A SEMINAR REPORT ON

MOVIE RECOMMENDATION SYSTEM

SUBMITTED TO THE SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE
IN THE PARTIAL FULFILLMENT FOR THE AWARD OF THE DEGREE

OF

BACHELOR OF ENGINEERING
IN
COMPUTER ENGINEERING

By

UNDER THE GUIDANCE OF
PROF.Nimbalkar.S

DEPARTMENT OF COMPUTER ENGINEERING
SVPM'S COLLEGE OF ENGINEERING.
MALEGAON(BK)
A/P-MALEGAON(BK)-413115, TAL. BARAMATI, DIST. PUNE (MS) INDIA

2021-2022

CERTIFICATE

This is to certify that the Seminar report entitled

Movie Recommendation System

is a bonafide work carried out by them under the supervision of Prof.Nimbalkar.S/ and it is approved for the partial fulfillment of the requirement of Savitribai Phule Pune University for the award of the Degree of **Bachelor of Engineering** (Computer Engineering)

This Seminar report has not been earlier submitted to any other Institute or University for the award of any degree or diploma.

Prof. Nimbalkar.S	**Prof. Dabhade S.A.**
Internal Guide	Head of Department
Department of Computer Engineering	Department of Computer Engineering
.....	**Dr. Mukane S.M.**
External Examiner	Principal
Date:	SVPM's College of Engg. Malegaon(BK)

Place: SVPM's College of Engineering,Malegaon(Bk).
Date:

Acknowledgements

I take this opportunity to thank our seminar guide Prof. Nimbalkar.S and Head of the Department Prof.Dabhade.S.A for their valuable guidance and for providing all the necessary facilities, which were indispensable in the completion of this seminar report. I am also thankful to all the staff members of the Department of Computer Engineering of SVPM's College of Engineering, Malegaon(Bk) for their valuable time, support, comments, suggestions and persuasion. I would also like to thank the institute for providing the required facilities, Internet access and important books.

khaire pritee
Karale Pratik
Shinde Rushikesh
Kamble Neha

Abstract

Abstract goes here : This is the era of information; Large amount of data is available. The availability of data is Only valid when it is beneficial to human in their work/ daily/labour. This system is a Personalized movie Recommendation which suggests the user which movie the individual should watch based on, the individual's previous interest, ratings interaction with the system. The system will play an important role especially when the user has no clear view/idea of which movie, he/she should watch. The system is designed and implemented with the help of Content-based filtering algorithm the system is built and tested on the available data set and the test results showed that the system has good recommendation effect. Keywords: content-based filtering, Movie Recommendation, NLP Algorithm

Contents

List of Figures

Chapter 1

Introduction

In the era of emerging technology, we have been living in the world where data is everywhere available about almost everything, the data only needs to be used properly and correctly in order and use the data making the users work more effortless. 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: Collaborative filtering approaches build a model from user's past behaviour (i.e. items purchased or searched by the user) as well as similar decisions made by other users. This model is then used to predict items (or ratings for items) that user may have an interest in. Content-based filtering: Contentbased filtering approaches 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.

1.1 Aim of Seminar

Design and Implementation of Movie Recommendation System based on NLP And Content-based Filtering algorithm

1.2 Seminar Objectives

Content based filtering uses item features to recommend other items similar to what the user likes, based on their previous action or explicit feedback.

1.3 Application

In this paper, the key research contents are to help users to obtain user-interested movie automatically in the massive movie information data using content-based filtering algorithm, and to develop a prototype of movie recommendation system based on content-based filtering algorithm.

Chapter 2

Literature Survey

1 NLP algorithm In Machine Learning Natural language processing (NLP) is a field in which computers understand, analyse, and derive meaning from the provided language or human language in a smart and useful and effective way. By using NLP, developers can learn, organize and structure knowledge which will help them to perform tasks such as automatic summarization, translation, named entity recognition, relationship extraction, speech recognition, and topic segmentation and many their task can be performed using NLP. I. NLP Examples • Use Summarizer to automatically summarize a block of text, exacting topic sentences, and ignoring the rest. 2. Content-based filtering Content-based filtering —Content based Filtering is to make recommendations based on similar products/and services according to their attributes Content based recommendation engines, which is the movie recommendation system that we developed, takes content or attributes of a product/movies you like, for example a movies genre, cast, director, keywords etc. , and then similarity matrix is calculated, based on the ranks other products/movies, on how similar they are to the liked product/movies, in this case we rank different movies based on how similar the recommended movies are to the liked movie using something called similarity scores

