

DEPARTMENT OF COMPUTER ENGINEERING AND APPLICATIONS



INDUSTRIAL PROJECT ON MOVIE RECOMMENDATION SYSTEM

Presented by: SHIVAM KUMAR
Univ. Roll NO.:171500320
GLA UNIVERSITY , Mathura

Contents:

1. About the company
2. Introduction to Machine Learning.
3. Types of Machine Learning.
4. Workflow Diagram.
5. What is Recommendation System
6. How Recommendation System Works.
7. Recommendation Filtering Techniques.
8. Algorithms.
9. The MovieLens Data Set.
- 10.Steps for making a recommendation system.
11. Calculating Results.
- 12.Conclusion.

About The Company:



CETPA INFOTECH PVT. LTD.

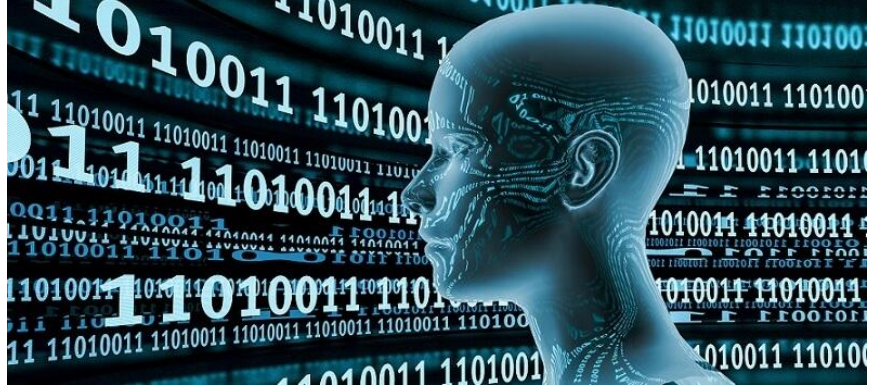
ADDRESS - D-58, Red FM Road, Sector 2, D Block,
Sector 2, Noida, Uttar Pradesh 201301

Development -

CETPA development deals in software as well as embedded production development.

CETPA INFOTECH is the training service provider in various engineering domains for all engineering students as well as for the working professionals.

WHAT IS MACHINE LEARNING:



Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves

For example, medical diagnosis, image processing, prediction, classification, learning association, regression etc. The intelligent systems built on machine learning algorithms have the capability to learn from past experience or historical data.

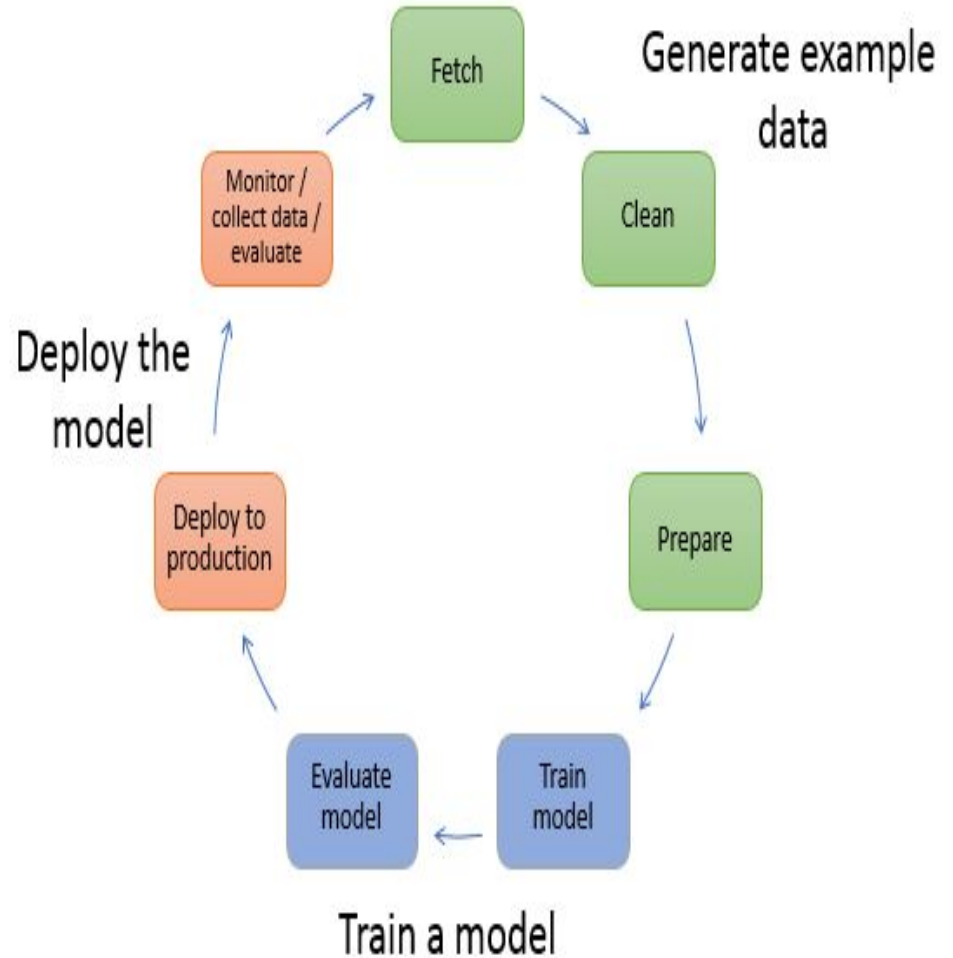
Types of Machine Learning:



