

Meticulous Technical Lead strictly adheres to both client-provided technical requirements and internal business goals. Balances project prioritization to fulfill both contractual obligations and internal revenue requirements. Provides foresight and management designed to result in satisfied customer bases and growing revenue figures.

## Work History

2020-03 -  
Current

### Senior Data Engineer

Gspann Technology, Gurgaon, Haryana

**Project1** : Expedia Group – hotwire.com

**Technology Stack:** Spark, Hive, AWS (EMR, EC2, S3, RDS, Datalpeline), Scala, Python, Apache Kafka, Unix, Jenkins, Splunk

#### Role & responsibilities:

- Maintain existing running email campaigns and modify based on requirement in optimal solution using Scala and spark.
- Used spark to process large Scala data set.
- Loaded final data to S3 and Redshift to build Datalake for Business and also for ML team to use as their source.
- Created processed layer using spark SQL to maintain high quality data usable for analytics/data researchers/scientists/ML team.
- Used Kafka and Spark to Publish messages on Kafka Topic for PN.
- Create and modify existing pipeline to schedule batch jobs using AWS data pipeline and Crontab.
- Worked on moving all jobs from Arch (unsecured domain) to Decaf (secure zone).
- Mentor juniors/ interns in their day to day work to bring them up to speed to work independently on any technology.
- Developed Unit Test Case in Spark Applications.
- Log analysis and monitoring using Splunk dashboard.
- Sonar integration to check code quality.
- Prepared Test Plan and Testing Strategies for Big Data Application
- Collaborated with Clients and developers and prepared test plans & Scripts for producing high

# Kishan

Senior Data Engineer

## Contact

### Address

Gurgaon, Haryana, 122017

### Phone

8378075089

### E-mail

kishansingh2411@hotmail.co  
m

## Skills

Hadoop  
 Very Good


SQL  
 Very Good

Apache Spark  
 Very Good

Hive, Sqoop, Impala  
 Very Good

No SQL(Hbase, MongoDB)  
 Good

Kafka  
 Good

AWS( S3, EC2, EMR, Redshift, Data Pipleine, SNS)  
 Good

Azure  
 Good

quality software applications.

- Developed, implemented, supported and maintained data analytics protocols, standards and documentation
- Communicated new or updated data requirements to global team

**Project 2:** Sephora

**Technology Stack:** Spark, Hive, Azure (ADF, DataBricks, Blob, Azure Portal etc)

**Role & responsibilities:**

- Understand business requirements and document it and get it approved from stakeholders.
- Write spark job to implement business logic as per documentation.
- Migrate data from EDW layer( oracle ) to Azure storage and create datalake to be used by Business and other dependent team.
- Addressed ad hoc analytics requests and facilitated data acquisitions to support internal projects, special projects and investigations
- Create Azure Data pipeline to schedule job on daily/weekly/monthly basis.

**Project 3:** Puresotage

**Technology Stack:** Snowflake

- Working on POC to validate computation and execution performance of query using Snowflake

2019-03 -  
2020-03

## Senior Data Engineer

*IRIS software, Haryana*

**Project :** Citi Bank

**Technology Stack:** Apache Spark, Hive, Oracle

**Role & responsibilities:**

- Develop a POC to understand the processing time between Oracle & Hive & Spark.
- Created Pipeline to load data from Oracle to Hadoop environment using sqoop or spark and analyze and test result and create audit report on daily basis

2018-06 -  
2019-03

## Data Engineer

*Intellicus Technologies, UP*

**Projects:** HEB & Home Depot for Kyvos

Scala, Python, Shell Script

Good

GCP

Good

CI-CD( jenkins, Gthub)

Good

Data analysis

Good

Leadership

Very Good

**Technology Stack:** Hadoop, Apache Spark, Apache Hive, AWS(EMR) **Role & responsibilities:**

- Create Data Ingestion framework for different sources like files and relational Databases using spark scala.
- Created datalake for Kyvos feed using hive/spark.
- Solution designing for building data pipelines having robust & flexible models so that data can be feed to Kyvos.
- Built data flow process and data lake management platform.
- Understand business requirement and underlying data and give one solution using Kyvos at minimum cost and minimum time.
- Support different workflows (oozie and Control-M), bug fixing, analysis and resolution of job failures.
- Leading offshore team technically and functionally to understand business requirements from clients and use kyvos to solve those problems.

2017-08 -  
2018-06

## Data Engineer

*Phenom People, Hyderabad, Telangana*

**Project:** Phenom Universal Profile (PUP) Enrichment & Data Lake

**Technology Stack:** Hadoop, Spark, Kafka, Hive, Mongo DB, Elastic Search, AWS

**Role & responsibilities:**

- Built and maintained data acquisition system from different sources like file system. API, Databases etc.
- Built data flow process and data lake management platform on HDFS using Hadoop tools like spark and Hive.
- Created processed layer using spark SQL to maintain high quality data usable for analytics/data researchers/scientists.
- Developed spark application based on requirement and run on EMR cluster.
- Understand source data and convert into standard format.
- Taking release plan and production deployment for every release.
- Automate DNA Standardisation process to ensure

80-90% Data

- Developed, implemented, supported and maintained data analytics protocols, standards and documentation
- Communicated new or updated data requirements to global team

2015-01 -  
2017-05

## Senior Data Engineer

*Datametica Solution Pvt Ltd, Pune, Maharashtra*

### **Project 1:** Cable Vision

**Technology Stack:** Hadoop (CDH), Apache sqoop, Apache Hive, Spark, Bash Scripting, Apache Oozie, Appworks

#### **Roles and Responsibilities:**

- Requirements gathering and requirement analysis.
- Data migration from Oracle/sql/netezza etc using sqoop to create data lake in Hadoop infrastructure.
- Analyzed source data, worked on data cleansing by means of removing null values, target table structures using hive and pig.
- Translate complex functional and business requirements into technical requirements.
- Wrote HIVE scripts and automate into Oozie workflow for business report.
- Prepared Functional document and Document of understanding for Engineering and Support services folks
- Developed, implemented, supported and maintained data analytics protocols, standards and documentation

### **Project2 : Office Depot - Clickstream and Coverage analytics (Data lake creation)**

**Technology Stack:** Apache Hadoop (CDH), Apache Sqoop, Apache Pig, Apache Hive, Bash Scripting, Oozie, Hbase , Phoenix

#### **Roles and Responsibilities:**

- Requirements gathering and requirement analysis.
- Analyzed source data, worked on data cleansing by means of removing null values, target table structures using hive.
- Created frameworks for DB2 data unload to hdfs, Files (fixed/delimited etc).

- Created generic framework for data ingestion from different Relational DBs like (Oracle, SqlServer, Netezza )After Data ingestion to Hadoop datalake used hive, impala to analyze data

2012-03 -  
2015-01

## **Systems Engineer**

*TCS, Pune, Maharashtra*

---

## **Education**

---

2007-07 -  
2011-06

## **Bachelor of Engineering: Electrical, Electronics and Communications Engineering**

*SIRT - Bhopal, MP*