

Advancing recommendation systems with DeepMF and hybrid sentiment analysis: Deep learning and Lexicon-based integration

ASBTRACT

Otel endüstrisinde müşteri memnuniyetinin sağlanması ve kişiselleştirilmiş önerilerin sunulması, unutulmaz bir misafir deneyimi oluşturma'nın temel unsurlarıdır. Ancak, geleneksel öneri sistemleri, etkinliklerini sınırlayan çeşitli zorluklarla karşı karşıya kalmaktadır. Bu zorluklar arasında, yeni veya az değerlendirilmiş öğelere yönelik önerilerde bulunmayı zorlaştıran soğuk başlangıç (cold start) problemi ile, ilgili bilgi eksikliğini ifade eden veri seyrekliği (data sparsity) yer almaktadır. Ayrıca, müşterilerin yorumlarında dile getirdikleri farklı duyguları (sentiment) doğru şekilde yorumlamak da önemli bir güçlük teşkil etmektedir.

Bu çalışma, müşteri yorumlarından elde edilen görüş ve duyguları yakalamak ve analiz etmek amacıyla, duygu analizi (sentiment analysis) yaklaşımını otel öneri sistemlerine entegre ederek bu zorlukları ele almaktadır. Araştırmanın temel amacı, hibrit bir duygu analizi yaklaşımını öneri sistemleriyle bütünleştirerek sistemlerin performansını artırmaktır.

Önerilen hibrit yaklaşım, sözcük temelli (lexicon-based) tekniklerle derin öğrenme yöntemlerini birleştirmekte; metin verisinin duygusunu analiz etmek için TextBlob, Bag of Words ve Çok Katmanlı Algılayıcı (Multilayer Perceptron - MLP) algoritmasını birlikte kullanmaktadır. Bu hibrit duygu analizi yaklaşımı, %88,63'lük etkileyici bir doğruluk oranı sergileyerek müşteri yorumlarındaki duyguları yakalamada yüksek bir etkinlik göstermektedir.

Bu entegrasyon sayesinde, öneri sistemleri müşteri duygularını daha iyi anlayarak kişiselleştirilmiş önerilerin kalitesini artırmaktadır. Ayrıca, hibrit duygu analizi yaklaşımı, işbirlikçi otel önerileri için DeepMF modeliyle birleştirilmiş ve 0.1 gibi dikkat çekici bir Kök Ortalama Kare Hatası (RMSE) değeri elde edilmiştir.

Duygu analizinin öneri sistemlerine entegre edilmesi, müşteri tercihleri hakkında değerli içgörüler sağlayarak önerilerin doğruluğunu ve kişiselleştirme düzeyini geliştirmektedir. Bu araştırma, otel endüstrisinde müşteri deneyimi yönetiminin optimize edilmesinde duygu analizinin potansiyelini vurgulamakta ve misafir memnuniyeti ile etkileşimini artırmak için değerli bir araç sunmaktadır.

1. INTRODUCTION

Bir seyahat planlamak karmaşık bir süreç olabilir ve bu sürecin en önemli unsurlarından biri, doğru oteli bulmaktır. Geniş bir yelpazede, farklı tercih ve ihtiyaçlara hitap eden oteller arasından seçim yapmak, gezginler için çoğu zaman zorlu bir görev haline gelir. Günümüz dijital çağında çevrimiçi yorumlar, bilgi edinme ve bilinçli seçimler yapma açısından paha biçilmez araçlar haline gelmiştir. Otel yöneticileri, bu yorumların işletmelerinin gelecekteki

başarısını şekillendirmedeki önemini iyi bilmektedir. Çevrimiçi yorumlar, geleneksel otel broşürlerine kıyasla daha özgün ve faydalı olarak görülmektedir. Ancak, çevrimiçi yorumların bolluğu zaman zaman bunaltıcı olabilmekte ve bu yorumlar arasından anlamlı bilgiler çıkarmak ciddi zaman ve çaba gerektirmektedir (Awotunde, Misra, Katta ve Adebayo, 2023).

Çevrimiçi yorumlar mevcut olmasına rağmen, gezginler sıklıkla bu yoğun geri bildirim yığını arasında etkili bir şekilde gezinmekte zorluk çekerler. Otel konaklamaları hakkında iyi bilgilendirilmiş kararlar almak için çok sayıda yorumu manuel olarak incelemek hem zaman alıcı hem de yorucu bir süreçtir. Bu durum, “karar yorgunluğu” olarak adlandırılan olguyu ortaya çıkarabilir; yani bilgi miktarının fazlalığı, bireyin güvenle karar verme yetisini zayıflatır. Dolayısıyla, yorum analiz sürecini sadeleştiren, gezginlerin ilgili bilgileri daha kolay ve verimli biçimde çıkarabilmesini sağlayan çözümlerin geliştirilmesi acil bir ihtiyaç haline gelmiştir.

Son yıllarda farklı alanlarda yaşanan ilerlemeler, bu soruna yönelik umut verici çözümler sunmaktadır. Örneğin, Roesser modelinde 2-Boyutlu Markov sığramalı sistemler için geliştirilen sonlu bölge asenkron H filtreleme kavramı, karmaşık ve dinamik veri ortamlarını yönetmeye yönelik gelişmiş teknikler sunmakta olup, bu durum çevrimiçi yorum verilerinin işlenmesi ve filtrelmesiyle benzerlik göstermektedir (Fang, Ren, Wang, Stojanovic ve He, 2024). Benzer şekilde, PDT-switched RDNN’ler için geliştirilen bozulma karşıtı durum kestirimi (anti-disturbance state estimation) teknikleri, çevrimiçi yorumlardaki değişkenlik ve gürültü gibi belirsizlikleri yönetmede dayanıklı ve uyarlanabilir sistemlerin önemini vurgulamaktadır (Song, Peng, Song ve Stojanovic, 2024). Ayrıca, iletişim kısıtlı sistemlerdeki nicemlenmiş yinelemeli öğrenme denetimi (quantized iterative learning control) ilkeleri, veri işlemede verimliliğin artırılması ve sistem performansının yinelemeli olarak geliştirilmesini ön plana çıkararak, kaynak kısıtlı ortamlarda öneri algoritmalarının iyileştirilmesine katkı sağlamaktadır (Tao, Tao, Zhuang, Stojanovic ve Paszke, 2024).

Gezginlerin sayısız çevrimiçi yorum arasında kaybolma sorunundan hareketle, bu çalışma yorumları analiz etme sürecini kolaylaştırmak amacıyla duygu analizini (sentiment analysis) öneri sistemine entegre etmeyi hedeflemektedir. Yorumlarda ifade edilen duyguların analiz edilmesiyle, her bir otelin olumlu ve olumsuz yönleri belirlenebilir ve gezginlere daha doğru, kişiselleştirilmiş öneriler sunulabilir. Bu yaklaşım, gezginlerin önceki misafirlerin duygu ve deneyimlerine dayanarak daha bilinçli kararlar almalarına olanak tanır; böylece tüm yorumları okumak için harcanan zamanı ve çabayı azaltır. Ayrıca, duygu analizinin öneri sistemine entegre edilmesiyle, otel işletmeleri müşteri geri bildirimleri hakkında değerli içgörüler elde ederek hizmet kalitesini artırabilir ve misafir memnuniyetini iyileştirebilir. Duygu analizinin öneri sistemine entegrasyonu, seyahat planlama sürecinin optimizasyonuna ve otel işletmelerinin genel başarısına önemli katkı sağlamaktadır.

Katkılar:

Bu çalışmanın katkıları aşağıdaki şekilde özetlenebilir:

- Duygu analizi için sözlük tabanlı teknikleri derin öğrenme modelleriyle birleştiren ve işbirlikçi filtreleme (collaborative filtering) temelli bir öneri sistemiyle entegre eden yeni bir yaklaşım önerdik. Özellikle, hibrit duygu analizi için TextBlob’u Çok Katmanlı Algılayıcı (Multi-Layer Perceptron - MLP) ile birlikte kullandık. Bu hibrit duygu analizi, ilgili parametreleri içerecek şekilde Derin Matris Faktörizasyonu (Deep Matrix Factorization - DeepMF) modeliyle entegre edilmiştir.
- Bag-of-Words, TF-IDF ve GloVe gibi çeşitli özellik çıkarım yöntemlerinin güçlü yönlerinden

yararlanan hibrit bir duygu analizi modeli uyguladık. Makine öğrenmesi ve derin öğrenme tekniklerini kullanarak bu yöntemlerin en uygun kombinasyonunu belirledik ve duygu analiz performansında kayda değer bir iyileşme sağladık.

- Duygu analizi sonuçlarını entegre ederek, modelimiz öneri sisteminin doğruluğunu önemli ölçüde artırmıştır. Hem modele dayalı (Non-negative Matrix Factorization - NMF, DeepMF) hem de belleğe dayalı (SlopeOne) yaklaşımları kullanarak, duygu analizinden elde edilen içgörülerini daha doğru ve kişiselleştirilmiş öneriler sunmak için değerlendirdik.
- Hibrit duygu analizi çıktısının, öneri sistemlerinin değerlendirilmesinde kullanıcı puanlarına alternatif bir metrik olarak kullanılabileceğini gösterdik. Duygu analizini öneri sürecine dahil ederek, sonuçların ve kullanıcı deneyimlerinin belirgin şekilde iyileştirilebileceğini ortaya koyduk.
- Önerilen modelimizi, alandaki diğer güncel yaklaşımlarla kapsamlı biçimde karşılaştırdık. Bulgularımız, duygu analizi ile entegre edilen öneri sistemi yaklaşımımızın üstün performansını açıkça ortaya koymuştur.

Makale Yapısı:

Makale beş ana bölümden oluşmaktadır: giriş, ilgili çalışmalar, önerilen sistem, sonuçlar ve tartışma ile sonuç bölümü. Giriş kısmı araştırmanın bağlamını ve problemini ortaya koyarken, ilgili çalışmalar bölümü alandaki önceki araştırmaları incelemektedir. Önerilen sistem bölümü, kullanılan yöntemleri ve teknikleri detaylandırmaktadır. Sonuçlar ve tartışma bölümü, deneysel bulguları ve analizlerini sunmaktadır. Son olarak, sonuç bölümü temel bulguları özetlemekte ve gelecekteki potansiyel araştırma yönlerini tartışmaktadır.

2. RELATED WORK

Bu literatür taramasının amacı, duygu analizinin öneri sistemlerine entegrasyonunu incelemek, bu alandaki gelişmeleri vurgulamak ve araştırmamızın ele almayı hedeflediği boşlukları belirlemektir. İlgili çalışmalar üç ana kategoriye ayrılmaktadır: **Öneri Sistemleri**, **Duygu Analizi** ve **Duygu Analizinin Öneri Sistemlerine Entegrasyonu**.

2.1 Recommendation system

Pan, Huo, Tang, Zeng ve Chen (2021), kullanıcı profillerinde etiket seyrekliği sorununu çözmek amacıyla **etiket farkındalıklı bir öneri sistemi (Tag-Aware Recommender System – TRS)** önermiştir. Geleneksel yöntemler benzer etiketleri genişletmiş ancak bu etiketleri ilgili kaynaklarla ilişkilendirememiştir; bu durum ise statik kullanıcı profillerine yol açmıştır. Yeni geliştirilen **sosyal etiket genişletme modeli (Social Tag Expansion Model – STEM)**, etiketleri kendi aralarındaki ilişkilere göre genişleterek ve seyrekliğe karşı yeterli sayıda etiketi dahil ederek dinamik kullanıcı profilleri oluşturur. STEM tabanlı TRS, üç temel işlem gerçekleştirir: ilgili etiketleri belirlemek için **etiket bulutu (tag cloud)** oluşturma, yeterli etiket kümesini bulmak için **etiket genişletme (tag expansion)** ve dinamik profiller inşa ederek genişletilmiş etiketlere ağırlık vermek için **kullanıcı profili yeniden yapılandırma (user profile refactoring)**. Analizler ve deneysel sonuçlar, bu tekniğin öneri doğruluğunu artırdığını ve mevcut etiket farkındalıklı öneri yöntemlerinden daha iyi performans gösterdiğini ortaya koymaktadır. Ancak STEM, etiket seyrekliğini etkili bir şekilde ele alırken, işbirlikçi filtrelemedeki veri seyrekliği gibi diğer seyreklik türlerini doğrudan ele almamaktadır.

Kuo, Chen ve Keng (2021), özellikle e-ticaret ve web hizmetleri alanlarında, **veri seyrekliği** ve **benzerlik seçim** konularındaki zorluklara odaklanmıştır. Çalışmada, veri seyrekliğinin etkisini azaltmak amacıyla **pertürbasyon tabanlı K-en yakın komşu (K-nearest neighbors)** ve **yoğunluk bazlı imputasyon (KDI-KNN)** algoritmasını birleştiren **hibrit bir meta-sezgisel (metaheuristic)** yöntem önerilmiştir. Benzerlik ölçümünü optimize etmek ve tahmin performansını artırmak için bir **benzerlik birleştirme fonksiyonu (similarities union function)** geliştirilmiştir. Deneysel sonuçlar, pertürbasyon tabanlı hibrit KDI-KNN algoritmalarının, temel KNN, orijinal KDI-KNN ve tek meta-sezgisel tabanlı KDI-KNN yaklaşımlarından daha iyi performans gösterdiğini ortaya koymaktadır. Bu araştırma, benzerlik ölçütlerinin veri kümesinin içeriğinden önemli ölçüde etkilendiğini vurgulamaktadır. Bununla birlikte, bu yaklaşım veri seyrekliğini ele alırken STEM modelinde görülen dinamik profil oluşturma özelliğini içermemektedir.

Bu yaklaşımlar önemli ilerlemeler sağlamış olsa da, literatürde hâlâ çeşitli boşluklar bulunmaktadır. İlk olarak, Pan'ın STEM modeli etiket seyrekliğini etkili biçimde ele almasına rağmen, **kullanıcı tercihlerini zamana bağlı olarak değişen dinamik bir biçimde** değerlendirmemektedir. Ayrıca, Kuo'nun hibrit meta-sezgisel yaklaşımı **benzerlik seçimine** odaklanmakta, ancak **kullanıcı çeşitliliği ve heterojen tercihlerin etkisini** göz ardı etmektedir. En önemlisi, her iki çalışma da **duygu analizini (sentiment analysis)** entegre etmemiştir; oysa duygu analizi, kullanıcı görüşleri hakkında daha derin içgörüler sağlayabilir ve öneri doğruluğunu artırabilir. Son olarak, her iki çalışma da yöntemlerini belirli veri kümeleri üzerinde değerlendirmiş olup, bu durum **bulguların genellenebilirliği** konusunda soru işaretleri doğurmaktadır. Bu boşlukların giderilmesi, daha **güçlü ve uyarlanabilir öneri sistemlerinin** geliştirilmesine katkı sağlayacaktır.

2.2. SENTIMENT ANALYSIS

Duygu analizi, özellikle konaklama sektöründe müşteri geri bildirimlerini anlamada hayati bir rol oynamaktadır. Kullanıcı yorumlarından içgörü elde etmek için çeşitli teknikler kullanılmış olsa da, duygu tespiti sürecinde doğruluk ve alaka düzeyinin artırılması konusunda hâlâ gelişime açık alanlar bulunmaktadır. Bu literatür incelemesi, mevcut duygu analizi yöntemlerini araştırmayı ve kullanıcı duygularını etkili bir şekilde yakalama ve sınıflandırmada karşılaşılan zorlukları belirlemeyi amaçlamaktadır.

Suryawanshi (2022), **konaklama sektörü yorumlarındaki temaları** incelemek amacıyla **doğal dil işleme (NLP)** tekniklerinden, özellikle konu modelleme ve duygu analizi yöntemlerinden yararlanmıştır. Çalışmada, tüketici davranışları üzerindeki farklı unsurların etkisini analiz etmek için **LDA (Latent Dirichlet Allocation)** konu modelleme yöntemi ve **TextBlob** Python kütüphanesi kullanılmıştır. Çalışma, tema tespiti ve duygu kategorilendirmede etkili sonuçlar üretmiş olsa da, TextBlob gibi **geleneksel duygu analizi yöntemlerine** dayanmakta ve bu yöntemler zaman zaman yorumlardaki **karmaşık bağlamsal nüansları** doğru şekilde anlamakta yetersiz kalabilmektedir.

Arroni, Galán, Guzmán-Guzmán, Núñez Valdez ve Gómez (2023), **sosyal ağlardaki görüş analizinde duygu analizinin önemini** vurgulayarak, **duygu tahmini için dönüştürücü (transformer)** mimarisi içerisinde **daha basit bir dikkat mekanizmasına (attention-based**

model) sahip bir model önermiştir. Model, temel duygu analizi tekniklerine göre üstün performans göstermiş ve hatta **DistilBERT** modelinden hız ve doğruluk açısından daha başarılı sonuçlar elde etmiştir. Bununla birlikte, bu yaklaşım verimli ve yalın bir yöntem sunarken, **model karmaşıklığı ile performans arasında bir denge gerekliliğini** de ortaya koymakta ve bu dengenin daha derinlemesine araştırılması gerektiğini göstermektedir.

