# Topic: Recommendation Engine

**Instructions:**

Please share your answers filled in-line in the word document. Submit code separately wherever applicable.

Please ensure you update all the details:

**Name: DHEERAJ MISHRA Batch ID:** DS\_01072021

**Topic: Recommender Engine**

**Hints:**

1. **Business Problem**
   1. **What is the business objective?**
   2. **Are there any constraints?**
2. **Work on each feature of the dataset to create a data dictionary as displayed in the image below:**



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.**
3. **Write about the benefits/impact of the solution - in what way does the business (client) benefit from the solution provided?**



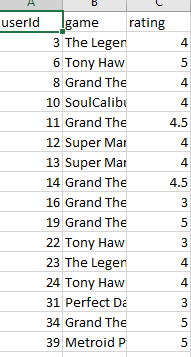
**Problem Statement: -**



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

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 DVDs. Snapshot of the dataset is given below. Build a Recommendation Engine and suggest top selling DVDs to the store customers.



1. **Business objective :**

**Max:- User based recommendation**

**Min:- Not popular items**

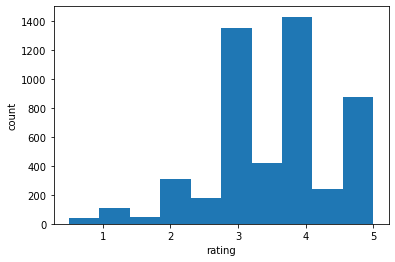
**Constraints:- Skill of employer**

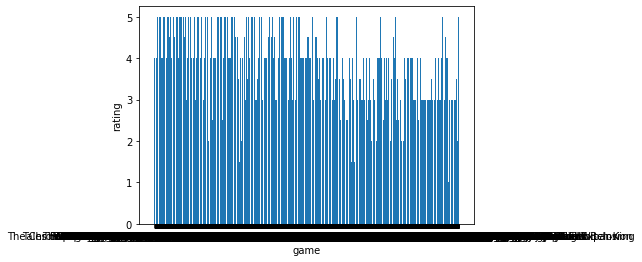
1. Data understanding :

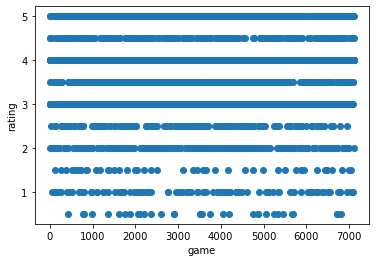
|  |  |  |  |
| --- | --- | --- | --- |
| NAME OF FEATURE | DESCRIPTION | TYPE | RELEVANCE |
| userid | Id of user | Discrete | Relevant |
| Game | Name of game | Strings , char | Relevant |
| rating | Ratings provided by the user | Continuous | Not relevant |

1. DATA CLEANSING

1. No null values found in each feature
2. All features are of type int64 , flot64 , object
3. Dataset consists of 5000 rows and 3 columns
4. TFIDF calculated for game column
5. Cosine similarity calculated for TFIDF sparse matrix
6. EDA:-
7. From histogram most of the users given rating 4
8. From bar graph we can observe for ratings with respect to game
9. From scatter plot rating is constant







1. MODEL BUILDING :-
2. User define model created for recommendation engine
3. Recommendation with respect to user id

1. **Impact :**

**From above information we can give recommendation according to user id . Also we can check which game has low ratings**

**Problem Statement: -**

The Entertainment Company, which is an online movie watching platform, wants to improve its collection of movies and showcase those that are highly rated and recommend those movies to its customer by their movie watching footprint. For this, the company has collected the 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. The ratings are between -9 and +9.

A screenshot of a cell phone

Description automatically generated

1. **Business objective :**

**Max:- User based recommendation**

**Min:- Not popular movies**

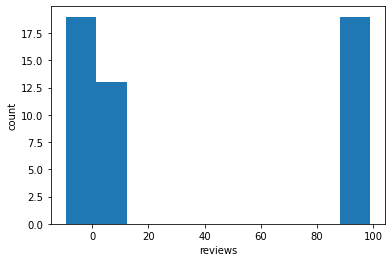
**Constraints:- Skill of employer**

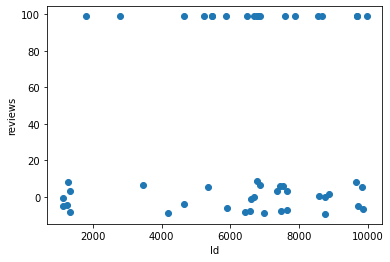
1. Data understanding :

|  |  |  |  |
| --- | --- | --- | --- |
| NAME OF FEATURE | DESCRIPTION | TYPE | RELEVANCE |
| Id | Id of user | Discrete | Not relevant |
| Titles | Name of movie | Strings , char | Relevant |
| Category | Characteristics of movies | Strings , char | Relevant |
| Reviews | Ratings provided by the user | Continuous | Not relevant |

1. DATA CLEANSING

1. No null values found in each feature
2. All features are of type int64 , flot64 , object
3. Dataset consists of 51 rows and 4 columns
4. TFIDF calculated for game column
5. Cosine similarity calculated for TFIDF sparse matrix
6. EDA:-
7. From histogram most of the users given rating -9 to 4 and 90 to 99
8. From scatter plot some rating is constant and some is scatter





1. MODEL BUILDING :-
2. User define model created for recommendation engine
3. Recommendation with respect to movies name

1. **Impact :**

**From above information we can give recommendation according to movies name . Also we can check which movies has low reviews**