Niraj Kumar

AI/ML Engineering Intern | Generative AI & RAG | Machine Learning | NLP | AWS
nirajofficial123@gmail.com | +91-7903292814 | GitHub | LinkedIn

Summary

AI/ML Engineering Intern with hands-on experience in Machine Learning, Deep Learning, NLP, and Generative AI. Built and deployed production-grade systems including RAG-based applications and ML pipelines. Strong foundation in EDA, model evaluation, API development, and cloud deployment using Docker and AWS.

Technical Skills

- **Programming:** Python, SQL
- **Machine Learning:** Regression, Classification, EDA, Feature Engineering, Model Evaluation
- **Deep Learning & NLP:** Neural Networks, Embeddings, Semantic Search
- **Generative AI:** LLMs, Prompt Engineering, RAG, Agentic AI
- **Frameworks & Tools:** FastAPI, LangChain, FAISS, Streamlit, Scikit-learn
- **Cloud & DevOps:** Docker, AWS (ECR, ECS Fargate), GitHub Actions

Experience

AI/ML Intern — Edunet Foundation (AICTE) *Jun 2025 – Jul 2025*

- Applied Machine Learning techniques to a real-world employee salary prediction problem.
- Performed data preprocessing, EDA, feature engineering, and model validation on structured datasets.

Projects

Document Portal — Generative AI RAG Platform GitHub *Dec 2025 – Jan 2026*
FastAPI, LangChain, FAISS, Docker, AWS ECS

- Built a production-grade document intelligence platform using Retrieval-Augmented Generation (RAG).
- Designed an end-to-end pipeline for document ingestion, chunking, embedding generation, and FAISS-based semantic retrieval.
- Integrated LLMs to enable contextual question answering across single and multiple documents.
- Developed scalable REST APIs using FastAPI and containerized the application with Docker.
- Deployed the system on AWS ECS Fargate with automated CI/CD pipelines using GitHub Actions.

Employee Salary Prediction System GitHub *Jun 2025 – Jul 2025*
Python, Machine Learning, EDA, Streamlit

- Designed an end-to-end Machine Learning pipeline to predict employee salaries from structured data.
- Performed extensive Exploratory Data Analysis (EDA) and feature preprocessing using Scikit-learn pipelines.
- Selected CatBoost and XGBoost as best-performing models based on evaluation results.
- Achieved an R^2 score of approximately 0.87, indicating strong predictive performance.
- Deployed the final model using Streamlit for real-time salary prediction.

Education

B.Tech — Sikkim Manipal Institute of Technology, Sikkim CGPA: 7.5
Class XII — Gyan Niketan, Patna 79.8%

Certifications

- Natural Language Processing Specialization — DeepLearning.AI
- Neural Networks and Deep Learning — DeepLearning.AI
- Artificial Intelligence Fundamentals — IBM
- Artificial Intelligence & Machine Learning Internship — Edunet Foundation (AICTE)