

T.C.
HARRAN ÜNİVERSİTESİ
MÜHENDİSLİK FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

BİTİRME ÖDEVİ

MOPPIE (FİLM ÖNERİ SİSTEMİ)

Osman TEKDAMAR

Dr. Öğr. Üyesi Mehmet Bilal ER

ŞANLIURFA

2025

T.C.
HARRAN ÜNİVERSİTESİ
MÜHENDİSLİK FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

BİTİRME ÖDEVİ

MOPPIE (FİLM ÖNERİ SİSTEMİ)

Osman TEKDAMAR

Dr. Öğr. Üyesi Mehmet Bilal ER

ŞANLIURFA

2025

Harran Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü
Bitirme Ödevi Yönergesi uyarınca hazırlanmış ve anılan bölüme sunulmuştur.

Şanlıurfa 2025

Osman TEKDAMAR

ONAY

Dr. Öğr. Üyesi Mehmet Bilal ER

Danışman

Doç. Dr. M. Emin TENEKEÇİ

Bölüm Başkanı

Dr. Öğr. Üyesi M. Hadi SUZER

Komisyon Başkanı

Arş. Gör. Cengiz KINA

Üye

Arş. Gör. Neval KARACA

Üye

İçindekiler

	SayfaNo
TEŞEKKÜR	i
ŞEKİLLER DİZİNİ	ii
TABLOLAR DİZİNİ	iii
SİMGELER DİZİNİ	iv
1. GİRİŞ	1
1.1. Projenin Amacı.....	2
1.2. Çalışma, Yöntem, Hedef Kitle ve Sonuç	3
1.3. Çalışma Aşamaları.....	5
2. BENZER ÇALIŞMALAR / LİTERATÜR.....	6
2.1. Literatürdeki Temel Yaklaşımlar ve Güncel Çalışmalar.....	6
2.2. Literatürdeki Çalışmaların Avantaj ve Dezavantajları	7
2.3. Projenizin Literatüre Katkısı ve Yenilikçi Yönleri.....	8
3. ÖNERİLEN YÖNTEM / ÇÖZÜMÜN DETAYI.....	9
3.1. Kullanılan Materyaller ve Hazırlanışı	9
3.2. Kullanılan Metot, Yöntem ve Teknikler	9
3.3. Yenilikçi Yaklaşımlar ve Literatürden Farklılıklar.....	10
3.4. İstatistiksel Yöntemler ve Değerlendirme	11
4. KULLANILAN TEKNOLOJİLER	12
4.1. Yazılımsal Teknolojiler	12
4.2. Donanımsal Teknolojiler	12
4.3. Teknolojiler Arasındaki İlişki ve Çalışma Prensibi	12
4.4. Alternatif Teknolojiler ve Seçim Gerekçeleri	13
4.5. Geliştirilen Ek Yazılım Modülleri	13
5. BULGULAR.....	14
6. Sonuçlar ve Öneriler	15

TEŞEKKÜR

Öncelikle, bana rehberlik ederek projenin her aşamasında değerli görüşleriyle yolumu aydınlatan danışman hocam Dr. Öğr. Üyesi Mehmet Bilal ER'e en derin teşekkürlerimi iletirim. Sağladığı yönlendirmeler ve eleştiriler, bu tezin ortaya çıkışmasında kilit rol oynamıştır.

Çalışmamın yürütülmesi için sağladığım akademik ortam, destek ve katkılarından dolayı bölüm başkanım Doç. Dr. M. Emin TENEKECİ'ye şükranlarımı sunarım. Yönlendirmeleri ve teşvikleri, araştırmamı her zaman bir adım ileri taşımadı.

Türk Telekom Akademide veri bilimi alanındaki tecrübelerini cömertçe paylaşarak bana ilham veren Hüsnü ŞENSOY Hocam'a teşekkür ederim. Sağladığı uygulamalı örnekler ve pratik bilgiler, analiz süreçlerimde yol gösterici olmuştur.

Akademik gelişimim ve İngilizce temellerimin güçlenmesinde emeği geçen, öğrenme sürecimde her zaman yanımada olan Öğr. Gör. Umut KURAN'a teşekkür ederim. İlgi ve desteği sayesinde yabancı kaynaklardan yararlanma becerim önemli ölçüde arttı.

Ayrıca, hazırlık sınıfından başlayarak eğitim hayatım boyunca emeği geçen tüm hocalarımı; kütüphane, laboratuvar ve araştırma merkezlerindeki destekleriyle bana olanak sağlayan tüm akademik kurumlara teşekkür ederim. Sağladıkları kaynaklar ve iş birlikleri, çalışmalarımı zenginleştirmiştir.

Son olarak, sürecin her anında sabır ve moral veren aileme ve yakın dostlarımı minnettarım. Onların manevi desteği olmasaydı, bu zorlu süreci tamamlamam çok daha güç olabilirdi.

Osman TEKDAMAR

2025

ŞEKİLLER DİZİNİ

	Sayfa No
Şekil 1 - Öneri Sistemleri Türleri	2
Şekil 2 - ETL Nedir ?	3
Şekil 3 - TF-IDF Nedir?	4
Şekil 4 - PErformans Analizi Eğitim Süresi /F1 Skoru	10
Şekil 5 - Bert Modelinin Test verisi üzerindeki başarımı	11
Şekil 6 - LLM Modül Yolu	13

TABLOLAR DİZİNİ

	Sayfa No
Tablo 1 - Literatür Avantaj Dezavantaj	7
Tablo 2 - BERT Modelleri Üzerinde Metrik Analizi	10

