

Arihant Jain

Mobile: +91- 8586970038

E-mail: arihant.jec2013@gmail.com

- ⇒ Experience: **4 Years of experience in Predictive Analytics and Machine Learning**
- ⇒ Designated as *Manager Data Analytics*–RBL Bank Gurgaon
- ⇒ previously designated as *Consultant* – Genpact India
- ⇒ seeking a position to utilize my skills and abilities in Data Analytics & Machine Learning

Core Competencies:

- ⇒ Have hands on experience and great Interests in building Predictive Model & Machine Learning & Deep Learning Models and use Analytics & Technology to solve Problems in different business domains.
- ⇒ Have great Interests in Developing and implementing new concepts in Distributed Computing technologies, Hadoop, Apache Spark, Storm, Map-Reduce and NoSQL databases
- ⇒ Hadoop Administration and development
- ⇒ Database Technologies -- MS SQL Server/ MySQL/NoSQL/HBase/Mongo DB
- ⇒ Big Data Technologies – Hadoop / Pig Latin / Hbase/ Hive / Flume / Sqoop/ Avro / Spark / Storm/ Tableau
- ⇒ Programming/Scripting Languages – R, Python (Numpy/Scipy/Matplotlib/Scikit-Learn), Basic knowledge of Java
- ⇒ Applied different Machine Learning Techniques (Supervised, Unsupervised learning) to develop models on the Data.
- ⇒ Have Hands on experience with UNIX and Shell Scripting.
- ⇒ Deep Learning Novice & knowledge of Deep Learning Applications in Banking/Finance

CAREER PROFILE

Manager Analytics-RBL Bank Gurgaon (Oct 2016 till Date)

▪ **Role & Responsibilities :**

- Building Predictive Model for Risk Prediction using Machine Learning Technique (Random Forest/XGBoost/SVM) in R and Python
- Customer Segmentation based on Analytics and Modeling
- Statistical modeling: Regression, clustering and segmentation
- Credit Card Activation and Engagement Propensity Models
- Risk Analysis: Evaluating several risk parameters and suggesting policy changes to reduce credit risk exposure. These parameters include defaults rates, repayment ratio, cibil band analysis, program wise delinquency analysis etc.
- Worked on score development for the existing customers on the portfolio based on their behavior trends. Same is used to define non-risky customers for cross-sell activities.
- Location Intelligence Solutions
- Social Media Analytics

Consultant-Genpact India (November 2013 till September 2016)

⇒ **Client:** GE Digital – Data Science Team : –

Solving Business use case with Big Data Analytics.

- Bringing data from different sources in to Pivotal Hadoop/Greenplum Data Lake from different sources like Historian, Enterprise data ware house, corporate data sources,

Tanium.

- Showcase the PGS use case for trending the max tagged event during the turbines startup over a 3 year period.
- Application portfolio management – Create a consolidated source of data for Application portfolio, CMDB, Risk management and application dependency mapping to facilitate cloud migration.
- Ingest the data from GE assets collected via Tanium and enabling for threat monitoring, inventory management.
- Next Gen Ops – Bringing and correlating data from configuration management, incident management for enabling operations and predicting the issues.
- Ingest the weather and atmospheric chemistry data to be used by various business units for analyzing different service patterns.

▪ ***Role & Responsibilities:***

- Responsible for the development activities of the Interfaces assigned, with the team aligned and as per the requirement specifications. Create architecture and lead the execution of multiple business usecases
- Work with business stakeholders and product owners for understanding the business requirement and translate to technical requirement.
- Architect and design the solution for data ingestion and consumption using REST.
- Develop some of the use cases while helping the team in resolving their issues.
- Ingest the data using Talend, Hive into HDFS, Greenplum etc.
- Conduct code reviews.
- Build the consumption model (REST API's, Tableau reporting, hive) for the data from the data lake.
- Educate and bring the teams onto following agile development methodology.

- ***Technology used*** – Pivotal Hadoop, HDFS, Hive, Oracle, Greenplum, HAWQ, Talend, Tableau, Java, Spring, Hibernate, JSON, JIRA, Github, Maven, Jenkins , Rally, Linux, REST, Spark , Map Reduce, Machine Learning

⇒ **Machine Learning Projects: -**

- Linear /Non –Linear Optimization Solver in Python (Scipy/numpy/Mystic)
- Rapid extraction Keyword algorithm in Python.
- Entity extraction algorithm in Python.
- Time Series forecasting with Python (Pandas).
- Implementing supervised and Unsupervised Learning.
- Predictive Analytics and Modeling for Propensity, Churn Modeling

⇒ **Hands on Big Data Analytics: -**

Description: I have successfully complete BIG DATA Analytics Certification from Jigsaw Academy in collaboration with Wiley (Mar 2015 to Sept 2015) which is considered to be Top certification in the Big data analytics Industry. During this rigorous certification period of six month I have done Hands on different technologies used in Big Data Ecosystem.

- Twitter Data Analysis
- Word count analysis with Map Reduce
- Analyze Email Marketing Campaigns of Magazine Publisher
- Parse user access log files and provide details about users accessing system within given time frame
- Scripting with PYTHON and used it for data analysis

- Installing and Configuration of Hadoop, hive and all Big data Environments frameworks
- Migration of data from RDBMS to NoSQL
- Various Machine Learning Algorithms on Hadoop including Clustering, Collaborative Filtering etc.
- Integration of R ,Hadoop and Tableau
- Log processing and Analytics

Technology used during Hands on in Big Data Analytics Ecosystem

- Hadoop Storage : HDFS/HBASE/SOLR
- Distributed Processing Framework : Map Reduce , Spark
- Data exchange : Sqoop
- Log Collector : Flume
- Scripting : Pig
- SQL Query : Hive
- Workflow : Oozie
- Database : NoSQL (Hbase , Mongo DB) and RDBMS (My SQL)
- ETL Concepts and Tools (Kettle)
- Hbase , Cloudera Basics and Manager ,Horton works and Green plum pivotal HD and Impala
- Python for data analysis and Scala
- Basic and Advance R Programming
- Data Visualization Tools (Tableau)
- Machine Learning
- Python
- Deep Learning

EDUCATION AND CERTIFICATIONS

- ⇒ Completed B.E in May 2013 with 78.88 % from Jabalpur Engineering College,Jabalpur ,Madhya Pradesh
- ⇒ Scored 80+ in 10th and 12th, among top 5 students in the class
- ⇒ **Wiley JIGSAW Certified Big Data Specialist (Mar – Sep 2015)**
- ⇒ A crash course in Data Science by John Hopkins (Coursera)
- ⇒ Machine Learning certification By Stanford University (Coursera)
- ⇒ Lean Trained
- ⇒ Awarded with Bronze award two times in Genpact
- ⇒ Got Multiple Appreciation Mail from the client and Manager

TECHNICAL SKILLS

- Programming/Scripting Languages – Python (Numpy/Scipy/Matplotlib) , Java , Shell Scripting and Ant Scripting
- Rdbms and NoSQL Databases
- IDE : My Eclipse, Eclipse, NetBeans 6.5
- Technologies and tools in **Big data Ecosystem**: Hadoop, Map Reduce, HDFS, Hive, Hbase, Sqoop, Pig, MongoDB, Mahout.
- Experience in objective oriented programming using Python and database integration with MySQL
- Experience in working with Basic and Advance R Programming
- Exploratory data analysis and Data Visualization Tools (Tableau)

