**Sravan Goud Lankoti**

**Linkedin:** [**Sravan Lankoti**](https://www.linkedin.com/in/sravan-lankoti/)

**Email: sravangoud.lankoti@gmail.com | Contact: (585) 748-2499**

**Professional Summary:**

* ML Engineer / Data Scientist with 7+ years of experience in **Finance** & **Healthcare**.
* Adept in utilizing tools such as **Python**, **SQL** and **AWS/Azure** services to extract insights from data and communicating findings to stakeholders.
* Strong expertise in **LLM techniques**, including **Prompt Engineering, RAG** (Retrieval-Augmented Generation), and AI agents, with hands-on cloud deployment experience (Azure, AWS).
* Experienced in building and deploying robust **time series** anomaly detection models and implementing **MLOps pipelines** using Jenkins and GitHub Actions for automated CI/CD.
* Strong expertise in **Python**, **Pandas**, **Pyspark** and other data processing libraries for handling and analyzing large datasets.
* **AWS Certified Cloud Practitioner, Developer Associate, Solutions Architect Associate.**
* Skilled in natural language processing (NLP) using tools such as **NLTK**, **Gensim** and **spaCy** for text analysis and classification tasks.
* Proficient in machine learning (ML) algorithms for classification and regression tasks**.**
* Working experience with deep learning frameworks such as **TensorFlow**, **Keras** and **PyTorch** for **image recognition** and **NLP** tasks.
* Experienced in developing end-to-end machine learning pipelines using **MLflow** & **Apache Airflow**.
* Skilled in statistical testing and **hypothesis testing** using tools such as **ANOVA, t-tests** and **chi-square** tests for data analysis and validation.
* Proficient in deploying and managing machine learning models in production environments.
* Experience with containerization technologies such as **Docker and container** orchestration tools like Kubernetes.
* Strong knowledge of version control systems particularly **Git** for tracking changes in code.
* Proficient in **SQL** for data extraction, transformation, manipulation and loading (**ETL**) operations on large datasets.
* Experience with **Apache Spark** and **Hadoop** for distributed computing and big data processing.
* Dedicated team player with excellent communication skills and the ability to work collaboratively with cross-functional teams.

**Technical Skills:**

* **Programming Languages:** Python, R
* **Algorithms:** Linear Regression, Lasso & Ridge, Logistic regression, KNN, naïve Bayes, Decision Tree, Random Forest, K means, K Means++, DBSCAN, PCA, XGBOOST, Gradient Descent, SVM, Deep Neural Networks, Transfer Learning, CNN, RNN, R-CNN, LSTM, BLSTM
* **Natural Language packages**: nltk, Gensim, Hugging face, TextBlob, Spacy, LangChain
* **Data Science Packages:** Pandas, NumPy, SciPy, Scikit-learn, xgboost, Keras, TensorFlow, PyTorch, shap, MLFlow, lightgbm, catboost, pycaret, OpenCV
* **Visualization:** Matplotlib**,** Seaborn, PowerBI, plotly, Grafana
* **Generative AI & LLMs:** LangChain, LangGraph, OpenAI API, RAG, LlamaIndex, Hugging Face Transformers, Vector DBs (FAISS, Pinecone)
* **Data Processing:** Image Augmentation, Annotation & Labeling, Synthetic Data Generation
* **Misc**: Fast API, Django, Apache Airflow, Object Oriented Programming
* **NVIDIA Technologies:** TensorRT, CUDA, DeepStream, Jetson, GPU Optimization
* **Cloud:** AWS (EC2, IAM, VPC, DynamoDB, RDS, Athena, Lambda, SNS, SQS, S3, SageMaker)
* **MLOps & CI/CD**: Jenkins, GitHub Actions, Azure DevOps, MLflow, Docker, Kubernetes
* **Time Series**: Anomaly Detection, Forecasting (ARIMA, Prophet, LSTM), Kaplan-Meier, survival models
* **Version Control:** Git, Azure Repos
* **Operating System:** Windows, Linux, Mac OS
* **Databases:** MS SQL, Oracle

**Certifications:**

AWS Certified Cloud Practitioner, valid until August 2026

AWS Certified Developer Associate, valid until September 2026

AWS Certified Solutions Architect Associate, valid until November 2026

**Professional Experience:**

***Client: United Therapeutics March 2024 – Present***

***Role: Sr.Data Scientist***

***Project: Patient Retention & HCP Targeting***

***Responsibilities:***

* Conducted **data cleaning, pre-processing** and **feature engineering** on patient shipment data using **Python** and related data processing libraries such as **Pandas**, **NumPy** and **Scikit-learn** to prepare the data for analysis and training.
* Deployed and fine-tuned **LLMs on AWS Bedrock**, integrating **foundation models (FM) from Anthropic, Meta and Amazon Titan** for enterprise applications.
* Designed **RAG-based AI solutions** using **Bedrock with vector databases (Amazon OpenSearch, Pinecone)** for semantic search and document retrieval.
* Developed and implemented machine learning models using **Kaplan Meier** algorithm to predict the likelihood of patient drop off and enhance patient retention strategies.
* Established rigorous **MLOps protocols via Azure DevOps** that enabled continuous integration/continuous delivery (CI/CD), achieving consistent deployments across environments, contributing to zero downtime.
* Performed data cleaning, transformation, manipulation & analysis on large scale data in **Azure databricks** using **Pyspark & SQL** to automate the generation of business critical reports.
* Transitioned legacy Python scripts to a streamlined **FastAPI architecture**, saving over $1M

in costs and increasing overall throughput of model deployment processes..

