

Sri Ram Sripada

(813)734-1536 | sriramed20@gmail.com | Tampa, Florida | [GitHub Portfolio](#) | [Linkedin](#)

WORK EXPERIENCE

Data Scientist Intern at [University of North Florida](#) | Jacksonville, FL Jun. 2023 – Aug 2023

- Designed and implemented **data pipelines** for critical stages: data ingestion, data validation, data transformation and streamlined model training using Python; optimizing model performance and reducing time-to-deployment by 30%
- Leveraged **Natural language Processing (NLP)** techniques like tokenization, topic modeling and text summarization using Python from unstructured text data resulting in 15% improvement in trend detection.
- Designed interactive **dashboards using Tableau** and **translated them into actionable data product user stories** enabling stakeholders to interpret data-driven insights; driving a 20% business growth.

Software Engineer at [Softility](#) | Hyderabad, India Jan. 2021 – Aug 2022

- Designed and implemented robust **Kafka topics using Python and Kafka for real-time data ingestion and validation**, resulting in a 30% reduction in latency.
- Developed **REST API's** using **Python** helped in reducing **data retrieval** and processing time for network devices by 10% through efficient utilization of SNMP protocol, resulting in enhanced **data accessibility**.
- Reduced software **deployment time by 25%** in a DevOps environment through **development of Linux/Unix shell scripts**; leveraged **Git version control** and **CI/CD pipelines** with Jenkins.

Machine Learning Engineer Intern at [SkillAscend](#) | Hyderabad, India Aug. 2020 – Dec 2020

- Performed thorough exploratory data analysis, **feature engineering**, **anomaly detection** and leveraged Python to implement Machine Learning algorithms, resulting in a significant 25% increase in model's accuracy for both classification and regression.
- Containerized machine learning models using Docker for model deployment and testing
- Applied **statistical techniques** (regression analysis, time series forecasting, clustering, A/B testing) contributing to a 15% improvement in data-driven decision accuracy.

PROJECTS

Credit Card Fraud Detection | Python, Machine Learning, Mlflow, Streamlit, API's

- Orchestrated robust **end-to-end data science pipelines using Python**, handling data ingestion, data validation, data preprocessing, **model training**, and **model deployment using Streamlit** to ensure accurate fraud prediction.
- Implemented real-time monitoring of ML models for experiment tracking on DagsHub's free hosted MLflow server.

Predicting Service Faults on Telecommunication Network Data | Python, Data Analysis [\(Link\)](#)

- Performed extensive data exploration, **exploratory data analysis**, and feature engineering, leading to a significant 10% improvement in model performance through hyper parameter tuning.
- Deployed XG Boost Classifier** model using Python- Streamlit increasing accuracy by 15%; reduced false positive rate by 20%.

Amazon Fine Food Reviews | Natural Language Processing | Data Extraction [\(Link\)](#)

- Engineered a **data extraction** utilizing SQLite database and Python, resulting in a 40% reduction in data processing time
- Transformed text data through the implementation of advanced text featurization techniques including TF-IDF, and Avg W2Vec, resulting in a 25% improvement in text analysis accuracy.

EDUCATION

University of South Florida | MS in **Business Analytics and Information Systems** Tampa, FL

- GPA: 3.71/4.0 Expected May. 2024

Osmania University | Bachelor of Engineering in **Electronics and Communication Engineering** Aug. 2020

SKILLS

Python | R | C | SQL | Excel | Descriptive and Inferential Statistics | Clustering | Hypothesis Testing | A/B testing | ANOVA | Time Series | Spark | Hadoop | MongoDB | Python (Pandas, NumPy, Scikit-learn, RESTful API Services, PySpark, NLTK, Tensorflow, Keras, Flask) | Classification | Decision Tree | Random Forest | Natural Language Processing | Topic Modeling | Plotly Dash | Supervised | Unsupervised Learning | Data Science Pipeline (cleansing, wrangling, visualization, modeling, interpretation) | VS Code | R studio | Git | Linux | Docker | Postman | Kafka | Spark | Tableau | Power BI | SQL Server | PostgreSQL | Splunk | Jenkins | Qlik

CERTIFICATION

Databricks Accreditation Lakehouse Fundamentals [\(Link\)](#)

Distributed Storage and Processing Fundamentals Certification [\(Link\)](#)

Presented Summer Internship Project at **DSSG 2023 Conference**, highlighting **social impact** through **data-driven solutions**.