Bu katkılara rağmen, literatürde hâlâ çeşitli boşluklar bulunmaktadır. **Suryawanshi'nin analizi**, duygu dinamiklerinin **zamansal değişkenliğini** dikkate almamaktadır; oysa duygular, olaylar ve eğilimlere bağlı olarak hızla değişebilir ve bu durum modelin **uygulanabilirliğini sınırlandırabilir**. **Arroni'nin çalışması** ise etkili bir yaklaşım sunmasına rağmen, **kullanıcı demografisi veya dış olaylar** gibi **bağlamsal bilgilerin entegrasyonunu** ele almamıştır; bu tür bilgiler, duygu tahmini doğruluğunu daha da artırabilir. Ayrıca, her iki çalışma da yöntemlerini **belirli veri kümeleri** üzerinde değerlendirmiştir; bu durum, bulguların **farklı bağlam ve sektörlerde genellenebilirliği** konusunda soru işaretleri doğurmaktadır.

Bizim çalışmamızda bu boşluklar, **hibrit duygu analizi (hybrid sentiment analysis)** gibi gelişmiş duygu analizi teknikleri kullanılarak ele alınmaktadır. Bu yaklaşım, birden fazla yöntemi birleştirerek **tüketici duygularını anlama konusunda doğruluk ve uyarlanabilirliği** artırmayı hedeflemektedir. Söz konusu yenilikçi yaklaşım, duygu analizi çerçevelerinin **farklı alanlarda daha sağlam ve uygulanabilir** hale gelmesini sağlamayı amaçlamaktadır.

2.3. Integration of sentiment analysis into recommendation system

Alotaibi ve diğerleri (2021) tarafından yapılan çalışma, sosyal medya verilerinin kullanıcı görüşlerinden öneriler çıkarmada kullanımını derinlemesine incelemektedir. Çalışmanın merkezinde, kelime gömme (word embedding) yaklaşımı ile **XGBoost sınıflandırıcısının** birlikte uygulanması yoluyla **öneri içeren cümlelerin sınıflandırılması** yer almaktadır. Araştırma, çevrimiçi otel yorumları ve Microsoft Windows App Studio'daki tartışmalardan elde edilen veri kümelerini ayrıntılı bir şekilde incelemiştir. Özellikle, **öneriyle ilişkili anahtar kelimeler ve ifadelerin belirlenmesinin** etkili bilgi çıkarımı açısından kritik özellikler olduğunun altı çizilmiştir. Bu titiz yaklaşım, yorumların analiz edilmesi ve öneri sistemlerinin geliştirilmesi için değerli içgörüler sunmaktadır.

Deac-Petruşel ve Limboi (2020), duygu analizi tekniklerinin öneri sistemlerine entegrasyonuna odaklanmıştır. Çalışmada, geleneksel **k-En Yakın Komşu (kNN)** algoritmasına iyileştirmeler getirilmiş; **duygu puanlama (sentiment rating)** yaklaşımı geliştirilmiş ve **duygu temelli kullanıcı benzerliği ölçütü (sentiment-based user similarity measure)** tanımlanmıştır.

Yadav (2022), **son teknoloji ürünü DeepFM modelini** kullanan yeni bir **hibrit öneri sistemi** önermektedir. Bu sistem, ürün yorumlarından elde edilen çoklu metinsel özellikleri — bağlamsal cümle gömme (contextual sentence embedding) vektörleri, ortalama duygu puanları ve olumsuzluk (negation) gibi dilbilimsel ipuçları — kullanıcı puanlarıyla birleştirerek tahmin doğruluğunu artırmaktadır. **Amazon** ve **Datafiniti** ticari veri kümeleri kullanılarak

değerlendirilen sistem, hem **puan tahmini (rating prediction)** hem de **öneri (recommendation)** görevlerini kapsamaktadır. Sonuçlar, **BERT** tarafından çıkarılan gömme vektörleri ve **VADER** tarafından türetilen duygu puanlarının entegre edilmesinin, yalnızca kullanıcı puanlarına dayalı sistemlere kıyasla tahmin doğruluğunu önemli ölçüde artırdığını göstermektedir.

Asani, Vahdat-Nejad ve Sadri (2021), **doğal dil işleme (NLP)** tekniklerini kullanarak, restoranlarla ilgili yorumlardan kullanıcıların yemek tercihlerini çıkaran **bağlama duyarlı bir öneri sistemi** geliştirmiştir. Sistem, yemek öğelerini kavramsal olarak gruplamak için **Wu-Palmer kümeleme yöntemini** kullanmakta ve yorumlarda bahsedilen yiyeceklerle yönelik **olumlu ve olumsuz görüşleri ayırt etmek amacıyla duygu analizi** uygulamaktadır. Kullanıcı tercihleri, konum, zaman ve diğer geri bildirimleri dikkate alarak, sistem kullanıcıların damak zevkine uygun yakındaki restoranları önermektedir. **TripAdvisor 2018 verileri** ile yapılan değerlendirmeler, sistemin önceki araştırmalara göre üstün performans gösterdiğini ortaya koymuş ve bu sistemi restoran öneri sistemleri alanında değerli bir katkı haline getirmiştir.

Elahi, Kholgh, Kiarostami, Oussalah ve Saghari (2023), kullanıcı yorumlarının duygu analizini öneri sürecine entegre ederek geliştirilmiş **yenilikçi bir hibrit öneri sistemi (RS)** önermiştir. Duygu analizi ve kişiselleştirilmiş önerilerin oluşturulması için gelişmiş algoritmalar kullanılmıştır. Hibrit öneri sistemi, **Amazon Digital Music** ve **Amazon Video Games** veri kümeleri üzerinde test edildiğinde, çeşitli temel modellerden daha iyi performans göstermiştir. Çalışmada, veri temizleme ve ön işleme için standart bir **NLP işlem hattı** kullanılmış; cümle gömme vektörleri **BERT** ve **Python Transformers kütüphanesi** aracılığıyla çıkarılmış; öneri amaçlı olarak **öğeler tabanlı işbirlikçi filtreleme, ALS, YoutubeRanker** ve **DeepFM** teknikleri uygulanmıştır.

Dang, Moreno-García ve De la Prieta (2021) tarafından yapılan bir diğer çalışmada, **duygu analizi (SA)** ile **tür (genre) tabanlı benzerliği** işbirlikçi filtreleme içerisinde birleştiren yeni bir öneri yöntemi önerilmiştir. Bu yöntem, tür ön işleme (genre preprocessing) için **BERT**'i ve duygu analizi için **derin öğrenme modellerini** kullanmaktadır. Film veri kümeleri üzerinde yapılan değerlendirmeler, bu yöntemin **SVD** ve **NMF** gibi geleneksel yöntemlere kıyasla önemli performans artışı sağladığını göstermektedir. **Kullanıcı tabanlı işbirlikçi filtreleme yaklaşımını** benimseyen bu yöntem, duygu analizi ve tür gömme (genre embedding) entegrasyonu yoluyla önerilerin güvenilirliğini artırmayı hedeflemektedir.

Ziani ve diğerleri (2017) tarafından gerçekleştirilen bir başka çalışmada, **duygu analizi (SA)** içeren **çok dilli bir öneri sistemi (RS)** geliştirilmiştir. Bu sistem, **Cezayirli kullanıcıların çevrimiçi yorumlara dayalı bilinçli kararlar almalarına yardımcı olmayı** amaçlamaktadır. Sistem, Cezayirli yorumları etkili bir şekilde analiz ederek duygu kutuplarını (polarity) tespit eder ve **Spearman benzerlik ölçütü** kullanan **kullanıcı tabanlı işbirlikçi filtreleme** aracılığıyla anlamlı öneriler üretir. RS ve SA'nın birleşimi, farklı dillerde yüksek doğrulukta öneriler sunmakta; veri kısıtlarını aşmak için **yarı denetimli Destek Vektör Makineleri (SVM)** kullanarak duygu kutbu puanlaması yapmaktadır.

Sahu ve diğerleri (2022) tarafından öne sürülen çalışmada, **duygu analizi (SA)** ile **hibrit bir öneri sistemi (RS)** birleştirilmiş ve henüz gösterime girmemiş filmleri önermek amacıyla bir çerçeve geliştirilmiştir. Başlangıçta, **Netflix'in resmi YouTube kanalındaki film**

fragmanlarına yapılan kullanıcı yorumları incelenerek film puanları tahmin edilmiştir. Ardından, **The Movie Database (TMDb)** verileri kullanılarak, bireysel kullanıcılar için **kişiselleştirilmiş, gelecekteki film öneri listeleri** oluşturulmuştur. Duygu temelli film puanlarının ve kullanıcı tercihlerinin birleştirilmesiyle **özelleştirilmiş öneriler** elde edilmiştir. Bu çalışma, özellikle **kullanıcıların sosyal medya yorumlarına dayalı olarak henüz yayımlanmamış filmlerin önerilmesine** odaklanmıştır; **VADER** ve **TextBlob**duygu analizi yöntemlerini puan tahmini için kullanmıştır.

Selmene ve Kodia (2020) tarafından yapılan bir diğer araştırmada ise, **TextBlob** tabanlı **Twitter verisi duygu analizi**, işbirlikçi filtreleme ile birleştirilmiştir. Bu yöntem, öneri sistemlerinin doğruluğunu artırmayı ve **seyreklik (sparsity)** ile **soğuk başlangıç (cold-start)** sorunlarını azaltmayı hedeflemektedir. Kullanıcı puanları ile tweet'lerdeki duygu analizinin birleştirilmesi sayesinde, öneri süreci daha zengin hale getirilmiştir.

Bizim çalışmamız, önceki araştırmaların üzerine inşa edilmekle birlikte, önemli yenilikler getirmektedir. Örneğin, biz **sözlük tabanlı teknikleri, makine öğrenmesi ve derin öğrenme modelleriyle** birleştiren özgün bir yaklaşım öneriyoruz. Bu yönüyle, duygu analizi yöntemlerini öneri sistemlerine entegre eden ancak **geleneksel tekniklerle sınırlı kalan** Deac-Petruşel ve Limboi (2020) ile Yadav (2022)'nin çalışmalarından ayrılmaktadır.

Duygu analizinin öneri sistemlerine entegrasyonu, Pan ve diğerleri (2021) tarafından da vurgulanmıştır. Bu araştırmacılar, **dinamik kullanıcı profil oluşturma** yoluyla öneri doğruluğunu artıran **etiket farkındalıklı bir öneri sistemi (tag-aware recommender system)** geliştirmiştir. Bu yaklaşım, etiket seyrekliğini etkili bir biçimde ele alırken, **işbirlikçi filtrelemedeki veri seyrekliğini** doğrudan ele almamaktadır. Benzer şekilde, Kuo ve diğerleri (2021) **veri seyrekliği** ve **benzerlik seçimi** konularındaki zorlukları ele almış; tahmin performansını artıran yöntemler önermiş, ancak **duygu analizini** entegre etmemiştir.

Bizim **hibrit duygu analizi modelimiz**, performansı artırmak için **bag-of-words, TF-IDF ve GloVe** gibi farklı özellik çıkarım yöntemlerini bir araya getirmekte; duygu analizi sonuçlarını ise **Deep Matrix Factorization (DeepMF)**, **Non-negative Matrix Factorization (NMF)** ve **SlopeOne** gibi gelişmiş tekniklerle öneri sistemine entegre etmektedir. Bu çok yönlü yaklaşım, **Suryawanshi (2022)** tarafından kullanılan geleneksel duygu analizi yöntemlerinin sınırlamalarını aşarak, önerilerin genel doğruluğunu ve kişiselleştirilmesini geliştirmektedir.

Tablo 1, bu ilgili çalışmalarda kullanılan duygu analizi teknikleri ve öneri sistemlerinin karşılaştırmalı bir genel görünümünü sunmakta; aynı zamanda **bizim katkılarımızı** ve **mevcut sınırlamaları** vurgulamaktadır.

Table 1

A comparative overview.

Paper	Sentiment analysis model	Recommendation system model	Proposed approach	Limit
Our work	Hybrid model (Lexicon-based with Deep Learning and Bag Of Words)	DeepMF with parameters	Combines lexicon-based methods (TextBlob) with ML (Random Forest, Naive Bayes) and DL (CNN, MLP) for sentiment analysis and select the best combination; integrates results into recommendation systems	Integrate transformer-based models to handle multi-class sentiment analysis and expand datasets across various domains
Deac-Petruşel and Limboi (2020)	Sentiment rating approach	Traditional k Nearest Neighbors	Introduces sentiment rating and sentiment-based user similarity measures to improve k-NN	Relies on traditional methods; may not capture complex sentiments
Yadav (2022)	BERT-Embedding and VADER	DeepMF	Proposes a hybrid recommendation system utilizing DeepFM; integrates multiple textual features and sentiment scores for accuracy	May not address data sparsity or the dynamic nature of user preferences.
Pan et al. (2021)	N/A	Tag-aware recommender system (TRS)	Proposes a tag-aware recommender system (TRS) that creates dynamic user profiles through tag expansion and user profile refactoring	Missing sentiment analysis; addresses only tag sparsity, not data sparsity in collaborative filtering
Kuo et al. (2021)	N/A	Hybrid metaheuristic with KDI-KNN for collaborative filtering	Reduces the impact of data sparsity and optimizes similarity measures for better performance	Missing sentiment analysis; limited to similarity measures.

3. PROPOSED MODEL

Önerilen metodolojimizde (Şekil 1), iki ana bileşene odaklanılmaktadır: **hibrit duygu analizi** ve **öneri sistemi**. Hibrit duygu analizi, kapsamlı bir yaklaşım elde etmek amacıyla birden fazla tekniğin entegrasyonunu içermektedir. Özellikle, **sözlük tabanlı yöntemleri** (örneğin **TextBlob**) **makine öğrenmesi algoritmaları** (örneğin **Random Forest** ve **Naive Bayes**) ve **derin öğrenme modelleri** (örneğin **CNN** ve **MLP**) ile birleştirmektedir. Bu entegrasyon, etkili duygu analizi modelini belirlemek için **bag-of-words (BoW)**, **TF-IDF** ve **GloVe** gibi çeşitli **özellik çıkarım yöntemlerinden** yararlanılmaktadır. Bu yaklaşım, metin verilerinden ayrıntılı duygu bilgilerini yakalamamızı ve daha derin içgörüler elde etmemizi sağlamaktadır.

Öneri sistemi bileşeninde ise **Negatif Olmayan Matris Faktörizasyonu (Non-Negative Matrix Factorization – NMF)**, **Derin Matris Faktörizasyonu (Deep Matrix Factorization – DeepMF)** ve **SlopeOne** yöntemlerini kullanılmaktadır. Hibrit duygu analizimizden elde edilen sonuçları entegre ederek **en yüksek performans gösteren modeli** seçmekteyiz; bu da öneri sisteminin doğruluğunu artırmaktadır. Duygu analizi sonuçlarını öneri çerçevesine dahil ederek, kullanıcıların **tercihlerini ve duygularını daha doğru yansıtan kişiselleştirilmiş öneriler** sunmayı hedeflemekteyiz.

3.1. TEXT DATA PRE-PROCESSING FOR REVIEWS

İlgili şekil (Şekil 2), duygu analizini kolaylaştırmada kritik bir aşama olan **veri ön işleme sürecinin adım adım ilerleyişini** göstermektedir. Veri ön işleme, analiz için uygunluğu sağlamak amacıyla yorum verilerini temizlemeye ve dönüştürmeye yönelik bir dizi işlemi içermektedir. Bu şekil, duygu analizi için metinsel verilerin hazırlanmasında yer alan çeşitli aşamaların görsel bir temsili sunmaktadır. Veri ön işleme, etkili bir duygu analizinin gerçekleştirilmesinde **temel bir rol** oynamaktadır. Bu süreçte, analiz için yorum verilerini hazırlamak açısından kritik öneme sahip bir dizi **veri temizleme işlemi** yer almaktadır. Bu işlemler arasında **tüm metnin küçük harfe dönüştürülmesi**, **noktalama işaretlerinin kaldırılması** ve gerektiğinde **veri özniteliklerinin dönüştürülmesi** bulunmaktadır. Metnin

bu şekilde standartlaştırılması, tutarlılığı sağlamanın yanı sıra duygu analizi sonuçlarının doğruluğunu da artırmaktadır (Darraz, Karabila, El-Ansari, Alami ve El Mallahi, 2025).

