RANJITHKUMAR B
Technology Specialist - Bigdata, AWS
Mail id: ranjith.baskaran05@gmail.com
D.O.B: 2nd May 1996

Mobile: +91 8754565484
linkedin.com/in/ranjithkumar-baskaran-301b0a231
1/439, Pillaiyar kovil street,
Kelambakkam, Ch-603103

CAREER OBJECTIVE

Passionate and Innovation oriented **Data Engineer** with 3 years of hands-on experience. Currently seeking a challenging position in a reputed technical organization where I can enhance my skills and strength in conjunction with the institution's goal and achieve excellence through experience.

WORK EXPERIENCE

Cognizant Technology Solutions   **December 2018 – Present**
Chennai, India
Associate**,**

PROFESSIONAL SUMMARY

➢ Design, Build and Deploy end to end Data pipelines for large scale data centric applications using distributed compute platforms with Spark and Hadoop frameworks.

➢ Hands on Experience in Bigdata technologies mainly in **Hive, Spark and Python.**

➢ Good understanding of Hadoop Ecosystem and its major components.

➢ Experience in working with data load from **HDFS** and **Linux** file system to **hive** tables

➢ Developed end to end Hive queries to parse the raw data, populated external & internal tables and store the refined data in partitioned external tables. Handled variety of data by using collection items like Struct, Map and querying from hive tables.

➢ Improvising the tuning options using HIVE functions such as **Partitioning** and **Bucketing**

➢ Experience in working with the Different file formats like **TEXTFILE, SEQUENCEFILE, ORC, Parquet and JSON**.

➢ Experience with **SQL** queries, creating the reports based on the requirement, filtering, grouping and modifying the data.

➢ Implemented solutions using **Apache Spark Core, SQL and Streaming**.

➢ Worked on creating **RDDs, DFs** for the required input data and performed the data transformations, actions.

➢ Having a very good knowledge on **AWS EC2, IAM, REDSHIFT and S3 Bucket**. Have created S3 bucket to push the data from HDFS to AWS S3 Bucket through Spark program

➢ Worked as an **individual contributor** for an end-to-end application delivery.

➢ Excellent communication, inter-personal skills, analytical skills and strong ability to perform as part of a team

EDUCATIONAL QUALIFICATION

LEVEL	BOARD OF STUDY/ UNIVERSITY	INSTITUTION	SCORE	YEAR OF PASSING
Under Graduation (B.E)	Anna University	Dhanalakshmi Srinivasan College Of Engineering and Technology	7.01	2017
Higher Secondary	State Board	St joseph Higher Secondary School	80	2013
SSLC	State Board	Government Higher Secondary School	79.2	2011

- ✓ Bigdata Ecosystem   : Hadoop (MapReduce, Hive), Spark (Batch Streaming, Spark SQL), YARN
- ✓ Cloud Platforms   : AWS (EC2, S3, IAM, Redshift, SNS, Lambda)
- ✓ Language   : Python, PYSpark
- ✓ Database   : MySQL, Oracle, Teradata
- ✓ Scripting   : Shell/Unix
- ✓ Development tools   : WinSCP, Putty
- ✓ Scheduling tool   : Autosys
- ✓ Bug Tracking tool   : Jira
- ✓ Operating system   : Windows, Linux

PROJECT EXPERIENCE

Project #1: Title: AEI Pharmacist Productivity
Client: ABBVIE

Roles and Responsibilities

- Gathered the business requirements from the Business Partners and Subject Matter Experts
- Developed **Hive** scripts to load create the table and load the files.
- Developed **Spark SQL** code in Spark as per client requirements.
- Experience in managing and reviewing **Hadoop log files**.
- Experience writing unit tests for Business requirements.
- Strong analytical, troubleshooting, and problem-solving skills
- Developed spark data frame to loading the **hive** aggregate table to **Aws redshift**
- Developed various **shell scripts** for running the spark submit to triggering the jar file
- Created jill file and scheduled the jobs in **Autosys** and monitored the result

Tools and Technologies:

- Big Data Ecosystems, Spark/PySpark, Hive, Unix, AWS REDSHIFT, Autosys

Project #2: Title: RPA bot command Centre
Client: ABBVIE

Roles and Responsibilities

- Worked closely with business customers for Requirement gatherings
- Bulk importing of data from various data sources like **Aws S3 and Oracle** into **HDFS** using **Spark JDBC** and **Boto API**
- Using HDFS for persisting the data for **primary staging**
- Creating **Hive** tables with appropriate **partitions** and mapping with ingested data.
- Developed **Spark SQL** to process the data and performing transformation and loading into the **hive table**
- Good Knowledge in writing different kinds of tests like **Unit test cases** In SQL and Hive
- Developed **Autosys jill file** for scheduling to run the **spark submit** command
- Coordinate with onsite business analyst and application developers, attend client meetings.

Tools and Technologies:

Big Data Ecosystems, Spark/PySpark, Hive, Oracle, Unix, AWS S3, Autosys