

AWANTIKA SRIVASTAVA

Data Scientist | Machine Learning, Data Science & NLP

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

PROFILE SUMMARY

Data Scientist with **2+ years of experience** in applying **machine learning, statistical analysis, and NLP techniques** to solve real-world business problems. Strong background in **Python-based data science workflows**, supervised and unsupervised learning algorithms, and end-to-end ML lifecycle. Experienced in working with **databases, cloud platforms, and analytics pipelines**, with the ability to learn and adapt to new AI/ML tools quickly.

CORE TECHNICAL SKILLS

- **Programming Languages & Tools:** Python, C, C++, SQL.
- **Statistics & Mathematics:** Statistical Modeling, Hypothesis Testing, A/B Testing, Confidence intervals, Bias, Probability, Linear Optimization, Trend Analysis, Risk Modeling.
- **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.
- **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 Handling & Databases:** Data Retrieval & Analysis, SQL Queries, MySQL, SQL Server, MongoDB.
- **Cloud & Data Science Platforms:** AWS, Sagemaker (Exposure), GCP (Exposure), Databricks (Exposure).
- **Tools & SDLC:** Git, GitHub, JIRA, Confluence, Structured Software Development.

EXPERIENCE

Data Scientist | PPS International Pvt. Ltd.

January 2024-Present

- Developed and optimized **CNN-based deep learning** models for real-time **computer vision** safety systems.
- Implemented the **complete AI/ML lifecycle**, including data collection, preprocessing, model training, evaluation, and deployment.
- Deployed optimized models using **TensorFlow Lite** on **edge devices**, achieving low-latency inference under constrained CPU and memory environments.
- Designed and maintained **end-to-end inference pipelines** for continuous video stream processing.
- Built a **Flask-based web application** to visualize AI detections and automatically recorded video events for monitoring and analysis.
- Optimized inference workflows to ensure reliable performance on **low-resource hardware platforms**.
- Worked independently and collaboratively to deliver production-ready AI solutions with minimal supervision.

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-ndzqo6r2mepjrtsdjmwx1.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)