RAAGA SINDHU MANGALAGIRI

Boston, MA(Open for relocation)| (857) 693-9649 | raagasindhu99@gmail.com | LinkedIn | GitHub

EXPERIENCE

Northeastern University, Boston   **Jan 2024 - Aug 2024**
Machine Learning Intern
Data Pipeline Development

- Engineered the CoreLogic tax record processing pipeline for a NASA-funded project, implementing end-to-end ETL processes using Python and Dask, leading to a 35% improvement in scalability
- Optimized workflows for data loading, preprocessing, and cleaning using advanced parallelization with Dask, reducing loading time by 85%
- Constructed modular data-handling tools like Python and Airflow, ensuring adaptability for future enhancements
- Analyzed geospatial analysis with ArcGIS and Python to study housing disparities and municipal services access, delivering actionable insights for resource allocation

Northeastern University, Boston   **Sep 2023 – Dec 2023**
Graduate Teaching Assistant – Computation and Visualization

- Guided over 100 students and collaborated with the faculty to develop course materials, boosting student engagement and project comprehension by 20%

Tata Consultancy Services, India   **Jul 2021 - Aug 2022**
Data Analyst

- Automated data workflows using Salesforce, streamlining data analysis and process improvement techniques to streamline data processes, reducing manual processing time by 30% and enhancing operational efficiency
- Integrated Salesforce with JIRA implementing ETL pipelines to automate demand planning, task management, and forecasting, resulting in 25% reduction in delays and improved business intelligence
- Utilized statistical analysis and trend analysis to identify bottlenecks, optimizing workflows for scalability and enhancing decision support capabilities across teams
- Developed dashboards and reports in Salesforce to provide actionable insights into key performance indicators (KPIs), improving transparency and team collaboration
- Delivered training sessions and detailed documentation, empowering teams to maximize the use of Salesforce tools for improved reporting, data visualization, and overall system adoption

PROJECTS

Telecom Device Upgrade Prediction – Python, GCP, Airflow | Link

- Designed and deployed scalable workflows for data preprocessing, feature engineering and model training using GCS, Airflow DAGs, and Vertex AI AutoML, automating predictions with SQL based logic for real-time decision-making

Breast Cancer - Python (Pandas, NumPy, Logistic Regression) | SQL | MS Excel | Link

- Implemented logistic regression on the UCI datasets with Python and SQL, achieving a 95% accuracy for early detection and improved planning

311 NYC Data Analysis - Google Big Query| Python| Tableau prep & Tableau | Link

- Analyzed 14 million NYC service requests with Python and Tableau, creating a dashboard to optimize city resources and enhance complaint resolution

EDUCATION

Northeastern University, College of Engineering, Boston, MA   **2022- 2024**
Master of Science in Data Analytics Engineering
Relevant Coursework – Machine Learning Operations, Foundation in Data Analytics, Statistical Learning, Computation and Visualization

Sreenidhi Institute of Science and Technology, Hyderabad, India   **2017-2021**
Bachelor of Science in Electronics and Computer Engineering
Relevant Coursework – Statistical Learning, Big Data Analytics, Machine Learning, Network Analysis

TECHINICAL SKILLS

Programming & Libraries	-	Python (Pandas, NumPy, Scikit-learn, Matplotlib, Plotly), MySQL, R,
Data Engineering	-	Dask, Airflow DAGs, MLflow
Cloud Platforms	-	Google Big Query, GCP, Docker, Kubernetes
Data Visualization	-	Tableau, Power BI, Flourish
Tools & Collaboration	-	Salesforce, JIRA, GitHub, Excel