AWANTIKA SRIVASTAVA

AI Engineer | Data Scientist | GenAI & Chatbot

+91-8920482037 |sawantika81@gmail.com| LINKEDIN |Github

PROFILE SUMMARY

AI Engineer with **2+ years of experience** building **chatbots, NLP systems, and GenAI solutions** using machine learning and LLMs. Strong in **Python, NLP, REST APIs, RAG, and cloud deployment**, with hands-on experience delivering **scalable, production-ready conversational AI systems**.

CORE TECHNICAL SKILLS

- **Programming & Data Science:** Python, C++, SQL, Numpy, Pandas, SciPy, scikit-learn
- **Statistics & Mathematics:** EDA, Statistical Modeling, Hypothesis Testing, Probability, Linear Optimization, Trend Analysis.
- **Machine Learning:** Supervised & Unsupervised Learning, Regression, Classification, Clustering, Random Forest, Decision Trees, SVM, KNN, K-Means, XGBoost, Hyperparameter Tuning.
- **Deep Learning:** Neural Networks, CNN, RNN, LSTM, BERT, Transfer Learning, Prompt Engineering, Transformer Models.
- **Frameworks & Tool:** TensorFlow, Keras, Pytorch, Flask, Git, GitHub.
- **GenAI & Chatbots:** LLMs, Prompt Engineering, RAG (chunking, retrieval, re-ranking), LLM-based chatbots, Structured Outputs.
- **NLP**: Text preprocessing, Tokenization, Sentiment Analysis, Transformer-based models, NLU, OpenCV.
- **Cloud & Deployment:** AWS (EC2, S3), Model Deployment, Monitoring, Scalability, Performance Optimization.
- **APIs & Backend:** REST APIs, FastAPI, Backend Integration, Third-party API Integration.
- **Soft Skills:** Problem Solving, Debugging, Cross-team Collaboration, Technical Documentation.

EXPERIENCE

Data Scientist / AI/ML Engineer | PPS International Pvt. Ltd.   **January 2024-Present**

- Designed and developed **end-to-end AI/ML pipelines** covering data ingestion, preprocessing, model training, evaluation, and deployment.
- Built and optimized **NLP and ML models** using Python, scikit-learn, TensorFlow, and PyTorch for chatbot and conversational AI use cases.
- Developed **LLM-powered chatbots** using prompt engineering, structured outputs, and **Retrieval-Augmented Generation (RAG)**.
- Implemented **LLM evaluation and scoring frameworks**, measuring performance using Recall@K, MRR, Faithfulness, and F1-score.
- Integrated chatbots with **backend systems and third-party REST APIs** to ensure seamless user experience.
- Built **RESTful AI services using FastAPI** for real-time inference and downstream integration.
- Deployed and monitored AI solutions on **AWS**, ensuring scalability, reliability, and performance.
- Collaborated with **product, engineering, and UX teams** and maintained clear technical documentation and workflows.

PROJECTS

Railway Driver Assistance System (RDAS) | Enterprise ML Project

- Built and deployed a **production-grade machine learning system** to detect unsafe driver behavior, translating real-world safety requirements into AI-driven solutions.
- Designed and owned **end-to-end ML pipelines** covering data ingestion, preprocessing, feature engineering, model training, evaluation, and inference for real-world deployment.
- Trained **CNN-based computer vision models (SSD MobileNet)** using **TensorFlow**, and optimized inference using **TensorFlow Lite** for low-latency environments.
- Achieved **real-time inference performance (20–25 FPS, <150 ms latency)**, enabling near real-time decision-making in operational settings.
- Integrated model predictions into **downstream systems via REST APIs**, delivering **actionable insights** to support operational and safety-related decision

Chatbot using LLM & RAG | Applied ML Project

- Built a **production-grade Generative AI (GenAI) application** using **Large Language Models (LLMs)** and **Retrieval-Augmented Generation (RAG)** for context-aware question answering.
- Designed and implemented **document ingestion, embedding generation, vector retrieval, and LLM inference orchestration pipelines**.
- Applied **prompt engineering, response evaluation, and reliability tuning** to improve output accuracy, consistency, and relevance.
- Deployed the solution as an **observable LLM inference service**, aligning with modern **GenAI and LLM production workflows**.

YouTube Comments Sentiment Analyzer | link - https://youtube-ai-analyzer-ndzqo6r2mepjrtsdjmwaxl.streamlit.app/

- Developed an **end-to-end NLP modeling system** using **BERT / DistilBERT** for large-scale sentiment classification.
- Performed **EDA, text preprocessing, feature engineering, model training, and statistical evaluation** on real-world user-generated data.
- Conducted **baseline comparisons and experiment tracking** to optimize model performance.
- Deployed the model as a **REST API-based inference service** to deliver actionable sentiment insights to business stakeholders.

CERTIFICATION

- IBM Data Science & AI Certification
- AWS Generative AI with Large Language Models
- OpenCV Computer Vision Certification

EDUCATION

IMS Engineering College, Ghaziabad   **September - 2020**
Bachelor of Technology (Electrical and electronics engineering)