AMALA GEORGE

Kottayam, Kerala

7736493641 amalageorge1805@gmail.com Amala George

OBJECTIVE

Data Science and AI enthusiast with hands-on experience in Machine Learning, Deep Learning, and Generative AI projects, including predictive modeling and RAG-based chatbots. Skilled in Python, TensorFlow, and LangChain, with a strong ability to design data-driven solutions that transform complex information into actionable insights.

EDUCATION

MSc.Data Science, University of Kerala, Kariavattom	8.98 CGPA	2023–2025
BSc.Mathematics, Kuriakose Elias College, Mannanam	8.93 CCPA	2020–2023

WORK EXPERIENCE

Data Science Intern, Ospyn Technologies,Thiruvananthapuram Mar 2025 – Jun 2025

- Implemented text extraction using Tesseract OCR and performed comprehensive data preprocessing.
- Developed RAG-based chatbot for M.Tech admission queries using Meta LLaMA 3-8B,enhancing contextual accuracy.

PROJECTS | GITHUB REPOSITORIES

- **Signature Forgery Prediction using DenseNet & Cyclical Learning Rate**
  - Designed a binary classification model using DenseNet121 with Laplacian edge-enhanced grayscale inputs, achieving 96.4% test accuracy in detecting forged signatures.
  - Employed a two-phase training strategy with Cyclical Learning Rate (CLR), leading to faster convergence and improved generalization over baseline CNNs.
- **RAG-Driven Chatbot for M.Tech Admissions**
  - Built a retrieval-augmented generation (RAG) chatbot using RASA, LangChain, and ChromaDB, integrated with LLaMA-3-8B via Groq API for high-precision, context-aware query responses.
  - Designed custom text chunking and PDF parsing workflows, resulting in more accurate context retrieval.
- **Enhanced Diabetes Risk Prediction with Random Forest Optimization**
  - Built an optimized Random Forest model for diabetes-risk prediction using SMOTE and hyperparameter tuning, achieving a strong 95.5% accuracy.
  - Designed an end-to-end ML pipeline with data preprocessing, resampling, and model optimization.
- **Sentiment-Driven Recommendation System for Personalized Suggestions**
  - Implemented sentiment classification using ML algorithms including Logistic Regression, Naïve Bayes, and XGBoost, with XGBoost providing high-quality sentiment predictions.
  - Developed a user-based recommendation model evaluated using RMSE, enabling personalized suggestions.
- **Predictive Modeling for Heart Failure Risk Assessment**
  - Built predictive models using Logistic Regression, Random Forest, and SVM, with Random Forest offering the most stable and reliable performance across evaluation metrics.
  - Performed comprehensive data analysis and visualization to highlight clinical features influencing heart-failure risk.

TECHNICAL SKILLS

- **Programming Languages:** Python(OOP, Data Structures), SQL
- **Data Analysis & Visualization:** Pandas, NumPy, Tableau, Excel, Matplotlib, Seaborn, Plotly
- **Machine Learning & Deep Learning:** Scikit-learn, TensorFlow, Neural Networks, Supervised & Unsupervised Learning, Regression, Classification, Clustering, CNNs, Transfer Learning, NLP, Computer Vision,
- **LLMs, Frameworks & Tools:** LLaMA & Qwen series, Hugging Face Transformers, Embedding Models, RAG (Retrieval-Augmented Generation), Flask, LangChain, ChromaDB, Git/GitHub, Jira

SOFT SKILLS

- Problem solving
- Adaptability & Continuous Learning
- Initiative & Ownership
- Storytelling with Data
- Communication & Collaboration
- Time Management

WORKSHOP ATTENDED

- Workshop on 'Machine Learning through python'in ICFOSS 15/12/2023–17/12/2023
- Workshop on Data Visualization using Power BI - ASAP 16/07/2024