Sanjay Dadadahalli Prasanna Kumar

Miami, FL | sdada004@fiu.edu | 786-784-9890 | https://github.com/sanjaydp | https://linkedin.com/in/SanjayDP

EDUCATION

Master of Science, Data Science & AI
Florida International University **GPA:** 3.70/4.00
Coursework: Machine learning, Data Mining, Data Science, AI, Data Modeling, Data Analytics.
Bachelor of Engineering, Electronics & Communication Engineering
SJB Institute of Technology **GPA:** 3.72/4.00

EXPERIENCE

Graduate Research Assistant, Florida International University Sep 2024 – Present

- Built and maintained automated ETL pipelines using Python and Pandas to clean, transform, and aggregate over 500,000+ rows of sensor data collected from embedded systems.
- Developed a **Streamlit-based** data interface integrated with **Flask APIs** to monitor and control live microcontroller data across 10+ devices, reducing manual intervention by **80%**.
- Deployed **AWS Lambda** and S3 for **real-time data ingestion** and ML inference, cutting report generation time from **20 minutes** to under **2 minutes**.

Senior Data Engineer, Brillio Technologies Pvt Ltd Nov 2021 – Jul 2024

- Built **scalable royalty application** for a leading music corporation, fully orchestrated on AWS to maximize data accessibility for over **1 million artists worldwide**.
- Spearheaded the development of an event-driven data platform using **AWS Lambda** and **Step Functions**, for real-time generation of royalty statements.
- Implemented a data validation microservice using **Great Expectations**, optimizing memory utilization by 40% and improving data integrity.
- Devised an automated mechanism to synchronize downstream **low-latency databases (Aurora)** and **real-time search (OpenSearch)** with highly updated data from the **data warehouse (Redshift)**.
- Optimized **Redshift query performance** by **35%** through table redesign, use of **sort/dist keys**, and **workload management tuning**, significantly reducing report latency for end users.
- **Awards** : Recognized as "Brillion of the Month" for two consecutive months.

Associate Professional Software Engineer, DXC Technologies Pvt Ltd Jun 2021 – Nov 2021

- Migrated legacy **COBOL-based ETL workflows** into Python- and SQL-based batch pipelines, improving maintainability and reducing processing time by **20%**.
- Refactored data ingestion and transformation logic for financial systems, increasing system uptime to **99.9%** and ensuring high availability for downstream reporting dashboards.

SKILLS

- **Programming Languages**: Python, C++, SQL
- **Big Data Technologies** : Hadoop, Apache spark, Map Reduce, kafka
- **Databases**: PostgreSQL, MySQL, Relational and NoSQL databases, Data Modeling (Snowflake Schema)
- **Machine Learning & AI:** Scikit-Learn, Pytorch, Tensorflow
- **Cloud Platforms**: AWS (Certified), Azure (Certified)
- **Data Visualization & DevOps Tools**: Power BI, Docker, Kubernetes, Terraform, Git, Microsoft Excel, Microsoft Office
- **Methodologies**: Object Oriented Programming (OOP), Agile (Scrum) SDLC
- **Operating Systems**: Unix/Linux, Windows

PROJECTS

Real-Time Retail Data Pipeline *(Individual)*
GitHub : https://github.com/sanjaydp/realtime-retail-data-pipeline

- Designed and implemented a comprehensive real-time retail **data pipeline** using **Apache Kafka, Spark Streaming, and Airflow**, featuring automated data quality validation, real-time monitoring with **Prometheus/Grafana, and RESTful API** integration for end-to-end retail analytics processing.

Driver Allocation Optimization using Machine Learning *(Team Lead)*
GitHub : https://github.com/sanjaydp/Driver_Allocation_Optimization

- Implemented **LSTM, XGBoost, and reinforcement learning** models to optimize driver allocation, **forecasting**, and real-time route decisions, reducing idle time by **20%** and improving trip efficiency by **10%**.

CERTIFICATIONS

AWS Certified: Cloud Practitioner

Microsoft Certified: Azure Data Fundamentals