

Film Öneri Sistemi

Movie Recommendation System

1. Miran Emre ESER
Yazılım Mühendisliği
Kocaeli Üniversitesi
Kocaeli, Türkiye
210229033
esermiran1@gmail.com

Özetçe—Bu proje, film izleme alışkanlıklarını analiz ederek kullanıcıya özel film önerileri sunmayı hedeflemektedir. Veri madenciliğinde kullanılan Apriori algoritması ile sık izlenen film kümeleri ve kullanıcı tercihleri arasındaki ilişki kuralları çıkarılmıştır. Tasarlanan kullanıcı arayüzü (GUI) ile kullanıcıların belirlenen popüler filmler veya kişisel seçimlerine dayalı önerilere, film türüne ya da film adına göre erişim sağlanmaktadır.

Anahtar Kelimeler — Veri Madenciliği, Apriori algoritması, GUI, birliktelik kuralları, Film Öneri Sistemi

Abstract—This project aims to analyze movie-watching habits and provide personalized movie recommendations to users. Using the Apriori algorithm, commonly applied in data mining, frequent movie sets and association rules among user preferences were derived. Through the designed Graphical User Interface (GUI), users can access recommendations based on popular movies or their personal selections, categorized by genre or movie title.

Keywords —Data Mining, Apriori Algorithm, GUI, Association Rules, Movie Recommendation System

I. GİRİŞ

Bu çalışmada, kullanıcıların izleme alışkanlıklarına ve beğenilerine dayanarak daha doğru ve etkili, kişiselleştirilmiş film önerileri sunmak için bu sistem geliştirilmiştir. Apriori algoritması kullanılarak, sık izlenen film kümeleri ve aralarındaki ilişkiler belirlenmiştir. Bu projenin amacı, kullanıcıların beğeni ve tercihlerine göre en uygun film önerilerini sunmak ve izleyici deneyimini daha keyifli hale getirmektir.

II. MATERYAL VE YÖNTEMLER

1. Sistem Tasarımı

a) Veri Seti

Projede, film derecelendirme ve izleme alışkanlıklarını içeren MovieLens 20M veri seti kullanılmıştır. Bu veri seti, kullanıcılar arasındaki beğeni ve izleme benzerliklerini analiz etmek için gerekli bilgiyi sağlamaktadır.

b) Veri Seti Temizliği

- En çok bulunan 10 kategori alınmış bunun dışındaki filmler silinmiştir.
- 3.5 altındaki reytingler silinmiştir. Bu sayede hem sevilen filmler hem de herkesin beğendiği filmler kalmıştır.
- 800'den az film izleyen kullanıcılar silinmiştir.
- 200'den az izlenen filmler silinmiştir.

c) Apriori Algoritması

Apriori algoritması, belirli bir filmi izleyen kullanıcıların diğer filmi izleme olasılıklarını belirlemek için kullanılmıştır. Algoritma, film kümelerinin birlikte izlenme sıklığını ve aralarındaki ilişkiyi destek, güven ve lift gibi değerlerle analiz ederek öneri sistemine entegre edilmiştir. Bu yöntemlerle elde edilen birliktelik kuralları, kullanıcı tercihleri doğrultusunda daha doğru ve etkili öneriler sunmak için kullanılmıştır.

d) Pandas Kütüphanesi

Veri setinin yüklenmesi, temizlenmesi ve işlenmesi için kullanılmıştır.

e) Rule Tree

Popüler türe göre önerim yaparken rulesların fazlalığından dolayı öneri süresi oldukça uzamaktaydı. Bu yüzden apriori algoritmasıyla elde edilen birliktelik kuralları, öneri sistemi içerisinde hiyerarşik bir yapı olan rule tree de tutulmakta. Diğer öneri sistemleri oldukça hızlı çalıştığından dolayı ruleslar direkt çekilmiştir.

Rule tree, rulesların ağaç yapısında saklanması sayesinde arama hızını arttırmakta ve ilişkili film kümelerinin daha verimli bir şekilde yönetilmesini sağlamakta. Bu yüzden rule tree ağacını kullanılmakta.

Popüler türe göre arama yaptığımızda rule tree kullanılmadan öneri yapılma süresi 48 sn sürerken, rule tree kullanarak arama süresi 0.2 sn'ye indirildi.

```
Romance türü için önerilen film: frozenset({'356'})
Seçim: Romance

***

İşlem süresi: 46.56872 saniye

Name: 988984, dtype: object
Önerilen film: frozenset({'356'}), Destek: 0.962134251290877
İşlem süresi: 0.20577 saniye
```

2. Kullanıcı Arayüzü (GUI) Tasarımı

a) Tkinter Kütüphanesi

Kullanıcı arayüzü (GUI) tasarımında, öneri sistemi için kullanıcı dostu ve etkileşimli bir ortam sağlamak amacıyla kullanılmıştır.

