

**METİN ANALİZİ VE KONUM TABANLI SOSYAL
AĞLARDA KONUM ÖNERİSİ**

**TEXT ANALYSIS AND LOCATION RECOMMENDATION
ON LOCATION BASED SOCIAL NETWORKS (LBSNS)**

FARHAD KHALILZADEH

PROF. DR. İLYAS ÇİÇEKLI

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

Bilgisayar Mühendisliği Anabilim Dalı için Öngördüğü

DOKTORA TEZİ olarak hazırlanmıştır.

ÖZET

METİN ANALİZİ VE KONUM TABANLI SOSYAL AĞLARDA KONUM ÖNERİSİ

Farhad KHALILZADEH

Doktora, Bilgisayar Mühendisliği Bölümü

Tez Danışmanı: Prof. Dr. İlyas ÇİÇEKLİ

Aralık 2024, 113 sayfa

Amaç: Bu tezin amacı, Konu-Bazlı Duygu Analizi (ABSA) ile entegre edilmiş Heterojen Bilgi Ağları (Heterogeneous Information Networks - HINs) kullanılarak öneri sistemlerini geliştirmektir. Kullanıcı-mekan ilişkilerinin inceliklerini yakalamada ve yorum ile zenginleştirilmiş meta-yol-tabanlı kelime gömmeleri aracılığıyla öneri doğruluğunu artırmada yaşanan zorlukları ele almayı hedeflemektedir. Yorumlar ve etkileşimlerden sözel ve yapısal iç görüler elde etmek için REHREC modeli (Yorumdan Etkilenen Heterojen Bilgi Ağı Tabanlı Öneri Sistemi) ve W-REHREC modeli (Ağırlıklı Yorumdan Etkilenen Heterojen Bilgi Ağı Tabanlı Öneri Sistemi) adlı iki yenilikçi model önerilmektedir. Çalışma, bu modellerin Yelp ve Foursquare gibi yorum açısından zengin ortamlar da dahil olmak üzere farklı veri setleri üzerindeki uyarlanabilirliğini de değerlendirmektedir. Bu araştırma, metinsel ve yapısal veri arasındaki boşluğu kapatmaya, Cold-Start sorunlarına, seyrek etkileşimlere ve ölçeklenebilir ve genellenebilir öneri yöntemlerine olan ihtiyacı ele alarak katkıda bulunmaktadır.

Yöntem: Önerilen REHREC modeli ve W-REHREC modeli iş akışları, kullanıcılar, mekanlar ve yorumlar arasındaki doğrudan ve dolaylı ilişkileri yakalamak için meta-yol-

tabanlı HIN kelime gömmeleri kullanmaktadır. W-REHREC modeli, ABSA duygu puanlarından türetilen kenar (edge) ağırlıklarını içerir ve etkili etkileşimleri önceliklendirir. Her biri farklı veri özelliklerini temsil eden Yelp ve Foursquare veri setleri üzerinde yapılan deneyler ile geliştirilen modeller test edilmiştir. Öneri doğruluğunu değerlendirmek için Root Mean Square Error (RMSE) ve Mean Absolute Error (MAE) gibi değerlendirme metrikleri kullanılmıştır. UBU, BUB, URRU, BRRB, UBRRBU, UBCaBU ve BCaB gibi meta-yollar kullanıcılar (U), işletmeler (B), Yorumlar (R) ve kategoriler (Ca) arasındaki karmaşık ilişkileri modellemek için kullanılmıştır. Modeller, veri setleri arasında en uygun performans için hiper-parametreler ayarlandıktan sonra, Random Walk algoritması ve Skip-Gram algoritması kelime gömmeleri (word embeddings) ile eğitilmiştir.

Çıktılar: Önerilen iş akışları, tüm veri setlerinde öneri doğruluğunda önemli iyileştirmeler gösterdi. Yorumlarla zenginleştirilmiş Yelp veri setinde, W-REHREC modeli modeli, temel HERec modeline kıyasla RMSE'yi %15.65 oranında azalttı ve MAE için de benzer iyileşmeler gözlemlendi. Foursquare veri setinde, W-REHREC modeli modeli, RMSE'yi %14.11 oranında iyileştirdi ve genelleme yeteneğini sergiledi. Sözel ve yapısal ilişkileri yakalayan meta-yollar bu iyileşmelerde önemli rol oynadı ve W-REHREC modeli, duygu-tabanlı bağlantıları vurgulayarak REHREC modeli modelini sürekli olarak geride bıraktı.

Sonuç: Bu tez, REHREC modeli ve W-REHREC modeli iş akışlarını geliştirilmesi yoluyla öneri sistemlerine güçlü bir katkı sunmaktadır. HIN kelime gömmeleri ile ABSA entegrasyonu sayesinde, bu modeller yorumlardan elde edilen verileri etkin bir şekilde kullanarak kullanıcı-mekan ilişkilerini yakalamaktadır. Duygu-ağırlıklı kenar kullanımı ile W-REHREC modeli, doğruluğu ve ölçeklenebilirliği önemli ölçüde artırır ve farklı veri setleri üzerinde uyarlanabilirliğini gösterir. Sonuçlar, modern öneri sistemleri için sözel ve yapısal iç görülerin birleştirilmesinin önemini vurgular. Bu çalışma, çok dilli duygu analizine, dinamik zamansal modellemeye ve gerçek zamanlı öneri uygulamalarına yönelik gelecekteki araştırmalar için bir temel sağlar ve kişiselleştirilmiş, bağlam odaklı ve ölçeklenebilir öneri sistemlerindeki gelişen zorluklara yanıt verir.

Anahtar Kelimeler: Öneri Sistemleri, Heterojen Bilgi Ağları, Konum Tabanlı Sosyal Ağlar, Konu-Bazlı Duygu Analizi, Meta-yol Kelime gömmeleri, İşbirlikçi Filtreleme

ABSTRACT

TEXT ANALYSIS AND LOCATION RECOMMENDATION ON LOCATION BASED SOCIAL NETWORKS (LBSNS)

Farhad KHALILZADEH

Doctor of Philosophy, Department of Computer Engineering

Supervisor: Prof. Dr. İlyas ÇİÇEKLI

December 2024, 113 pages

Purpose: The purpose of this thesis is to enhance recommendation systems by leveraging Heterogeneous Information Networks (HINs) integrated with Aspect-Based Sentiment Analysis (ABSA). It aims to address challenges in capturing nuanced user-business relationships and improving recommendation accuracy through review-enriched meta-path-driven embeddings. Two innovative frameworks, REHREC and W-REHREC (Weighted REHREC), are proposed to incorporate semantic and structural insights from reviews and interactions. The study also evaluates the adaptability of these models across diverse datasets, including review-rich environments (Yelp, Foursquare). This research contributes to bridging the gap between textual and structural data while addressing cold-start issues, sparse interactions, and the need for scalable and generalizable recommendation methods.

Methods: The proposed REHREC and W-REHREC frameworks utilize meta-path-driven HIN embeddings to capture direct and indirect relationships among users, businesses, and reviews. W-REHREC incorporates edge weights derived from ABSA sentiment scores, prioritizing impactful interactions. Experiments were conducted on two datasets: Yelp and

Foursquare, each representing varying data characteristics. Evaluation metrics, including Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), assessed recommendation accuracy. Meta-paths such as UBU, BUB, URRU, BRRB, UBRRBU, UBCaBU, and BCaB were utilized to model complex relationships between users (U), businesses (B), reviews (R) and categories (Ca). The models were trained with random walks and skip-gram embeddings, with hyperparameters tuned for optimal performance across datasets.

Results: The proposed frameworks demonstrated significant improvements in recommendation accuracy across all datasets. On the review-rich Yelp dataset, W-REHREC achieved an RMSE reduction of 15.65% compared to the baseline HERec model, with similar improvements observed for MAE. On the Foursquare dataset, W-REHREC improved RMSE by 14.11%, showcasing its generalization capability. Meta-paths capturing semantic and structural relationships were instrumental in these improvements, with W-REHREC consistently outperforming REHREC by emphasizing sentiment-driven connections.

Conclusion: This thesis presents a robust contribution to recommendation systems through the development of REHREC and W-REHREC frameworks. By integrating ABSA with HIN embeddings, these models effectively leverage review data to capture nuanced user-business relationships. W-REHREC, with its sentiment-weighted edges, significantly enhances accuracy and scalability, demonstrating adaptability across diverse datasets. The results highlight the importance of combining semantic and structural insights for modern recommendation systems. This work provides a foundation for future research in multi-language sentiment analysis, dynamic temporal modeling, and real-time recommendation deployment, addressing evolving challenges in personalized, context-aware, and scalable recommendation technologies.

Keywords: Recommendation Systems, Heterogeneous Information Networks, Location Based Social Networks, Aspect-Based Sentiment Analysis, Meta-Path Embeddings, Collaborative Filtering.

TEŞEKKÜR

Öncelikle, doktora yolculuğum ve ilgili araştırmalarım boyunca gösterdiği sarsılmaz destek için danışmanım Prof. Dr. İlyas Çiçekli'ye içten teşekkürlerimi sunmak isterim. Sabır, teşvik ve engin bilgisi paha biçilmezdi. Çalışmamın başarısını sağlamak için her zaman rehberlik, düşünceli geri bildirim ve motivasyon sağlamak için yanımdaydı. Onun samimi ve destekleyici tutumu, ilerlememde hem ilham verici hem de belirleyici olmuştur ve buna minnettarlığım sonsuzdur.

Tez jüri üyelerim, Prof. Dr. Pınar Karagöz ve Dr. Öğr. Üyesi Burcu Can'a da özel teşekkürlerimi sunmak isterim. Kapsamlı yorumları, teşvikleri ve yapıcı eleştirileri için minnettarım. Zorlayıcı soruları ve önerileri, araştırmamı daha geniş açılardan keşfetmeme teşvik etti ve bu çalışmanın derinliğini önemli ölçüde zenginleştirdi.

Son olarak, bu yolculuk boyunca gösterdikleri sarsılmaz sevgi, destek ve motivasyon için aileme, özellikle sevgili eşim Maryam ve oğlum Alp'e derin minnettarlığımı sunarım. Onların pozitifliği ve teşvikleri, bana güç veren ve bu tezi tamamlamamı sağlayan sürekli bir destek kaynağı oldu.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET.....	i
ABSTRACT	iii
TEŞEKKÜR	v
İÇİNDEKİLER.....	vi
ŞEKİLLER DİZİNİ.....	x
ÇİZELGELER DİZİ.....	xi
KISALTMALAR	xii
1. GİRİŞ	1
1.1. Tezin Amacı	1
1.2. Problem Tanımı.....	2
1.3. Amaçlar ve Hedefler	3
1.3.1. Amaçlar	4
1.4. Katkıları.....	5
1.4.1. REHREC modeli Model Geliştirme.....	5
1.4.2. W-REHREC modeli: Duygu-Ağırlıklı Yorum Entegrasyonu.....	5
1.4.3. Duygu ile Güçlendirilmiş Meta-yol Kelime Gömmesi Yaklaşımı.....	6
1.4.4. Ağırlıklı PageRank ile Geliştirilmiş Öneri Sıralaması	6
1.4.5. Deneysel Doğrulama: Diğer Modellere Karşı Kazanımlar	7
1.4.6. LBSN'ler için Ölçeklenebilir Model Oluşturma.....	7
1.5. Tezin Ana Hatları	7
2. LİTERATÜR TARAMASI.....	10
2.1. Genel Bakış	10
2.2. Öneri Sistemleri.....	11
2.2.1. Öneri Sistemlerine Yaklaşımlar	12
2.3. Heterojen Bilgi Ağı (HIN)	15
2.3.1. HIN'lerin Yapısı ve Bileşenleri	15
2.3.2. Öneri Sistemlerinde HIN'lerin Uygulamaları.....	16

2.3.3. Makine Öğrenme Modelleriyle Entegrasyon	17
2.3.4. HIN'ler için Kelime Gömmesi ve Boyut Azaltma.....	18
2.4. Konu-Bazlı Duygu Analizi (ABSA).....	21
2.4.1. ABSA'ya Genel Bakış	21
2.4.2. ABSA'daki Teknikler	22
2.4.3. ABSA'daki Zorluklar ve Tekniklerindeki Gelişmeler.....	22
2.4.4. ABSA'nın Öneri Sistemlerinde Uygulamaları.....	24
2.5. Graf-bazlı Sıralama Algoritmaları: PageRank	24
2.5.1. PageRank Algoritmasına Genel Bakış	24
2.5.2. Çeşitli Alanlarda PageRank Uygulamaları.....	25
2.5.3. PageRank'in Varyantları ve Geliştirmeleri.....	26
2.5.4. PageRank Kullanımında Karşılaşılan Zorluklar ve Dikkat Edilmesi Gerekenler	27
2.5.5. Modern Sistemlere PageRank'in Entegrasyonu.....	28
3. YÖNTEM.....	30
3.1. Giriş	30
3.2. HIN Tabanlı Sistemlerde Kullanıcı Yorumlarının Önemi	30
3.3. Meta-yollar ve Önerileri Geliştirmedeki Rollerini	31
3.4. Cold-Start ve Veri Seyrekliği Sorunlarının Ele Alınması	31
3.5. Duygu Analizi ve HIN Kelime Gömmelerinin Birleştirilmesi.....	32
3.6. Gelişmiş Öneriler için Birleşik Bir Model	32
3.7. Yorum Grafi Oluşturma.....	33
3.7.1. REHREC'te Yorum Grafi Oluşturma	33
3.7.2. W-REHREC'te Yorum Grafi Oluşturma	35
3.7.3. REHREC ve W-REHREC'te Yorum Graflarının Birleşik Faydaları	37
3.8. HIN Ağ Üretimi.....	38
3.8.1. REHREC'te HIN Ağı Oluşturma.....	38
3.8.2. W-REHREC'de HIN Ağı Üretimi	39
3.8.3. REHREC ve W-REHREC'de HIN Ağ Üretimini Birleşik Faydaları	41
3.9. HIN Kelime gömmesi.....	42
3.9.1 REHREC'e HIN Kelime gömmesi	42
3.9.2. Ağırlıklı Yorumlarla Meta-Yol Odaklı HIN Kelime Gömmesi	43
3.9.3. HIN Kelime Gömmesi Yaklaşımlarının Birleşik Faydaları	45

3.10. Yorum Ağırlıklandırması için Konu-Bazlı Duygu Analizi (ABSA)	45
3.10.1. ABSA Metodolojisine Genel Bakış	46
3.10.2. ABSA'nın Ağırlıklı Yorum Graflarına Entegrasyonu	46
3.10.3. Yorum Ağırlıklandırması için ABSA Kullanmanın Faydaları	48
3.11. Değerlendirme notu Tahmin Fonksiyonu ve Öneri Çıktısı	49
4. DENEYLER VE DEĞERLENDİRME	51
4.1. Genel Bakış	51
4.2. Veri kümeleri.....	51
4.2.1. Yelp Veri Kümesi.....	52
4.2.2. Foursquare Veri Kümesi	59
4.3. Meta Yollar	61
4.3.1. Random Walk Serileri İçin Filtreleme Aşaması.....	63
4.4. Değerlendirme Ölçütleri.....	64
4.4.1. Kök Ortalama Kare Hatası (RMSE).....	64
4.4.2. Ortalama Mutlak Hata (MAE)	65
4.4.3. RMSE ve MAE'nin Karşılaştırılması	66
4.5. Deneysel Kurulum.....	66
4.5.1. Veri Kümesi Bölme ve Eğitim Oranları.....	66
4.5.2 Parametre Yapılandırması	66
4.5.3. Yapılandırmanın Amacı	68
4.6. Deneylerde Tanıtılan Yöntemler	68
4.6.1. HERec: Heterojen Bilgi Ağı Kelime Gömmesi Önerisi	68
4.6.2. NCF: Sinirsel İşbirlikçi Filtreleme	69
4.6.3. MCRc: Meta-yol Tabanlı Evrişimsel Öneri	69
4.6.4. NGCF: Sinirsel Graf İşbirlikçi Filtreleme.....	69
4.6.5. Bu Yöntemlerin Çalışmadaki Rolü	70
4.7. Veri Kümesindeki Etkinlik Deneylerinin Sonuçları	70
4.8. Veri Kümeleri Arasında Yorum Dahil Meta-Yollarla RMSE Sonuçları	74
4.9. Sonuçlar ve Gözlemler: Yelp Veri Kümesi.....	78
4.10. Sonuçlar ve Gözlemler: Foursquare Veri Kümesi	81
4.11. Sonuçlar ve Gözlemler: Meta-Yollar	83
4.11.1. "URRU" Meta-Yolunun Veri Kümeleri Üzerindeki Etkisi.....	84
4.11.2. "BRRB" Meta-Yolunun Veri Kümeleri Üzerindeki Etkisi	86

4.11.3. "UBRRBU" Meta-Yolunun Veri Kümeleri Üzerindeki Etkisi.....	89
4.11.4. "BURRUB" Meta-Yolunun Veri Kümeleri Üzerindeki Etkisi.....	91
4.11.5. Veri Kümeleri Arasında Birleştirilmiş Meta-Yolların Etkisi	94
5. TARTIŞMA, SONUÇ VE GELECEĞE YÖNELİK DÜŞÜNCELER.....	97
5.1. Tartışma.....	97
5.2. Sonuç ve Geleceğe Yönelik Düşünceler	99
KAYNAKLAR.....	102
EKLER	113
EK 1 - Tezden Türetilmiş Yayınlar	113
ÖZGEÇMİŞ.....	114

ŞEKİLLER DİZİNİ

	<u>Sayfa</u>
Şekil 3.1 Yaklaşımın iş akışını gösteren diyagram.	50
Şekil 4.1 YELP'te Yorumların dağılımı.	55
Şekil 4.2 YELP'te Yorumların örnek kelime bulutu.	56
Şekil 4.3 Yelp Veri Setinin Örnek HIN Ağı.	63
Şekil 4.4 Örnek Mata Yolu tabanlı Filtreleme İşlemleri.	65
Şekil 4.5 YELP'de RMSE.	72
Şekil 4.6 YELP'te MAE.	73
Şekil 4.7 FOURSQUARE'de RMSE.	74
Şekil 4.8 FOURSQUARE'de MAE.	75
Şekil 4.9 YELP'de Matapath'ın RMSE Üzerindeki Etkisi.	77
Şekil 4.10 FOURSQUARE'de RMSE Üzerindeki Matapath Etkisi.	78

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 4.1 Veri kümesi üzerindeki etkililik deneylerinin sonuçları.	71
Çizelge 4.2 %90 eğitimde farklı meta-yolları kullanılarak elde edilen RMSE sonuçları.	76

KISALTMALAR

REHREC	Review Effected Heterogeneous Information Network Recommendation
W-REHREC	Weighted Review Effected Heterogeneous Information Network Recommendation
HIN	Heterogeneous Information Networks
ABSA	Aspect-Based Sentiment Analysis
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
LBSN	Location Based Social Networks

1. GİRİŞ

1.1. Tezin Amacı

Son yıllarda, Konum Tabanlı Sosyal Ağlar'ın (Location Based Social Networks - LBSN) alanındaki hızlı gelişmeler, insanların mekansal bilgiyi kullanma ve keşfetme biçimini devrim niteliğinde değiştirmiştir ve bu da öneri sistemi iyileştirmeleri için bir çalışma alanı sunmaktadır. Konum tabanlı sosyal ağlar, kullanıcıların, konum bilgisiyle birleştirilmiş metin biçimindeki girişler (check-ins) ve yorumlar gibi konum tabanlı etkinlikleri kaydetmelerine olanak tanır. Bu veri karışımının teknik çeşitliliği, kullanıcı deneyimini iyileştiren, bağlam odaklı (context-aware) ve yüksek derecede kişiselleştirilmiş öneriler açısından öneri alanında yeni fırsatlar ve zorluklar sunar.

Geleneksel öneri yöntemleri, kullanıcı-öğe etkileşim verilerini kullanarak İşbirlikçi Filtreleme (Collaborative Filtering) veya İçerik tabanlı (content-based) yaklaşımlar gibi oldukça başarılı sonuçlar elde etmiştir; ancak, LBSN'e özgü olanlar sınırlıdır. Bu yöntemler, veri eksikliği, Cold-Start sorunları ve daha önemlisi, LBSN'deki zengin mekansal-zamansal ve metinsel bilgiden tam anlamıyla yararlanamama gibi sorunlarla karşılaşmaktadır. Bu zor bir problemdir çünkü bu veri sadece parçalanmış değil, aynı zamanda mekânsal bilgi ve kullanıcı tercihleri de, mekanlar, kullanıcılar ve yorumlar içeriği arasındaki ilişkileri kuracak daha güçlü, daha uyumlu bir model gerektiren karmaşık yapılardır.

Ayrıca, doğal dil işleme (NLP) ve heterojen bilgi ağlarındaki (HIN) büyük gelişmeler, bu aşamada geleneksel öneri yaklaşımlarını sorunlu hale getirirken, bu sınırlamalara yönelik bazı umut verici çözümler sunmaktadır. NLP teknikleri, özellikle Konu-Bazlı Duygu Analizi (Abstract-Based Sentiment Analysis - ABSA) sayesinde, konunun belirli konuları hakkında (örneğin, ambiyans, hizmet kalitesi gibi) daha ince bir duygu analizi yapılabilen ve kullanıcı önerilerini ek kişiselleştirme ile zenginleştirilebilmektedir. Öte yandan, HINs, heterojen verilerin entegrasyonunu kolaylaştırır (kullanıcı yorumları ve sosyal veri setlerini tek bir modelde birleştirme olanağı sağlar). Bu yaklaşım, öneri sistemlerinin kullanıcıların spesifik ihtiyaçlarına daha iyi bir bağlam sunmasını sağlamakla kalmaz, aynı zamanda öneri sistemlerinde çok belirgin olan standart sorunları (örneğin veri eksikliği) da azaltır.

LBSN verisinin bu çok boyutlu doğasından yola çıkarak, bu tez, LBSN'den faydalı bilgi çıkarma amacıyla bir öneri modeli önermektedir. Özellikle, belirli HIN yapısı sayesinde metin ve konum tabanlı bilgilere dayalı gelişmiş meta-yol-tabanlı ağ kelime gömmeleri gerçekleştiren ve kullanıcı tercihlerini duygu farkındalıklı yorum ağırlıklandırma

mekanizmaları ile yakalayan öneri modeli tasarlamayı hedeflemektedir. Geliştirdiğimiz Ağırlıklı Yorumdan Etkilenen Heterojen Bilgi Ağı Tabanlı Öneri Sistemi (W-REHREC modeli) olarak adlandırdığımız bu model, kullanıcı tarafından oluşturulan/öngörülen yorumlara dayalı önerilerin yorumlanabilirlik, ölçeklenebilirlik ve kişiselleştirilme açısından daha iyi doğruluğunu sağlayan mevcut öneri modellerindeki önemli bir boşluğu doldurmayı hedeflemektedir.

Bunu başarmak için, tez, etki analizine dayalı duygu analizi, ağ kelime gömmeleri ve duyguya dayalı izlenim yararlarını tartan bir sıralama prosedürünü içeren LBSN'ler için öneri tekniklerinde bir iyileştirme sunmaktadır. Bu araştırma, bu teknikleri, hem mekansal hem de metinsel verileri içeren dijital toplumda kullanıcıların memnuniyetini ve katılımını optimize eden yeni bir öneri sistemi biçiminde uygulayarak, LBSN'lerde mevcut öneri sistemlerini genişletmeyi amaçlamaktadır.

1.2. Problem Tanımı

Son zamanlarda Konum Tabanlı Sosyal Ağlar (LBSN) çok yaygın kullanılmasıyla, konum bazlı öneri sistemleri daha fazla önem kazanmıştır. Ancak, kullanıcı tarafından üretilen içeriğin büyüklüğü— milyonlarca konum etiketli, her mekan etrafında yer alan binlerce metin tabanlı yorum, puanlamalar ve sosyal etkileşimler—mevcut öneri sistemleri için bir dizi karmaşık sorunu beraberinde getirmektedir. Ancak, geleneksel öneri yaklaşımları— genellikle İşbirlikçi (Collaborative) filtreleme (CF) ve İçerik tabanlı (content-based) filtreleme (CBF)—LBSN'nin ürettiği devasa çok yönlü verileri kullanmasını engelleyen birkaç önemli sınırlamaya sahiptir.

