

M.PRADEEP KUMAR

Phone: (+91)-7502505750
pradeepkumar1998@gmail.com

PROFESSIONAL SUMMARY

- ❖ Having 4.2 years of IT experience and 2.3 years of experience working in **Big data Ecosystem** such as **Spark and Hadoop**.
- ❖ Exposure in design and development of solutions for Big Data using the Hadoop Eco-system technologies (**HDFS, Hive, Sqoop, Apache Spark**)
- ❖ Worked extensively on Hadoop migration project and POCs.
- ❖ Capable of processing large sets of **Structured** and **Semi-structured** Hands on experience on Data ingestion tools.
- ❖ Able to assess business rules, collaborate with stakeholders and perform **source-to-target** data mapping, design and review.
- ❖ Replaced HQL scripts with Spark scripts with various dataframe operations
- ❖ Strong data base experience in **MS SQL Server**.
- ❖ Experience in transferring data from RDBMS to HDFS and HIVE table using **SQOOP**.
- ❖ Worked in **Agile Methodology**.
- ❖ Performed **AWS** reads and writes using Spark Scala.
- ❖ Performed Data Enrichment, cleansing and common data aggregations through RDD transformations
- ❖ Deep knowledge in incremental imports, **partitioning and bucketing** concepts in **Hive** and **Spark SQL** needed for optimization.
- ❖ Experience collecting log data from various sources and integrating it into **HDFS** using Flume; **staging data in HDFS for further analysis**.
- ❖ Have good problem solving and analytical skills and love to innovate in order to perform better.
- ❖ Have strong Interpersonal skills, communication skills and people skills to manage a team.

TECHNICAL SKILLS

Data Eco System	:	Hadoop, Sqoop, Hive, Apache Spark,
Cloud Skills	:	AWS
Databases	:	MS SQL Server, MySQL
Languages	:	SQL, UNIX Shell Script, Spark Scala
Operating Systems	:	Linux and Windows

PROFESSIONAL EXPERIENCE

Project Name: Montra

Nov 2021 – Present

Montra is the flagship performance brand in TI Cycles' Track and Trail range, created to bring an international standard of performance to Indian riders, with cycles that are faster, lighter and stronger than anything these roads have ever seen.

Project Role: Big Data Developer.

Technologies: Hadoop, HDFS, Hive, Sqoop, Spark, SQL, AWS.

Responsibilities:

- ❖ Performed **Import and Export** of data into **HDFS and Hive** using **Sqoop** and managed data within the environment.
- ❖ Implement predefined operators such as **GroupBy** and **ReduceBy** in **Scala Spark**.
- ❖ Exported necessary **spark Jars** to run in the cluster.
- ❖ Performed **Import and Export** of remote data to **AWS S3**.
- ❖ Performed Data enrichments such as **filtering, sorting** and **aggregation** using Hive.

- ❖ Exported necessary spark Jars to run in the cluster.
- ❖ Used JIRA for bug tracking and Bit bucket to check-in and checkout code changes..
- ❖ **Loaded and transformed** large sets of semi structured data likes **CSV, XML, JSON, Avro, Parquet**.
- ❖ Loaded and transformed large sets of flat files data
- ❖ Assisted QA team during testing and defects fixes
- ❖ Batch script analysis and execution in putty tool.
- ❖ Developed workflow in control-m to automate the tasks of loading the data into Target.
- ❖ Preparation of Documentations for CR and attain signoff from different business levels for deployment.

Project Name: Bharat Matrimony

June 2020 – Oct 2021

Bharat Matrimony, India's leading provider of online matchmaking services delivers matchmaking services to users in India. Hundreds of thousands of people have met their life partners through our revolutionary matchmaking service and countless others have made some very special friends.

Project Role: Hadoop Developer

Technologies: Hadoop, HDFS, Hive, Sqoop, Oozie, Map Reduce, Spark Sql

Responsibilities

- ❖ Performed **Import and Export** of data into **HDFS and Hive** using **Sqoop** and managed data within the environment.
- ❖ Involved in **creating Hive tables, data loading** and writing **hive queries**.
- ❖ Involved in Writing DDL and DML scripts to transform Data and populate in Target table.
- ❖ Was responsible for **Optimizing Hive** queries that helped in saving Cost to the project.
- ❖ Handled **Hadoop Map Reduce jobs** to process large data sets.
- ❖ **Managed Hive Tables** and created child tables based on partitions.
- ❖ **Loaded and transformed** large sets of semi structured data.
- ❖ **Code & peer review** of assigned task. Unit testing and Bug fixing.
- ❖ Wrote Hive queries for transformations on the data to be used by downstream models.
- ❖ Used JIRA for bug tracking and Git hub to check-in and checkout code changes.

WORK EXPERIENCE

- ❖ Worked as an **Associate Engineer in Envensys Technologies** from August 2018 – Present.

EDUCATION

- ❖ Bachelor in Arts and Science, The American College, Madurai, 2018.

DECLARATION

I hereby declare that all the above information is true to the best of my knowledge and belief.

M.Pradeep kumar