Veri temizleme işlemlerine ek olarak, ön işleme aşaması, verinin kalitesini daha da artırmak için çeşitli **Doğal Dil İşleme (NLP)** görevlerini de içermektedir. Bu görevlerden biri olan **tokenizasyon (sözcükleme)**, metnin bireysel birimlere veya sözcüklere ayrılması işlemidir. Bu adım, metinsel verinin daha iyi analiz edilmesini ve anlaşılmasını kolaylaştırır. Ayrıca, “a”, “the”, “is” gibi **sık kullanılan ve anlam taşımayan durak sözcüklerin (stop words)** kaldırılması, gürültüyü azaltarak daha anlamlı içeriklerin vurgulanmasına olanak tanır.

Bunun yanı sıra, veri ön işleme aşamasında **lema çıkarımı (lemmatization)** yöntemi kullanılmaktadır. Bu yöntem, kelimeleri kök veya temel biçimlerine indirger. Böylece kelimeler standartlaştırılır, veri boyutu (dimensionality) azaltılır ve duygu analizi algoritmalarının verimliliği artırılır. Bu NLP görevlerinin veri ön işleme aşamasına dâhil edilmesiyle, ilgisiz gürültülerin ortadan kaldırılması ve metnin standartlaştırılması sayesinde duygu analizinin kalitesi önemli ölçüde artırılmaktadır. Sonuç olarak, bu süreç **daha hassas duygu sınıflandırması ve analizinin** yapılmasını mümkün kılmaktadır (Darraz, Karabila, El-Ansari ve diğerleri, 2024).

Şekil 3, bir kullanıcının yorumuna uygulanan **ön işleme hattının (preprocessing pipeline)** somut bir örneğini sunmaktadır. Şekilde, kullanıcının **orijinal, işlenmemiş yorum metni** ile **ön işleme adımlarından geçirildikten sonra dönüştürülmüş hali** yan yana gösterilmektedir.

3.2. HYBRID SENTIMENT ANALYSIS METHOD PROPOSAL

Duygu analizi metodolojimiz, **doğruluğu ve etkinliği artırmak** amacıyla **sözlük tabanlı teknikleri, makine öğrenmesi ve derin öğrenme** yöntemleriyle birleştirmektedir. **Sözlük tabanlı yöntemler**, önceden tanımlanmış duygu sözlüklerini kullanarak metindeki olumlu veya olumsuz duyguları hızlı bir şekilde belirler. **Makine öğrenmesi algoritmaları**, etiketlenmiş verilerdeki örüntüleri (pattern) tanımlayarak yeni metinlerdeki duyguları tahmin edebilir. **Derin öğrenme** ise çok katmanlı sinir ağları aracılığıyla kelimeler ve ifadeler arasındaki karmaşık ilişkileri ortaya çıkarır; böylece metinlerdeki **ince ve nüanslı duygu ifadelerini** yakalayabilir. Bu entegre yaklaşım, metinsel verilerdeki duyguların **kapsamlı ve hassas bir analizini** sağlamaktadır (bkz. Şekil 4).

Makine öğrenmesi ve derin öğrenme tekniklerini uygulamadan önce, metinsel verilerimizin **nicel verilere dönüştürülmesi** gerekmektedir. Bu dönüşüm, **özellik çıkarımı (feature extraction)** süreciyle gerçekleştirilir. Özellik çıkarımı, metni bu algoritmaların anlayabileceği **sayısal temsillere dönüştürme** işlemidir. Metnin matematiksel olarak analiz edilebileceği bir biçime dönüştürülmesi sayesinde, makine öğrenmesi ve derin öğrenme modelleri verideki örüntüleri ve ilişkileri etkin bir biçimde öğrenebilir. Bu adım, **doğru ve anlamlı duygu analizi sonuçları** elde etmek için kritik bir öneme sahiptir.

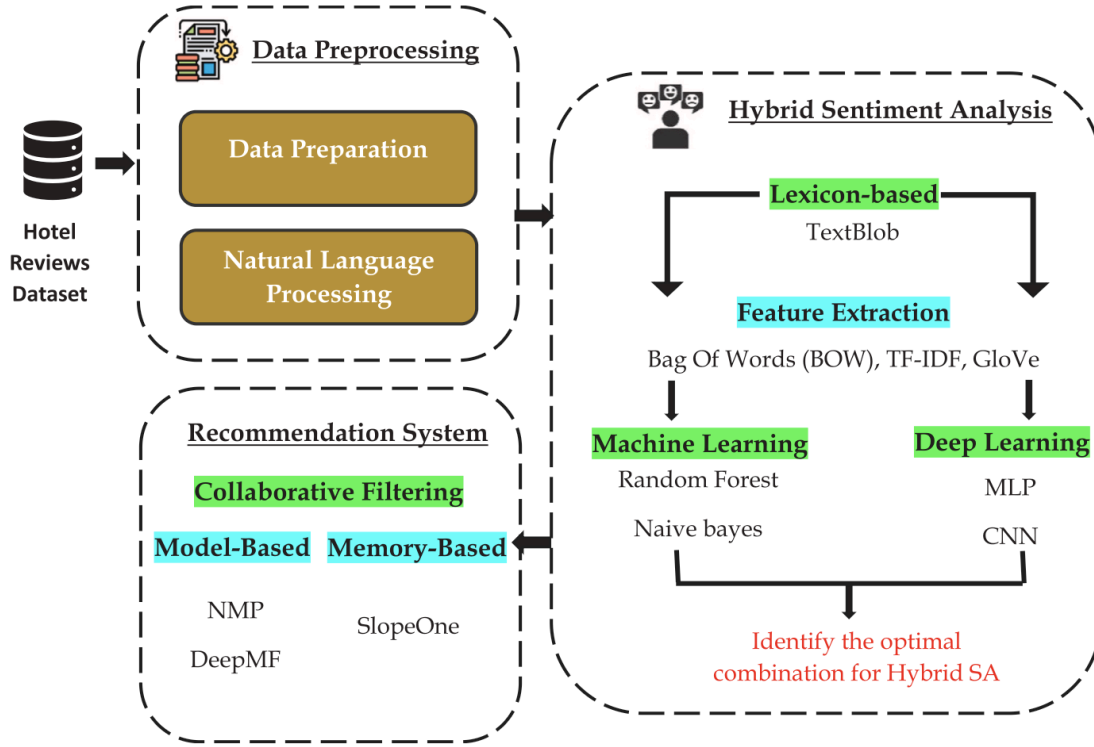


Fig. 1. Proposed methodology: Hybrid sentiment analysis and recommendation framework.

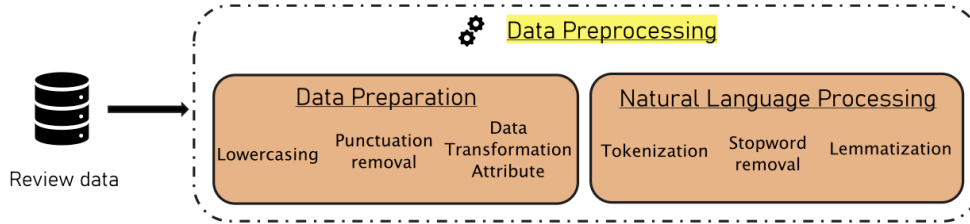


Fig. 2. Data preprocessing steps for analysis.

3.2.1. FEATURE EXTRACTION

Özellik çıkarımı (feature extraction) bölümümüzde, metinsel verileri hem makine öğrenmesi hem de derin öğrenme algoritmalarına uygun **nicel temsillere** dönüştürmek için çeşitli tekniklerden yararlanmaktayız. İlk olarak, **Bag of Words (BoW)** yaklaşımı, metni cümle yapısını dikkate almadan bir **kelime kümesine** dönüştürür. Bu yöntem sonucunda, her bir girdinin metindeki belirli bir terimin geçiş sıklığını temsil ettiği bir **özellik gösterimi** elde edilir. $m \times n \times n$ boyutlarında bir özellik matrisi oluşturularak, burada m cümle sayısını ve n ise derlemdeki (corpus) benzersiz kelimelerin sayısını ifade eder. Bu sayede BoW, metin verilerinin **sayısal biçimde analiz edilmesine** olanak tanır (Soumya & Pramod, 2020).

Bunun yanında, **Terim Frekansı – Ters Belge Frekansı (Term Frequency-Inverse Document Frequency – TF-IDF)** yöntemini de entegre etmekteyiz. TF-IDF, bir terimin bir belge koleksiyonu içindeki önemini değerlendiren bir tekniktir. Bu yöntem, bir terimin bir belge içindeki **geçiş sıklığını (TF)** ve tüm belgeler arasındaki **nadirliğini (IDF)** birleştirerek hesaplama yapar. **Bilgi erişimi, metin madenciliği ve belge sınıflandırma** gibi alanlarda

yaygın olarak kullanılan TF-IDF, metinlerdeki **önemli terimleri belirlemeye** ve belgeleri içeriklerine göre sıralamaya yardımcı olur (Kaiser & Ali, 2018).

Bu yöntemlere ek olarak, **GloVe (Global Vectors for Word Representation)** gömme (embedding) modelini de sistemimize entegre etmekteyiz; özellikle **GloVe-Wiki-Gigaword-300** modelinden yararlanmaktayız. **Kelime gömme modelleri**, kelimeler arasındaki **sözdizimsel (syntactic)** ve **anlamsal (semantic)** ilişkileri yakalayabilme yetenekleri nedeniyle, **metin sınıflandırma** ve **duygu analizi** çalışmalarında büyük ilgi görmektedir (Rezaeina, Rahmani, Ghodsi ve Veisi, 2019). **Wikipedia** ve **Gigaword** derlemelerinin birleşiminden elde edilen **küresel kelime birlikte bulunma istatistikleri** üzerinde eğitilen GloVe-Wiki-Gigaword-300 gömmeleri, zengin anlamsal bilgileri kapsayan yoğun vektör temsilleri sunar. Bu sayede metinsel verileri daha derinlemesine **anlama ve analiz etme** kapasitemiz artmaktadır.

Bu özellik çıkarım yöntemlerini bir araya getirerek, duygu analizi görevlerinde gerekli olan **karmaşıklık ve anlamsal incelikleri** etkili bir biçimde yakalayabilmekteyiz.

3.2.2. LEXICON BASED

Sözlük tabanlı duygu analizi çerçevemizde, metinsel verilerden duygu analizi yapmak ve duygu bilgilerini çıkarmak için **VADER (Valence Aware Dictionary and Sentiment Reasoner)**, **TextBlob** ve **SentiWordNet** araçlarının gücünden yararlanmaktayız.

VADER, özellikle **sosyal medya içeriklerindeki duyguları analiz etmek** üzere tasarlanmış, özel bir duygu analizi aracıdır. Duygu sözlüğünü (sentiment lexicon) sözcüksel özelliklerle birleştirerek, kelimeleri **anlamsal yönelimlerine (semantic orientation)** göre **olumlu** veya **olumsuz** olarak etiketler. VADER, dört farklı duygu ölçütü üretir: **olumlu**, **nötr**, **olumsuz** ve **bileşik (composite)** skor. Bileşik skor, genel duygu yoğunluğunu temsil eder ve tüm sözlük skorlarının toplamının normalize edilmesiyle -1 ile 1 arasında bir aralıkta hesaplanır. VADER'in önemli bir özelliği, **noktalama işaretlerini ve büyük harf kullanımını** dikkate almasıdır; böylece özel karakterleri filtrelemeden veya metni küçük harfe dönüştürmeden metindeki **duygunun tamamını** yakalayabilir (Chiny, Chihab, Bencharef ve Chihab, 2021).

Öte yandan, **TextBlob**, kapsamlı **metin işleme** ve **duygu analizi** yetenekleri sunan, yaygın olarak kullanılan bir **Python kütüphanesidir**. Çerçevemiz içinde, TextBlob'un sunduğu **sentiment (duygu)** fonksiyonunu kullanmaktayız. Bu fonksiyon iki temel duygu özelliği üretir: **kutupsallık (polarity)** ve **öznelik (subjectivity)**. Kutupsallık, metinde ifade edilen duygunun yönünü gösterir ve -1 ile 1 arasında değişen bir değer alır; sırasıyla **olumsuz**, **nötr** ve **olumlu** duyguları temsil eder. Öznelik ise metnin **kişisel görüş** içerme derecesini ölçer; 0 ile 1 arasındaki değerler, sırasıyla **yüksek nesnellik** ve **yüksek öznellik** anlamına gelir. TextBlob'un kullanımı sayesinde, metinlerimizde ifade edilen duyguları bu özellikler aracılığıyla kolaylıkla **analiz edebilir ve yorumlayabiliriz** (RamyaSri, Niharika, Maneesh ve Ismail, 2019).

Buna ek olarak, **SentiWordNet**'i de sistemimize entegre etmekteyiz. SentiWordNet, **WordNet**'teki **synset'lere (eşanlamlı kelime kümeleri)** duygu puanları atayan bir sözcüksel kaynaktır. Bu yapı, her bir synset için **olumluluk**, **olumsuzluk** ve **nötrlük** derecelerini sayısal olarak belirlememizi sağlar ve WordNet'in **hiyerarşik yapısı** temelinde duygu analizi yapılmasına imkân tanır.

Bu sözlük tabanlı araçlardan yararlanarak, **makine öğrenmesi** ve **derin öğrenme** çerçevelerine entegre edilebilecek **en uygun duygu analizi modelini** belirleyebilmekteyiz. Böylece, duygu analizinin **doğruluğunu ve etkinliğini** önemli ölçüde artırmaktayız.

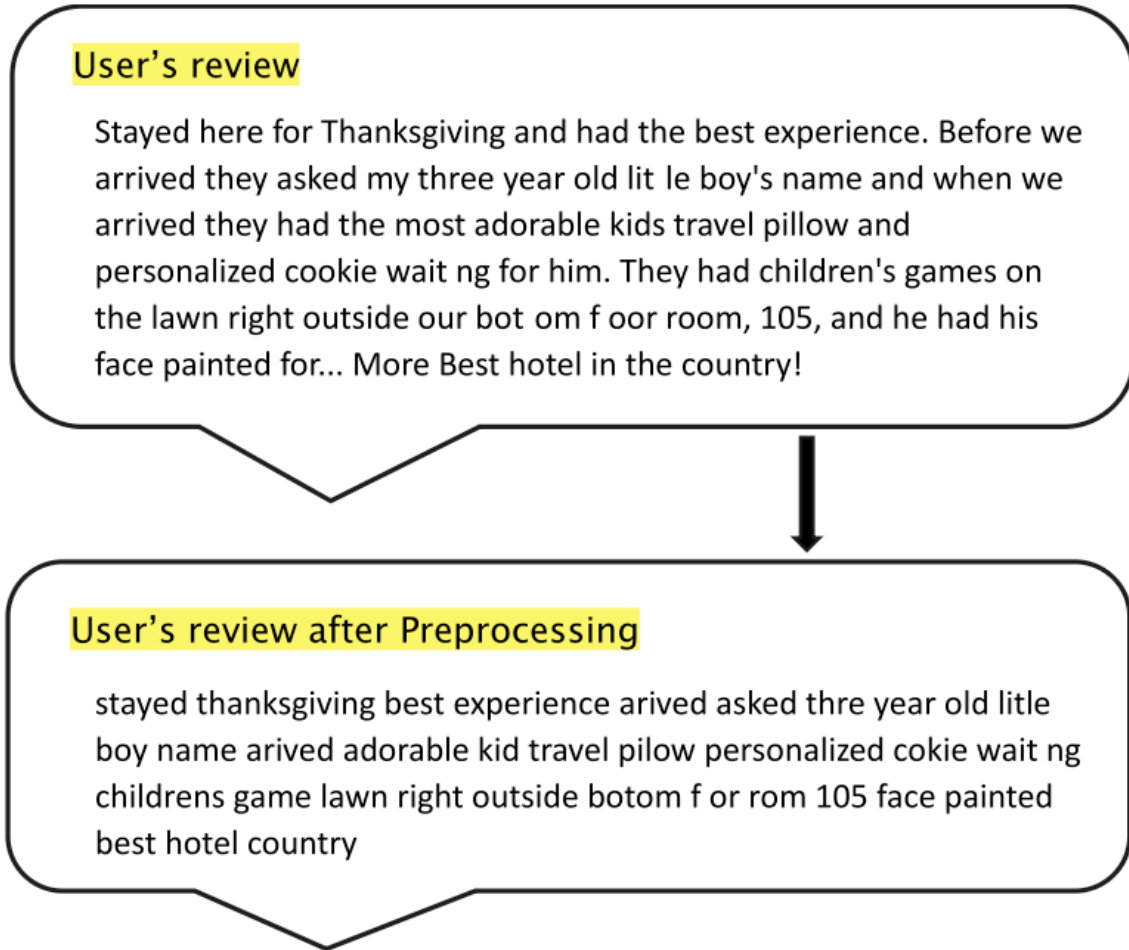


Fig. 3. Example of review preprocessing: Original and processed text.