Öncelikle, geleneksel öneri sistemleri genellikle heterojen bilgi ağlarının doğasını göz ardı eder. Kullanıcılar, konumlar, yorumlar, kategoriler ve diğerlerinin iç içe geçtiği LBSN'lerin temel heterojenliğini ve zengin bağlantılarını sıkça göz ardı ederler. Bu sistemler genellikle homojen ağları kullandığından, LBSN'lerde gömülü dinamik etkileşimleri karakterize edemezler—bu da heterojen bilginin ihmal edilmesi nedeniyle düşük öneri doğruluğuna yol açar.

İkinci olarak, mevcut öneri yöntemleri Cold-Start ve veri eksikliği (data sparsity) sorunlarıyla karşı karşıya kalmaktadır. Bu, yeni kullanıcılar için ya da tarihsel etkileşim verisinin eksik olduğu durumlarda doğru tahminler yapmayı zorlaştırır, dolayısıyla kullanıcı deneyimleri ve memnuniyetini ve etkileşimi büyük ölçüde etkiler. Geleneksel öneri

teknikleri, kullanıcıların ziyaret ettikleri/yorum yazdıkları yerlerle olan etkileşimleri nedeniyle LBSN'lerde daha fazla zorluklarla karşılaşır; bu etkileşimler, insanların LBSN'leri daha az düzenli kullanması nedeniyle öğrenilecek davranış setini daha da seyrek hale getirir.

Bir diğer ve daha büyük sorun, yorumların potansiyel duygu bilgisi içeren çok sayıda metin verisi içermesidir, ancak bu bilgi yeterince kullanılmaz. Kullanıcı yorumlarında, konuların belirli konularının (örneğin, ambiyans, hizmet veya fiyat) kalitesiyle ilgili önemli bilgiler sunulmakta ve bu da önerilerin kişiselleştirilmesini artırabilir. Bu ince ayrıntılı duygu analizlerinden büyük bir zenginlik olmasına rağmen, mevcut çoğu model bu bilgileri nadiren kullanır, yorumları yerel (ince ayrıntılı) tercihlerden oluşan bir akış veya diziyi değil de tek bir monolitik öge (item) olarak ele alır.

Ayrıca, Heterojen Bilgi Ağları (HINs) ve Konu-Bazlı Duygu Analizi (ABSA) üzerine yapılan son çalışmalara rağmen, bu sorunlara yönelik teorik olarak iyi çözümler geliştirilmiş olsada, bunları daha iyi LBSN önerileri için birleştiren pratik uygulamalar çok azdır. HIN tabanlı modelleme ve duygu ağırlıklı yorum mekanizmalarını çalışma üzerine sınırlı sayıda model olması, öneri sistemlerinin yüksek derecede kişiselleştirilmiş, bağlam farkındalığına sahip (context-aware) öneriler sunma spektrumunu (tayfını) kısıtlamaktadır.

Bu boşlukları kapatmak için, bu tez, LBSN'lerin heterojen, seyrek ve duygu yüklü verilerinden öğrenebilen bir öneri modeline odaklanmaktadır. Özellikle, bu tezin amacı, iki araştırma alanını: HIN tabanlı ve ABSA teknikleri — entegre eden yepyeni bir öneri modeli geliştirmektir. Böylece model, bunlar arasındaki farklı türdeki varlıkları ve ilişkileri kullanabilir, Cold-Start ve veri eksikliği sorunlarını ele alabilir; daha doğru ve alakalı bir şekilde LBSN önerilerini iyileştirmek için kullanıcıların ince ayrıntılı duygu verilerini kullanır.

1.3. Amaçlar ve Hedefler

Bu tez çalışmasının ana amacı, Konum Tabanlı Sosyal Ağlar (LBSN) içindeki heterojen, duygu zengin ve konum odaklı veriye dayalı olarak son derece doğru, bağlam farkındalığına sahip ve kişiselleştirilmiş öneriler sunmaktır. Mevcut öneri yöntemlerinin karşılaştığı sınırlamaları hafifletmek için, Heterojen Bilgi Ağları (HINs) ve Konu-Bazlı Duygu Analizi (ABSA) yöntemlerini entegre eden, bu yaklaşımları etkili bir şekilde birleştirerek veri eksikliği sorunlarını aşmayı ve kullanıcı yorumlarında mevcut olan ince ayrıntılı bilgiden faydalanmayı hedefleyen bir model geliştiriyoruz.

1.3.1. Amaçlar

Tezin amacı aşağıdaki öğeler olarak özetlenebilir:

- LBSN için çok sayıda veri ögesini (örneğin, kullanıcılar, konular, yorumlar, kategoriler) ve ilişkileri entegre edebilen, bu varlıklar arasındaki karmaşık bağlantıları kodlamak için temel oluşturacak birleşik bir ağ yapısına dayalı HIN tabanlı bir model tasarlamak.
- Sosyal ağ içindeki kullanıcı tarafından üretilen yorumlar üzerinde Konu-Bazlı Duygu Analizi (ABSA) uygulayarak ve konu-bazlı duygular (örneğin, ambiens, yemek kalitesi, hizmet) çıkararak bu duyguları HIN'den türetilmiş grafta gömülü hale getirerek daha kullanıcıya özel ve ilişkilendirilebilir öneriler sağlamak.
- HIN modelini duygu gücüyle zenginleştirmek için, yorumlar üzerine ek bir duygu ağırlıklı Random Walk algoritması ekleyen bir öneri sistemi sunuyoruz. Bu algoritma, yorumları güçlerine göre önyargılı hale getirerek, kullanıcıların beğenip beğenmediği konularda yüksek değerli kaynaklara odaklanmayı sağlayan bir öğrenme algoritmasını etkinleştirir.
- Veri eksikliği olan belirli durumlarda kullanıcı-öge etkileşimlerini nasıl temsil edeceğini öğrenmek için yapılandırılmış HIN'den yararlanan ve duygu farkındalığına sahip meta-yolları, örneğin, kullanıcı-yorum ve mekan-yorum yollarını içeren meta-yol tabanlı bir kelime gömmesi (embedding) süreci öneriyoruz.
- Ağ bağlantısı ve duygu tabanlı girdinin bir kombinasyonuna dayalı olarak önerileri belirlemek için önerilen ağırlıklı PageRank algoritması ile kenar ağırlıklarının kullanılmasını sağlıyor ve benzer kullanıcı profili gruplarındaki yüksek alaka düzeyine ve yüksek kaliteye sahip öğeleri tanımlamayı sağlar.
- Önerilen modelin, ünlü bir veri seti (Yelp) üzerindeki temel öneri sistemleriyle karşılaştırılarak doğruluk, kullanıcı memnuniyeti ve veri eksikliği sorunundaki etkinliğini göstererek iyileştirmelerini kanıtlar.

Bu hedeflere ulaşarak, bu tez LBSN tarafından sunulan benzersiz zorluklara ve fırsatlara hitap eden, ölçeklenebilir, duygu farkındalığına ve bağlam duyarlılığına sahip bir modeli geliştirerek öneri sistemleri alanına katkıda bulunmayı amaçlamaktadır.

1.4. Katkılar

Bu tez, konum tabanlı sosyal ağlarda (LBSN) heterojen bilgi ağlarının (HIN'ler) ve duygu temelli yorum ağırlıklandırmasının gücünden tam anlamıyla yararlanan yeni bir model önererek en son teknolojiyi iletir. Bu çalışmada gerçekleştirilen iki temel çalışmadan, W-REHREC modeli ve REHREC modeli modellerinden elde edilen bilgilerden yararlanarak, bu tez, doğruluk, kişiselleştirme ve yorumlanabilirlik açısından LBSN önerilerini geliştirmek için duygu analizini, Random Walk algoritması süreçlerini ve ağ kelime gömmesi tekniklerini kullanan yaklaşımlar önermektedir. Bu tez aşağıdaki temel katkıları sağlamaktadır.

1.4.1. REHREC modeli Model Geliştirme

Tez, kullanıcı tarafından üretilen yorumların doğrudan HIN'e düğüm olarak dahil edildiği Yorumdan Etkilenen Heterojen Bilgi Ağı Tabanlı Öneri Sistemi (REHREC modeli) adlandırdığımız modelini içeren bir Heterojen Bilgi Ağı (HIN) öneri çözümünü sunmaktadır. Bu model, bağlam farkındalığı sağlayan özelliklerin dahil edilmesiyle, geleneksel kullanıcı-öge etkileşimlerinin ötesinde kullanıcılar, mekanlar ve yorumlar arasında daha derin ilişkiler öğrenebilme yeteneğine sahiptir.

REHREC modeli, HIN'ler içinde kullanıcıları, yorumlar ve mekanları bağlamak için meta-yol-ları benzersiz bir şekilde kullanarak, özellikle düğümler arasındaki yorum ile üretilen ek ilişkilerden etkilenen veri seyrekliği ortamlarında, kamuya açık LBSN veri kümelerinde mevcut en iyi öneri sistemlerini aşan bir model oluşturur.

1.4.2. W-REHREC modeli: Duygu-Ağırlıklı Yorum Entegrasyonu

Ağırlıklı Yorumdan Etkilenen Heterojen Bilgi Ağı Tabanlı Öneri Sistemi (W-REHREC), yorum grafına duygu gücünü ekleyerek REHREC modeli'in bir genişletmesidir; bu modelde, belirli bir konuma yönelik yorum duygusuna dayalı olarak düğüm kenarlarına ek ağırlıklar uygulanır. Bu modeldeki Konu-Bazlı Duygu Analizi (ABSA), duygu gücüne ve mekanın belirli konuları ile (örneğin, ambiyans, yemek kalitesi, hizmet seviyesi vb.) ne kadar ilgili olduklarına bağlı olarak, yorumlara farklı ağırlıklar atar.

W-REHREC modeli, HIN'ler içinde yollar keşfederken yorumların yüksek duygu gücünü dikkate alan duygu-ağırlıklı bir Random Walk algoritması kullanır. Kullanıcılar için en

önemli olan mekanın belirli niteliklerini içeren bu iyileştirme, kullanıcı tercihleriyle uyumlu daha doğru öneriler sağlanmasını mümkün kılar.

1.4.3. Duygu ile Güçlendirilmiş Meta-yol Kelime Gömmesi Yaklaşımı

Bu gözlemlerden yola çıkarak, kullanıcılar, mekanlar, yorumlar arasındaki duygu yüklü bağlantıları özel olarak hedefleyen öneri doğruluğunu artırmak için HIN'lerde meta-yol tabanlı bir kelime gömmesi yaklaşımı öneriyoruz. Meta-yol tabanlı kelime gömmesi alanında bu yöntem, modeli, duygu açısından zengin yorumları seçici bir şekilde vurgulayarak, veri kıtlığı olan ortamlarda karmaşık kullanıcı-öge ilişkilerini öğrenmeye olanak tanır.

Duygu-farkındalıklı meta-yol-lar (örneğin, Kullanıcı-Mekan-Yorum-Yorum-Mekan-Kullanıcı (UBRRBU) ve Mekan-Kullanıcı-Yorum-Yorum-Kullanıcı-Mekan (BURRUB)) kullanılarak model, eş-değerlendirilmiş nesnelere üzerinde kullanıcıya özel pattern madenciliğini pattern tabanlı bir yöntemle dönüştürür. Bu yöntem, geleneksel yöntemlerin gözden kaçırdığı, kullanıcılar arasında bulunan karmaşık kalıpların daha hassas bir şekilde belirlenmesine olanak tanır ve böylece önerilerin kişiselleştirilmesini ve alakalılığını artırır.

1.4.4. Ağırlıklı PageRank ile Geliştirilmiş Öneri Sıralaması

Bu tezde sunulan, duygu gücü ve ağ bağlantısına dayalı olarak mekanları sıralamak için ağırlıklı bir PageRank algoritması uygulanır. Genel duygu puanları, sıralama algoritmasındaki kenarlara dayalı olarak hesaplanır, bu da modelin pozitif yüksek kaliteli, tutarlı ve kaliteli yorumlara sahip mekanları önceliklendirerek önerilerin güvenini ve güvenilirliğini artırmasını sağlar.

Bu duygu-ağırlıklı PageRank tabanlı yaklaşım, kalite odaklı önerilerin üretilmesi gereksinimini karşılar; zira öğeler, kullanıcı yorumlarından yüksek duygu yoğunluğuna sahip olmalarının yanı sıra HIN'de güçlü bağlantılar kurmuşlarsa öneri listesinin üst sıralarına yerleştirilir. Bu iki aşamalı sıralama, kullanıcı tercihlerine daha yakın öneriler oluşturarak daha iyi kalite sonuçlar sağlar.

1.4.5. Deneysel Doğrulama: Diğer Modellere Karşı Kazanımlar

Yelp veri setinde yapılan tüm deneyler, REHREC modeli ve W-REHREC modeli modellerinin, birçok temel yöntemler karşısında tahmin doğruluğunda öne çıktığını gösteriyor. Özellikle yeni W-REHREC modeli modelimiz, RMSE açısından orijinal HEREC'e kıyasla %15,65'ye kadar daha iyi performans sergiliyor. Deneyler, bu tezin katkılarını da doğrulamakta ve duygu-ağırlıklı HIN'lerin öneri metriklerinde sürekli olarak güçlü iyileştirmeler sağladığını göstermektedir.

Bu tez, HIN'ler içinde ABSA'nın dahil edilmesinin, tahmin hatalarını sadece azalttığını değil, aynı zamanda Cold-Start kullanıcıları ve veri eksikliği gözlemlenen mekanlar için öneri kalitesini artırdığını gösteriyor. Bu da Cold-Start ve veri eksikliği senaryolarında önemli bir iyileştirmeyi işaret etmektedir.

1.4.6. LBSN'ler için Ölçeklenebilir Model Oluşturma

Bu tez, ABSA ve HIN tabanlı öneri yöntemlerini birleştirerek Yelp dışında diğer sosyal platformlar ve konum tabanlı uygulamalar için genelleştirilebilen ölçeklenebilir bir model oluşturur. Bu yaygınlık, REHREC modeli modellerinin kullanıcı profilleri ve metinsel duygulara dayalı kişiselleştirilmiş önerilere ihtiyaç duyan daha geniş bir alan yelpazesine uygulanabilir olmasını sağlamak için çok önemlidir.

Genel olarak, bu tez, duygu-ağırlıklı HIN tabanlı modeller aracılığıyla LBSN'ler için öneri sistemlerini iyileştirme konusunda kapsamlı bir bakış sunar. REHREC modeli ve W-REHREC modeli modelleri, veri eksikliği, Cold-Start ve duygu kullanımının etkili bir şekilde ele alınması gibi zorlukları aşarak öneri araştırmalarında yenilikçi bir paradigma sunar ve bu alanda gelecekteki yenilikler için bir yol haritası oluşturur.

1.5. Tezin Ana Hatları

Bu tezin yapısı, Konum Tabanlı Sosyal Ağlar (LBSN'ler) için öneri modelinin tasarımı, uygulanması ve değerlendirilmesine yönelik olarak düzenlenmiştir. Tez aşağıdaki bölümlere ayrılmıştır:

Mevcut bölüm olan Bölüm 1, araştırma alanımızdaki en güncel durumu özetler ve özellikle LBSN'lerdeki öneri sistemlerine odaklanarak zorlukları ve fırsatları belirler. Tezin arka planını, motivasyonunu ve problem beyanını, ayrıca amaç ve hedeflerini sunar. Ayrıca, araştırma çalışmasının ana katkılarını ve önemini tanımlar, sonraki bölümler için temel bir bağlam oluşturur.

Bölüm 2: Literatür Taraması

Bu bölüm, öneri yaklaşımları, Heterojen Bilgi Ağları (HIN'ler), Konu based Duygu Analizi (ABSA) ve LBSN uygulamalarındaki önceki çalışmalara kapsamlı bir Yorum sunar. Literatür taraması, klasik ve daha yeni öneri tekniklerini eksikliklerini belirterek inceler ve REHREC modeli ve W-REHREC modeli modellerinin yeniliklerine yol açar. Ayrıca, duygu analizinin ve HIN'lerin öneri doğruluğunu artırmadaki değişen rolüne de değinir.

Bölüm 3: Yöntem

Bu bölüm, önerilen yeni REHREC modeli ve W-REHREC modeli modellerinin teorik temellerini ve tasarım felsefelerini açıklar. LBSN'lere uygun bir heterojen bilgi ağı inşa etmek, HIN'e konu based duygu ağırlıkları eklemek ve duygu ağırlıklı Random Walk algoritması geliştirmek gibi teknik konuları kapsar. Meta-yol tabanlı kelime gömmesi ve öneri sıralaması için ağırlıklı PageRank yaklaşımı da metodolojiye dahil edilmiştir, bu bileşenlerin nasıl etkileşerek öneri uygunluğunu ve kullanıcı memnuniyetini artırdığı gösterilir.

Bölüm 4: Deneyler Ve Değerlendirme

Bu bölüm, REHREC modeli ve W-REHREC modeli modellerinin pratik olarak nasıl gerçekleştirildiğine dair bir kılavuz niteliğindedir. LBSN veri yapısı, ABSA tabanlı duygu analizi ve duygu ağırlıklı meta-yol-ların inşası gibi yöntemler açıklanır. Ayrıca, kullanılan yazılım araçları ve programlama kütüphaneleri ile modelin etkin bir şekilde çalıştırılması için optimizasyon teknikleri de bu bölümde açıklanır. Bu bölüm, önerilen modellerin doğruluğunu test etmek için Yelp veri setinde yapılan deneysel analizleri ele alır. Veri seti hazırlığı, değerlendirme ölçütleri ve temel karşılaştırmalar hakkında bilgi verilir. Sonuçlar, REHREC modeli ve W-REHREC modeli modellerinin temel yöntemlerden daha iyi performans gösterdiğini ve veri eksikliği sorunlarında daha üstün olduğunu gösterir. Birkaç temel model ile yapılan kapsamlı karşılaştırma, önerilen modelnin avantajlarını ve dayanıklılığını vurgular

Bölüm 5: Tartışma, Sonuç Ve Geleceğe Yönelik Düşünceler

Bu bölüm, deneysel sonuçların tartışıldığı bölümdür. Önerilen metodolojilerin geleneksel öneri stratejilerinin ana zayıflıklarını nasıl etkili bir şekilde aştığını açıklar. Duygu odaklı yorumların etkileri, meta-yol tabanlı HIN kelime gömmesi yöntemlerinin avantajları ve modelin gerçek dünya LBSN'lerde uygulanmasına ilişkin pratik hususlar ele alınır. Ayrıca, gelecekteki iyileştirmeler için olasılıklar ve modelin diğer alanlarda nasıl uygulanabileceği de vurgulanır. Bu bölüm, tezin tüm katkılarını ve bulgularını özetler. Araştırma hedeflerinin hangi yollarla gerçekleştirildiğini ve REHREC modeli ve W-REHREC modeli modellerinin ne kadar yenilikçi olduğunu vurgular. Son olarak, gelecekteki araştırmalar için daha fazla veri kaynağının entegrasyonu, yöntemimizin diğer LBSN'lerdeki uygulanabilirliğinin incelenmesi veya duygu ağırlıklı HIN'lerin öneri sistemlerinde ölçeklenebilirlik ve çok konuluğunun artırılması gibi öneriler sunar.

Bu yapı, okuyucunun araştırma problemini, alınan yenilikçi yaklaşımı ve önerilen çözümlerin pratik etkilerini giderek daha iyi anlamasını sağlayarak tezin bütünsel bir şekilde sunulmasını hedefler.

2. LİTERATÜR TARAMASI

2.1. Genel Bakış

Literatür taraması bölümü, öneri sistemlerinin gelişiminde temel ve ileri düzey fikirleri açıklar. Bu genel bakışta, heterojen bilgi ağlarına (HIN) uygulanan öneri modelleri ve bunların duygu analizi ile graf-bazlı algoritmaları ile ilgili önemli yaklaşımları, yenilikçi gelişmeleri ve açık sorunları özetliyoruz. Aşağıda, bu panoramik yapının ana içerik alanlarını özetlemekteyiz.

Öneri Sistemleri alt başlığı, en yaygın öneri sistemi türlerinin ve bunların nasıl çalıştığının yanı sıra farklı kullanım durumlarının bir genel bakışıyla başlar. Öge özellikleri arasındaki benzerliği hesaplayarak kullanıcı tercihine göre bir ögeyi önermek için kullanılan kişiselleştirilmiş bir öneri yaklaşımı olan İçerik Tabanlı Filtrelemeyi tanıtır. Ardından, benzer kullanıcıların tercihlerini kullanarak etkili ve sınırlı öneriler oluşturmak için alternatif bir yöntem olarak İşbirlikçi Filtrelemeyi açıklar. Hibrit Öneri Sistemleri, bireysel yaklaşımların zayıflıklarını en aza indiren daha zengin, daha uyarlanabilir modeller oluşturmak için hem içerik tabanlı hem de işbirlikçi yöntemlerin avantajlarından yararlanır.

Heterojen Bilgi Ağı (HIN) inceleme alt başlığı, HIN mimarisini ve karmaşık veri ilişkilerini modellemelerine olanak tanıyan çeşitli bileşenleri açıklar. Adından da anlaşılacağı gibi, aralarında heterojen düğümler ve kenarlar içerebilen HIN'lerin Yapısını ve Bileşenlerini açıklar. Öneri Sistemleri bölümündeki HIN'ler, HIN'lerin yapısını ve bileşenlerini açıklar. Daha sonra, öneri sistemlerinde HIN'lerin veri odaklı uygulamalar ile daha zengin etkileşimleri ve çok türdeki ilişkileri nasıl desteklediği anlatılır. Makine Öğrenme Modelleriyle Entegrasyon alt başlığı, iç görüler veya geliştirilmiş öneriler elde etmeye yardımcı olmak için HIN'lerin nasıl kullanılacağını ve daha gelişmiş makine öğrenme modelleriyle nasıl birleştirileceğini açıklar. HIN'ler için Gömme ve Boyut Azaltma alt başlığı, HIN'lerdeki yapıları çok daha yoğun yapıya sahip bir vektör uzayına eşleme tekniklerini kapsar, böylece büyük veri işleme görevlerini ele alırken kullanılabilir hale getirir. Ayrıca, HIN'lerden gömme sırasında ağı bazı yapısal özelliklerini koruma tekniklerini de ayrıntılı olarak açıklar.

Abstract based duygu analizi (ABSA) alt başlığı, mekan duygu analizini ve öneri modellerindeki rolünü inceleyerek daha fazla ayrıntı ve doğruluk sağlamak için önemini

tartıřır. ABSA'nın genel bir yorumunu ve belirli kenarlara iliřkin kullanıcı tarafından üretilen içeriğin olumsuz, olumlu veya nötr olduđunu ıkaran veya belirleyen yöntemleri açıklar.

Literatür taramasının son bölümü, ađ içindeki düđümlerin sıralanmasında önemli bir rol oynayan PageRank ve onun önemini tartıřır. PageRank'in temel kavramları ve ađ analizindeki etkisiyle bařlar. PageRank Varyantları ve İyileřtirmeleri alt bařlığı, belirli uygulamalar için yapılan adaptasyonları ve daha iyi hesaplama verimliliđi için optimizasyonları açıklar. PageRank Kullanımında Karřılařılan Zorluklar ve Dikkat Edilmesi Gerekenler alt bařlığı uygulama ile ilgili olabilecek sınırlamaları ve pratik sorunları anlatır. Son olarak, Modern Sistemlerde PageRank Entegrasyonu alt bařlığı PageRank'in diđer algoritmalar ve modellerle nasıl entegre edildiđini ve öneri modellerini nasıl iyileřtirdiđini açıklar.

Bu literatür yorumları, bu alanların nasıl bir araya gelerek entegre ve sofistike bir öneri sistemi oluřturduđunu anlamak için sađlam bir temel oluřturur. Öneri yöntemleri, HIN řeması, duygu analizi özellik ıkarımı ve sıralama algoritmalarının kavramları bir araya getirilerek, bu metodolojinin sonraki bölümleri REHREC modeli ve W-REHREC modeli modellerinin bu alanlardan nasıl faydalandıđını, daha kiřiselleřtirilmiř ve ilgili öneriler üretecek yeni bir perspektife nasıl entegre edildiđini gösterir.

