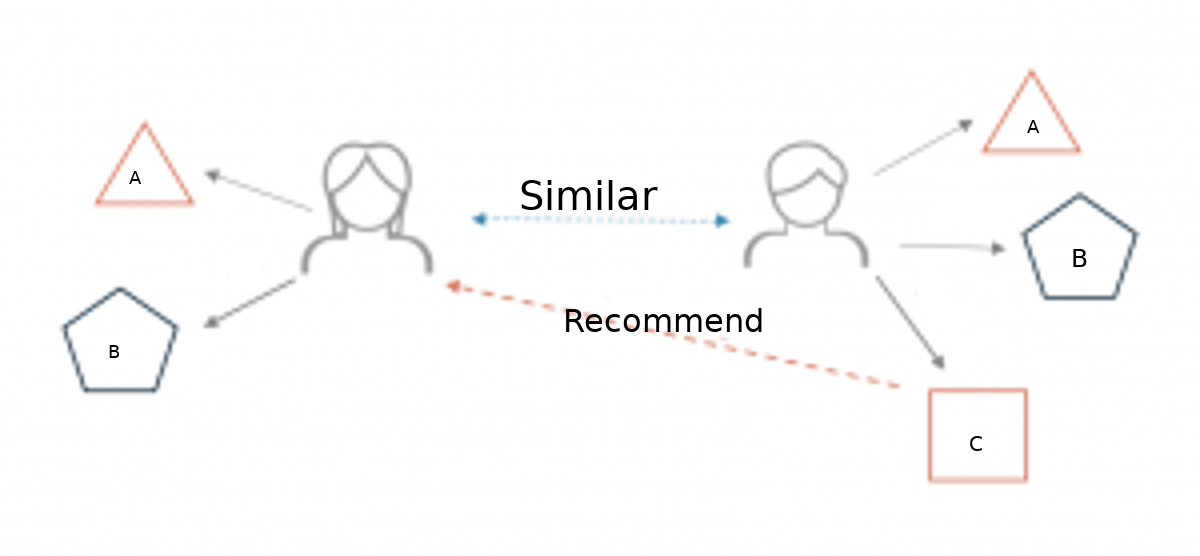
# [SYNOPSIS](https://www.wikihow.com/Write-a-Synopsis)

A Synopsis Is An In-Depth Summary Of A Written Work That Describes The Content Of That Work From Beginning To End. Unlike A Summary, Which Just Gives A General Overview Of A Story, A Synopsis Contains All Of The Plot Details, Including The End. Typically, Synopses Are Submitted To Publishers Or Agents After You Have Written A Novel, Screenplay, Or Other Long Work. A Good Synopsis Will Cover The Main Conflict And The Resolution Of The Story While Describing The Emotional Development Of The Main Character. It Is Important To Edit Your Synopsis Carefully, As It Will Usually Be Included As Part Of A Larger Proposal

## **INTRODUCTION**

A recommendation system or recommendation engine is a model used for information filtering where it tries to predict the preferences of a user and provide suggests based on these preferences. These systems have become increasingly popular nowadays and are widely used today in areas such as movies, music, books, videos, clothing, restaurants, food, places and other utilities. These systems collect information about a user's preferences and behaviour, and then use this information to improve their suggestions in the future. Movies are a part and parcel of life. There are different types of movies like some for entertainment, some for educational purposes, some are animated movies for children, and some are horror movies or action films. Movies can be easily differentiated through their genres like comedy, thriller, animation, action etc. Other way to distinguish among movies can be either by releasing year, language, director etc. Watching movies online, there are a number of movies to search in our most liked movies . Movie Recommendation Systems helps us to search our preferred movies among all of these different types of movies and hence reduce the trouble of spending a lot of time searching our favourable movies. So, it requires that the movie recommendation system should be very reliable and should provide us with the recommendation of movies which are exactly same or most matched with our preferences. A large number of companies are making use of recommendation systems to increase user interaction and enrich a user's shopping experience. Recommendation systems have several benefits, the most important being customer satisfaction and revenue. Movie Recommendation system is very powerful and important system. But, due to the problems associated with pure collaborative approach, movie recommendation systems also suffers with poor recommendation quality and scalability issues.



