Vijay Kumar

+91 8239886299| E-Mail | LinkedIn

Education

Indian Institute of Technology, Delhi 7.03/10
Bachelor of Technology in Electrical Engineering *July 2019 – 2023*

Vidhur Navodya Sr Sec School, Kota 87.00 %
Class XII *April 2017 – April 2018*

Work Experience

Software Engineer | *Jio Platforms Limited, Mumbai* Aug 2023 - Present

Extract Transform Load Tool

- Connectors integration of storage accounts Azure, GCP bucket and AWS s3 as database sources
- Iteration functionality feature enhancement which also enables user to dynamically replace the variables
- Column rename feature which enables the user to rename the Data Frame column names with prefix, post-fix, date-time stamp and regular expressions
- Feature of join/union/intersection integration with the help of recursive function that allows user to read multiple databases in one job and increased the automation resources efficiency

Spark Libraries and Kubernetes Jobs

- Implemented the Spark streaming classes for the live streaming of data which removes the latency of minutes to the nearly real time data processing
- Configured Spark jobs using YAML files in a Kubernetes cluster, optimizing driver and executor memory settings and leveraging dynamic resource allocation for efficient processing
- Apache Spark libraries and functions for data parsing, filtering, DML execution, re-partitioning and utilized SparkSession and SparkContext for optimized data processing

Neo4J Graph Database

- Designed a topological graph setting properties of nodes and relationships with the help of Cypher query language
- Created and deployed Spark job on Kubernetes cluster for reading data from ElasticSearch into Spark Dataframe and writing it to Neo4J data base
- Performance tuning of Spark Job using aggregation functions, improved data quality and resource optimisation for processing over 80 million entries every day

Light Weight ETL

- Optimised resource utilisation with new ETL tool design implementation for small data processing
- Replication of the Scala Spark ETL tool into python module that runs on local system efficiently

Projects

Dynamic Memory Allocator | *Prof. Rahul Garg* | *Course Project* 2020

- Developed an efficient JAVA based system to allocate/free memory as per requirement using linked lists and trees
- Implemented Doubly Linked List data structure using First Split Fit algorithm to track free and allocated memory
- Implemented Best Split Fit algorithm to optimally perform allocate and free operations while minimizing fragmentation

Graph Topology Analysis | *Prof. Rahul Garg* | *Course Project* 2020

- Implemented bi-directed graph using two csv files having data regarding storylines of characters in Marvel comics.
- Implemented DFS on the graph to generate independent storylines by utilizing Hash-Map and Array-List data structures

Technical Skills

Languages: C++, Scala, Java, SQL, Cypher, Python

Softwares, Libraries, Tools and Frameworks: Spark Core, Spark SQL, Spark Streaming, Neo4J, SQLDeveloper, Kubernetes, Kafka, ElasticSearch, Git/GitHub, VS Code, IntelliJ, MATLAB, Overleaf, Microsoft Office