2.2. Öneri Sistemleri

ok sayıda seeneđin olduđu durumlarda, insanlar istek listelerini hızla oluřturmak için yardıma ihtiya duyarlar. Öneri sistemleri, insanların gerekten keyif alacakları farklı nesnelere veya ilgi noktalarını keřfetmelerine yardımcı olmak için oluřturulmuřtur. Bu tür sistemler, dijital arayüzlerle eřitli řekillerde sürekli etkileřimde bulunduđumuz modern teknolojinin ayrılmaz bir parası haline gelmiřtir ve kullanıcı deneyimlerini etkilemektedir. Öneri sistemleri, internet tabanlı sosyal ađlar gibi yaygın olarak kullanılan uygulamalardan biri olarak öne ıkmakta ve hedeflenmiř reklamlar aracılıđıyla evrimii dijital platformlarda bařarıyla kullanılmaktadır [1].

2.2.1. Öneri Sistemlerine Yaklaşımlar

Öneri sistemleri, kullanıcı tercihleri doğrultusunda önerilerde bulunmak için kullanıcıya ait mevcut verileri kullanır. Temel ilke, kullanıcılara ilgili içerik veya öğeler sunarak dijital platformlarla olan etkileşimlerini ve deneyimlerini zenginleştirmektir. Öneri sistemleri, esas olarak iki ana yaklaşıma dayanır — İşbirlikçi Filtreleme (CF) ve İçerik Tabanlı (content-based) Filtreleme (CBF).

2.2.1.1. İçerik Tabanlı (content-based) Filtreleme

İçerik tabanlı (content-based) filtreleme öneri sistemleri, kullanıcıların geçmişte beğendiği öğelere benzer öğeleri önermek için öğe özelliklerini ve kullanıcı etkileşimlerini kullanır. Bu yöntem, bir kullanıcının öğeler üzerindeki tercihlerinin, gelecekte de benzer öğeleri beğeneceğini gösterdiği ilkesine dayanır. Bu yaklaşım, öğelerin özelliklerini, yani anahtar kelimeleri/kategorileri/özellikleri inceleyerek ve bunları kullanıcı eğilimleri veya davranışlarıyla karşılaştırarak çalışır. Örneğin, bir kullanıcı makine öğrenmesi hakkında makaleler okuduğunda, İçerik tabanlı (content-based) filtreleme yaklaşımı daha fazla ilgili etiketlenmiş (veya konuyla ilgili) makale önerecektir [2].

İçerik (content) filtreleme, diğer kullanıcıların puanlamalarına dayanmadığı için üstün kişiselleştirme sağlar. Bu, sınırlı kullanıcı etkileşimi olan yeni bir ürün için idealdir. Dahası, İçerik tabanlı (content-based) sistemler, bilgi yayılımı ve kabulü üzerine kurulu öneri profillerini sürekli olarak analiz ederek güncellediğinden, zamanla kullanıcı tercihlerindeki değişikliklere uyum sağlama yeteneğine sahiptir [3].

Ancak, içerik tabanlı (content-based) filtrelemenin bazı dezavantajları da vardır. Büyük bir dezavantaj, aşırı özelleşmedir ve bu, sistemin yalnızca kullanıcının tükettiği öğelere benzer öğeleri önermesi anlamına gelir. Bu, kullanıcı keşfini sınırlayabilir ve onları farklı şeylerle tanışmaktan alıkoyabilir. Son bir endişe ise, sistemin önerilerini sağlamak için ayrıntılı öğe meta verilerine dayanması gerektiğidir; öğelerin özellikleri eksik veya iyi tanımlanmamışsa, sistem tarafından sağlanan öneriler etkili olmayabilir [4].

İçerik tabanlı (content-based) filtrelemeyi geliştirmek için karar ağaçları, k-NN ve SVM gibi farklı makine öğrenimi teknikleri kullanılmıştır. Bu sistemler ayrıca, öneri kalitesini artıran, yapılandırılmamış metni analiz etmek için terim frekansı-ters belge frekansı (term frequency-inverse document frequency TF-IDF) ve kelime gömmesi gibi bir dizi Doğal Dil İşleme (NLP) tekniğini uygular [5, 6].

2.2.1.2. İşbirlikçi Filtreleme (Collaborative Filtering)

İşbirlikçi Filtreleme (Collaborative Filtering), öneri sistemleri için en çok kullanılan ve araştırılan algoritmalardan biridir. Temel mantık diğer kullanıcılarla benzer davranışlar sergileyen bir kullanıcının tercihlerinin diğer kullanıcıların tercihlerine benzer olacaktır. İşbirlikçi filtrelemenin iki ana türü vardır: kullanıcı tabanlı ve öge tabanlı [4].

Kullanıcı Tabanlı İşbirlikçi Filtreleme yöntemde, hedef kullanıcıyla benzer puanlar veren veya benzer şekilde davranan kişiler bulunur ve bu kullanıcıların beğendiği öğeler, hedef kullanıcıya önerilir. Örneğin, Kullanıcı A ve Kullanıcı B birçok aynı filme benzer şekilde puan vermişse, Kullanıcı A'ya, Kullanıcı B'nin yüksek puan verdiği ancak henüz izlemediği filmler önerilir. Kullanıcılar arasında davranışsal bir korelasyon üzerinden yapılan öneriler, topluluk içinde bol miktarda kullanıcı etkileşimi varsa oldukça iyi sonuçlar verir [7].

Öge Tabanlı İşbirlikçi Filtreleme yöntemi, kullanıcı etkileşimlerine dayanarak öğeler arasındaki benzerliği ölçer. İki öğenin ilişkisi, aynı kullanıcı grubundan aynı puanı alıp almadıkları ile belirlenir. Örneğin, bazı kullanıcılar iki kitabı birden yüksek puanladıysa, bir başka kullanıcı bu kitaplardan birini beğendiyse, diğer kitap da önerilir. Daha büyük ölçeklerde İşbirlikçi filtreleme gerektiğinde, öge tabanlı İşbirlikçi filtreleme, kullanıcı tabanlı yöntemlere göre daha iyi ölçeklenebilir [8].

İşbirlikçi filtreleme avantajlarından biri, öge özelliklerine bağlı kalmak yerine, kullanıcıların davranış ve tercihlerini toplu olarak kullanır. Kategoriler arası öneriler yapabilir; örneğin, benzer zevklere sahip anonim kullanıcılar arasındaki korelasyonları bularak, bir kullanıcıya yeni bir restoran önerisi yapabilir [9].

Ancak işbirlikçi filtreleme sistemlerinde şu sorunlar gözlemlenebilir:

Cold-Start sorunu, yeni kullanıcıların veya öğelerin çok az etkileşim geçmişine sahip olması durumunda ortaya çıkar. Kullanıcı veya öğeler hakkında yeterli veri olmadığı için sistem

faydalı bir öneri üretemez. Sistem, benzerlikleri veya tercihleri geçmiş veri olmadan hesaplamak zorunda kalır ve bu da sistemi etkisiz hale getirir [10].

Veri Seyrekliği: Kullanıcı-öge etkileşim matrisi, geniş öge katalogları ve çok sayıda kullanıcı içeren bazı sistemlerde oldukça seyrek olabilir. Bu tür bir seyrekliğin olduğu durumlarda, sistem güvenilir komşular bulmakta zorlanır ve yeterli öneri üretemez [11].

Ölçeklenebilirlik: Kullanıcı-kullanıcı benzerliğinden İşbirlikçi filtrelemeye geçişle birlikte, kullanıcı tabanı genişledikçe ve dikkate alınacak öge sayısı önemli ölçüde arttıkça performans sorunları ortaya çıkmıştır. Bu sorunu çözmek için, boyut azaltma ve dağıtık hesaplama gibi verimli hesaplama stratejileri genellikle gereklidir. Bu yöntemler, sistem performansını zaman içinde korumak için kullanılmaktadır [12].

2.2.1.3. Hibrit Öneri Sistemleri

Hibrit öneri sistemleri, hem İşbirlikçi hem de içerik tabanlı (content-based) filtrelemenin sahip olduğu eksiklikleri gidermek için tasarlanmıştır. Hibrit sistemler, her iki yöntemden gelen güçlü yönlerini birleştirerek daha doğru ve bütünsel öneriler sunar. Örneğin, hibrit bir model başlangıçta içerik tabanlı (content-based) filtrelemeyi kullanarak yeni bir kullanıcıya öneriler sağlayabilir ve daha fazla etkileşim verisi mevcut olduğunda İşbirlikçi filtrelemeye geçebilir [13].

Hibrit Sistem Türleri:

Ağırlıklı Hibrit: Bu yaklaşımda, içerik tabanlı ve işbirlikçi filtreleme algoritmalarından elde edilen puanlar, her bileşene belirli ağırlıklar verilerek birleştirilir. Böylece, bireysel yöntemlerin güçlü yönlerinden daha fazla, zayıf yönlerinden ise daha az yararlanır.

Değişken Hibrit: Öneri yöntemlerini bağlama göre değiştiren bir yaklaşımdır. Örneğin, kullanıcı verilerinin az olduğu durumlarda içerik tabanlı filtreleme kullanılabilir, yeterli veri olduğunda ise işbirlikçi filtreleme öneri çıkar [14].

Özellik Artırımı: Bu yöntemde, bir öneri yönteminden elde edilen çıktı, diğer bir yöntem giriş verisi olarak kullanılır. Örneğin, işbirlikçi filtreleme çıktıları, içerik tabanlı algoritmalara kullanıcı tercihlerine dair daha doğru bilgi sağlamak için kullanılabilir [15].

E-ticaret, video akışı ve sosyal medya platformları gibi alanlarda hibrit modeller son derece etkili olduğunu kanıtlamıştır. Bu sistemler, çeşitli öneri yaklaşımlarının avantajlarını bir araya getirerek Cold-Start ve veri seyrekliği sorunlarının üstesinden gelebilir. Bu da daha iyi bir kullanıcı deneyimi ve etkileşimine yol açabilir [16].

Geleneksel öneri sistemleri olan içerik tabanlı filtreleme, işbirlikçi filtreleme ve hibrit modeller dijital platformlarda kullanıcı deneyimini büyük ölçüde geliştirmiş olsada, ölçeklenebilirlik ve Cold-Start sorunları gibi sınırlamaları bulunmaktadır [17]. Bu tür zorluklar, farklı bağlantılı veri kaynaklarını ele alabilecek ve bunlardan yararlanabilecek ileri modellerin gerekliliğini vurgulamaktadır. Heterojen Bilgi Ağları (HIN) bu noktada avantajlarını ortaya koymaktadır. HIN'ler, farklı türde varlıklar ve bunlar arasındaki ilişkileri temsil eden karmaşık ağ yapıları kullanarak önerileri daha zengin, çok boyutlu bir yaklaşımla sunar. HIN'ler, heterojen veri türlerini entegre ederek üstün tahmin doğruluğu, kullanıcı memnuniyeti ve kapsamlı değerlendirmeler sağlar.

2.3. Heterojen Bilgi Ağı (HIN)

Dünyamız etkileşimde bulunan ve kendilerini bir sistem olarak adlandırılan çeşitli ve heterojen varlık türlerinden oluşur. Heterojen Bilgi Ağı (HIN) olarak adlandırılan bir ağ, bu varlıklar ve bunların entegrasyonu ile oluşturulan bir ağdır. HIN'lerin içeriklerinde, daha derinlerde anlamsal bağlantılar ima edildiğinden, ortaya çıkarılmayı bekleyen çok sayıda güçlü bilgi vardır. Ve bu tür ağlar ayrıca e-ticaret siteleri ve Facebook gibi sosyal ağlar gibi platformlar tarafından doğal olarak oluşturulur ve gerçek hayatta sıklıkla karşımıza çıkar.

2.3.1. HIN'lerin Yapısı ve Bileşenleri

HIN'ler bir düğüm kümesi ve bir kenar kümesi içeren graflar olarak tanımlanabilir. Her düğüme ve her kenara farklı bir tür atanabilir ve bu da ağın ilişkileri heterojen bir şekilde

modellemesine olanak tanır [18]. Öneriler için kullanılan bir HIN'de temsil edilebilecek örnek varlıklar şunları içerir:

- Kullanıcı Düğümleri (U): Ağdaki her kullanıcılar.
- Öge veya mekan düğümleri (B): Kullanıcıların etkileşimde bulunduğu ürünler, hizmetler ve konumlar.
- Yorum Düğümleri (R): Metin dizisi olarak, kullanıcıların derinlemesine yorum ve geri bildirimlerde bulunmasıyla üretilen içeriklerdir.
- Kategori Düğümleri (Ca): Ögelerle ilgili çeşitli kategorileri içeren nitelikler.

Heterojen Bilgi Ağları (HIN) düğümlerin ve kenarların aynı tipte olduğu homojen ağların aksine düğümlerin ve bağlantıların farklı tiplerde olabildiği heterojen verileri temsil edebilirler [19]. Homojen ağlarda, benzerlik ölçüsü Öklid mesafesi, Kosinüs benzerliği ve Jaccard endeksi gibi mesafe ölçütlerine dayanır. Ancak HIN farklı tipte düğümler (node)den ve ilişkilerden oluşur, bu nedenle bu ölçütler HIN'ler için uygun değildir. k-means gibi mevcut standart kümeleme algoritmaları doğrudan HIN'e uygulanır, ancak varlıklar farklı kategorilere ve tiplere ait olduğundan uygulanabilir değildir [20].

2.3.2. Öneri Sistemlerinde HIN'lerin Uygulamaları

HIN'lerin farklı türdeki varlıklardan ve ilişkilerden gelen verileri entegre etme yeteneği, bunların öneri sistemlerinde yaygın olarak kullanılmasına yol açmıştır. Bir öneri modeline kolayca entegre edilebilen HIN'ler aşağıdaki bazı avantajları sunabilir.

Daha Fazla Bağlam Farkında Öneriler: HIN'ler, farklı varlıklar arasındaki daha karmaşık ilişkileri modellememize olanak tanır ve bu da daha fazla bağlam farkında öneriyle sonuçlanır [21]. HIN'ler, basit kullanıcı-öge eşleştirmesinden ziyade bağlamın zenginliğine dayalı öneriler sağlamak için kullanıcı etkileşimleri, öge kategorileri ve sosyal bağlantılar gibi ek veri etkileşimlerini kullanabilir [21].

Veri Seyrekliği İçin Bir Çözüm: HIN'ler genellikle işbirlikçi filtrelemede ortaya çıkan veri seyrekliği sorununu azaltabilir. Doğrudan etkileşim verileri kullanıcı-öge düzeyinde önerilerde bulunmak için çok seyrek olsada, yardımcı veriler ve çeşitli düğüm türleri bunları kolaylaştırır [22].

Cold-Start Çözümleri: HIN'ler, yeni kullanıcılar veya öğeler için tahminlerde bulunmak amacıyla kullanıcı nitelikleri ve öğe meta verileri gibi yardımcı bilgileri entegre edebilir ve böylece Cold-Start sorunlarının etkisini etkili bir şekilde azaltabilir [23].

Verilerin Zengin Temsili: Heterojen Bilgi Ağları (HIN'ler) çeşitli düğümlerin yanı sıra ilişkileri de içerir. . Burada her düğüm kullanıcı, konum veya kategori gibi çeşitli varlık türlerini temsil edebilir ve iki farklı düğüm türü arasındaki bir ilişki (yani kullanıcı-öğe ilişkisi) farklı türdeki düğümler arasındaki çok değişik ilişkilendirmeleri yakalayabilir [24].

Karmaşık Etkileşimleri Yakalamada Esneklik: Meta-yollar varlıklar arasındaki anlamlı ilişkileri temsil eden düğüm ve kenar türlerinin dizileridir. Meta-yollar, yüksek düzeyli etkileşimleri yakalamak ve öneri sisteminin model temsil yeteneğini geliştirmek için esnek bir araç sağlar [25].

Gelişmiş Kişiselleştirme: HIN'ler aracılığıyla çok türdeki verilerin dahil edilmesiyle, öneri sistemleri bağlama duyarlı kişiselleştirilmiş öneriler sağlayabilir [26].

2.3.3. Makine Öğrenme Modelleriyle Entegrasyon

Makine öğrenimi algoritmaları HIN'lerin analizini iyileştirmeye yardımcı olabilir. HIN kelime gömmeleri — [1]'deki gibi — düğümleri ve ilişkileri sürekli bir vektör uzayına atar ve benzer düğümler/ilişkiler birbirine yakın olur. Sonuç olarak, lojistik regresyon, sinir ağları veya gradient boosting gibi standart makine öğrenimi modelleri öneriye uygulanabilir [27].

Daha yakın zamanda, çeşitli derin öğrenme yaklaşımları ve özellikle graf sinir ağları (GNN'ler), HIN'lerdeki öğelerin karmaşık etkileşimlerinin daha fazla modellenmesi için kullanılmıştır. GNN'ler, ağın heterojenliğini korurken yakındaki düğümlerden bilgi toplar ve böylece daha zengin bilgilere sahip düğüm temsillerini öğrenir [28].

HIN'in temeli, farklı türdeki ilişki anlayış yapısı ve onların bağlam farkında öneriler üretilmesinde kullanımı üzerine kurulmuş olsada, HIN'lerin bu potansiyelinden yararlanmak için bilgi gösterimi için gelişmiş yöntemler esastır. Kelime gömmesi yaklaşımları, yapısal ve anlamsal benzerlikleri korurken bir HIN'deki düğümleri ve ilişkileri sürekli bir alana dönüştürür [29]. Bu kelime gömmeleri benzerlikleri hızlı bir şekilde hesaplamak, makine öğrenme modelleri oluşturmak ve kümeleme gibi diğer alt akış görevleri için gereken temel kelime gömmeleridir. Aşağıdaki bölümde, öneri sistemlerinin performansının iyileştirmesini

ve ölçeklenebilirliğini sağlayan farklı HIN kelime gömmesi yöntemlerinin detaylarını tanıtacağız.

2.3.4. HIN'ler için Kelime Gömmesi ve Boyut Azaltma

Heterojen Bilgi Ağlarından (HIN) ek iç görüler elde etmek için düğüm kelime gömmeleri gereklidir. Düğüm kelime gömmesi yaklaşımları, ağdaki düğümleri sürekli bir vektör uzayına kelime gömmesi yapmak için kullanılır ve ağda bulunan yapısal ve anlamsal bilgileri korur. Kelime gömmeleri öneri sistemlerinin karmaşık ilişkileri daha etkili bir şekilde tanımlamasına olanak tanırken veri alma ve analizinin genel verimliliğini artırır.

2.3.4.1. HIN'ler için Kelime Gömmesi

Kelime gömmesi tabanlı teknikler ve benzerlik algılama yöntemleri, HIN'lerden yararlı bilgiler çıkarmak için çeşitli anlamlı yaklaşımların tanımlanmasına yol açmıştır. Bu yaklaşım, yalnızca düğümler arasındaki doğrudan ilişkileri yakalayan geleneksel benzerlik tabanlı stratejilerin temel sınırlamaları olan veri seyrekliği ve gürültüsü problemini çözüm üretir.

Bu konuyla ilgili makalelerden birinde [30], HIN'deki hem alt öğeler hem de kullanıcılar esasen karşılaştırılabilir olacak şekilde sınırlandırılmıştır ve bu karşılaştırmalı düğümler bir bütün olarak öneri modelini geliştirmek için bir matris çarpanlarına ayırma (Matrix Factorization - MF) yaklaşımına dahil edilmiştir. Bu entegrasyon, modelin verilerden daha doğru sonuçlar çıkarma kapasitesini daha da iyileştirir. Meta-yol tabanlı yöntemler [31], graf üzerindeki ilişki benzerliğini algılar ve düğüm kelime gömmesi yoluyla karmaşık ilişkileri daha iyi temsil etmek için bunların üzerine matris çarpanlarına ayırma modelleri oluşturur.

Seyreklik ve gürültü sorunları yaşayan geleneksel benzerlik tabanlı yaklaşımların aksine, kelime gömmesi tabanlı yöntemler, yüksek boyutlu ağ alanını düşük boyutlu yoğun bir gösterime sıkıştırma yeteneğine sahiptir. Bu türe bir örnek, öge değerlendirme prosedürleriyle bütünleştirilmiş meta-yol sonuçlarına dayalı stratejilerin kelime gömmesi yapmaktır [14]. Bu yaklaşım, meta-yolları kullanarak, varlıklar arasındaki karmaşık ilişkileri yansıtabilecek ve daha iyi öneriler önerecek kelime gömmeleri öğrenilmesini garanti eder.

Daha yakın zamanda, öneri performansını artırmak için, gelişmiş kelime gömmesi tekniği, graf verilerine dayalı yapay sinir ağlarını kullanmıştır. Örneğin, bir çalışma [33] ayrıca graf-bazlı yapıdan öğrenilen bir sinir ağı kelime gömmelerine dayalı öneri sistemleri sunmuştur. Farklı bir yöntem [34], bir dizi meta-yol tarafından dolaşan düğüm kelime gömmelerini öğrenen bir algoritmadan oluşur ve bu da etkili bilgi çıkarımına ve yoğunlaştırılmış bilgi çıktılarına izin verir. Bu tür modeller, düğümler arasındaki doğrusal olmayan etkileşimleri ve karmaşık bağımlılıkları modellemede yetenekler göstermiştir.

Heterojen ağların gizli yapıları içine gömülmüş bilgileri çıkarmaya çalışan işlevsel yöntemler üzerine yapılan araştırmalar, otomatik kodlayıcılar (autoencoders) [35] gibi derin öğrenme yöntemleri, ağlar arasındaki doğrusal olmayan ilişki özelliğinin keşfedilmesinde önemli bir rol oynamıştır. Bu, ağın önemli niteliklerini ve ilişkilerini korurken düğümlerin kompakt temsillerini öğrenmemizi sağladığı için yüksek kaliteli öneriler sunabilmeye yönelik önemli bir adımdır [36]. Dahası, birkaç Random Walk algoritmasına dayalı yaklaşım, ağlarda örnekleme ve bu örneklere dayalı düğümler üretmeye başarıyla uygulanmıştır. Düğüm dizileri üretmek için ağlarda Random Walk algoritması gerçekleştiren [37, 38] çalışmalarında olduğu gibi, bu diziler Skip-Gram algoritması [39] gibi kelime gömmesi modelleri için girdi olarak kullanılmıştır. Bunu yaparak, kelime gömmeleri ağın hem yerel hem de küresel yapısını yakalayabilir ve düğümler arasındaki ilişkilerin daha iyi anlaşılmasını sağlayabilir.

2.3.4.2. Boyutu Azaltmak İçin Bir Ağ Kelime Gömmesi

Kelime gömmesi yaklaşımı hem düğümlerin boyutlarını azaltmayı hem de ağdaki önemli bilgileri tutarken yine de iyileştirilmiş alan karmaşıklığına izin vermeyi amaçlamaktadır. Bu nedenle, özellikle öneri sistemi tasarımları için yararlı olan başarılı bir yöntem olmuştur [40, 41].

Tarihsel olarak, ilk kelime gömmesi algoritmaları düğüm boyutluluğunun azaltılmasına odaklanmıştır [42]. Yavaş yavaş, dinamik graf kelime gömmelerini de içerecek şekilde ilerlemiştir. Örneğin bir çalışma [43], iki düğüm birbirine bağlı olduğunda komşu düğümlerin temsil alanında birbirine yakın kalmasını sınırlayarak her düğümün temsilinin graf çarpanlarına ayrılmasını önermiştir, ancak bunun tersi geçerli değildir; buda ağdaki bu düğümler arasındaki bazı ilişkileri yansıtır. Daha sonraki bir çalışmada [44], bu yaklaşım,

kelime gömmesi sürecini daha iyi temsil eden düğümlerdeki ikinci dereceden benzerliklerin toplanmasıyla iyileştirilmiştir.

Otomatik kodlayıcıların bildirilen bir diğer kullanımı, kelime gömmesinde doğrusal olmayan bilgileri bulmaktır [45]. Otomatik kodlayıcılar yüksek boyutlu giriş verilerini daha düşük boyutlu bir alana sıkıştırarak ve daha sonra orijinal verileri yeniden oluşturan bir tür sinir ağıdır; buda doğrusal yöntemlerle görünmeyen kelime gömmesi alanındaki doğrusal olmayan yapıyı yakalar.

