

Brijesh Sharma

Delhi, India

Mobile: +91-7015345752

Email: brijesh.sharma121102@gmail.com

Date of Birth: 20 Aug 1994

Big Data Developer – IT Delivery and Customer Service

Seeking challenging assignments for a career encompassing personal & professional advancement.

PROFILE

- Presently working for **Bold sytem Tecnology Privated Ltd**, Noida as **Big data developer**
- 3 Years and 6 months of Software Development/Bug fix experience in BigData (Hadoop, Pig, Hive, sqoop,apache kafka Spark,Airflow and other Hadoop components), python
- Working as a **Big Data Developer** in **Bold system Tecnology Private Ltd.**
- Experience in **Hadoop Eco System ,Google cloud services and knowledge of AWS**
- **Confident and poised in interacting with individuals at all levels** coupled with excellent coordination and liaison skills with the Management of the Client to accomplish the project
- Highly organized to work at various levels in a project.
- **Machine Learning skills:** Machine Learning Techniques like Linear Regression, Logistic Regression, Support Vector Classifier (Type of SVM), Decision Tree (CART), K-NN, Naive Bayes, K-Means Clustering, Dimensionality Reduction, Principle Component Analysis (PCA), Content based Recommender Systems, Predictive analytics, Sentiment

CORE COMPETENCIES

🔧 BigData 🔧 Python 🔧 Hadoop Eco System 🔧 Airflow 🔧 SQL and RDBMS 🔧 Project Development 🔧 Interpersonal & Leadership Skills
🔧 Process Analysis & Enhancement 🔧 Google cloud Services 🔧 Bigquery 🔧 Google App Engine 🔧 Google pubsub 🔧 Google Dataproc 🔧
Clevertap

TECHNICAL SKILLS

Enterprise Software	:	Hadoop EcoSystem
Languages	:	Python and Scala
Framework	:	Hadoop Framework
Scripting Language	:	Shell Script, Python
IDE's	:	Eclipse ,jupyter Notebook and Pycharma
Database	:	Mysql, MongoDB, Cassandra, Hbase
Operating System	:	UNIX, Windows, MS-DOS, LINUX, Mac Os
Data Ingestion Tools	:	Sqoop and Apache Nifi
Job scheduler	:	Airflow and oozie
Visualization	:	Matplotlib, zeppelin and Redash
Architectures	:	Worked on microservices and data lake architecture
Deployment	:	Docker and Google App Engine
Cloud managed service	:	Big Query, Pubsub, Google app engine
Certification	:	Coursera machine learning

WORK EXPERIENCE

COMPANY	DESIGNATION	DURATION
Bold System Technology private ltd	Big data developer	July 2019 to till date

Gradeup(Grade Stack private Ltd)	Data Engineer	March 2019 to June 2019
----------------------------------	---------------	-------------------------

Ethonix infotech Pvt. Ltd	Big Data Developer	May 2017 to Feb 2019
---------------------------	--------------------	----------------------

PROJECTS

Bold System Technology Pvt Ltd.

Title : Cxreport
Client : Bold,Noida
Location : Noida,India
Designation : Big data developer
Tools : spark sql ,airflow,sqoop,spark,mysql and redash

About the Project:

A cloud based data analytics platform which is used for user engagement data process This analytics platform is used for user profile reporting.This platform used by internal team. Revenue report also generated using this platform

- Involved in architecture of data platform analytics platform
- Planned and Designed the architecture data from web sent data in azure event bus service .
- Prepare data for cxreport from different data sources Hbase and Mysql
- write hive script which process the data from raw to meaning full data
- Process data visualise in redash visualisation tools
- Airflow is used for schedule cxreports job
- Replace oozie with airflow schedule All legacy job on airflow
- Monitoring spark streaming job
- Handle daily adhoc query related to data

Gradeup(Grade Stack Private Limited)

Title : Udofy
Client : Gradeup,Noida
Location : Noida,, India
Designation : Data Engineer.
Tools :google pubsub,google app engine,Bigquery,airflow and redash

About the Project:

A cloud based data analytics platform which is used for user engagement data process This analytics platform is used for user profile reporting

- Involved in architecture of data platform analytics platform
- Planned and Designed the architecture data from android app and web sent data in Google Pubsub topic .
- Develop subscriber which sink data from pubsub to Big query
- write batch script which process the data from raw to meaning full data
- Process data visualise in redash visualisation tools
- Upload data on clevertap analytics tools for user notification track
- Airflow is used for managed scripts jobs

Ethonix Infotech Private Limited

Title : Datalake for healthcare data
Client : Sir H.N HOSPITAL, Mumbai and Paras Hospital, gurugram
Location : Noida, India
Designation : Big Data Developer.
Tools : jupyter,zepplien,spark,confluent kafka,spark and Hive

About the Project:

A Real-Time Streaming solution to integrate data from different data Source and store on HDFS and then process that data for the machine learning model and reporting. The solution was designed to deliver the Doctor Recommendation and Chest X ray classification

- Designed the end-to-end architecture on Pivotal stack to handle the real-time streaming and batch processing.
- Apache Kafka and Hadoop Ecosystem was used to collect the data in real time streams or scheduled batch loads.
- Data from all sources was stored and processed in HDFS.
- Real time/batch data analysis and analytics was handled by apache Kafka , spark
- zeppelin is used for data visualisation and reporting
- An in-house Android App was developed to deliver real-time alerts (notifications) to the mobile devices using Google Cloud Messaging.

Ethonix Infotech privited Limited

Title : **Fraud detection in WCPF**
Client : Zurich, New York US
Location : Noida, India
Designation : Big Data Developer

About the Project:

Product development – fraud detection in WCPF :- product developed for the insurance company in which we detect the fraud on the basis of some rules for the US client and deals with the healthcare data

- Hadoop Ecosystem components Installation.
- Nifi is used for raw data ingestion from mysql into hive tables
- Creation of hive table and loading the data into hive tables
- curate the data by using spark
- Insurance Fraud Detection – Finding out fraud using K nearest neighbor algorithm and svm
- Building Docker Images of big data tools for deploymet
- worked on google cloud platform and also on google dataproc cluster
- hbase and hive integration for data analysis
- spark and cassandra integration for data analysis

EDUCATIONAL QUALIFICATIONS

- **B.Tech (computer science and Engineering)** from Galgotias University ,Greater Noida , Uttar Pradesh, 2017 with **CGPA 6.5**
- **Higher Secondary** from Bhawna senior secondory School, palwal, Haryana, HBSE Board, 2013 with **74.5%**
- **High School** from BPS Public School, Palwal, Haryana, CBSE Board, 2010 with **68%**

ACHIEVEMENTS

- Received **Applause Award** for a client-facing project in **Ethonix infotech Pvt. Ltd.**
- Active member of the Interview panel representing **Ethonix infotech Pvt. Ltd.**
- **Won the best player award on sports day in Ethonix infotech Pvt. Ltd**

PERSONAL DETAILS

Date of Birth : 20 aug 1994
Nationality : Indian
Marital Status : Single
Passport Status : Valid
Languages Known : English and Hindi

