Naga Sai Gopi Krishna Lingamallu

+917989715372   lnsgopikrishna@gmail.com   www.linkedin.com/gopikrishnal

SUMMARY

6+ years experienced, meticulous & result-oriented Senior Data Engineer armed with a proven track record of analytical acumen in developing (deploying) complex data pipelines (machine learning and statistical modeling algorithms/techniques for identifying patterns and extracting valuable insights). Possesses diverse experience in planning & executing multiple projects and liaising with the key stakeholders to identify & resolve business problem statement and deliver excellent results.

EDUCATION

IIIT Bangalore | Bengaluru, IN   *Jul '23 - Aug '24*

Post Graduation Diploma in AI and ML (Specialization in MLOps)

Birla Institute of Technology & Sciences (BITS)- Pilani | Hyderabad, IN   *Aug'14 – Jun'18*

B.E (Hons) in Mechanical Engiineering

TECHNICAL SKILLS

Languages: Python, Java, HTML, JavaScript

Tools & Technologies: GCP, Cloud Dataflow, Composer, Kubernetes, Apache Spark, Apache NiFi, Docker, Jenkins, Git, Terraform, NoSQL (MongoDB), BigQuery

Machine/Deep Learning: Linear/Logistic Regression, Decision Trees, Random Forests, CNN, RNN, LSTM, MLOps

Certifications: Machine Learning Certificate | Coursera

PROFESSIONAL EXPERIENCE

Zebra Technologies (formerly Reflexis Systems) | Senior Software Engineer   *Dec '22 -Present*

Next-Gen Analytics & Reporting Platform

A common platform to receive & process the data from all the customers
- Conducted POC's to assess the correct tool for *capturing CDC data* from existing IBM DB2 to migrate to cloud native SQL offerings
- Tested with CDC tools like QLIK replicate, Debezium for migrating the data from DB2 to PostgresSQL
- Developed Airflow jobs using Google Composer to write the migrated data to BigQuery for analytical purpose
- Analyzed customers data on Cognos custom report to determine the frequently used fields in their reports to determine the Partioning & Clustering strategy of data in BigQuery

Migration to Google Cloud Platform

- Conducted POC's to assess compatibility of legacy pipelines with Google Cloud Platform (GCP)
- Dockerized existing ETL and streaming solutions to maintain continuity before moving to Cloud based solutions
- Migrated & optimized the spark streaming jobs to be compatible with Google Cloud Dataproc.
- Developed Cloud Native Google Dataflow (powered by Apache Beam) to replace both the existing ETL, streaming & aggregation pipelines
  - Enabling dual functionality for batch and streaming, to streamline data pipelines and implement a serverless solution

Zebra Technologies (formerly Reflexis Systems) | Software Engineer I & II   *Jul '18 - Nov '22*

Data Streaming pipeline for AI Performance (a retail product offering of Zebra)

- Architectured a near real-time data streaming pipeline for the application leveraging Apache Spark Structured streaming for reading and aggregating real-time data from Mongo utilizing the Spark framework in Scala, with the goal of reducing latency and aggregating data at varying time intervals upon arrival
- Created a custom Microbatch streaming source for Mongo to aggregate data according to specified granularity, leveraging Apache Spark in Scala
- Improved Spark performance by Fine-Tuning to identify the optimum parameters like number of cores, partitions, memory etc. for cluster and standalone mode
- Implemented custom check pointing for Spark using MongoDB. Scaled modules to handle data spikes and to minimize

downtime
- Designed and executed custom check pointing for Spark utilizing MongoDB's cursor to enhance system reliability and manage downtime and unexpected outages

Data ingestion pipeline for AI Performance
- Designed real-time data ingestion ETL pipeline using Apache NiFi to read data from multiple sources
- Built pipelines/flows using Apache NiFi as a part of ETL solution and Apache Spark for streaming & aggregating data to handle more than 1 GB of data per minute.
- Formed ETL groups to collect data from files in CSV, XML, JSON and other formats transforming the data and stored into Mongo/ HDFS/ FileSystem/pub-sub.
- Enforced secure NiFi to be able to login using via LDAP and implemented repository-based system to track changes using NiFi.
- Integrated Apache NiFi (an ETL tool) with UI using API calls to deploy flows on the go. Written custom ETL processors to cater the special needs of clients.

VOLUNTEER EXPERIENCE & POSITION OF RESPONSIBILITY

Volunteered at Nirmaan Organization & NSS-College chapter

Educational Initiatives Head (Nirmaan Organization,an NGO, BITS-Pilani Hyderabad Chapter)   Aug 2016 – May 2017