

# Jiya Mary Joseph

Data Scientist — AI/ML Engineer

☎ +91-8086027316 — ✉ [josephjiyamary@gmail.com](mailto:josephjiyamary@gmail.com) — [in linkedin](#) — [🐦 DataBirdSavvy](#)  
[🐦 DataBird](#) — [🔗 LeetCode](#) — [k kaggle](#)

— AI/ML enthusiast with hands-on experience in **Generative AI**, **NLP**, and **deep learning**. Skilled in building and deploying **end-to-end Generative AI solutions** and **machine learning** applications. Focused on creating scalable, data-driven systems using modern AI techniques.

## SKILLS

**Machine Learning / Deep Learning, Generative AI & Artificial Intelligence:** TensorFlow, Keras, PyTorch, Scikit-learn, Machine Learning Algorithms, Artificial Neural Networks, LSTM Networks, Recurrent Neural Networks, Fine-tuning, Inference, Retrieval-Augmented Generation, Prompt Engineering, LangChain, LangGraph, CrewAI, LlamaIndex.

**Computer Vision:** OpenCV, Image Classification, Object Detection, Transfer Learning, Convolutional Neural Networks(CNN), Image Augmentation.

**Statistics:** Probability Distributions, Hypothesis Testing, Regression Analysis, Statistical Inference, Time Series Analysis.

**Data Science & Analytics:** Data Structures and Algorithms(DSA), Exploratory Data Analysis(EDA), Hypothesis Testing, Pandas, MLflow.

**Cloud & Machine Learning Operations:** Amazon Web Services Elastic Compute Cloud(EC2), AWS S3, AWS Bedrock, AWS ECR, AWS ECS, Machine Learning Operations, Apache Kafka.

**Development Tools:** Microsoft Visual Studio, Jupyter Notebook, Docker, Google Colab, GitHub, Git, Postman.

**Visualization Tools:** Microsoft Power BI, Tableau, Chart.js.

**Web & Frontend Development:** React.js, Next.js, Tailwind CSS, JavaScript, CSS, HTML, Bootstrap.

**Frameworks:** Streamlit, Flask, FastAPI, Django, Nginx.

**Cloud & Databases:** MySQL, MongoDB Atlas, SQL Server, Apache Cassandra, Pinecone, Facebook AI Similarity Search, SQLite, PostgreSQL, Weaviate.

**Programming Languages:** Python, C#, Visual Basic, C++, JavaScript.

## PROJECTS

### AI OET Mentor

[www.oetmentor.xyz](http://www.oetmentor.xyz) [GitHub](#)

AI-powered platform for Occupational English Test (OET) preparation, tailored for healthcare professionals.

- Integrated **fine-tuned models** (OpenAI, Llama) with **RAG** using **LangChain** and **Pinecone** for domain-specific task generation.
- Developed a **context-aware retrieval pipeline** using **MongoDB Atlas** and **Pinecone** for efficient knowledge retrieval.
- Dockerized **Ollama inference pipelines** for seamless local execution of fine-tuned models.
- Backend deployed on **AWS EC2 (Ubuntu OS)** with **Nginx** for optimal HTTP handling.

**Tech Stack:** LangChain, Ollama, Groq, MongoDB Atlas, Pinecone, Unsloth, Gemini, AWS EC2, Docker, OpenAI.

**Impact:** Personalized and adaptive OET learning solutions for healthcare professionals.

### APS Failure Prediction

[GitHub](#)

Predictive maintenance model for Air Pressure Systems in heavy-duty vehicles.

- Tackled a **binary classification problem** to detect failures in **Air Pressure Systems (APS)** for heavy vehicles.
- Built a **real-time data pipeline** using **Kafka Producer-Consumer architecture**.
- Utilized **FastAPI** to serve predictions and **MongoDB Atlas** for data persistence.
- Deployed model using **Docker** on **AWS EC2 (Ubuntu)** for real-time prediction serving.
- Utilized **AWS S3** for scalable storage and **AWS ECR** for container image management.
- Utilized **GitHub Actions** for CI/CD automation.

**Tech Stack:** Python, FastAPI, Apache Kafka, MongoDB Atlas, Docker, AWS EC2, AWS S3, AWS ECR, GitHub Actions.

**Impact:** Reduced repair costs through accurate fault detection; improved vehicle safety via predictive analytics.

### AI-Driven E-Commerce Assistant

[GitHub](#)

AI-powered chatbot for intelligent product recommendations and real-time negotiations.

- Implemented a recommendation system using **Sentence-Transformers** and **FAISS** for similarity-based searches.
- Developed an AI negotiation chatbot with **GPT-4** and a **Random Forest model** (91% accuracy) for discount prediction.
- Designed a retrieval-based chatbot leveraging web scraping (**BeautifulSoup**) and **RAG** for product information queries.

