**AMOL SHINDE**

Hadoop Developer

C:\Users\ananu\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Personal.png

**Personal info**

**Address**

Lohegoan, Pune, 411047

**Phone**

+91-8484-844445

**Email**

amol.shinde@amolshinde.in

**Date of birth**

08th July 1991

**Github**

<http://github.com/amolshinde146>

C:\Users\ananu\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Skills.png

**Key Skills**

Hadoop

Map Reduce

Hive

HBase

Spark

Sqoop

Shell Script

Python

Scala

C:\Users\ananu\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Software.png

**Software**

Windows

Ubuntu Linux

JetBrains PyCharm

JetBrains Intellij

Eclipse

Phoenix

DB Analyser

C:\Users\ananu\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Language.png

**Languages**

English

Hindi

Marathi

C:\Users\ananu\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Objective.png

**Objective**

Looking forward to a challenging job in a reputable organization to utilize my Hadoop development skills for the growth of the organization and myself.

C:\Users\ananu\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Overview.png

**Professional Summary**

* IT Professional with over 3 + years of experience specializing in big data technology components like HDFS, Hive, HBase, Sqoop, and Spark
* Experience in analysing data using HiveQL
* Hands-on experience in importing and exporting data using Sqoop from relational database to HDFS and vice versa
* Good understanding of Spark Architecture including structured and low-level API's
* High technical and analytical skills with a clear understanding of design goals
* Good analytical mind with willingness and ability to quickly learn new technologies
* Worked on different RDBMS databases like SQL Server, Oracle.
* Experience in different file formats like text, ORC, Parquet, Avro.
* Expert in working with Hive data warehouse tool like creating tables, data distribution by implementing partitioning and bucketing, writing and optimizing the HiveQL
* Preparing data for analysis: profiling, cleaning, joining, transforming, aggregating, and filtering large and varied data sets.
* Creating reusable custom scripts, queries, and code commands for most of data processing tasks.
* Acquire, ingest and process data from multiple sources and systems into Big Data platforms.
* Knowledge of NoSQL database.

C:\Users\ananu\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Certification.png

**Work Experiences**

• HCL Technologies Ltd, Hadoop Developer, *21-Aug-2016 to Present*

**C:\Users\ananu\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Education.png**

**Education**

• MCA from DR. DY Patil School of MCA, PUNE(MAHARASTRA) 2014.

C:\Users\ananu\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Interests.png

**STRENGTH**

* Leadership
* Quick Learner
* Punctuality
* Goal Oriented Approach

C:\Users\ananu\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Projects.png

**Projects**

##### Project: E-commerce Data Migration & Data Analysis (E-commerce)

**Project Details:**

The PoC was based on a customer case study to find out which two items are sold together and hence enable the client to sell these products together at a discounted rate. We also analysed more than two products together. Extensively, used the Public datasets for the PoC data. Processed data using shell script spark programs and stored/retrieved the data using Hive queries.

**Environment**: Spark, Hive, Python, Shell Scripting, Sqoop.

**Role** & **Responsibilities**

* Analysed the requirements and document them in the functional specification document
* Involved in loading data from multiple sources into HDFS
* Developed the script & Spark Code for analysis of the data
* Loaded the result into Hive table for querying.
* Code and deployment guide review with peers and operations team.
* Developed Simple to complex Spark Jobs on hive operation.
* Involved in Spark and Spark SQL.
* Processing of data in Spark by creating data frame, performing filtering and applying
* various actions to achieve the business requirement.

**Project: Telecom Data Migration & Data Analysis (Telecom)**

**Environment:**  Spark, Hive, Sqoop, Shell Script, Python

**Role** & **Responsibilities**

* Perform Data Ingestion using SQOOP into HDFS and staging tables in HIVE.
* Worked on compression technique for compressing the data before loading into Core tables.
* Developed shell script which will import data.
* Developed shell script for sending automated email with status after the batch run.
* Worked on SCD-1, SCD-2.
* Experience in working with different file formats like PARQUET, ORC, JSON.
* Creation of RDD, Data frame as per use cases using Spark Framework.
* Created Partitioning in hive for better performance.
* Fixing existing issues in the system and also writing code for enhancement.
* Unit testing and integration testing of owned use cases.
* Writing different kind of analysis, design, testing and Business validations.
* Developed Map Reduce jobs for data cleaning and pre-processing.

**Project: - Data Migration (Telecom)**

**Environment**: Hive, Shell Scripting, Sqoop.

**Roles and Responsibilities:**

* Creation of all necessary HIVE tables such as staging table, Core table based on

requirement.

* Perform Data Ingestion using SQOOP into HDFS and staging tables in HIVE. The data

import is incremental in nature.

* Using proper compression technique for compressing the data before loading into Core tables.
* Developed shell script which will import data