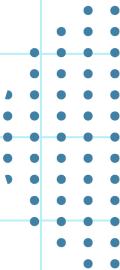
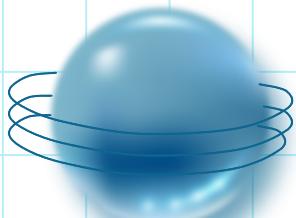


6TH GRADE

MOVIE
RECOMMENDATION
ENGINE

Table of contents

- 1. INTRODUCTION**
- 2. DATASET**
- 3. PROPOSED SYSTEM**
- 4. SYSTEM DESIGN**
- 5. ALGORITHMS**
- 6. SYSTEM REQUIREMENT**
- 7. RESULT**
- 8. CODING**
- 9. SNAPSHOT**
- 10. CONCLUSION**
- 11. FUTURE WORK**
- 12. REFERENCE**



ABSTRACT

This project is planned by including a machine learning approach to suggest movies to the users using K-Means clustering algorithm, K Nearest neighbours algorithm and Affinity propagation clustering algorithm to recommend movies to the users

INTRODUCTION *

In our project, by exploring different Machine learning algorithm such as K-Means clustering algorithm, K Nearest Neighbors algorithm and Affinity propagation clustering algorithm, we recommend top 20 movies to users based on the rating given by users to the movies.

DATASET

The Movie Lens dataset contains a movie list of nineteen genres. The opposite file consists of: user id, movie id, ratings. These two files are pre-processed and manipulated therefore to produce our system.



PROPOSED SYSTEM

In this project of Movie recommendation system using machine learning algorithms we will be exploring different Machine learning algorithm such as K-Means clustering algorithm, K Nearest Neighbors algorithm and Affinity propagation clustering algorithm we recommend top 20 -movies to users based on therating given by users to the movies.

SYSTEM DESIGN

In our system, the primary and foremost is knowledge gathering the dataset are pre-processed and provided to following steps. Now, the data is ready for applying the desired algorithmic rule, here we use K-Means clustering algorithm, K nearest Neighbors algorithm and Affinity propagation clustering algorithm. Once after applying the algorithm separately for the dataset we will predict and recommend 20 movies for the users who are yet to watch the movie, and finally we conclude by comparing the results of each algorithm.

ALGORITHMS

K MEANS CLUSTERING ALGORITHMS

The algorithm works as follows:

1. First we initialize k points, called means, randomly.
2. We categorize each item to its closest mean and we update the mean's coordinates, which are the averages of the items categorized in that mean so far.
3. We repeat the process for a given number of iterations and at the end, we have our clusters.

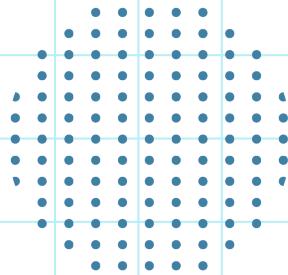


K NEAREST NEIGHBOR ALGORITHMS

KNN is a machine learning algorithm to find clusters of similar users based on co ratings, and make predictions using the average rating of top-k nearest neighbors.

NEAREST NEIGHBORS ALGORITHMS

- * **BRUTE FORCE**
- * **K-D TREE**
- * **BALL TREE**



SYSTEM REQUIREMENT

HARDWARE

- System : i5 Processor
- Ram : 16 GB

SOFTWAARE

- Operating system : Windows 10
- Tool: Anaconda Navigator – 64bit
- Scripting Tool: Jupyter Notebook

JUPYTER NOTEBOOK

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

- Language: Python3.8

PYTHON

Python's design philosophy emphasizes code readability with its notable use of significant whitespace. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python is dynamically typed and garbage-collected. It supports multiple programming paradigms, including procedural, object-oriented, and functional programming.

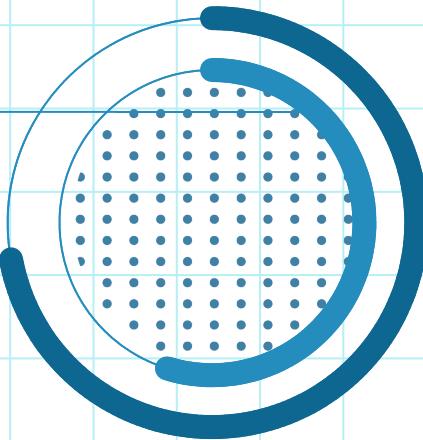


IMPORTANT PERCENTAGES

53%

JUPITER

Jupiter is a gas giant and
the biggest planet



60%

SATURN

Saturn is the second-largest
planet in the Solar System

RESULTS



The dataset used is Rate.csv and Movie.csv which is taken from MovieLens. This dataset consists of 100004 ratings given by 671 users for 9077 movies, each user rated nearly minimum of 20 movies. The below image shows the movies recommended to the users. After the result we check which system is more efficient for the recommendation by comparing the result