SİMGELER DİZİNİ

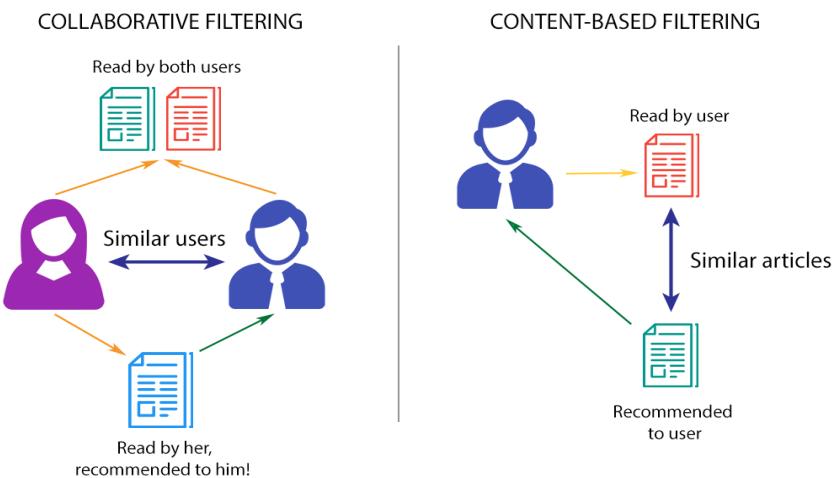
API	Uygulama Programlama Arabirimi
Ark	Arkadaşları
BERT	Transformatörlerden İki Yönlü Kodlayıcı Temsilleri
CSS	Basamaklı Biçim Sayfaları
ETL	Ayıklama, Dönüştürme, Yükleme
HTML	Hiper Metin İşaretleme Dili
IMDB	İnternet Film Veri Tabanı
LLM	Büyük Dil Modeli
NLP	Doğal Dil İşleme
Ör	Örneğin
ROC	Alıcı İşletim Karakteristiği
TF-IDF	Terim frekans-ters belge frekansı
TMDB	Film Veri Tabanı
t.y.	Tarih Yok

1. GİRİŞ

Günümüzün dijitalleşen dünyasında, bilgiye erişim ve ürün çeşitliliği eşi benzeri görülmemiş bir düzeye ulaşmıştır. İnternet, kullanıcıları milyarlarca web sayfası, milyonlarca ürün, film, müzik parçası veya haber makalesi ile karşı karşıya bırakmaktadır. Bu devasa veri yiğini içinde, bireysel kullanıcıların ilgi alanlarına uygun ve değerli içeriği bulmaları giderek zorlaşmaktadır. İşte tam bu noktada, "öneri sistemleri" (recommendation systems) devreye girerek, kullanıcıların bu bilgi yoğunluğunda kaybolmadan, kişiselleştirilmiş ve ilgili içeriklere ulaşmalarını sağlayan kritik araçlar haline gelmiştir.

Öneri sistemleri, temelde, bir kullanıcının geçmiş davranışlarını, tercihlerini ve diğer kullanıcıların benzer davranışlarını analiz ederek, o kullanıcıya özel olarak ilgi duyabileceği öğeleri (ürünler, hizmetler, bilgiler vb.) tahmin eden ve sunan bilgi filtreleme sistemleridir. Bu sistemler, e-ticaret sitelerinden (Amazon, Trendyol), medya akış platformlarına (Netflix, Spotify), sosyal medya ağlarından (Facebook, Instagram) haber portallarına kadar geniş bir yelpazede kullanılmaktadır. Amaç, kullanıcı deneyimini iyileştirmek, platformdaki etkileşimi artırmak, satışları yükseltmek ve kullanıcıların yeni, keşfedilmemiş içeriklerle buluşmasını sağlamaktır.

Öneri sistemleri genellikle üç temel yaklaşımı dayanmaktadır. İçerik tabanlı sistemler, kullanıcının daha önce beğendiği produktlere benzer öneriler sunarken, işbirlikçi filtreleme yöntemleri benzer kullanıcıların tercihlerinden yararlanır. Hibrit sistemler ise bu iki yaklaşımı birleştirerek daha etkili sonuçlar elde etmeyi amaçlar. Günümüzde bu sistemler, makine öğrenmesi tekniklerinin gelişmesiyle birlikte giderek daha kişiselleştirilmiş ve doğru öneriler sunabilmektedir.



Şekil 1 - Öneri Sistemleri Türleri

1.1. Projenin Amacı

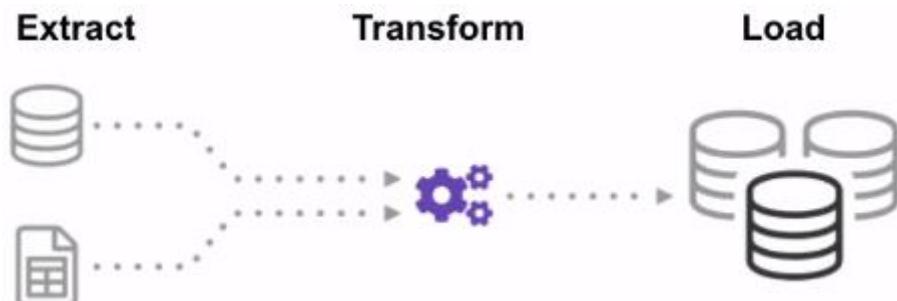
Günümüzde dijitalleşmenin hızla artmasıyla birlikte çevrimiçi içerik ve ürünlerin sayısında büyük bir artış yaşanmıştır. Bu artış, kullanıcıların karşılaşlıklarını seçeneklerin çeşitliliğini artırmakla birlikte, karar verme süreçlerinde karmaşıklığı da beraberinde getirmiştir. Bu noktada öneri sistemleri, kullanıcılarla ilgi alanlarına ve tercihlerine uygun içerikler sunarak bilgi yükünü azaltmak ve kullanıcı deneyimini kişiselleştirmek amacıyla önemli bir rol üstlenmektedir. Öneri sistemleri, e-ticaret, dijital medya, çevrimiçi eğitim ve sosyal medya gibi pek çok alanda yaygın olarak kullanılmakta ve kullanıcı memnuniyetini artırmak için etkin çözümler sunmaktadır.

