

映画コンテンツ推奨システムに関する研究

平石研究室

S20534

RUMON_MD

目次

- レコメンデーションとは
- 推薦手法の分類
- 研究の内容
- データの前処理段階
- 属性選んだ理由
- 実行動画
- 関連研究
- 終わりに

Welcome.

Millions of movies, TV shows and people to discover. Explore now.

Search for a movie, tv show, person.....

Search

Trending

Today

This Week



Fast X

<https://www.themoviedb.org/movie/840326>



The Super Mario Bros. Movie



White Men Can't Jump



Sisu

Jan 27, 2023



Ant-Man and the Wasp: Quantumania



Asterix & Obelix: The Middle Kingdom



Dungeons & Dragons: Honor Among Thieves

はじめに

- 映画レコメンデーションシステムは、何らかのデータセットに基づいてユーザーに映画の提案を提供するシステムであり。また、ユーザーに対して彼らの好みや興味に基づいて映画を提案するシステムのことです。これは、ユーザーが視聴する可能性の高い映画を見つけるのに役立ちます。

例えば、ユーザーが自分の好きな映画あるいは自分の好きな俳優の映画を見ると、この推薦システムは同じジャンルの映画や、同じ俳優が出演する映画、またはその両方を推奨します。

推薦手法の分類:

- コンテンツベースのフィルタリング: 映画の属性や特徴(ジャンル、監督、出演者など)を分析し、ユーザーの好みとの一致度を評価して映画を推薦します。
- 協調フィルタリング: 複数のユーザーの映画視聴履歴や評価データを収集し、その情報をもとに似た傾向や興味を持つユーザーを特定し、彼らが視聴した映画を推薦します。
- ハイブリッドフィルタリング: 複数の手法やアルゴリズムを組み合わせる手法です。たとえば、コンテンツベースのフィルタリングと協調フィルタリングを組み合わせることで利点を結集することがあります。

研究の内容:

- 本研究ではコンテンツベースのフィルタリングシステム手法で5000映画のデータセットを使います。映画の属性としては、映画のID,ジャンル、キーワード、タイトル、キャスト、出演者などデータを収集します。そして、ユーザーの検索クエリや視聴履歴などから、ユーザーの好みや嗜好を反映した映画の属性との間で類似性を計算し、類似性の計算結果をもとに、映画をランキング付けし、ユーザーに推薦します。

データ前処理の段階

- この段階では収集したメタデータセットから必要なデータを分析し、データの前処理段階を行います。

```
movies.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
Int64Index: 4809 entries, 0 to 4808  
Data columns (total 7 columns):  
#   Column      Non-Null Count  Dtype  
---  ---  
0   movie_id    4809 non-null   int64  
1   title       4809 non-null   object  
2   overview    4806 non-null   object  
3   genres      4809 non-null   object  
4   keywords    4809 non-null   object  
5   cast        4809 non-null   object  
6   crew        4809 non-null   object  
dtypes: int64(1), object(6)  
memory usage: 300.6+ KB
```

