**Krishna Rao  
401-210-0387**

**ajay.kumar@busitantsinc.com**

2. **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

***Professional Summary:***

* 11+ years of professional IT experience on Big data, ETL, AWS, Informatica, Mainframe Technologies and worked for Banking, Credit Card, Insurance, and Financial Domain, with the extensive experience in Coding, Unit testing, analyzing tests results, Preparing Estimation, JOB scheduling Documents and JOB Flow Documents.
* 5+ years of experience on Hadoop ecosystem i.e., HDFS, Spark, Scala, Python, Kafka, Hive, Sqoop, Pig, Oozie, RDBMS and Map Reduce.
* 3+ years of AWS experience on S3, EMR, Redshift, SQS, SNS, EBS.
* Worked on analysis of data collected from Kenesis stream.
* Used Spark Streaming API for reading the batches of data from Kenesis Stream.
* Used Spark programming concepts with Scala for collecting and analyzing the data.
* Worked on Oozie workflows for triggering a spark application, which reads the data from HDFS or S3 location for processing.
* Good working Knowledge on UNIX Commands.
* 6+ months experience on SNOWFLAKE.
* Worked on Selenium automation testing
* Worked on CI/CD pipeline.
* Developed end-to-end ETL using technologies like INFORMATICA, Unix, Python and Autosys
* Implemented Proofs of Concept on Hadoop for analysis of different types of data from different sources like oracle, Json, XML and storing to Hadoop HDFS.
* Loaded the data set in to HDFS, Hive and processed the batch data using Hive, Pig and automated the workflow using Oozie.
* Experience in using AWS cluster.
* Worked on PySpark DataFrames, RDDs to process and integrate the data from heterogeneous sources.
* Effectively worked on Amazon RDS (Oracle, MySQL, Amazon Aurora) and Non-relational Database - Amazon DynamoDB
* Performed complex querying analytics with Amazon Redshift.
* Worked on creating pipeline process to gather real time streaming data with Amazon Kinesis.
* Strong experience on conversion OS mainframes code to information Datawarehouse.
* Strong experience on documenting Low Level Design, Unit test plan, Unit test cases and Deployment documents Auto flow diagrams, Auto flow templates, UTC, UTR, LDMS and CII.
* Have Strong experience in Agile methodology.
* Involved in sprint planning, stand-ups, retrospective meetings
* Worked on Informatica 9.5 & 10x, Putty, WinSCP, IBM Datastudio.
* Certified and attained the following professional designations:
* AWS certified solution Architect associate.
* Azure fundamental certified
* ISTQB Foundation Level Certified.
* LOMA Level 1 Certified.
* Interacting with Client’s for their issues and Business requirements.
* Have performed various testing activities like Unit Testing, system testing, Regression Testing, to ensured that the code is ideal. Experienced in production support activities.
* Self-motivated and highly adaptable to the demands of the project in terms of new requirements, meeting deadlines, delivering quality output.
* Worked on Insurance, Banking Domains.

***Education:***

1. Master of Computer Application (MCA) from, Jawaharlal Nehru Institute of Technology, Hyderabad, India, 2009.

***Career Profile***

1. **Tech. Lead (Hadoop) Client: UBS Salt Lake City, UT Sept, 2021 - Till Date**

Sr. Tech. Lead -Hadoop Developer

**Description: UBS AG** is a [Swiss](https://en.wikipedia.org/wiki/Switzerland) global [financial services](https://en.wikipedia.org/wiki/Financial_services) [company](https://en.wikipedia.org/wiki/Company). The company provides [wealth management](https://en.wikipedia.org/wiki/Wealth_management), [asset management](https://en.wikipedia.org/wiki/Asset_management), and [investment banking](https://en.wikipedia.org/wiki/Investment_banking) services for private, corporate, and institutional clients worldwide, and is generally considered to be a [bulge bracket](https://en.wikipedia.org/wiki/Bulge_bracket) bank.