2 Finding Similarity 2.1 How similar are the text from each movie and how do we find the similarity between them? First let us analyse the text, Text From Movie A: The word "Amazing" is present 2 times and the word "Spiderman" is present 1 time. Text From Movie B: The word "Amazing" is present 1 time and the word "Spiderman" is present 2 times. Now, let's go and plot this on a 2-Dimensional

graph. Text from Movie A will have the point (1,2) and The Text from Movie B will have the point (2,1) where the X-axis on the graph indicates the number of times the word “Spiderman” appears and the Y-axis indicates the number of times word “Amazing” appears. The origin point for both vectors is (0,0). We can change text to a similar vector of word counts by using a Count Vectorizer function or just by doing what we did above.

2.2 Create vector representation for Bag of words, and create the similarity matrix The recommender model can only read and compare a vector (matrix) with another, so we need to convert the ‘Bag of words’ into vector representation using Count Vectorizer, which is a simple frequency counter for each word in the ‘Bag of words’ column. Once I have the matrix containing the count for all words, I can apply the cosine similarity function to compare similarities between movies. Cosine Similarity formula to calculate values in Similarity

Chapter 3

Problem Statement

Design and Implementation of Movie Recommendation System based on NLP And Content-based Filtering algorithm

3.1 Future scope

There is a vast amount of data available in the digital era, in which we are living now. Most of the data i.e. about 79branch of data science, which helps to extract meaningful data and insight from, the raw available data. Here, NLP play's an important role in Data Science, especially in the field of text data which provides insight from text data. Experts have predicted that the demand for the NLP experts will grow exponentially in the near feature. Future scope of NLP explains that, In NLP machines are thought to process and interpret text as it is done by the humans. NLP is considered as the "text analysis enabler and speech recognition applications. " NLP having the capability of interpreting text makes the task of analysing of Large amount of data simple and productive.

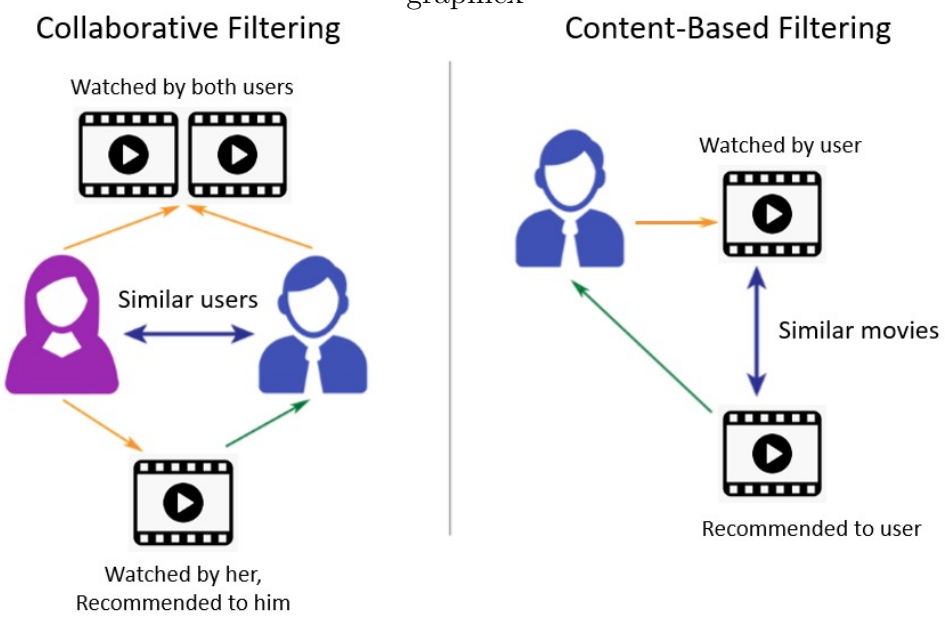
Chapter 4

Details of Design of System

4.1 Overview of System Architecture

After Pre-processing, testing and training the module, we moved on to design the Front-end for the developing the frontend we used Flask frame work. In the front-end design, the user is asked to enter the Hollywood movie on which they would like to have the recommendation. If the movie is present in to the database then based on the similarity matrix top 10 most similar movies are recommended by the system. / If the entered movie is not present the system shows notification as "The Movie You Entered is not Present in the Database Please Enter another Hollywood Movie". i. This figure shows that the Movie iOS not in the data base or it's spelled wrong

