AWANTIKA SRIVASTAVA

Data Science Engineer | Machine Learning & AI

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

PROFILE SUMMARY

Data Scientist with **2+ years of professional experience** applying **statistical analysis, machine learning, and data science techniques** to solve real-world business and engineering problems. Strong foundation in **mathematics, statistics, and data modeling**, with hands-on experience working on **structured and unstructured datasets**. Skilled in collaborating with software and engineering teams to translate analytical solutions into **reliable, scalable applications**, and communicating insights clearly to both technical and non-technical stakeholders.

CORE TECHNICAL SKILLS

- **Programming Languages & Tools:** Python, C++, Numpy, Pandas, Scikit-learn, SQL.
- **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 & AI:** Neural Networks, CNN, RNN, LSTM, Transformers (BERT), Transfer Learning, Model Fine-tuning, RAG, LLMs, Prompt Engineering.
- **Frameworks:** TensorFlow, Keras, Pytorch, Flask.
- **Computer Vision**: Image Classification & Preprocessing, Object detection (YOLO, SSD, MobileNet, ResNet), Video Analytics, OpenCV.
- **NLP**: Text preprocessing, Tokenization, Transformers, Sentiment Analysis, Topic Modeling (LSA, LDA), Transformer-based models.
- **Data Analysis & Visualization:** EDA, Feature Engineering, Matplotlib, Seaborn.
- **APIs & Integration:** REST APIs, Model Integration, Inference Services.
- **MLOps & Cloud:** Docker, MLflow, CI/CD, AWS (EC2, S3), Cloud ML Workflows.

EXPERIENCE

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

- Designed, developed, and deployed **machine learning and AI models** to solve real-world business problems.
- Performed **data collection, cleaning, feature engineering, and exploratory data analysis (EDA)** on structured and unstructured datasets.
- Built and optimized **classification, predictive, and NLP models** using Python and ML frameworks.
- Developed **data pipelines** for model training, evaluation, and deployment.
- Integrated trained models with applications using **REST APIs** for real-time inference.
- Collaborated with software engineers to **productionize models** following basic **MLOps practices**.
- Monitored model performance, accuracy, and drift to ensure reliability and scalability.
- Documented methodologies, results, and insights for technical and business stakeholders.

PROJECTS

Railway Driver Assistance System (RDAS) | Enterprise ML Project

- Designed and deployed a **real-time unsafe behavior detection system** using **CNN-based SSD MobileNet (TensorFlow)** models on large-scale video data.
- Built **end-to-end ML pipelines** covering data ingestion, preprocessing, model training, evaluation, and production inference.
- Optimized models using **TensorFlow Lite**, achieving **20–25 FPS real-time performance** with **<150 ms inference latency** on edge/production environments.
- Developed a **Flask/FastAPI-based web dashboard** to visualize detections and automatically record **30-second event clips**, reducing manual review effort by **~60%**.
- Performed **error analysis, bias checks, and continuous performance monitoring** to improve robustness and operational reliability.

Chatbot using LLM & RAG | Applied ML Project

- Developed a **production-ready chatbot** using **Large Language Models (LLMs)** and **Retrieval-Augmented Generation (RAG).**
- Built pipelines for **document ingestion, embedding, retrieval, and context-aware question answering**.
- Applied **prompt engineering** to improve response accuracy and relevance.
- Deployed the solution using **Python-based inference services** and Streamlit.

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)