AG team handles the Loan processing and customer verification data base & applications

**Role: Sr. Developer**

Environment: Hadoop, Spark, Scala, Hive, HDFS, Cloudera, Unix, SQL Server 2015

**Responsibilities:**

* Meeting various stakeholders for requirement gathering and data analysis for new business insights.
* Developed the Application using Scala and Spark based on the existing functionality.
* Analysis and feasible study of requirements.
* Run Proof of Concepts to validate the feasibility of the design.
* Implemented Spark Core in Python to process data in memory.
* Performed job functions using Spark APIs in Python for real time analysis and for fast querying purposes.
* Developed Theia Pyspark compute.
* Building and deploying pypark compute in artifact.
* Scheduling and Executing computes in airflow.
* Developed SPARKOLA computes by creating inputs, outputs and pipelines.
* Involved in creating Spark applications using cache, map, reduceByKey etc. functions to process data.
* Importing and exporting data from different databases like Mainframes RDBMS, Oracle, MySQL into HDFS and Hive using Sqoop
* Written PySpark scripts to load data from different type of formats.
* All the transformation related error handling logic is written in the Python scripts.
* Written Hive queries and UDF's for few of the business requirements.
* Involved to write Shell scripts to run the PySpark script file and Error handling in the Shell script.
* Worked on PySpark DataFrames, RDDs to process and integrate the data from heterogeneous sources.
* Created Staging tables in Hive and lookup tables are created in HBase.
* Established Integration between Hive and HBase.
* Helped team understanding the architecture of the project and analysing
* Understanding and reviewing the business requirements and Mapping documents
* Importing and exporting data into HDFS and Hive using Sqoop
* Creating Hive tables and working on them using Hive QL
* Created Hive queries and scripts to implement business transformation rules.
* Created Pig scripts to process HDFS data
* Export the data into Oracle using Sqoop
* Involved in Technical feasibility discussions, design review meetings, peer reviews, Code reviews.

**2. Tech. Lead (Hadoop) Client: Confidential Sept, 2020 - Sept, 2021**

Client: Confidential

**Responsibilities:**

* Worked on analysing Hadoop stack and different big data analytic tools including Pig and Hive, HBase database and Sqoop.
* Written multiple MapReduce programs to extract data for extraction, transformation and aggregation from more than 20 sources having multiple file formats including XML, JSON, CSV &other compressed file formats.
* Implemented Spark Core in Scala to process data in memory.
* Performed job functions using Spark APIs in Scala for real time analysis and for fast querying purposes.
* Involved in creating Spark applications in Scala using cache, map, reduceByKey etc. functions to process data.
* Created Oozie workflows for Hadoop based jobs including Sqoop, Hive and Pig.
* Created Hive External tables and loaded the data in to tables and query data using HQL.
* Performed data validation on the data ingested using MapReduce by building a custom model to filter all the invalid data and cleanse the data.
* Handled the importing of data from various data sources, performed transformations using hive, Map-Reduce, loaded data into HDFS and extracted data from MySQL into HDFS using Sqoop.
* Wrote HiveQL queries by configuring number of reducers and mappers in the query needed for the output.
* Transferred data between Pig Scripts and Hive using HCatalog, transferred relational database using Sqoop.
* Configured and maintained different topologies in Storm cluster and deployed them on regular basis.
* Responsible for building scalable distributed data solutions using Hadoop. Installed and configured Hive, Pig, Oozie, and Sqoop on Hadoop cluster.
* Developed simple to complex Map-Reduce jobs using Java programming language that was implemented using Hive and Pig.
* Ran many performance tests using the Cassandra -stress tool in order to measure and improve the read and write performance of the cluster Configuring the Kafka, Storm and Hive to get and load the real time messaging.
* Supported MapReduce Programs that are running on the cluster. Cluster monitoring, maintenance and troubleshooting. Analysed the data by performing Hive queries (HiveQL) and running Pig Scripts (Pig Latin).
* Cluster coordination services through Zookeeper. Installed and configured Hive and also written Hive UDFs.
* Worked on the Analytics Infrastructure team to develop a stream filtering system on top of Apache Kafka and Storm.
* Worked on a POC on Spark and Scala parallel processing. Real streaming the data using Spark with Kafka.

**3. Hadoop Developer Client: Experian Aug, 2018 - Sept, 2020**

Experian is Credit reporting system, it gets data in different types of forms like JSON, CSV, Key-Value pairs and perform transformation and store in Hive finally. We are dealing with both Historical and Incremental data. Earlier the whole logic regarding data transformation is developed in Perl script, stores in Teradata. Now we have developed the transformations logic in PySpark and finally storing the data into Hive tables (Hadoop)..