*In Supervised learning, you train the machine using data which is well "**labeled**." It means some data is already tagged with the correct answer. It can be compared to learning which takes place in the presence of a supervisor or a teacher.

*Unsupervised learning is a machine learning technique, where you do not need to supervise the model. Instead, you need to allow the model to work on its own to discover information. It mainly deals with the unlabelled data.

WorkFlow Diagram:



What is a Recommendation System?

What is a Recommendation System?

Recommendation system is an information filtering technique, which provides users with information, which he/she may be interested in.

Examples:



Video-on-demand provider in North America and UK

- Matches 23 million customers with a huge inventory of movies according to their tastes
- 60-70% of views result from the recommendations⁴



Gold standard of e-commerce. Pioneer in using recommendations

- Sits on a huge volume of collective information of its customers
- Customers can view what people with similar tastes viewed or purchased
- Customers can ask the recommendations engine to ignore selected purchases



Social and professional networking sites

- Sits on a huge volume of collective information of its customers
- Customers can view what people with similar tastes viewed or purchased
- Customers can ask the recommendations engine to ignore selected purchases



Music station. Offers music suggestions based on ratings

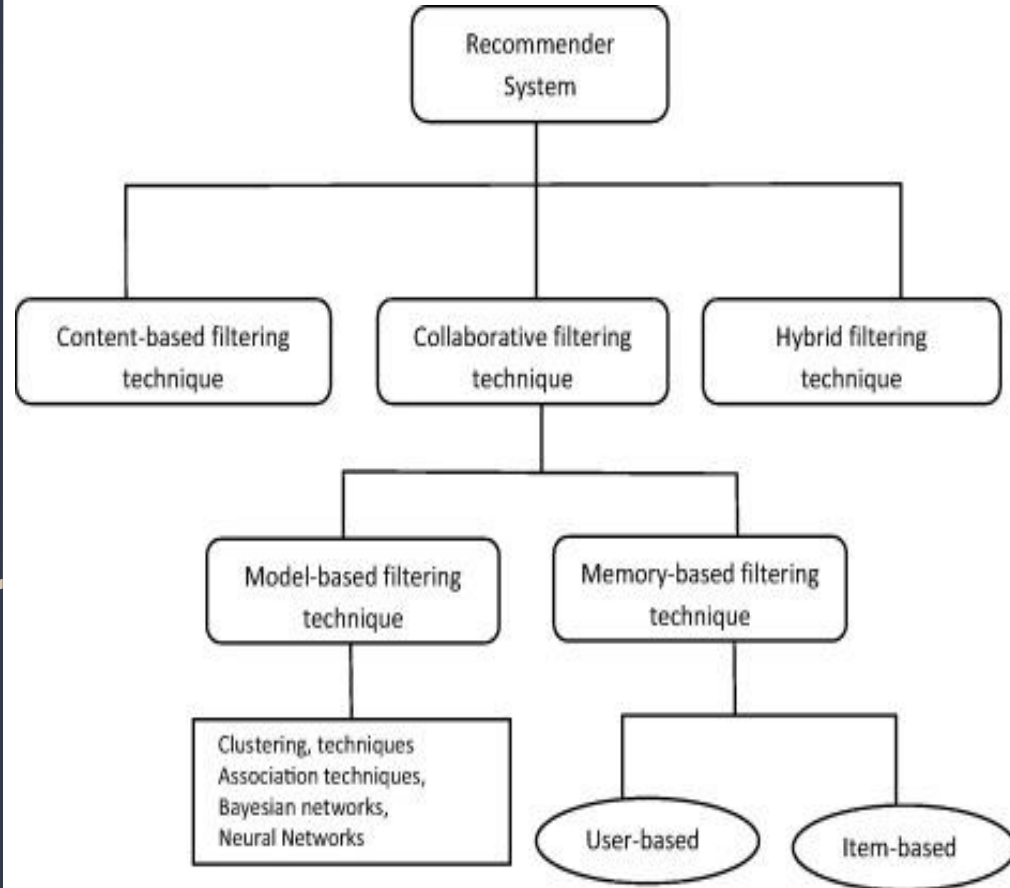
- Sits on a huge volume of collective information of its customers
- Customers can view what people with similar tastes viewed or purchased
- Customers can ask the recommendations engine to ignore selected subscriptions³