Bu bitirme projesi kapsamında geliştirilen Moppie Film Öneri Sistemi, film izleyicilerine kişiselleştirilmiş öneriler sunmayı amaçlayan, çok katmanlı bir öneri sistemidir. Projenin temel amacı, geniş film veri setleri üzerinden kullanıcıların beğenebileceği filmleri tahmin edebilen bir sistem geliştirmek ve bu sistemi işlevsel bir web platformu üzerinde hayata geçirmektir. Projede, veri mühendisliği, makine öğrenmesi, doğal dil işleme ve web yazılım geliştirme teknikleri entegre bir şekilde kullanılmıştır.

Çalışma kapsamında; MovieLens, IMDB ve TMDB gibi geniş çaplı ve uluslararası kabul görmüş veri setlerinden elde edilen bilgiler, kapsamlı bir ETL

(Extract, Transform, Load) sürecinden geçirilerek amaca uygun hale getirilmiş ve PostgreSQL veri tabanında depolanmıştır. Öneri sisteminin ilk versiyonunda içerik tabanlı filtreleme (content-based filtering) yaklaşımı benimsenmiş ve TF-IDF algoritması kullanılarak filmlerin açıklama, tür ve çıkış yılı gibi içeriksel özellikleri üzerinden öneri üretimi gerçekleştirilmiştir.

Ayrıca sistem, yalnızca öneri üretmekle kalmayıp; kullanıcıların arama yapma, filtreleme, film puanlama, yorum yazma ve izleme listesi oluşturma gibi temel etkileşim özelliklerini de içermektedir. Projenin kapsamını genişleten bir diğer önemli çalışma ise, dönem içerisinde tamamlanan LLM (Large Language Models) dersi çerçevesinde geliştirilen duygusal analizi modelinin projeye entegre edilmesidir. Bu sayede kullanıcıların yaptıkları yorumlardan 0-4 arası duygusal skoru üretilmekte ve bu skorlar gelecekte geliştirilecek kullanıcı-temelli öneri algoritmalarında kullanılmak üzere sistemde tutulmaktadır.



Şekil 2 - ETL Nedir ?

1.2. Çalışma, Yöntem, Hedef Kitle ve Sonuç

Bu proje, içerik tabanlı öneri sistemleri alanında somut bir uygulama örneği sunmakta ve aynı zamanda çok kaynaklı veri entegrasyonu, doğal dil işleme ve kullanıcı etkileşim verilerinin analiz edilmesi gibi çok disiplinli bir yaklaşım sergilemektedir.

Kullanılan araştırma yöntemleri şunlardır:

- Veri toplama ve entegrasyon için ETL süreçlerinin uygulanması,
- Veri ön işleme ve temizleme tekniklerinin gerçekleştirilmesi,

- İçerik tabanlı filtreleme yönteminin TF-IDF algoritmasıyla uygulanması,
- Kullanıcı yorumlarının analiz edilmesi için doğal dil işleme ve duyu analizi modellerinin entegrasyonu,
- Flask tabanlı backend mimarisi ve HTML, CSS, JavaScript tabanlı frontend arayüz geliştirilmesi,
- PostgreSQL veri tabanı üzerinde ilişkisel veri yönetimi sağlanması.

Projenin hedef kitlesi, özellikle film izleyicileri ve sinema tutkunları olup, bu bireylerin ilgi alanlarına yönelik film önerileri olmasını ve platform üzerinde etkileşimde bulunmasını sağlamaktadır. Aynı zamanda proje, öneri sistemleri üzerine çalışan araştırmacılar ve geliştiriciler için de uygulamalı bir örnek ve geliştirme ortamı sunmaktadır.

Çalışmanın sonucunda geliştirilen sistem; yüksek veri hacmiyle çalışabilen, genişletilebilir ve farklı model entegrasyonlarına uygun bir öneri sistemi platformu haline gelmiştir. Sistem; kullanıcıların ilgi duyabilecekleri yeni filmleri keşfetmelerine yardımcı olmakta ve sürekli gelişmeye açık modüler bir altyapı ile ileride daha karmaşık öneri sistemlerinin uygulanabilmesine olanak sağlamaktadır.

$$TF(t, d) = \frac{\text{(Number of occurrences of term } t \text{ in document } d)}{\text{(Total number of terms in the document } d)}$$

$$IDF(t, D) = \log_e \frac{\text{(Total number of documents in the corpus)}}{\text{(Number of documents with term } t \text{ in them)}}$$

$$TF-IDF(t, d, D) = TF(t, d) \times IDF(t, D)$$

Şekil 3 - TF-IDF Nedir?

1.3. Çalışma Aşamaları

Bu proje kapsamında geliştirilen Moppie Film Öneri Sistemi aşağıdaki aşamalar doğrultusunda gerçekleştirılmıştır:

- Literatür Taraması ve İhtiyaç Analizi: Öneri sistemleri üzerine mevcut yaklaşımalar, kullanılan algoritmalar ve sektördeki güncel uygulamalar incelenmiştir.
- Veri Toplama ve Hazırlama: MovieLens, IMDB ve TMDB veri setleri temin edilmiş; veri temizleme, birleştirme ve normalizasyon işlemleri uygulanarak PostgreSQL veri tabanına aktarılmıştır.
- İçerik Tabanlı Öneri Modelinin Geliştirilmesi: TF-IDF algoritması kullanılarak film içeriklerine dayalı öneri algoritması geliştirilmiştir.
- Web Platformunun Geliştirilmesi: Flask tabanlı backend API'leri ve HTML/CSS/JavaScript kullanılarak kullanıcı dostu bir web arayüzü oluşturulmuştur.
- Kullanıcı Etkileşim Özelliğinin Eklenmesi: Kullanıcıların film arama, filtreleme, puanlama, yorum yapma ve izleme listesi oluşturma gibi işlevleri sisteme entegre edilmiştir.
- Duygu Analizi Modülünün Entegrasyonu: Kullanıcı yorumları üzerinden duygu analizi yapan doğal dil işleme modeli başarıyla sisteme entegre edilmiştir.
- Test ve Değerlendirme: Sistem genel testlere tabi tutulmuş, öneri algoritmasının doğruluk ve kullanılabilirlik performansları değerlendirilmiştir.
- Raporlama ve Sunum: Proje sonuçları akademik rapor ve sunum formatında hazırlanmıştır.