**Responsibilities:**

* Developed the Application using Scala and Spark based on the existing functionality.
* Analysis and feasible study of requirements.
* Run Proof of Concepts to validate the feasibility of the design.
* Importing and exporting data from different databases like Mainframes RDBMS, Oracle, MySQL into HDFS and Hive using Sqoop
* Written PySpark scripts to load data from different type of formats.
* All the transformation related error handling logic is written in the Python scripts.
* Written Hive queries and UDF's for few of the business requirements.
* Involved to write Shell scripts to run the PySpark script file and Error handling in the Shell script.
* Worked on RDD's and DF in transformation process.
* Created Staging tables in Hive and lookup tables are created in HBase.
* Established Integration between Hive and HBase.
* Performed transformations on data as per the requirement.

**4. Hadoop Developer Client: NLY –TSSC Nov, 2015 - Sept, 2018**

**Technologies :** Spark, Scala, Hadoop, Hive, Java, MapR, Mainframes, COBOL, JCL

**Client :** Newyork Life Insurance, US

**Role :** I.T Analyst.

**Description:**

New York Life Insurance Company (NYLIC) is the largest mutual life-insurance company in the United States, and one of the largest life insurers in the world, ranking #80 on the 2015 Fortune 500 lis with about $512 billion in total assets under management, and more than $19 billion in surplus and AVR. Inforce team handles the insurance processing system application.

**Responsibilities:**

* Developed the Application using Scala and Spark based on the existing functionality.
* Reading the input raw file and validating the Header, Data rows and footer.
* Separating parent and child records
* Eliminating parent duplicates and count the total number of records.
* Compare the input file with existing IDM file and identify new records.
* Analysed the Hive UDF’s used for data ingestion from Oracle to Hadoop and sqooping back to Oracle.
* Resolved long standing Hadoop UDF related application issues which required a code fix related to Distributor rebate percentage.
* Pro-Actively and successfully Improved the performance of the application for calculation process, by implementing a Java utility to filter the required data even before the analysis is being done.
* Single point of Contact for Hadoop related issues in the application.
* Trained the team in Hadoop related project specific utilities to reduce the dependency on a single resource and to strengthen the team technical caliber.
* Helped team understanding the architecture of the project and analyzing.

**March 2013 - Nov 2015 HTC Global Services**

**3. Software engineer**

**Client: State Farm Insurance, Bloomington, Illinois, US.**

State farm is one of the largest insurance companies in the world. It offers a wide variety of properties, Auto, Health, life, casualty insurance, surety products and services to business, organizations and individuals in the United States and in selected international markets.

**Project-1: SAMS Maintenance.**

Sales and associate management system (SAMS) is associates and agents’ information maintenance system. It holds all the employee’s information (external and internal employee’s interns, agent's, SFO's, AFO's, SL's information). SAMS is OLTP system. Developing or modification of Programs or JCL.

**Environment:**  Lotus notes, Testlink, Control-M**, FILE**-AID, FILE-AID-DB2, COBOL, JCL, DB2, PANVALET, IBM DEBUGGER**.**

**Responsibilities:**

* Understanding the HDD, preparing Specification documents. Interacting with the client to get clarifications and acceptance. Impact Analysis of existing source mappings/scripts.
* Designed and developed complex mappings by using Lookup, Expression, Update, Sequence generator, Aggregator, Router, Stored Procedure etc., transformations to implement complex logics while coding a mapping.
* Design & develop highly efficient/high performance ETL mappings/workflows.
* Code changes in existing mappings/workflows and Unit testing.
* Performing Configuration Management by using the PANVALET.
* Preparing unit test cases and unit test results. Peer reviewing of Mappings and test results.
* Executing the queries by using IBM DB2 Develop centre tool and Pocket tool.
* Uploading test cases, test results in Testlink.
* Preparing AUTO Flow diagrams, Power exchange, AUTO flow templates, UTC's, UTR's, CII, LDMS specs.
* Uploading all the documents to share folder and lotus notes.

**Project-2: FDW strategic Support Team.**

This is a migration projectfromMainframes to Unix and Informatica. All the OLAP mainframes programs/modules are converted to Unix/Informatica scripts and mappings. It also involves developing mappings as per new business requirements. Testing, debugging the mappings /workflow/scripts. Support on Pre and post implementation.

