Navyasri Muppera
Data Engineer
Phone: (913) 972-0385 | Email: navyasrimuppera5@gmail.com | Location: TX

SUMMARY

- Data Engineer with Around 4 years of experience in Data Extraction, Data Modelling, Statistical Modeling, Data Mining and Data Visualization.
- Proficient in Machine Learning algorithms and Predictive Modeling including Linear Regression, Logistic Regression, Naive Bayes, Decision Tree, Random Forest, Gradient Boosting, SVM, ANN, CNN, RNN, KNN, and K-mean clustering.
- Expert in Python libraries such as NumPy, SciPy for mathematical calculations, Pandas for data preprocessing/wrangling, Matplotlib, Seaborn for data visualization, Sklearn for machine learning, TensorFlow, Keras for Deep learning, and NLTK for NLP.
- Experience in Hadoop utilities such as Hive, HDFS, HBase, Oozie, Kafka, Sqoop SQL, MapReduce, Hive, Pig and Zookeeper.
- Working on designing data pipeline architecture in the Google Cloud Platform (GCP).

TECHNICAL SKILLS

Languages: Python, SQL, Scala
IDE's: Visual Studio Code PyCharm, Jupyter Notebook
Big Data Ecosystem: Hive, HDFS, HBase, Oozie, Kafka, Sqoop SQL, MapReduce, Pig, Zookeeper
ETL Tools: Apache Spark, Apache Airflow, Apache Oozie, SSIS, Informatica
Visualization Tool: Tableau
Cloud Technologies: AWS (S3, EMR, Redshift, Athena, EC2, Glue, Lambda, RDS), GCP
Packages: NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, TensorFlow
Database: MySQL, PostgreSQL, MongoDB, HBase, Cloudera
Operating System: Windows, Linux, MacOS

EXPERIENCE

JPMorgan Chase & Co., TX | Data Engineer **Mar 2022 – Present**

- Improving database models and querying techniques, increasing query efficiency by 20%.
- Install and configure Apache Airflow for the S3 bucket and Snowflake data warehouse and create dags to run the Airflow.
- Designing and Developing ETL Processes in AWS Glue to migrate data from external sources like S3, and Parquet/Text Files into AWS Redshift.
- Importing and exporting data jobs, to perform operations like copying data from HDFS and to HDFS using Sqoop and developed Spark code and Spark-SQL/Streaming for faster testing and processing of data.
- Prepare Data models and schema on GCP for different projects based on star and snowflake schema designs
- Developing models relying on Linear Regression, Multiple Regression, Decision Trees, Random Forest, Logistic Regression, and Naive Bayes.
- Implementing PySpark for data processing to handle data from various RDBMS and streaming sources; used Snowflake for the Data Warehouse service.

Kpit Technologies, India | Data Engineer **Mar 2019 – July 2021**

- Imported real-time data to Hadoop using Kafka and implemented the Oozie job for daily imports
- Implemented automation processes using AWS tools like CloudFormation and Elastic Beanstalk, reducing infrastructure setup time by 60% and improving team productivity by 25%.
- Achieved 90% customer monthly retention by predicting the likelihood of returning customers using a Random Forest, XG Boost algorithms.
- Worked on Pandas, NumPy, Seaborn, matplotlib, Scikit-learn, SciPy, and NLTK in Python for developing various machine learning algorithms.
- Utilized Apache Spark with Python to develop and execute Big Data Analytics and Machine learning applications, executed Machine Learning use cases under Spark ML and MLlib.
- Developed Spark streaming pipeline to batch real-time data, detect anomalies by applying business logic, and write the anomalies to the HBase table.
- Worked on CI/CD solution, using Git and Jenkins to set up and configure the big data architecture on the AWS cloud platform.
- Crafted and modified complex SQL queries, leveraging indexing and query optimization techniques to enhance database efficiency, leading to a 30% improvement in application response time.
- Loaded data into the cluster from dynamically generated files using Flume and from relational database management systems using Sqoop.

EDUCATION

Master in Computer Science **Dec 2022**
University of Missouri Kansas City (UMKC), MO

Bachelor's in Computer Science **May 2019**
Velagapudi Ramakrishna Siddhartha Engineering College, Andhra Pradesh, India