**Gokulankannan Arumugam**

**Phone: +65-94663223**.

**Email: gokulit09@gmail.com**

**Objective :**

Business minded AI based Data Analyst with demonstrated ability to deliver valuable insights from Data Lake on the Advanced Big Data Analytics technology via analytics and advanced data driven methods.

**Summary:-**

* **4+ years of experience in Datawarehousing Business Intelligent Platform and Spark based Big Data Development and Analysis in Advancing DBS with AI team.**
* Implemented **Spark Analytical Datamart** using Scala and Java for **Bank Loyalty Program**
* Established the DBS BankWide Enterprise DataIngestion Framework to ingest data from various source to Staging using **TALEND Tool for ETL Operations and Big Data Computations.**
* Implemented **Spark Analytical Datamart** using Scala and Java. Build the **Star schema and SnowFlake schema based OLAP Cubes** accordingly for the reports to be generated for the projects
* **Schema generation and Optimized queries** are generated in Hive on MapReduce, Hive on Spark, Impala, SparkSQL. Readily available for the loaded tables in Impala and Spark
* Enhanced **Automated Encryption of HIVE** user layer sensitive column information
* Enabled **Realtime Ingestion** on Framework using **Flume**
* Designed Hadoop Stage history Maintenance on **HIVE for Analytical DataModel.**
* Designed Framework to **Parse QlikView log** files and provided user layer on top of **IMPALA**
* Fine tune and optimized **HIVE and Impala performance**
* **Automated Loading of Transactional / Non-Transactional** table Data into HIVE from Teradata using SQOOP by reducing time latency of data loading and increased performance by handling multiple scenarios
* Defined & Automated **WEB HDFS and HDFS JAVA API** Framework using **Spring Java.**
* Designed **Java xml VTD Parser Framework** to increase the performance and reduce time latency using **SpringBoot**
* Experience in **Spark, Scala and HDFS** to manage Extraction, Transformation, Load by the action of Cleansing, Validation, Transformation and handling for SCD types.
* Designed **End-to-End Airflow Scheduling and Monitoring** Tool Server Setup, Development and **Integrated with Talend**.
* Configured the **Airflow cluster using MariaDB** as a MetaStore and **RabbitMQ** for the Scheduler Messaging on Top of 3 Clusters.
* Ingested Logs flow from Talend, Spark, Hadoop and AirFlow to **ELK**(LogStash, ElasticSearch and Kibana)
* Expertise in **Database Design**, Development and Management of Schemas, automation using **Stored Procedures, Functions, DDL, DML SQL queries.**
* Proficient in **RDBMS concepts with Hadoop Hive, Impala, Teradata, DB2.**
* Expertise in Data Warehousing & Data Integration tools and Extensive experience **in ETL (Informatica, Talend and Shell Scripting).**
* Provided Solutions, Designed and Automated **Datawarehousing SCD Types for Hive.**
* Worked on **Batch and Realtime ingestion** with various Data Source and Data platforms using **IBM** **Tivoli, Flume, JMS, Kafka and Shell Scripts.**
* Extraordinary **POC** onFinding sensitive attributes based on column level and row level analysis and build predictive model using **R.**
* POC on **NoSQL Databases** like **HBase, Accumulo, Casandra and Elastic Search.**
* Integrated Hadoop technologies as Hive, Sqoop, Linux & Unix Scripting, Java, Custom MapReducer files deployment **via Continuous Integration/Continuous Delivery CI/CD Method using Git, Bitbucket, Jenkins, Docker, Nexus, CI/RA and Artifactory**

**Professional Experience:-**

* Working as a “**Senior Data Engineer**” in Accision Labs, Singapore for the Client ‘**UOB**’ for the “**Power to Predict**” from Jun 2018 till date.
* Played role as a “**Junior Big Data Architect**” in **Optimum Technologies** **(P) Ltd**, **Singapore** for the Client **‘DBS’** for the **Business Intelligence Platform** from Sep 2016 to May 2018.
* Played role as a “**Big Data Developer**” in **Optimum Technologies** **(P) Ltd**, **Singapore** for the Client **‘DBS’** for the project **of ACS and Digital Banking** from Jul 2015 to Aug 2016.
* Played role as a “**Software Engineer**” in **Skava Systems**, Coimbatore from May 2014 to May 2015.

**Certifications:-**

* Course Completion Certification on **“Cloudera Spark and Hadoop Developer” at Cloudera on Demand Training.**
* Course Completion Certification on **“Scala” at Cloudera on Demand Training.**

**Educational Profile: -**

* **Master of Engineering Computer Science (M.E)** in Sri Krishna College of Engineering and Technology, Coimbatore during 2012 till 2014.
* **Bachelor of Information Technology (B. Tech)** in SNS College of Engineering, Coimbatore during 2008 till 2012.

**Technical Skills:-**

**Cloudera Ecosystems**:Cloudera Hadoop, Spark with Scala, HDFS, Hive, Impala, Zookeeper, Hue, Sqoop, Oozie.

**NoSQL Database**: HBase, Accumulo, Cassandra, Elastic Search

**Databases**:Hive, Impala, Teradata, Hadoop, DB2, MariaDb, Oracle

**Languages**:Core Java, Scala, Python, Spring Boot, JScript’s, Shell Scripts

**Elastic Logging&Search**:ELK and Solar Index

**Real-time Ingestion**: Flume, kafka, Custom JAVA Messaging Queue

**Reporting & Monitoring**: R, Qlikview, Kibana, AirFlow, IBM Tivoli

**ETL Tools**: Informatica, Talend, Informatica BDM

**DevOps Tools**: Jenkins, GitHUB, Artifactory, CI/RA, JIRA

I hereby assure you that the information furnished above by me is true to the best of my knowledge.