Vijay Kumar

+91 8239886299| E-Mail | LinkedIn

Education

Indian Institute of Technology, Delhi 7.03
Bachelor of Technology in Electrical Engineering *July 2019 – April 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

- Designed and Implemented an iteration mechanism that reduced plugin usage in the pipeline by 23%
- Connectors integration of 3 storage accounts Azure, GCP bucket and AWS S3 as database sources
- 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 and replaced **100%** custom jobs
- Column rename feature which enables the user to rename the Data Frame column names with prefix, post-fix, date-time stamp and regular expressions

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 configuration and leveraging dynamic resource allocation for efficient processing
- Utilized Spark Session API's for data parsing, filtering, SparkSQL, re-partitioning and optimized processing

Neo4J Graph Database

- Implemented a topological graph with property attributes of nodes and relationships using Cypher query language
- Optimized performance by 70% using Spark job deployment on Kubernetes cluster compared to python custom job
- Tuned Spark job performance via Spark SQL aggregations, enabling processing of 20M+ records/day

Light Weight ETL

- Designed and developed a lightweight ETL module for client-side, low-volume data processing, reducing resource usage by 35% and eliminating the need for a scrapping cluster
- Configured and deployed data pipelines using the new ETL module, improving performance, maintainability, and client usability
- Performed end-to-end testing and benchmarking on local and cluster environments; implemented debugging features including logger integration and config refresh, and resolved critical bugs

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 story-lines by utilizing Hash-Map and Array-List data structures

Technical Skills

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

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