

RAGHAVENDER REDDY SRIPATHI

+1 (940)-331-8374

Denton, TX | Raghureddy170454@gmail.com | [linkedin.com/in/raghavender-reddy-sripathi-262b00203](https://www.linkedin.com/in/raghavender-reddy-sripathi-262b00203)

SUMMARY

Software and Data Engineer with 3 years of experience in designing and optimizing high-quality software and data solutions. I am proficient in Python, C++, SQL, Java, and ETL processes, with expertise in machine learning, cloud computing (AWS, Azure), and data visualization (Tableau, Power BI). I am skilled in developing RESTful APIs, database optimization, and microservices architecture, leading to enhanced system efficiency and scalability. Passionate about leveraging data analytics and automation to drive business insights and streamline operations. Adept at collaborating in agile environments, mentoring junior developers, and implementing CI/CD pipelines to improve deployment workflows.

TECHNICAL SKILLS

Programming Languages	Python, SQL, C, C++, Java, R-Programming.
ML/DL Libraries	Scikit-learn, NumPy, Pandas, Matplotlib, SciPy, Keras, TensorFlow, OpenCV, Pytorch.
Technical Tools	Tableau, Power BI, Spark, PySpark, Hadoop, Big Query, Jenkins, Git, Wireshark.
Cloud Technologies	AWS, Microsoft Azure, Redshift, S3, SageMaker.
Web-Applications	HTML, CSS, JavaScript.
Databases	SQL, MySQL, PostgreSQL, MongoDB.
IDE	PyCharm, Visual Studio, IntelliJ IDEA, Android Studio.
Game Engines	Pygame, Pyglet, Panda3D, Ursina.

WORK EXPERIENCE

Open group Inc.

Software Engineer

Irving, Texas | Feb 2024 – Present

- Designed and managed large-scale training datasets and pipelines for ML models using Spark and **Big Query**.
- Developed, tested, and maintained high-quality software solutions using Python, enhancing overall system efficiency by 20%.
- Designed and optimized database schemas using **PostgreSQL** and **Amazon RDS**, improving query performance by 30%.
- Built and deployed **FastAPI + Pydantic** microservices for internal ML models, integrated into GitHub Actions CI/CD workflows.
- Created **ETL pipelines** to process raw CSV/JSON from third-party sources; profiled and cleaned data using **Pandas**, flagged outliers proactively.
- Conducted **exploratory data analysis (EDA)** to spot performance bottlenecks and usage trends without prompting from PMs.
- Implemented CI/CD workflows using **Jenkins** and **GitHub Actions**, reducing deployment downtime by 40%.
- Managed containerized applications using **Docker** and **Kubernetes**, enabling scalable microservice deployment.
- Conducted data security reviews, implementing RBAC and encryption protocols to safeguard sensitive information.
- Integrated **monitoring tools** and performed system optimizations, improving user experience by 30%.
- Mentored junior developers and led **Agile Scrum** meetings, fostering a collaborative and high-performance engineering culture.
- Contributed to deep learning proof-of-concepts using **TensorFlow** and **Keras** for internal AI model development.
- Used **Jupyter Notebooks**, **Matplotlib**, and **Pandas** for data exploration and analysis of system logs and performance metrics.

Pupil Share-Software

Data Engineer

Hyderabad, India | sept 2021 – June 2022

- Built ETL pipelines and automated data ingestion workflows using Apache Airflow and Spark.
- Designed and maintained **ETL pipelines** using **Apache Spark** and **Apache Airflow**, significantly enhancing data throughput and reliability.

- Built scalable data storage and processing solutions using **AWS Redshift, S3**, and data lakes.
- Applied data cleaning and transformation techniques using **Python, Pandas**, and **NumPy**, improving data integrity.
- Worked with both **SQL (PostgreSQL, Amazon RDS)** and **NoSQL (MongoDB)** databases for diverse data storage needs.
- Automated data workflows with **Airflow**, ensuring smooth and timely execution of data pipelines.
- Monitored data infrastructure performance and implemented optimizations that reduced data processing latency.
- Supported **ML model deployments** on **AWS SageMaker**, contributing to real-time and batch inference capabilities.
- Ensured **data security** and compliance by applying encryption, IAM, and secure access controls.
- Built actionable dashboards and reports using **Power BI** and **Tableau** for internal stakeholders.
- Gained hands-on experience with **cloud platforms (AWS)** and exploratory work with **Big Data** tools such as **Apache Spark** and **Databricks**.
- Familiar with event-streaming tools such as **Kafka**, used for log data ingestion and processing.
- Participated in full **software development lifecycle (SDLC)**, from gathering requirements to deployment and monitoring of ETL pipelines.

Tech Stack: Object-Oriented Programming (OOP), AWS (S3, Lambda, EC2, CloudWatch), Git, Jenkins, Agile (Scrum), CI/CD, SDLC, C#, .NET MVC, JavaScript, HTML, Software Development, Data Structures, Python, C++, Java, Node.js, REST APIs, ETL.

EDUCATION

University of North Texas, The State University of Texas

DENTON, TX

Master of Science in Computer Science, CGPA: 3.54/4.0

August 2022 – December 2023

Courses: Computer Algorithms, Machine Learning, Information Retrieval, Computer Networking, Graph Theory, fundamentals of AI, Computer Architecture, Big data, Software Engineer, Data Analytics.

CMR Technical Campus, Hyderabad

August 2017 – sept 2021

Bachelor of Engineering in Electronics and communication Engineering.

Courses: Computer programming in C, Network Analysis, Analog Electronics, Electrical Technology, Control systems, Analog communication, Operating systems, VLSI Design, Java programming, Computer Networks, Digital image processing, Digital signal processing, Microprocessors and microcontrollers.

PROJECTS

- **Customer Segmentation and Lead Optimization:** Analyzed customer data to segment users based on behavior and demographics using Python (Pandas, Scikit-learn) and SQL.
- **ETL Pipeline for Sales and Marketing Analytics:** Built and maintained ETL pipelines using Apache Spark and Apache Airflow to automate data ingestion and processing for sales and marketing teams.
- **Financial Data Dashboard Development:** Extracted and processed data using SQL to create real-time visual reports for executives, aiding data-driven decision-making.
- **Automated Testing Framework:** Built an automated testing framework using Selenium and Python, increasing test coverage and reducing manual testing time by 50%.
- **Information Retrieval Project:** Designed and implemented a text parser, indexer, and retrieval engine using TREC Data for efficient information retrieval.
- **Distributed Systems Research:** Conducted research on enhancing training performance and robustness of distributed systems for IoT-based robotics applications.
- **Learning Management System (LMS):** Led the design and development of an LMS that organized educational courses, training resources, and learning programs. Enhanced user experience with tracking, reporting, and automation functionalities, advancing e-learning through practical software applications.
- **Hospital Length of Stay Prediction:** Created a machine learning model to predict hospital length of stay using Decision Tree, Random Forest, KNN, SVM, and Logistic Regression algorithms.

CERTIFICATIONS

- Programming for Everybody (Python).
- DeepLearning.AI.
- Internet of Things workshop.
- Fundamentals of Analog electronics.