snapshot

The screenshot shows a Jupyter Notebook interface with three tabs: 'Home', 'flix.flow', and 'Home'. The 'flix.flow' tab is active and displays the following content:

jupyter flix.flow Last Checkpoint: 10 days ago

File Edit View Run Kernel Settings Help Trusted

```
[24]: movies['crew']=movies['crew'].apply(collapse)
movies['cast']=movies['cast'].apply(collapse)
movies['genres']=movies['genres'].apply(collapse)
movies['keywords']=movies['keywords'].apply(collapse)
```

```
[25]: movies.head(15)
```

	movie_id	title	overview	genres	keywords	cast	crew
0	19995	Avatar	[In the, 22nd, century, a, paraplegic, Marin...]	[Action, Adventure, Fantasy, ScienceFiction]	[cultureclash, future, spacewar, spacecolony, ...]	[SamWorthington, ZoeSaldana, SigourneyWeaver]	[JamesCameron]
1	285	Pirates of the Caribbean: At World's End	[Captain, Barbossa, long, believed, to, be, d...]	[Adventure, Fantasy, Action]	[ocean, drugabuse, exoticisland, eastindiatriad...]	[JohnnyDepp, OrlandoBloom, KeiraKnightley]	[GoreVerbinski]
2	206647	Spectre	[A, cryptic, message, from, Bond's, past, send...]	[Action, Adventure, Crime]	[spy, basedonnovel, secretagent, sequel, mib, ...]	[DanielCraig, ChristophWaltz, LéaSeydoux]	[SamMendes]
3	49026	The Dark Knight Rises	[Following, the, death, of, District, Attorney...]	[Action, Crime, Drama, Thriller]	[dcomics, crimefighter, terrorist, secretiden...]	[ChristianBale, MichaelCaine, GaryOldman]	[ChristopherNolan]
4	49529	John Carter	[John, Carter, is, a, war-weary, former, mil...]	[Action, Adventure, ScienceFiction]	[basedonnovel, mars, medallion, spacetravel, p...]	[TaylorKitsch, LynnCollins, SamanthaMorton]	[AndrewStanton]
5	559	Spider-Man 3	[The, seemingly, invincible, Spider-Man, goes,...]	[Fantasy, Action, Adventure]	[dualentity, amnesia, sandstorm, loveofone's...]	[TobeyMaguire, KirstenDunst, JamesFranco]	[SamRaimi]
6	38757	Tangled	[When, the, kingdom's, most, wanted-and-most...]	[Animation, Family]	[hostage, magic, horse, fairytale, musical, pr...]	[ZacharyLevi, MandyMoore, DonnaMurphy]	[ByronHoward, NathanGreno]
7	99861	Avengers: Age of Ultron	[When, Tony, Stark, tries, to, jumpstart, a, d...]	[Action, Adventure, ScienceFiction]	[marvelcomic, sequel, superhero, basedoncomics...]	[RobertDowneyJr., ChrisHemsworth, MarkRuffalo]	[JossWhedon]

A screenshot of a Jupyter Notebook interface running in a browser window titled "flix.flow". The notebook URL is "localhost:8889/notebooks/flix.flow.ipynb". The code cell contains the following Python code:

```
[34]: from sklearn.metrics.pairwise import cosine_similarity
[35]: similarity = cosine_similarity(vector)
[36]: similarity
[37]: new[new['title'] == 'The Lego Movie'].index[0]
[38]: def recommend(movie):
    index = new[new['title'] == movie].index[0]
    distances = sorted(list(enumerate(similarity[index])), reverse=True, key=lambda x: x[1])
    for i in distances[1:5]:
        print(new.iloc[i[0]].title)
[39]: recommend('Avatar')
Gandhi, My Father
The Wind That Shakes the Barley
```

The code imports the cosine_similarity function from sklearn.metrics.pairwise, calculates the similarity matrix for the vector, prints it, finds the index of 'The Lego Movie' in the new dataset, defines a recommend function that takes a movie title as input, finds the indices of the top 5 most similar movies, and prints their titles. Finally, it calls the recommend function for 'Avatar'.



11:22
17-05-2024

CONCLUSION

Our system will calculate the common in predicting rating from user information which may be used to analyze that movie ought to suggest to new users using three Machine Learning Algorithm

FUTURE WORK

Neural Networks and Deep Learning have been all the rage the last couple of years in many different fields, and it appears that they are also helpful for solving recommendation system problems.