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
- Spearheaded **data quality initiatives** by identifying and resolving data concerns; executed **data cleaning and preprocessing** operations using **Python programming**, leading to a significant **40% decrease in data errors** and enhanced data integrity for business analyses
- Led team of up to 2 Interns in executing a comprehensive strategy for **leading Non-Government Organization** using Python for data analysis and **Tableau** for impactful visualizations **driving business growth by 20%.**
- Leveraged **Natural language Processing (NLP)** techniques **like tokenization, topic modeling and text summarization** using Python to extract valuable insights from text data, resulting in **15% improvement in trend detection**.
- Designed interactive **dashboards using Tableau**, enabling stakeholders to interpret data-driven insights, leading to **2x faster response** to market trends across industries.

Software Engineer at Softility | Hyderabad, India Jan. 2021 – Aug 2022
- Developed **Python scripts** that reduced **data retrieval** and processing time for various network devices by 40% through efficient utilization of SNMP protocol and **REST API's**, resulting in enhanced **data accessibility** and operational efficiency for a leading US broadband company.
- Optimized and worked on validating quality assurance of data polled from **Kafka topics** which maximized data quality by 20%.
- Reduced software **deployment time by 50%** 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.
- Implemented data analysis using **statistical techniques** to monitor key performance metrics like $R2$ score and log loss.
- Designed **database schemas** to store scraped data and unified data storage resolving 30% of data quality issues faced in modelling.

PROJECTS

Predicting Service Faults on Telecommunication Network Data (Python, Machine Learning) (Link)
- Performed extensive data exploration, **exploratory data analysis**, and feature engineering, leading to a significant 10% improvement in model performance through hyper parameter tuning.
- Achieved remarkable log-loss of 0.52, placing within the **Top 300** on the highly competitive **Kaggle Leadership Board**.
- **Developed and deployed** an XG Boost Classifier model which increased accuracy by 15%; reduced false positive rate by 20%.

Amazon Fine Food Reviews (Text Classification, Natural Language Processing) (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.
- Optimized **machine learning model** selection through **Hyper-parameter tuning**, evaluated and compared performance using train and test AUC scores, accuracy, and F1 scores.

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| MS 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 | Ploty 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 Internship Project at **DSSG 2023 Conference**, highlighting **social impact** through **data-driven solutions**.