**Environment:** IBMDatastudio, Unix, HPSM, RMS, Track tool, Lotus notes, Testlink, Control-M, PANVALET, FILE-AID, FILE-AID-DB2, BMCDB2, COBOL, JCL, DB2, IBM RDZ DEBUGGER, TSO

**Responsibilities:**

* Understanding the HDD, preparing Specification documents. Interacting with the client to get clarifications and acceptance. Impact Analysis of existing source mappings/scripts.
* Designed and developed complex mappings by using Lookup, Expression, Update, Sequence generator, Aggregator, Router, Stored Procedure etc., transformations to implement complex logics while coding a mapping.
* Design & develop highly efficient/high performance ETL mappings/workflows.
* Code changes in existing mappings/workflows and Unit testing.
* Performing Configuration Management by using the PANVALET.
* Preparing unit test cases and unit test results. Peer reviewing of Mappings and test results.
* Executing the queries by using IBM DB2 Develop centre tool and Pocket tool.
* Uploading test cases, test results in Testlink.
* Preparing AUTO Flow diagrams, Power exchange, AUTO flow templates, UTC's, UTR's, CII, LDMS specs.
* Uploading all the documents to share folder and lotus notes.

**Aug 2011 - Feb 2013 HTC Global Services**

1. **Software engineer**

**Client: MEIJER, Michigan, USA.**

Meijer is developing retail store market. It is having 250 stores in 5 regions in USA and each store contain 500 employees.

**Project: STORE LABOR (ABL-ABS-REPORTS).**

The current forecasting process at Meijer is a patch-work collection of manual processes. Meijer has chosen to replace these methods by automating with the reports and repeatable process that can be driven by business customers with little to no IT intervention. This involves forecasting & Analysis.

**Environment** COBOL, JCL, VSAM, DB2, SQL, Change men, IBM DEBUGGING, FILE-AID, FILE-AID-DB2, TSO/ISPF.

**Responsibilities:**

* Understanding the HDD and Specs.
* Impact Analysis of source programs.
* Developed the COBOL, COBOL + DB2, JCL programs.
* Analysing the existing Queries and query tuning.
* Data import and export using FILE-AID-DB2.
* Peer Review of programs and test results.
* Executing the queries by using IBM DB2 Develop centre tool and Teradata tool.
* Preparing unit test cases and unit test results.
* Performing Configuration Management by using the Change men.

**Project-2:** Pharmacy Transmission

The pharmacies are processing fills within the store EPS system for patients with insurance that is to be billed electronically. Migrate from VSAM files to DB2 tables

**Environment: COBOL, JCL, VSAM, DB2, SQL, PANVALET, IBM DEBUGGING**, **FILE-AID, FILE-AID-DB2, TSO/ISPF**

**Roles:**

1. Developer
2. Team Member

**Responsibilities:**

* Understand the HLD, work request and Specs.
* Impact Analysis of source programs.
* Developed the CICS, COBOL, DB2 Code & DB2 queries.
* Prepare unit test plan and test cases.
* Performing Configuration Management buy using the PANVALET.

**Feb, 2011 - July 2011 HTC Global Services**

* **Software engineer**

**Client: AOI, Michigan, USA.**

**Environment:** COBOL, JCL, VSAM, DB2, SQL, Changemen, IBM DEBUGGING, FILE-AID, FILE-AID-DB2, TSO/ISPF

**Project: Migration from ADABAS to DB2**

This Project is being executed for Auto-Owners Insurance Group (A-O). This is a conversion project which modernizes the existing legacy system and converts the calls/commands related to ADABAS file system to equivalent DB2 SQL statements (Z/OS).

**Roles:**

1. Developer
2. Team Member

**Responsibilities:**

* Understanding the HDD and Specs.
* Impact Analysis of source programs.
* Developed the COBOL, COBOL + DB2, JCL programs.
* Analysing the existing Queries and query tuning.
* Data import and export using FILE-AID-DB2.
* Peer Review of programs and test results.
* Executing the queries by using IBM DB2 Develop centre tool and Teradata tool.
* Preparing unit test cases and unit test results.
* Performing Configuration Management by using the Changeset