Ayrıca, bazı araştırmalar Random Walk algoritması yöntemini kullanarak ağdan düğümleri örneklemiş ve üretilen örneklere göre kelime gömmeleri çıkarmıştır [46]. Belirtilen meta-yollara göre Random Walk algoritmasına dayalı düğüm dizileri üreterek Skip-Gram algoritması modelini heterojen bilgi ağlarına (HIN'ler) uyarlayabilir [47]. Bu yaklaşım kelime gömmesinin bir düğümden diğerine (Random Walk algoritması yoluyla) öğrenmesini sağlar ve hem ağı yerel yapısını hem de küresel meta-yol odaklı yapısını korur.

Bununla birlikte, düğümlerin farklı özellikleri ve heterojen graflardaki anlamsal farklılıklar nedeniyle mevcut tüm kelime gömmesi teknikleri HIN'ler için uygulanabilir değildir. Çeşitli düğüm tipleri arasındaki meta-yolların anlamsal anlamı çeşitlilik olduğundan, tüm düğümleri aynı alana tekdüze bir şekilde kelime gömmesi yapmak mantıklı olmayacaktır. Bu nedenle, araştırmacılar gösterimleri heterojen graflara daha iyi uyarlamak için farklı algoritmalar önerdiler [48, 49, 50]. Örnekler arasında, kelime gömmelerinin meta-yollar arasında dolaşırken öğrenildiği ve vektör gösterimlerinin ağı heterojenliğini ve anlamsal çeşitliliğini yansıttığından emin olunan derin yürüme teknikleri yer alır [28].

Burada önemli olan birincil hedefi boyut azaltma olan HIN'in önemli özelliklerini koruyan kelime gömmelerini öğrenmek olacaktır. Bu sorunların üstesinden gelmek için, önerideki etkenliği artırmak için çeşitli anlamsal somutlaştırılmış ilişkilerin bilgilerini korumayı amaçlayarak kelime gömmesi ağları için son zamanlarda Skip-Gram algoritması [39] ve meta-yol tabanlı derin yürüme yaklaşımları önerilmiştir [28].

Kelime gömmesi yöntemlerinin temel bileşeni boyutsal karmaşıklığı azaltmak ve HIN'lerde ağ özelliklerini korumak olsa da, öneriler birden fazla biçimde geldiği için, bu öneri sistemleri için, nihayetinde daha ince taneli bağlam farkında öneriler yapmak, kelime gömmelerinin bir hedefidir. Kullanıcı tarafından oluşturulan içeriği ayrıntılı bir düzeyde anlamak, bu önerileri daha alakalı ve kişiselleştirilmiş hale getirmeye yardımcı olabilecek

temel anahtar faktörlerden biridir. İşte bu noktada, Konu-Bazlı Duygu Analizi (ABSA) bize yardımcı olur. ABSA, kullanıcı memnuniyetini artıran ve HIN kelime gömmeleri tarafından tahmin edilen modele ince taneli tamamlayıcılık sağlayan yorumlanabilirlik ve yüksek ayrıntı düzeyi içgörülerini ekleyerek, bir yorumdan çıkarılan belirli niteliklerin bağımsız konu çıkarımına izin verir.

2.4. Konu-Bazlı Duygu Analizi (ABSA)

Kullanıcı tarafından oluşturulan içerikten bireysel konulara veya özelliklere göre daha ince ayrıntılı bir şekilde duyguyu anlamak için, Konu-Bazlı Duygu Analizi (ABSA) gibi daha gelişmiş doğal dil işleme (NLP) tekniklerini kullanabiliriz. Tüm metin için genel bir duygu puanı veren geleneksel duygu analizinin aksine, ABSA hem olumlu hem de olumsuz duyguları elde ederek bir yorumda bulunan belirli özellikler hakkında konuşmaya olanak tanır. Bu tür ince ayrıntılı yorumlama, daha fazla müşteri memnuniyeti verisi oluşturulmasını sağlar ve önerileri müşterilerin tercihleriyle daha kesin bir şekilde uyumlu hale getirdiği için öneri kalitesinin iyileştirilmesine yardımcı olur [51].

2.4.1. ABSA'ya Genel Bakış

Konu-Bazlı Duygu Analizi (ABSA), bir metinde belirtilen farklı konuları çıkarmak ve onlara duygu değerleri (olumlu, olumsuz veya nötr) vermektir oluşur. Bu nedenle, örneğin bir restoran yorumunda, "yemek kalitesi", "hizmet" vb. gibi kendi duygusu ile etiketlenebilen farklı özellikler olabilir. Bu konu odaklı analiz, kişiselleştirme ve alaka düzeyini iyileştirmek için öneri sistemlerine yararlı bilgiler sağlar [52].

ABSA genellikle şu adımları içerir:

- Konu Çıkarımı: Metindeki ilgili nitelikleri veya temaları bulma.
- Duygu Sınıflandırması: Çıkarılan her bir konunun belirli bir duygu kutupluluğu ile sınıflandırılması.
- Konu-Duygu Eşleştirmesi: Metnin her bir duygu puanı için, metin hakkında ayrıntılı bir bilgiye sahip olmak amacıyla ilgili konu ile bağlantı kurulması [53].

2.4.2. ABSA'daki Teknikler

ABSA'nın metodolojileri, kural tabanlı yaklaşımlardan makine öğrenme yaklaşımlarına doğru muazzam bir şekilde deđiřti. Daha eski kural tabanlı yaklaşımlar, konu çıkarma ve duygu belirleme için önceden tanımlanmış sözlükler ve sözdizimsel (syntactic) kalıplar temelinde çalışıyordu. Basit olmalarına rağmen, bu teknikler esnek olmama eğilimindeydi ve incelikli veya alan bağımlı dile uyarlanması zordu [54].

ABSA'daki son gelişmeler, makine öğreniminin, özellikle de derin öğrenmenin yükselişine bağlanabilir. Evrişimsel sinir ağıları (Convolutional Neural Networks - CNNs) [54] ve tekrarlayan sinir ağıları (recurrent neural networks - RNNs) [55] gibi son konu çıkarma ve sınıflandırma yöntemleri, büyük derlemlerden bağlam öğrenerek daha doğru sonuçlar elde etmek için uygundur. Son yıllarda, dönüřtürücü (transformer) tabanlı mimariler (BERT: Bidirectional Encoder Representations from Transformers), metin içindeki bağlamsal ilişkileri ve bağımlılıkları modelleme yeteneklerinden yararlanarak çeşitli ABSA ön ayarlarında çığır açan performans elde etti. İlk olarak, BERT tabanlı ABSA modelleri, bağlamı daha iyi anladıkları için cümlelerdeki anlamlı kelime anlamlarını tanımada iyidir [56].

Graf sinir ağıları (GNN'ler), kelimeleri ve konu lerini düğümler ve kenarlardan oluşan bir grafa dönüřtürerek Konu-Bazlı Duygu Analizi ile iyileřtirmenin başka bir yolu olarak kabul edilir. Bu yöntem, kelimelerin bir cümlede düzenlendiđi karmaşık biçime dayalı olarak bilgilendirilmiş kararlar alarak konu çıkarma ve duygu sınıflandırmasını mümkün kılar ve tahmin performanslarını iyileřtirir [57].

2.4.3. ABSA'daki Zorluklar ve Tekniklerindeki Gelişmeler

ABSA'nın avantajları olmasına rağmen bazı zorlukları da vardır. Zorluklar řu şekilde sıralanabilir:

Konu Çıkarma Karmaşıklığı: Yapılandırılmamış bir metnindeki konuları doğru bir şekilde tanımlamak zordur (örneğin örtülü konular olduđu veya metnin alan özgü bir dilde yazılmış olduđu durumlarda). Bu genellikle farklı alanlar arasında iyi genelleme sađlayan karmaşık modeller gerektirir [61].

Duygu Belirsizliđi: İnsanların çeşitli ruh halleri ve hisleri vardır, bu da duygu sınıflandırmasını zorlařtırır. Aynı insan yorumlarında ayırıcı duygu örnekleri ve karışık

duygulara sahip alaycılık veya deyimlerin kullanımı bu süreci karmaşık hale getirebilir ve bağlamsal anlayış gerektirir [62].

Eğitim İçin Veri Sıkıntısı: Bir ABSA modelini eğitmek için açıklamalı veri kümelerine ihtiyaç duyulur, ancak bunlar kolayca bulunamayabilir veya gerekli kalitede yeterli olmayabilir [60]. Bu veri eksikliği, ABSA modellerinin ölçeklenebilirliğini ve geliştirilebilirliğini sınırlar [63].

ABSA'nın son gelişimi esas olarak belirsizliğin ve yorumlanabilir modellerin etkili tasarımına odaklanmıştır. Konuya özel veri kümelerinde ince ayarlı önceden eğitilmiş dil modellerinden transfer öğreniminin ortaya çıkmasıyla, özellikle BERT [63], RoBERTa [64] ve ALBERT [65] gibi varyantları olmak üzere son teknoloji dönüştürücü (transformer) mimarileri daha doğru konu çıkarma ve duygu sınıflandırmasını kolaylaştırmıştır. Bu tür modeller, bir cümledeki kelimeler arasındaki ilişkiyi veya alaka düzeyini öğrenmelerine olanak tanıyan ve böylece verilerin daha bağlamsal anlaşılmasına ve dolayısıyla geleneksel RNN, CNN tabanlı yaklaşımlardan daha iyi performansa yol açan öz-dikkate (self-attention) dayanmaktadır.

Ek olarak, konu çıkarma ve duygunun doğruluğunu artırmak için kural tabanlı yöntemleri ve geleneksel makine öğrenimi modellerini entegre eden hibrit yöntemler araştırılmıştır. Bu yaklaşımlarda, konular önce kural tabanlı dilbilimsel yöntemlerle tanınır ve daha fazla makine öğrenimi modeli çıktılarına kategorilere göre rafine etmek üzere eğitilir [65]. Dahası, sinir ağı modelleri, konu veya duyguyu tespit ederken kelimelere ve ifadelere özel dikkat göstermeye yardımcı olan bir dikkat (attention) mekanizmasıyla donatılmıştır [66].

Bir diğer umut verici yön ise tanımlanan konu ile ilgili metnin bazı bölümlerine dikkat etmeyi öğrenen konuya özel dikkat modelleridir. Bu bağlamda, bu modeller dikkatlerini incelenen konuya göre belirler [67], bu da daha etkili duygu sınıflandırmasını kolaylaştırır.

ABSA'yı geliştirmek için, harici bilgi kaynakları olarak alan özgü sözlükler ve ontolojilerin kullanımını da içeren çeşitli teknikler önerilmiştir. Bu tür bilgiler belirsizliği önlemeye ve ek duygu bağlamı sağlamaya yardımcı olabilir, nüans duygularını açıklığa kavuşturabilir [68]. Diğer modeller, örneğin konu modellemesini ve konu ortak çıkarımını ortaklaşa eğiterek [69], duygu sınıflandırıcılarının sağlamlığını ve geliştirilebilirliğini artırabilecek ABSA için çok-görevli öğrenme (multi-task learning) stratejileri önermektedir.

2.4.4. ABSA'nın Öneri Sistemlerinde Uygulamaları

E-ticaret, misafirperverlik ve eğlence gibi çok sayıda kullanıcı tarafından oluşturulan içerik barındıran alanlarda, ABSA'nın öneri sistemleriyle birleştirildiğinde etkili olduğu gösterilmiştir [58]. ABSA, sistemlerin belirli bir ürün veya hizmetin her bir özelliğiyle ilişkili duyguları analiz ederek kullanıcı tercihleriyle yakından eşleşen daha kişiselleştirilmiş öneriler sunmasını sağlar. Örneğin, restoranlarla ilgili yorumlarda hızlı hizmeti yüksek öncelikli niteliklerinden biri olarak ilişkilendiren bir kullanıcıya, yorumların hizmetin hızıyla ilgili olumlu özetler bildirdiği yerler önerilecektir [59].

ABSA ayrıntılı duygu bilgisi içeren öneri sistemleri sağlasa da, bir ağdaki varlık sıralaması hala gerçek bir bekleyen çalışmadır. Graf-bazlı sıralama algoritmaları, düğümleri ağ yapısı hakkında temel bilgiye göre sıralayan değerli tekniklerdir. PageRank, ağ/graf-bazlı verilerde yaygın olarak kullanılan klasik bir algoritmadır. Başlangıçta web sayfalarını sıralamak için geliştirilen PageRank'in icadından bu yana, bu kavram da öneri sistemini iyileştirmek için etkili düğümleri belirlemek ve önerilen öğelerin daha alakalı olmasını sağlamak gibi bir dizi uygulamaya dönüştürülmüştür. Bir sonraki bölümde graf-bazlı sistemlerde PageRank'in teorisi ve pratik uygulaması tartışılmaktadır.

2.5. Graf-bazlı Sıralama Algoritmaları: PageRank

Graf-bazlı sıralama algoritmaları, graf yapısı ve bağlantılarına dayanarak bir düğümün ağdaki önemini tahmin etmek için olmazsa olmazdır. Bu algoritmalar arasında, PageRank tartışmasız en ünlüsüdür ve kolayca en yaygın olanlardan biridir. PageRank, Larry Page ve Sergey Brin tarafından web sayfalarını sıralamak için yaratılmış bir algoritmadır ancak bilgi çıkarımı, sosyal ağ analizi ve öneri sistemi çevrelerinde çeşitli değiştirilmiş biçimlerde de kullanılmıştır.

2.5.1. PageRank Algoritmasına Genel Bakış

PageRank, bir grafa atanan sayı ağırlıklarına göre her düğümün önemini sıralamak için kullanılan yinelemeli bir algoritmadır. PageRank'in ardındaki sezgi, bir düğümün öneminin yalnızca dış değerlendirme notune, yani bağlandığı düğümlerin miktarına değil, aynı zamanda kenarları ona doğru yönlendiren düğümlerin kalitesine (önemi) de bağlı olduğudur.

Bu yinelemeli özellik, PageRank'in önemli düğümleri etkili bir şekilde keşfetmesini sağlar [70].

Bir düğümün puanı PageRank tarafından şu şekilde hesaplanır:

$$PR(r_i) = (1 - d) + d \sum_{r_j \in M(r_i)} \frac{PR(r_j)}{|M(r_j)|} \times PR(w_{ij}) \quad (1)$$

Formüdeki ifadeler şunları gösterir:

- $PR(r_i)$, r_i düğümün PageRank değeri.
- d , genellikle 0,85 ile 0,95 arasında seçilen sönümlenme faktörüdür; bu d olasılığıyla kullanıcının yeni bir rastgele yerleştirilmiş düğüme atlamak yerine diğer düğümlere geçerek Random Walk'a devam edeceği anlamına gelir [71].
- $|M(r_j)|$, r_j düğümünden çıkan bağlantıların sayısıdır.
- $M(r_j)$, r_j düğümüne bağlanan düğümlerin kümesidir.
- w_{ij} , r_i ve r_j yorumların arasındaki kenarların ağırlığıdır .
- $PR(w_{ij})$, r_j den r_i ye olan bağlantının önemini ayarlayan bir ağırlık faktörüdür.

Algoritmanın bu döngü yapısı, PageRank değerlerinin yinelemeler arasında değişmeyi bıraktığı yakınsamaya kadar devam eder [72].

2.5.2. Çeşitli Alanlarda PageRank Uygulamaları

2.5.2.1. Web Sayfası Sıralaması

PageRank ilk olarak web sayfalarını sıralamak için tanıtıldı ve en çok bilineni oldu. PageRank, arama motorlarının sayfaları yalnızca içeriklerine göre değil, aynı zamanda ne kadar alakalı ve bağlantılı olduklarına göre de sıralamasına olanak sağladı; bu, web sayfalarını düğümler ve hyperlinkleri yönlendirilmiş kenarlar olarak ele alarak elde edildi. Bu sıralama yöntemi yalnızca sayfaların iyi bağlantılı olduğundan emin olmakla kalmadı, aynı zamanda diğer alakalı önemli sayfalarla da bağlantılı olduğu ve böylece arama sonuçlarının kalitesi iyileştirildi [73].

2.5.2.2. Sosyal Ağ Analizi

PageRank algoritması, sosyal ağları, bir kişinin ağ üzerinde ne kadar etkisi olduğunu bulmak için, birbiriyle bağlantılılık açısından incelemek üzere değiştirildi. Örneğin, düğümler kullanıcıları temsil edebilirken, kenarlar arkadaşlığı veya takibi ifade edebilir. PageRank'in kullanıcıları bağlantılarına göre bir puan verdiğini bilerek, platform bu tekniği uygulayabilir ve bilgiyi yayma konusunda en yetenekli kişileri veya seçici pazarlama kampanyalarında ana odak noktaları olacak kişileri bulabilir [74].

2.5.2.3. Öneri Sistemlerinde PageRank

PageRank, heterojen bir ağdaki merkeziliklerine göre öğeleri, kullanıcıları veya diğer ilgi çekici varlıkları sıralamak için öneri sistemlerinde kullanılır. Bu bağlamda, PageRank'in, kullanıcı etkileşimleriyle sınırlı olmayan ancak ağ yapısı üzerindeki dolaylı birliktelikleri de içerebilen ağlara öneri modellerini uygulamak için kullanıldığı ilgili bir örneği gösteriyor [75]. Bu tür ilişkileri bulmak için tasarlanmış sistemler için veri kümesinde birçok eş zamanlı faktörden oluşan karmaşık kullanıcı-öge ilişkileriyle kullanıcıya yönelik önerilerin sağlanmasını ve alakalılığını artırır.

PageRank algoritmasına dayalı birkaç algoritma, kullanıcı yorum grafi ve kullanıcı-öge etkileşimi grafi üzerinde analiz gerçekleştirerek en iyi dereceli ürün veya hizmeti belirlemek için kullanılmıştır. Bu algoritmalar, kenar ağırlıkları gibi ek faktörleri hesaba katarak kullanıcıya daha kişiselleştirilmiş ve bağlam farkında bir şekilde önerilerde bulunabilir. Bu durumda, kullanıcı güveni veya yorum duygusu hakkında bilgi verir [76].

2.5.3. PageRank'in Varyantları ve Geliştirmeleri

Yıllar geçtikçe, orijinal PageRank algoritmasının birçok varyasyonu, ilk versiyonun bazı sınırlamalarını gidermek ve çeşitli uygulamalar için verimliliğini daha da artırmak amacıyla ortaya çıktı.

2.5.3.1. Ağırlıklı PageRank

Ağırlıklı PageRank, kullanıcı değerlendirme notları, güven puanları veya duygu gücü gibi belirli özelliklere bağlı olarak grafin kenarlarına farklı ağırlıklar uygular. Bu ayarlama, diğer

ilişkileri ağda daha derine gömdüğü için sıralama sürecine yardımcı olur. Örneğin, daha yüksek duygu puanı veya değerlendirme notuya sahip bir öneri sisteminde yorum, bir ögenin sıralanması sırasında daha fazla ağırlığa sahip olabilir [77].

2.5.3.2. Kişiselleştirilmiş PageRank

Kişiselleştirilmiş PageRank, kişiselleştirilmiş rastgele hareketleri ayarlayabilir ve sıralama sonuçlarını belirli kullanıcılara veya gruplara yansıtabilir. Algoritma daha sonra eşit olasılıkla rastgele bir düğüm seçmez, bunun yerine etki alanı tarafından belirlenen bir düğüm seçimi yapar (yani, kullanıcı geçmişine karşılık gelen düğümler). Bu yöntem, özellikle öneri sistemlerinde olduğu gibi kullanıcı başına öge sıralamaları üreten modeller için uygundur [78].

2.5.3.3. Konuya Duyarlı PageRank

Topikal bilgi içeren PageRank, konuya duyarlı PageRank olarak adlandırılır. Bu algoritma, konuya göre bölümlendirerek düğümleri belirli konular ve kullanıcı ilgi alanları için daha yüksek alaka ile puanlar. Bu varyant, belirli kategorilere veya konulara yönelik kullanıcı önyargısını dikkate alması gereken öneri sistemleri için yararlıdır [79].

2.5.3.4. Dinamik PageRank

Dinamik PageRank, grafin zamanla değişen doğasını hesaba katar ve bu da onu ağın her zaman statik olmadığı uygulamalar için iyi bir model haline getirir. Bu yöntem, PageRank'i uyarlanabilir olduğu gerçeğine dayanarak hesaplar; bu PageRank'in tam hesaplamasını sıfırdan çalıştırmaz, bunun yerine yalnızca bağlantı değişikliklerine ve ağ yapısına göre halihazırda var olan puanları uyarlar [80].

2.5.4. PageRank Kullanımında Karşılaşılan Zorluklar ve Dikkat Edilmesi Gerekenler

PageRank güçlü bir algoritma olmasına rağmen, etkili bir uygulama için ele alınması gereken zorlukları da beraberinde getirir:

Hesaplama Karmaşıklığı: Algoritmanın yinelemeli yapısı, özellikle büyük ölçekli ağlar için hesaplama açısından yoğun olabilir. Yakınsamayı hızlandırmak için genellikle optimizasyonlar ve paralel hesaplama teknikleri kullanılır [81].

Sönümlenme Faktörü Hassasiyeti: Sönümlenme (Damping) faktörünün seçimi sonuçları önemli ölçüde etkileyebilir. Kötü seçilmiş bir faktör, optimum olmayan sıralamalara veya daha yavaş yakınsamaya yol açabilir [82].

Çıkma Sokaklar ve Örümcek Tuzaklarıyla Başa Çıkma: Çıkış kenarı olmayan düğümler (çıkma sokaklar) veya kapalı bir döngü oluşturan düğüm grupları (örümcek tuzakları) PageRank hesaplamasında sorunlara neden olabilir. Bağlantı budama ve kenar yönlendirme gibi teknikler bu sorunları hafifletmek için kullanılır [83].

2.5.5. Modern Sistemlere PageRank'in Entegrasyonu

Mevcut öneri sistemleri ve birçok veri analizi uygulaması, en iyi sonucu elde etmek için diğer algoritmalar ve veri işleme stratejileriyle zenginleştirilmiş bir PageRank çeşidi kullanır. Örneğin, PageRank'i ABSA ile birleştirebiliriz, böylece algoritmamız düğümleri yalnızca PageRank puanına göre değil, aynı zamanda duygu odaklı ağırlıklara göre de sıralar. Başarılı olursa, bu hibrit yöntemler duygu farkındalığına sahip ve bağlam özelinde önerilerin önemli olduğu uygulamalarda faydalı olur [84].

Ayrıca, GNN'ler, düğümler ile topolojik ve özelliklere duyarlı tasarımları ilişkilendirerek PageRank'i iyileştirmek için kullanılmıştır. Düğüm niteliklerini ve bunların graf-teorik önemini hesaba katarak daha zengin bir öneri kümesi sağlar [85].

Literatür, yapılan ilerlemeleri vurgulayan ancak aynı zamanda hala yapılması gereken çalışmalarını da gösteren öneri sistemleri, HIN'ler, kelime gömmesi yöntemleri ve duygu analizi hakkında kapsamlı bir genel bakış sağlamıştır. Bu çalışmalar kullanıcı etkileşimleri, duygu ve ağ yapılarının kullanımının en genel geçerliliğini kanıtlaya da, veri seyrekliğini, Cold-Start sorunlarını ve kişiselleştirilmiş önerileri kapsayacak entegre metodolojiler hala önemlidir. Bu temel keşiflerin üstüne, aşağıdaki metodoloji bölümünde yaklaşımımızı başlangıçta iki yeni yöntem geliştirerek uygulayarak açıklıyoruz: Yorumdan Etkilenen Heterojen Bilgi Ağı Tabanlı Öneri Sistemi (REHREC) ve bunun daha da gelişmiş bir

versiyonu olan W-REHREC. Bu yaklaşımlar, iyi temellendirilmiş esnek bir öneri modeli için kullanıcı yorumlarını, duygu analizini ve meta-yol tabanlı yöntemleri entegre eder.

3. YÖNTEM

3.1. Giriş

Etkili öneri sistemi geliştirme, yalnızca kullanıcı-öge işbirlikli bilgilerini değil aynı zamanda değerlendirme notları, öge özelliklerini, bağlamları ve yarı yapılandırılmış kaynakları da tanımlayan karmaşık ağ yapılarına dayanır. Heterojen bilgi ağları (HIN'ler), bu senaryolarda çok konulu verileri modellemek ve analiz etmek için etkili bir tekniktir. Yorumdan Etkilenen Heterojen Bilgi Ağı Tabanlı Öneri Sistemi System (REHREC) ve onun ağırlıklı uzantısı Ağırlıklı Yorumdan Etkilenen Heterojen Bilgi Ağı Tabanlı Öneri Sistemi System (W-REHREC) olarak adlandırılan iki tamamlayıcı yöntem öneriyoruz. Bunlar kullanıcı yorumlarını meta-yol tabanlı olanlarla birleştirerek öneri sistemlerinin genel performansını iyileştirir.

