AWANTIKA SRIVASTAVA

Data Research Analyst | Data Analyst

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

PROFILE SUMMARY

Data Research Analyst with 2+ years of experience in data analysis, research, reporting, and dashboarding. Strong expertise in Python, SQL, Excel, Power BI, and Tableau for handling large datasets, building MIS reports, and delivering actionable business insights. Experienced in primary and secondary research, data validation, and collaborating with stakeholders to support data-driven decision making.

CORE TECHNICAL SKILLS

- **Programming Languages:** Python, C++, SQL..
- **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.
- **Data Analysis & Research:** Primary and secondary research, data validation, data summarization, statistical analysis, insight generation.
- **Databases & Querying:** PostgreSQL, MySQL, MongoDB, joins, subqueries, aggregations, data extraction for MIS and management reporting
- **Reporting & Business Intelligence:** Power BI, Tableau, Excel dashboards, KPI tracking, analytical reports.
- **Excel & Documentation:** Excel (Pivot Tables, VLOOKUP/XLOOKUP, Power Query), Excel-based dashboards, Microsoft Word reporting and documentation
- **Data Engineering & Pipelines:** Data intake, validation, mining, integration, process optimization across multiple stages of data lifecycle.
- **Visualization & Insights Delivery:** Data visualization best practices, dashboard design, storytelling with data, presenting insights to stakeholders
- **Communication & Collaboration:** Strong written and verbal communication, stakeholder coordination, cross-functional collaboration, project ownership

EXPERIENCE

Data Analyst / Data Research Analyst | 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)