**SAMENA MRUDULA  
mrudulasamena7@gmail.com  
+91- 7032529286**

**Synopsis**

4.5 years of work experience in IT Industry. Handled Big Data and Data Warehousing projects with extensive with Bigdata like **Apache Hadoop, Apache Spark, Apache Hive, Scala, Python, Pyspark, Sqoop, AWS** Good Team player as well able to priorities own workload & work under own initiative. Having good leaning attitude towards versatile technology. Adaptable to change.

**Professional Summary**

* Good personal computer and business solutions software skills.
* Hands on Experience on HDFS, Pyspark, Spark, Hive, Python, Experience with all flavor of Hadoop distributions, including Cloudera and Elastic MapReduce (EMR).
* Experience in importing and exporting data using Apache Sqoop from Relational Data Base Systems to HDFS and vice-versa.
* Expertise in developing production ready Spark applications utilizing **Spark-Core, Data frames, Spark-SQL and Spark-Streaming API's.**
* Strong experience working with both batch processing using Spark framework.
* Experience in handling different file formats **(CSV, JSON, ORC and PARQUET)** in Spark.
* Experience in writing HiveQL queries to store processed data into Hive tables for analysis. Having good exposer on Spark (RDD) and its working functionality.
* Strong Knowledge on Spark architecture and its integrations like Spark SQL, Data Frames and Datasets API.

**Skill Set**

|  |  |
| --- | --- |
| Operating Systems | Windows XP/NT, Unix |
| Languages | Hive, SQL, Scala, Python |
| Storage System | HDFS |
| Tools | Spark, Sqoop, Hive, Scala, Python, SQL |

**Employment Summary**

* Worked with **OASIS INFOTECH LIMITED**, Hyderabad from **Feb-2018 to July-2022**.

Qualification

* BTech - **Vignan’s Institute of Engineering for Women** – 2016

**Project 1:**

**Client :** AT&T, Atlanta

**Role :** Data Engineer

**Methodology :** Agile (Scrum)

**Environment :** AWS,Spark, Python, Pyspark,Hive,MySQL.

**Duration** **:** March 2020 to July 2022

**Description:**

AT&T is a leader in telecommunication services, including cell phones, wireless, U-verse, digital TV, high speed internet, DSL, home phone and many other services. ABS sales portal is a single point for managers to submit requests for system access and responsibility assignment for new hires and job changes. ABS sales portal contains over 100 ports let, which are used by thousands of AT&T employees.

**Responsibilities:**

* Responsible to manage data coming from different sources and involved in HDFS maintenance and loading of structured and unstructured data.
* Performed data migration from on premises oracle to aws cloud.
* Importing the huge amounts of retail data into HDFS from various sources and processing it CLOUDERA.
* Triggering the lambda function to collect the information.used glue jobs for applying transformations.
* Used spark transformations while moving the data to different layers.
* Worked on different file formats like Text file, Sequence file, Avro, Json, Parquet, Orc, Csv.
* Created Partitions, Buckets based on State to further process using Bucket based Hive joins.
* Optimizing the Hive queries using Partitioning and Bucketing techniques, for controlling the data distribution.
* Created Partitions, Buckets based on State to further process using Bucket based Hive joins.
* Storing and loading the data from HDFS to AWS S3 and backing up and Created tables in AWS cluster with S3 storage.
* Developing Spark Programs using Python APIs to compare the performance of Spark with Hive and SQL

**Project 2:**

**Client :** IKEA

**Role :** SQL Developer

**Methodology :** Agile (Scrum)

**Environment :** MySQL

**Duration** : March 2018 to Feb 2020

**Description:**

IKEA is a Swedish-founded multinational group, headquartered in the Netherlands, that designs and sells ready-to-assemble furniture, kitchen appliances and home accessories. The company uses approximately 1% of the Earth's Wood supply, making it one of the largest users of wood in the retail sector. It is the world's largest furniture retailer with 300+ stores worldwide. Operating in 41 countries.

We used to analyze historical sales information and generate reports on monthly basis and compare sales between months, compare sales of different product categories, analyze growth and decline in sales, generating weekly, monthly and yearly reports for the same. The Final reports we are providing to BI team for sales forecasting purpose.

**Responsibilities:**

* Strong proficiency in MySQL database management
* Decent experience with recent versions of MySQL
* Experience with replication configuration in MySQL
* Knowledge of de-facto standards and best practices in MySQL
* Proficient in writing and optimizing SQL statements
* Knowledge of MySQL features, such as its event scheduler
* Ability to plan resource requirements from high level specifications
* Knowledge of limitations in MySQL and their workarounds in contrast to other popular relational databases.

I assure that the above information is true to the best of my knowledge.

Mrudula