Name of the Project – Recommendation System for OTT Platform.

Objectives/Vision

During this time of the pandemic, we have seen tremendous growth in the OTT Platforms. The global over-the-top (OTT) market is expected to observe remarkable growth owing to the rising adoption of OTT platforms as more people are subscribing to these entertainment platforms during the Covid-19 pandemic lockdown. It is expected to grow from 104.11 billion dollars in 2019 to 161.37 billion dollars in 2020 at a compound annual growth rate of a whopping 55%. This exponential growth, again, is mainly due to the worldwide lockdown caused by the COVID-19 outbreak, during which subscription to various OTT streaming channels and viewership has increased. Customer's social behavior is shifting from traditional subscriptions to broadcasting services and to over-the-top (OTT) on-demand video. This will drive the OTT streaming market in the forecast period at a very fast pace. Various segments of the population have started using video streaming services instead of regular television for entertainment, due to added benefits such as on-demand services and ease of access. The dramatic rise in demand for live streaming channels and the ongoing creation of cloud-based OTT services would drive substantial market growth in the following years. Audiences are shifting to online streaming platforms like Hotstar, Hulu, Netflix, and Amazon Prime Video. Not only users are binging movies and series from such platforms, but the businesses are also making huge profits from their exclusive content.

Functional Requirements

1. Understand Data Engineering Discipline and can work on Azure Cloud services.
2. Design and develop a complete ETL process based on the analysis of input data from various sources using ETL tools like Spark.
3. Get exposure to data migration into the cloud.
4. Understand monitoring and security aspects and implementation.
5. Understand Hadoop big data, Hadoop fundamental, and Data warehouse fundamental.
6. Overview of a real-time ETL pipeline with Kafka.

7. Overview of PySpark, if there is a scope in the respective use case.
8. Get exposure to data visualization and a corresponding tool like Tableau.

Non-Function Requirement:

1. Analyze the problem domain
2. Analyze the input data formats and design for data clean-up
3. Address data migration from the current system to the cloud (Can Include selecting and applying an ETL Tool)
4. Identifying the optimal cloud services aligned to the business commercial goal
5. Define and document your proposed architecture
6. Create and present a Minimal Viable Release (MVP) to the stakeholders.

Tools and Technologies to be used

- Data Extraction, Data transformation, data loading, and migration tools,Hadoop,spark

- Visualization tool: Tableau

- Cloud components