REHREC ve ağırlıklı varyantı (W-REHREC), HIN'lerin kavramlarını, kullanıcı yorumlarını öneri üzerinde bir etki ölçüsü olarak içerecek şekilde genişletir. Veri seyrekliği, Cold-Start, daha kişiselleştirilmiş ve bağlam farkında bir öneriye ihtiyaç duyma gibi iyi bilinen zorlukları ele alırlar. Bu yöntemler, duygu analizini ağ yapısına yerleştirir ve kullanıcı geri bildirimine ağırlık getirir, bu da kullanıcıların yorumlarını ağırlıklı toplama yoluyla önerilere dönüştürmeye olanak tanır.

3.2. HIN Tabanlı Sistemlerde Kullanıcı Yorumlarının Önemi

Kullanıcı yorumları gibi yapılandırılmamış veriler, kullanıcıların beğenileri ve beğenmediklerine derinlemesine odaklanma sağlayabilir. Birçok geleneksel öneri sistemi, yalnızca sayısal değerlendirme notları veya belirli bir öge ile kaç kez etkileşim kurulduğunu dikkate aldıkları için kullanıcı geri bildirimlerinden gelen nitel bilgileri genellikle göz ardı eder. Ancak bu yöntem, kullanıcı yorumlarından önemli bağlamı yakalayamaz. Bunun yerine, REHREC ve W-REHREC, modellerine yerleştirilmiş metinsel verileri en iyi şekilde kullanarak etkili kullanıcı duygusunu ve tercihlerini çıkarmak için yorumları bir HIN'e dahil eder. Konu-Bazlı Duygu Analizi (ABSA) ile yorumlar, kullanıcı duygularının ayrıntılı bir şekilde anlaşılmasına yardımcı olan bir duygu puanı ile birlikte konulara ayrıştırılır.

REHREC, kullanıcı geri bildirim düğümleri oluşturarak ve bunları sırasıyla kullanıcı düğümlerine ve öge düğümlerine eşleyerek bu zenginleştirilmiş yorumları ağa aktarır. Bu yapılandırma, bu nedenle yorumlardaki paylaşılan duygulardan kaynaklanan kullanıcılar ve

öğeler arasındaki meta-yolların çıkarılmasını sağlar. Kullanıcı-Yorum-Öge veya Kullanıcı-Yorum-Yorum-Kullanıcı gibi meta-yollar, öğeler hakkında fikir paylaşan kullanıcıları açığa çıkarmak ve daha iyi öneriler için örtük benzerlikleri belirtmek için kullanılabilir.

3.3. Meta-yollar ve Önerileri Geliştirmedeki Roller

Meta-yollar, iki farklı düğüm türü arasındaki korelasyonu temsil eder ve hem REHREC hem de W-REHREC'te önemli bir rol oynarlar. Bu rotalar, sistemin, bir kullanıcı ile bir öge arasındaki doğrudan olmayan ancak çeşitli araçları içeren ilişkileri ifade eden çoklu-atlamalı ilişkileri keşfetmesini sağlar. Örneğin, REHREC modelinde bulunan diğer meta-yolları göz önünde bulundurarak, Kullanıcı-Yorum-Öge-Yorum-Kullanıcı, kullanıcıların ortak yorumlar ve duygular perspektifinden birbirleriyle nasıl bağlantı kurduğuna ışık tutabilir.

W-REHREC modeli, ağırlıklı meta-yolları tanıtarak bu kavramı iletir. Bu ağırlıklar, meta-yolun bir parçasını oluşturan yorumların duygu yoğunluğu tarafından belirlenir. Olumlu ve yüksek puanlı yorumlar daha büyük ağırlıklara katkıda bulunur ve daha güçlü, daha etkili bağlantıları işaret eder. Bu yaklaşım yalnızca tarafsız ve etkili yorumlar arasında ayırım yapmakla kalmaz, aynı zamanda öneri sisteminin kullanıcı tercihlerini daha iyi yansıtan yollara öncelik vermesine de olanak tanır.

3.4. Cold-Start ve Veri Seyrekliği Sorunlarının Ele Alınması

Cold-Start, yeni kullanıcılar veya öğeler hakkında yeterli veri toplanamadığında işbirlikçi filtreleme ve içerik tabanlı öneri sistemlerinde karşılaşılan bir diğer büyük sorundur. Bu sorunu ele almak için REHREC ve W-REHREC, kullanıcı yorumlarından duygu bilgilerini kullanır. Kullanıcı-öge etkileşimleri aracılığıyla açık bağlantı her zaman gerçekleştirilemediğinden, kullanıcı tarafından oluşturulan yorum bilgilerinden örtük bağlantılar ortaya çıkarabiliriz.

REHREC'te bu işlem, yeni kullanıcıları (veya öğeleri) bir yorum puanına göre aynı duyguyu paylaşır paylaşmaz arka plan bilgileriyle mevcut düğümlere bağlayarak yapılır. Bir örnek vermek gerekirse, yeni bir kullanıcı olumlu duyguya sahip bir restoran hakkında bir yorum

yazarsa, bu şekilde model bu yeni kullanıcıyı yorumlarında benzer değeri önceliklendirmiş kullanıcılara atayabilir ve bu çok az etkileşime göre öğeler önerebilir.

W-REHREC daha sonra bu bağlantıları duygu büyüklüğüne göre ağırlıklandırarak bunu bir adım öteye taşır. Bu, bu tür güçlü görüşlere sahip yorumların önerileri tahmin ederken en önemli şey olmasını sağlar ve böylece sistemin bunu kullanıcılarla yüksek oranda uyumlu içeriğe göre tartmasına olanak tanır. W-REHREC yalnızca veri seyrekliğini çözmekle kalmaz, aynı zamanda önemli yorumların etkisine vurgu yaparak Cold-Start'ta önerilerin doğruluğunu da artırır.

3.5. Duygu Analizi ve HIN Kelime Gömmelerinin Birleştirilmesi

Hem REHREC hem de Ağırlıklı REHREC, ağ içindeki karmaşık ilişkileri yakalamak için HIN kelime gömmelerinin gücünden yararlanır. Bu amaçla, düğüm kelime gömmesi yaklaşımları, bir grafın potansiyel olarak yüksek boyutlu öznitelik yapısını, düğümler ve benzerliğin bir alt akış görevinin parçası olarak hesaplanması gereken durumlar içindeki önemli ilişkileri koruyabilen düşük boyutlu sürekli bir vektör gösterimine dönüştürür [6]. REHREC'te, düğüm kelime gömmeleri, kullanıcı yorumlarının düğüm benzerliklerini hesaplamak için kenar oluşturmada rol oynamasına izin veren duyguları içerir.

Öte yandan W-REHREC modeli duygu ağırlıklı yolları kelime gömmesine dahil etmeyi göz önüne alır, yani kelime gömmesi sürecinde kullanıcı görüşlerinin önemini dahil etmeyi ele alır. Bu şekilde, düğümleri yalnızca doğrudan bağlantılarına değil, aynı zamanda konuya-özel duygu verileri aracılığıyla bağlantıların gücüne ve duygusuna göre kelime gömmesi işlemini verimli bir şekilde gerçekleştirilebilir. Hem kullanıcıyı hem de öğeyi aynı anda ortak bir gizli (latent) özellik alanına yerleştirerek, bu genel yaklaşım öneri sisteminin yalnızca açık kullanıcı tercihlerini değil aynı zamanda örtük olanları da yakalamasını sağlayarak önerilerin genel kalitesini artırır.

3.6. Gelişmiş Öneriler için Birleşik Bir Model

REHREC ve W-REHREC, ortak temelleri ve farklı yenilikleri nedeniyle tamamlayıcılık sergiler. REHREC, kullanıcı tarafından oluşturulan yorumları kullanmak ve duygu tabanlı bağlantıları tespit etmek için esnek bir model görevi görür, ancak W-REHREC ile, gerçek değerli duygu yoğunluğundan ağırlıklar olarak başka bir işlem çıktısı katmanı ekleyerek katkıda bulunuruz ve daha ince öneri mekanizmalarına olanak tanırız. Bu metodolojilerin

birleşimi, hedeflenen, bağlamsallaştırılmış öneriler sağlamak için kullanıcı tarafından oluşturulan içeriğin zenginliğini gelişmiş ağ analiziyle bütünleştiren bütünsel bir yaklaşım üretir.

Aşağıdaki bölümlerde, hem REHREC hem de W-REHREC'in tam olarak nasıl uygulandığına dair daha fazla tartışma sunacağız; böylece her yöntemin, veri seyrekliği ve kişiselleştirme zorluklarıyla aynı anda etkili bir şekilde yüzleşmek için daha derin, duyguya duyarlı bir öneri sistemi oluşturmak üzere nasıl tasarlandığını göstereceğiz.

REHREC ve W-REHREC'in temel kavramlarını ve hedeflerini aklımızda tuttuğumuzda, bunların perde arkasında nasıl işlediğini tanımlayan bir model oluşturmamız gerekir. Kullanıcılar, öğeler ve farklı duygu verileri arasındaki bağlantıları içeren bir yorum grafi, her iki modelin de merkezinde yer alır. Bu, yalnızca kullanıcının duygudaki incelikleri görüşlere geri bildirimde bulunmasını sağlamakla kalmaz, aynı zamanda karmaşık bağlantıların temeli olarak da işlev görür. Bir sonraki bölümde, bu yorum grafinin nasıl oluşturulacağını ele alarak, belirli bir tür öneri modeli oluşturan her adımı ve yöntemi açıklıyoruz.

3.7. Yorum Grafi Oluşturma

Yorum grafini oluşturma süreci, iki REHREC ve W-REHREC modelini oluşturmanın önemli bir parçasıdır. Bir sonraki adım, kullanıcı-öge etkileşimlerini, öge özelliklerini ve daha sonra incelenecek ve öneri oluşturmada kullanılacak duygu verilerini içerecek modeli oluşturmaktır. Sonraki birkaç bölüm, yorum graflarının bu iki yöntem için nasıl oluşturulduğunu açıklar ve benzerlikleri ve iyileştirmeleri belirtir.

3.7.1. REHREC'te Yorum Grafi Oluşturma

Yorumdaki Etkilenen Heterojen Bilgi Ağı Tabanlı Öneri Sistemi (REHREC), kullanıcı tarafından oluşturulan içerik ve ilişkiyi birleştiren yorum grafi oluşturmak için daha ayrıntılı bir yaklaşım izler. Bu son grafi, yorum bilgileriyle zenginleştirilmiş kullanıcı-öge-e etkileşimlerindeki gizli bağlantıları araştırmak istiyoruz. Bu, aşağıdaki gibi birkaç önemli adımı içerir.

Düğüm Gösterimi: Heterojen bilgi ağlarında aşağıdaki düğümler bulunur.

- Kullanıcı Düğümleri (U): Sistemdeki her kullanıcı, etkileşimleri ve ilişkili meta veriler (örneğin kullanıcı kimliği, etkileşim geçmişi) hakkında bilgi içeren bir düğüm olarak temsil edilir.
- Öge Düğümleri (I): Öge kategorisi, değerlendirme notu geçmişi ve ürün/hizmet gibi kullanıcıların etkileşimde bulunduğu diğer önemli özellikleri içeren ayrı düğümler olarak temsil edilir.
- Yorum Düğümleri (R): Kullanıcı yorumları farklı düğümler olarak ele alınır ve kullanıcıları ögelere bağlar. Yorum yaklaşımı, kullanıcıların görüşlerinin ögelerin bu belirli konulara bağlandığı anlamına gelen Konu-Bazlı Duygu Analizi (ABSA) kullanılarak öge yorumlarıyla ilişkili duygudan oluşur.

Doc2Vec Kullanarak Özellikleri Çıkarma: Yorumlar, daha fazla işleme veya modelleme için uygun hale getirilmesi gereken metinsel bilgiler içerir ve bu, içeriğin dilsel bağlamını yakalayan Doc2Vec adı verilen bir kelime gömmesi tekniği kullanılarak yapılır. Yorum metinlerini işlemek için sürekli bir kelime torbası (CBOW) yaklaşımı kullanan iki katmanlı bir sinir ağına sahip Doc2Vec modelidir. Bu yaklaşım, temel anlamsal bilgileri korumayı garanti eder. Model eğitildikten sonra, her yorum için daha sonra sınıflandırma işlemleri için özellik vektörleri olarak kullanılacak bir vektör gösterimi oluşturur.

Duygu Sınıflandırması: Her bir yorum olumlu, olumsuz veya nötr olarak sınıflandırılabilir ve ikili sınıflandırma problemi olarak ele alınır. Eğitim verileri 1 ile 5 arasında puanlarla işaretlenmiş yorumlardan oluşur. 3'ten düşük puanlı yorumlar olumsuz olarak sınıflandırılır ve 3 veya üzeri puanlı yorumlar olumlu olarak sınıflandırılır. Daha sonra, yorumların vektörleştirilmiş gösterimlerini olumlu veya olumsuz duygu olarak kategorize etmek için, doğru sınıflandırmaları elde etmek amacıyla önce bir SVM sınıflandırıcısı kullanırız.

Kenar Yapısı: Hetrojen bilgi ağlarında aşağıdaki kenar türleri bulunur.

- Kullanıcı Yorum Kenarı (U-R): Bir kullanıcı düğümü ile onun yazdığı yorum düğümleri arasındaki kenarlar. Bu bağlantılar, yorumun nereden geldiğini gösterir ve sistemin her bir görüşü tek tek kullanıcıya kadar izlemesini sağlar.
- Yorum-Öge Kenarları (R-I): Yorum ve öge düğümleri arasındaki kenarlar, burada yorumun hedefi bir ögedir. Bu tür bir bağlantı, ağın kullanıcı duygularını belirli ögelerle ilişkilendirmesine olanak tanır.

- Kullanıcı-Öge Kenarları (U-I): Ayrıca, bir kullanıcının açık yorum verilerinde ögeye atıfta bulunmadan bir ögeyle etkileşime girdiği durumu karşılamak için doğrudan kullanıcı ve öge düğümleri olan kullanıcı-öge kenarlarımız vardır.
- Yorum-Yorum Kenarları (R-R): Bu kenarlar anlamsal olarak benzer olan ve aynı duygu sınıfına sahip olan yorumlar arasında oluşturulur. Doc2Vec tarafından üretilen vektör gösterimleri kullanılarak, bu, benzer anlamsal özelliklere sahip yorumları eşleştirmek için kullanılır.

Meta-yol Entegrasyonu: REHREC, ağ içi analitik yetenekleri için meta-yol kavramını kullanır. Örneğin, Kullanıcı-Öge-Yorum-Yorum-Öge-Kullanıcı (U-I-R-R-I-U) gibi bir meta-yol, kullanıcıların verilen öğelere benzer görüşler gösterdiği bağlantılar oluşturur. URRU: Bu, kullanıcılar hakkında, yorum düzenlerine göre diğer yarı benzer kullanıcılara bilgi toplayan bir diğer önemli meta-yoldur (işbirlikçi filtreleme işlevi için).

3.7.2. W-REHREC'te Yorum Grafi Oluşturma

Ağırlıklı Yorumdan Etkilenen Heterojen Bilgi Ağı Tabanlı Öneri Sistemi System (W-REHREC), daha yüksek karmaşıklıkta bir yorum grafi tanımlama yolunda çalışmayı genişletir. REHREC modelinin üzerine inşa ederek, kenarlara duygu yoğunluklarına göre ağırlıklar atarlar. Bu iyileştirme, sistemin bazı bağlantıları diğerine tercih etmesini ve böylece öneri doğruluğunu artırmasını sağlar.

Ağırlıklı Düğüm Gösterimi:

- Kullanıcı düğümleri (U): REHREC ile aynı şekilde, kullanıcı düğümü gerçek dünyadaki bir müşteriye karşılık gelir; meta veriler, profilleri ve etkileşim geçmişleri hakkında özellikleri tanımlayabilir.
- Öge Düğümleri (I): Öğeler grafin çıkarılamayan bir parçasını oluşturur ve daha iyi önerilerde bulunmak için açıklayıcı ve tanımlayıcı nitelikler içeren düğümler içerir.
- Yorum Düğümleri (R): İkinci düğüm türü, bizim durumumuzda W-REHREC'te ek duygu ağırlıklandırma bilgisi içeren yorum düğümleridir. Her düğüm yalnızca konu tabanlı duygu puanlarını değil, aynı zamanda yorumun graf geçiş sürecini ne kadar etkileyeceğine karar veren benzersiz duygu yoğunluğunu da içerir.

BERT tabanlı ABSA (Konu-Bazlı Duygu Analizi) modeli, kullanıcı yorumlarından duygu ve anlam çıkarmak için kullanılıyor. Yorum metinlerini ayrıştırıyor ve ardından restoranın

yemek, servis, ambiyans ve maliyet gibi her önemli konusuna karşı duygu puanı sağlıyor. Bu tür zengin duygu açıklamaları, yorum grafının son derece ve oldukça alakalı öneriler üretmek için anahtar olan çok zengin kullanıcı görüşlerini ifade etmesine yardımcı oluyor.

Geliştirilmiş Kenar Yapısı:

- Yorum-Yorum Kenarları (RR): W-REHREC modelindeki yeniliklerden biri, konu tabanlı duygulardaki eşleşmelere göre yorum düğümleri arasında kenarlar oluşturmaktır. Örneğin, iki yorum bazı konular hakkındaki duygularında senkronize ise, bu iki yorum arasında bir kenar oluşur. Kenara, kaç özelliğin hizalandığına bağlı olarak ağırlık verilir.

Her Çift Yorumun Yorum Düğümleri Arasındaki Kenar Ağırlığı Hesaplaması: Genel olarak, her çift yorum düğümü arasındaki kenar ağırlığı hesaplaması aşağıda ifade edildiği şekildedir:

Aynı Konuda Eşleşen Her Duygu İçin +1: Her iki yorumda aynı konudan olumlu veya olumsuz bir duyguyla bahsediyorsa (örneğin "iyi yemek kalitesi" +1 verir).

Bağlantısız Duygular için 0: Bir konu hakkındaki duygu yorumları arasında farklılık gösteriyorsa ek bir ağırlık atanmaz.

Paylaşılmayan Konular Arasında Kenar Yok: Farklı konularla ilişkili iki yorum arasında kenar yoktur.

Örneğin, aşağıdaki değerlendirmeleri göz önünde bulundurun:

- Yorum 1: “Restoran düşük fiyatlar ile harika yemekler sunuyor, ancak ambiyansın iyileştirilmesi gerekiyor.”
- Yorum 2: “Buradaki yemekler birinci sınıf ve uygun fiyatlı, ayrıca hoş bir ambiyansa sahip.”

Bu iki yorum arasındaki kenar ağırlığı şu şekilde hesaplanır:

- Yiyecek: Her iki yorumda da olumlu duygu → +1
- Fiyat: Her iki yorumda da olumlu duygu → +1
- Ortam: Farklı duygular (biri olumlu, biri olumsuz) → 0

Toplam kenar ağırlığı = +2. Bu yaklaşım, grafın yorumlar arasındaki nüanslı ilişkileri yansıtmasını, karşıt duyguları hesaba katarken paylaşılan bakış açılarını vurgulamasını sağlar.

Kenar Ağırlıklarını Normalleştirme: Tüm kenar ağırlıklarınızı hesapladıktan sonra, 0 ile 1 arasında son ağırlıklar vermek için normalleştirme uygulanır. Bu adımın önemi, ağırlıklı kenarlara göre Random Walk algoritmasını hesaba katacak olması ve daha sonra bu kelime gömmeleri üretim için kullanacak olmasıdır.

HIN Entegrasyonu: Ağırlıklı yorum grafi oluşturulduktan sonra, temel Heterojen Bilgi Ağı (HIN) ile entegre edilir. Bu tür bir entegrasyon daha sonra meta-yol tabanlı graf kelime gömmesi süreçleri için kullanılacak ve daha doğru önerilerle sonuçlanacak ve ayrıca öneride bulunan iki tarafın duygusunu da alacaktır.

3.7.3. REHREC ve W-REHREC'te Yorum Graflarının Birleşik Faydaları

Özellikle, hem REHREC hem de W-REHREC, duygu verilerinden yorum grafi oluşturmayı, onu güçlü bir HIN yeteneğine entegre etmek için kullanır. Bu modellerde, şunlar için yorum grafları oluşturmayı kolaylaştırırlar:

Gelişmiş Kullanıcı Temsili: Yorumlar düğümler aracılığıyla kullanıcılar ve öğelerle bağlandığından, bu tür yöntemler kullanıcının daha iyi temsil edilmesine yol açar.

Duyguya dayalı kişiselleştirilmiş öneriler: Konuya dayalı duygu analizi, kullanıcıların belirli seçimleri hakkında içgörüler sağlayacak ve tercihlerine göre daha kişiselleştirilmiş önerilerde bulunmalarına olanak tanıyacaktır.

Daha İyi Kişiselleştirme: Ağırlıklı REHREC durumunda, yüksek yoğunluklu ağırlıklara sahip yorumlar önem taşıdığından, çok sayıda güçlü ağırlıklı yorum yazılarak belirli kullanıcı tercihleri karşılanabilir.

Bu tür veriler, yalnızca veri seyrekliği ve Cold-Start sorunlarının ele alınmasında mevcut yöntemlerden çok daha iyi olmakla kalmayacak, aynı zamanda bu kapsamlı yorum graflarından gelen ilişkileri kullanarak son derece kişiselleştirilmiş, duyguya duyarlı öneriler de sağlayabilecek daha gelişmiş öneri sistemlerine yaratabilir.

Yorum grafi öncelikle kullanıcı tarafından oluşturulan duyguyu ve onun anlamsal ilişkilerini yakaladığından, bu yapıyı daha geniş bir Heterojen Bilgi Ağı'na (HIN) entegre ediyoruz. Daha sofistike öneri teknikleri için ayrıntılı bir ağ oluşturmak üzere kullanıcılar, öğeler ve yorumlar gibi çeşitli düğümleri birbirine bağlayan yorum grafindan HIN oluşturulması. Aşağıda, REHREC ve W-REHREC modellerimizde duygu verilerini ve kullanıcı

etkileşimlerini temsil etmek için HIN'in oluşturulmasını açıklıyoruz, böylece onlar öneri sistemlerinde kelime gömmesi kaynakları olarak kararlı bir şekilde hizmet edebilirler.

3.8. HIN Ağ Üretimi

Heterojen Bilgi Ağı (HIN) oluşturma adımı hem REHREC hem de W-REHREC modellerini oluşturmada önemli bir adımdır. Bu aşamada, düğümler ve kenarlar kullanıcıların sistemle nasıl etkileşime girdiğine, öğelerin özelliklerine ve yorum verilerine göre anlamlı bir şekilde düzenlenir [14]. Sonraki alt bölümler, okuyucunun benzersiz niteliklerinden bazılarını gözlemleyebilmesi ve ortak süreci geri alabilmesi için hem REHREC hem de W-REHREC için HIN ağ oluşturma yöntemini kısaca açıklar.

3.8.1. REHREC'te HIN Ağı Oluşturma

Yorundan Etkilenen Heterojen Bilgi Ağı Tabanlı Öneri Sisteminde (REHREC), HIN oluşturma süreci duygu bilgisini kullanmak için kullanıcı, öge ve yorum düğümleri arasında alana-özü heterojen bir bilgi ağı oluşturur. Bu tür bir yaklaşım kullanıcı-öge ilişkisini açıklayabilir ve böylece daha fazla bağlam farkında önerilere izin verebilir. REHREC için HIN oluşturma adımları aşağıdaki gibidir:

Düğüm Oluşturma:

Kullanıcı Düğümleri (U): Bu düğümler ilişkili kullanıcı meta verileriyle (kullanıcı kimliği ve etkileşim geçmişi) birlikte bireysel kullanıcıları temsil eder. Kullanıcı düğümleri, tercih ve davranış analizinin temelini oluşturur.

Öge Düğümleri (I): Bir olay grafında, her kullanıcı eylemine karşılık gelen bir öge düğümü (I) vardır. Öge düğümlerin kategori, değerlendirme notu geçmişi ve ilgili öznitelikleri gibi özellikler, kullanıcı düğümlerini ve yorum düğümlerine bağlamak için kullanılır.

