**PROJECT PROPOSAL (GROUP E)**

**Problem Definition:**

Movies are a popular form of entertainment, and there is a need for personalized movie recommendations to enhance the user experience. The main problem we are trying to solve is providing accurate and relevant movie recommendations to users based on their preferences.

**Description of the Dataset:**

The dataset used in this project was obtained from Kaggle under the movie industry. Below shows a detailed description of the dataset.

# Column Non-Null Count Dtype Missing Values

--- ------ -------------- ------ ---------------

0 name 7668 non-null object 0

1 rating 7591 non-null object 77

2 genre 7668 non-null object 0

3 year 7668 non-null int64 0

4 released 7666 non-null object 2

5 score 7665 non-null float64 3

6 votes 7665 non-null float64 3

7 director 7668 non-null object 0

8 writer 7665 non-null object 3

9 star 7667 non-null object 1

10 country 7665 non-null object 3

11 budget 5497 non-null float64 2171

12 gross 7479 non-null float64 189

13 company 7651 non-null object 17

14 runtime 7664 non-null float64 4

No. Of Rows – 7668 No. Of Columns – 15

**Proposed Solution:**

The proposed solution is a movie recommendation system that utilizes genre, casting crew, movie released company, year released, director, country, and ratings as features for similarity-based recommendations. It utilizes a content-based filtering approach to identify movies with similar attributes to the user’s searched movie. We propose to utilize a classification-based machine learning algorithm as part of our recommendation system.