

# SHWETANK YADAV

**Address:** 933, ghamapur, Jabalpur (M.P)

**Mobile:** +91 9406885198

**E-Mail:** shwkntyadav3091@gmail.com

---

## PROFILE

**Objective:** A qualified Post Graduate (M.E) and a dynamic **IT Professional**, having 2.5 year total professional experience with 1.5 years in Big Data processing and UNIX in big data environment and now seeks challenging opportunities to improve leadership skills/proficiencies & experience in an open environment to drive growth for organization as well as for self

### Big Data / Hadoop Engineer

**About Me:** A self-motivated and analytical Software Solution Development Professional, with successful track record of analysis/proper elicitation of business requirements, Documenting the requirement in the form of Specifications for further Design/Development/Testing so as to comply with business needs and provide valuable effort/time estimate for individual tasks, keeping constant focus on the project scope and requirements.

#### **Special Mention:**

- Experienced with study & analysis client business requirements/processes/workflow, document same for development team and provide estimate (effort/time) for assigned tasks
- Similar experience of over sight of project requirements/life cycle by tracking the same, with a constant focus on the project scope and requirements.
- Proficient further in Big Data (Hadoop) Management and System (UNIX)/Services Management
- An innovative **Team Player** with Fast Learning curve on new technologies and possess analytical, problem solving, decision making, planning, communication and interpersonal skills

---

## QUALIFICATION

#### **Education:**

- **ME.**, from Jabalpur engineering college, Jabalpur, M.P (2016-2018) with **78%**
- **B.E.**, from LNCT Engineering college, Jabalpur, M.P (2009-2013) with **70%**

#### **Accreditation:**

- Oracle pl/sql, UNIX, Scala,python training
- Hadoop Ecosystem training

#### **Certification:**

**Hadoop with spark and scala, Besent technology,Bangalore.**

#### **Technical Skills:**

- **Operating Systems:** UNIX & Linux, Windows
- **Languages:** Python, Scala
- **Databases:** Oracle 10g, Hive
- **IDE:** IntelliJ Idea, Eclipse,pycham
- **Tool:** Putty, Jira, Bitbucket,Ctrl-M.
- **Big Data Technologies:** Apache Hadoop – MR, Hive, Sqoop, Pig, Spark, SparkSql
- **Hadoop Distributions:** Apache, Cloudera, Horton

---

## PERSONAL DOSSIER

- Date of birth: 05-Oct-1991
- Languages known: English,Hindi
- Location Preference: Pune
- Nationality: Indian
- Passport Details: U5774335 (valid till 04-02-2030)

**PROFESSIONAL EXPERIENCE**  
**IBM,INDIA(26-MAY-2021 TO13-AUG-2021)**

**PROJECT SUMMARY**

<b>Project</b>	<b>Transcate enrichment</b>
<b>Client</b>	Westpac banking corporation, Australia
<b>Duration</b>	2.5 month
<b>Team Size</b>	12
<b>Environment</b>	Cloudera
<b>Details</b>	Project for linking connection between customers and banking system.
<b>Role</b>	Getting data in Excel file format need to convert it to csv file using python program ,read those data using pyspark put necessary transformation according to logic ,upload to hive table process, put some check and balance, hive queries send data to archive location.
<b>Responsibilities</b>	Requirement gathering from Product Owner and Analyzing them. <ul style="list-style-type: none"> <li>• Use sql terradata for testing logic Involved in implementing code logic in PySpark</li> <li>• Involved in implementing code logic in PySpark.</li> <li>• Extensively worked in Hive ,shell script and HBase data processing</li> <li>• Exposure in writing Hive queries</li> </ul>
<b>Skill used</b>	Python,BASH scripting, Control -M

**ELIXIR ENGINEERING,(17-MARCH2019 TO 17-MAY-2021)**

**PROJECT SUMMARY**

<b>Project</b>	<b>Incident ticket analysis</b>
<b>Client</b>	Dell Emc
<b>Duration</b>	7 month
<b>Team Size</b>	12
<b>Environment</b>	Dell Emc (Cloudera)
<b>Details</b>	Project for Customer help desk analysis scala for shell scripting and converting unstructured data into structured data compatible for apache spark framework.
<b>Role</b>	Getting data in Excel file format need to convert it to csv file using scala program ,read those data upload to hive table process, put some check and balance, hive queries send data to archive location.worked as hadoop/spark developer for huge amount of data
<b>Responsibilities</b>	Requirement gathering from Product Owner and Analyzing them. <ul style="list-style-type: none"> <li>• Involved in implementing code logic in Spark and Scala/Python environment</li> <li>• Created Archival frame work to archive old partitions.</li> <li>• Created Data capping framework for downstream consumers.</li> <li>• Extensively worked in Hive ,shell script and HBase data processing</li> <li>• Exposure in writing Hbase and Hive queries</li> <li>• Developed Spark applications using Spark-Core, Spark SQL and Hive</li> <li>• Extensively worked on Hadoop Stack Such as HDFS,HIVE,HBASE, Spark, Scala, Python,BASH scripting, Control -M .</li> </ul>
<b>Skill used</b>	Hadoop,hive,sqoop,scala

<b>Project</b>	<b>Critical Business Analytics</b>
<b>Client</b>	Legato
<b>Duration</b>	8 months
<b>Team Size</b>	8
<b>Environment</b>	Cloudera
<b>Details</b>	Business analysis of all insurance holder and claim processing work on spark sql,python to provide Batchid, reading and writing data from csv file, developing hadoop/spark program

	using scala/python.
<b>Role</b>	Integration design and pipelining data using spark-Scala hive and control-M
<b>Responsibilities</b>	<p>Involved in major enhancements to convert Hive/SQL queries into Spark Transformations using Spark SQLs and Spark Dataframes RDDs in Scala.</p> <ul style="list-style-type: none"> <li>• Worked on reading multiple data formats like CSV, ORC on HDFS.</li> <li>• Experienced in handling large datasets using Partitions, Spark in Memory capabilities, Broadcasts in Spark, Transformations and other during ingestion process itself.</li> <li>• Classified threshold value to identify duplicates, repeated and similar.</li> </ul>
<b>Skills used</b>	Spark-sql, hive, scala

**DECLARATION:** I hereby declare that all the information furnished above is correct to the best of my knowledge.

(Shwetank Yadav)