3.2.3. MACHINE LEARNING

Duygu analizi çerçevemizde, makine öğrenmesi algoritmalarını sözlük tabanlı analizle entegre ederek hibrit bir yaklaşım benimsiyoruz. Özellikle, duygu sınıflandırmasını geliştirmek amacıyla Random Forest ve Naive Bayes algoritmalarını kullanıyoruz. Random Forest, birden fazla karar ağacını birleştirerek tahminler yapan güçlü bir topluluk (ensemble) öğrenme yöntemidir ve verideki karmaşık örüntüleri yakalamamıza olanak tanır. Öte yandan, Naive Bayes, özellikler arasında bağımsızlık varsayımına dayanan olasılıksal bir sınıflandırıcıdır ve duygu analizi görevlerinde etkili olduğu kanıtlanmıştır.

Modelin güvenilir bir şekilde değerlendirilmesi ve parametre ayarlamalarının optimize edilmesi için 5 katlı ve 10 katlı çapraz doğrulama (cross-validation) yöntemlerini kullanıyoruz. Çapraz doğrulama, veri kümesini birden fazla alt kümeye bölerek modelin eğitim ve test aşamalarını yinelenmeli biçimde gerçekleştirir. Bu yöntem sayesinde, daha sağlam

performans tahminleri elde edebilir ve aşırı öğrenme (overfitting) veya yetersiz öğrenme (underfitting) risklerini azaltabiliriz.

Makine öğrenmesi algoritmalarını sözlük tabanlı analizle entegre ederek ve çapraz doğrulama yönteminden yararlanarak geliştirdiğimiz duygu analizi çerçevemiz, metin verilerinden duyguları etkili biçimde yakalayıp analiz edebilmekte; böylece daha doğru ve güvenilir duygu sınıflandırma sonuçları sunmaktadır.

3.2.4. DEEP LEARNING

Duygu analizi çerçevemizde, doğruluk ve etkinliği artırmak amacıyla derin öğrenme modellerini sözlük tabanlı yöntemlerle entegre ediyoruz. Özellikle, sözlük tabanlı araçların sağladığı içgörülerini tamamlamak için **Evrışimsel Sinir Ağları (Convolutional Neural Networks – CNN)** ve **Çok Katmanlı Algılayıcılar (Multi-Layer Perceptrons – MLP)** kullanıyoruz.

Evrışimsel Sinir Ağı (CNN), çok katmanlı algılayıcı yapısını temel alan, ileri beslemeli derin bir sinir ağı türüdür ve minimum ön işleme gerektirir. Başlangıçta görüntü sınıflandırma ve bilgisayarlı görü (computer vision) görevleri için geliştirilmiş olan CNN'ler, zamanla çeşitli **doğal dil işleme (NLP)** uygulamalarına uyarlanmıştır. Metin üzerinde uygulandığında, metin bir boyutlu bir dizi (1D array) olarak temsil edilir ve CNN mimarisi, **1D evrişim (convolution)** ve **havuzlama (pooling)** işlemlerini içerir. CNN'ler, n-gram'leri dikkate alarak cümleleri önceden tanımlanmış kategorilere etkili biçimde sınıflandırabilir ve metin içindeki yerel bağımlılıkları ile örüntüleri yakalayabilir (Sharma, Chaurasia ve Srivastava, 2020).

Çok Katmanlı Algılayıcı (MLP) ise herhangi bir fonksiyonu yaklaşıklayabilme özelliğine sahip güçlü ve esnek bir sinir ağı modelidir ve bu yönüyle geniş bir uygulama yelpazesinde kullanılabilir. En az bir gizli katman içeren çoklu katmanlardan oluşur ve bu katmanlarda çok sayıda doğrusal olmayan birim bulunur. MLP, **evrensel bir fonksiyon yaklaşırtıcı (universal function approximator)** olarak çalışır ve girdi değişkenleri arasındaki karmaşık ilişkileri öğrenme yeteneğine sahiptir. MLP'de veri akışı tek yönlüdür: giriş katmanındaki her düğüm bir tahmin değişkenini temsil eder ve veriler gizli katmanlar aracılığıyla ilerleyerek çıktı katmanına ulaşır. Çıktı katmanının yapısı tahmin görevine bağlıdır; ikili sınıflandırma (binary classification) durumlarında tek bir nörondan oluşurken, çok sınıflı (non-binary) tahminlerde çıktı katmanı, tahmin görevindeki sınıf sayısını (N) temsil eden **N nöron**dan oluşur (Ahmad, Aftab, Muhammad ve Ahmad, 2017).

Modellerimizin dayanıklılığını sağlamak ve aşırı öğrenmeyi (overfitting) önlemek için **5 katlı ve 10 katlı çapraz doğrulama (cross-validation)** tekniklerini kullanıyoruz. Veri kümesinin birden fazla alt kümesi üzerinde model doğrulaması yaparak, modellerimizin genelleme yeteneklerini güvenle değerlendirebilir ve duygu analizi görevleri için performanslarını optimize edebiliriz.

Bu şekilde, **derin öğrenme modellerinin sözlük tabanlı yöntemlerle entegrasyonu ve titiz çapraz doğrulama süreçleri** sayesinde, duygu analizi sonuçlarında **üstün doğruluk ve güvenilirlik** elde ediyoruz.

3.3. Collaborative filtering-based recommendation system

Öneri sistemimizde, hibrit duygu analizi yaklaşımımızdan elde edilen sonuçları **işbirlikçi filtreleme (collaborative filtering)** yöntemine entegre ettik. **Bellek tabanlı (memory-based)** ve **model tabanlı (model-based)** yöntemleri birlikte kullanarak, önerilerimizin doğruluğunu ve etkinliğini artırıyoruz. Kullanıcı-öğe benzerlik hesaplamalarına duygu bilgilerini dâhil ediyor ve ayrıca kullanıcı tercihlerini ve duygularını daha iyi yakalayabilmek için bu bilgileri öneri modelimize ek özellikler (features) olarak entegre ediyoruz. Özellikle, **Negatif Olmayan Matris Faktörizasyonu (Non-negative Matrix Factorization – NMF)** ve **Derin Matris Faktörizasyonu (Deep Matrix Factorization)** gibi model tabanlı yaklaşımların yanı sıra, **SlopeOne** gibi bellek tabanlı bir yöntemi de kullanıyoruz.

Ayrıca, sistemin dayanıklılığını ve güvenilirliğini sağlamak amacıyla, öneri sistemimizin eğitim ve değerlendirme süreçlerinde **8 katlı çapraz doğrulama (cross-validation)** yöntemi uyguluyoruz. Duygu analizinin gücünü işbirlikçi filtrelemeye entegre ederek ve titiz doğrulama tekniklerinden yararlanarak, hem kullanıcıların açık derecelendirmelerini (explicit ratings) hem de kullanıcı yorumlarında ifade edilen duygularını dikkate alan **kişiselleştirilmiş öneriler** sunmayı hedefliyoruz.

Bu entegrasyon, özellikle **otelcilik sektöründe**, daha alakalı ve tatmin edici öneriler sunarak **genel kullanıcı deneyimini** önemli ölçüde geliştirmemize olanak tanımaktadır.

Şekil 5, hibrit duygu analizinin (SA) öneri sistemi (RS) içine nasıl entegre edildiğini ayrıntılı biçimde göstermektedir. Süreç, duygu analizi için en uygun kombinasyonu belirlemek amacıyla **sözlük tabanlı yöntemler**, **makine öğrenmesi (ML)** ve **derin öğrenme (DL)** tekniklerini birleştiren hibrit SA ile başlar. Bu hibrit SA'nın çıktısı, yeni bir sütun olarak kaydedilir ve öneri sistemi için **tahmini derecelendirme (predicted rating)** girdisi olarak kullanılır. Entegrasyon adımında, bu duygu temelli derecelendirme, öneri sistemine ek bağlamsal bilgi sağlayarak modelin performansını artırır. Bu sütun, tahmini derecelendirme olarak **NMF**, **DeepMF** ve **SlopeOne** gibi işbirlikçi filtreleme yöntemlerine aktarılır ve önerilerin üretilmesinde kullanılır.

Öneri sisteminin sağlamlığını ve doğruluğunu garanti altına almak için **8 katlı çapraz doğrulama** uygulanmıştır. Böylece, duygu analizinin öneri sistemi yapısına entegrasyonunun etkinliği doğrulanmıştır. Bu şekil, **hibrit SA'dan öneri sisteminin geliştirilmesine uzanan veri akışını** vurgulayarak sistemin tüm mimarisini görsel olarak ortaya koymaktadır.

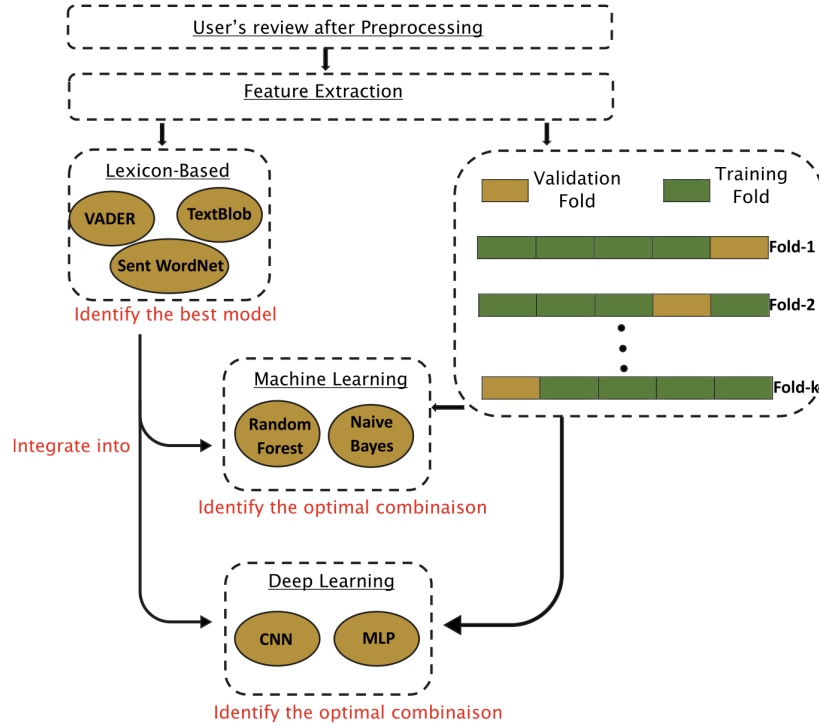


Fig. 4. Hybrid sentiment analysis: Integrating machine learning and deep learning approaches with lexicon-based methods for enhanced accuracy.

3.3.1. DeepMF

Derin Matris Faktörizasyonu (Deep Matrix Factorization – DMF), kullanıcılar ile öğeler arasındaki doğrusal olmayan ilişkileri yakalamayı amaçlayan, **işbirlikçi filtreleme (collaborative filtering)** için geliştirilmiş bir **derin öğrenme yaklaşımıdır**. Geleneksel matris faktörizasyonu tekniklerinden farklı olarak, DMF **derin sinir ağı mimarisi** kullanarak kullanıcı-öge etkileşim matrisinden doğrudan bir temsil (representation) öğrenir. Bu yaklaşım, basit gizil faktör (latent factor) modellerine kıyasla verideki **daha karmaşık örüntüleri modellemeye** olanak tanır.

Diğer sinir ağları tabanlı işbirlikçi filtreleme yöntemlerinden temel farkı, DMF'nin nihai tahmin aşamasında, öğrenilen kullanıcı ve öge temsilleri arasındaki **kosinüs benzerliğini (cosine similarity)** kullanmasıdır. Bu yönüyle DMF, nokta çarpımı (dot product) veya diğer benzerlik ölçütlerine dayanmak yerine, **kullanıcılar ve öğeler arasındaki ince ilişkiler**in sinir ağı aracılığıyla daha etkili biçimde yakalamayı hedefler.

Genel olarak, DMF, giriş verilerinde yer alan doğrusal olmayan kullanıcı-öge etkileşimlerini doğrudan modelleyerek **derin öğrenmenin işbirlikçi filtreleme görevlerine sistematik biçimde uygulanmasını** sağlar (Zhou, Wen, Li ve Zhou, 2019).

Bu şekil, **derin matris faktörizasyonunun mimarisini (Şekil 6)** göstermektedir.

Table 2

Statistics of the DataFiniti hotel dataset.

Dataset metric	Value
Number of Unique Hotels	1853
Number of Unique Users	6942
Total Number of Reviews	10,000
Rating Range	1 to 5 stars

4. Results and discussion

4.1. DataSet description

Araştırmamızda kullanılan veri kümesi, **DataFiniti'nin Business Database** veritabanından alınmış olup, özellikle **otel yorumlarına** odaklanmaktadır. Veri kümesi, her biri bir kullanıcının kiraladığı otelle ilgili **puanlamasını ve yorumunu** içeren toplam **10.000 satırlık kapsamlı bir derlemeden** oluşmaktadır. Veri kümesinde; **otel kimliği (ID)**, **adres**, **kullanıcı adı**, **otelin enlem (latitude)** ve **boylam (longitude)** bilgileri gibi çeşitli temel ayrıntılar yer almaktadır.

Veri kümesi, toplamda **1.853 benzersiz otel** ve **6.942 farklı kullanıcı** içermektedir. Bu zengin veri kümesi, **duygu analizi** yapılması ve **öneri sistemlerinin** geliştirilmesi için değerli bir kaynak sağlamaktadır. Duygu analizi, öznel bilgilerin sınıflandırılmasında kritik bir rol oynar; böylece müşteri deneyimleri ve görüşleri hakkında anlamlı içgörüler elde edilmesini sağlar. Öte yandan, öneri sistemleri kullanıcı puanlarını ve yorumlarını kullanarak **kişiselleştirilmiş otel önerileri** sunar; bu da kullanıcı memnuniyetini artırır ve **otel seçim sürecini optimize eder**.

Tablo 2, veri kümesine ilişkin temel metrikleri sunmakta olup, benzersiz otel ve kullanıcı sayısı, mevcut toplam yorum sayısı ve **1 ile 5 yıldız arasındaki puan aralığı** gibi önemli bilgileri içermektedir.

Bu çalışmada kullanılan veri kümesinin, öneri sisteminin etkinliğini etkileyebilecek **birkaç sınırlılığı** bulunmaktadır. Özellikle, bazı sütunlarda **eksik değerlerin (missing values)** bulunması, duygu analizinin doğruluğunu azaltabilir ve genel öneri kalitesini olumsuz etkileyebilir. Bu sorunu hafifletmek için, eksik değerlere sahip kayıtları kaldıran **“dropna” yöntemi** kullanılmıştır.

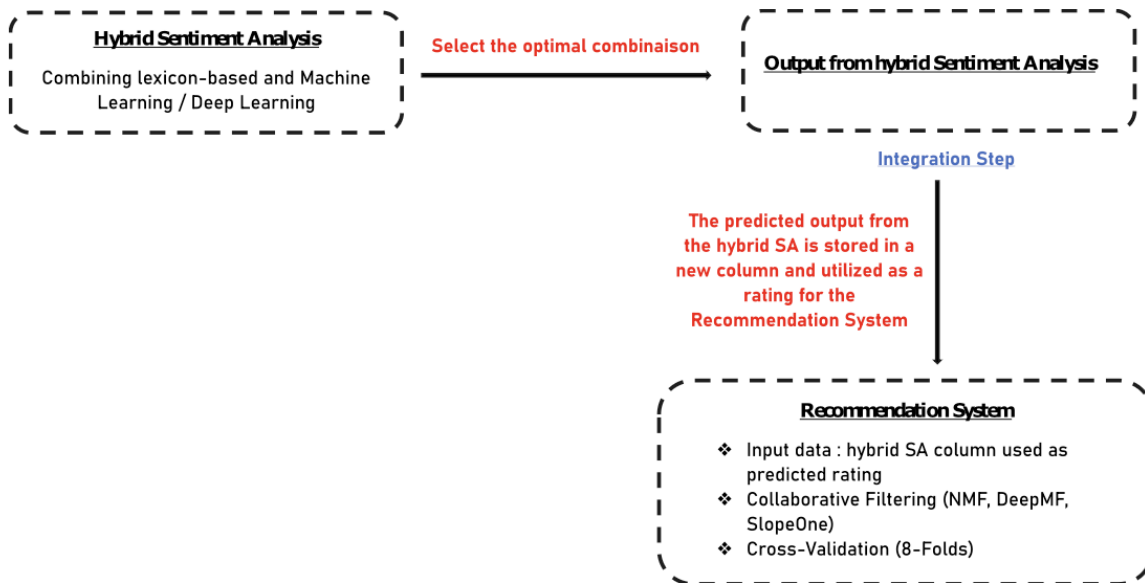


Fig. 5. Integrate hybrid SA into RS architecture.

