AWANTIKA SRIVASTAVA

Full-Stack Data Scientist | ML & MLOps Engineer

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

PROFILE SUMMARY

Full-Stack Data Scientist with **2+ years of experience** owning the **end-to-end ML lifecycle**, from translating business problems into analytical formulations to building statistically sound models and deploying them into **production-grade ML systems**. Strong hands-on expertise in **Python, machine learning, NLP, GenAI/LLM workflows, CI/CD for ML, model monitoring, and cloud-based ML platforms**. Experienced in operating production ML with a focus on **reliability, observability, and continuous improvement**.

CORE TECHNICAL SKILLS

- **Programming Languages & Tools:** Python, C++, Numpy, Pandas, Scikit-learn, SQL., PySpark.
- **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 & GenAI:** 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.
- **MLOps & Engineering:** MLflow (experiment tracking, model registry), CI/CD for ML, Model Versioning, Monitoring & Drift Detection, Docker.
- **Cloud & Big Data:** AWS (S3, EC2, SageMaker exposure), Databricks (workflow exposure).

EXPERIENCE

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

- Owned the **end-to-end ML lifecycle** from problem framing and feature engineering to deployment and production monitoring.
- Designed and implemented **predictive, classification, and NLP models** using Python and modern ML frameworks.
- Built **data pipelines** for training, validation, and deployment with clean separation of data, features, and model logic.
- Productionized ML models using **CI/CD practices**, automated testing, and versioned deployments.
- Implemented **model monitoring and drift detection**, defining retraining triggers aligned with business SLAs.
- Tracked experiments and managed model artifacts using **MLflow**, following disciplined promotion workflows.
- Collaborated with product, engineering, and business stakeholders to deliver **decision-ready insights**.
- Operated production ML systems with a focus on **observability, alerts, and continuous improvement**.

PROJECTS

Railway Driver Assistance System (RDAS) | Enterprise ML Project

- Built and deployed a **production ML system** to detect unsafe driver behavior using deep learning models.
- Designed **data ingestion, preprocessing, feature engineering, training, and inference pipelines** for real-world deployment.
- Trained **CNN-based models (SSD MobileNet)** using TensorFlow and optimized inference using **TensorFlow Lite**.
- Achieved **real-time inference (20–25 FPS, <150 ms latency)** in production environments.
- Implemented **model monitoring, performance regression checks, and alerting** post-deployment.
- Integrated ML outputs into downstream systems via **REST APIs** for operational decision-making.

Chatbot using LLM & RAG | Applied ML Project

- Built a **production-grade Generative AI (GenAI) application** using **Large Language Models (LLMs)** and **Retrieval-Augmented Generation (RAG)** for context-aware question answering.
- Designed and implemented **document ingestion, embedding generation, vector retrieval, and LLM inference orchestration pipelines**.
- Applied **prompt engineering, response evaluation, and reliability tuning** to improve output accuracy, consistency, and relevance.
- Deployed the solution as an **observable LLM inference service**, aligning with modern **GenAI and LLM production workflows**.

YouTube Comments Sentiment Analyzer | link - https://youtube-ai-analyzer-ndzqo6r2mepjrtsdjmwaxl.streamlit.app/

- Developed an **end-to-end NLP modeling system** using **BERT / DistilBERT** for large-scale sentiment classification.
- Performed **EDA, text preprocessing, feature engineering, model training, and statistical evaluation** on real-world user-generated data.
- Conducted **baseline comparisons and experiment tracking** to optimize model performance.
- Deployed the model as a **REST API-based inference service** to deliver actionable sentiment insights to business stakeholders.

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)