

Big Data technologies:

SPARK (core/sql/streaming/Mlib)
FLUME-1.6.0(with Spark)
HADOOP
Map-Reduce(basics)
Kafka(basics)
Hive(basics)

Machine Learning Algo&Tech:

XGBOOST
Logistic Regression
Decision Tree
Pandas, NumPy
Sklearn, StatsModel
Spark-Mlib

Language:

Core-Java
Python
Scala(basics)

No-sql Databases:

Mongo, Cassandra
Druid, Hdfs

Sql-Databases:

SQL/ORACLE

Web Technologies:

REST, SOAP, HTML
Django, Flask

Basic Knowledge:

NLP, TF-IDF, Text mining
Naive Bayes, Linear Regression
Messaging Queue, Solace, IOT
AWS, Docker, Random Forest

Others: Json/Unix/XML

Build Tool: Maven/Gradle

Server: Tomcat, WebLogic

Operating System :

Windows/Linux/IOS (mac)

CAREER-HIGHLIGHTS:

- Extensive knowledge and 7+ years of professional experience in analysis, Design, Development, customizations, implementation of Product (Bigdata, Machine-Learning) in Banking & Financial, Telecom, Insurance.
- Extensive knowledge in SPARK(2.0/2.2.0)(Core/Sql/Streaming/Structured-streaming), Spark-MLib(Logistic-Regression,Decision-Tree),Spark-Mongo,Spark-cassandra,spark-flume,using JAVA/SCALA/PYTHON (depends on project architecture).
- Professional experience in Machine-Learning Algorithms like XGBOOST, Regression (Logistic Regression), Classification (Decision Tree), EDA(Data cleaning, Univariate Analysis, Bivariate Analysis),Feature Selection(RFE, PCA) ,Dummy variable creation(One hot encoding, Level encoding)and design ,implement data models, perform statistical analysis and scoring .
- Professional experience in Core-Java (Multithreading, Concurrency, Collections, Exception Handling, OOPS), Data structure, Design-pattern, Web-service (SOAP, REST).
- Experience in Python, Django template, Flask, Pandas, NumPy, Sklearn, matplotlib, StatsModels.
- Having good experience on all flavors of Big-Data Technologies like HDFS, FLUME (with Spark push and pull base model) ,Hadoop Distribution (Horton works and Cludera), Docker.
- Extensive knowledge on No Sql database like MONGO DB, CASSANDRA DB, HDFS, Druid DB, Netezza DB and Sql db. like My Sql Db, Oracle Db.
- Having basic knowledge on Naïve Bayes algo, Unsupervised learning, NLP (Text Mining, lexical Processing, Syntactic Processing, Lemmatization, TF-IDF, LDA), NLTK, Scala (for Spark), shell scripting, UNIX commands, XML, KAFKA, HIVE, Map-Reduce.
- Worked on Agile Methodology (Scrum, sprint-planning, sprint-demo, retrospective) and Software Development Life Cycle & LEAN Methodologies.

PROFESIONAL SYNOPSIS:

Machine Learning product → Machine Learning Studio (MLS).

Big-Data product → STREAMANALYTIX (Real time analysis), Clear-Trail (Entity-Profiling).

Telecom product → BSS (Business support System/Unified Billing system).

ACADEMIC RECORDS:

- B. Tech in Electronics & Communication Engineering from Haldia Institute of Technology in the year of 2011.
- Currently pursuing PGD in Machine Learning & AI from International Institute of Information Technology Bangalore (IIIT-B).

PROJECT EXPERIENCE:

#Machine Learning Studio(Domain-Banking, organization-IBM, methodology-Agile):

Technology Skills: Machine Learning (xgboost, Decision Tree, Logistic Regression), Spark 2.2.0, Java, Python, Pandas, NumPy, Sklearn, Netezza

Description:

MLS is the analytical and modeling ecosystem with an approve suite of machine-learning algorithms and analytical tools. This is the entry point for all Data scientists and Analytical users.

Roles & Responsibilities:

- ✚ Scoring model ideation and implementation and Root Cause Analysis and Predictive modelling implementations using machine learning algorithms (Xgboost) for fraud analysis.
- ✚ Solving use cases for building models like Logistic regression and Decision Tree using pandas, NumPy, sklearn, spark-mllib, statsmodel, Dask.
- ✚ Working on EDA(Data cleaning, Univariate Analysis, Bivariate Analysis),Feature Selection(RFE, PCA) ,Dummy variable creation(One hot encoding, Level encoding) for provided use case data.
- ✚ Understanding the requirements from the product team and developing these features and contributing back to the platform
- ✚ Design and development of xgboost end to end for MLS using Python, Django Template, Java (concurrency, multithreading, collections).
- ✚ Developed analytical tool like Netezza for fetching data using PySpark.
- ✚ Fetching data from Data-Lake (Cornerstone) using Pyspark-Hive connector and processed them for particular requirement .
- ✚ Developed Leatherman framework for easily deploying data model to the cluster using Docker, flask.

