

AWANTIKA SRIVASTAVA

AI/ML Engineer – Time Series Forecasting & Production ML

+91-8920482037 | sawantika81@gmail.com | [LINKEDIN](#) | [Github](#)

PROFILE SUMMARY

AI/ML Engineer with 2+ years of hands-on experience in building production-grade machine learning and GenAI systems. Strong background in Python, deep learning, time-series forecasting, and LLM-based applications. Experienced in developing end-to-end AI pipelines for unstructured data, document understanding, forecasting, and conversational AI, with a focus on scalable, production-ready solutions.

CORE TECHNICAL SKILLS

- **Programming:** Python, C++, SQL, Numpy, Pandas, Scikit-learn, TensorFlow, Keras, PyTorch.
- **Statistics & Mathematics:** Statistical Modeling, Descriptive & Inferential Statistics, Hypothesis Testing, Probability, Sampling, Scenario Analysis.
- **Machine Learning:** Supervised & Unsupervised Learning, Reinforcement Learning, Regression, Classification, Clustering, Random Forest, Decision Trees, SVM, KNN, K-Means, XGBoost, Hyperparameter Tuning.
- **Deep Learning:** Neural Networks, CNN, RNN, Transformers (BERT), Transfer Learning, Model Fine-tuning.
- **GenAI & LLMs:** LangChain, RAG pipelines, embeddings, semantic search, vector databases, prompt engineering, LLM fine-tuning
- **Time Series & Forecasting:** LSTM, ARIMA, SARIMA, Prophet, Demand Forecasting, Trend & Seasonality Analysis
- **NLP & Conversational AI:** Text preprocessing, intent classification, information retrieval, chatbot development
- **Document AI & Data Processing:** OCR pipelines, unstructured document parsing, data extraction, transformation
- **MLOps & Deployment:** Docker, ML pipelines, CI/CD basics, model monitoring, retraining workflows
- **Databases & Tools:** SQL, Git/GitHub, REST APIs, cloud fundamentals (AWS/Azure basics)

EXPERIENCE

AI/ML Engineer | PPS International Pvt. Ltd.

January 2024-Present

- Designed, trained, and **deployed machine learning models into production** using Python and cloud platforms, improving **model inference reliability by 30%**.
- Built **end-to-end ML pipelines** covering data ingestion, preprocessing, feature engineering, model training, validation, and automated deployment.
- Implemented **containerized ML services using Docker**, enabling scalable and reproducible deployments across cloud environments.
- Performed **exploratory data analysis (EDA)** and applied statistical techniques on **large-scale datasets** to uncover trends and optimize feature selection.
- Collaborated closely with **data analysts, product managers, and stakeholders** to convert business problems into effective ML solutions.
- Optimized model performance through **hyperparameter tuning and evaluation metrics**, resulting in measurable accuracy improvements.
- Integrated trained models with **Python-based APIs** to support real-time and batch inference use cases.
- Followed best practices for **production ML**, including code versioning, experiment tracking, and deployment monitoring.

PROJECTS

Railway Driver Assistance System (RDAS) | Enterprise ML Project

- **Designed and deployed** a real-time computer vision-based ML system **to detect unsafe driver behaviors from continuous video streams**.
- **Trained** and optimized **CNN-based object detection models (SSD MobileNet architecture)** to perform real-time inference on video data.
- Implemented **end-to-end ML pipelines** covering data ingestion, preprocessing, model training, evaluation, and production inference.
- Achieved **20–25 FPS real-time processing** with **<150 ms inference latency** by optimizing models for deployment.
- Built and integrated a **Flask-based web interface** to visualize detections and automatically recorded **30-second event clips**, reducing manual review effort.

Amazon Stock Price prediction | Applied ML Project

- Developed an **LSTM-based time series forecasting model** using historical data of **50K+ records** to predict future trends.
- Applied **data normalization, sequence modeling, and feature selection**, improving forecast accuracy by **18–22%**.
- Optimized model performance through **hyperparameter tuning**, reducing validation loss by **20%**.
- Evaluated models using **RMSE and MAE**, ensuring reliable performance on unseen data.
- Built reusable **Python ML pipelines** for training, evaluation, and inference.

LLM and RAG Based Chatbot | Applied ML Project

- Designed and deployed a **Retrieval-Augmented Generation (RAG) based chatbot** to answer resume-specific queries with **90%+ response relevance**.
- Implemented **document chunking, embeddings, and vector similarity search**, reducing incorrect responses by **35%**.
- Integrated **LLM APIs with Python backend services**, enabling real-time, context-aware question answering.
- Optimized prompt engineering and retrieval logic, improving answer precision by **25%**.
- Built a scalable architecture suitable for **production-ready GenAI applications**.

CERTIFICATION

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

EDUCATION

IMS Engineering College, Ghaziabad

Bachelor of Technology (Electrical and electronics engineering)

September - 2020