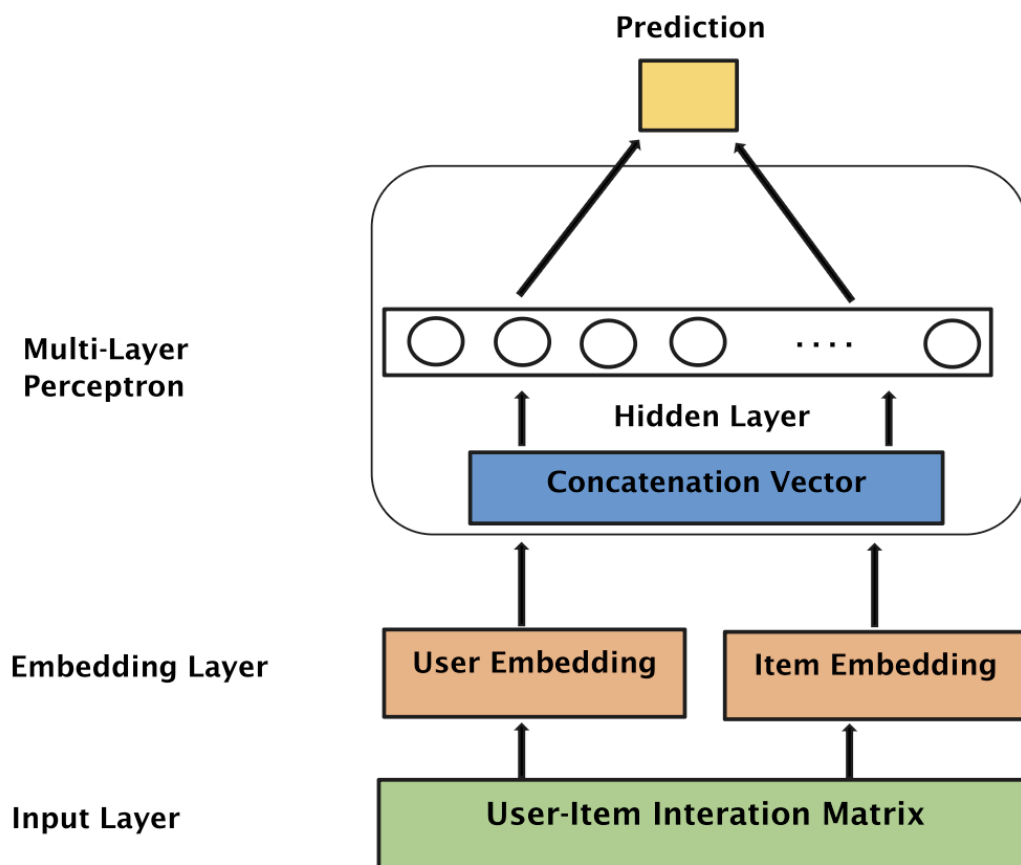


Fig. 6. Deep matrix factorization (DMF) architecture.

4.2. Model evaluation metrics

Önerilen modelimizin doğruluğunu ve güvenilirliğini değerlendirmek için kullanılan ölçütler şunlardır:

1. **Doğruluk (Accuracy):**

Doğruluk, doğru tahmin edilen örneklerin toplam örnek sayısına oranını ölçer.

Sınıflandırma modellerinin değerlendirilmesinde temel bir ölçüttür.

$\text{Doğruluk} = \frac{\text{Doğru Tahmin Sayısı}}{\text{Toplam Tahmin Sayısı}}$ (1) $\text{Doğruluk} = \frac{\text{Toplam Tahmin Sayısı}}{\text{Doğru Tahmin Sayısı}}$ (1)

2. **Matthews Korelasyon Katsayısı (MCC):**

MCC, ikili sınıflandırmaların kalitesini ölçen bir metriktir ve hem doğru hem de yanlış pozitif ve negatifleri dikkate alır. Değer aralığı -1 ile +1 arasındadır; +1 mükemmel sınıflandırmayı, 0 rastgele tahmini, -1 ise tahmin ile gözlem arasında tam bir uyumsuzluğu ifade eder (Chicco & Jurman, 2023).

$\text{MCC} = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$ (2) $\text{MCC} = \frac{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}{(TP \times TN) - (FP \times FN)}$ (2)

Burada, TP doğru pozitifleri, TN doğru negatifleri, FP yanlış pozitifleri ve FN yanlış negatifleri temsil eder.

3. **ROC-Eğrisi Altındaki Alan (ROC-AUC):**

ROC-AUC, Gerçek Pozitif Oranı (TPR) ile Yanlış Pozitif Oranı (FPR) arasındaki ilişkiyi gösteren ROC eğrisinin altında kalan alanı ölçer. Bu metrik, tüm olası eşik değerleri boyunca sınıflandırma performansının genel bir değerlendirmesini sunar.

4. **Kök Ortalama Kare Hata (RMSE):**

RMSE, tahmin edilen ve gerçek değerler arasındaki ortalama farkı ölçer. Kare hata ortalamasının karekökü alınarak hesaplanır. Düşük RMSE değerleri, daha yüksek doğruluk ve tahminlerle gözlemler arasındaki daha iyi uyum anlamına gelir (Katarya & Verma, 2017).

$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$ (3) $\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$ (3)

Burada n toplam gözlem sayısını, y_i gerçek değerleri, \hat{y}_i ise tahmin edilen değerleri ifade eder.

5. **Ortalama Mutlak Hata (MAE):**

MAE, tahmin edilen ve gerçek değerler arasındaki ortalama mutlak farkı ölçer. Model doğruluğuna ilişkin doğrudan bir değerlendirme sunar; daha düşük MAE değerleri, daha iyi performans ve tahminlerle gözlemler arasında daha yakın bir uyum anlamına gelir (Nawara & Kashef, 2020).

$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$ (4) $\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$ (4)

Burada n toplam gözlem sayısını, y_i gerçek değerleri, \hat{y}_i ise tahmin edilen değerleri temsil eder.

6. **Normalize Edilmiş İndirimli Kümülatif Kazanç (NDCG):**

NDCG, sıralama listelerinin veya öneri sistemlerinin kalitesini değerlendiren bir metriktir. Hem öğelerin ilgililik düzeyini hem de listedeki konumlarını dikkate alır.

Daha yüksek NDCG değerleri, daha iyi sıralama kalitesi ve daha alakalı öneriler anlamına gelir (Giabelli, Malandri, Mercorio, Mezzanzanica & Seveso, 2021).

$\text{NDCG} = \frac{\text{DCG}}{\text{IDCG}}$ (5) $\text{NDCG} = \frac{\text{IDCG}}{\text{DCG}}$ (5)

Burada DCG (İndirimli Kümülatif Kazanç) şu şekilde hesaplanır:

$\text{DCG} = \sum_{i=1}^n \frac{2^{\text{rel}_i}}{i+1}$ $\text{DCG} = \sum_{i=1}^n \frac{2^{\text{rel}_i}}{i+1}$

Burada rel_i , i . sıradaki sonucun ilgililik derecesini ifade eder. $IDCG$ (İdeal İndirimli Kümülatif Kazanç) ise sonuçların ilgililiğe göre mükemmel şekilde sıralandığı ideal durumu temsil eder.

Tablo 3.

5 katlı değerlendirme: Lexicon tabanlı tekniklerle ve teknikler olmadan Rastgele Orman (Random Forest) modellerinin karşılaştırması.

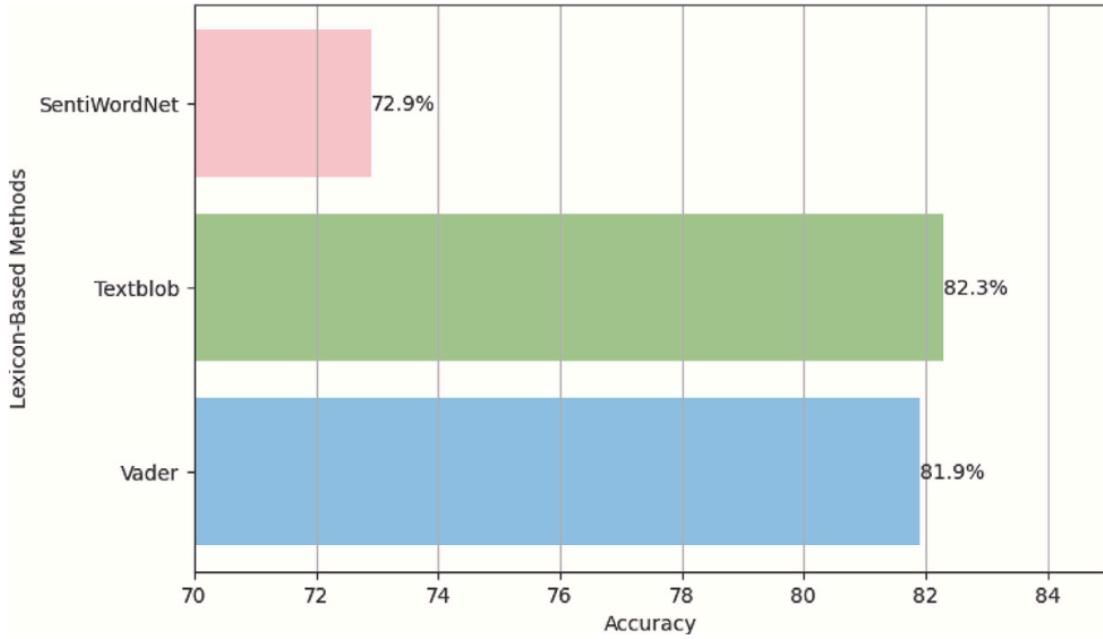


Fig. 7. Comparison of accuracy among lexicon-based sentiment analysis methods.

4.3. Sentiment analysis evaluation

4.3.1. Lexicon-based results

Lexicon tabanlı duygu analizi yöntemleri arasındaki doğruluk karşılaştırması (Şekil 7), özellikle **Datafiniti Otel** verisi alanında, performans açısından dikkate değer farklılıklar ortaya koymaktadır. **Textblob**, %82,3 doğruluk oranı ile en yüksek performansı sergilerken, onu %81,9 ile **Vader** yakından takip etmektedir. Buna karşılık, **SentiWordNet** %72,9 doğruluk oranı ile geride kalmaktadır. Bu farklılıklar, özellikle otel alanı gibi belirli görev ve veri kümesi özelliklerine uygun bir lexicon tabanlı yöntemin seçilmesinin önemini vurgulamaktadır.

Textblob'un üstün performansı, Vader ve SentiWordNet'e kıyasla otel verileri bağlamında duygu nüanslarını daha iyi yakalama yeteneğine sahip olduğunu göstermektedir. Bununla birlikte, en uygun yöntemin seçimi sırasında **hesaplama verimliliği** ve **alana uyarlanabilirlik** gibi diğer faktörlerin de dikkate alınması gerekmektedir.

Lexicon tabanlı yöntemlerin yanı sıra, **Textblob**'un **makine öğrenimi ve derin öğrenme modelleriyle bütünleştirilmesi**, özellikle **hibrit duygu analizi** bağlamında, performansı

daha da artırabilir. Textblob'un dil işleme yeteneklerinden, makine ve derin öğrenme algoritmalarının öğrenme kapasitesiyle birlikte yararlanılması, daha ince ayrıntılara ve bağlama duyarlı bir duygu analizi elde edilmesine olanak tanır.

Table 3

5-fold evaluation: Random Forest with and without lexicon-based techniques.

Model	Accuracy	MCC	ROC-AUC
BOW+RandomForest	84.56%	0.514	66.6%
BOW+RandomForest+TextBlob	86.3%	0.574	71.2%
TF-IDF+RandomForest	84%	0.49	65.9%
TF-IDF+RandomForest+TextBlob	85.54%	0.547	69.4%
GloVe+RandomForest	84.6%	0.51	68.8%
GloVe+RandomForest+TextBlob	85.24%	0.53	71%

4.3.2. Integrating lexicon-based with machine learning

Tablo 3, Random Forest algoritmasını kullanan çeşitli makine öğrenimi modelleri için, **TextBlob** lexicon tabanlı tekniğinin entegrasyonu ile ve entegrasyon olmadan elde edilen değerlendirme metriklerini sunmaktadır. Sonuçlar, TextBlob'un makine öğrenimi modellerine entegre edilmesinin, doğruluk (accuracy), **Matthews Korelasyon Katsayısı (MCC)** ve **Alıcı İşletim Karakteristiği Eğrisi Altındaki Alan (ROC-AUC)** metrikleri açısından genel olarak performans iyileşmesine yol açtığını göstermektedir.

Örneğin, **BOW+RandomForest** modeli %84,56 doğruluk oranı elde ederken, **TextBlob** eklendiğinde bu oran %86,3'e yükselmektedir. Benzer şekilde, **MCC** skoru 0,514'ten 0,574'e, **ROC-AUC** değeri ise %66,6'dan %71,2'ye çıkmaktadır. Benzer eğilimler **TF-IDF+RandomForest** ve **GloVe+RandomForest** modellerinde de gözlemlenmekte, TextBlob'un eklenmesi genel performansı artırmaktadır.

Bu bulgular, **TextBlob** gibi lexicon tabanlı tekniklerin makine öğrenimi modelleriyle bütünleştirilmesinin, yalnızca makine öğrenimi kullanılmasına kıyasla daha sağlam ve daha doğru tahminler üretebileceğini göstermektedir. Lexicon tabanlı bilginin, Random Forest algoritmasının öğrenme kapasitesini tamamlayarak çoklu değerlendirme metrikleri açısından performansı artırdığı görülmektedir.

Tablo 4, Random Forest algoritmasını kullanan çeşitli makine öğrenimi modelleri için, **TextBlob** lexicon tabanlı tekniğiyle ve bu teknik olmadan, **10 katlı çapraz doğrulama (10-fold cross-validation)** kullanılarak elde edilen değerlendirme metriklerini sunmaktadır. 5 katlı sonuçlarla benzer biçimde, TextBlob'un entegrasyonu doğruluk, MCC ve ROC-AUC metrikleri açısından performans artışı sağlamaktadır.

Örneğin, **BOW+RandomForest** modelinde doğruluk %84,9'dan %86'ya yükselmiştir. **MCC** skoru 0,525'ten 0,567'ye, **ROC-AUC** değeri ise %67,6'dan %71,4'e çıkmıştır. Benzer bir örüntü **TF-IDF+RandomForest** ve **GloVe+RandomForest** modellerinde de görülmektedir; TextBlob'un eklenmesi genel performansı artırmaktadır.

10 katlı sonuçlar, 5 katlı bulgularla tutarlıdır ve lexicon tabanlı tekniklerin (örneğin TextBlob) makine öğrenimi modelleriyle bütünleştirilmesinin, yalnızca makine öğrenimi kullanımına kıyasla daha güçlü tahmin yetenekleri sağladığı sonucunu pekiştirmektedir.

Tablo 5, Naive Bayes makine öğrenimi modellerinin, **TextBlob** lexicon tabanlı tekniğinin entegrasyonu ile ve entegrasyon olmadan gerçekleştirilen **5 katlı değerlendirme** sonuçlarını sunmaktadır.

Sonuçlara göre, **BOW+NaiveBayes** modelinde doğruluk oranı zaten oldukça yüksek olup %87,73'tür. TextBlob'un eklenmesiyle bu oran hafifçe artarak %87,88'e çıkmakta ve bu değer, tüm modeller arasında en yüksek doğruluk oranını temsil etmektedir. **MCC** skoru 0,63'te sabit kalırken, **ROC-AUC** değeri %79,4 olarak korunmaktadır.

Buna karşın, **TF-IDF+NaiveBayes** modeli çok daha düşük bir performans göstermekte; doğruluk %78,4, **MCC** 0,2 ve **ROC-AUC** %52,5'tir. Ancak, **TextBlob** entegre edildiğinde performans belirgin biçimde iyileşmekte; doğruluk %85,3'e, **MCC** 0,54'e ve **ROC-AUC** %74'e yükselmektedir.

GloVe+NaiveBayes modelinde ise doğruluk %82,5, **MCC** 0,52 ve **ROC-AUC** %77,4'tür. TextBlob'un eklenmesiyle doğruluk %82,6'ya, **MCC** 0,53'e yükselirken, **ROC-AUC** değeri sabit kalmaktadır.

Bu sonuçlar, **TextBlob** gibi lexicon tabanlı tekniklerin Naive Bayes modelleriyle entegrasyonunun etkisinin, kullanılan metin gösterim biçimine (BOW, TF-IDF veya GloVe) bağlı olarak değişebileceğini göstermektedir. **BOW+NaiveBayes** modeli zaten güçlü bir performans sergilerken, **TF-IDF+NaiveBayes** modeli TextBlob'un eklenmesiyle öngörü gücünde önemli bir artış göstermektedir. **GloVe+NaiveBayes** modeli ise yalnızca küçük bir gelişme göstermektedir.

En dikkat çekici sonuç, tüm modeller arasında en yüksek doğruluk oranına (%87,88) ulaşan **BOW+NaiveBayes+TextBlob** modelidir.