**Tech Stack:** Streamlit, OpenAI, FAISS, Hugging Face, LangChain, BeautifulSoup, Scikit-learn.

**Impact:** Enhanced e-commerce UX with intelligent product discovery and dynamic negotiation.

## MINI PROJECTS

---

<b>AI-Powered HR Chatbot</b> — HR support with LLM fine-tuning & conversational memory. <b>Skills:</b> LLM fine-tuning (Unsloth, BERT), multi-model integration, Streamlit deployment.	<a href="#">HrChatbot</a>
<b>Patient Health Monitoring</b> — Real-time LSTM predictions from wearable data. <b>Skills:</b> Time-series modeling, real-time ingestion, automated health alerts.	<a href="#">Demo</a> — <a href="#">GitHub</a>
<b>Chat with PDF</b> — RAG-based PDF query assistant. <b>Skills:</b> RAG pipeline, embeddings, enterprise knowledge assistant.	<a href="#">GitHub</a>
<b>Chest Cancer Detection</b> — CNN + MLOps deployment. <b>Skills:</b> VGG16, MLflow, DVC, CI/CD with Docker & AWS.	<a href="#">GitHub</a>
<b>Face Recognition</b> — Celebrity classification (98% accuracy). <b>Skills:</b> Wavelet features, SVM, Flask deployment.	<a href="#">GitHub</a>
<b>Laptop Price Prediction</b> — Random Forest regression API. <b>Skills:</b> Regression (88%), Flask API, AWS hosting.	<a href="#">GitHub</a>
<b>Web Designing</b> — Static site cloning & UI customization. <b>Skills:</b> Responsive UI/UX, Bootstrap, modern web design.	<a href="#">GitHub</a>
<b>Sales Dashboard</b> — BI dashboards for sales insights. <b>Skills:</b> Data visualization, forecasting, segmentation.	<a href="#">Tableau</a> — <a href="#">Power BI</a>

## EXPERIENCE

---

<b>Baby Memorial Hospital, Calicut</b> <i>HR Associate</i>	<b>18/03/2020 – 23/12/2020</b>
<ul style="list-style-type: none"> <li>– Facilitated seamless <b>onboarding processes</b>, ensuring new hires quickly integrated into the <b>company culture</b>.</li> <li>– Actively contributed to <b>training initiatives</b> for onboarding and ongoing <b>professional development</b>.</li> <li>– Played a key role in <b>recruitment activities</b>, especially in <b>training</b>.</li> </ul>	
<b>Advait Trade and Consultancy Pvt. Ltd., Cochin</b> <i>Programmer</i>	<b>27/03/2008 – 30/11/2011</b>
<ul style="list-style-type: none"> <li>– Implemented <b>web</b> and <b>Windows applications</b> using <b>.NET</b> and <b>VB</b>, delivering robust and scalable solutions.</li> <li>– Developed <b>user-friendly interfaces</b> using <b>HTML</b>, <b>CSS</b>, and <b>JavaScript</b>.</li> <li>– Worked extensively with <b>SQL Server</b> for backend data management and query optimization.</li> <li>– Managed <b>file deployments</b> for the <b>Waterflow application</b> using <b>WinSCP</b> to transfer and deploy files across servers.</li> <li>– Conducted thorough <b>testing</b> and managed <b>deployment</b> of applications across platforms.</li> <li>– <b>Key Projects:</b> <ul style="list-style-type: none"> <li>– <b>Infobeats</b> — Developed Windows app with SMS-based enquiry and reservation features, providing instant consumer responses. <i>Tech:</i> ASP.Net, VB, SQL Server, XML, JavaScript. <i>Role:</i> Coding, SQL queries, data validation, functional spec updates, deployment support.</li> <li>– <b>AyurRekha</b> — Developed web app for modular hospital management including health info, resources, and patient services. <i>Tech:</i> ASP.Net, VB, SQL Server, JavaScript. <i>Role:</i> Coding, data validation, form testing/upload, SQL queries.</li> </ul> </li> </ul>	

## EDUCATION

---

<b>Prototype</b> <i>AI/ML/Data Science — Bootcamp (IT simulated Infrastructure)</i>	<i>May 2023 – Present</i> <b>Certificate</b>
<b>M. G University, Kottayam</b> <i>Bachelor of Technology (B.Tech) in Electronics and Communication Engineering (ECE)</i>	2007
<b>Balikamotom H.S.S, Pathanamthitta</b> <i>Higher Secondary Examination (XII)</i>	2003