# AWANTIKA SRIVASTAVA

**AI/ML Engineer | LLM | MLOps**

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

## PROFILE SUMMARY

AI / Machine Learning Engineer with **2+ years of hands-on professional experience** in designing, building, and deploying **data-intensive, scalable AI systems**. Strong expertise in **Python**, end-to-end **ML/LLM pipelines**, and production deployment of **Large Language Models (LLMs)**. Experienced in building **backend services, data platforms, and inference systems**, collaborating cross-functionally to deliver reliable AI products. Adept at translating complex requirements into **production-ready, cost-efficient, and scalable solutions**.

## CORE TECHNICAL SKILLS

- **Programming Languages:** Python, C, C++, SQL
- **Machine Learning:** Supervised & Unsupervised Learning, Regression, Classification, Clustering, Random Forest, Decision Trees, SVM, KNN, K-Means, XGBoost.
- **Deep Learning & AI:** Neural Networks, CNN, RNN, LSTM, Transformers (BERT), Transfer Learning, Model Fine-tuning.
- **Computer Vision**: Image Classification & Preprocessing, Object detection (YOLO, SSD, MobileNet, ResNet), Video Analytics, OpenCV
- **NLP**: Text preprocessing, Tokenization, Transformers, Sentiment Analysis, Transformer-based models, LLMs, RAG, Prompt Engineering.
- **Data & ML Platforms:** End-to-End ML Pipelines, Feature Engineering, Model Training & Inference, Data Processing Workflows.
- **MLOps & Deployment:** Docker, AWS (EC2, S3), MLflow, CI/CD for ML, Model Versioning, Monitoring.
- **Backend & Systems:** FastAPI, Flask, REST APIs, HTTP, Client–Server Architecture, Scalable System Design.
- **Tools & Collaboration:** Git, GitHub, Experiment Tracking, Documentation, Cross-functional Collaboration.

## EXPERIENCE

**AI/ML Engineer | PPS International Pvt. Ltd.**                                        January 2024-Present
- Designed and developed **AI-powered backend services** using **Python and FastAPI**, serving ML and LLM models via RESTful APIs.
- Built and maintained **end-to-end ML and LLM pipelines**, covering data ingestion, preprocessing, training, inference, and monitoring.
- Deployed and managed containerized AI services using **Docker** on **AWS**, focusing on scalability, reliability, and cost efficiency.
- Implemented **LLM-based production workflows**, including prompt engineering, retrieval pipelines, and inference orchestration.
- Designed data processing pipelines to support **large-scale data ingestion, feature engineering, and model evaluation workloads**.
- Collaborated cross-functionally with product and engineering teams to translate requirements into **sustainable AI solutions**.
- Applied MLOps best practices such as **model versioning, experiment tracking, CI/CD, and performance monitoring**.
- Contributed to architectural decisions and best practices to improve **system scalability, fault tolerance, and observability**.

## PROJECTS

**Railway Driver Assistance System (RDAS) | Enterprise ML Project**
- Designed and deployed a **real-time ML systems** for unsafe behavior detection using **CNN-based object detection** models.
- Implemented **end-to-end ML pipelines** from data ingestion to production inference.
- **Deployed** optimized models using **TensorFlow Lite** for continuous, **low-latency inference.**
- Developed a **Flask-based** web interface to display detections and recorded video clips in **real time**.
- **Deployed** the complete solution on **edge devices**, enabling reliable on-device inference and **real-time event recordings**.

**Amazon Stock Price prediction | Applied ML Project**
- **Built batch-oriented ML workflows** for time-series forecasting using **LSTM** models.
- Implemented data **preprocessing**, feature engineering, and model **evaluation pipelines** on **large** historical datasets.
- Designed **sliding-window** based **sequence generation** and trained LSTM model 5-year stock price data.
- Evaluated model performance using appropriate regression **metrics** and **trend-based analysis** for short-term forecasting.

**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)**