Tablo 6, farklı özellik temsilleri ve **TextBlob** kullanılarak, **Naive Bayes** modellerinin **10 katlı çapraz doğrulama** sonuçlarını sunmaktadır. **BOW+NaiveBayes** modeli %87 doğruluk oranına ulaşmakta, TextBlob eklendiğinde bu oran %87,54'e yükselmektedir. **TF-IDF** için doğruluk %78,8'den %85,4'e çıkmakta ve bu, kayda değer bir iyileşme anlamına gelmektedir. **GloVe+NaiveBayes** modelinde doğruluk %82,64'ten %82,8'e hafifçe yükselmektedir.

Bu sonuçlar, özellikle **TF-IDF** özellikleriyle birlikte **TextBlob** entegrasyonunun **Naive Bayes** model performansını artırmadaki faydasını açıkça ortaya koymaktadır.

Şekil 8, Random Forest ve Naive Bayes modellerinin, farklı özellik gösterim yöntemleriyle, **TextBlob** entegrasyonu öncesi ve sonrası performans karşılaştırmasını, **5 katlı ve 10 katlı çapraz doğrulama** sonuçları üzerinden göstermektedir.

Table 4

10-fold evaluation: Random Forest with and without lexicon-based techniques.

Model	Accuracy	MCC	ROC-AUC
BOW+RandomForest	84.9%	0.525	67.6%
BOW+RandomForest+TextBlob	86%	0.567	71.4%
TF-IDF+RandomForest	84.4%	0.5	66%
TF-IDF+RandomForest+TextBlob	85.9%	0.562	71%
GloVe+RandomForest	84.8%	0.52	69%
GloVe+RandomForest+TextBlob	85.4%	0.54	71.8%

Table 5

5-fold evaluation: Naive Bayes with and without lexicon-based techniques.

Model	Accuracy	MCC	ROC-AUC
BOW+NaiveBayes	87.73%	0.63	79.4%
BOW+NaiveBayes+TextBlob	87.88%	0.63	79.4%
TF-IDF+NaiveBayes	78.4%	0.2	52.5%
TF-IDF+NaiveBayes+TextBlob	85.3%	0.54	74%
GloVe+NaiveBayes	82.5%	0.52	77.4%
GloVe+NaiveBayes+TextBlob	82.6%	0.53	77.4%

Table 6

10-fold evaluation: Naive Bayes with and without lexicon-based techniques.

Model	Accuracy	MCC	ROC-AUC
BOW+NaiveBayes	87%	0.62	79%
BOW+NaiveBayes+TextBlob	87.54	0.628	79.5%
TF-IDF+NaiveBayes	78.8%	0.23	53%
TF-IDF+NaiveBayes+TextBlob	85.4%	0.55	75%
GloVe+NaiveBayes	82.64%	0.5	77%
GloVe+NaiveBayes+TextBlob	82.8%	0.53	77.7%

Table 7

5-fold evaluation: CNN with and without lexicon-based Techniques.

Model	Accuracy	MCC	ROC-AUC
BOW+CNN	77%	0.04	50%
BOW+CNN+TextBlob	77.3%	0.03	50.3%
TF-IDF+CNN	77%	0.02	50%
TF-IDF+CNN+TextBlob	77.3%	0.03	50.3%
GloVe+CNN	83%	0.5	73%
GloVe+CNN+TextBlob	84%	0.52	74%

4.3.3. Integrating lexicon-based with deep learning

Tablo 7’de sunulan sonuçlar, farklı öznitelik gösterim yöntemleri kullanılarak gerçekleştirilen Evrişimli Sinir Ağı (CNN) modellerinin performansını ve TextBlob duygu analizi entegrasyonunun etkisini göstermektedir. Değerlendirilen modeller arasında Kelime Torbası (Bag of Words – BOW), TF-IDF ve GloVe gömme (embedding) temsilleri yer almaktadır. TextBlob’un eklenmesi, çoğu modelde sınırlı da olsa bir iyileşme sağlamıştır.

BOW ve TF-IDF modelleri sırasıyla %77 ve %77,3 doğruluk oranlarına ulaşmış; TextBlob entegrasyonu yapıldığında MCC ve ROC-AUC değerlerinde yalnızca küçük farklar gözlenmiştir. Buna karşın, GloVe tabanlı CNN modelleri belirgin bir performans artışı göstermiştir: TextBlob olmadan %83 doğruluk, 0,5 MCC ve %73 ROC-AUC değerleri elde edilirken, TextBlob’un eklenmesiyle doğruluk %84’e, MCC 0,52’ye ve ROC-AUC %74’e yükselmiştir.

Genel olarak, TextBlob’un entegrasyonu BOW ve TF-IDF yöntemlerinde sınırlı bir etki gösterse de, GloVe gömmeleriyle performansı önemli ölçüde artırmaktadır. Bu durum, sözlük temelli duygu analizi ile derin öğrenme modellerinin birleştirilmesinin duygu sınıflandırma başarımını iyileştirmede etkili olduğunu ortaya koymaktadır.

Tablo 8’de sunulan sonuçlar, 10 katlı çapraz doğrulama yöntemi kullanılarak değerlendirilen farklı derin öğrenme modellerinin performansını göstermektedir. GloVe kelime gömmelerini CNN mimarisine girdi olarak kullanan derin öğrenme modelleri, genel olarak en güçlü performansı sergilemiştir. GloVe+CNN modeli %83,8 doğruluk, 0,52 MCC ve %74,8 ROC-AUC skorlarına ulaşmıştır. TextBlob duygu analiz aracının GloVe+CNN modeline entegre edilmesi doğruluğu %84,5’e, MCC’yi 0,53’e çıkarmış, ROC-AUC değerini ise %73 olarak kaydetmiştir.

Buna karşın, BOW veya TF-IDF temelli derin öğrenme modelleri, CNN ve TextBlob ile birleştirildiğinde dahi, %77,3 doğruluk, 0,03 MCC ve %51 ROC-AUC değerlerinin altında kalmıştır. Bu bulgular, GloVe gibi önceden eğitilmiş kelime gömmelerinin CNN gibi derin öğrenme modellerinde kullanılmasının etkinliğini vurgulamakta ve sözlük temelli tekniklerin entegrasyonunun performansı daha da iyileştirebileceğini göstermektedir.

Tablo 9’da sunulan sonuçlar, çok katmanlı algılayıcı (MLP) mimarisi kullanılarak gerçekleştirilen farklı derin öğrenme modellerinin 5 katlı çapraz doğrulama yöntemiyle değerlendirilmesini göstermektedir. BOW özniteliklerini kullanan modeller, genel olarak en yüksek performansı sergilemiştir. BOW+MLP modeli %88 doğruluk, 0,64 MCC ve %79

ROC-AUC değerlerine ulaşmıştır. TextBlob'un entegrasyonu doğruluğu %88,22'ye, MCC'yi 0,65'e, ROC-AUC'yi ise %80'e yükseltmiştir.

Buna karşın, TF-IDF veya GloVe gömmelerini kullanan MLP modelleri görece düşük performans sergilemiştir. Özellikle, TextBlob'un GloVe+MLP modeline entegre edilmesi performansta belirgin bir düşüşe neden olmuştur: doğruluk %73'e, MCC 0,01'e, ROC-AUC ise %50'ye gerilemiştir. Bu sonuçlar, BOW özniteliklerinin MLP modelleriyle birlikte kullanılmasının bu metin sınıflandırma görevi için oldukça etkili olduğunu, ancak önceden eğitilmiş gömmelerin (örneğin GloVe) belirli derin öğrenme mimarilerinde, özellikle de sözlük tabanlı tekniklerle birlikte kullanıldığında, zorluklar yaratabileceğini göstermektedir.

Tablo 10'da sunulan sonuçlar, MLP mimarisine sahip derin öğrenme modellerinin 10 katlı çapraz doğrulama yöntemiyle değerlendirildiği performansı göstermektedir. BOW özniteliklerini kullanan modeller yine en yüksek performansı sergilemiştir. BOW+MLP modeli %88 doğruluk, 0,63 MCC ve %79 ROC-AUC skorlarına ulaşmıştır. TextBlob'un entegrasyonu doğruluğu %88,63'e, MCC'yi 0,6637'ye ve ROC-AUC değerini %81,45'e yükseltmiştir.

Buna karşın, TF-IDF veya GloVe gömmelerini kullanan MLP modelleri daha düşük performans göstermiştir. Önceki bulgulara benzer şekilde, TextBlob'un GloVe+MLP modeline entegre edilmesi performansta belirgin bir düşüşe neden olmuştur: doğruluk %70'e, MCC 0,04'e, ROC-AUC ise %51'e gerilemiştir. Bu sonuçlar, BOW özniteliklerinin MLP modelleriyle birlikte kullanılmasının metin sınıflandırma görevlerinde etkili olduğunu bir kez daha doğrulamakta; buna karşın, önceden eğitilmiş gömmelerin (özellikle GloVe) sözlük temelli tekniklerle birleştirildiğinde bazı performans kayıplarına yol açabileceğini göstermektedir.

Duygu analizi için derin öğrenme yöntemlerinde, Rastgele Arama (Random Search) algoritması, hiperparametre uzayını rastgele örnekleyerek model performansını optimize eder; böylece doğruluk ve verimliliği artırır. Tablo 11 ve 12, sırasıyla CNN ve MLP modelleri için kullanılan hiperparametreleri göstermektedir. Bu tablolar, her model için duygu analizi amacıyla belirlenen hiperparametre yapılandırmalarını özetlemektedir. Hiperparametreler, doğrudan veriden öğrenilmeyen ancak eğitim sürecinden önce ayarlanan konfigürasyon değerleridir.

Şekil 9, CNN ve MLP modellerinin, farklı öznitelik gösterim yöntemleriyle ve TextBlob entegrasyonu olup olmadan, 5 katlı ve 10 katlı çapraz doğrulama yöntemleriyle değerlendirilen performans karşılaştırmasını göstermektedir.

TextBlob'un derin öğrenme modelleriyle entegrasyonu duygu sınıflandırma performansında umut verici iyileşmeler gösterse de, bir sınırlama olarak yürütme süresinin uzadığı gözlemlenmiştir. Derin öğrenme algoritmaları, özellikle TextBlob gibi sözlük temelli yöntemlerle birleştirildiğinde, daha uzun işlem sürelerine yol açabilir. Bunun nedeni, hem derin öğrenme modellerinin eğitiminin hesaplama açısından yoğun olması hem de sözlük tabanlı duygu analizinin her bir metin örneği için ek işlem gerektirmesidir. TextBlob'un dahil edilmesi, her metin örneğine ek bir analiz katmanı eklediği için işlem süresini artırmaktadır. Bu sınırlama, özellikle büyük veri kümeleri veya gerçek zamanlı uygulamalarda, hızın kritik olduğu durumlarda daha belirgin hale gelir.

Dolayısıyla, hibrit yaklaşım doğruluk ve yorumlanabilirliği artırsa da, daha fazla hesaplama kaynağı ve zaman gerektirebilir; bu da pratik uygulamalarda dikkate alınması gereken bir denge unsurudur.

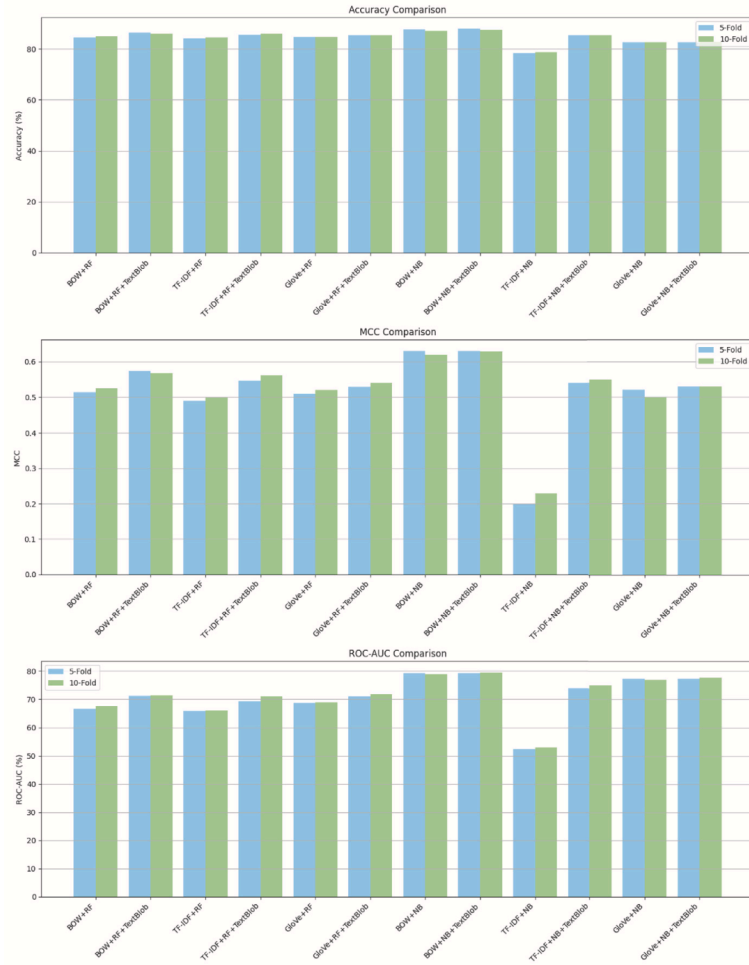


Fig. 8. Performance comparison of machine learning models with lexicon-based techniques (5-fold and 10-fold cross-validation).

Table 8

10-fold evaluation: CNN with and without lexicon-based Techniques.

Model	Accuracy	MCC	ROC-AUC
BOW+CNN	77.23%	0.02	50%
BOW+CNN+TextBlob	77.26%	0.01	50%
TF-IDF+CNN	77.23%	0.02	50%
TF-IDF+CNN+TextBlob	77.3%	0.03	51%
GloVe+CNN	83.8%	0.52	74.8%
GloVe+CNN+TextBlob	84.5%	0.53	73%

Table 9

5-fold evaluation: MLP with and without lexicon-based Techniques.

Model	Accuracy	MCC	ROC-AUC
BOW+MLP	88%	0.64	79%
BOW+MLP+TextBlob	88.22%	0.65	80%
TF-IDF+MLP	87%	0.62	79%
TF-IDF+MLP+TextBlob	87.4%	0.62	79%
GloVe+MLP	86%	0.58	76.4%
GloVe+MLP+TextBlob	73%	0.01	50%

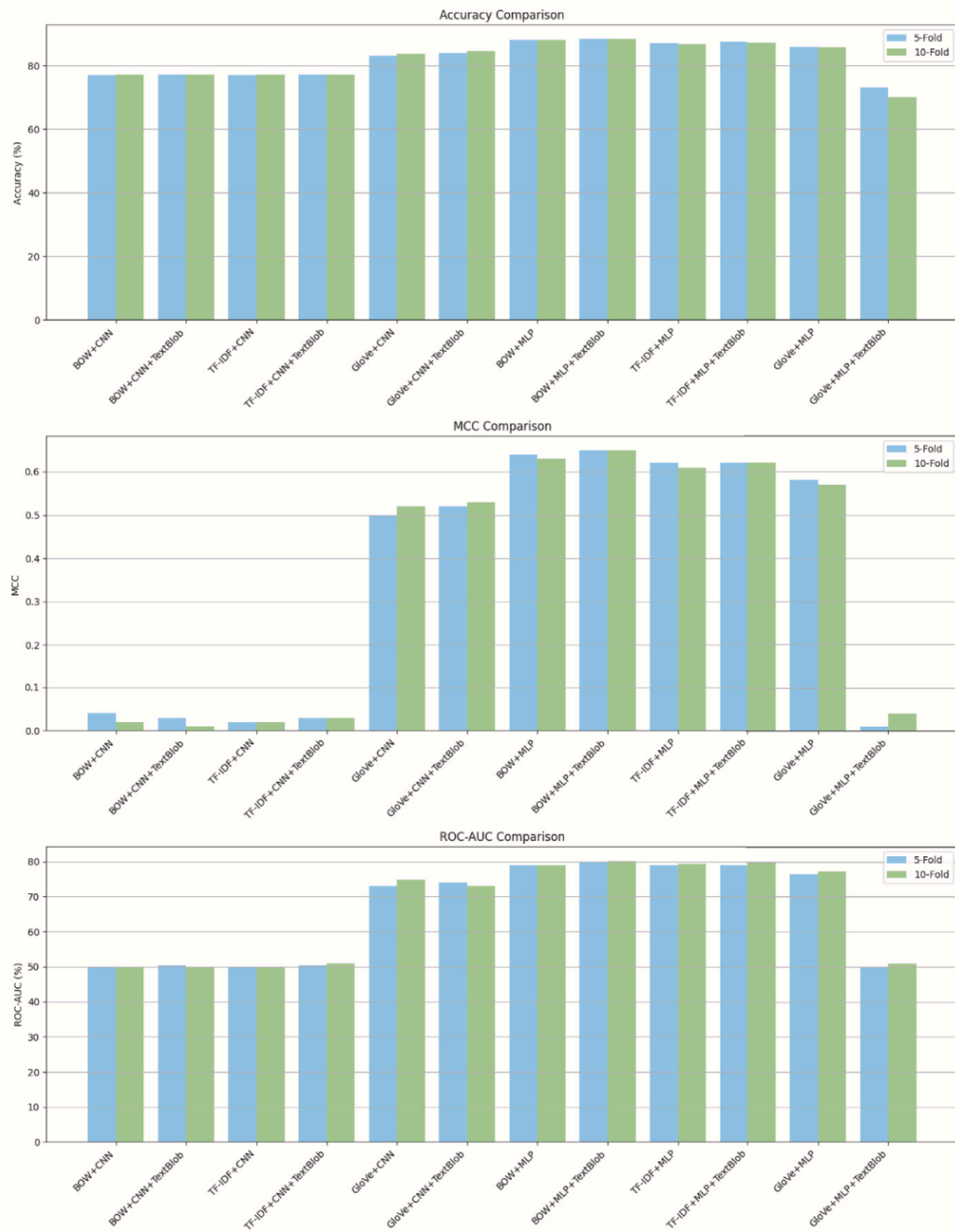


Fig. 9. Performance comparison of deep learning models with lexicon-based techniques (5-fold and 10-fold cross-validation).

