AWANTIKA SRIVASTAVA

NLP Engineer | Data Scientist | LLMs & AWS

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

PROFILE SUMMARY

Data Scientist with 2+ years of experience in predictive modeling, NLP, vector space modeling, machine learning, and scalable data systems. Strong expertise in Python, statistical modeling, ensemble methods, and deep learning. Hands-on experience in building LLM-powered search systems using Retrieval-Augmented Generation (RAG), embeddings, semantic retrieval, and reranking techniques. Proven experience in developing scalable ML pipelines, optimizing classifiers, and deploying production-grade AI systems on AWS and Databricks.

CORE TECHNICAL SKILLS

- **Programming & Data Science:** Python, C++, SQL., PySpark, Pandas, Numpy, Scikit-learn
- **Statistics & Mathematics:** EDA, Statistical Modeling, Descriptive & Inferential Statistics, Bayesian Statistics, Hypothesis Testing, Probability, A/B Testing
- **Machine Learning:** Supervised & Unsupervised Learning, Regression, Classification, Clustering, Random Forest, Decision Trees, SVM, KNN, K-Means, XGBoost, LightGBM, Model Evaluation Metrics (Accuracy, Precision, Recall, F1-score, ROC-AUC).
- **Deep Learning:** ANN, CNN, RNN, LSTM, BERT, Transfer Learning, Self-Attention, multi-head Attention, Time-Series Forecasting, Tensorflow, PyTorch, Keras
- **NLP:** Text preprocessing, Tokenization, Chunking, Sentiment Analysis, Transformer-based models, NLTK, TextBlob, SpaCy
- **Generative AI & LLMs:** LLMs, Prompt Engineering, RAG, FAISS, Embeddings, LLM Evaluation, LangChain, HuggingFace
- **Computer Vision:** Image Classification & Preprocessing, Object detection (YOLO, SSD, MobileNet, ResNet), Video Analytics, OpenCV
- **Search & Indexing Technologies:** Elasticsearch, OpenSearch, Solr, Embedding-based Vector Search
- **MLOps & Deployment:** MLflow, Docker, REST API , AWS (EC2, S3), CI/CD Pipelines, Scalable ML Deployment, Model Monitoring
- **Tools & Version Control:** Git, GitHub, Jupyter, Streamlit

EXPERIENCE

Data Scientist | PPS International Pvt. Ltd. **January 2024-Present**

- Developed **predictive modeling solutions** using regression, decision trees, random forest, gradient boosting, and deep learning models for datasets.
- Built **scalable ML pipelines** using PySpark and Spark SQL for large-scale distributed data processing and model training.
- Designed and optimized **classifiers using feature selection, hyperparameter tuning, and ensemble methods**, improving model performance by 15–20%.
- Applied **Bayesian statistics and advanced evaluation metrics** to enhance model robustness and decision accuracy.
- Built **LLM-powered search and automation systems** using **Retrieval-Augmented Generation (RAG)** pipelines with embedding-based semantic retrieval.
- Fine-tuned **LLM models using PEFT/LoRA techniques** for domain-specific query understanding and improved contextual response generation.
- Designed and implemented **reranker models (cross-encoder based reranker design)** to improve search relevance and ranking quality.
- Integrated **Elasticsearch / OpenSearch** for semantic indexing, hybrid search, and vector search capabilities.
- Developed **scalable data systems in collaboration with product and engineering teams** to enhance user experience and search quality.
- Managed model lifecycle using **MLflow**, and deployed **real-time inference APIs on AWS**, ensuring low-latency production systems.

PROJECTS

Railway Driver Assistance System (RDAS) | Enterprise ML Project

- **Designed and deployed a** real-time computer vision–based ML system **for unsafe** driver behavior detection using **CNN-based SSD MobileNet** model.
- Trained and optimized models on large-scale video datasets, achieving **20–25 FPS** real-time processing with **<150 ms inference latency**.
- Implemented **end-to-end ML pipelines** for data ingestion, preprocessing, model training, evaluation, and production inference.
- Deployed optimized models using **TensorFlow Lite** on edge/production environments for continuous monitoring.
- Built a **Flask-based web dashboard** to visualize detections and automatically record **30-second event clips**, reducing manual review effort.

Amazon Stock Price Prediction | Applied ML Project

- Developed an **LSTM-based time series forecasting** model using historical data of **50K+** records to predict future trends.
- Applied data **normalization**, **sequence** modeling, and **feature** selection, improving forecast accuracy by **18–22%.**
- Optimized model performance through hyperparameter tuning, reducing validation loss by 20%.
- **Evaluated** models using **RMSE** and **MAE**, ensuring reliable performance on unseen data.
- Built reusable Python **ML pipelines** for training, **evaluation**, and inference.

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

- Developed **LSTM-based time-series forecasting** model trained on 5+ years historical stock data reducing **RMSE** by 18**%.**
- Engineered lag **features**, **rolling statistics**, and volatility indicators improving prediction stability.
- **Implemented** custom **BPTT training** loop with adaptive learning rate scheduling for **stable convergence**.
- Applied **regularization techniques** reducing overfitting gap by **20**%.

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 (Electricaland electronics engineering)