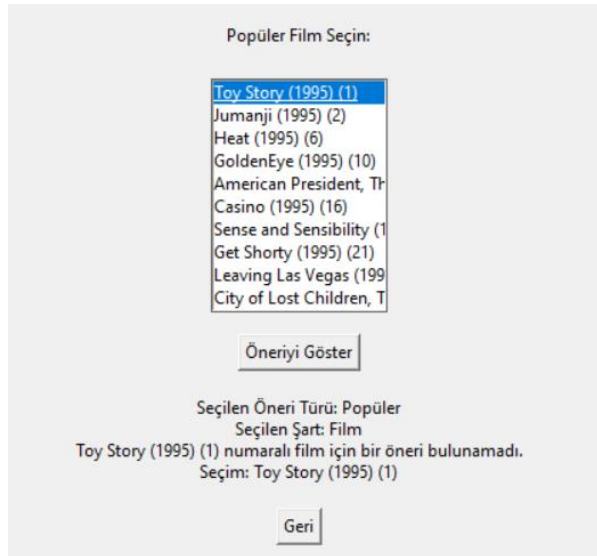
b) Kullanılan Bileşenler

- Buton :** Kullanıcılara, öneri türünü (Popüler Film Önerileri veya Kişiselleştirilmiş Film Önerileri) ve öneri ölçütünü (Film Türüne Göre veya Film İsmine Göre) sunmak için kullanılmıştır.
- ListBox :** Kullanıcılar, belirlenen film türleri ve filmleri gösterirken kullanılmaktadır.

c) FİLM ÖNERİ SİSTEMİNİN TASARIMI

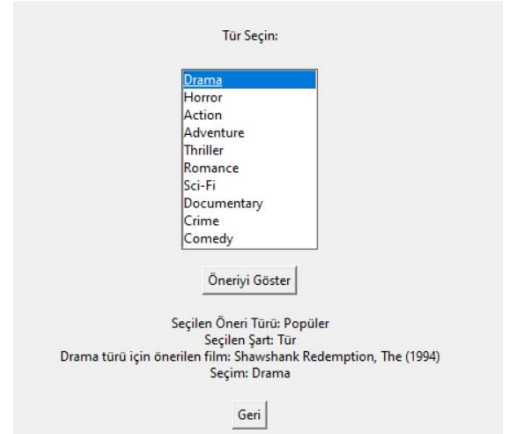
i. Popüler Filme Göre Öneriler

Kullanıcının, seçtiği filmle beraber, en çok izlenen filmi birliktelik kurallarına dayanarak alıyoruz. Bu sayede, popüler filmler arasından en çok izlenen filme göre öneri sunulmaktadır.



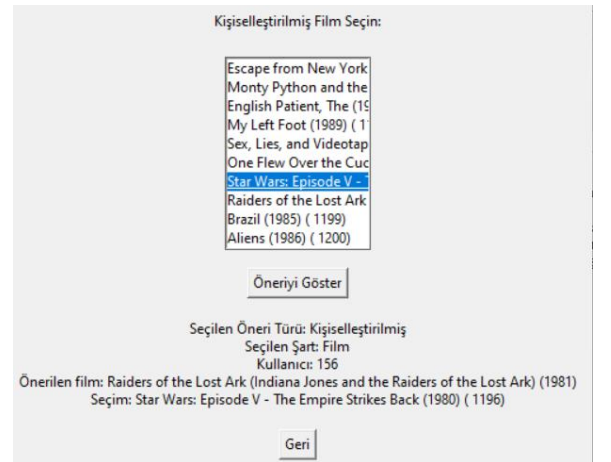
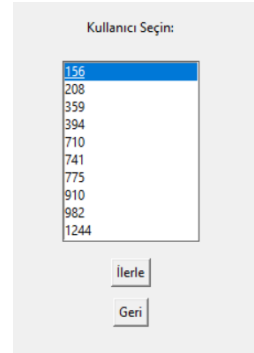
ii. Popüler Türe Göre Öneriler

Tüm filmler içinden kullanıcının seçtiği türdeki filmleri alıyoruz sonra bu filmlerden en çok izlenen filmi alıp kullanıcıya öneriyoruz.