Yorum Düğümleri (R): REHREC modelinin yeni bir özelliği, kullanıcı tarafından oluşturulan içeriği yansıtan yorum düğümlerin tasarımıdır. Bu yorum düğümleri duygu ve kullanıcı-öge etkileşim bilgilerinin analitik olarak araştırılabileceği üçlü bir yapıyı oluşturan kullanıcılara ve öğelere bağlıdır.

Kenar Yapısı:

- Kullanıcı Öge Kenarları (U-I): Kullanıcı düğümleri kullanıcıların satın alma veya görüntüleme gibi öğelerle etkileşime girmesi arasındaki doğrudan etkiyi gösteren öge düğümlerine doğrudan bağlıdır.
- Kullanıcı Yorum Kenarları (U-R): Kullanıcıları yazdıkları yorumlara bağlar ve her yorumun nereden geldiğini açıkça gösterir.
- Yorum-Öge Kenarları (R-I): Yorumları öğelere bağlar ve her yorumun hangi öge hakkında olduğunu gösterir. Bu nedenle, kullanıcı duygularını bağlamak için kullanılabilir.

Ek Duygu Bilgileri: İnceleme düğümlerine eklenen incelemelerden duygu puanları elde etmek için Konu Tabanlı Duygu Analizi (ABSA) kullanılır. Duygu verileri kullanıcı kutupluluğuna dair ek bir bakış açısı sunar ve düz etkileşim verilerine dayanarak aksi takdirde fark edilmeyecek ilişkileri keşfetmeye yardımcı olabilir.

Meta-yol Kullanımı: REHREC'teki HIN, sistemin içinde gizli karmaşık düğüm ilişkilerini ortaya çıkaran meta-yollara dayanır. Örneğin, Kullanıcı-Öge-Yorum-Yorum-Öge-Kullanıcı (UIRRIU) ve Kullanıcı-Yorum-Yorum-Kullanıcı (URRU) gibi meta-yollar, kullanıcıları benzer zevklere veya ortak deneyimlere sahip olan kişilerle birleştirir. Bu yollar, doğrudan kullanıcı-öge etkileşiminin ötesine geçen ilişkiler oluşturmaya yardımcı olur ve böylece işbirlikçi filtrelemeyi iyileştirir.

3.8.2. W-REHREC'de HIN Ağı Üretimi

W-REHREC, kullanıcı duygusunun gücünü ve duygusal doğasını daha iyi göstermek için her ilişkiye doğru ağırlıklı bağlantıları birleştirerek temel yapı temelinde HIN'i genişletir. Bu yöntem, duygunun gücüne dayalı olarak öneriyi daha yüksek öncelikli ilişkilere daralttı. W-REHREC için HIN oluşturma süreci aşağıdaki gibi özetlenmiştir:

Düğüm Oluşturma:

Kullanıcı Düğümleri (U): REHREC'e benzer şekilde, kullanıcı düğümleri meta veri görevi görür ve kullanıcıların etkinliklerine açılan kapıdır.

Öge Düğümleri (I): Bu düğümler geleneksel öneri modellerinde olduğu gibi öğeleri temsil etmek ve bunların niteliklerini ayrıntılı olarak açıklamak için kullanılır.

Yorum düğümleri (R): W-REHREC'te, yorum düğümleri yalnızca duygu bilgisini değil aynı zamanda duygu yoğunluğunu da içerir. Bu bilgileri dahil ederek, ağ bir kullanıcının ne kadar

güçlü bir görüşe sahip olduğunu gösterebilir, bu da ağırlıklı kenarlar oluşturmak için önemlidir.

Ağırlıklı Kenar Yapısı:

U-I: Ağırlıklı Kullanıcı-Öge Kenarları (bu kenarlar örneğin yorum duygu verilerinin toplamını tutabilir, vb.) Aynı ögeyi birden fazla kez inceleyen ancak her zaman olumlu duyguya sahip bir kullanıcı söz konusu olduğunda, kenar ağırlığı kullanıcı ile öge arasındaki ilişki gücünü yansıtmalıdır.

Ağırlıklı Kullanıcı-Yorum Kenarları (U-R): Bunlar, bir kullanıcıyı yorumlarına bağlamada REHREC'e benzerdir ancak yorumun duygu yoğunluğuna göre ağırlık taşır. Daha güçlü görüşlere sahip yorumlar (olumlu veya olumsuz) kullanıcı düğümü üzerinde daha yüksek bir etkiye sahip olacaktır.

Ağırlıklı Yorum-Öge Kenarları (R-I): Yorum düğümleri ile öge düğümleri arasındaki bağlantılar, yorumun duygusuna göre ağırlıklandırılır. Güçlü görüşlü yorumlara karşılık gelen kenarlar daha yüksek ağırlıklara sahip olacak ve bu da öneriler söz konusu olduğunda daha fazla görüş bildiren yorumların daha önemli olduğunu gösterecektir.

Ağırlıklı Yorum-Yorum Kenarları (R-R): Yorum-Yorum Kenarları (görüntü tabanlı) duygu eşleşmelerine göre oluşturulur. “Yemek kalitesi” veya “hizmet” gibi bir konu hakkında aynı görüşe sahip olan yorumlar, bu kenarlar aracılığıyla birbirine bağlanacaktır. Bu kenarların ağırlıkları, konu eşleşme sayıları ve duygu güçleri kullanılarak belirlenir ve bu da yorumlar arasındaki korelasyonu anlamaya yardımcı olur.

Duygu analizi ve ağırlıklandırma mekanizması: BERT tabanlı bir ABSA modeli, kullanıcı yorumlarında ayrıntılı duygu analizi yapmak için kullanılır. Her yorum için konuları çıkarır ve duygu puanları ve yoğunlukları atar. Sistem, bu duygu yoğunluklarını ağdaki kenarlara ağırlık vermek için kullanır, böylece hangi yorumların daha yüksek düzeyde etkiye sahip olduğunu açıkça görürsünüz.

Ağırlık Normalizasyonu ve Entegrasyonu: Kenar ağırlığı daha sonra tanımlanmış bir aralığa uyacak şekilde normalleştirilir (örneğin, [0,1]). Bu normalizasyon, bu ağırlıkların küresel HIN'e enjekte edilmesine olanak tanır ve bunu sonraki işlem adımlarıyla (örneğin Random Walk algoritması, kelime gömmesi oluşturma) uyumlu hale getirir.

Ağırlıklı Meta-yol Tabanlı REHREC: W-REHREC modeli öneri sürecinde ağırlıklı meta-yolları kullanır. Güçlü, olumlu geri bildirimlere sahip öğelere dikkat eder, buna bir örnek

Kullanıcı-Yorum-Öge (URI) meta-yoludur. Alternatif olarak, Kullanıcı-Yorum-Yorum-Kullanıcı (URRU) gibi diğer meta-yollar, kullanıcıların ortak yüksek yoğunluklu yorumlarına göre benzerliklerini vurgular. Sistem, kişiselleştirilmiş önerilerde bulunurken daha güçlü ilişkilere odaklanmasını sağlayan ağırlıklı meta-yolları öğrenir; böylece bunların alakalarını zenginleştirir.

3.8.3. REHREC ve W-REHREC'de HIN Ağ Üretiminin Birleşik Faydaları

Kullanıcı etkileşimlerinin ve duygularının karmaşıklıklarını etkili bir şekilde yakalamak için hem REHREC hem de W-REHREC'teki HIN oluşturma süreçleri aşağıdaki gibidir:

Gelişmiş İlişki Haritalaması: İki model, kullanıcılar, öğeler veya yorumlar arasındaki doğrudan veya dolaylı ilişkileri bulmayı kolaylaştıran karmaşık bir ağ yapısı oluşturur.

Tercih Metrikleri: Duygu verilerinin (ve W-REHREC durumunda duygu yoğunluğunun) dahil edilmesiyle, bu ağlar kullanıcıların neyi beğenip neyi beğenmediği konusunda daha fazla farkındalığa sahip olur.

Meta-yol Kullanımı: Meta-yolların benimsenmesi, doğrudan bağlantıların ötesinde olan karmaşık geçiş stratejilerini kolaylaştırır, daha yüksek düzeyde analiz ve kişiselleştirilmiş önerilere olanak tanır.

Toplu olarak, bu HIN oluşturma metodolojileri, basit etkileşim geçmişinden öznel duygu analizine kadar kullanıcı girdisinin tüm konularını kullanan süreçlerin ve öneri algoritmalarının yerleştirilmesi için temel oluşturur.

Heterojen Bilgi Ağı (HIN) oluşturulduktan sonra, duygu bilgileriyle zenginleştirilmiş öğeler ve yorumlardan gelen ilişkilerle birlikte kullanıcı etkileşimlerini temsil eden bir diğer önemli görev, bunu makine öğrenimi modelleri için uygun bir biçime dönüştürmektir. Bu, HIN'deki karmaşık ilişkileri sürekli vektörler olarak ele alan kelime gömmesi mekanizması aracılığıyla yapılır. Bu kelime gömmeleri, kümeleme veya hatta öneri oluşturma gibi aşağı akış görevlerini iyileştirmeye yardımcı olabilecek ağ hakkında önemli yapısal ve anlamsal bilgiler içerir. Bu bölümün geri kalanında, HIN'in yerleştirilmesi hem REHREC hem de W-REHREC modelleri için açıklanmaktadır.

3.9. HIN Kelime gömmesi

Kelime gömmesi teknikleri, Heterojen Bilgi Ağı'nın (HIN) karmaşık yapısını makine öğrenimi algoritmaları tarafından kolayca işlenebilecek bir biçime dönüştürülmesinde önemli bir bileşendir. HIN kelime gömmesi sürecinde yer alan ilişkileri ve anlamsal anlamları sürekli vektör uzaylarına eşliyoruz. Bu alt bölümde REHREC modeli için kelime gömmesi sürecini aşağıda ayrıntılı olarak açıklıyoruz.

3.9.1 REHREC'e HIN Kelime gömmesi

REHREC'in kelime gömmesi gösterimi, ağı ilişkisel ve anlamsal özelliklerinin korunacağı şekilde her türlü düğümü (kullanıcılar, öğeler ve yorumlar) yerleştirir. Bu süreçteki adımlar az ama önemlidir.

Düğümün Vektör Uzayı Temsili: Bir $G=(V,E)$ grafi olan bir HIN'deki her düğüm (kullanıcılar, öğeler ve yorumlar), ilişkilerini göz önünde bulundurarak bir vektör olarak temsil edilir. Burada, düğümleri meta-yollardaki yakınlıklarına ve komşuluklarına göre kelime gömmesine çevirmek için Node2Vec veya Metapath2Vec gibi yöntemler kullanılır [1].

$$f(v): V \rightarrow R^d \quad (2)$$

Bu, V 'nin düğüm kümesi ve d 'nin kelime gömmesi alanının boyutu olduğu $f(v)$ eşleme fonksiyonunu öğrenmek istediğimiz anlamına gelir.

Kelime gömmesi, kullanıcı-yorum-öge (URI) ve URRU gibi çoklu-atlamalı ilişkilerin meta-yollar aracılığıyla keşfedilebildiği yollar gibi Meta-yolları takip eder. REHREC, düğümleri bu iki taraflı graflar boyunca yollar üzerine yerleştirerek kullanıcı tercihlerini, paylaşılan görüşleri ve öge benzerliklerini vektör gösterimlerine korur.

Aşağıdaki meta-yol (mp), v_1, \dots, v_{t+1} düğümlerin T_1, \dots, T_{t+1} düğüm tipleri arasındaki R_1, \dots, R_t ilişkilerini açıklar :

$$mp: T_1 \xrightarrow{R_1} \dots T_k \xrightarrow{R_k} T_{k+1} \dots \xrightarrow{R_t} T_{t+1} \quad (3)$$

Bir meta-yol boyunca düğümden düğüme geçiş şu şekilde ifade edilebilir:

$$P(v_{k+1}|v_k) = \begin{cases} \frac{1}{|N(v_k)|} & (v_k, v_{k+1}) \in E \wedge \varphi(v_{k+1}) = T_{k+1} \\ 0 & otherwise \end{cases} \quad (4)$$

Yukarıdaki formülde dağılım olasılığı P hesaplanır; burada v_k ve v_{k+1} , mp meta-yoluna ilişkin k ve $k+1$ inci düğümleri T_k ve T_{k+1} türleri ile temsil eder. $N(v_k)$, T_{k+1} tipi ile v_k düğümüne bağlanan ilk kenarda bulunan komşular kümesini temsil eder.

Denklem yapısına göre, yürüyüşün uzunluğunu ve başlangıç noktasını hesaba katarak bir dizi Random Walk algoritması oluşturabiliriz. Kelime gömmesi modeli, belirli meta-yollar boyunca Random Walk algoritması yoluyla düğüm dizilerini örnekleme yoluyla gözetimsiz bir şekilde eğitilir. Daha sonra, maksimum pencere boyutu içinde birlikte bulunma olasılığını en üst düzeye çıkaran kelimelerin vektör gösterimlerini öğrenmek üzere uyarlanmış bir Skip-Gram algoritması modeline girdi olarak verilir. Skip-Gram algoritması modelinin hedefi en üst düzeye çıkarmaya çalıştığı yer:

$$\max \sum_{(v,u) \in D} \log Pr(u|v) \quad (5)$$

Burada Skip-Gram algoritması modeli, merkez düğüm v verildiğinde, bir pencere boyutu w içinde komşu düğümleri bağlam gözleme olasılığını en üst düzeye çıkarmayı hedefler. D , Random Walk algoritmasından üretilen düğüm-bağlam çiftlerinin kümesidir. $Pr(u | v)$, merkez düğüm v verildiğinde bağlam düğümünün olasılığıdır.

Bir HIN kelime gömmesinin sonuçları, graftaki her düğüm için vektörler olarak temsil edilir ve doğrudan alt akış görevlerinde (kümeleme, sınıflandırma, benzerlik tabanlı öneri vb.) kullanılmaya hazırdır. Bu tür kelime gömmeleri anlamlıdır ve orijinal HIN'de kalan bağlamsal veya ilişkisel bilgileri kapsüller ve böylece öneri performansını iyileştirmeye çok yardımcı olur.

3.9.2. Ağırlıklı Yorumlarla Meta-Yol Odaklı HIN Kelime Gömmesi

REHREC ile karşılaştırıldığında, bir sonraki model W-REHREC, kelime gömmesi vektörleri oluştururken hem HIN'lerin hiyerarşik bilgi karakteristiğini hem de kullanıcı duygularının yoğunluğunu hesaba katarak kelime gömmesi gösterimini genişletir. Bu yöntem, sistemin öncelikle doğru ve kişiselleştirilmiş öneriler oluşturmak için daha fazla bağlam sağlayan sağlam, duygu odaklı bağlantılara odaklanmasını sağlar.

Temel REHREC modelinin aksine, W-REHREC, kullanıcı yorumlarından çıkarılan duygu gücüne bağlı olarak düğümler arasında ağırlıklı kenarlar içerir. Bu ağırlıkları, Random Walk algoritması sırasında örnekleme yaparken kenar ağırlıklarını dahil etmek için standart kelime

gömmesi yaklaşımlarını değiştirerek ağa yerleştiririz. Bu, kelime gömmeleride daha güçlü hale getirmek için daha yüksek duygu yoğunluklarına sahip yollardan geçmeyi teşvik eder. Daha sonra ağırlıklı bir meta yol üzerinden düğümden düğüme geçiş olasılığı şu şekilde güncellenir:

$$P(r_j \rightarrow r_i) = \frac{w_i}{\sum_j w_j} \quad (6)$$

Burada r_j Random Walk algoritmasındaki geçerli bir düğümdür ve r_i ise bir kenar aracılığıyla bağlı olan r_j 'nin komşu yorum düğümüdür. w_i kullanıcı r_j 'ya bağlı olan r_i yorumunun duygu tabanlı ağırlığını temsil eder. Bu, iki düğüm arasındaki ilişkinin gücünü, önemini veya sıklığını gösterir. Ve formülün sigma kısmı, r_j ile komşuları arasındaki tüm kenar ağırlıklarının toplamıdır $\{r_1, r_2, \dots\}$.

Ağırlıklı Meta-Yol Örneklemesi: Kelime gömmesi işlemi ağırlıklı meta-yollar tarafından konu edilir. Bu izlerdeki yollar, daha güçlü duygu hizalamalarına sahip yorumları barındırır ve bu da modelin benzer kullanıcı tercihlerini ve ilişkilerini yakalamasını sağlar. Bu kenarlar, Random Walk algoritması sırasında daha yüksek ağırlıklara sahip olacak ve yürüyüşünün bunları seçme olasılığını artıracak ve böylece duygu açısından zengin bağlantıların potansiyel katkılarını artıracaktır.

Değiştirilmiş Skip-Gram algoritması Modeli: W-REHREC'te öğrenmeyi kelime gömmesi için olan Skip-Gram algoritması modelinin bir çeşidini benimsiyoruz ve kenar ağırlıklarını dikkate alarak değiştiriyoruz. Ağırlıklı kelime gömmesi için amaç fonksiyonu şu şekilde yazılabilir:

$$\max \sum_{(v,u) \in D} w_{vu} \cdot \log Pr(u|v) \quad (7)$$

Bu formülde:

D: Random Walk algoritması ile üretilen düğüm-bağlam çiftlerinin kümesi.

w_{vu} : v düğümü ile u düğümü arasındaki kenarın ağırlığı, ilişkinin gücünü veya önemini yansıtır.

$Pr(u | v)$: Merkez düğüm v verildiğinde bağlam düğümü u 'nun koşullu olasılığı.

Bu nedenle, daha yüksek kümülatif ağırlıklar üreten yollar, kelime gömmesi üzerinde daha büyük bir etkiye sahip olacak ve duygu odaklı ilişkileri son vektör uzayında daha iyi yansıtacaktır.

Ağırlıklı kelime gömmesi sürecinin önemli bir parçası, duygu yoğunluğu etkisinin dengelenmesidir. Güçlü duygunun daha fazla sayılması gerekse de, diğer önemli bağlantıları da bastırmamalıdır. Bu ağırlıklı yolların temsile orantılı bir şekilde katkıda bulunduğundan emin olmak için normalleştirme tekniklerini uyguluyoruz ve ağır somutlaştırılmış küresel dengesini koruyoruz. Normalleştirme, ağırlıkları yorum grafi kenarlarına ayarladığımız kenar ağırlıklandırma adımıyla yapılır, ağırlıklar 0 ile 1 arasında olacak şekilde normalleştirilir.

Meta-yol odaklı ağırlıklı kelime gömmesi, kullanıcı yorumlarına duygu yerleştirir. Meta-yol odaklı, ağırlıklı kelime gömmesinin çıktısı, yalnızca ağdaki yapısal ilişkileri değil aynı zamanda kullanıcı yorumlarından açık ve örtük içeriği yakalayan başka bir boyutu da yansıtan bir düğüm vektörü kümesi olarak sunulur. Vektörler, kullanıcıların en uç görüşlere ve beğenilere sahip olduğu rotaları güçlendirerek kişiselleştirilmiş önerileri belirlemek için öneri sistemini etkili bir şekilde iyileştirir.

3.9.3. HIN Kelime Gömmesi Yaklaşımlarının Birleşik Faydaları

REHREC ve W-REHREC'teki iki kelime gömmesi işlemi, zengin anlamı yakalamak için meta-yolların yeteneğinden yararlanır. Genel olarak, W-REHREC'in ağırlıklı konu yapısı, iki dizi yapısal olarak benzer olduğunda duygu yoğunluğunu da dikkate aldığı için daha yüksek düzeyde kişiselleştirme sunar. Bu kelime gömmeleri, kişiselleştirilmiş, bağlam duyarlı önerileri destekleyen sonraki makine öğrenimi görevlerinde kullanılan girdi vektörleri olarak işlev görür.

3.10. Yorum Ağırlıklandırması için Konu-Bazlı Duygu Analizi (ABSA)

ABSA, kullanıcıların mekanlar hakkında ürettiği yorumların nasıl değerlendirilmesi veya kullanılması gerektiği konusunda ayrıntılı bilgi edinmemizi ve ölçmemizi sağladığı için W-REHREC modelinin önemli bir bileşenidir. Konu-Bazlı Duygu Analizi (ABSA) aracılığıyla, ürün veya hizmetlerin çeşitli konuları ile ilişkili ayrıntılı bir görüş puanı elde edebiliriz, ABSA, yorum düğümlerin ve kenarlarının bir miktar ağırlığını sağlamak için hem düğümlere hem de kenarlara ek ayrıntılar ekleyerek HIN'i zenginleştirir. Bu şekilde, kullanıcılardan gelen geri bildirim derinliğinden yararlanmak, sistem tarafından üretilen önerilerin kişiselleştirilmesini ve doğruluğunu artırır. Bu bölümde, ABSA için kullanılan yöntemleri ve W-REHREC modelindeki uygulamalarını avantajlarıyla birlikte açıklıyoruz.

3.10.1. ABSA Metodolojisine Genel Bakış

Konu-Bazlı Duygu Analizi (ABSA), kullanıcı yorumlarını farklı konulara ayırmayı ve ifade edilen bir duygunun her bir konu için olumlu mu yoksa olumsuz mu olduğunu belirlemeyi içerir. Bunun aksine, küresel duygu analizi tüm bir metin gövdesinin kutupluluğuna dair genel bir bakış sağlar ve genellikle çok daha az ayrıntılıdır. ABSA, belirli niteliklere (örneğin yemek kalitesi, servis, fiyatlandırma ve ambiyans vb.) daha derinlemesine dalmanıza olanak tanır. Bu nedenle, bu derinlemesine düzey analizi duygunun nasıl yakalandığı ve ağa nasıl dahil edilebileceği ve daha iyi öneriler için kenar ağırlıklarının nasıl geçirilebileceği ile ilgili çok daha derin içgörüler sağlar.

ABSA, kullanıcı yorumlarında bulunan hedef konuların çıkarılmasıyla başlar. W-REHREC modeli bağlamında, konular incelenen öğelere karşılık gelen önceden tanımlanmış kategorilerdir, örneğin ürün özellikleri/hizmet nitelikleri. Bu çıkarma, sözdizimsel ayrıştırma ve konu göstergelerini tanımlamak üzere eğitilmiş geleneksel makine öğrenimi modellerine dayanır.

Konular çıkarıldıktan sonra, farklı konulara bağlı bir dizi duyguyu sınıflandırmak bir sonraki aşamadır. Bu, duygunun olumlu, olumsuz veya nötr olup olmadığını bulmayı içerir. Özellikle, W-REHREC modeli, her yorumun karmaşık temsilini öğrenmek için BERT tabanlı (derin öğrenme tabanlı) bir ABSA modeli kullanır. Herhangi bir verilen girdi için, duygu sınıflandırması -1 ile +1 (dahil) arasında bir sayı sağlar; burada -1'i çok olumsuz, 0'ı nötr ve +1'i çok olumlu olarak kabul ederiz.

Konu çıkarma ve duygu sınıflandırma adımından sonra, her yorumu konu-duygu çiftlerine karşılık gelen bir vektör olarak düşünebiliriz. Örneğin, buna bir örnek olarak şu şekilde bir restoran yorumu verilebilir:

Yorum Vektörü: [Yemek: +1, Servis: -1, Ortam: +1, Fiyatlar: +1]

Bu, sisteme, yorumdaki çeşitli bileşenlere göre uygulanacak duygu konusunun ve ağırlığının yoğunlaşmasını sağlar.

3.10.2. ABSA'nın Ağırlıklı Yorum Graflarına Entegrasyonu

Bu konu tabanlı duygu verileri, yorum grafinin ağırlıklı kenarlarını elde etmek için kullanılabilir. Bu entegrasyonun adımları şunlardır:

Ağırlıklı Kenarların Oluşumu: Yorum düğümleri birlikte aynı konuyu paylaştıklarında ağırlıklı bir kenar aracılığıyla birbirine bağlanır; burada, duygu puanları birbirleriyle karşılaştırılır. İki yorum aynı konuyu paylaşıyorsa ve aynı görüşe sahipse, aralarındaki kenara pozitif bir ağırlık ekleriz. Öte yandan, yorumlar aynı konuda farklı görüşe sahip ise negatif ağırlık ekleriz veya aynı konudan bahsetmiyorlarsa 0 ekleriz. Burada, iki yorum düğümü arasındaki ve eşleşen konu duyguları arasındaki ağırlığı türetmek için aşağıdaki formülü kullanırız.

$$w_{ij} = \begin{cases} \frac{1}{|A|} \sum_{a \in A} s_i(a) \cdot s_j(a), & \text{eğer } \exists a \in A, i \text{ ve } j \text{ yorumları ortak bir konuya sahipler ise} \\ 0, & \text{aksi takdirde} \end{cases} \quad (8)$$

Bu formülde:

w_{ij} : Yorum düğümleri i ve j arasındaki kenarın ağırlığı.

