Pratyush Panigrahi

Mobile: **+91-** **8297354847**

Email: **hadooppratyush@gmail.com**

# Career Objective: -

To be associated with an organization that gives me scope to update my knowledge an accordance with the latest trends in IT industry and be a part of a team that dynamically works towards the growth of organization.

# Key Skills: -

* **Hadoop Ecosystem Skills: ( HDFS, MapReduce, Pig, Hive, Sqoop, Flume, HBase, Oozie ,scala)**
* **Hadoop Distributions : ( Cloudera )**
* **Java Key Skills : ( Core Java)**

# Professional Summary: -

* Over around **3+ Years of experience in BigData(Hadoop) framework.**
* Having **2.5 years’ experience in major components in Hadoop Ecosystem including Hive, HBASE, PIG, SQOOP, Flume & Map-Reduce/HDFS framework.**
* Responsible for writing Map Reduce programs.
* Logical Implementation and interaction with HBASE.
* Perform data analysis using Hive and Pig.
* Analysing the process to load log data into HDFS using Flume.
* Support development, testing, and performance test teams during project life cycle
* Experience in using SQOOP to import data into HDFS.
* Good Knowledge on Hadoop Cluster architecture.
* Excellent Problem solving skills, detail orientation, and diagnostic abilities.
* Ability to coordinate and work successfully with a global team.
* Good Interpersonal skills and hard working to learn new technologies.

# Work Experience: - 5+ Years

* Working as **Big Data Developer** in  **Arileo Technology Noida since** since Dec 2013 to till March 2019

# Qualification: -

* **B.Tech(Electronics And Electrical Engineering )** from Centurion University, Odisha With Aggregate Percent of **65 %.**
* **12th** from Saint Paul’s School Rourkela With Aggregate Percent of **67 %.**
* **10th** from Saint Paul’s School Rourkela With Aggregate Percent of **72 %.**

# Technical Skillset: -

|  |  |
| --- | --- |
| **Languages** | Java ,Scala |
| **Framework** | Hadoop (1.x ,2.x) ,Spark |
| **Hadoop Technologies** | Map-Reduce, Hive, Pig, Sqoop, HBase and Flume, Oozie |
| J**ava IDEs** | Eclipse, IntelliJ |
| **Hadoop Distributions** | Cloudera |
| **Web Server** | Tomcat |
| **Version Control** | SVN ,BitBucket |
| **Methodologies** | Agile |
| **Databases** | Oracle10g, MySQL,DB2 |
| **Operating Systems** | Windows (10,7), Linux (Centos, Ubuntu) |

# Project Details: -

**Current Project:**

**Project Name : Auto-reconciliation(Swift Messages)**

**Client : The Northern Trust Bank Chicago(US)**

**Role : Hadoop Developer**

**Environment : Cloudera ,HDFS ,Hive , Sqoop, MySQL, Oozie ,HBase ,Tableau**

**Description:**

Have been working on data migration project

RCN auto reconciliation

Table from archival databases were transferred from the archival databses onto the HDFS

The project consisted of multiple layers

Staging Layer

Trans Layer

RDW layer

HUB layer

So we were writing hive insertion script(DML and DDL ) to transfer data from one layer to the another layer and hence storing the data in an optimized manner

The insertion script was placed in the Oozie workflow Job

Staging layer – consisted of the raw Data

Trans Layer – the data was migrated from the staging layer on to the trans layer after having optimized the data

RDW – various business logic was performed onto the trans layer data and was stored onto the RDW layer

HUB layer- the data was replicated on the Hub layer and the access was given to the Tablue team for generation of reports

**Roles and Responsibilities:**

* Involved in loading data from Staging tables to HDFS using Oozie workflow jobs
* Developed multiple Hive script to analyze the data set.
* Developed Hive Scripts to process the data and generate reports.
* Exported the patterns analysed back into RDBMS using SQOOP.
* Created sqoop job to export the data into RDBMS
* Gave KT session on Hadoop Ecosystem to team members.

**Project #2**

**Have developed scala codes to divide the file(client’s) into multiple files and storing the data in individual files**

**POC – have done a data analysis of medical report of an hospital of USA, the amount of patient admitted,reason,their details,admission year,…..the details were in the table format in CSV file… used to upload the table to hue…then import the file to the local machine shell by forming rdd then …splitting the columns we required along with its delimiter according to its requirement then anayzed it…and forming dataframes out of rdd…and then writing SQL scripts to extract the data required…..**

**Project: #3**

**Project Name : PERFORMANCE ANLYTICAL CALCULATION ENGINE**

**Client : THE NORTHERN TRUST**

**Role : Hadoop Analyst**

**Team Size : 3**

**Environment : Hortoworks ,HDFS, Pig ,Hive , Sqoop, MySQL, Oozie**

**Description:**

The data was pulled out from the Bloomberg database and was provided to the downstream team

For calculation of performance of Custodian Accont

**Roles and Responsibilities:**

* Involved in loading data from client shared location to HDFS using Oozie.
* Tuned Hive performance and maintained its performance/stability by developing repair tools to better manage hive operations in a real-world production environment
* Responsible for building scalable distributed data solutions using Hadoop.
* Developed Hive queries to process the data and generate reports.
* Exported the patterns analysed back into RDBMS using SQOOP.
* Created sqoop job to export the data into RDBMS

# Personal Credentials :-

Father’s Name : Mr. Sharat Kumar Panigrahi

Mother’s Name : Mrs. Minati Panigrahi

Hobbies : Travelling, Reading books

Marital status : Single

Languages known : English, Hindi

Strength : Optimistic, a good team player & Commitment to my work.

Permanent Address : Rourkela,Odisha

Date of Birth : 20-08-1991

I hereby declare that all the information mentioned above is true to the best of my knowledge and belief.

Place: Hyderabad Pratyush Panigrahi