HARSHIT TEWARI

Specialist Programmer at Infosys · BTech (CS) · Jaypee Institute of Information Technology - Noida
+91 7905177037 · htewari20@gmail.com · linkedin.com/in/hat20

PROFESSIONAL EXPERIENCE

Oct 2021 – Ongoing

Asset Data Repository – REALM Azure Migration   *Specialist Programmer*

- Led a team of 5 on the migration of entire deployment structure from **legacy systems to Azure**.
- Responsible for migrating the **data warehousing** services from **Hive** to **Azure Synapse**.
- Worked together with the team to implement all the necessary **Scala** code changes.
- Improved the data ingestion time significantly by implementing **T-SQL COPY INTO** approach for our existing **Spark Scala** code base.
- Built the data validation strategies and implemented testing at functional, unit and integration level.

Asset Data Repository – LDTI   *Specialist Programmer*

- Automated the process of loading a file into Azure Synapse from an existing ADLS location.
- Designed and integrated the **T-SQL COPY INTO** approach to our existing code base.
- Responsible for the creation of **ADF pipelines** for development environment.
- Provided last minute enhancements when needed and giving support to the production application after successful deployment.
- Worked together with the Business Analytics team for the creation of **Kibana** dashboards.

Skills Used – *Hadoop, Spark Scala, Spark SQL, Python, Hive, SQL Server, Azure - ADLS, ADF, Synapse & Databricks*

Aug 2019 – Sept 2021

Data Operations Workbench   *Systems Engineer*

- Created customized tools in **Python** such as **Cost Management** tool, **data ingestion and monitoring** tool, **CloudFormation** templates generator for **AWS**, **job scheduler for Azure** tool for stakeholders.
- Translated and transformed the raw data gathered from various tools into dashboards in **Splunk** for the analysis of the data produced.
- Responsible for migrating all dashboards from **Splunk** to **Kibana**.
- Provided online demonstrations of the tools to the management for current and prospective clients.

Big Data Training Capstone Project   *Systems Engineer Trainee*

- Led the team of 4 as a Team Leader and Scrum Master.
- Performed data cleaning and transformation of **IPL Ball by Ball and IPL all statistics** dataset by using **Pig** along with **HBase** and then proceeded to perform in-depth analysis by applying sets of business rules.
- Provided insights to the team for the creation of dashboard in **PowerBI**.
- Applied **Random Forest** and **N-fold cross validation** machine learning techniques to predict the outcome for the 2017 and 2018 seasons with **82% accuracy**.

Skills Used – *Hadoop, Machine Learning, Python, Hive, SQL Server, HBase, Pig, Splunk, Kibana, Azure - ADLS, ADF*

ACHIEVEMENTS

- Qualified Infosys Power Programmer Bridge Program to accelerate career progression from Systems Engineer to Specialist Programmer. Only 174 selected out of roughly 10000 developers.
- Capstone Project declared as the best project of our Big Data training batch.