**Movie Recommendation System using K Nearest Neighbors**

**Submitted By:**

Mehak Saleem 2017-CE-54

Shanza Gul 2017-CE-58

Ayesha Akram 2017-CE-60

**Authors:**

Mehak Saleem (saleem.mehak266@gmail.com)

Shanza Gul ([shanzagul2525@gmail.com](mailto:shanzagul2525@gmail.com))

Ayesha Akram (ayeshaakram192@gmail.com)

**Abstract:**

We propose a movie recommendation system based on genre correlations. The movie recommender system is built using the K-Means Clustering and K-Nearest Neighbor algorithms and content based filtering. The dataset is taken from google class. The system is implemented in python programming language. The proposed work deals with the introduction of various concepts related to machine learning and recommendation system.  For different cluster values, different values of Root Mean Squared Error are obtained and movies are recommended with minimum distance.

**INTRODUCTION:**

Recommender System is a system that seeks to predict or filter preferences according to the user’s choices. Recommender systems are utilized in a variety of areas including movies, music, news, books, research articles, search queries, social tags, and products in general.

Recommender systems produce a list of recommendations in any of the two ways:

* **Collaborative filtering**
* **Content-based filtering**

Our project is **“Movie Recommendation System”** based on Content-based filtering ,this approach uses a series of discrete characteristics of an item in order to recommend additional items with similar properties. Content-based filtering methods are totally based on a description of the item and a profile of the user’s preferences. It recommends items based on user’s past preferences.

**DATA SET DETAILS:**

Our data set contains information of movies including movie id ,name and category of movie like adventure, crime, thriller, romance etc.

**DATA PREPROCESSING DETAILS:**

Data has been converted in 0 n 1 form on the basis of the genre. The attributes which are present in the movie are marked as 1 otherwise 0.The genre column is converted into the list of 0 and 1.

**METHOD:**

KNN Algorithm is one the simplest classification algorithm that attempts to determine what group of data point is in by looking at the data points around it. The KNN algorithm assumes that similar things exist in close proximity. In other words, similar things are near to each other.

**STEPS:**

1. Load the data (training+test)
2. Initialize K (any integer) to your chosen number of neighbors

3. For each example in the data

* Calculate the distance between the query example and the current example from the data.
* Add the distance and the index of the example to an ordered collection

4. Sort the ordered collection of distances and indices from smallest to largest (in ascending order) by the distances

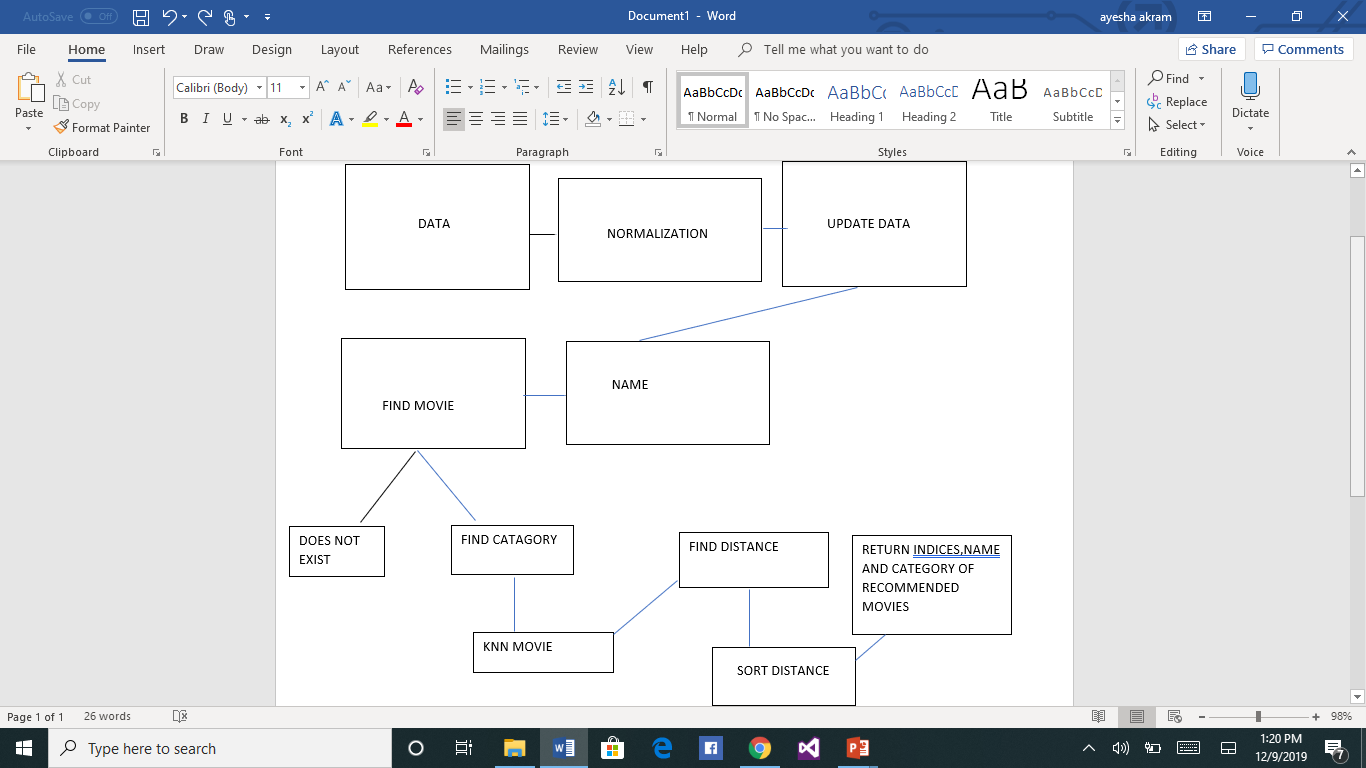
5. Pick the first K entries from the sorted collection

6. Get the labels of the selected K entries

7. If regression, return the mean of the K labels

8. If classification, return the mode of the K labels

**ARCHITECTURE:**

****

**IMPLEMENTATION DETAIL:**

First we have performed preprocessing on the raw data and converted it in 0 and 1 form on the basis of genre. and created a new csv file which is further used in the project Then we created a template of interface in html and linked it to our main knn.py file in which we rendered the template and performed the execution.

**A screenshot of a social media post

Description automatically generatedSCREENSHOTS:**

**A screenshot of a social media post

Description automatically generated**

**EXPERIMENT AND RESULTS:**

When we enter the name of movie in the Search box and hit the search button ,10 movies display on the console whom distance to the movie name entered is shortest. The movies recommendation take place on basis of genre (category).

# **References:**

<https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761>

<https://www.geeksforgeeks.org/python-implementation-of-movie-recommender-system/>