

# Deepti Mishra

8755310290 | mishradeepti55@gmail.com

## Profile Summary

- Currently working in Infosys with overall experience of 2.9 years in Big Data, ETL, Hadoop, Python, Spark, SQL & Azure.

## Skills

- Tools : Microsoft Sql Server, MySql, Spark, Spark sql, HIVE, SQOOP, Hadoop, Linux, Azure data factory, Synapse Analytics, DataLake, Databricks
- Programming languages : SQL, Python, HQL (Hive Query Language), Scala

## Professional Experience

### • Infosys Limited

Nov, 2018 - Present

Systems Engineer

- Got training in Infosys in Big Data, Hadoop, Python, SQL.
- Good understanding of Hadoop Architecture & components.
- Experience in Mysql & SQL Server.
- Hands on experience on RDD Architecture, implementing Spark operations on RDD and optimizing transformations in Spark.
- Experience in creating Hive tables and queries using HQL.
- Knowledge of Hive partitioning and bucketing.
- Experience in Python.
- Experience in Data Transformation using Hive, Spark, Sqoop.
- Identifying and resolving performance issues in various levels like Hive Performance issues. Good understanding of Spark transformations and Actions, DataFrames and Spark Sql.
- Azure Services like Data factory, synapse analytics, data bricks, power BI

## Projects

### • ETL & Data Analysis

- Involved in converting Hive queries into Spark transformations using Spark RDD and Scala / Python Use
- Spark SQL for analysis of captured data.
- Develop program in Python and Scala to perform transformation, create Dataframes and using Spark Sql. Used Sqoop for transferring Data.

- **ETL - Azure**

- Using Azure data factory for transforming data using adf pipelines / activities & fetching data from on - premise sources.
- Using Azure synapse analytics for data warehousing.
- Using various azure services for developing complete ETL Pipelines for data warehousing purpose and Analytics.
- Analysis of datasets using Spark & Spark Sql with Databricks.
- Databricks notebooks integration with adf pipeline.
- Python-Azure SDK for automating platform activities on Data Lake (Storage Account) like Creating containers, directories, applying ACLs, collecting meta data of files, generating checksum etc. using Python scripts.
- Deploying application with Azure Devops Pipeline using .yaml files.

- **Data Cleansing, Analysis & Report Generation**

- Datasets cleansing using Python.
- MongoDB used for storage purposes.
- Report generation done by using Power BI. Used Git for version control and Jenkins for Continuous Integration.

## Education

- **Maharishi Vidya Mandir Sr. Sec. School , Nainital**  
HSE 2012
- **Maharishi Vidya Mandir , Sr. Sec. School , Nainital**  
SSE 2014
- **Dr. Ram Manohar Lohia Avadh University , Uttar Pradesh**  
B.Tech. ( Computer Science & Engineering ) 2014-2018

## Certifications

- Azure Fundamentals - AZ-900  
Infosys Certified Big Data Developer  
Infosys Certified Spark Developer  
Infosys Certified Python Developer
-