

Mrugesh Patel

Jersey City, New Jersey

Hadoop Data Engineer at Deutsche Bank

Summary

- # Extensive experience in Requirements gathering, Impact Analysis, High-level design, Programming, testing, JOB scheduling and implementation of Big Data Applications using Hadoop, AWS
 - # Involved in designing, building and establishing Enterprise Big data solutions for huge amount of data into Datalake to preserve history and valuable data of decades for long time
 - # Experience with different data formats like Json, Avro, parquet
 - # Good knowledge of Data warehousing concepts and ETL processes
 - # Hands on experience on AWS cloud services (EC2, S3, RDS, EMR, Lambda, API Gateway)
 - # Proficient in Hbase, Hive, Sqoop, Spark, AWS, streamsets, strong understanding of HDFS architecture
 - # Good knowledge on various scripting languages like Linux/UNIX shell-scripting and Python
 - # Good understanding of Waterfall and Agile methodologies
 - # Strong team player, ability to work independently and in a team as well, ability to adapt to a rapidly changing environment, commitment towards learning, Possess excellent communication, project management, documentation, interpersonal skills
-

Experience

Hadoop Data Engineer at Deutsche Bank

October 2018 - Present (1 year)

Big Data Consultant at Accenture

February 2018 - September 2018 (8 months)

- # Data migration from on premise database to cloud based data store over AWS
- # Analyzed data which need to be loaded into Hadoop MapR and contacted with respective source teams to get the table information and connection details
- # Used Sqoop to import data from different RDBMS systems like Oracle, DB2 and MySQL and loaded into HDFS
- # Created Hive tables and partitioned data for better performance
- # Used Spark RDD's to perform transformation on datasets as well as to perform actions like count, reduce, first.
- # Used AWS services like EC2 and S3 for small data sets
- # Used Cloud watch logs to move application logs to S3 and create alarms based on a few exceptions raised by applications

- # Developed shell scripts to move files (received through SFTP) from landing zone server to HDFS, update the file tracker and send mails after the execution is complete
- # Worked in Agile development environment having KANBAN methodology. Participated in daily scrum and other design related meetings

Big Data Developer at UnitedHealth Group

February 2016 - January 2018 (2 years)

- SME on Data ingestion in to Hadoop projects.
- Extensive knowledge of ETL tools like Streamset, Talend and Tivoli.
- Created Data Pipelines using streamsets to move data from traditional databases to Hadoop Platform
- Hands on experience of setting up ACE security on Hadoop cluster for business users
- Developed reports using Hive, Python and Shell-script as per client's business requirement
- Ingested data from one tenet to other tenet using hive, sqoop, pig.
- Implemented Partitioning, Bucketing in Hive for better organization of the data.
- Used Sqoop efficiently to bring RDBMS sources like mysql, oracle and DB2 to Hadoop.
- Used shell script to load the incremental data to hadoop.
- Developed pig scripts to load the large volume of data into Hbase.
- Developed ETL components to load log files into Elastic Search and send the JSON files to the customer.
- Designed data flow architecture for ingestion projects according to the requirement.
- Moved data from hadoop to Elastic Search using ES hadoop connector.
- Developed spark application to invoke the spark property file and load data into Elastic Search.
- Loaded data from Hadoop to Elastic Search using Hive context in Spark.
- Optimising the spark jobs by setting memory and executor configurations.
- Experience in performance tuning pig and hive jobs.
- Extensively used Sense to create indexes and mapping.
- Experience in administering Kibana and sense services that are installed on top of Elastic Search.
- Monitored Elastic Search Cluster using Kibana and sense.
- All the projects are developed and deployed under Agile methodologies with everyday standups.
- Used codehub to checkin and checkout the production code and automated deployments with Jenkins.

Rating System Analyst / Developer at ISO

July 2015 - January 2016 (7 months)

- Develop and maintain code in C# using .NET framework.

- Understanding of the Insurance Industry, especially within the General Liability & Commercial Umbrella.
- Analyze the Insurance Circulars and Implement changes to the Rating System Software.
- Conduct the daily migration of development on Subversion (SVN) and SharePoint.
- Analyze actuarial changes and their impact to the Rating System Software.

Technical Analyst / Developer at Provenir

January 2014 - June 2015 (1 year 6 months)

- Work as part of a team providing support to customer business units and the technology community.
- Produce functional specification documentation based upon client business requirements.
- Analyze, design, configure software and coordinate with clients and assist with project leadership.
- Perform and participate in design and code reviews and rigorous testing of solutions prior to delivery.
- Managing back office configuration and testing.
- Supporting client testing and implementations.
- Understanding of the Financial Services Industry, especially within the credit lifecycle (acquisition, origination, account management, portfolio management, collections and/or recovery).

Device Test Engineer at Tech Mahindra

September 2013 - December 2013 (4 months)

- Support Carrier Validation Team to test 2G/3G/4G (LTE) mobile handsets to verify Carrier's Requirement.
- Experience with wireless communication test sets (Anite network simulators).

Education

New Jersey Institute of Technology

Masters of Science (MSc), Computer Science, 2014 - 2015

Activities and Societies: Volley Ball, Badminton, Cricket

New Jersey Institute of Technology

Bachelors of Science (BSc), Computer Science, 2010 - 2014

Mrugesh Patel

Jersey City, New Jersey

Hadoop Data Engineer at Deutsche Bank

Linked  **Recruiter**