A : İki yorum arasındaki ortak konuların kümesi.

$s_i(a)$ ve $s_j(a)$: Sırasıyla i ve j incelemesindeki a yönü için duygu puanlarıdır. Bu puanlar, olumlu duygu için +1, olumsuz duygu için -1 veya duygu ifade edilmediğinde 0 gibi normleştirilmiş değerler olabilir.

$|A|$: i ve j yorumları arasındaki paylaşılan konuların toplam sayısı.

Koşullar: Paylaşılan konu yoksa ($A = \emptyset$), w_{ij} ağırlığı 0 olarak ayarlanır.

Ağırlık Normalizasyonu: Hesaplanan ağırlıklar, kenar ağırlıklarını düşük aralıkta tutmak ve kelime gömmesi işlemini eşit oranda etkilemek için normleştirilmelidir. Aşağıdaki min-max normalizasyon formülünü kullanırız:

$$w_{ij}^{norm} = \frac{w_{ij} - w_{min}}{w_{max} - w_{min}} \quad (9)$$

Bu formülde:

w_{ij} : i ve j düğümleri arasındaki kenarın orijinal ağırlığı.

w_{min} : Tüm ağırlık kümesindeki en düşük ağırlık değeri.

w_{max} : Tüm ağırlık kümesindeki maksimum ağırlık değeri.

w_{ij} norm : Formül uygulandıktan sonra $[0,1]$ aralığında olması garanti edilen normalize edilmiş ağırlık.

Eğer $w_{\max} = w_{\min}$ ise (tüm ağırlıklar aynıysa), w_{ij} norm = 0,5 olarak ayarlanır. Bu, sıfıra bölmeyi önler.

Yorumların ortak bir konusu olmadığında, ilgili düğümler arasında hiçbir kenar eklenmez. Bu nedenle, ağırlıklı bütünlüğünü koruyan öneri çalışırken, ilgisiz veya çelişen yorumların birbirlerini etkilemesine izin verilmez.

Örnek Senaryo:

İki restoran yorumunu ele alalım, her biri farklı bir kullanıcıya ait:

- Yorum 1: “Restoranda yemek ve servis iyi, ancak iç mekan daha iyi olabilir”
- Yorum 2: “Servis hızlıydı, fiyatlar yüksekti ve bu dünyanın dışında bir yemek deneyimiydi. Ortam keyifliydi.”

Bu yorumlara ABSA'yı uygulamak aşağıdaki konu-duygu çiftlerini verebilir:

- Yorum 1: [(yemek, +1), (hizmet, +1), (ortam, -1)]
- Yorum 2: [(yemek, +1), (hizmet, +1), (fiyatlandırma, -1), (ortam, +1)]

Ortak nitelikler için kenar ağırlığını hesaplamak için aşağıdaki yöntem kullanılır:

- Yemek: Her ikisi de olumlu $\rightarrow +1$
- Hizmet: Her ikisi de olumlu $\rightarrow +1$
- Ortam: Farklı duygular $\rightarrow -1$

Bu, +1'lik toplam normalleştirilmemiş kenar ağırlığını verir ve daha sonra normalleştirilip HIN'de kullanılır.

3.10.3. Yorum Ağırlıklandırması için ABSA Kullanmanın Faydaları

Burada ABSA ile yorum ağırlıklandırması olan W-REHREC modelinin bazı faydalarını listeleyeceğiz:

Daha İyi Öneriler: Genelleştirilmiş duygular yerine, model, kullanıcı için hangi özellik konu ünün yararlı olduğuna dayanarak önerilerde bulunabilir.

Öneri Doğruluğunda İyileşme: Daha ayrıntılı duygu analizi, benzer genel puanlara sahip ancak belirli konularda farklı görüşlere sahip yorumları ayırt edebilmemiz sayesinde çok daha iyi önerilerde bulunmayı mümkün kılıyor.

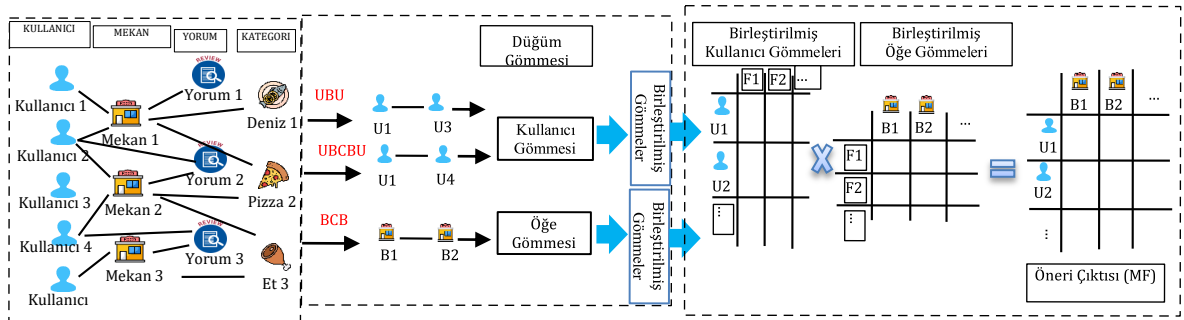
Bağlam-farkında eşleştirmeler: Bu konu tabanlı yöntem, modelin yalnızca bağlama özgü nitelikteki yorumlar arasındaki ilişkileri öğrenmesini ve paylaşılan kullanıcı bakış açılarından faydalanmasını sağlar.

W-REHREC modelinde ABSA kullanımı, yorum bilgilerini, heterojen bilgi ağı (HIN) inşası üzerinde ağırlıklı kenar somutlaşmasını daha da kolaylaştıran bağlamsal içgörülerin bir tutkalına dönüştürür. Bu tür bir konu düzeyinde analiz, kullanıcı tercihi hakkında içgörü sağlar ve hedeflenen, duyguya duyarlı önerilere olanak tanır. Duygu gücünü ve konu eşleşmesini birleştirerek ABSA, modelin yalnızca önemli ilişkilere odaklanmasını sağlar ve bu da daha iyi kişiselleştirilmiş önerilerle sonuçlanır.

3.11. Değerlendirme notu Tahmin Fonksiyonu ve Öneri Çıktısı

Heterojen bilgi ağlarından faydalı bilgiler, meta-yollar kullanılarak kullanıcı tipi yerleştirmeler ve iş tipi yerleştirmeler olarak çıkarılır. Çıkarılan bu bilgiler, bir öneri sistemi geliştirilmesinin temelini oluşturur.

Bu aşamada, bu yerleştirmeler, öneri sistemi içinde değerlendirme notu tahminleri oluşturmak için birleştirilir. HERec modelinin [14] uzantıları olan önerilen REHREC ve ayrıca W-REHREC modelleri, Matris çarpanlarına ayırma (MF) yöntemini kullanır. Şekil 3.1'de gösterildiği gibi, heterojen ağdan meta-yol tabanlı Random Walk yaklaşımı kullanarak düğüm Gömmeleri türetiliyoruz. Aynı düğüm tipiyle başlayan birden fazla meta-



Şekil 3.1 Yaklaşımın iş akışını gösteren diyagram

yol, düğüm gömme aşamasında gösterildiği gibi tek bir düğüm için birden fazla Gömmenin oluşturulmasıyla sonuçlanır. Farklı meta-yollar tarafından oluşturulan benzersiz bir tek

düğüm gömmesi elde etmek için, gömme iş akışının son aşaması olarak bir birleştirme adımı tanıtılır. Bu birleştirilmiş gömmeler daha sonra öneri görevini gerçekleştirmek için matris çarpanlarına ayırmada kullanılır.

Matris Çarpanlarına Ayırmada, kullanıcının ilgi noktasına (POI) verilen değerlendirme notu (burada bir POI, mekan (business) olarak geçiyor) kullanıcı matrisi ve POI (mekan) matrisini kullanılarak elde edilir. Bu düşük boyutlu matrislere gizli faktör matrisleri adını verebiliriz. Bu iki matrisin satır endeksleri, gizli faktörlerdeki kullanıcılar ve POI'ler arasındaki ilişkiyi gösterir. Belirli bir kullanıcının (u) bir mekan (b) üzerindeki POI değerlendirme notuna ait matematiksel ifade aşağıdaki gibidir,

$$Rating_{u,b} = U_u^T \cdot V_b \quad (10)$$

burada U_u kullanıcının gizli faktörünü ve V_b POI ile ilgili gizli faktörü gösterir. Öte yandan, HERec [29] modeli, MF'ye kullanıcı ve öğenin gömülmesini ekler. Değerlendirme notların tahmini, REHREC ve W-REHREC algoritmalarında aşağıdaki işlev kullanılarak yapılır.

$$Rating_{u,b} = U_u^T \cdot V_b + \lambda \cdot e_u^{(U)T} \cdot \Omega_u^{(U)} + \mu \cdot \Omega_b^{(I)T} \cdot e_b^{(I)} \quad (11)$$

Denklem içinde, Ω_u kullanıcı yerleştirmesine ait ağırlığı gösterir ve Ω_b mekan yerleştirmesine ait ağırlığı gösterir. Ayrıca, formülde λ ve μ katsayıları yerleştirmelerden matris çarpanlarına ayırma işlevine entegrasyon rolünü oynar.

Öneri, 0 ile 5 arasında bir sayı olan öngörülen değerlendirme notlarına dayanarak yapılır. Bu amaçla bir eşik değeri koyarız. 3 ve üzeri olarak öngörülen değerlendirme notları kullanıcıya önerilir ve diğerleri elenir.

4. DENEYLER VE DEĞERLENDİRME

4.1. Genel Bakış

Bu bölümde hem REHREC hem de W-REHREC'in etkinliği ve gerçek dünya uygulanabilirliği hakkında derinlemesine bir değerlendirme sunmayı amaçlıyoruz. Bu ilgili ve kişiselleştirilmiş öneriler sunma yeteneklerini göstermeyi ve ayrıca diğer algoritmalara göre rekabet avantajlarını gösterecek olan temel ve son teknoloji yöntemlerle karşılaştırma yapmayı içerir. Mevcut yöntemlerden daha iyi performans gösteren yeni yöntemleri belirlemek için, öneri etkinliğini değerlendirmek üzere Kök Ortalama Kare Hatası (RMSE) gibi iyi bilinen metriklerden yararlanıyoruz.

Bir diğer önemli odak noktası, farklı büyüklük ve karmaşıklığındaki veri kümeleri üzerinde test edilen modellerin ölçeklenebilirliği ve verimliliğidir. Bu, modellerin artan kullanıcı-öğe etkileşimlerini ve kapsamlı yorum verilerini ne kadar etkili bir şekilde ele aldığını değerlendirmeyi içerir. Konu-Bazlı Duygu Analizi (ABSA), ağırlıklı yorum grafları ve meta-yol odaklı kelime gömmeleri gibi temel bileşenlerin bireysel katkıları da etki analizi yoluyla incelenir. Her bir bileşenin modellerin genel performansını artırmadaki önemini izole etmek ve anlamak için ablasyon çalışmaları yürütülür.

Bu bölüm, teorik geliştirme ile pratik uygulama arasındaki boşluğu kapatmayı, önerilen modellerin yalnızca kontrollü ortamlarda mükemmel olmasını değil, aynı zamanda gerçek dünya senaryolarına da etkili bir şekilde uyum sağlamasını sağlamayı amaçlamaktadır. Bu deneylerden elde edilen sonuçlar, modellerin değerine dair kritik kanıtlar sağlayacak ve gelişmiş öneri sistemlerinde gelecekteki araştırmalar için yol açacaktır.

4.2. Veri kümeleri

Önerilen REHREC ve W-REHREC modellerinin değerlendirmesi iki farklı veri kümesi üzerinde gerçekleştirilir: Yelp Veri Kümesi ve Foursquare Veri Kümesi. Her veri kümesi, öneri sistemi performansının farklı konularını vurgulayan benzersiz özellikler sunarak önerilen metodolojiler için kapsamlı bir test ortamı sağlar.

Metinsel yorum verilerindeki zenginliğiyle bilinen Yelp Veri Kümesi, Heterojen Bilgi Ağlarında Konu-Bazlı Duygu Analizi (ABSA) ve yorum odaklı meta-yolların entegrasyonunu doğrulamak için ideal bir ortam görevi görür. Veri kümesi, anlamsal bilgilerle zenginleştirilmiş sağlam bir HIN'in oluşturulmasını sağlayan kapsamlı kullanıcı

tarafından oluşturulmuş yorumlar, kullanıcı profilleri ve mekan nitelikleri içerir. Konum tabanlı sosyal ağ verilerinden kaynaklanan Foursquare Veri Kümesi, kullanıcı girişlerine, ipuçlarına ve kullanıcı bağlantılarına odaklanır. Açık yorumlardan yoksun olsa da, metinsel bilgi olan ipuçları sağlar ve her ipucunu kullanıcı tercihleri, coğrafi konumlar ve sosyal etkileşimler hakkında değerli bilgiler içeren bir yorum olarak kullanırız.

Bu veri kümeleri birlikte, ayrıntılı yorum verilerinden yararlanma ve yorumdan esinlenen teknikleri genelleştirmeden örtük kullanıcı etkileşimlerine (Yelp ve Foursquare) kadar bir dizi öneri zorluğunu temsil eder. Bu veri kümelerinin dahil edilmesi, önerilen modellerin çeşitli koşullar altında test edilmesini, çok konuluğun, uyarlanabilirliğinin ve çeşitli gerçek dünya senaryolarında etkililiğinin doğrulanmasını sağlar.

4.2.1. Yelp Veri Kümesi

Yelp [-] Veri Kümesi, öneri sistemlerini değerlendirmek için zengin bilgi sağlayan kapsamlı, herkese açık bir veri kümesidir. Zengin kullanıcı etkileşimleri, ayrıntılı yorumlar ve çeşitli ürün kategorileri nedeniyle akademik ve endüstri araştırmalarında yaygın olarak benimsenmiştir. Bu bölüm, Yelp Veri Kümesinin, yapısının ve REHREC ve W-REHREC modellerine uygulanmasının derinlemesine bir analizini sağlar.

Yelp Veri Kümesi, restoranlar, barlar, mağazalar ve hizmet sağlayıcıları gibi çeşitli mekanları yorum için yaygın olarak kullanılan Yelp platformundan alınan verileri içerir. Veri kümesi birkaç temel bileşen içerir. Kullanıcı veri bölümü, anonimleştirilmiş kullanıcı kimlikleri de dahil olmak üzere kullanıcılar için benzersiz tanımlayıcılar sağlar. Ayrıca, kullanıcı profili bilgileri, her kullanıcı tarafından katkıda bulunulan yorum sayısı, ortalama değerlendirme notları ve sosyal graf dinamiklerini analiz etmek için kullanılabilen kullanıcı bağlantıları hakkında bilgiler gibi meta verileri de içerir.

İşletme verileri, benzersiz mekan kimlikleri, kategoriler ve fiyat aralığı, mevcut olanaklar ve açılış saatleri gibi özellikler de dahil olmak üzere mekanlar hakkında ayrıntılı bilgiler içerir. Ayrıca, şehir ve eyalet gibi coğrafi konum ayrıntıları, kullanıcı geri bildirimlerine dayalı toplu mekan değerlendirme notlarıyla birlikte sağlanır. Yorum verileri, her biri 1'den 5'e kadar bir ölçekte yıldız değerlendirme notusu ile birlikte gelen zengin bir kullanıcı tarafından oluşturulan yorum koleksiyonundan oluşur. Bu yorumlar, gönderim tarihini belirten zaman damgalı kayıtları ve genellikle kullanıcı görüşleri ve açıklamaları içeren ayrıntılı metinsel içeriği içerir.

Son olarak, giriş verileri bilgileri aracılığıyla kullanıcıların mekanlara yaptığı ziyaretlerin sıklığı ve zamanlaması hakkında içgörüler sağlar. Veri kümesi ayrıca, kullanıcıların belirli mekanlar hakkında sunduğu özlü öneriler veya tavsiyelerden oluşan bahşış verilerini de içerir.

Yelp Veri Kümesi, mekanlar, kullanıcılar, yorumlar, ipuçları ve girişler dahil olmak üzere birden fazla varlığı kapsayan zengin bir bilgi içerir. Bu veri kümesi JSON formatında yapılandırılmıştır ve aşağıdaki temel dosyaları içerir:

İşyeri Verileri: Çeşitli kategorilerdeki 144.000 mekanın ayrıntılarını içerir. İşyeri saatleri, park yeri müsaitliği, ambiyans, fiyat aralığı ve coğrafi konum (enlem ve boylam) gibi nitelikler dahildir.

Yorum Verileri: Kullanıcılar tarafından yazılmış 4,1 milyon yorum içerir. Her yorum bir kullanıcı kimliği, mekan kimliği, yıldız değerlendirme notu, yorum metni ve yayınlanma tarihini içerir.

Kullanıcı Verileri: Etkileşim geçmişleri, yorum sayıları ve sosyal bağlantıları (arkadaşlar) dahil olmak üzere 1 milyon kullanıcının profillerini yakalar. Ortalama yıldız değerlendirme notları ve alınan iltifat sayısı gibi ölçümleri içerir.

Giriş Verileri: 125.000 mekandeki girişleri toplar ve zaman içindeki ziyaret kalıplarına ilişkin içgörüler sağlar.

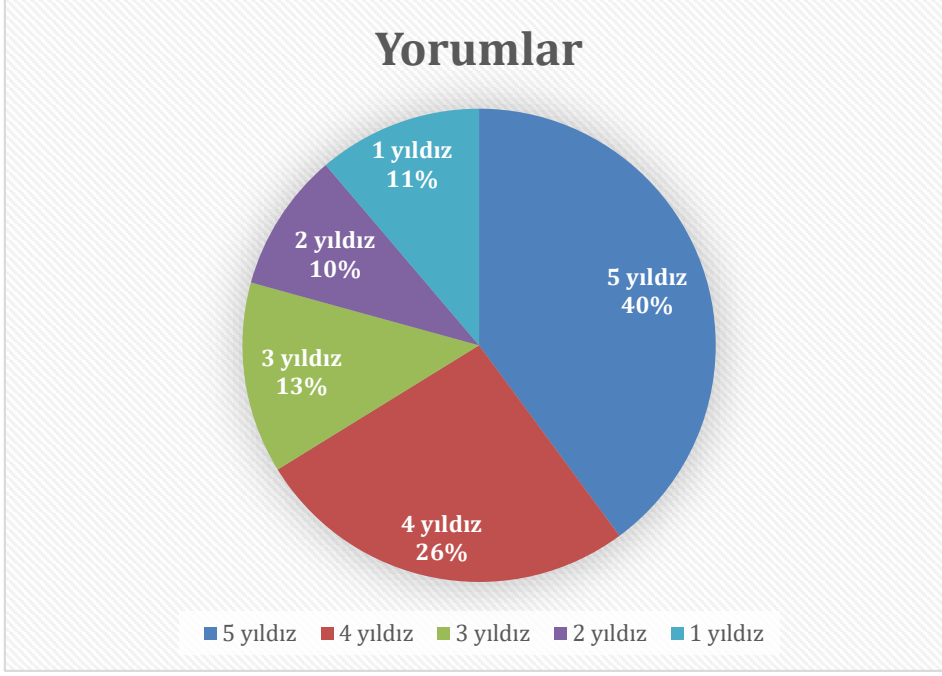
İpucu Verisi: İşletmeler hakkında 947.000 kullanıcı ipucunu içeren, özlü görüş ve tavsiyeleri özetleyen veridir.

Yıldız Değerlendirme notlarına Göre Yorum Sayıları:

Boyutu 2,13 GB olan yelp_academic_dataset_review.json dosyası, kullanıcı yorumlarını ve değerlendirme notlarını analiz etmeye yardımcı olmak için oylar, kullanıcı_id, yorum_id, yıldızlar, tarih, metin, tür ve mekan_id gibi verileri içerir.

Bu veri dosyası, 1'den 5 yıldıza kadar yıldız değerlendirme notlarına dağıtılmış yorumları içerir:

- 5 yıldız: 744.609 yorum (olumlu duygu).
- 4 yıldız: 491.297 yorum (orta derecede olumlu duygu).
- 3 yıldız: 244.067 yorum (nötr duygu).
- 2 yıldız: 177.509 yorum (orta derecede olumsuz duygu).
- 1 yıldız: 209.333 yorum (olumsuz duygu).

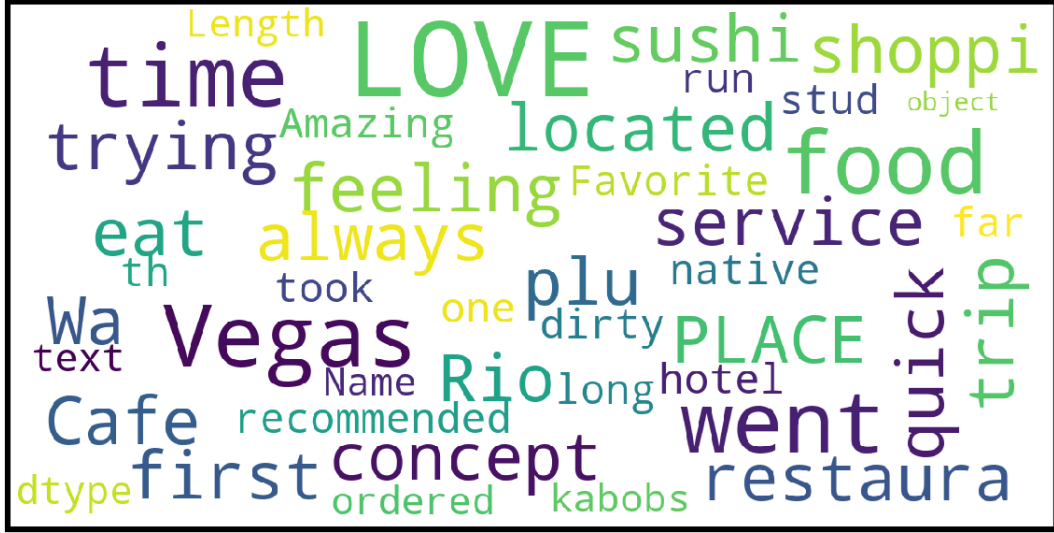


Şekil 4.1 YELP'te Yorumların dağılımı

Bu dağılımı görselleştiren Şekil 4.1'deki pasta grafi, yorumların çoğunluğunun olumlu (5 yıldız) olduğunu, en az sayıda yorum ise olumsuz puan (1 yıldız) verildiğini ortaya koymaktadır.

Metinsel Veri Analizi:

Şekil 4.2'de bulunan, yorum metninden oluşturulan bir kelime bulutu, kullanıcıların sıklıkla dile getirdiği popüler konuları ve duyguları yansıtan, yaygın olarak kullanılan terimlerin bir anlık görüntüsünü sunar.



Şekil 4.2 YELP'te yorumların örnek kelime bulutu

Metin yorumları genellikle, Konu-Bazlı Duygu Analizi nin merkezinde yer alan yiyecek, servis, ambiyans ve fiyatlandırma gibi konulara vurgu yapar.

Kullanıcı Katılımı:

- 273 MB boyutundaki yelp_academic_dataset_user.json dosyası, Yelp'teki kullanıcı davranışları ve etkileşimleri hakkında fikir vermek için yelping_başlama, oylar, yorum_sayısı, isim, kullanıcı_id, arkadaşlar, hayranlar, ortalama_değerlendirme notu , tür, övgüler ve elit statüsü gibi bilgileri içerir.
- Yüksek aktiviteli kullanıcılar veri kümesine önemli katkılarda bulundu; bazıları 1.000'den fazla yorum yazdı ve geniş arkadaş ağlarını korudu.
- Seçkin kullanıcılar (tutarlı, yüksek kaliteli katkılarıyla tanınanlar) daha yüksek düzeyde etkileşim ve etki göstermektedir.

