

MIHIR SINGH

PYSPARK_2+, HADOOP AND ML DEVELOPER

The best of work comes to those who earn it.

PROFILE

Big Data Specialist at IBM with total experience of 3.3 years in Big Data.

AWARDS

Awarded Excellence award for exceptional performance in Big Data Development from IBM.

Won ZONAL ROUND from INSTITUTE OF USA in

Robotics and MATLAB for Hand gesture recognition.

CERTIFICATIONS

- | Certified JAVA Programmer awarded

by HP with grade A

- | Certification from Udemy on Python, Hive, Pyspark, Scala, ML using Python and Spark.

- | Acclaim Badge holder for Certified SPARK and SPARK ML developer.

Hacker-rank certification in Python.

WORK EXPERIENCE

IBM- BIG DATA and ML engineer Apr,2020-Present

Accenture -Big Data and ML Engineer Feb,2018–Apr,2020

PROJECTS

- | **Data Warehousing and Aggregation:** Worked on end to end data ingestion techniques from various source using **SPARK JDBC** and in various **SCD types (Type 1 and 2)**. Have worked on complex problem and delivered solutions with optimized results in **spark sql and Structured Streaming**.
- | **Big Data Developer — Data migration and Ingestion in Hadoop** Developed Data Ingestion techniques in Pyspark to extract and transform using Spark, calculate KPI's and pushing the data to Hive table. The process was automated using Oozie workflow and Scheduled via a coordinator. The infrastructure used from the project was **Azure HDInsight (Cloud services)**.
- | **Suggested order model development using SparkML:** Have developed model in **SparkML using inbuilt K-means clustering and Apriori Algorithm(custom algo to promote products in particular category)** for a recommender system that used unsupervised learning to classify and recommend focused products based on previous purchase.

EDUCATION

Amity University B. Tech CS&E: 2013 – 2017 CGPA: 6.53 Grade A

ISC: 2012 – 2013 79% Computer Science and Mathematics

ICSE: 2010-2011 80.59% All subjects

SKILLS