How Recommendation System Works :

* **Recommendation engines** basically are data filtering tools that make use of algorithms and data to recommend the most relevant items to a particular user. Or in simple terms, they are nothing but an automated form of a “shop counter guy”. You ask him for a product.

* Such systems are recommending items similar to those a given user has liked in the past, regardless of the preferences of other users. ...

Recommendation Filtering Techniques:

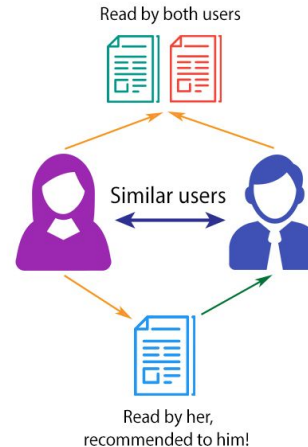


ALGORITHMS :

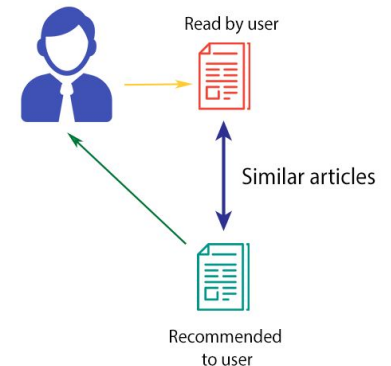
1. Collaborative Filtering: It makes recommendations for a user based on ratings and preferences data of many users.

2. Content based Recommendations: It takes into account the likes and dislikes of the user and generates a User Profile. It works solely with the past interactions of a given user and do not take other users into consideration.

COLLABORATIVE FILTERING



CONTENT-BASED FILTERING



The Movie Lens Dataset:

The dataset that I'm working with is [MovieLens](#), one of the most common datasets that is available on the internet for building a Recommender System. The version of the dataset that I'm working with ([1M](#)) contains 1,000,209 anonymous ratings of approximately 3,900 movies made by 6,040 MovieLens users who joined MovieLens in 2000.

Importing Python Libraries in Jupyter Notebook:



1. Numpy:

NumPy is the fundamental package for scientific computing in **Python**. ... **NumPy** arrays facilitate advanced mathematical and other types of operations on large numbers of data.



2. Pandas:

Pandas is the most popular **python library** that is used for data analysis.



Reading the CSV file using Pandas Library :

```
# Import pandas
import pandas as pd

# reading csv file
pd.read_csv("filename.csv")
```

Now let's get the movie titles:

```
In [4]: movie_titles = pd.read_csv("Movie_Id_Titles.csv")
movie_titles.head()
```

Out[4]:

	item_id	title
0	1	Toy Story (1995)
1	2	GoldenEye (1995)
2	3	Four Rooms (1995)
3	4	Get Shorty (1995)
4	5	Copycat (1995)

We can merge them together:

```
In [5]: df = pd.merge(df, movie_titles, on='item_id')
df.head()
```

Out[5]:

	user_id	item_id	rating	timestamp	title
0	196	242	3	881250949	Kolya (1996)
1	63	242	3	875747190	Kolya (1996)
2	226	242	5	883888671	Kolya (1996)
3	154	242	3	879138235	Kolya (1996)
4	306	242	5	876503793	Kolya (1996)

Data Visualization Using Matplotlib :

Exploring the Data a bit and get a look at some best rated movies by creating a ratings dataframe with average ratings and number of ratings.

