

# Salma S

## AI/ML & Data Science Engineer | Python | LLMs | ML Pipelines | NLP & CV

[Email: syedmannu2003@gmail.com](mailto:syedmannu2003@gmail.com) | [Phone: +91-7397762195](#) | [LinkedIn](#) | [GitHub](#) | [Portfolio Website](#)

### SUMMARY

Innovative AI/ML Engineer with 1+ years of hands-on experience designing, developing, and deploying end-to-end AI/ML solutions across NLP, computer vision, and predictive modeling. Skilled in LLM pipelines, RAG systems, feature engineering, ML lifecycle management, and MLOps, with a track record of building production-ready applications for enterprise and research projects. Adept at translating complex datasets into actionable insights and collaborating with cross-functional teams to deliver scalable, reliable solutions.

### EDUCATION

#### Bachelor of Technology in Artificial Intelligence and Data Science

2022 - 2026

**Dhirajlal Gandhi College of Technology, Tamil Nadu** - CGPA: 8.29

#### Credit-Linked Certification in Artificial Intelligence & Machine Learning

MAR 2025 - DEC 2025

**Indian Institute of Technology (IIT) Guwahati** - CGPA: 6.7

### WORK EXPERIENCE

#### AI/ML Internship Trainee – Apex (JUL 2025 - Aug 2025)

- Developed LLM-powered candidate matching platform with multi-faceted scoring and RAG pipelines.
- Implemented explainable AI dashboards and automated resume entity extraction using FAISS, Google Gemini LLM, PyTorch, LangChain.
- Built interactive Streamlit UI with insights, visualizations, and real-time candidate analysis.

#### AI-Green Skills Intern – Edunet Foundation with Shell (JUL 2025 - AUG 2025)

- Built EV Adoption Forecast tool predicting 3-year adoption trends using regression, feature engineering, and cumulative growth metrics.
- Designed multi-county comparison dashboards using Streamlit, Pandas, NumPy, Matplotlib.
- Developed a predictive ML pipeline for actionable insights to support urban infrastructure planning.

#### Artificial Intelligence Intern - EI Systems (JUNE 2025 - AUG 2025)

- Developed Keystractor OCR & NLP tool for extracting structured data from resumes, PDFs, and images.
- Built Flask-based backend and UI, integrated spaCy NLP for entity recognition and skill extraction.

#### Data Science Intern - Zephyr Technologies and Solutions Pvt LT (SEP 2024 - OCT 2024)

- Developed Keystractor OCR & NLP tool for extracting structured data from resumes, PDFs, and images.
- Built Flask-based backend and UI, integrated spaCy NLP for entity recognition and skill extraction.

#### Data Science Intern - Skill Vertex (JAN 2024 - MAR 2024)

- Developed SentinelX Fraud Guard, a real-time fraud detection system using Gradient Boosting and LIME explainable AI.
- Built Streamlit dashboard for live transaction monitoring with predictive insights.

## WORK EXPERIENCE

### **RentAI-Pro – Your AI-Powered Document Explainer**

- Engineered an end-to-end RAG pipeline using LangChain, Ollama (TinyLLaMA), and FAISS for context-aware document retrieval.
- Developed a Flask-based backend with Gunicorn for scalable API serving and Dockerized deployment on Hugging Face Spaces.
- Integrated advanced query understanding for PDF and text documents, reducing response latency by 30%.

### **Manas AI Wellness**

- Built a FastAPI backend integrated with Google Gemini API and Hugging Face Transformers for mental-wellness text analytics.
- Designed interactive sentiment and stress-level detection pipelines leveraging PyTorch and NLP transformers.
- Deployed the system with Docker + Hugging Face Spaces, ensuring 99% uptime and real-time emotion tracking.

### **Natural Language Financial Q&A**

- Developed an intelligent financial chatbot using FinBERT, Google Generative AI, and Streamlit for real-time insights.
- Implemented RAG-based summarization combining data from Yahoo Finance, Alpha Vantage, and Reuters News Feed.
- Designed dynamic visualizations with Plotly and deployed the solution on Streamlit Cloud for investor-ready analytics.

### **AI Semantic Candidate Matcher**

- Created an AI-driven hiring assistant leveraging Sentence Transformers, TensorFlow/Keras, and Scikit-learn for candidate-job matching.
- Integrated Google Generative AI for contextual resume interpretation and skill-gap analysis.
- Built and deployed an interactive dashboard with Streamlit, achieving 91% semantic match accuracy.

### **SentinelX Fraud Guard**

- Designed a fraud detection system using Scikit-learn, Imbalanced-Learn, and LIME for explainable AI.
- Developed class-imbalance correction techniques improving fraud detection precision by 18%.
- Built an interactive Streamlit dashboard with Plotly visualizations for risk scoring and interpretability.

### **SentinelX Fraud Guard**

- Designed a fraud detection system using Scikit-learn, Imbalanced-Learn, and LIME for explainable AI.
- Developed class-imbalance correction techniques improving fraud detection precision by 18%.
- Built an interactive Streamlit dashboard with Plotly visualizations for risk scoring and interpretability.

### **AstroWave Intelligence**

- Developed a CNN-based signal classification system using TensorFlow/Keras and gwpy (Q-Transform) for gravitational-wave analysis.
- Integrated Flask backend and Hugging Face Spaces for real-time signal visualization and prediction.
- Deployed a Dockerized inference pipeline, reducing model latency by 40% on 256x256 signal maps.

---

## Technical Skills

- Programming & Scripting: Python, R, SQL, NoSQL, JavaScript, PySpark
- Machine Learning & AI: Deep Learning, Transfer Learning, Reinforcement Learning, Optimization, Predictive Analytics
- Frameworks & Libraries: TensorFlow, PyTorch, Scikit-learn, Keras, XGBoost, LangChain, LangGraph, FAISS, Hugging Face Transformers
- Data Engineering & Processing: Pandas, NumPy, Dask, Feature Engineering, ETL Pipelines, Data Wrangling
- Visualization Tools: Matplotlib, Seaborn, Plotly, Power BI, Tableau, Streamlit
- Backend & APIs: Flask, FastAPI, Streamlit, RESTful APIs, GraphQL, Gunicorn, Nginx
- NLP & LLMs: BERT, GPT, T5, Llama, FinBERT, Gemma, Prompt Engineering, RAG, Knowledge Graphs
- Cloud & MLOps: AWS (S3, EC2, SageMaker), Google Cloud (Vertex AI), Azure ML, Docker, Kubernetes, MLflow, Airflow, Kubeflow
- Databases: MySQL, PostgreSQL, MongoDB, Cassandra, Firebase, ArangoDB
- DevOps & Lifecycle: CI/CD (GitHub Actions), Model Versioning, Experiment Tracking, Explainable AI (SHAP, LIME)
- Tools & Platforms: Git, GitHub, Jupyter Notebook, Google Colab, VS Code, Hugging Face Spaces