SHIVAM GARG

Hadoop Developer

📞 **+91 9911223678**

✉ **shivamgarg24572@gmail.com**

OBJECTIVE

Seeking for a Hadoop Developer position to utilize expertise in Big Data, Pyspark, HiveQL, Sqoop.

TECHNICAL SKILLS

- ❖ Knowledge of Hadoop ecosystem.
- ❖ Worked on Hadoop components like Pyspark, Hive, Sqoop, Oozie
- ❖ Experience of working on Linux
- ❖ Language: SQL, Python, Shell scripting

CERTIFICATES AND TRAININGS

- ❖ Cloudera certified Spark and Hadoop Developer **(CCA-175)**, ID: **100-024-658**
- ❖ Global Agile Developer Certification (Infosys)
- ❖ ITIL 2011 Foundation Certificate in IT Service Management

PROFESSIONAL EXPERIENCE

Infosys (June 2020 – Present)

- ▪ Hadoop Development –
  - o Utilized Pyspark to distribute data processing on large dataset to improve ingestion and processing speed of data by 80%.
  - o Designed and implemented data pipelines to process data by integrating 10+ million row records from different data sources using Hive and pyspark and stored the result in HDFS.
  - o Worked on Python using Pandas, numpy libraries to analyze data.
  - o Created notebooks to insert data to Oracle and populate the tables.

United health group (July 2017 – April 2020)

- ▪ Hadoop Development –
  - o Analyzed datasets using HiveQL to recommend business improvements.
  - o Developed Spark scripts by using Pyspark shell commands as per the requirement.
  - o Optimizing of existing algorithms in Hadoop using Spark Context, Data Frames and Pair RDD's.
  - o Import data from MySQL to HDFS, using Sqoop to load data into hive tables.
  - o Worked on Migrating jobs from other platforms like oracle, MySql to Hadoop platform for faster execution of jobs and save cost.
  - o ETL, Data cleansing and transformation of data using sqoop and Hive for managing data from disparate sources.
  - o Transforming the data from datalake into meaningful form as per the business requirement.
  - o Used Hive data warehouse to store structured and filtered data into HDFS.

- ▪ Hadoop Admin -
  - o Managing multiple hadoop clusters formed with several number of servers.
  - o Installing Mapr, Hortonworks clusters comprising several number of servers to be used by the developers to run their jobs.

- o Hadoop Cluster capacity planning, performance tuning, cluster Monitoring, Troubleshooting.
- o Configuring different hadoop tools like Hive, Hbase, pig, oozie, spark etc. for different users.
- o Experience in monitoring and troubleshooting issues with Linux memory, CPU, OS, storage and network.

PROJECTS

- ❖ Hadoop related projects
  - Worked on ETL projects and built data pipelines using sqoop, hive and Pyspark.
  - Designed oozie workflows which comprised of sqoop and hive jobs.
  - Volume provisioning and queue creation automation using shell scripting.
  - Other automations to fix and self-heal different issues in the cluster using Python.

- ❖ College projects -
  - Bluetooth controlled fire Extinguisher robot using "Arduino microcontroller"
  - Automatic temperature sensor and cut off mechanism"

ACADEMIC QUALIFICATION

- ❖ **B.Tech** (Electronics and Communication Engineering) – NIET Gr.Noida (UPTU) - 68.22%.
- ❖ **Intermediate-** CBSE- 74.2%
- ❖ **High school-** CBSE-8.2(CGPA)

STRENGTHS	HOBBIES
• Good presentation skills	• Music, Movies
• Good listening skills	• Swimming
• Keen to learn new things	• Playing badminton
• Strong passion of new technology	• Reading news

PERSONAL DETAILS

- ❖ **Address** :- **A-265**, Beta-1st, Greater Noida
- ❖ **Date of Birth** :- 02 / JAN / 1995
- ❖ **Language proficiency** :- English, Hindi
- ❖ **Marital status**   :- Unmarried