Visualization Imports

```
In [6]: import matplotlib.pyplot as plt
import seaborn as sns
sns.set_style('white')
%matplotlib inline
```

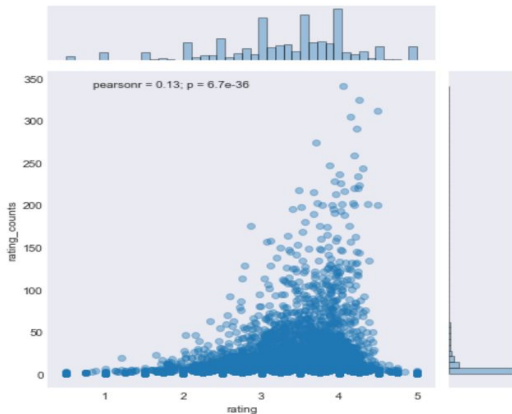
Let's create a ratings dataframe with average rating and number of ratings:

```
In [7]: df.groupby('title')['rating'].mean().sort_values(ascending=False).head()
```

```
Out[7]: title
Marlene Dietrich: Shadow and Light (1996)    5.0
Prefontaine (1997)                          5.0
Santa with Muscles (1996)                   5.0
Star Kid (1997)                             5.0
Someone Else's America (1995)              5.0
Name: rating, dtype: float64
```

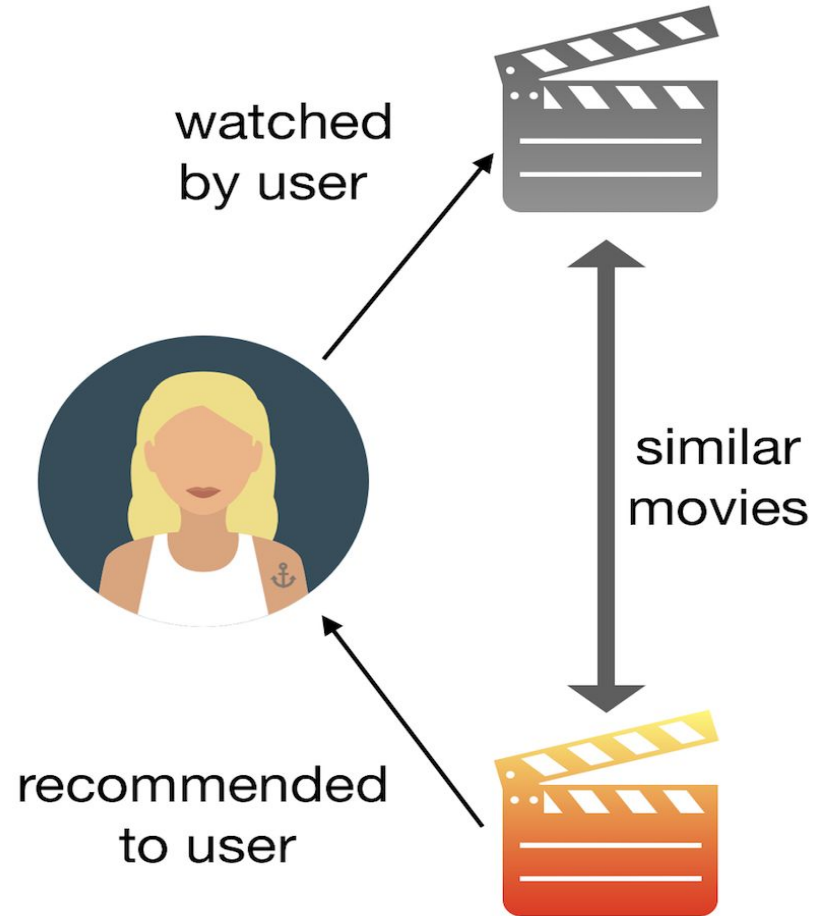
```
In [8]: df.groupby('title')['rating'].count().sort_values(ascending=False).head()
```

```
Out[8]: title
Star Wars (1977)          583
Contact (1997)            509
Fargo (1996)              508
Return of the Jedi (1983) 507
Liar Liar (1997)          485
Name: rating, dtype: int64
```



CALCULATING RESULTS:

Movies having similar Ratings.



Conclusion:

Recommender systems open new opportunities of retrieving personalized information on the Internet. It also helps to alleviate the problem of information overload which is a very common phenomenon with information retrieval systems and enables users to have access to products and services which are not readily available to users on the system.



**THANK
YOU**

FOR

LISTENING TO

MY PRESENTATION