### Different types of recommendation engines

The most common types of recommendation systems are content-based and collaborative filtering recommender systems. In collaborative filtering, the behavior of a group of users is used to make recommendations to other users. The recommendation is based on the preference of other users. A simple example would be recommending a movie to a user based on the fact that their friend liked the movie. There are two types of collaborative models Memory-basedmethods and Model-based methods. The advantage of memory-based techniques is that they are simple to implement and the resulting recommendations are often easy to explain. They are divided into two:

* **User-based collaborative filtering**: In this model, products are recommended to a user based on the fact that the products have been liked by users similar to the user. For example, if Derrick and Dennis like the same movies and a new movie come out that Derick like, then we can recommend that movie to Dennis because Derrick and Dennis seem to like the same movies.
* **Item-based collaborative filtering**: These systems identify similar items based on users’ previous ratings. For example, if users A, B, and C gave a 5-star rating to books X and Y then when a user D buys book Y they also get a recommendation to purchase book X because the system identifies book X and Y as similar based on the ratings of users A, B, and C.

Model-based methods are based on Matrix Factorization and are better at dealing with sparsity. They are developed using data mining, machine learning algorithms to predict users’ rating of unrated items. In this approach techniques such as dimensionality reduction are used to improve accuracy. Examples of such model-based methods include Decision trees, Rule-based Model, Bayesian Model, and latent factor models.

**Content-based systems** use metadata such as genre, producer, actor, musician to recommend items say movies or music. Such a recommendation would be for instance recommending Infinity War that featured Vin Diesel because someone watched and liked The Fate of the Furious. Similarly, you can get music recommendations from certain artists because you liked their music. Content-based systems are based on the idea that if you liked a certain item you are most likely to like something that is similar to it.

### PROJECT MODEL

# **Iterative Model**



In this Model, you can start with some of the software specifications and develop the first version of the software. After the first version if there is a need to change the software, then a new version of the software is created with a new iteration. Every release of the Iterative Model finishes in an exact and fixed period that is called iteration.

The Iterative Model allows the accessing earlier phases, in which the variations made respectively. The final output of the project renewed at the end of the Software Development Life Cycle (SDLC) process.

### The various phases of Iterative model are as follows:

**1. Requirement gathering & analysis:** In this phase, requirements are gathered from customers and check by an analyst whether requirements will fulfil or not. Analyst checks that need will achieve within budget or not. After all of this, the software team skips to the next phase.

**2. Design:** In the design phase, team design the software by the different diagrams like Data Flow diagram, activity diagram, class diagram, state transition diagram, etc.

**3. Implementation:** In the implementation, requirements are written in the coding language and transformed into computer programmes which are called Software.

**4. Testing:** After completing the coding phase, software testing starts using different test methods. There are many test methods, but the most common are white box, black box, and grey box test methods.

**5. Deployment:** After completing all the phases, software is deployed to its work environment.

**6. Review:** In this phase, after the product deployment, review phase is performed to check the behaviour and validity of the developed product. And if there are any error found then the process starts again from the requirement gathering.

**7. Maintenance:** In the maintenance phase, after deployment of the software in the working environment there may be some bugs, some errors or new updates are required. Maintenance involves debugging and new addition options.

### OBJECTIVE

The primary goal of movie recommendation systems is to filter and predict only those movies that a corresponding user is most likely to want to watch. The ML algorithms for these recommendation systems use the data about this user from the system's database.

Excellent film analysis will explain how a film has been made: which filmmaking techniques have been chosen and why, how the visual storytelling supports the narrative, and the effect that filmmaking elements have on the viewer.

### SCOPE

* **Work on several numbers of data:**

The number of choices for anything on internet is very high and it’s tedious to refine most wanted data by self while searching. The scope of this proposal system includes working within numerous data, with ease.

* **Saving of time:**

Many people have problem selecting the alternative item of movie due to lack of time and due to search issues. Also movie recommendations from friends can be time consuming. The system helps in saving lots of time.

### Methodology

* Collecting the data sets : Collecting all the required data set from Kaggle web site.in this project we require movie.csv,ratings.csv,users.csv.
* Data Analysis : make sure that that the collected data sets are correct and analysing the data in the csv files. i.e. checking whether all the column Felds are present in the data sets.
* Algorithms : in our project we have only two algorithms one is cosine similarity and other is single valued decomposition are used to build the machine learning recommendation model.
* Training and Testing the model : once the implementation of algorithm is completed . we have to train the model to get the result. We have tested it several times the model is recommend different set of movies to different users.

### HARDWARE AND SOFTWARE REQUIREMENT

* HARDWARE REQUIREMENTS:-

The most common set of requirements defined by any [operating system](https://en.wikipedia.org/wiki/Operating_system) or [software application](https://en.wikipedia.org/wiki/Software_application) is the physical computer resources, also known as [hardware](https://en.wikipedia.org/wiki/Computer_hardware), A hardware requirements list is often accompanied by a [hardware compatibility list](https://en.wikipedia.org/wiki/Hardware_compatibility_list) (HCL), especially in case of operating systems. An HCL lists tested, compatible, and sometimes incompatible hardware devices for a particular operating system or application. The following sub-sections discuss the various aspects of hardware requirements.