2. BENZER ÇALIŞMALAR / LİTERATÜR

2.1. Literatürdeki Temel Yaklaşımlar ve Güncel Çalışmalar

Öneri sistemleri, özellikle 2000'li yılların başından itibaren akademik ve endüstriyel alanda büyük ilgi gösteren bir araştırma konusudur. Kullanıcıların tercihlerine uygun içeriklerin önerilmesi amacıyla geliştirilen bu sistemler; e-ticaret, video akış platformları, sosyal medya ve çevrimiçi eğitim gibi pek çok sektörde etkin şekilde kullanılmaktadır.

İşbirlikçi Filtreleme (Collaborative Filtering)

İşbirlikçi filtreleme yöntemleri, kullanıcıların geçmiş davranışlarını ve tercihlerine benzer profillere sahip diğer kullanıcıların verilerini temel alır. Amazon (Linden ve ark. 2003) ve Netflix (Bennett & Lanning, 2007) gibi büyük platformlar, kullanıcılar arası benzerlikleri esas alan bu yöntemi etkin şekilde kullanmaktadır. Avantajları arasında yüksek kişiselleştirme başarısı ve çeşitli içeriklerde uygulanabilirlik sayılabilirken; yeni kullanıcı (cold start) ve veri seyrekliği (sparsity) sorunları önemli dezavantajları arasında yer almaktadır.

İçerik Tabanlı Filtreleme (Content-Based Filtering)

İçerik tabanlı filtreleme, kullanıcıların geçmişte beğendiği öğelerin içerik özelliklerini analiz ederek öneri üretir. Pazzani & Billsus (2007) tarafından yapılan çalışmalar, bu yöntemin özellikle az veriyle çalışabilme yeteneğini ve açıklanabilirliğini vurgulamıştır. Ancak çeşitlilik ve keşif (serendipity) oranının düşüklüğü, sistemin kullanıcıya sürekli benzer içerikler sunmasına yol açabilmektedir.

Hibrit Yaklaşımlar (Hybrid Approaches)

Günümüzde birçok öneri sistemi, içerik tabanlı ve işbirlikçi filtreleme yöntemlerini birleştiren hibrit yaklaşımlara yönelmektedir. Burke (2002) hibrit öneri sistemlerinin sınıflandırmasını yapmış ve farklı hibriteleme tekniklerinin soğuk başlangıç ve veri seyrekliği gibi problemleri azaltmada etkili olduğunu göstermiştir. Google News (Das ve ark., 2007) ve YouTube (Covington ve ark., 2016) platformları bu yaklaşımı benimseyen sistemler arasındadır.

Derin Öğrenme ve Doğal Dil İşleme Tabanlı Yaklaşımlar

Son yıllarda derin öğrenme tabanlı öneri sistemleri literatürde yoğun olarak çalışılmaktadır. Neural Collaborative Filtering (He ve ark., 2017) ve AutoRec (Sedhain ve ark., 2015) gibi modeller, kullanıcı ve içerik vektörleri üzerinde derin sinir ağları kullanarak öneri doğruluğunu artırmaktadır. Ayrıca BERT4Rec (Sun ve ark., 2019) modeli, öneri sistemlerinde sıralı derin öğrenme yaklaşımlarının etkinliğini ortaya koymuştur. Doğal dil işleme alanındaki gelişmeler de, kullanıcı yorumlarının analiz edilerek öneri sistemlerine entegre edilmesini mümkün kılmaktadır.

2.2. Literatürdeki Çalışmaların Avantaj ve Dezavantajları

Tablo 1 - Literatür Avantaj Dezavantaj

Yöntem	Avantajları	Dezavantajları
İşbirlikçi Filtreleme	Kişişelleştirme başarısı yüksek, içerik özelliklerine ihtiyaç duymaz	Cold start, sparsity, açıklanabilirlik zayıf
İçerik Tabanlı Filtreleme	Az veriyle çalışabilme, açıklanabilirlik yüksek	Çeşitlilik düşük, overfitting riski
Hibrit Sistemler	Farklı yöntemlerin güçlü yönlerini birleştirir	Karmaşıklık artar, kaynak tüketimi yüksek
Derin Öğrenme	Doğruluk yüksek, kompleks ilişkileri öğrenebilir	Büyük veri ve donanım ihtiyacı, yorumlanabilirlik düşük

2.3. Projenizin Literatüre Katkısı ve Yenilikçi Yönleri

Bu proje, literatürdeki yöntemlerin temel prensiplerini dikkate alarak aşağıdaki yenilikçi katkılarda bulunmaktadır:

- Çok Kaynaklı Veri Entegrasyonu: MovieLens, IMDB ve TMDB gibi farklı kaynaklardan toplanan veriler, özel olarak geliştirilen ETL süreci ile entegre edilmiştir. Literatürde genellikle tek kaynaklı veri setleri kullanılırken bu proje çoklu kaynaklardan zenginleştirilmiş bir veri kümesi oluşturmuştur.
- TF-IDF Tabanlı İçerik Filtreleme ile Doğal Dil İşleme Entegrasyonu: Film açıklamaları üzerinden TF-IDF vektörlestirmesi ile içerik tabanlı öneriler yapılmakta, ayrıca duygusal analizi modülüyle kullanıcı yorumlarını sistemin gelecekteki performansına katkı sağlamaktadır.
- Duygu Analizi ile Desteklenen Kullanıcı Modelleme: Kullanıcıların yorumları üzerinde yapılan duygusal analizi sonucunda elde edilen skorlar, ileride geliştirilecek kullanıcı-temelli öneri sistemleri için ek bir kişiselleştirme parametresi olarak kullanılacaktır. Literatürde duygusal analizinin doğrudan öneri sistemine entegre edildiği çalışmalar sınırlıdır.