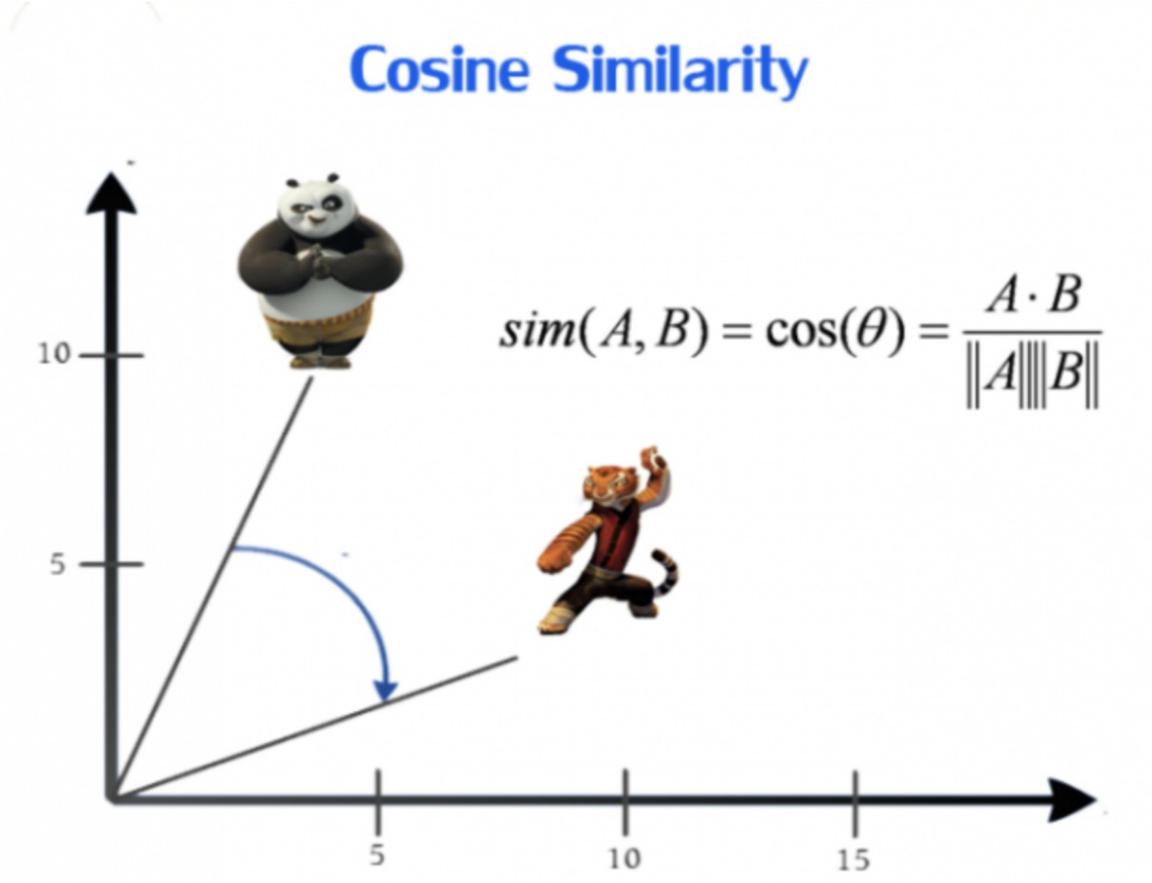
graphicx



ii. **This figure shows top 10 movie's which are similar to the users entered movie.**

Content-based filtering —Content based Filtering is to make recommendations based on similar products/and services according to their attributes.Content based recommendation engines, which is the movie recommendation system that we developed, takes content or attributes of a movies you like, for example a movies genre, cast, director, keywords etc. , and then similarity matrix is calculated, based on the ranks other movies, on how similar they are to the liked movies, in this case we rank different movies based on how similar the recommended movies are to the liked movie using something called similarity scores.

graphicx



Cosine similarity is a metric used to measure how similar two items are. Mathematically, it measures the cosine of the angle between two vectors projected in a multi-dimensional space. The output value ranges from 0–1. 0 means no similarity, where as 1 means that both the items are 100

Chapter 5

Conclusions

The model has recommended vary similar movies. From my “domain knowledge”, We can see some similarities mainly based on directors, actor’s and other plot’s. We trained and tested the recommendation system. The accuracy rate is 80provides very good prediction rate and is more reliable then the recommendation system based on collaborative filtering algorithm. Which is explained in content-based filtering (2.2) why we have used collaborative filtering? We have built the Recommendation system using content-based filtering with the of natural language processing .and conclude that the movie recommendation system which uses content-based filtering is more reliable and provides more accurate prediction and doesn’t shows a new users/item problem when new one is added