Table 10

10-fold evaluation: MLP with and without lexicon-based techniques.

Model	Accuracy	MCC	ROC-AUC
BOW+MLP	88%	0.63	79%
BOW+MLP+TextBlob	88.63%	0.6637	81.45%
TF-IDF+MLP	86.7%	0.61	79.4%
TF-IDF+MLP+TextBlob	87.3%	0.62	79.7%
GloVe+MLP	85.7%	0.57	77.2%
GloVe+MLP+TextBlob	70%	0.04	51%

Table 11

Hyperparameter configuration for CNN.

Module	Description	Parameters	Value
Dropout	Regularization layer to prevent overfitting	Rate	0.2
Conv1D	Convolutional layer for feature extraction	Filters kernel_size Activation	50, 100 3 'relu'
Dense	Fully connected layer	Units Activation	200 'relu'
Dense	Output layer for binary classification	Units Activation	1 'sigmoid'
Compile	Compilation step for the Keras model	Loss Metrics Optimizer	'binary_crossentropy' ['acc'] 'adam'

Table 12

Hyperparameter configuration for the MLP model.

Parameter	Description	Value
layer_1_units	Number of units in the first hidden layer	200
activation_1	Activation function for the first hidden layer	'relu'
dropout_1_rate	Dropout rate for the first hidden layer	0.2
layer_2_units	Number of units in the second hidden layer	100
activation_2	Activation function for the second hidden layer	'relu'
dropout_2_rate	Dropout rate for the second hidden layer	0.2
output_units	Number of units in the output layer	1
activation_output	Activation function for the output layer	'sigmoid'
loss	Loss function used for binary classification	'binary_crossentropy'
optimizer	Optimizer for training the model	Adam(lr=0.0001)
batch_size	Number of samples per gradient update	16
epochs	Number of epochs to train the model	5

4.3.4. Optimal sentiment analysis models for recommendation systems

Sonuçlar (Şekil 10), daha gelişmiş teknikler kullanıldıkça duygu analizi doğruluğunda kademeli bir artış olduğunu göstermektedir. Başlangıçta, TextBlob %82,3'lük tatmin edici bir doğruluk oranı elde ederek temel düzeyde duygu sınıflandırmasında etkinliğini ortaya koymaktadır. Devamında, Kelime Torbası (Bag-of-Words – BOW) yaklaşımının Naive Bayes ile birleştirilmesi belirgin bir iyileşme sağlamış ve %87,73 doğruluk oranına ulaşmıştır. Bu iyileşme, TextBlob'un duygu analizi çıktılarının BOW ve Naive Bayes modeliyle entegre edilmesiyle daha da artmış, doğruluk oranı %87,88'e yükselmiştir.

BOW çerçevesine Çok Katmanlı Algılayıcı (Multilayer Perceptron – MLP) mimarisinin eklenmesi, doğruluğu %88'e çıkararak sinir ağı tabanlı yaklaşımların metinsel verilerdeki karmaşık örüntüleri yakalamadaki gücünü vurgulamaktadır. Dikkat çekici biçimde, BOW,

MLP ve TextBlob'un bir araya getirilmesi en yüksek doğruluk oranı olan %88,63'e ulaşmış ve bu yöntemlerin birbirini tamamlayıcı niteliğini açıkça göstermiştir.

BOW + MLP + TextBlob modelinin üstün performansı göz önüne alındığında, bu çıktının bir öneri sistemine girdi özelliği olarak sorunsuz biçimde entegre edilmesi mümkündür. Duygu analizi sonuçlarının öneri sistemine dahil edilmesiyle, sistem kullanıcıların tercihlerini ve duygularını daha iyi anlayabilir; bu da daha kişiselleştirilmiş ve isabetli önerilere yol açar. Bu entegrasyon, genel kullanıcı deneyimini güçlendirmenin yanı sıra, öneri sisteminin etkinliğini de artırabilir.

4.4. Collaborative filtering results in recommendation system analysis

Sunulan tablolar, duygu temelli ve orijinal puanların tahmininde kullanılan çeşitli modellerin üç performans metriği — Kök Ortalama Kare Hatası (Root Mean Square Error, RMSE), Ortalama Mutlak Hata (Mean Absolute Error, MAE) ve Normalize Edilmiş İndirgenmiş Kümülatif Kazanç (Normalized Discounted Cumulative Gain, NDCG) — açısından değerlendirme sonuçlarını göstermektedir.

RMSE tablosunda (Tablo 13), DeepMF, duygu temelli puanlamalarda en düşük RMSE değerine ulaşarak, NMF ve SlopeOne modellerine kıyasla duygu temelli puanları tahmin etmede üstün doğruluğunu ortaya koymuştur. Ancak, orijinal puanlar dikkate alındığında, DeepMF diğer modellerden daha iyi performans göstermeye devam etse de RMSE değeri nispeten daha yüksektir. Benzer şekilde, MAE tablosunda (Tablo 14) DeepMF, duygu temelli puanlamalarda en düşük MAE değerine ulaşarak etkinliğini göstermekte ve orijinal puanlamalar için de diğer modellere göre üstünlüğünü korumaktadır. NDCG tablosuna (Tablo 15) geçildiğinde, tüm modellerin yüksek NDCG skorları elde ettiği görülmektedir; bu durum, modellerin duygu temelli sıralamada etkili olduklarını göstermektedir. Dikkat çekici bir şekilde, tüm modeller orijinal puanlamalar için aynı NDCG skorunu üretmiştir, bu da bu bağlamda benzer sıralama performansına işaret etmektedir. Bu bulgular, DeepMF modelinin farklı değerlendirme metrikleri boyunca sağlam bir performans sergilediğini, özellikle duygu temelli puanları tahmin etmede öne çıktığını ve duygu analizi ile öneri sistemleri için umut verici bir model olduğunu vurgulamaktadır.

Tablo 16, DeepMF modelinin temel parametrelerini özetlemektedir; bu parametreler arasında kat sayısı (fold), dönem (epoch) sayısı, yığın boyutu (batch size) ve öğrenme oranı (learning rate) bulunmaktadır.

Orijinal puanlar ve duygu temelli puanlarla eğitilmiş NMF modellerinin karşılaştırılması, tahmin doğruluğu ve sıralama kalitesi açısından önemli farklılıkları ortaya koymaktadır. Orijinal puanlarla eğitilen NMF modeli (Şekil 11), duygu temelli puanlarla eğitilen NMF modeline (Şekil 12) kıyasla belirgin şekilde daha yüksek RMSE ve MAE değerleri üretmiştir. Orijinal puanlı NMF modeli ortalama yaklaşık 1,2 RMSE ve 0,9 MAE değerlerine ulaşırken, duygu temelli NMF modeli bu hata metriklerinde belirgin düşüşler göstermiştir: ortalama RMSE yaklaşık 0,4 ve ortalama MAE yaklaşık 0,31 olarak elde edilmiştir. Bununla birlikte, orijinal puanlı NMF modeli neredeyse kusursuz bir sıralama performansı göstermekte (ortalama NDCG: 0,9998), duygu temelli NMF modeli ise biraz daha düşük bir ortalama NDCG değeri (0,8037) sergilemektedir. Bu sonuçlar, sıralama kalitesinde küçük bir düşüşe

rağmen, duygu bilgisinin entegrasyonunun öneri doğruluğunu artırmadaki önemini vurgulamaktadır.

Orijinal puanlarla eğitilen DeepMF modeli (Şekil 13'te gösterildiği gibi), duygu temelli puanlara kıyasla daha düşük RMSE ($0,3193 \pm 0,0133$) ve MAE ($0,2559 \pm 0,0102$) değerleriyle üstün performans sergilemiştir. Ayrıca, NDCG skoru da daha yüksek ($0,9349 \pm 0,0091$) olup daha iyi sıralama doğruluğuna işaret etmektedir. Buna karşın, duygu temelli model (Şekil 14), daha yüksek RMSE ($0,4241 \pm 0,0205$) ve MAE ($0,3383 \pm 0,0233$) değerleriyle birlikte daha düşük bir NDCG skoru ($0,8103 \pm 0,0291$) göstermekte; bu da tahmin doğruluğu ve sıralama hassasiyetinde azalmaya işaret etmektedir.

SlopeOne modellerinin orijinal ve duygu temelli puanlara göre performans karşılaştırması, duygu bilgisinin dahil edilmesiyle belirgin bir iyileşme olduğunu göstermektedir. Orijinal puanlarla eğitilen model (Şekil 15'te gösterildiği üzere), farklı katlarda RMSE değerleri 1,1121 ile 1,2668 arasında, MAE değerleri ise 0,8308 ile 0,9781 arasında değişmektedir. Buna karşın, duygu temelli SlopeOne modeli (Şekil 16), belirgin şekilde daha düşük RMSE ($0,3806-0,4178$) ve MAE ($0,2926-0,3239$) değerleriyle daha iyi performans sergilemiştir. Ayrıca, duygu temelli model, tüm katlarda orijinal puan modelinden sürekli olarak daha iyi sonuçlar elde etmiştir. Ortalama NDCG değeri de belirgin şekilde daha yüksek olmakla birlikte, standart sapmanın biraz daha fazla olması modelin tahmin doğruluğundaki artışa rağmen küçük bir değişkenlik gösterdiğini ortaya koymaktadır.

Table 13

RMSE: Sentiment-based vs. Original ratings.

Model	Sentiment rating	Original rating
NMF	0.3975	1.1931
DeepMF	0.1126	0.2633
SlopeOne	0.3961	1.1895

Table 14

MAE: Sentiment-based vs. Original ratings.

Model	Sentiment rating	Original rating
NMF	0.3115	0.9239
DeepMF	0.075	0.21
SlopeOne	0.3066	0.91

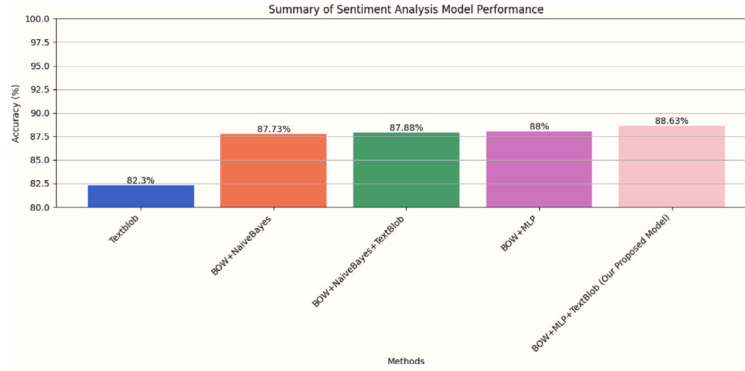


Fig. 10. Performance comparison of sentiment analysis models.

Table 15

NDCG: Sentiment-based vs. Original ratings.

Model	Sentiment rating	Original rating
NMF	0.8037	0.99
DeepMF	0.8099	0.99
SlopeOne	0.8037	0.99

Table 16

Parameters of the DeepMF model.

Parameter	Value
Number of folds	8
Number of epochs	20
Batch size	50
Learning rate	0.002

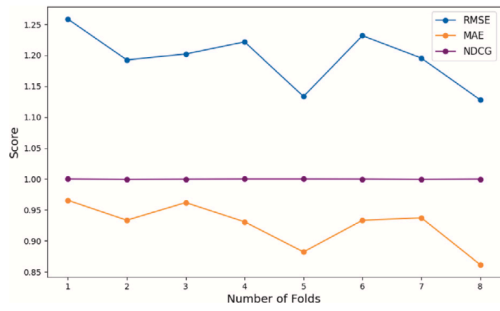


Fig. 11. Performance comparison of NMF with original ratings: RMSE, MAE, NDCG.

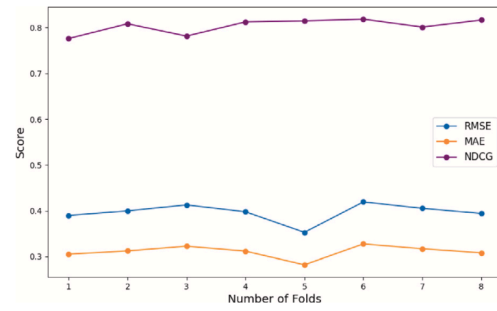


Fig. 12. Performance comparison of NMF with sentiment ratings: RMSE, MAE, NDCG.

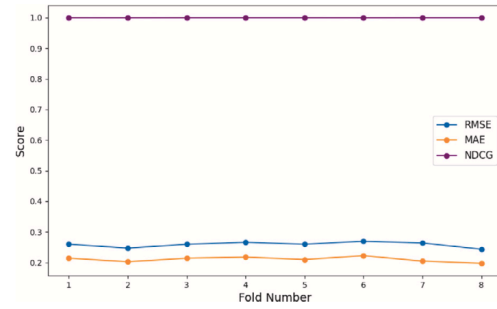


Fig. 13. Performance comparison of DeepMF with original ratings: RMSE, MAE, NDCG.

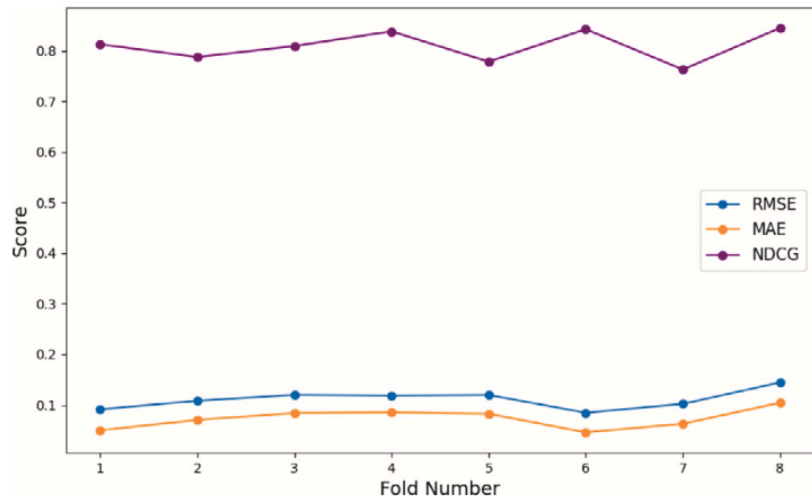


Fig. 14. Performance comparison of DeepMF with sentiment ratings: RMSE, MAE, NDCG.

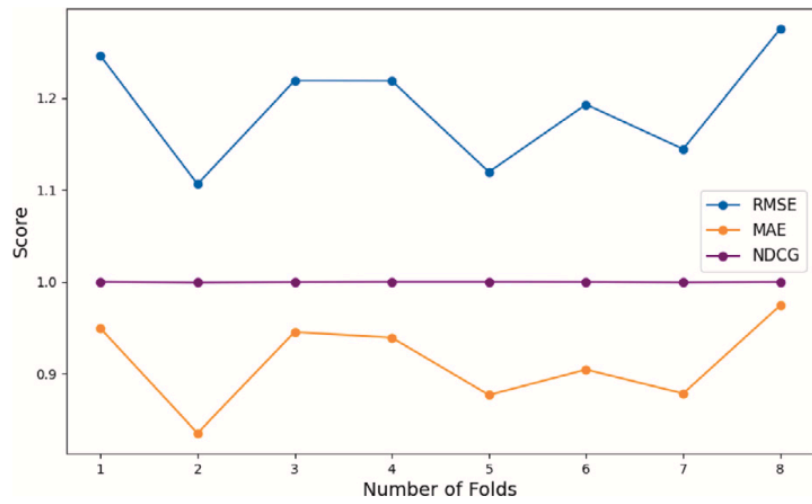


Fig. 15. Performance comparison of SlopeOne with original ratings: RMSE, MAE, NDCG.

