

nagapavan.data@gmail.com

+918179363863

Hyderabad

India

Skills:

- Spark, Hadoop, Kafka, Spark Streaming, Apache Phoenix, Apache Flume, Sqoop, Apache HDFS, Yarn, Kerberos.
- Pandas, NumPy, Fast Api, Py-Spark
- Data Quality, Data Pipeline, System Design, ETL & ELT framework
- AWS, HDP, CDP, Airflow, Azkaban
- Python, SQL, HQL,
- Windows, Linux & Mac
- Delta Table, Hive, HBase, Redshift, RDS, NoSQL, Data Lake. MySQL
- IntelliJ, Eclipse, Spyder, Jupyter, VS Code, GitLab, VSTS, Bitbucket, GitHub

Education:

Btech:

JNTU 2017

Nagapavan Manuri

Summary:

- Offering 7+ years of experience as a data engineer with understanding of large volume data processing and ingestion framework design.
- Has strong experience in real time and batch mode data processing and load handling on cloud and on premises cluster.
- Has analytical ability, communication & coordination skills and ability to learn new things quickly.
- Expert on high and optimized quality of technical work, self-driven and dedication towards work and highly self-motivated.

Experience:

JP Morgan Chase – Lead Data Engineer (Banking domain)

Dec 2021 - Current

- Designed and develop real time and near real time data ingestion, Processing pipeline.
- Handling Growth team and provided technical guidance to developers & team.
- Managed the full life cycle of software development projects.
- Identified technology trends and recommended system improvements.
- Collaborated with stakeholders to define business requirements for applications.
- **Project:** In this project I am working a Tech Lead Data Engineer and managing Paytm **growth market** data ingestion and processing pipeline. We get data from different Paytm business and apply transformation, aggregation to build some meaningful feature which could hold some customer information and running our growth campaign on it.
- **Tech Used:** Spark, Kafka, AWS, S3, Athena, Data Lake, Starburst, Scala, Python, Kafka, Spark Streaming, Airflow, Azkaban, Delta Table, Pandas, NumPy, Fastapi.

Deloitte Consulting - Senior Developer

Apr 2019 – Nov 2021

- Migration and Optimization of existing queries to improve query performance
- Analyzed user requirements, designed and developed ETL processes to load enterprise data into the Data Warehouse.
- **Project:** Change Healthcare's Patient Identity Management solution uses demographic information from payer rosters, eligibility requests and claims, and, applying a proprietary set of matching rules and processes
- **Tech Used:** AWS, S3, Athena, Spark, Kafka, Spark Streaming, Scala, Python, Pandas.

Data Intensity - Senior Developer

Bangalore, Karnataka

Jan 2017 – Mar 2019

- Provide overall leadership to the entire module and including managing deliverable of other functional team leaders
- Collaborated with project managers to define project scope and timeline.
- Maximized performance through testing and tuning of applications.
- Developed updates to existing software to meet changing customer demands.
- ***Project:*** It is a complete package of all the features which are required to keep the network secure and active. It is having the capability to understand the security breach and generate the alarms for that to the responsible.
- ***Tech Used:*** Spark, Kafka, Hadoop, Yarn, Hive, HBase, Java, Scala, HDP, Spark Streaming, Sqoop, Flume, Phoenix.