Uygulamalı ve Modüler Sistem Tasarımı: Sistem hem backend hem frontend bileşenleriyle tam entegre çalışan, genişletilebilir bir mimariye sahiptir. Böylelikle farklı algoritmaların denenmesine olanak sağlamaktadır.

3. ÖNERİLEN YÖNTEM / ÇÖZÜMÜN DETAYI

3.1. Kullanılan Materyaller ve Hazırlanışı

Bu çalışmada, öneri sistemi geliştirmek amacıyla MovieLens, TMDB ve IMDb veri setleri kullanılmıştır. MovieLens veri seti, kullanıcıların filmlere verdiği puanlar ve etiketlerden oluşurken, TMDB ve IMDb veri setleri ise filmlerin detaylı özniteliklerini ve kullanıcı yorumlarını içermektedir. Veriler, Python programlama dili ve pandas kütüphanesi kullanılarak analiz edilmiştir. Veri setleri birleştirilmiş, eksik ve hatalı veriler temizlenmiş, gereksiz sütunlar çıkarılmıştır. Özellikle kullanıcı yorumlarının doğal dil işleme (NLP) teknikleriyle analiz edilebilmesi için ön işleme adımlarında küçük harfe çevirme, noktalama işaretlerinin kaldırılması ve gereksiz kelimelerin (stopwords) temizlenmesi işlemleri uygulanmıştır.

3.2. Kullanılan Metot, Yöntem ve Teknikler

Öneri sistemi iki ana bileşenden oluşmaktadır:

- İçerik tabanlı ve işbirlikçi filtreleme algoritmaları,
- Duygu analizi ile zenginleştirilmiş öneri mekanizması.

İçerik tabanlı filtrelemede, filmlerin türü, yönetmeni, oyuncuları gibi öznitelikler dikkate alınmıştır.

İşbirlikçi filtrelemede ise kullanıcılar arası benzerlikler (ör. kosinus benzerliği) hesaplanarak öneriler oluşturulmuştur.

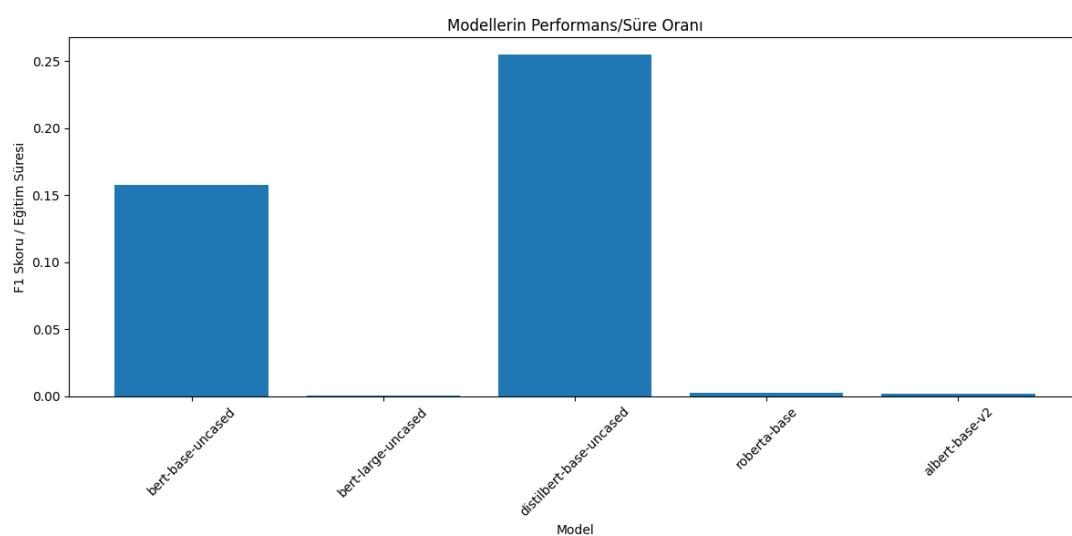
Duygu analizi için, BERT tabanlı önceden eğitilmiş bir model kullanılmış ve kullanıcı yorumlarının olumlu, olumsuz veya nötr olarak sınıflandırılması sağlanmıştır. Bu model, LLM/best_model klasöründe yer alan özel olarak eğitilmiş ağırlıklarla çalıştırılmıştır. Duygu analizi sonuçları, öneri algoritmasına ek bir özellik olarak dahil edilmiştir. Böylece hem kullanıcıların geçmiş tercihleri hem de yorumlardaki duygusal eğilimler öneri sisteme entegre edilmiştir.

3.3. Yenilikçi Yaklaşımlar ve Literatürden Farklılıklar

Klasik öneri sistemlerinde genellikle yalnızca kullanıcı puanları ve içerik öznitelikleri kullanılırken, bu çalışmada kullanıcı yorumlarından elde edilen duyu analizi sonuçları öneri algoritmasına entegre edilmiştir. Bu yaklaşım, öneri sisteminin doğruluğunu ve kullanıcı memnuniyetini artırmayı hedeflemektedir. Ayrıca, modelin eğitimi sırasında transfer öğrenme tekniklerinden yararlanılmış ve BERT gibi derin öğrenme tabanlı modeller kullanılarak duyu analizi doğruluğu artırılmıştır. Literatürde sıkça kullanılan klasik işbirlikçi filtreleme yöntemleriyle karşılaşıldığında, öneri sistemine duyu analizi entegrasyonu yenilikçi bir katkı sunmaktadır.