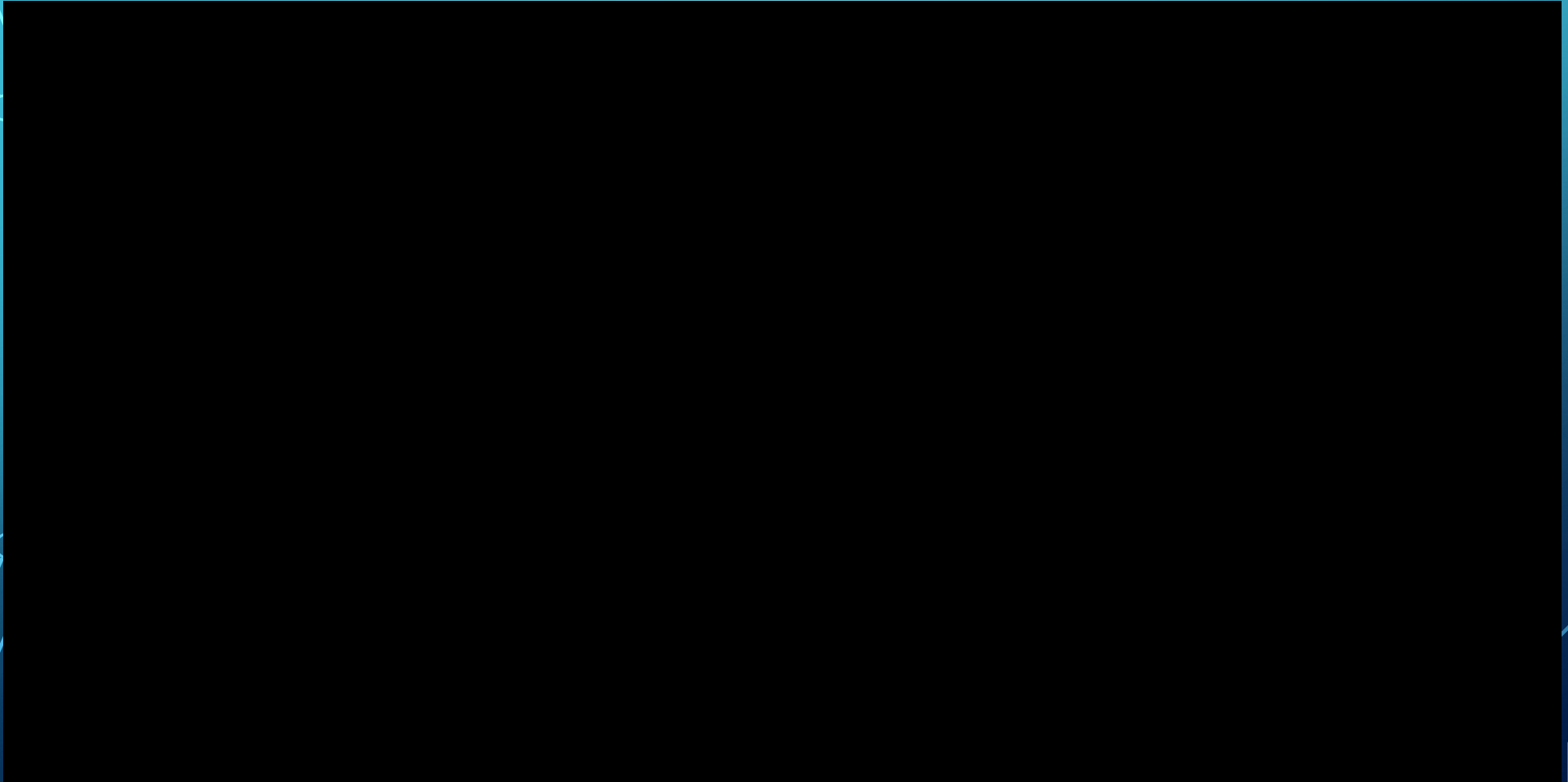
```
movies.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
Int64Index: 4809 entries, 0 to 4808  
Data columns (total 23 columns):  
#   Column      Non-Null Count  Dtype  
---  ---  
0   budget      4809 non-null   int64  
1   genres      4809 non-null   object  
2   homepage    1713 non-null   object  
3   id          4809 non-null   int64  
4   keywords    4809 non-null   object  
5   original_language  4809 non-null   object  
6   original_title  4809 non-null   object  
7   overview    4806 non-null   object  
8   popularity  4809 non-null   float64  
9   production_companies  4809 non-null   object  
10  production_countries  4809 non-null   object  
11  release_date  4808 non-null   object  
12  revenue       4809 non-null   int64  
13  runtime       4807 non-null   float64  
14  spoken_languages  4809 non-null   object  
15  status        4809 non-null   object  
16  tagline       3965 non-null   object  
17  title         4809 non-null   object  
18  vote_average  4809 non-null   float64  
19  vote_count    4809 non-null   int64  
20  movie_id     4809 non-null   int64  
21  cast         4809 non-null   object  
22  crew         4809 non-null   object  
dtypes: float64(3), int64(5), object(15)  
memory usage: 901.7+ KB
```

なぜデータの前処理段階では7つ属性を選んだ

- 5000本映画のデータセットには属性として映画の様々な情報を含まれており、その中から映画のID、タイトル、あらすじ、ジャンル、キーワード、キャスト、クルーの7つの属性を選んだ。映画のID属性を使用することで、特定の映画を一意に特定することができる。そして、ジャンル情報を使用することで、特定のジャンルの映画を抽出したりすることができる。キーワード情報を使用することで、特定のテーマや要素を持つ映画を検索したり、関連する映画を推薦したりすることができる。これらの属性は、映画の基本的な情報や内容、関連要素、制作に関わる人々に焦点を当てており、映画の分析や検索、推薦などの目的に適している。

実行動画：



終わりに

- 本研究では5000本の映画データを収集して、データの前処理を実施した。データの前処理では映画データベースやメタデータから必要なデータの収集を行なった。
- 今後、前処理ステージから次のデータモデリングステージを実施し、自分ローカルホストでウェブサイトを作成し実施を行い。

関連研究

- 参考文献1:
 - ミュージックレコメンダーシステムに関して行った研究があり、ミュージックレコメンダーシステムとは、ユーザーの最近の視聴履歴から学習し、彼らがおそらくそうするだろう曲をお勧めし、将来的に聞きたいことで表す。音楽レコメンダーシステムを使用することにより、音楽配信者は、配信した曲を再生して提供することができる。ユーザーは、かつて聞いたことのある音楽のことである。
- 参考文献2:
 - 機械学習アルゴリズムを用いた映画推薦システムに行った研究があり、映画推薦システムは、協調フィルタリングアプローチに基づいている。

関連研究の流れ：

- ユーザーから提供された情報の使用、分析し、次に、ユーザーに最適な映画を勧める。その時、おすすめの映画のリストは、以前のユーザーがこれらの映画に付けた評価と、この目的のためのさまざまな機械学習アルゴリズム使う
- 参考文献 3:
- 機械学習を利用した図書推薦システム：
- 書籍レコメンデーションシステムを使用して行った研究があり、ユーザーはウェブ上で利用可能な多数のオプションから書籍を検索して選択でき、または他の場所の電子情報源。それらは、オブジェクトの大きなグループとユーザーのニーズの説明が与えられた場合に、説明に適合する製品の選択をユーザーに提供する。このシステムは、推奨事項を提供するだけであり、おすすめは、購入、習慣、レビュー、いいね など、以前のユーザー アクティビティに基づいている。



ご清聴ありがとうございました。