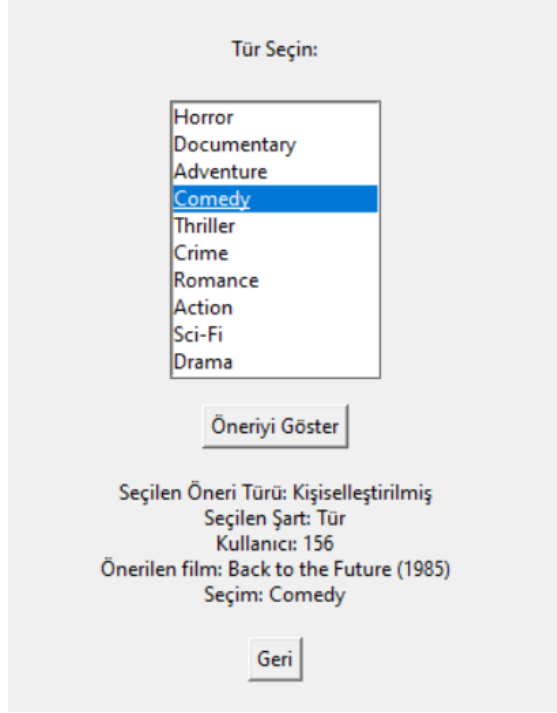
iii) Kişiselleştirilmiş Filme Göre Öneriler

Önce kullanıcı seçilir sonra, seçilen kullanıcının izlediği filmler içerisinde bir film seçilmesi istenmektedir. Birliktelik kuralları kullanılarak kullanıcı tarafından seçilen filmle beraber en çok izlenen film seçilir. Sonra bu filmi kullanıcının izleyip izlemediğine bakılır eğer izlememişse önerilir izlemişse 2. En çok izlenen film seçilir.



IV) Kişiselleştirilmiş Türe Göre Öneriler

Önce kullanıcı seçilir arından. Kullanıcının belirlenen film türleri arasından seçim yapması istenmektedir. Kullanıcının izlediği filmlerle beraber en çok izlenen filmler alınır. Bu filmlerin tür kontrolü yapılır ve seçtiği türdeki filmler alınır. Sonra bu filmlerden kullanıcının izlemediği o türdeki en çok izlenen film seçilerek kullanıcıya gösterilir.



III. Değerlendirme Ölçütleri

Bu projede film öneri sisteminin doğruluğunu, performansını ve kullanıcı deneyimini değerlendirmek için çeşitli ölçütler kullanılmıştır.

Doğruluk : Kullanıcının önceki tercihlerine ve seçtiği türe göre sunulan önerilerin kullanıcıya uygun olup olmadığı kontrol edilmiştir. Bu doğruluğu ölçmek için doğruluk oranı, güven ve lift gibi Apriori algoritması kullanılarak oluşturulan birliktelik kuralları incelenmiştir.

Performans : Apriori algoritması büyük veri setlerinde çalışırken algoritmanın hızlı ve verimli çalışması önemlidir. Veri kümesi büyüdükçe yanıt süresinde değişiklikler görülmektedir. Bu süreyi en aza indirmek için Rule Tree ağaç yapısı kullanılmıştır. Rule tree ile önerilerin kullanıcıya hızlı sunulması sağlanmıştır.

Kullanıcı Deneyimi : Kullanıcı arayüzü basit, anlaşılır ve kullanıcı dostu olacak şekilde tasarlanmıştır. Arayüz, kullanıcıların popüler ve kişisel film önerilerini kolayca almasını sağlamaktadır.

IV. Sonuçlar ve Tartışma

Bu çalışmada, kullanıcıların isteklerine göre popüler film önerileri ve kişiselleştirilmiş film önerileri sunan bir uygulama geliştirilmiştir. Apriori algoritması kullanılarak büyük bir veri setinden çıkarılan kurallar kullanıcıların daha doğru ve verimli bir öneri alabilmesi için kullanılmıştır. Geliştirilen kullanıcı arayüzü sayesinde, kullanıcılar kolayca film türüne veya film ismine göre öneriler alabilmektedir.

Gelecek çalışmalarda, kullanıcı geri bildirimine göre arayüz üzerinde geliştirilmeler yapılabilir, daha fazla veri kaynağı eklenerek sistemin farklı kullanıcı profillerine göre uygun öneriler sunma yeteneği geliştirilebilir.

V. Kaynakça

- <https://pandas.pydata.org/docs/index.html>
- <https://www.kaggle.com/datasets/grouplens/movielens-20m-dataset>
- <https://medium.com/@ekrem.hatipoglu/machine-learning-association-rule-mining-birliktelik-kural-%C3%A7%C4%B1kar%C4%B1m%C4%B1-apriori-algorithm-4326b8f224c3>
- <https://docs.python.org/3/library/tkinter.html>