Raj Kumar Bhardwaj

Email – raaj.nghani@gmail.com
Contact No. : +91-990-784-9535
Linkedin : linkedin.com/in/raj-kumar-bhardwaj-212364188

OBJECTIVE

To join a company that offers me a stable and positive atmosphere and inspires me to enhance my skills and utilize my knowledge and experience for the growth of the organization.

PROFESSIONAL SUMMARY
Having 2 years of experience working in **Hadoop Ecosystem,** design/developing applications.   - Domain experience includes **Education and Supply and Chain Management**   - Exposure in design and development of solutions for Big Data using the Hadoop eco system technologies (**HDFS, Hive, Sqoop, Apache Spark, AWS)**   - Handling **External** and **Managed** Table in **Hive.**   - Capable of processing large sets of **Structured** and **Semi-structured.**   - Brought in simplification process and **Optimization** initiatives to bring efficiency into applications.   - Having good practice on Data ingestion tools like **sqoop**.   **Sqoop** imports from MySql to HDFS and exports data from HDFS to MySql.   - Hands On knowledge of **Sqoop** performance tuning.   - Controlling imports from MySql.   - Capable of handling Sqoop with **incremental**, and **password files**.   **Hive** table management.   **External** and **Managed** tables in Hive.   - Manned versatile roles across diverse applications as **Data Engineer, Developer**..   - Able to assess business rules, collaborate with stakeholders and perform source-to-target data mapping, design and review   - Working on AWS Components like **S3, EMR, EC2.**   - Strong database experience in **MS SQL Server**.   - Handling of **RDD** and **Dataframes**.   - Read different file formats like **CSV, Parquet, ORC, AVRO, JSON, XML.**   - Having good knowledge for **DSL operations** in datasets through **Spark**.   - Read **Complex Data** import from AWS S3, HDFS.   - Import json & xml data from webapi and make it readable format by flattening it.   - Flatten the **Complex data** and make it readable format.   - Processing of **Complex Array** and save as Complex array in HDFS.   - Hands On knowledge of **PySpark.**   - Having handson experience in deploying **spark jobs** over **EMR** cluster as a **step execution**   - Have good problem solving and analytical skills and love to innovate in order to perform better.   - Have strong Interpersonal skills, communication skills and people skills to manage a team.

Company Name: Mass Express Cargo Pvt Ltd. February, 2022 - till date
Project: Data Migration to HDFS

Data coming from different sources like web server, clients, vehicle gps to mysql server, import this data to hdfs directory and process. After processing it is put to the downstreams like Hive and s3 location. Create and manage data pipeline from upstream (database) to s3 location for business analytics.

Project Role: Developer

- Performed **Import and Export** of data into **HDFS and Hive** using **Sqoop** and managed data within the environment.
- Involved in **creating Hive tables**, **data loading** and writing **hive queries**.
- Was responsible for **Optimizing Hive** queries that helped in saving Cost to the project.
- Handled **Hadoop Map Reduce jobs** to process large data sets.
- **Managed Hive Tables** and created child tables based on partitions.
- Created Workflow to **automate data loading** into the Hadoop Distributed File System.
- And **Spark** to pre-process the data and to apply complex transformation
- Involved in working on the Data Analysis, Data Quality and data profiling for handling the business that helped the Business team.
- **Loaded and transformed** large sets of semi structured data.
- **Code** & **peer review** of assigned tasks. Unit testing and Bug fixing.
- Have done POCs using **Apache Spark sql**

Technologies: Hadoop, HDFS, Hive, Sqoop, Map Reduce, **Spark sql**, Cloudera CDH, AWS s3.

Company Name: Yes India Logistics. *October 2020– February, 2022*
Project: Data integration from SqlServer to Hive and s3 RDS.

Data coming from different sources like web server, clients, vehicle gps to mysql server, import this data to hdfs directory and process. After processing it is put to the downstreams like Hive and s3 location.

