Sowmya Raghavendra

P: +1 (470)-680-5113 | sowmyaraghu1901@gmail.com
LinkedIn: https://www.linkedin.com/in/sowmya-raghavendra | Portfolio: https://sowmyaraghu19.github.io

WORK EXPERIENCE

GEORGIA STATE UNIVERSITY — **Atlanta, GA**
Graduate Research Assistant — **August 2024 – May 2025**

- Architected a machine learning model that increased business accreditation accuracy by 25%.
- Performed sentiment analysis using Natural Language Processing (NLP) on company news articles, strengthening fraud detection capabilities.
- Led a team of 10+ researchers to implement data-driven accreditation processes, streamlining operations and reducing manual verification by 40%.

MERCEDES BENZ RESEARCH & DEVELOPMENT INDIA — **Bangalore, KA, India**
Software Consultant — **August 2022 – August 2024**

- Developed and optimized ETL pipelines in PySpark, and SQL improving data processing efficiency by 20%.
- Engineered real-time data streaming solutions using Apache Kafka, enhancing system scalability and reducing data ingestion latency by 40%.
- Designed data models for microservices and structured data storage to support both transactional processing and downstream analytics.
- Managed and deployed containerized microservices via Kubernetes and Docker, improving system reliability and uptime by 30%.
- Automated CI/CD pipelines on Azure DevOps, reducing deployment time by 35% and improving system maintainability.
- Led cross-functional collaboration with product managers and analysts to align data strategies with business goals.

PROJECTS

NETFLIX DATA ENGINEERING PIPELINE | *Azure, PySpark, Airflow, Delta Live Tables, SQL, Databricks*

- Built an end-to-end data pipeline using Azure Data Factory and Databricks, reducing manual intervention by 80%.
- Optimized data ingestion workflows with validation checks, boosting query performance by 50%.

LEGALDOC AI ASSISTANT | *Streamlit, OpenAI, LangChain, FAISS, Python*

- Built an AI tool to auto-generate legal documents from case chats, reducing drafting time by 90%.
- Enabled context-aware Q&A using RAG with FAISS, achieving 85% semantic accuracy on legal queries.

US ACCIDENTS SEVERITY PREDICTION | *Python, Scikit-learn, Pandas, Voting Classifier*

- Analyzed 2.8M+ U.S. traffic accident records to address the challenge of predicting accident severity using environmental, time, and location features.
- Deployed Voting Classifier achieving 81% accuracy and 78% F1 score, improving predictive reliability for emergency response and traffic safety planning.

SKILLS

Programming: Python, SQL (MySQL, PostgreSQL), PySpark, Java
Big Data & Analytics: Apache Spark, BigQuery, Databricks
ETL & Data Pipelines: Apache Airflow, Kafka, Data Factory, Blob Storage, Pipelines
Cloud & DevOps: Azure (Data Factory, Databricks, Redis Cache, Log Analytics), GCP
Visualization & Tools: Power BI, Looker Studio, Jupyter Notebook
Machine Learning Competencies: Regression, Clustering, Decision Trees, Random Forest, Neural Networks
Version Control & Workflow: Git, Jira, CI/CD (Azure DevOps, Docker, Kubernetes)

EDUCATION

GEORGIA STATE UNIVERSITY — **Atlanta, GA**
Master of Science in Data Science and Analytics, GPA – 4.0/4.0 — **August 2024 – December 2025**
Relevant Coursework: Data Analysis, Machine Learning, Artificial Intelligence

DAYANANDA SAGAR ACADEMY OF TECHNOLOGY & MANAGEMENT — **Bangalore, KA, India**
Bachelor of Engineering in Computer Science, GPA – 3.87/4.0 — **August 2018 – July 2022**
Relevant Coursework: Data Management, Big Data Analytics, Data Programming