* Engineeredanefficient **serverless architecture** hosting Python APIs on **Azure Functions** while integrating the **Angular UI** onto Azure App Service, resulting in enhanced app responsiveness and user satisfaction.
* Designed and implemented an automated workflow using **Azure Logic Apps** for monthly runs,

incorporating alert email notifications and saving on billing.

* Utilized **Azure Key Vault** service to securely store and manage sensitive credentials, including database passwords and API keys, ensuring a zero-breach environment for application deployment.
* Developed custom data pipelines using tools such as **Apache Airflow** to automate data processing and model training.
* Developed custom visualizations and dashboards using tools such as **Matplotlib**, **Seaborn and PowerBI** to present analysis results to stakeholders.
* Continuously monitored model performance and updated the models as new data and insights became available.

***Client: UBS Nov 2020 – March 2024***

***Role: ML Engineer***

***Project: Margin Call Automation***

***Responsibilities:***

* Performed data analysis on margin call email data using **Python**, **Pandas** and **NumPy** to identify patterns and trends.
* Developed **custom embeddings** for **document retrieval and RAG pipelines**, improving AI-driven insights.
* Built predictive classifier models using **scikit-learn** and **TensorFlow** to classify an email as margin call email or not. Separate models are trained to further classify the margin call email as demand or antic demand.
* Utilized pre-trained deep learning models such as **Semantic Text Similarity and DeepSet** to extract domain specific **named entities** from various pdf attachments sent by clients.
* Converted **pdf attachments** to structured text documents and leveraged **SOTA pre-trained transformer** based Question & Answering to extract key information.
* Utilized historical email data to train predictive models and evaluated their performance using metrics such as **Precision, Recall, ROC-AUC curve, Precision-Recall curve and F1-score**.
* Utilized **TF-IDF vectorizer** to convert email subject into vectors and feed that data to a **LinearSVC** with linear kernel to build a binary classifier.
* Utilized AWS services such as **S3, Sagemaker** and **Lambda** to store and process data, **monitor and deploy models**.
* Conducted **hyperparameter** **tuning** using techniques such as **Grid Search** and **Random Search** to **optimize** model performance.
* Utilized ensemble learning techniques such as **Bagging** and **Boosting** to improve model performance.
* Implemented **blue-green deployments** and canary releases for minimizing downtime and risks during model updates.
* Utilized **Docker** containers and **Kubernetes** for efficient deployment and scaling of the system on cloud platforms.
* Implemented **infrastructure as Code** using tools like Terraform and Ansible to automate the provisioning and configuration of infrastructure components.
* Continuously monitored model performance and updated the models as new data and insights became available.

***Client: American Express May 2018 – Nov 2020***

***Role: Data Scientist***

***Project: Credit Card Fraud Detection***

***Responsibilities:***

* Conducted data analysis on credit card transaction data using **Python**, **Pandas**, and **NumPy** to identify patterns and trends.
* Built predictive models using **scikit-learn** and **TensorFlow** to identify fraudulent transactions.
* Developed custom **feature engineering** techniques to identify **potential fraud patterns**, such as unusual transaction frequency or high-value transactions.
* Built and evaluated classification algorithms such **as Logistic Regression, Random Forest and XGBoost** to predict fraudulent transactions.
* Utilized AWS services such as **S3, EC2, Sagemaker** and **Lambda** to store and process data, **monitor and deploy models**.
* Conducted **hyperparameter** **tuning** using techniques such as **Grid Search** and **Random Search** to **optimize** model performance.
* Utilized ensemble learning techniques such as **Bagging** and **Boosting** to improve model performance.
* Optimized model inference and data processing pipelines for maximum efficiency and reduced latency.
* Implemented security best practices for machine learning infrastructure, ensuring compliance with industry regulations.
* Developed custom visualizations and dashboards using tools such as **Matplotlib** and **Seaborn** to present analysis results to stakeholders.
* Orchestrated containerized applications using Kubernetes to manage deployments, scaling, and self-healing.
* Contributed to the development of best practices and standard operating procedures for **credit card fraud detection**.
* Conducted regular security audits and vulnerability assessments to mitigate potential risks.
* Continuously monitored model performance and updated the models as new data and insights became available.