AWANTIKA SRIVASTAVA

AI / Machine Learning Engineer | LLM | Chatbot | Automation

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

PROFILE SUMMARY

AI / Machine Learning Engineer with **2+ years of experience** in building and deploying **ML, deep learning, and LLM-powered applications**. Strong expertise in **chatbot development, LLM integration, RAG pipelines, backend APIs, automation, and testing workflows**. Hands-on experience with **Selenium / Playwright-based automation**, large dataset processing, predictive modeling, and production ML systems. Comfortable working with **cross-functional teams** to deliver scalable, secure, and high-performance AI solutions.

CORE TECHNICAL SKILLS

- **Programming & Data Science:** Python (Pandas, NumPy, Scikit-learn), C++, REST APIs, FastAPIs, Flask, Git, GitHub
- **Statistics & Mathematics:** Statistical Modeling, Hypothesis Testing, Confidence Intervals, A/B Testing, Bias & Variance Analysis, Scenario Analysis.
- **Machine Learning:** Supervised & Unsupervised Learning, Regression, Classification, Clustering, Random Forest, Decision Trees, SVM, KNN, K-Means, XGBoost, Model Evaluation Metrics, Hyperparameter tuning, Model Evaluation
- **Deep Learning :** Neural Networks, CNN, RNN, LSTM, Transformers (BERT), Transfer Learning, Model Fine-tuning , PyTorch, TensorFlow, Keras
- **Computer Vision:** Image Classification & Preprocessing, Object detection (YOLO, SSD, MobileNet, ResNet), Video Analytics, OpenCV
- **NLP:** Text preprocessing, Tokenization, Chunkings, Sentiment Analysis, Transformer-based models, NLTK, Spacy
- **LLMs & Chatbots:** LLMs, OpenAI APIs, LLaMA, Prompt Engineering, RAG, Vector Databases, Chatbot Architecture, LLM Integration, LangChain
- **Automation & Testing:** Selenium, Playwright, Automated Test Suites, UI & API Testing, Test Automation for Frontend & Backend Applications
- **Model Optimization & Systems:** Model Quantization, Low-Latency Inference, Edge Deployment, Performance Optimization, Scalability & Reliability
- **Cloud & MLOps:** AWS EC2, AWS SageMaker, Azure, Model Deployment, ML Pipelines, Experiment Tracking

EXPERIENCE

Data Scientist / Machine Learning Engineer | PPS International Pvt. Ltd.   **January 2024-Present**

- Designed and deployed **end-to-end ML and deep learning solutions** for production use.
- Built **backend APIs** to integrate ML and LLM models into applications.
- Developed **automation scripts and test workflows** to validate ML-driven features and APIs.
- Applied **Selenium-based automation** to test web interfaces and dashboards.
- Worked with **large datasets**, handling preprocessing, feature engineering, and validation.
- Optimized ML inference pipelines for **performance, latency, and reliability**.
- Collaborated with **product, UX, and engineering teams** to deliver integrated solutions.
- Maintained technical documentation and followed **software development best practices**.

PROJECTS

Railway Driver Assistance System (RDAS) | Enterprise ML Project

- **Designed and deployed a** real-time **computer vision–based ML system** for unsafe driver behavior detection using **CNN-based SSD MobileNet** model.
- Trained and optimized models on large-scale video datasets, achieving **20–25 FPS** real-time processing with **<150 ms inference latency**.
- Implemented **end-to-end ML pipelines** for data ingestion, preprocessing, model training, evaluation, and production inference.
- Deployed optimized models using **TensorFlow Lite** on edge/production environments for continuous monitoring.
- Built a **Flask-based web dashboard** to visualize detections and automatically record **30-second event clips**, reducing manual review effort.

Chatbot Using LLM & RAG | Applied ML Project

- Built a Lightweight **LLM-Powered chatbot** using **TinyLLaMA** to answer user queries over content.
- Implemented a **Retrieval-Augmented Generation (RAG) pipeline** to retrieve relevant resume sections for contextual question answering.
- Selected **TinyLLaMA** to ensure **low memory footprint and fast inference**, making the solution suitable for resource-constrained environments.
- Applied **prompt engineering techniques** to improve response relevance and consistency.
- Deployed the chatbot as an interactive **Streamlit web application** for real-time user interaction.

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

- Deployed transformer-based **NLP models** as **production-ready** services with **REST APIs**.
- **Fine-tuned** and served a **DistilBERT**-based **sentiment** classification model for large-scale text inference.
- **Built** and deployed an interactive **streamlit web application** to perform real-time **sentiment analysis** on YouTube comments.
- **Processed** high-volume text **data** with sub-second inference **latency** for real-time sentiment analysis.

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)