AWANTIKA SRIVASTAVA

Machine Learning Engineer | AI & Machine Learning Solution

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

PROFILE SUMMARY

Machine Learning Engineer with 2+ years of hands-on experience in designing, developing, and operationalizing machine learning solutions with strong exposure to **AWS-based ML pipelines**. Experienced across the full ML lifecycle including **data ingestion, feature engineering, model training, evaluation, CI/CD integration, and production deployment**. Proven ability to work on scalable ML systems, automate workflows, and collaborate with cross-functional teams while adhering to enterprise-grade security and engineering standards.

CORE TECHNICAL SKILLS

- **Programming Languages:** Python, C++, SQL., PySpark
- **Statistics & Mathematics:** Statistical Modeling, Descriptive Statistics, Hypothesis Testing, Probability, Sampling, Scenario 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, TensorFlow, Keras, PyTorch.
- **MLOps:** ML Pipelines, Model Training & Evaluation, CI/CD (Exposure), Model Versioning
- **NLP**: Text preprocessing, Tokenization, Sentiment Analysis, Topic Modeling (LSA, LDA), Transformer-based models.
- **Cloud (AWS):** S3, SageMaker (Training & Endpoints – Exposure), Lambda (Exposure)
- **Workflow Orchestration:** Apache Airflow (DAG concepts, scheduling – Exposure)
- **Data Engineering:** ETL Pipelines, Batch Processing, Data Validation, SQL Server
- **Tools:** Git, GitHub, GitLab (Exposure), Docker (Basics)

EXPERIENCE

Machine Learning Engineer | PPS International Pvt. Ltd. **January 2024-Present**

- Designed and supported **AWS-aligned ML pipelines** covering data preprocessing, model training, evaluation, and deployment workflows.
- Worked on **end-to-end ML lifecycle automation**, ensuring reproducibility, versioning, and validation of models.
- Contributed to PoCs using **SageMaker-style training and inference patterns**, focusing on scalability and reliability.
- Assisted in integrating ML workflows with **CI/CD pipelines**, including unit testing and pipeline validation.
- Collaborated with data engineering teams on **ETL pipelines and batch data processing**.
- Followed enterprise engineering practices including **Git-based workflows, branching strategies, and documentation**.
- Supported monitoring and troubleshooting of ML systems in production-like environments.

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 **September - 2020**

Bachelor of Technology (Electrical and electronics engineering)