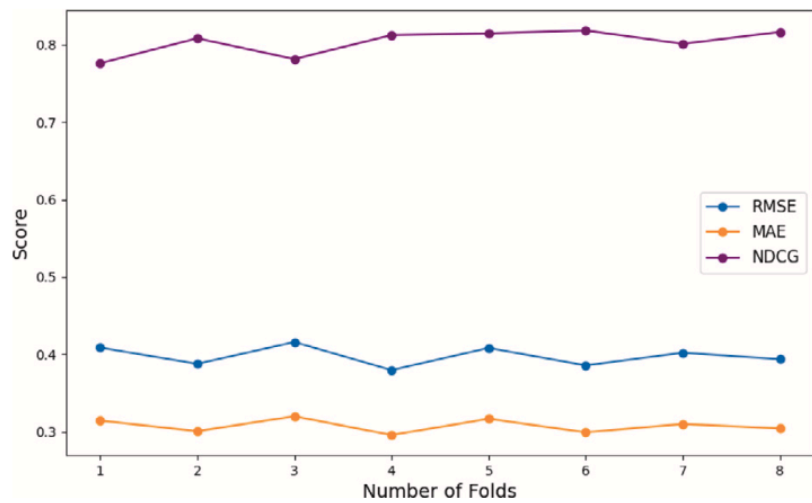


Fig. 16. Performance comparison of SlopeOne with sentiment ratings: RMSE, MAE, NDCG.

4.5. Impact of hyperparameters on sentiment analysis and recommendation system performance

Bu bölümde, modellerimizin hiperparametrelere olan duyarlılığı, duygu analizi ve öneri sistemimizin hiperparametre ayarlaması yapılmış ve yapılmamış hâlleri arasındaki performans karşılaştırmasıyla incelenmektedir. Bu parametrelerin etkisini anlamak, onların sonuçları ve modellerin genel etkinliğini nasıl etkilediğine dair önemli içgörüler sunmaktadır.

Duygu analizi, **Kelime Torbası (Bag of Words – BOW)** temsiline dayalı bir **Çok Katmanlı Algılayıcı (Multi-Layer Perceptron – MLP)** modeli ve **TextBlob** kullanılarak gerçekleştirilmiştir. Duygu analizi modelinin performansı, hiperparametre ayarlaması yapılmış ve yapılmamış durumda ayrı ayrı değerlendirilmiştir.

Tablo 17 sonuçları özetlemektedir. Tablo, hiperparametre ayarlamasının duygu analizi modelinin performansını önemli ölçüde artırdığını göstermektedir. Özellikle, doğruluk oranı %2,6'dan fazla iyileşmiş, ayrıca **MCC** ve **ROC-AUC** gibi diğer metriklerde de belirgin artışlar gözlemlenmiştir. Bu durum, MLP için seçilen hiperparametrelerin (örneğin gizli katman birim sayısı, dropout oranları ve öğrenme oranı gibi) modelin duyguları doğru sınıflandırma yeteneğini optimize etmede kritik bir rol oynadığını göstermektedir.

Öneri sistemi için **Derin Matris Ayrıştırma (Deep Matrix Factorization – DeepMF)** modeli uygulanmış ve benzer şekilde hiperparametre ayarlaması yapılmış ve yapılmamış versiyonları karşılaştırılmıştır.

Tablo 18, elde edilen sonuçları ayrıntılı olarak göstermektedir. Tabloya göre, hiperparametrelerle optimize edilen model, hiperparametre ayarlaması yapılmamış olana kıyasla belirgin şekilde daha iyi performans sergilemiştir. Hiperparametre ayarlı modelin **RMSE** ve **MAE** değerleri daha düşük çıkmış, bu da puanları daha doğru tahmin etme yeteneğinin arttığını göstermektedir. Ancak, varsayılan ayarlara sahip modelin **NDCG** skoru daha yüksek bulunmuştur. Bu da, hiperparametre ayarlamasının tahmin doğruluğunu artırırken sıralama kalitesini etkileyebileceğini göstermektedir. Bu fark, hiperparametre ayarlamalarında dengeli bir yaklaşımın gerekliliğini ortaya koymaktadır; zira farklı performans metrikleri üzerinde farklı etkiler yaratmaktadır.

Karşılaştırmalı analiz, her iki modelin de hiperparametre yapılandırmalarına karşı duyarlı olduğunu ortaya koymaktadır. Duygu analizinde, yüksek doğruluk ve güvenilir performans metriklerine ulaşmak için hiperparametre ayarlaması büyük önem taşımaktadır. Buna karşılık, öneri sisteminde hiperparametre ayarlaması yapılmadığında NDCG skorunda hafif bir artış gözlemlense de, **RMSE** ve **MAE** tarafından gösterilen genel tahmin kalitesi belirgin şekilde düşmektedir.

Table 17

Sentiment analysis performance with and without hyperparameters.

Model	Accuracy	MCC	ROC-AUC
With hyperparameters (BOW + MLP + TextBlob)	88.63%	0.6637	81.45%
Without hyperparameters (Default settings)	86.00%	0.5900	73.00%

Table 18

Recommendation system performance with and without hyperparameters.

Model	RMSE	MAE	NDCG
With hyperparameters (DeepMF)	0.1126	0.075	0.80
Without hyperparameters (Default settings)	0.3700	0.330	0.88

4.6. Comparative evaluation: Our proposed model versus prior research on the Datafiniti Hotel Dataset

Önerilen modelimizin performansının önceki araştırma modelleriyle karşılaştırmalı analizi (Şekil 17), otel verisi analizinin gelişen yapısını ortaya koymaktadır. **Suryawanshi (2022)**, konu modelleme, metin sınıflandırma ve duygu analizi gibi çeşitli doğal dil işleme (NLP) tekniklerini kullanarak otel yorumları üzerindeki temaların etkisini kapsamlı bir şekilde incelemiştir. Bu çalışma, önemli içgörüler sunmuş ve veri kümesi üzerinde **%64 doğruluk oranı** elde etmiştir.

Bu temeller üzerine inşa eden **Arroni ve arkadaşları (2023)**, sosyal ağlarda kamuoyunu anlamada duygu analizinin artan önemini vurgulamışlardır. Otel ile ilgili tweet'ler üzerinde yapılan çalışmada, daha basit duygu analizi yaklaşımlarını geride bırakan **özel bir dikkat (attention) mekanizmasına dayalı derin öğrenme modeli** önermişlerdir ve **%71 doğruluk oranı** elde etmişlerdir.

Alotaibi'nin modeli (Alotaibi ve ark., 2021) ise, kelime gömme (word embedding) yaklaşımlarını **XGBoost sınıflandırıcısı** ile entegre ederek **%84 doğruluk oranı**yla daha iyi bir performans göstermiş, ancak yine de önerdiğimiz modeli geçememiştir.

Genel olarak, önerilen modelimizin üstün performansı, duygu analizi görevlerinde sağlamlığını ve etkinliğini ortaya koymaktadır. Önceki modeller değerli katkılar sağlamış olsa da, çalışmamız onların temelleri üzerine inşa edilmiş ve **model performansında kayda değer bir artış sağlayan yeni teknikler** sunmuştur. **Datafiniti Hotel veri kümesi** üzerinde **%88,63 doğruluk oranı** elde edilmiştir. Yaklaşımımız, daha gelişmiş metodolojiler ve optimizasyonlardan yararlanarak, alanda doğruluk açısından yeni bir ölçüt (benchmark) ortaya koymuştur. Bu sonuçlar, daha yüksek doğruluk elde etmek için **sürekli yenilik ve güncel tekniklerin benimsenmesinin önemini** vurgulamaktadır.

Önerilen **öneri modelimizin**, **Deac-Petruşel ve Limboi (2020)** ile **Yadav (2022)** modelleriyle karşılaştırmalı analizi (Şekil 18), **Kök Ortalama Kare Hatası (Root Mean Squared Error – RMSE)** açısından performansta önemli bir iyileşme olduğunu

göstermektedir. Önerilen modelimiz, **0.1126 gibi son derece düşük bir RMSE** değeri elde ederek, önerilerinde yüksek düzeyde doğruluk sağlamıştır. Buna karşılık, **Deac-Petruşel modeli** belirgin şekilde daha yüksek bir **0.37 RMSE** değerine sahipken, **Yadav modeli 1.01 RMSE** değeriyle en zayıf performansı göstermiştir.

Bu belirgin RMSE farkı, önerdiğimiz yaklaşımın tahmin hatalarını en aza indirme ve daha doğru öneriler sunmadaki etkinliğini açıkça ortaya koymaktadır. Bu sonuçlar, modelimizin **önceki yaklaşımların sınırlamalarını aşan gelişmiş metodolojiler ve optimizasyonlar** içerdiğini ve **öneri sistemlerinde doğruluk açısından yeni bir standart oluşturduğunu** göstermektedir. Modelimizin düşük RMSE değeri, **daha güvenilir ve hassas öneriler** sunma potansiyelini kanıtlayarak, **pratik uygulamalar için yüksek uygunluk** taşıdığını ortaya koymaktadır.

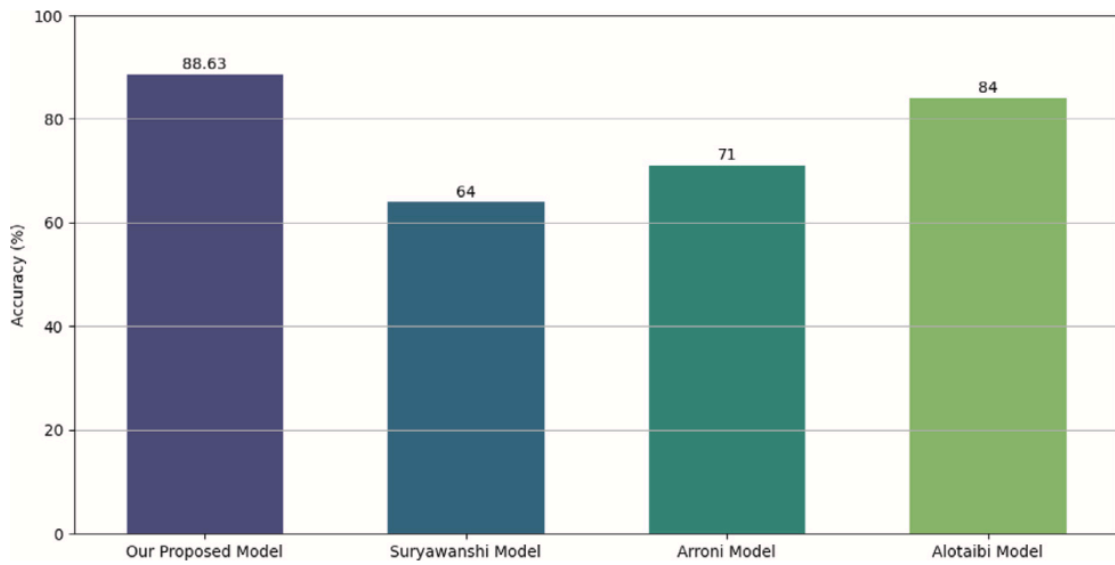


Fig. 17. Analyzing sentiment analysis: Our model Vs. Existing research.

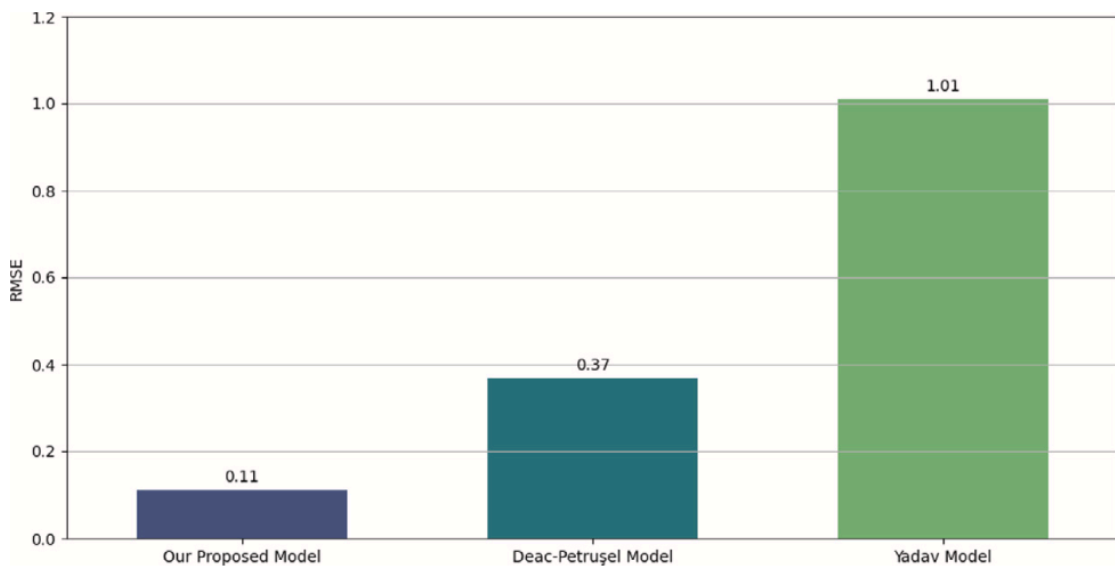


Fig. 18. Analyzing recommendation system: Our model Vs. Existing research.

5. Conclusion and future work

Sonuç olarak, çalışmamız; sözlük tabanlı teknikleri makine öğrenimi ve derin öğrenme yöntemleriyle birleştiren hibrit bir duygu analizi yaklaşımını entegre ederek öneri sistemlerini başarıyla geliştirmiştir. Datafiniti Hotel veri seti kullanılarak gerçekleştirilen kapsamlı deneyler sonucunda, en uygun kombinasyonun **TextBlob** ile **Bag of Words** ve **Multilayer Perceptron (MLP)** algoritmasının **10 katlı çapraz doğrulama** yöntemiyle kullanılması olduğu belirlenmiştir. Bu yaklaşım, %88,63'lük etkileyici bir doğruluk oranına ulaşmış ve müşteri yorumlarından duygu tespitinde yüksek bir doğruluk elde edildiğini göstermiştir. Bu gelişmiş duygu analizinin öneri sistemlerine entegre edilmesi, sistemlerin müşteri duygularını daha iyi anlamasını ve bu içgörülerini önerilere dahil etmesini sağlayarak daha **kişiselleştirilmiş ve anlamlı öneriler** sunmasına imkân tanımaktadır.

Ayrıca, bu hibrit duygu analizinin **DeepMF**, **NMF** ve **SlopeOne** gibi işbirlikçi filtreleme teknikleriyle entegrasyonunu da araştırdık. Bulgularımız, belirli parametrelerle optimize edilmiş ve **8 katlı çapraz doğrulama** ile test edilmiş **DeepMF** tekniğiyle hibrit duygu analizinin kombinasyonunun en etkili yöntem olduğunu göstermektedir. Bu yöntem, **0.1 RMSE (Root Mean Square Error)** değeri elde ederek önerilerin doğruluğu ve kalitesinde önemli bir iyileşme sağlamıştır.

Yeni kullanıcı (cold start) sorununu çözmek amacıyla, işbirlikçi filtreleme ile duygu analizini birleştiren hibrit bir yaklaşım uyguladık. Kullanıcıların ilk etkileşimlerinden ve yorumlardan çıkarılan duygu verilerinden yararlanarak, etkileşim geçmişi çok az olan yeni kullanıcılar için dahi anlamlı öneriler üretebildik. Böylece, kullanıcıların sisteme ilk girişlerinden itibaren **akıcı ve tatmin edici bir deneyim** yaşamaları sağlanmıştır.

Veri seyrekliği (data sparsity) sorununa yönelik olarak, matris faktörizasyon tekniklerini uyguladık ve duygu analizi içgörülerıyla güçlendirilmiş hibrit modeller kullandık. Bu yöntemler sayesinde, seyrek veri kümelerindeki boşluklar etkili biçimde doldurulmuş, sistemin **dayanıklılığı ve güvenilirliği** artırılmıştır.

Gelişmiş duygu analizinin ileri seviye işbirlikçi filtreleme teknikleriyle bütünleştirilmesi, yalnızca öneri sistemlerinin performansını artırmakla kalmamış, aynı zamanda **kullanıcı memnuniyetini** de önemli ölçüde yükseltmiştir. Çalışmamızdan elde edilen umut verici sonuçlar, bu hibrit yaklaşımın öneri sistemlerinde devrim yaratma potansiyeline sahip olduğunu, sistemleri kullanıcı duyarlılıklarına ve tercihlerine daha duyarlı hâle getirebileceğini göstermektedir.

Gelecekteki çalışmalarımızda, çok sınıflı duygu analizini gerçekleştirmek için **transformer tabanlı modellerin entegrasyonunu**, farklı alanlarda veri setlerinin genişletilmesini, **gerçek zamanlı öneri sistemlerinin geliştirilmesini**, **bağlamsal bilgilerin dahil edilmesini** ve sistemin daha da iyileştirilmesi için **kullanıcı çalışmaları yürütülmesini** hedefliyoruz.