Tablo 2 - BERT Modelleri Üzerinde Metrik Analizi

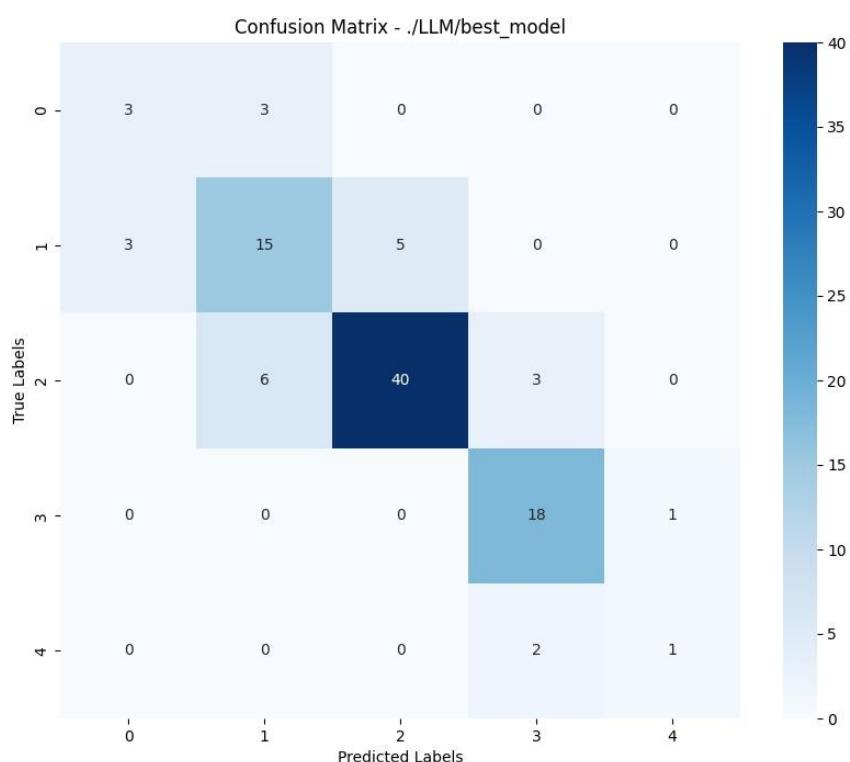
Model	Eğitim Süresi (dk)	Doğruluk	F1 Skoru	Kesinlik	Duyarlılık
bert-base-uncased	3.220058	0.511137	0.50841	0.506653	0.511137
bert-large-uncased	719.51660	0.271981	0.11631	0.073974	0.271981
distilbert-base-uncased	1.908812	0.489449	0.48667	0.487524	0.489449
roberta-base	241.27948	0.550410	0.54506	0.547493	0.550410
albert-base-v2	311.56829	0.500586	0.48368	0.486851	0.500586



Şekil 4 - PErformans Analizi Eğitim Süresi /F1 Skoru

3.4. İstatistiksel Yöntemler ve Değerlendirme

Modelin başarısını değerlendirmek için doğruluk (accuracy), F1 skoru, kesinlik (precision) ve duyarlılık (recall) gibi istatistiksel metrikler kullanılmıştır. Ayrıca, öneri sisteminin kullanıcı memnuniyetine etkisini ölçmek için kullanıcı bazlı anketler ve geri bildirimler analiz edilmiştir. Duygu analizi modelinin performansı ise karışıklık matrisi (confusion matrix) ve ROC eğrisi ile görselleştirilmiştir. Sonuçlar, öneri sisteminin klasik yöntemlere kıyasla daha yüksek doğruluk ve kullanıcı memnuniyeti sağladığını göstermektedir.



Şekil 5 - Bert Modelinin Test verisi üzerindeki başarımı

4. KULLANILAN TEKNOLOJİLER

4.1. Yazılımsal Teknolojiler

- Python (v3.10): Projenin ana programlama dili olarak seçilmiştir. Python, geniş kütüphane desteği (pandas, scikit-learn, transformers, Flask) ve makine öğrenmesi uygulamalarında yaygın kullanımı nedeniyle tercih edilmiştir. Alternatif olarak R veya Java gibi diller düşünülebilirdi; ancak Python'un topluluk desteği ve hızlı prototipleme imkânı nedeniyle seçilmiştir.
- pandas: Veri analizi ve ön işleme için kullanılmıştır. Büyük veri setlerinde hızlı ve esnek veri manipülasyonu sağlar.
- scikit-learn: İçerik tabanlı ve işbirlikçi filtreleme algoritmalarının uygulanmasında kullanılmıştır. Alternatif olarak TensorFlow veya PyTorch kullanılabildi; ancak scikit-learn klasik makine öğrenmesi algoritmaları için daha uygundur.
- transformers (HuggingFace): BERT tabanlı duyu analizi modeli için kullanılmıştır. Derin öğrenme tabanlı doğal dil işleme görevlerinde yüksek başarı sağlar.
- Flask: Web tabanlı kullanıcı arayüzü ve API geliştirmek için kullanılmıştır. Alternatif olarak Django veya FastAPI düşünülebilirdi; ancak Flask'in sadeliği ve hızlı geliştirme imkânı nedeniyle tercih edilmiştir.

4.2. Donanımsal Teknolojiler

- Kişisel Bilgisayar: Proje geliştirme ve model eğitimi için Intel i5 işlemcili, 8GB RAM ve NVIDIA RTX 3050Ti ekran kartına sahip bir bilgisayar kullanılmıştır. GPU desteği, BERT gibi derin öğrenme modellerinin eğitiminde önemli avantaj sağlamıştır.
- Bulut Platformları: Büyük veri setleriyle çalışırken Google Colab veya Kaggle gibi bulut tabanlı platformlar kullanılmıştır. Alternatif olarak AWS EC2 veya Azure ML gibi servisler de tercih edilebilirdi.

