



# Hybrid Recommender System

miuul

# İş Problemi



ID'si verilen kullanıcı için item-based ve user-based recommender yöntemlerini kullanarak 10 film önerisi yapınız.

# Veri Seti Hikayesi

Veri seti, bir film tavsiye hizmeti olan MovieLens tarafından sağlanmıştır. İçerisinde filmler ile birlikte bu filmlere yapılan derecelendirme puanlarını barındırmaktadır. 27.278 filmde 2.000.0263 derecelendirme içermektedir. Bu veri seti ise 17 Ekim 2016 tarihinde oluşturulmuştur. 138.493 kullanıcı ve 09 Ocak 1995 ile 31 Mart 2015 tarihleri arasında verileri içermektedir. Kullanıcılar rastgele seçilmiştir. Seçilen tüm kullanıcıların en az 20 filme oy verdiği bilgisi mevcuttur.

## movie.csv

3 Değişken

27278 Gözlem

1.5 MB

<b>movieid</b>	Eşsiz film numarası.
<b>title</b>	Film adı
<b>genres</b>	Tür

## rating.csv

4 Değişken

20000263 Gözlem

690.4 MB

<b>userid</b>	Eşsiz kullanıcı numarası. (UniqueID)
<b>movieid</b>	Eşsiz film numarası. (UniqueID)
<b>rating</b>	Kullanıcı tarafından filme verilen puan
<b>timestamp</b>	Değerlendirme tarihi



# Proje Görevleri

---

# User Based Recommendation

## Görev 1: Veri Hazırlama

**Adım 1:** **movie**, **rating** veri setlerini okutunuz.

**Adım 2:** **rating** veri setine Id'lere ait film isimlerini ve türünü **movie** veri setinden ekleyiniz.

**Adım 3:** Toplam oy kullanılma sayısı 1000'in altında olan filmlerin isimlerini listede tutunuz ve veri setinden çıkartınız.

**Adım 4:** index'te userID'lerin sutunlarda film isimlerinin ve değer olarak ratinglerin bulunduğu dataframe için pivot table oluşturunuz.

**Adım 5:** Yapılan tüm işlemleri fonksiyonlaştırınız.

## Görev 2: Öneri Yapılacak Kullanıcının İzlediği Filmlerin Belirlenmesi

**Adım 1:** Rastgele bir kullanıcı id'si seçiniz.

**Adım 2:** Seçilen kullanıcıya ait gözlem birimlerinden oluşan **random\_user\_df** adında yeni bir dataframe oluşturunuz.

**Adım 3:** Seçilen kullanıcıların oy kullandığı filmleri **movies\_watched** adında bir listeye atayınız

## Görev 3: Aynı Filmleri İzleyen Diğer Kullanıcıların Verisine ve Id'lerine Erişilmesi

**Adım 1:** Seçilen kullanıcının izlediği filmlere ait sutunları `user_movie_df`'ten seçiniz ve **`movies_watched_df`** adında yeni bir dataframe oluşturunuz.

**Adım 2:** Her bir kullancının seçili user'in izlediği filmlerin kaçını izlediğini bilgisini taşıyan **`user_movie_count`** adında yeni bir dataframe oluşturunuz.

**Adım 3:** Seçilen kullanıcının oy verdiği filmlerin yüzde 60 ve üstünü izleyenlerin kullanıcı id'lerinden **`users_same_movies`** adında bir liste oluşturunuz.

## Görev 4: Öneri Yapılacak Kullanıcı ile En Benzer Kullanıcıların Belirlenmesi

**Adım 1:** `user_same_movies` listesi içerisindeki seçili user ile benzerlik gösteren kullanıcıların id'lerinin bulunacağı şekilde `movies_watched_df` dataframe'ini filtreleyiniz.

**Adım 2:** Kullanıcıların birbirleri ile olan korelasyonlarının bulunacağı yeni bir `corr_df` dataframe'i oluşturunuz.

**Adım 3:** Seçili kullanıcı ile yüksek korelasyona sahip (0.65'in üzerinde olan) kullanıcıları filtreleyerek `top_users` adında yeni bir dataframe oluşturunuz.

**Adım 4:** `top_users` dataframe'ine `rating` veri seti ile merge ediniz.

## Görev 5: Weighted Average Recommendation Score'un Hesaplanması ve İlk 5 Filmin Tutulması

**Adım 1:** Her bir kullanıcının **corr** ve **rating** değerlerinin çarpımından oluşan **weighted\_rating** adında yeni bir değişken oluşturunuz.

**Adım 2:** Film id'si ve her bir filme ait tüm kullanıcıların weighted rating'lerinin ortalama değerini içeren **recommendation\_df** adında yeni bir dataframe oluşturunuz.

**Adım 3:** **recommendation\_df** içerisinde weighted rating'i 3.5'ten büyük olan filmleri seçiniz ve weighted rating'e göre sıralayınız.

**Adım 4:** movie veri setinden film isimlerini getiriniz ve tavsiye edilecek ilk 5 filmi seçiniz.

# Item Based Recommendation

**Görev 1:** Kullanıcının izlediği en son ve en yüksek puan verdiği filme göre item-based öneri yapınız.

**Adım 1:** **movie, rating** veri setlerini okutunuz.

**Adım 2:** Seçili kullanıcının 5 puan verdiği filmlerden puanı en güncel olan filmin id'sinin alınız.

**Adım 3:** User based recommendation bölümünde oluşturulan **user\_movie\_df** dataframe'ini seçilen film id'sine göre filtreleyiniz.

**Adım 4:** Filtrelenen dataframe'i kullanarak seçili filmle diğer filmlerin korelasyonunu bulunuz ve sıralayınız.

**Adım 5:** Seçili film'in kendisi haricinde ilk 5 film'ı öneri olarak veriniz.

**miuul**

[miuul.com](http://miuul.com)