# Topic: Recommendation Engine

**Instructions**

Please share your answers filled inline in the word document. Submit Python code and R code files wherever applicable.

Please ensure you update all the details:

**Name: Shahina Athar**

**Batch Id:** 10122020

**Topic: Recommender Engine**

1. **Business Problem**
   1. **Objective**
   2. **Constraints (if any)**
2. **Work on each feature of the dataset to create a data dictionary as displayed in the below image:**



**Using R and Python codes perform:**

1. **Data Pre-processing**

**2.1 Data Cleaning and Data Mining.**

1. **Exploratory Data Analysis (EDA):**
   1. **Summary**
   2. **Univariate analysis**
   3. **Bivariate analysis**
2. **Model Building**
   1. **Build the Recommender Engine Model on the given data sets.**
   2. **Use UBCF technique for getting recommendations.**
3. **Share the benefits/impact of the solution - how or in what way the business (client) gets benefit from the solution provided.**

# Note:

The assignment should be submitted in the following format:

* R code
* Python code
* Code Modularization should be maintained
* Documentation of the model it can be word of PDF format (elaborating on steps mentioned above)

Q) Build a recommender system with the given data using UBCF.



**Problem Statement: -**

This dataset is realted to the video gaming industry and a survey was coducted to build recommendation engine so that the store can improve the sales of its gaming DVD’s. Snapshot the dataset is given below build a recommendation engine and suggest top selling dvds to the store customers.



**1.Business Problem:**

**Objective:** This dataset is realted to the video gaming industry and a survey was coducted to build recommendation engine so that the store can improve the sales of its gaming DVD’s. We as a Data Scientist suggest top selling dvds to the store customers.

**2. Work on each feature of the dataset to create a data dictionary as displayed in the below image:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name of Feature** | **Description** | **Types** | **Relevance** |
| userId | User Id of the customer | Quantitative, Discrete, Count | Relevant |
| game | Name of the game | Quantitative, Discrete,  Nominal | Relevant |
| rating | Ratings of the game | Quantitative, Continuous,  Interval | Relevant |

1. **Share the benefits/impact of the solution - how or in what way the business (client) gets benefit from the solution provided.**

**Ans:** This dataset is related to the video gaming industry and a survey was conducted to build a recommendation engine so that the store can improve the sales of its gaming DVD's. We as a Data Scientist suggest top-selling DVD's store customers. The benefits are we suggest top-selling DVD’s to the store customers so that it’s benefits for both customer as well as the video gaming industry, customers get best choices because of recommendation and their time is also saved, whereas video gaming industry gains their profits. I generate a recommendation that recommends to the customer top-selling DVD’s.

**Problem Statement: -**

The Entertainment Company, which is a startup online movie watching platform, wants to improvise its collection of movies and showcase those that are highly rated, and recommend those movies to its customer by their movie watching footprints. For this the company has collected its data and shared it with you to provide some analytical insights and also to come up with a Recommendation Algorithm so that it can automate its process for effective recommendations based on Users Interest and behavior patterns.

Remember the ratings are between -9 to +9.

**1.Business Problem:**

**Objective:** The Entertainment Company, which is a startup online movie watching platform, wants to improvise its collection of movies and showcase those that are highly rated, and recommend those movies to its customer by their movie watching footprints. For this the company has collected its data, we as a Data Scientist provide some analytical insights and also to come up with a Recommendation Algorithm so that it can automate its process for effective recommendations based on Users Interest and behavior patterns.

**2. Work on each feature of the dataset to create a data dictionary as displayed in the below image:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name of Feature** | **Description** | **Types** | **Relevance** |
| Id | Customer Id’s | Quantitative, Discrete, Count | Relevant |
| Titles | Title of the movies | Qualitative, Discrete,  Nominal | Relevant |
| Category | Category of the movies | Qualitative, Discrete,  Nominal | Relevant |
| Reviews | Reviews of the movies | Quantitative, Continuous,  Interval | Relevant |

1. **Share the benefits/impact of the solution - how or in what way the business (client) gets benefit from the solution provided.**

**Ans:** Entertainment Company, which is a startup online movie watching platform, wants to improvise its collection of movies and showcase those that are highly rated, and recommend those movies to its customer by their movie watching footprints. For this the company has collected its data, we as a Data Scientist provide some analytical insights and also come up with a Recommendation Algorithm so that it can automate its process for effective recommendations based on User's Interests and behavior patterns. The benefits are we suggest customers those movies according to their interest so that it’s benefits for both customer as well as the Entertainment Company, customer get best choices because of recommendation and their time is also saved, whereas Entertainment Company gains their profits. I generate a recommendation that recommends to the customer top-10 movies according to their interest.

A screenshot of a cell phone

Description automatically generated