SRILOKA SOMULA | AI/ML Engineer

Orlando, Florida | (215) 681-0106 | srilokanagaseshu@gmail.com

SUMMARY

Experienced **AI/ML Engineer** with **3 years** of delivering robust, scalable **machine learning models** and data-driven solutions. Proficient in **Python, SQL, and Shell Scripting,** specializing in **supervised and unsupervised learning**, **deep learning, and NLP** with **TensorFlow, PyTorch,** and **Hugging Face Transformers**. Skilled in managing the full model lifecycle — from data preprocessing and feature engineering to training, evaluation, and deployment using **MLflow, Docker, and Apache Airflow**. Hands-on experience with **AWS SageMaker and Azure ML Studio** for deploying real-time AI applications.

TECHNICAL SKILLS

Programming Languages: Python, SQL, Java, Shell Scripting, Scala, R
Machine Learning & Deep Learning: Supervised Learning, Unsupervised Learning, Classification, Regression, Clustering, Ensemble Methods (Random Forest, XGBoost, LightGBM), Neural Networks (ANN, CNN, RNN, LSTM), and Transformers
NLP & LLMs: Text Classification, Named Entity Recognition (NER), Sentiment Analysis, Hugging Face Transformers, LangChain, Retrieval- Augmented Generation (RAG), BERT, GPT
Frameworks & Libraries: Scikit-learn, TensorFlow, Keras, PyTorch, XGBoost, LightGBM, OpenCV,
Data Processing & Analysis: Pandas, NumPy, PySpark, Feature Engineering, Data Cleaning, EDA,
Model Deployment & MLOps: Flask, FastAPI, Docker, MLflow, DVC, Git, GitHub Actions, Apache Airflow,
Cloud Platforms: AWS SageMaker, Azure ML Studio, Jupyter Notebooks,
Visualization: Matplotlib, Seaborn, Plotly, Power BI, Amazon QuickSight
Version Control & DevOps: Git, GitHub, Jenkins, Azure DevOps
Agile & Project Tools: JIRA, Trello, Scrum, Kanban, Confluence

WORK EXPERIENCE

AI/ML Engineer — Apr 2024 - Present

Capital One, McLean - VA

- Ingested and preprocessed large-scale structured and unstructured data using Python and Airflow, boosting data quality by 22%.
- Engineered features with TF-IDF and embeddings, increasing model accuracy by 15%.
- Developed and evaluated ML/DL models (XGBoost, CNN, LSTM) with TensorFlow and PyTorch, achieving 88% precision.
- Leveraged Hugging Face Transformers and LangChain for LLM-powered NLP, enhancing text processing speed and context understanding by 30%.
- Optimized models via KPIs, cross-validation, GridSearchCV, and SHAP explainability; validated through A/B testing.
- Managed ML lifecycle with MLflow; automated CI/CD pipelines using GitHub Actions and Docker for robust deployments.
- Built MLOps workflows with Kubeflow and Airflow, automating retraining triggered by data drift, improving update frequency by 40%.
- Deployed FastAPI inference services containerized with Docker, maintaining sub-150ms latency.
- Created dashboards with QuickSight and Python visualization libraries to convey business impact.
- Documented workflows and models in Confluence; maintained version control with Git.
- Followed Agile Scrum practices, collaborating across teams in sprint planning and reviews.

Machine Learning Engineer — Jul 2021 – Feb 2023

Value Momentum, India

- Developed supervised ML models (XGBoost, CNNs) for claims classification, achieving 92% accuracy and reducing manual review by 30%.
- Applied NLP techniques (NER, tokenization, BERT, RoBERTa) to extract critical data from unstructured text, boosting extraction speed by 25%.
- Built feature pipelines with PyTorch and TensorFlow, enhancing model generalization via attention mechanisms and dropout.
- Deployed REST APIs with FastAPI on AWS SageMaker; monitored via CloudWatch, ensuring 99.9% uptime.
- Implemented MLOps using MLflow and automated SageMaker deployment pipelines, cutting deployment time by 40%.
- Tuned hyperparameters and performed cross-validation, improving precision and recall by 15%.
- Integrated model outputs into claims workflows, reducing processing time by 30% and improving compliance.

ACADEMIC PROJECT

Machine Learning Methods for Attack Detection in the Smart Grid

- Developed a machine learning system to detect false data injection attacks in smart grids using **supervised/semi-supervised learning**, validated on IEEE test systems.
- Analysed large synthetic datasets using **NumPy, Pandas, and Scikit-learn** to extract key features and optimize model performance.
- Improved detection accuracy and reduced false positives through fusion techniques and visualized results using **Matplotlib**.

EDUCATION

Webster University — **St. Louis, MO**
Masters in Cybersecurity ***(GPA:3.6/4.0)*** — Jan 2023- Mar 2024

Usha Rama College of Engineering and Technology — **Gannavaram, IN**
Bachelors in Computer Science ***(GPA:3.0/4.0)*** — *Jun 2018 – May 2022*