

# Amal S

 amalsreekumar94@gmail.com  +91 7356902532  Kollam, Kerala  
 [linkedin.com/in/amal-s-3b32a8264](https://linkedin.com/in/amal-s-3b32a8264)  [github.com/Amalsreekumar1](https://github.com/Amalsreekumar1)  
 <https://amalsreekumar1.github.io/Amal-S-Portfolio/>

## Summary

---

BCA Data Science student with strong skills in ML, EDA, and explainable AI. Built and presented a SHAP/GraphCodeBERT-based code similarity system at ICCCNT 2025 and developed an end-to-end educational RAG application using LangChain, BM25 and FAISS. Seeking an entry-level Data Scientist role to deliver actionable insights and build transparent, innovative ML solutions.

## Experience

---

- Data Analyst Intern, Unified Mentor** 11/2024 – 12/2024  
• Collected, cleaned, and transformed datasets using Python and Excel for business analysis. Remote  
• Developed interactive dashboards in Power BI to visualize trends and KPIs, enabling data-driven management decisions.  
• Delivered insights that helped optimize business processes and improve efficiency.

## Education

---

- Bachelor of Computer Applications in Data Science,** 08/2023 – 07/2026  
Amrita Vishwa Vidyapeetham, Amritapuri, Kollam, Kerala  
Expected Graduation: July 2026  
CGPA: 8.45  
Relevant Coursework: Data Structures, Data Mining, Machine Learning
- 12th Grade, Science,** GOVT. Model Higher Secondary School, Kulasekharapuram 11/2021 – 03/2023  
Percentage: 87% Kollam, Kerala
- 10th Grade,** GOVT. Model Higher Secondary School, Kulasekharapuram 06/2020 – 03/2021  
Percentage: 98% Kollam, Kerala

## Skills

---

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

## Projects

---

### **SHAP-Coded Insights: Interpretable Structural Similarity of Programs with GraphCodeBERT,**

*Presented at ICCNT 2025 Conference* ↗

- Researched code similarity analysis using GraphCodeBERT, AST parsing, token overlap, and SHAP explainability.
- Developed a working model to assess similarity between source code submissions.
- Utilized Python, GraphCodeBERT, Abstract Syntax Tree (AST) parsing, SHAP, and difflib unified diff to conduct research and analyze program similarity.

### **Explainable RAG System for Educational Q A** ↗

- Built a end-end Explainable RAG app for educational question-answering and learning support.
- Converted PDFs into a structured CSV knowledge base and used vector-based retrieval (FAISS) along with BM25 keyword search for hybrid document retrieval.
- Added SHAP-based explainability and document traceability to help students understand why each answer was generated.
- Tools: Python, LangChain, FAISS (vector retrieval), 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 ↗