Hardware Requirements :

• A PC with Windows/Linux OS

• Processor with 1.7-2.4gHz speed

• Minimum of 8gb RAM

• 2gb Graphic card

Software Specification :

• Text Editor (VS-code/WebStorm)

• PyCharm Editor

• Python libraries

* SOFTWARE REQUIREMENTS:-

[Software requirements](https://en.wikipedia.org/wiki/Software_requirements) deal with defining software resource requirements and prerequisites that need to be installed on a computer to provide optimal functioning of an application. These requirements or prerequisites are generally not included in the software installation package and need to be installed separately before the software is installed.

**Python libraries:** For the computation and analysis we need certain python libraries which are used to perform analytics. Packages such as SKlearn, Numpy, pandas, Matplotlib are needed.

* **SKlearn:** It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy.

* **NumPy:** NumPy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays. It is the fundamental package for scientific computing with Python. Pandas: Pandas is one of the most widely used python libraries in data science.

### MODULE DESCRIPTION

The proposed system consists of two modules :-

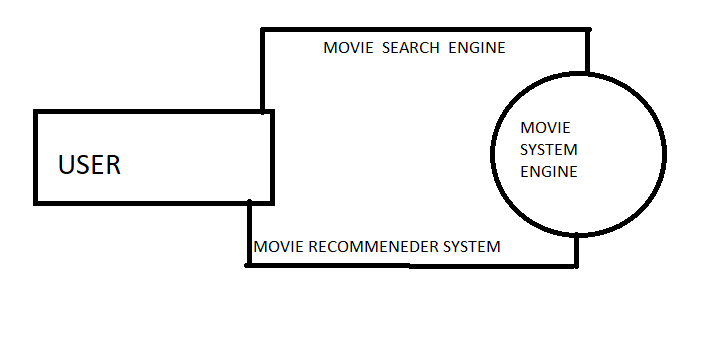
* API MODULE
* SEARCH MODULE
* API MODULE : I made two function , first function fetches poster from TDMN database and second function is fetches video key from TDMN database. Video key is used to you tube url.In the project, YouTube url is used for to recomend which type of movie you want.
* SEARCH MODULE : This module searches the movie which the user types and lists the details about that particular movie like its popularity, rating ,trailer and many more.

### DESIGN OF SOLUTION

* DFD (DATA FLOW DAIGRAM)

A data flow diagram (DFD) is a graphical or visual representation using a standardized set of symbols and notations to describe a business's operations through data movement. They are often elements of a formal methodology such as Structured Systems Analysis and Design Method ([SSADM](https://www.techtarget.com/searchsoftwarequality/definition/SSADM)). Superficially, DFDs can resemble flow charts or Unified Modeling Language ([UML](https://www.techtarget.com/searchsoftwarequality/definition/Unified-Modeling-Language)), but they are not meant to represent details of software logic.

* + - * DFD Level 0 is also called a Context Diagram. It’s a basic overview of the whole system or process being analyzed or modeled.



* + - * DFD Level 1 provides a more detailed breakout of pieces of the Context Level Diagram. You will highlight the main functions carried out by the system

