Tools:** Python (Pandas, Matplotlib, Seaborn), Power BI, MS Excel.

SHAP-Coded Insights: Interpretable Structural Similarity of Programs 2025

github

Presented at ICCNT 2025 Conference

- Designed a machine learning-based code similarity system using GraphCodeBERT and AST representations.
- Integrated SHAP for model interpretability, enabling transparent similarity scoring.
- **Tools:** Python, GraphCodeBERT, AST parsing, SHAP, difflib unified diff.

Explainable RAG System for Educational Q&A 2025

github

- Built an AI-powered Retrieval-Augmented Generation (RAG) system using LangChain and FAISS.
- Implemented hybrid search (vector + BM25) to improve retrieval relevance.
- Developed explainability layer using SHAP for transparent AI responses.
- **Tools:** Python, LangChain, FAISS, BM25, SHAP, Streamlit.

Certifications

- Machine Learning with Python: Foundation – LinkedIn, 2024
- Certificate of Presentation: ICCNT 2025 – IEEE, 2025
- Excel Essential Training (Microsoft 365) – LinkedIn, 2024
- Introduction to AI Agents – DataCamp, 2025