Problem: MySql server is not capable of processing and storing all datas in the database.

Solution: We need to import all datas to HDFS and process it through spark and store saved data to Hive table and s3 location for downstream use.

Project Role: Hadoop Developer

- Performed **Import and Export** of data into **HDFS and Hive** using **Sqoop** and managed data within the environment.
- Was responsible for **Optimizing Hive** queries that helped in saving Cost to the project.
- **Managed Hive Tables** and created child tables based on partitions.
- Created Workflow to **automate data loading** into the Hadoop Distributed File System.
- And **Spark** to pre-process the data and to apply complex transformation
- Involved in working on the Data Analysis, Data Quality and data profiling for handling the business that helped the Business team.
- **Loaded and transformed** large sets of semi structured data.
- **Code** & **peer review** of assigned tasks. Unit testing and Bug fixing.

- Have done POCs using **Apache Spark sql**

Technologies: Hadoop, HDFS, Hive, Sqoop, Map Reduce, **Spark sql**, Cloudera CDH, AWS s3.

Project Name: Wheels Logistics *March-2018 to October-2020*
Wheels Logistics. offers Complete Set of national logistics service by roadways, railways and Airfreight including Customs Clearance in association with partners. Delivering through best distribution channel networks in India.

Project Role: Executive

Responsibilities

- answering the phone, taking messages and redirecting calls to respective offices
- organising and maintaining files and records and updating them when necessary
- creating and maintaining updated documents and spreadsheets
- overseeing the sorting and distribution of incoming mails
- preparing outgoing mail items such as envelopes or packages
- operating office equipment such as photocopier and printers
- bookkeeping and issuing invoices
- recording meeting minutes and dictations
- performing an inventory of office supplies and order of need

Technologies: Ms-Excel, MS-PowerPoint, MS-Word, Outlook

Computer Instructor (April, 2013 – Feb, 2018)
Kendriya Vidyalaya (NTPC Shaktinagar, Distt-Sonebhadra (U.P.))
Responsibilities:
- Teach computer applications to students studying in class III to Class X.
- Responsible for installation of different educational software.
- Preparation of internal and external exams.
- Invigilation in Board exams.
- Co-ordinate in Co-Curricular Activities.
- Assist in the Account department for preparation of school budget, internal audit and AG audit.
- Generating report of day-to-day activities and sent to the Ministry of HRD, Govt. of India.

Computer Instructor (Apr, 2012 – Feb, 2013)
Kendriya Vidyalaya (CWS Jayant, Distt-Singrauli (M.P.))
Responsibilities:
- Responsible for installation of different educational software.
- To teach computer applications to students studying in class VI to Class X.
- Preparation of internal and external exams.

- Invigilation in Board exams.
- Co-ordinate in Co-Curricular Activities.
- Generating report of day-to-day activities and send to Ministry of HRD, Govt. of India.

Computer Teacher (July, 2009 – Feb, 2012)
DAV Sr. Sec. Public School, Dudhichua Singrauli (M.P.)
Responsibilities:
Teach students up to Std. 10th Computer as a complementary subject.

TECHNICAL SKILLS		
Data Eco System	:	Hadoop, Sqoop, Hive, HBASE, Apache Spark, AWS
Distribution	:	Cloudera, Hortonworks
Databases	:	MS SQL Server, MySQL, Mongodb , MS-Access, OpenOffinceBase
Languages	:	**Scala, Python**, PySpark, SQL, C, C++
Operating Systems	:	Linux, Cent OS and Windows
IDE	:	Intellij, Eclipse, Netbeans

ACADEMIC QUALIFICATION			
Degree/Diploma	Board/University	Passing Year	%age
BCA	Punjab Technical University, Jalandhar	2009	79%
Inter	MP Board	2002	53%
Matric	MP Board	2000	56%

Certification
Business Analytics with Excel (Skillup in April, 2022)
Data Science with Python (Skillup in April, 2022)