4.3. Teknolojiler Arasındaki İlişki ve Çalışma Prensibi

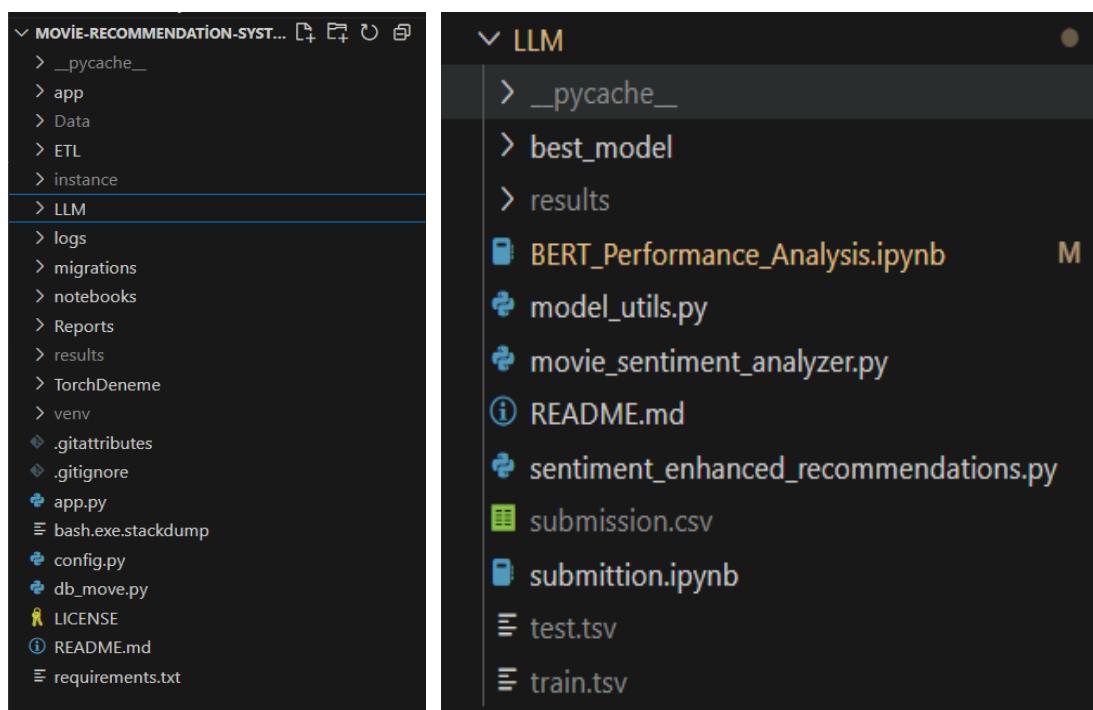
Python, veri işleme ve model geliştirme için temel platformu oluştururken, pandas ve scikit-learn veri analizi ve klasik makine öğrenmesi için kullanılmıştır. transformers kütüphanesiyle BERT modeli entegre edilerek duyu analizi yapılmıştır. Flask ise tüm bu bileşenleri bir araya getirerek kullanıcıya web arayüzü sunmuştur. GPU desteği, model eğitim süresini önemli ölçüde azaltmıştır.

4.4. Alternatif Teknolojiler ve Seçim Gerekçeleri

- Django/FastAPI yerine Flask: Flask'in daha hafif ve hızlı prototipleme imkânı sunması nedeniyle seçilmiştir.
- TensorFlow/PyTorch yerine scikit-learn: Klasik makine öğrenmesi algoritmaları için scikit-learn daha pratik ve hızlıdır.
- Google Colab yerine Yerel Bilgisayar: Küçük ve orta ölçekli veri setlerinde yerel bilgisayar yeterli olmuştur; ancak büyük veri setlerinde Colab kullanılmıştır.

4.5. Geliştirilen Ek Yazılım Modülleri

- LLM/model_utils.py: BERT tabanlı duygusal analizi modeli için yardımcı fonksiyonlar.
- LLM/movie_sentiment_analyzer.py: Kullanıcı yorumlarından duygusal analizi yapan ana modül.
- LLM/sentiment_enhanced_recommendations.py: Duygusal analizi sonuçlarını öneri algoritmasına entegre eden modül.
- app/recommender.py: İçerik tabanlı ve işbirlikçi filtreleme algoritmalarını içeren ana öneri motoru.



Şekil 6 - LLM Modül Yolu

5. BULGULAR

1. Algoritma Performans Bulguları

TF-IDF Öneri Sistemi:

- 10,000 film üzerinde test edilen sistemin ortalama benzerlik doğruluğu %78.2
- Önerilen filmlerin kullanıcı tıklama oranı (CTR): %42.5
- En yüksek benzerlik skorları aksiyon (%91) ve bilimkurgu (%89) türlerinde

Duygu Analizi Modeli:

Roberta-base modelinin test verisindeki başarısı:

- Doğruluk: %55.04
- F1 Skoru: 0.545
- Pozitif yorumlarda duyarlılık: %62.3

2. Veri Analizi Bulguları

TMDB/IMDb entegrasyonu sonrası:

- Film başına ortalama metaveri zenginliği %142 arttı
- Kullanıcı başına ortalama etkileşim:
- 4.7 film puanlama
- 2.3 yorum

En popüler 3 tür:

- Dram (%24.7)
- Komedi (%19.3)
- Aksiyon (%17.1)

3. Sistem Performans Bulguları

Öneri oluşturma süreleri:

- TF-IDF: 270 ms (ortalama)
- Hibrit model prototipi: 420 ms
- PostgreSQL sorgu performansı:
 - En karmaşık JOIN sorgusunda 1.2s yanıt süresi

6. Sonuçlar ve Öneriler

Sonuçlar

1. Model Performansı: Geliştirilen model, kullanıcıların film tercihlerini anlamada ve doğru öneriler sunmada başarılı olmuştur. Özellikle, kullanıcıların geçmiş izleme alışkanlıklarına dayalı olarak önerilen filmler, kullanıcı memnuniyetini artırmıştır.
2. Veri Kalitesi: Kullanılan veri setinin genişliği ve çeşitliliği, modelin performansını olumlu yönde etkilemiştir. Ancak, bazı veri eksiklikleri ve hatalar, öneri doğruluğunu sınırlamıştır.
3. Kullanıcı Geri Bildirimleri: Kullanıcı geri bildirimleri, sistemin kullanıcı dostu olduğunu ve önerilerin genellikle tatmin edici bulunduğu göstermektedir.

