Vishal Mishra

Senior Big Data Engineer Data Analyst Machine Leaning Enthusiast Data Science

Singapore

- Email me on Indeed: http://www.indeed.com/r/Vishal-Mishra/882fcceeff7ea831

To continue the career in Big Data domain and to excel, experience and enjoy its flavor.

Work Experience

Senior Data Engineer

Visa Worldwide

February 2019 to Present

Account Level Management (Reporting Module Analytics)

ALM (Credit Card Product) is a project comes under Loyalty programme. The idea behind this project is to migrate older Ab-initio ETL framework to open source technologies like Hadoop/spark. This framework will help to generate the interchange product id which will determine the interchange fee for any credit card transaction. The objective of this module is to generate various statistics report (card level) which will help business to predict/change card holder's interchange product id.

As a Team Member, I am responsible for:

• Development of spark application jobs based on the business requirement.
• Reading data from hive using Spark Dataset API, performing business logic on the data and then writing data to MY SQL tables for the reporting purpose.

Technologies used: Spark Core, Spark Dataset API, Hadoop, Hive, Python, Oracle MY SQL

Big Data Engineer

Standard Chartered Bank - Singapore

March 2017 to January 2019

Single Customer View

The objective of this project is to create a Data Lab which will help to enable 360-degree view of the customers by joining data across systems and countries. It will also help to capture the large volume of data in a single repository.

As a Team Member, I was responsible for:

• Development of Apache SQL framework which read property files and execute the queries in the property files using Spark abstraction API's DataFrame.
• Development of property files which includes joining of multiple tables in Hive
• Development of nifi jobs for copying data from one cluster to the new HaaS cluster
• Control-M jobs to automate the process - nifi file copy and transformation job.

Technologies used: Spark Core, Spark SQL, Hadoop, Hive, HDFS, Apache Nifi, Control-M

Big Data Engineer
Standard Chartered Bank - Singapore
May 2016 to March 2017

Real time Data Ingestion Framework

The objective of this project is to develop a real time streaming application in which XML messages will be sourced from various sources and the messages will be parsed into different types of fields and then store those fields into various hive tables. The table data will be further used by the downstream application for different types of analytics.

As a Team Member, I was responsible for:

• Development of Kafka producer which transfers the xml message and place it in Kafka topic.
• Development of Spark streaming application in Scala which reads from the Kafka topic
• Development of programs to perform column and table shredding of the xml messages using Scala/ java.
• Development of programs which loads shredded data into mapped table in Hbase.
• Development of automated jobs which reads Hbase tables at the end of the day and store daily files in the hive tables
• Developments of jobs which perform various types of data validations on the daily files and store the validated result in the hive tables

Technologies used: Spark, Scala, Kafka, Zookeeper, Oozie, Hbase, Hadoop, Hive, Java

Big Data Engineer
Standard Chartered Bank
December 2015 to May 2016

Data Ingestion Framework

The objective of this project is to develop a common data lake where data from various payment sources will be sources using hadoop technologies.

As a Team Member, I was responsible for:

• Development of Map Reduce programs which will perform different types of file and data validations.

• Development of falcon/oozie workflow to integrate map reduce programs.
• Development of unix scripts to perform some specific task like new line removal
• Development of control-m jobs to schedule the data ingestion on daily basis based on the frequency of files
• Development of hive views in the common data lake in order to provide access to different teams

Technologies used: MapReduce, falcon, Zookeeper, Oozie, Hadoop, Hive, Impala, Java, control-m

Senior Software Engineer
Accenture - Bengaluru, Karnataka

June 2014 to December 2015

Revenue Optimization System

The objective of this project is to do log file analysis of the various applications which involved different competitor rates from various hotels. The huge amount of log file data will be used to determine future rates for the Starwood hotels and its brands. We have written set of multiple Map Reduce jobs to parse the log files and converted the data to csv format. We have loaded the data into HDFS and then into Hive tables where we have written Hive join queries to fetch information from multiple tables to perform various analytics.
As a Team Member, I was responsible for:

• Development and support of Map Reduce Programs those are running on the cluster.
• Configuring hadoop eco system like Sqoop, pig, flume, hive.
• Import & Export of tables/file using Sqoop
• Design and develop applications involving large data using Hadoop eco system
• Manage data like log files coming from different sources
• Load and transform large sets of structured and semi structured data.
• Loading data from UNIX file system to HDFS.

