
PROFESSIONAL SUMMARY

Microsoft Azure certified Data Engineer with a strong background in building robust **ETL** pipelines, optimizing data processing, and data modeling. Demonstrated success in solving complex problems, optimizing existing workflows, reducing processing times, and delivering data-driven insights. Skilled in **Python**, **PySpark**, **Kubernetes**, **SQL** and cloud technologies.

TECHNICAL SKILLS

Certifications Microsoft Certified Azure Data Engineer Associate, Talend Certified Data Integration Developer

Skills Python, SQL, Java, AWS, Azure, Kafka, ActiveMQ, Azure Data Factory, Azure Kubernetes Services, Jira, Git, Docker

PROFESSIONAL EXPERIENCE

Data Engineer

Pepsico

May 2023-Present

Plano, TX

- Engineered a high-velocity Python **ETL** pipeline, capturing **30k** change records per second from SQL Server **supply chain** tables to EventHub/**Kafka**, achieving a **10%** reduction in latency.
- Stream data from **EventHub** to Data Lake storage leveraging Spark Streaming in **Python** resulting in a 15% reduction in processing latency.
- Spearheaded the **Kubernetes** migration of a critical application, resulting in a remarkable **60%** improvement in **multiprocessing** performance.
- Validated data integrity post-data ingestion, ensuring the reliability of information between source data and the bronze layer.
- Executed **PySpark** transformations on **sales** data, seamlessly transitioning information from the silver layer to the gold layer.

Data Engineer

Artha Solutions

May 2022-May 2023

Remote

- Developed the back-end of a data governance tool in **Python**, ingesting data, validating against predefined rules, and implemented 12 metrics for data-driven insights into the quality and compliance of ingested data.
- Built an **ETL** pipeline using Python and **ActiveMQ** to ingest 100GB of data from source files or relational tables to a destination **SQL** database, achieving 4 times faster processing.
- Designed and implemented Talend jobs for seamless integration of data from four different sources into SQL Server tables.
- Successfully migrated data from PostgreSQL to SQL Server, encompassing 150 tables, stored procedures, and views.

Data Analyst

Deloitte

Jul 2019-Dec 2020

Hyderabad, Telangana

- Conducted data aggregation, transformation, analysis and interpretation of logs data from AWS cloud platform stored in S3 buckets using **Elastic Search** tools, resulting in a 20% reduction in response time for vulnerability security operations.
- Developed a vulnerability security operation automation workflow using Python-based serverless APIs in **AWS Lambda** and **NoSQL database**, leading to a 25% increase in security monitoring efficiency and a 15% improvement in incident response time.
- Automated cloud technical security controls and managing IAM roles and permissions using **Python** reducing security risks by 30%.

ACADEMIC PROJECTS

Self-supervised Detection of Spine Anomalies: Implemented novel **self-supervised learning** methods on spine X-ray images for classification and achieved an improved AUROC metric of 0.85 from baseline on Vindr-SpineXR dataset.

Recommendation System: Conducted feature engineering and text preprocessing to implement a hybrid recommendation system with **XGBoost** and recorded an RMSE value of 0.977 on Yelp reviews dataset of 6 million records.

Sentiment Analysis with Transformers: Constructed **Transformer** blocks for **BERT LLM** model to pre-train and fine-tune a model for **sentiment analysis** and attained 90% validation accuracy on IMDB reviews dataset with 25k training and 25k test examples

Fake News Detection: Executed sophisticated Fake News detection on the LIAR dataset, analyzing real political statements. Implemented Siamese **BERT** models with shared weights and applied domain adaptive pre-fine tuning, resulting in a remarkable **6.25%** accuracy improvement.

EDUCATION

Master of Science in Computer Science

University of Southern California, Los Angeles, CA, USA

Bachelor's of Tech in Computer Science Engineering

Sreenidhi Institute of Science and Technology, Hyderabad, India