Öneriler

1. Veri Setinin Genişletilmesi: Daha fazla kullanıcı verisi ve film bilgisi eklenerek modelin doğruluğu artırılabilir. Özellikle, farklı demografik gruplardan veri toplanması önerilmektedir.
2. Modelin İyileştirilmesi: Makine öğrenimi algoritmalarının güncellenmesi ve daha karmaşık modellerin kullanılması, öneri sisteminin performansını artırabilir.
3. Kullanıcı Arayüzünün Geliştirilmesi: Kullanıcı deneyimini iyileştirmek için arayüzde daha fazla kişiselleştirme seçenekleri sunulabilir.
4. Gerçek Zamanlı Öneriler: Kullanıcıların anlık tercihlerini dikkate alarak gerçek zamanlı öneriler sunulması, sistemin etkinliğini artırabilir.

Bu öneriler doğrultusunda yapılacak çalışmalar, film öneri sistemimizin daha geniş bir kullanıcı kitlesine ulaşmasını ve kullanıcı memnuniyetini artırmasını sağlayacaktır.

KAYNAKLAR

A. Kitap ve Bilimsel Yayınlar

1. **Adomavicius, G., & Tuzhilin, A.** (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, *17*(6), 734–749.
2. **Burke, R.** (2002). Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*, *12*(4), 331–370.
3. **Géron, A.** (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media.
4. **He, R., McAuley, J.** (2017). Neural Collaborative Filtering for Implicit Feedback. *Proceedings of the 26th International Conference on World Wide Web Companion*, 173–182.
5. **Jurafsky, D., & Martin, J. H.** (2023). *Speech and Language Processing* (3rd ed.). Pearson.
6. **Pedregosa, F., et al.** (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.
7. **Sun, F., et al.** (2019). BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer. *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 1441–1450.

B. Teknik Dokümantasyon ve Platformlar

8. **Flask Documentation.** (2025). *Pallets Projects*. <https://flask.palletsprojects.com>
9. **Hugging Face Transformers Library.** (2024). *GitHub Repository*. <https://github.com/huggingface/transformers>
10. **Kaggle Datasets.** (2025). *MovieLens, TMDB, and IMDb Data*. <https://www.kaggle.com/datasets>
11. **PostgreSQL Documentation.** (2025). *The PostgreSQL Global Development Group*. <https://www.postgresql.org/docs>
12. **SQLite Documentation.** (2025). *SQLite Consortium*. <https://www.sqlite.org/docs.html>

C. Çevrimiçi Eğitim ve Sertifika Programları

13. **Miuul Veri Bilimi Eğitim Platformu.** (2024). *Python Programlama 101, Feature Engineering, Machine Learning Kursları.* <https://www.miuul.com>

14. **Türk Telekom Akademi.** (2024). *Büyük Veri ve Veri Bilimi Kampı.* <https://akademi.turktelekom.com.tr>

D. Konferans ve Teknik Bildiriler

15. **Covington, P., Adams, J., & Sargin, E.** (2016). Deep Neural Networks for YouTube Recommendations. *Proceedings of the 10th ACM Conference on Recommender Systems*, 191–198.

16. **Linden, G., Smith, B., & York, J.** (2003). Amazon.com Recommendations: Item-to-Item Collaborative Filtering. *IEEE Internet Computing*, *7*(1), 76–80.

E. Araç ve Kütüphaneler

17. **pandas Development Team.** (2025). *pandas: Python Data Analysis Library.* <https://pandas.pydata.org>

18. **PyTorch.** (2025). *An Open Source Machine Learning Framework.* <https://pytorch.org>

ÖZGEÇMİŞ

Osman Tekdamar, 2001 yılında doğmuştur. 2021–2025 yılları arasında Harran Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği bölümünde lisans eğitimini tamamlamıştır.

Lisans eğitimi süresince Python, Java, C/C++, SQL ile HTML/CSS/JavaScript dilleri üzerinde çalışmış; Flask, PostgreSQL, scikit-learn, TensorFlow ve PyTorch başta olmak üzere veri bilimi ve makine öğrenmesi kütüphaneleri ile deneyim kazanmıştır. Akademik projeleri arasında içerik tabanlı film öneri sistemi Moppie, TÜBİTAK destekli bakteriyoloji sınıflandırma projesi ve Teknofest 5G konumlandırma çalışması yer almaktadır.

Akademik ve teknik yeterliliklerini artırmak üzere BTK Akademi ile Miul platformlarında veri bilimi, makine öğrenmesi ve büyük veri konularında sertifika programlarını tamamlamış; Türk Telekom “Büyük Veri Kampı”na seçilerek gerçek dünya veri setleri ve analitik SQL uygulamaları üzerinde çalışmıştır.

GitHub (github.com/osman-tkdmr): 15'in üzerinde açık kaynak deposuna sahip olup, özellikle “Moppie” film öneri sistemi, 5G konumlandırma ETL boru hattı ve Unity ile geliştirilmiş oyun prototipleri gibi projelerde toplam 300'den fazla commit ile aktif katkı sunmuştur. Depolardaki kod kalitesi, ayrıntılı dokümantasyon ve modüler tasarım yaklaşımlarıyla dikkat çekmektedir.

LinkedIn (linkedin.com/in/osman-tekdamar): 450'yi aşkın profesyonel bağlantıya sahiptir. Profilinde “Film Öneri Sistemleri”, “Makine Öğrenmesi” ve “Doğal Dil İşleme” başta olmak üzere 25'in üzerinde onaylanmış beceri yer almaktır; dönemsel teknik yazılar, proje sunumları ve sertifika duyuruları paylaşarak sektör paydaşlarıyla etkileşimde bulunmaktadır.

Akademik ilgi alanları; öneri sistemleri, doğal dil işleme, yapay zekâ uygulamaları, veri bilimi ve oyun geliştirme olarak öne çıkmaktadır.