

AWANTIKA SRIVASTAVA

Technical Analyst – Data Scientist | Python | SQL | Statistics | ML

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

PROFILE SUMMARY

Data Science professional with **2+ years of experience** in **Python, SQL, Statistics, and Machine Learning**, with hands-on exposure to **data analysis, model building, and learner support** activities. Strong ability to **explain complex data science concepts** clearly, assist in **assignment evaluation**, and support learners through **doubt resolution and project guidance**. Experienced in working with real-world datasets, building ML models, and collaborating with academic and technical teams.

CORE TECHNICAL SKILLS

- **Programming Languages:** Python (Numpy, Pandas), C++, SQL, Excel.
- **Statistics & Mathematics:** EDA, Statistical Modeling, Descriptive Statistics, Inferential Statistics, Hypothesis Testing, Probability, Sampling.
- **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, Time Series Forecasting, TensorFlow, Keras, PyTorch.
- **BI & Reporting (Exposure):** Power BI / Tableau (Basic), Reporting Dashboards
- **Teaching & Academic Support:** Doubt Resolution, Assignment Evaluation, Learner Tracking, Documentation
- **NLP:** Text preprocessing, Tokenization, Sentiment Analysis, Topic Modeling (LSA, LDA), Transformer-based models, NLP Pipelines
- **Computer Vision:** Image Classification & Preprocessing, Object detection (YOLO, SSD, MobileNet, ResNet), Video Analytics, OpenCV
- **Tools & Platforms:** Git, GitHub, Streamlit, LMS Exposure, Zoom
- **Communication & Collaboration:** Strong written and verbal communication, stakeholder coordination, cross-functional collaboration, project ownership

EXPERIENCE

Data Scientist | PPS International Pvt. Ltd.

January 2024-Present

- Analyzed large, complex datasets to uncover patterns, trends and insights supporting data-driven decision-making
- Designed and executed statistical analyses including hypothesis testing and confidence interval estimation to validate business assumptions.
- Built and evaluated machine learning models (classification and regression) to support predictive and analytical use cases
- Wrote complex SQL queries to extract, transform, and analyze data from relational and NoSQL data sources
- Conducted exploratory data analysis (EDA) to identify data quality issues, bias, and anomalies prior to modeling
- Partnered with product, compliance, and engineering teams to translate analytical findings into clear business recommendations
- Designed and analyzed A/B tests and controlled experiments to evaluate changes in user workflows and system configurations, assessing impact using statistical significance and confidence intervals.

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

- **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/>

- Built a **batch ML pipeline** to analyze large-scale YouTube data using Python and SQL.
- Performed ETL, EDA, and **feature engineering on structured and unstructured data**.
- Trained NLP-based ML models for sentiment classification and trend analysis.
- Designed modular pipelines suitable for **Airflow-based orchestration and scheduling**.
- Generated actionable insights by analyzing audience engagement and content performance.

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