AWANTIKA SRIVASTAVA

AI/ML Engineer | AI & Machine Learning Solution

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

PROFILE SUMMARY

AI/ML Engineer with 2+ years of hands-on experience in designing, developing, and deploying end-to-end AI/ML solutions. Strong expertise in statistical modeling, machine learning algorithms, deep learning, data preprocessing, and exploratory data analysis (EDA). Proven ability to build scalable, production-ready ML pipelines using Python, cloud platforms, and containerization, delivering measurable business impact across predictive analytics and intelligent systems.

CORE TECHNICAL SKILLS

- **Programming & ML Frameworks:** Python, C++, SQL., Numpy, Pandas, Scikit-learn, TensorFlow, Keras, PyTorch.
- **Statistics & Mathematics:** EDA, Statistical Modeling, Descriptive Statistics, Hypothesis Testing, Probability, Sampling, Scenario Analysis.
- **Machine Learning:** Supervised & Unsupervised Learning, Reinforcement 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, Computer Vision, TensorFlow, PyTorch, Keras
- **NLP:** Text preprocessing, Tokenization, Chunkings, Sentiment Analysis, Transformer-based models, NLTK, Spacy
- **Computer Vision:** Image Classification & Preprocessing, Object detection (YOLO, SSD, MobileNet, ResNet), Video Analytics, OpenCV
- **Generative AI & LLMs:** LLMs, OpenAI APIs, Hugging Face, LangChain, Prompt Engineering, Context Management, LLM Integration
- **Databases:** SQL, MySQL, PostgreSQL, MongoDB
- **Cloud & Deployment:** AWS (EC2, S3), Docker, REST APIs, Model Deployment & Monitoring
- **Visualization & Reporting:** Power BI, Tableau, Matplotlib, Seaborn
- **Tools & Practices:** ML Pipelines, Model Optimization, A/B Testing, Version Control (Git), Agile collaboration

EXPERIENCE

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

- Designed, trained, and **deployed machine learning models into production** using Python and cloud platforms, improving **model inference reliability by 30%**.
- Built **end-to-end ML pipelines** covering data ingestion, preprocessing, feature engineering, model training, validation, and automated deployment.
- Implemented **containerized ML services using Docker**, enabling scalable and reproducible deployments across cloud environments.
- Performed **exploratory data analysis (EDA)** and applied statistical techniques on **large-scale datasets** to uncover trends and optimize feature selection.
- Collaborated closely with **data analysts, product managers, and stakeholders** to convert business problems into effective ML solutions.
- Optimized model performance through **hyperparameter tuning and evaluation metrics**, resulting in measurable accuracy improvements.
- Integrated trained models with **Python-based APIs** to support real-time and batch inference use cases.
- Followed best practices for **production ML**, including code versioning, experiment tracking, and deployment monitoring.

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

- 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.

LLM and RAG Based Chatbot | Applied ML Project

- Designed and deployed a **Retrieval-Augmented Generation (RAG) based chatbot** to answer resume-specific queries with **90%+ response relevance.**
- Implemented **document chunking, embeddings, and vector similarity search**, reducing incorrect responses by **35%**.
- Integrated **LLM APIs with Python backend services**, enabling real-time, context-aware question answering.
- Optimized prompt engineering and retrieval logic, improving answer precision by **25%**.
- Built a scalable architecture suitable for **production-ready GenAI applications**.

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)