USER

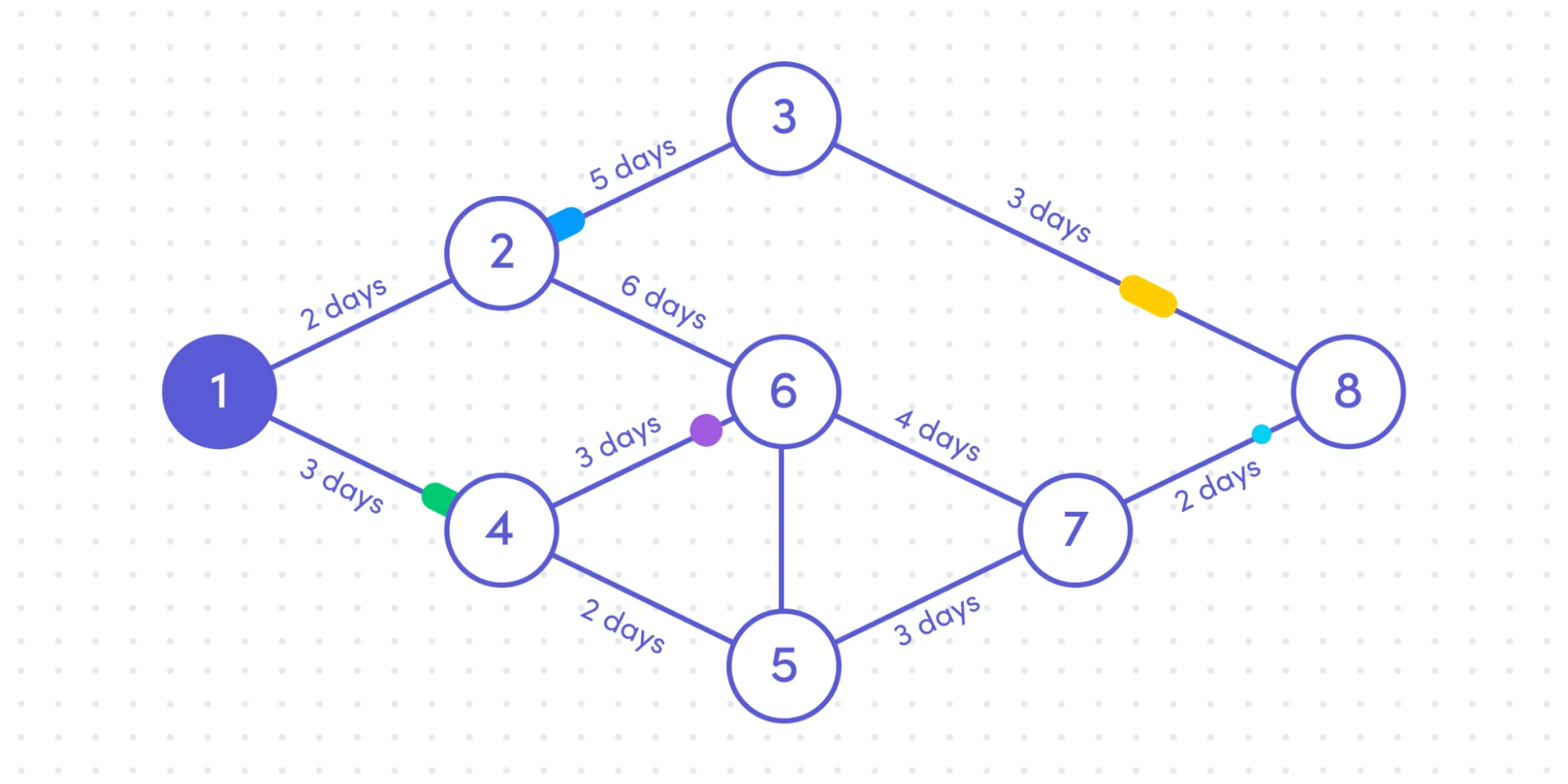
### Planning and Scheduling

The objective of software project planning is to provide a framework that enables the manger to Make reasonable estimates of resources , cost and schedules . these estimate are made within a Limited time frame at the beginning of a software project and should be update regulary as the Project progresses . in addition , estimates should attempt tp define best case and worst caseScenarios sp thst project outcomes can be bounded.

* Pert chart

A PERT chart, also known as a PERT diagram, is a tool used to schedule, organize, and map out tasks within a project.

PERT stands for program evaluation and review technique. It provides a visual representation of a project's timeline and breaks down individual tasks.



* GANTT CHART

A Gantt chart is a project management tool that illustrates work completed over a period of time in relation to the time planned for the work. A Gantt chart can include the start and end dates of tasks, milestones, dependencies between tasks, assignees, and more.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| TASK | 16-18 oct | 18-21 oct | 21-26 oct | 26-30 oct | 30-1 oct -nov |
| DEVELP PROJECT | 2 days |  |  |  |  |
| Analysis |  | 2 days |  |  |  |
| Designing |  | 2 days |  |  |  |
| Coding |  |  | 5 days |  |  |
| Unit Testing |  |  |  | 1 days |  |
| Beta Testing |  |  |  |  | 2 days |

### Future Scope

* In the proposed approach, It has considered Genres of movies but, in future we can also consider age of user as according to the age movie preferences also changes, like for example, during our childhood we like animated movies more as compared to other movies. There is a need to work on the memory requirements of the proposed approach in the future. The proposed approach has been implemented here on different movie datasets only. It can also be implemented on the Film Affinity and Netflix datasets and the performance can be computed in the future.
* Studying more Content-based approach and their scalability
* Decrease the RMSE
* Impment Hybrid CF techniques and run on larger data sets

.

### Conclusion

* Cosine Similarities and Person Coefficient Correlation overall produce similar result Pearson Coefficient is only slightly better .
* In ALS , as expected the time taken increase with the size of the data and the RMSE decreases with increase in the size of the data.
* In all the methods studied , time taken on cluster is considerably less than the time taken on the local machine

### 