50 GB Data processed and tested for XGBOOST

#StreamAnalytix(<https://www.streamanalytix.com/>)(organization-Impetus,methodology-Agile):

Technology Skills: SPARK 2.0, Core Java, Cassandra, HDFS

Description:

StreamAnalytix is product of multi-engine, Open Source based platform with enterprise-grade features, flexibility, extensibility, ease-of-use, and full enterprise class support.

Roles & Responsibilities:

- ✚ Impact Analysis Cater requirements for new component of StreamAnalytix like **Mongo DB** and **Flume**.
- ✚ Implemented **Spark Streaming** (with **JAVA**) Connector allows on-the-fly analysis of live data streams, batch data, normal data using Mongo Db.
- ✚ Implemented Channel, Emitter, Lookup and sink functionality using **Spark2.0, Java** for supporting Mongo Db functionalities (**Indexing, Authentication, Sharding (Hashed and Ranged), HA**)
- ✚ Implemented **Apache Flume** (1.6.0) client using **Spark 2.2 (Pull base and Push Base approach)** using **SCALA** allows on-the-fly analysis of live data streams using Flume.
- ✚ Develop custom Receiver for pub-sub model of **Spark-Redis connector** for supporting streaming.

#Entity Profiling (organization-Impetus, methodology-Agile):

Technology Skills: Spark 2.0, Core Java, Mongo, Flume 1.6.0, Scala (Spark)

Description:

Entity Profiling is a generic product which creates a complete profile for an entity. An entity can be a person, Customer Number, Account Number etc.

Roles & Responsibilities:

- ✚ Implementation of **SPARK - Cassandra** Connector to fetch data from **Cassandra** base on key Space and table using **Scala**.
- ✚ Developed code for Token-Range base approach using **multithreading** and **Cassandra-java connector** for some exceptional tables where we can't able to query on the basis of use cases.
- ✚ Developed **Hdfs-reader, Hdfs-writer** using **Java** for solving the product requirement and creating streaming data and dumped data into **HBASE**.

2 TB data migrated from
Cassandra to HBase

#VOICE MONITORING (organization-HARMAN, Domain-Telecom, Methodology-Agile):

Technology Skills: Core Java, Mongo, Druid, Restful webservice, json, sql.

Description:

“VOICE MONITORING is responsible for capturing network elements (SIP, RTP, RTCP) from server and show on the client side

Roles & Responsibilities:

- ✚ Worked on **SIP, RTP, RTCP** protocol and MOS calculation using **Jetpcap**.
- ✚ Worked on **Core Java** basically on **Multithreading, Collections, Exception Handling, Concurrency, Spring-Core**.
- ✚ Worked on No Sql DB like **Mongo DB, Druid** (Time series DB), **Unix, Sql**.

#BSS)Business Support System (organization-HARMAN, Domain-Telecom, Methodology-Agile):

Technology Skills: Core Java (Multithreading, collections), RMI call, Sql, Unix, XML, json, SOAP.

Description:

Unified Billing has evolved into a real-time critical operation that impacts the profitability of an operator and the satisfaction of customers

Roles & Responsibilities:

- ✚ Impact Analysis cater requirements for new feature of BSS product and Development using **Core Java, sql, xml**.
- ✚ Developing API (**SOAP wsdl**) and **RMI call** for interacting with another component (ESS or DCRM).
- ✚ Worked on **Design Pattern** and **Data-Structure**.

#Vantiv Middleware (Fifth Third Bank) (organization-WIPRO, Domain-Banking, Methodology-Agile):

Technology Skills: Core Java (Multithreading, collections), RMI call, Sql, Unix, Xml, json SOAP.

Description:

The Primary Objective of the project is to provide Integration Services to Vantiv Payment Processing Solution using Integration tool

Roles & Responsibilities:

- ✚ Impact Analysis, Test plan, **HLD** documentation & Development.
- ✚ Worked on Core Java basically on **Multithreading, Collections, OOPS** concept, **Exception Handling, Sql, IBPM**.

#Offer Treatment Service(CAPITAL ONE) (organization-WIPRO,Domain-Banking-Creditcard,Methodology-Agile):

Technology Skills: Core Java (Multithreading, collections), Sql, Unix, SOAP, Blaze-Advisor.

Description:

Unified Billing has evolved into a real-time critical operation that impacts the profitability of an operator and the satisfaction of customers

Roles & Responsibilities:

- ✚ Development of business rules in Blaze Business Rule Engine IDE and deployment of these rules into production.
- ✚ Working on Core Java (Multithreading, Collections, Exception handling, OOPS) for develop rules and fixing bug.

#HOME_PCC(TH THE HARTFORD) (organization-WIPRO, Domain-INSURANCE, Methodology-Agile):

Technology Skills: Core Java (Multithreading, collections), Unix, Xml, json, Blaze-Advisor (FICO-PRODUCT).

Description:

Home Policy Change Centre (HPCC) is a new application to allow agent self-service submission for certain policy changes.

Roles & Responsibilities:

- ✚ Developing of **Code template, Ruleset template, Group template, Loop template** using **BLAZE-ADVISOR** and **JAVA**.
- ✚ Creating dummy XML for testing those Code template and rules.