

SAI KALYAN VANAMA

+1 - 619. 622. 8258 | saikalyanvanama@gmail.com | Arlington, TX, USA | [LinkedIn](#)

TECHNICAL SKILLS

- **Proficient:** Python (including libraries like PySpark, Pandas, NumPy, Matplotlib, Seaborn, NLTK), SQL, Microsoft Azure (includes Databricks, Key Vault, Synapse Analytics, Data Factory, Data Storage), Azure Devops, SQL Server (SSDT, SSMS, SSIS, SSAS), Data Extraction, Transformation, and Loading (ETL), Data Modeling, Data Warehousing, Data Pre-processing, Machine Learning Algorithms, Data Visualization (Power BI), Apache Airflow, AWS (S3, Glue, RDS, Redshift, Athena, Lambda), PowerBI, Git, Snowflake

ACADEMICS

University of Texas at Arlington, Arlington, Texas, USA

Masters in Computer Science

Visvesvaraya Technological University, India

Bachelors of Engineering – Information Science and Engineering

GPA : 3.8/4

Graduated: May 2024

CGPA: 7/10

Year : 2019

PROFESSIONAL EXPERIENCE

Mindtree | Data Engineer | Bangalore, India | 3 Years

- Authored and managed DAGs in Apache Airflow to automate ETL pipelines, successfully integrating data from multiple sources into Azure Synapse Analytics. This initiative led to a 40% reduction in data processing time and significantly boosted data analytics efficiency.
- Enhanced ETL workflows within Azure Synapse Analytics to ensure efficient data storage and accessibility, resulting in a 30% increase in report generation speed and improved data accuracy, thereby supporting better decision-making.
- Orchestrated the development of ETL pipelines leveraging Azure Data Factory and Databricks; streamlined data migration from on-premises SQL Server to Azure Data Lake Gen2, resulting in a 40% reduction in data processing time.
- Enhanced data transformation and analysis processes through Databricks, optimizing data storage in Parquet format which slashed query execution time by an impressive 50%, boosting overall data processing efficiency and accuracy.
- Established Snowflake as the data warehouse and leveraged Power BI for real-time reporting and analysis, providing actionable insights to stakeholders resulted in streamlined data access, actionable insights, cost optimization, and improved stakeholder satisfaction.
- Implemented CI/CD practices using Azure devops to automate the build, test, and deployment of ETL pipelines, resulting in a 50% reduction in deployment time and 80% improvement in data quality.
- Demonstrated leadership by managing a team of four individuals, overseeing testing activities within the Agile framework, and achieving a 30% improvement in project efficiency.
- Facilitated communication between stakeholders, managers, and team members, contributing valuable suggestions that resulted in a more robust and user-friendly application.

ACADEMIC PROJECTS

Real Time Data Analysis on Airbnb Dataset | Data Engineer

Developed a scalable and efficient data pipeline to ingest, transform, and analyze Airbnb listings data using AWS services

- Utilized AWS Lambda to ingest real-time Airbnb data into Amazon S3, ensuring efficient storage with date-based partitioning.
- Implemented AWS Glue ETL jobs to clean, transform, and normalize data, leveraging Glue Data Catalog for metadata management.
- Loaded data into Amazon Redshift for complex queries and used AWS Athena for ad-hoc analysis; visualized insights with Amazon QuickSight dashboards.

Sentiment Analysis on Twitter Dataset | Data Science

Implemented sentiment analysis on Twitter data using Recurrent Neural Networks

- Spearheaded the development of a cutting-edge Sentiment Analysis model for Twitter data using advanced deep learning techniques, showcasing expertise in Natural Language Processing (NLP) and enabled accurate sentiment analysis for actionable insights and decision-making.
- Executed comprehensive data pre-processing, encompassing text cleaning, tokenization, and word embedding, on a Kaggle Twitter dataset to ensure pristine input quality for the sentiment analysis model
- Prepared an optimized Recurrent Neural Network (RNN) architecture with embeddings and dropout layers, significantly improving the efficiency of sentiment analysis by 80%

LEADERSHIP + AWARDS

Got 4 times 'SpotOn' Award for showing exceptional performance in the project, Mindtree

Secured 2nd spot in Hackathon, Sri Venkateshwara College of Engineering