Technologies used: Hadoop, HDFS, MapReduce, Hive, Pig, Flume, Sqoop, Oozie, Zookeeper

Senior Systems Engineer
Infosys Limited - Bengaluru, Karnataka
July 2012 to May 2014

Bankline

The project was intended to credit related changes in the Bankline Screens including payment, Bulk Transfer, Manage Payments.

As a Team Member, I was responsible for:

• Development of Beans, JSPs, class files and validations flow
• Unit Testing and System testing
• Support during User Acceptance Testing

Technologies used: Java, Servlet, JSP, Struts, DB2, Websphere
Client: Royal Bank of Scotland, UK

Systems Engineer
Infosys - Bengaluru, Karnataka
July 2011 to March 2012

Telephony Transformation Programme

The aim of the project is Telephony Automation. The new telephony application estate will offer mostly the same business functionality and services as the existing system. The CICS Services will be re-engineered as reusable multi-channel services or decommissioned.

As a Team Member, I was responsible for:

- Development of Component Design and Unit Test Script
- Development of beans, JSPs and class files
- Unit Testing and System Testing
- Support during Acceptance Testing and Implementation

Technologies used: Core Java, DB2, JSPs, Websphere

System Engineer
Infosys - Bengaluru, Karnataka
January 2011 to June 2011

Rainbow-Account Segregation and APS Enquiry - is part of client initiative to re-structure their presence across globe.

This project intended towards two main functionalities:
1. Data Segregation of Moving accounts and Branches.
2. Required changes in back office Account Opening.

As a Team Member, I was responsible for:

- Development of Component Design and Unit Test Script
- Development of screen design and coding
- Unit Testing and work package testing
- Configuration Control of Project

Technologies used: COBOL, CICS, DB2, JCL MF Utilities- FILEAID, SDSF, TSO, VSAM, SPUFI

Software or O/P System: Z/OS 390, MVS, Endevor, Xpeditor, File Aid.
Client: Royal Bank of Scotland, UK

Trainings Undertaken:
- Machine Learning Bootcamp from Google.
- Java Technology and Mainframes from Infosys Limited.

Education

B. Tech in Computer Science & Engineering
Ajay Kumar Garg Engineering College Ghaziabad - Ghaziabad, Uttar Pradesh
June 2006 to July 2010

Skills

Hadoop (5 years), Hive (5 years), Apache Spark (4 years), Java (7 years), Apache Kafka (2 years), Apache Spark (4 years), Hbase (2 years), Python Programming (2 years), Scala (2 years), Cloud Computing AWS (Less than 1 year), Jenkins (1 year)

Links

Additional Information

Technical Profile:

• Having 8.5 years of professional experience in IT industry as Spark, Hadoop and Java Developer.
• Experience in Big Data Hadoop Technologies like Apache Spark, Kafka, Hbase, Map Reduce, Hive, Pig, Sqoop, Flume, Oozie, Impala & Apahce Nifi etc for 4.5 years.
• Experience in Dev ops tools like github, Jenkins and Ansible.
• Strong Programming and debugging skills in Core Java, Python.
• Experience in Agile methodology using Scrum.
• Experience of almost complete SDLC cycle (Design, Coding, System Testing, UAT Support and Implementation).
• Extensive experience of development in banking and Hospitality domain.
• Experience in leading teams and coordinating.
• Software Configuration Management (SCM) using git (bitbucket), VSS, clearcase and SVN.
• Competent, Co-operative and result driven IT Professional who can work effectively in a stand-alone as well as in a team environment.
• Strong interpersonal communication skills.
• Actively involved in knowledge sharing sessions and trainings.

My Strengths:
• Self Confidence and Positive Approach
• Adaptable to changes
• Ability to learn new Technologies
• Good Team Player