Mekan Nitelikleri:

- 73,6 MB boyutundaki yelp_academic_dataset_business.json dosyası, Yelp'teki işletme bilgilerini ve müşteri puanlarını analiz etmeye yardımcı olmak için mekan_id, kategoriler, ad, tür, detaylı_adresi, nitelikler, mahalleler, şehir, saatler, review_count, boylam, enlem, açık durum, eyalet ve yıldızlar gibi verileri içerir.
- Veri kümesindeki mekanlar restoranlar, barlar ve hizmet sağlayıcılar gibi çeşitli kategorilere yayılıyor ve analiz için geniş bir alan sunuyor.
- Ortam, park yeri ve fiyat aralığı gibi özellikler, önerileri geliştirmek için bağlamsal bilgiler sağlar.

Giriş Bilgileri:

- 26,7 MB boyutundaki yelp_academic_dataset_checkin.json dosyası, Yelp'teki farklı işletmelerdeki kullanıcı girişlerini ve kalıplarını analiz etmek için giriş_bilgisi, gitiş tipi ve mekan_idsi gibi verileri içerir.
- Girişlerden elde edilen zamansal veriler, müşteri ziyaretlerindeki kalıpları ortaya çıkarır ve bu, yoğun iş saatlerini ve kullanıcı tercihlerini belirlemek için kullanılabilir.

Yelp Veri Kümesinin REHREC ve W-REHREC modellerinde kullanıma uygun hale getirilmesi için aşağıdaki ön işleme adımları uygulandı:

Veri Temizleme: Eksik metin veya değerlendirme notu içeren yorumlar gibi yinelenen ve eksik kayıtların kaldırılması.

Anlamli analizlerin yapılabilmesi için yeterli yoruma sahip olmayan mekanların filtrelenmesi.

Konu Çıkarımı ve Duygu Açıklaması: İnceleme metinlerinden belirli yönleri (örneğin, yiyecek kalitesi, servis, ambiyans) çıkarmak için Konu Tabanlı Duygu Analizi'nin (ABSA) uygulanması. Her bir konu için -1 (negatif) ile +1 (pozitif) arasında değişen değerlere sahip duygu puanlaması.

Normalizasyon: Değerlendirme notları ve inceleme sayıları gibi sayısal özelliklerin ortak bir ölçeğe göre standartlaştırılması. Kelimeleştirme, durdurma sözcüğü kaldırma ve köklendirme/lemmatizasyon yoluyla metin normalizasyonu.

Gömme Oluşturma: Metinsel incelemeler için anlamsal gömmeler oluşturmak amacıyla BERT veya Doc2Vec gibi tekniklerin kullanılması.

Kodlama: Meta veri kodlaması, mekan kategorileri gibi kategorik özellikler için kullanılır ve bunları makine öğrenimi modelleri için uygun bir biçime dönüştürür.

Yelp Veri Kümesi, çeşitli ve ayrıntılı verileri nedeniyle önerilen öneri modelleri için mükemmel bir test ortamı görevi görür. Önemli uygulamalardan biri, kullanıcı düğümlerin bireysel değerlendiricileri ve mekan düğümlerin değerlendirilen öğeleri temsil ettiği

Heterojen Bilgi Ağı'nın (HIN) oluşturulmasıdır. Değerlendirme düğümleri kullanıcıları mekanlara bağlayan ve ağı duygu verileriyle zenginleştiren araçlar olarak hizmet eder.

Bir diğer önemli uygulama ise konu tabanlı yorum ağırlıklandırmasıdır. Konu-Bazlı Duygu Analizi'nden (ABSA) elde edilen sonuçlar, kenarlara duygu odaklı ağırlıklar atamak için yorum grafinin entegre edilir. "Hizmet kalitesi" veya "fiyatlandırma" gibi belirli konular için duygu yoğunluğu, önerileri iyileştirmek ve doğruluklarını artırmak için kullanılır.

Veri kümesi ayrıca, nüanslı ilişkileri ortaya çıkarmak ve işbirlikçi filtrelemeyi geliştirmek için Kullanıcı-Yorum-Mekan ve Kullanıcı-Yorum-Yorum-Kullanıcı gibi meta-yolları kullanarak meta-yol analizini de destekler. Bu analiz, kullanıcılar ve öğeler arasındaki kalıpları ve tercihleri belirlemeye yardımcı olur.

Son olarak, milyonlarca yorum ve etkileşim içeren Yelp Veri Kümesinin büyük boyutu, modellerin hesaplama verimliliğini ve ölçeklenebilirliğini değerlendirme fırsatı sunar. Bu ölçeklenebilirlik testi, modellerin kapsamlı ve karmaşık verileri etkili bir şekilde işleyebilmesini sağlar.

Yelp Veri Kümesi, gelişmiş öneri sistemleri geliştirmek ve değerlendirmek için onu son derece değerli bir kaynak haline getiren çok sayıda güçlü konu sunar. Başlıca avantajlarından biri de verilerinin zenginliğidir. Veri kümesi, kullanıcıların deneyimleri hakkında derinlemesine görüşler sunduğu, genellikle etkileşimde buldukları mekanların belirli konularını vurgulayan ayrıntılı metinsel yorumlar içerir. Bu yorumların yanı sıra veri kümesi, mekan kategorileri, nitelikler (örneğin, fiyat aralığı ve olanaklar) ve kullanıcı demografisi gibi yapılandırılmış meta verileri içerir. Metinsel ve yapılandırılmış verilerin bu kombinasyonu, etkileşimlerdeki kullanıcı davranışlarını, tercihlerini ve kalıplarını analiz etmek için sağlam bir temel sağlar ve araştırmacıların müşteri deneyimleri hakkında değerli içgörüler elde etmesini sağlar.

Veri kümesinin bir diğer önemli gücü de çeşitliliğidir. Yelp Veri Kümesi, restoranlar, barlar, perakende mağazaları ve hizmet sağlayıcıları gibi çok çeşitli mekan kategorilerini kapsar. Bu çeşitlilik, birden fazla alanda genelleme yapılmasına olanak tanır ve bu veriler kullanılarak geliştirilen öneri modellerinin belirli bir sektörle sınırlı olmamasını, ancak çeşitli bağlamlara uyum sağlayabilmesini sağlar. Farklı endüstrilerden etkileşimleri dahil ederek, veri kümesi çok konulu ve heterojen verileri işleyebilen modellerin oluşturulmasını destekler.

Veri kümesinin gerçek dünyayla alakalı olması bir diğer kritik faktördür. Veriler doğrudan milyonlarca kullanıcı tarafından küresel olarak kullanılan Yelp platformundan kaynaklandığı için gerçek tüketicilerin otantik etkileşimlerini ve görüşlerini yansıtır. Bu gerçek dünya uygulanabilirliği, veri kümesinden türetilen tüm bulguların veya modellerin pratik, gerçek yaşam senaryolarına dayanmasını sağlar. Bu veri kümesinde eğitilen ve değerlendirilen modeller, gerçek dünya uygulamalarında dağıtıldıklarında etkili bir şekilde performans göstermek için daha iyi konumlandırılmıştır çünkü gerçek kullanıcı davranışlarını ve geri bildirim modellerini ele almaya zaten alışmışlardır.

Bu güçlü konuların birleşimi—zenginlik, çeşitlilik ve gerçek dünyayla alakalılık—Yelp Veri Kümesini öneri sistemlerinde, duygu analizinde ve kullanıcı tercihi modellemesinde yeni metodolojileri keşfetmek için ideal bir test ortamı haline getirir. Ayrıntılı ve çok konulu yapısı, kapsamlı analizleri destekleyebilmesini, inovasyonu konulandirebilmesini ve modern, veri odaklı uygulamaların taleplerini karşılayan sistemler geliştirmenin önünü açabilmesini sağlar.

Yelp Veri Kümesi zengin ve çeşitli olsa da, doğru ve tarafsız analiz sağlamak için ele alınması gereken çeşitli zorluklar ve sınırlamalar sunar. Dikkat çeken bir sorun, verilerdeki dengesizliktir. Restoranlar gibi belirli kategoriler, perakende veya hizmet sağlayıcılar gibi diğerlerine kıyasla önemli ölçüde fazla temsil edilmektedir. Bu dengesizlik sonuçları çarpıtabilir ve daha az temsil edilen alanlarda düşük performans gösterirken iyi temsil edilen kategorileri tercih eden öneri modellerinde önyargılara yol açabilir. Bu sınırlama, tüm kategorilerde adil temsili sağlamak için model eğitimi sırasında dikkatli örnekleme veya ağırlıklandırma tekniklerini gerektirir.

Bir diğer zorluk ise kullanıcı etkileşimlerinin seyrekliğidir. Veri kümesindeki bazı kullanıcılar çok aktifken, çok sayıda yorum ve değerlendirme notu ile katkıda bulunurken, kullanıcıların önemli bir kısmının etkileşimleri sınırlıdır. Bu seyreklik, anlamlı kalıpları belirlemek için yeterli miktarda kullanıcı-öge etkileşimine dayanan işbirlikçi filtreleme modelleri için zorluklar yaratır. Seyrek veriler, seyrek kullanıcılar için daha zayıf önerilere yol açabilir ve bu sorunu hafifletmek için matris çarpanlarına ayırma, öğrenme akatarılması (transfer learning) veya yardımcı verilerin dahil edilmesi gibi gelişmiş teknikler gerektirebilir.

Veri kümesi ayrıca önemli ön işleme zorlukları oluşturan gürültülü metinsel veriler içerir. Yorumlar genellikle resmi olmayan dil, argo, kısaltmalar ve ara sıra genel duyguya veya

konuya katkıda bulunmayan alakasız içerikler içerir. Örneğin, kullanıcılar konu dışı yorumlar, emojiler veya yazım hataları ekleyebilir ve bu da metin analizini karmaşıklştırabilir ve doğal dil işleme (NLP) modellerinin etkinliğini azaltabilir. Belirteçleme, durdurma sözcüğü kaldırma, yazım düzeltme ve duygu açıklaması gibi sağlam ön işleme adımları, metinsel verileri temizlemek ve bunların alt akış görevleri için uygunluğunu sağlamak için gereklidir.

Bu zorluklara rağmen, Yelp Veri Kümesi değerli bir kaynak olmaya devam ediyor. Ancak, sınırlamalarının üstesinden gelmek, sağlam ve doğru öneri sistemleri oluşturma potansiyelini tam olarak değerlendirmek için dikkatli veri işleme, ön işleme ve model optimizasyonu gerektirir. Bu sorunların ele alınması, veri kümesinden elde edilen içgörülerin yalnızca güvenilir değil, aynı zamanda çeşitli gerçek dünya senaryolarında da uygulanabilir olmasını sağlar.

Yelp Veri Kümesi, gelişmiş öneri sistemleri geliştirmek ve değerlendirmek için değerli bir kaynak sağlar. Zenginliği, çeşitliliği ve gerçek dünya uygulanabilirliği, onu REHREC ve W-REHREC modellerini test etmek için ideal hale getirir. Ayrıntılı yorumlarından, yapılandırılmış meta verilerinden ve kullanıcı-öge ilişkilerinden yararlanarak, bu modeller veri seyrekliği ve ölçeklenebilirlik gibi zorlukları ele alırken doğru ve kişiselleştirilmiş öneriler sunma yeteneklerini gösterebilir.

4.2.2. Foursquare Veri Kümesi

Foursquare Veri Kümesi, New York City'deki restoran mekanlarıyla kullanıcı etkileşimlerine odaklanan Foursquare'den alınan ayrıntılı bir veri koleksiyonudur. Girişler, ipuçları ve kullanıcı sosyal ağ verilerinin bir kombinasyonunu içerir ve bu da onu konum tabanlı sosyal ağlar (LBSN'ler) ve öneri sistemleri konusunda araştırma yapmak için mükemmel bir kaynak haline getirir. Bu bölüm, veri kümesinin ayrıntılı bir açıklamasını, uygulamalarını, istatistiksel özelliklerini ve ilişkili zorlukları sağlar.

Veri kümesi, restoran alanındaki kullanıcı etkinliğinin kapsamlı bir konusunu sunarak kullanıcı tercihlerinin ve sosyal dinamiklerin çeşitli konularını yakalar. Kullanıcı verileri, Foursquare platformundaki 3,112 kullanıcı için benzersiz tanımlayıcıların yanı sıra belirli mekanlar için yazılmış ipuçları gibi kullanıcı tarafından oluşturulan içerikle ilgili bilgileri içerir. Ayrıca kullanıcılar arasındaki sosyal ağ bağlantılarını da içerir ve işbirlikçi filtrelemeyi ve sosyal etki modellemesini destekleyen bir graf oluşturur. Mekan verileri,

mutfak türlerine ve restoran özelliklerine göre kategorize edilmiş New York City'deki 3,298 restoran mekanını kapsar. Bu veri kümesi mekan popülerliği ve kullanıcı değerlendirme notları gibi meta verileri içerir.

Giriş verileri, her biri bir zaman damgasına bağlı olan restoran mekanları genelindeki 27,149 kullanıcı girişini içerir. Sıralı giriş bilgilerini, zamansal kullanıcı davranışı ve konum tercihleri hakkında içgörüler sağlar. Bahşiş verileri, restoranlar için özlü yorumlar ve öneriler sunan ve İngilizce dilinde 10,377 kullanıcı tarafından oluşturulmuş veriler önemli ipuçları verir. Bu ipuçları belirli mekanlara bağlanır ve veri kümesini duygu ve tanımlayıcı içerikle zenginleştirir. Sosyal ağ verileri, kullanıcı arkadaşlıklarını temsil eder ve sosyal bağlantıları ve bunların kullanıcı davranışı üzerindeki etkilerini keşfetmek için kullanılabilir bir ağ oluşturur. Bu, akran etkilerinin ve sosyal gruplar içindeki tercihlerin yayılmasının incelenmesini sağlar.

Veri kümesi, 3,298 mekana dağılmış 27,149 girişi, mekan başına ortalama yaklaşık sekiz giriş ile yakalar. Kullanıcı etkinliği değişir; bazı kullanıcılar düzinelerce giriş yaparken diğerleri daha az etkileşime girer. Mekanlar, kaliteli yemek, rahat restoranlar ve fast food gibi türe göre kategorilere ayrılır ve analiz için çeşitli veriler sunar. Kafeler ve pizzacılar gibi popüler kategoriler daha fazla giriş alır ve kullanıcı tercihlerini vurgular. Zaman damgalarının analizi, öğle ve akşam yemeklerinde en yoğun etkinliğin yanı sıra hafta sonları artan girişleri ortaya koyar. Mevsimsel değişiklikler, tatillerden veya hava koşullarından etkilenen kullanıcı davranış değişikliklerini yansıtır.

Veri kümesi, ortalama bir kullanıcının birkaç arkadaşının olduğu bir graf oluşturan 3,112 kullanıcı için sosyal bağlantılar içerir. Sosyal ağ içindeki kümeleme, benzer giriş bilgilerine sahip kullanıcıların genellikle aynı alt gruplara ait olduğunu ve paylaşılan tercihleri gösterdiğini gösterir.

Veri kümesi, özellikle konum tabanlı öneri sistemlerinin geliştirilmesi ve değerlendirilmesinde çeşitli araştırma uygulamalarını destekler. Kullanıcı düğümlerin bireyleri, mekan düğümlerin restoranları ve kenarların girişleri veya sosyal bağlantıları gösterdiği heterojen bir bilgi ağının (HIN) oluşturulmasını sağlar. Zamansal ve mekansal nitelikler ağı geliştirerek karmaşık ilişki modellemesine olanak tanır. Kullanıcı tarafından oluşturulan ipuçları, yiyecek kalitesi, servis veya ambiyans gibi konu düzeyinde analiz edilebilen duygu verileri sağlar. İpuçlarından gelen duygu puanları, yorum grafiğindeki kenarlara ağırlık atayarak önerileri iyileştirmek için kullanılır.

Kullanıcı sosyal ağlarının dahil edilmesi, akran etkisinin öneri sürecine dahil edilmesine olanak tanır. Modeller, bir kullanıcının sosyal çevresinde popüler olan mekanları önermek için arkadaşlıklardan yararlanabilir. Sıralı giriş kalıpları, öğle saatlerinde öğle yemeği mekanları önermek gibi zaman tabanlı tercihleri hesaba katan modellerin geliştirilmesini sağlar. Veri kümesinin orta büyüklükte olması ve ayrıntılı yapısı, gelişmiş öneri algoritmalarının ölçeklenebilirliğini ve verimliliğini test etmek için uygun hale getirir.

Veri kümesi, onu araştırma için değerli bir kaynak haline getiren çeşitli güçlü konular sunar. Girişlerin, ipuçlarının ve sosyal ağ bağlantılarının birleşimi, kullanıcı davranışına dair çok konulu bir ortam sağlar. Restoran mekanlarına odaklanması, yiyecek ve içecek sektöründe hedefli araştırmalara olanak tanır ve diğer alanlara genelleştirilebilen içgörüler sunar. Zaman damgalarının ve sosyal bağlantıların dahil edilmesi, dinamik kullanıcı davranışının ve sosyal etkinin araştırılmasını sağlar. Foursquare gibi popüler bir platformdan doğrudan alınan veriler, veri kümesinin pratik öneri senaryolarına uygulanabilirliğini garanti eder.

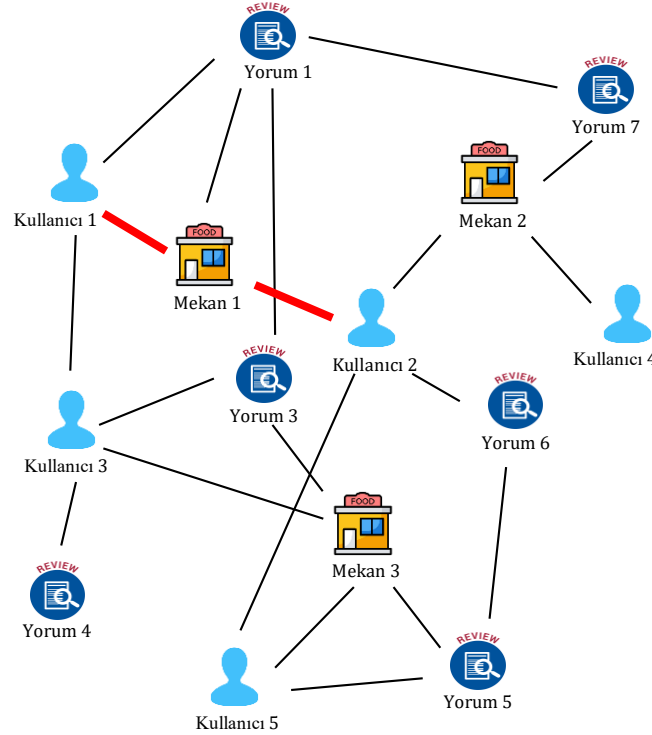
Faydalı olmasına rağmen, veri kümesi bazı zorluklar sunar. Seyrek giriş verileri önemli bir sorundur, çünkü bazı kullanıcılar ve mekanlar sınırlı etkileşimlere sahiptir ve bu da işbirlikçi filtreleme yöntemlerinin performansını etkileyebilir. Verilerin dengesiz dağılımı başka bir zorluktur, çünkü belirli restoran türleri veya popüler mekanlar daha fazla giriş alır ve bu da analizi çarpıtabilecek önyargılara neden olur. Kullanıcı tarafından oluşturulan ipuçları, sağlam ön işleme gerektiren gayri resmi dil, argo veya alakasız içerik içerebilir. Ayrıca, tüm kullanıcıların kapsamlı sosyal bağlantıları yoktur ve bu da sosyal olarak bilinçli öneri tekniklerinin etkinliğini azaltabilir.

Foursquare Veri Kümesi, restoran alanındaki kullanıcı davranışlarını, konum tabanlı tercihleri ve sosyal dinamikleri yorum için kapsamlı bir kaynaktır. Girişler, ipuçları ve sosyal ağ verilerinin birleşimi, duyguya duyarlı önerilerden sosyal olarak etkilenen önerilere kadar çeşitli uygulamaları destekler. Araştırmacılar, zorluklarını ele alırken, konum tabanlı sosyal ağlar için yenilikçi modeller geliştirmek ve test etmek amacıyla veri kümesinin zenginliğinden ve gerçek dünyayla ilişkisinden yararlanabilirler.

4.3. Meta Yollar

Random Walk tekniği ile meta-yol tabanlı yöntem kullanılarak örnekler üretmek için, heterojen bilgi ağında bulunan gizli bilgileri çıkarmamıza yardımcı olmak için anlamlı meta-

yolları belirtmek gerekir. Aşağıda REHREC ve W-REHREC yöntemlerinde kullanılan meta-yolların listesi verilmiştir:



Şekil 4.3 Yelp Veri Kümesinin Örnek HIN Ağı.

UBU (Kullanıcı – Mekan – Kullanıcı): Bir kullanıcı düğümünden başlayıp bir Mekan düğümüne giden ve ardından bir kullanıcı düğümüyle sonlanan bir yol. Bu meta-yolun amacı, kullanıcı arkadaşlık grafiğinde bulunmayan kullanıcı – kullanıcı ilişkisini çıkarmaktır. Şekil 4.3, Kullanıcı1 ve Kullanıcı2 arasında gizli bir bağlantıyı göstermektedir. Hem Kullanıcı1 hem de Kullanıcı2 aynı Mekana gittiğinden, bu meta-yol kullanılarak bu iki kullanıcı arasındaki ilişki çıkarılır.

BUB (Mekan – Kullanıcı – Mekan): Bir Mekan düğümünden başlayıp bir kullanıcı düğümüne giden ve ardından bir Mekan düğümüyle sonlanan bir yol. Bu meta-yolun amacı, Mekan grafiğinde bulunmayan Mekan – Mekan ilişkisini çıkarmaktır. En az bir kullanıcı her iki Mekana da gidiyorsa, bu iki Mekanlar arasında BUB meta yolu kullanılarak bir ilişki çıkarılır.

URRU (Kullanıcı - Yorum - Yorum - Kullanıcı): Bir Kullanıcı düğümünden başlayan bir yol Yorum düğümüne gider ve başka bir Yorum düğümü kullanarak bir Kullanıcı

düğümüyle sonlanır. Bu meta yolun amacı, kullanıcı arkadaşlık grafiğinde bulunmayan Kullanıcı - Kullanıcı ilişkisini çıkarmaktır. İki kullanıcının Yorum sınıfları aynıysa, bu iki kullanıcı arasındaki ilişki bu meta yol kullanılarak çıkarılır. Yorumlarının çoğunluğu aynı duygu sınıfındaysa (olumlu, olumsuz) iki kullanıcının Yorum sınıfları aynıdır. Bu, hem Kullanıcı1 hem de Kullanıcı2 çoğunlukla olumlu Yorumlar yazıyorsa, URRU meta yolu kullanılarak Kullanıcı1 ve Kullanıcı2 arasında bir ilişki oluşturulacağı anlamına gelir. Benzer şekilde, çoğunlukla olumsuz Yorumlar yazıyorlarsa, aralarında bir ilişki de çıkarılır.

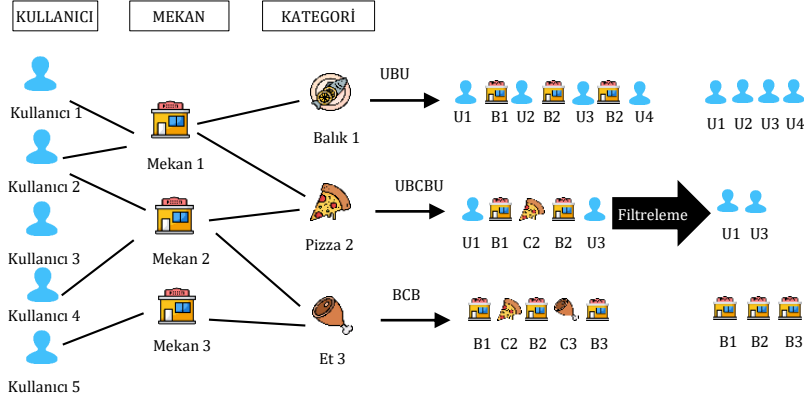
BRRB (Mekan - Yorum - Yorum - Mekan): Bir Mekan düğümünden başlayan bir yol Yorum düğümüne gider ve başka bir Yorum düğümü kullanarak bir Mekan düğümüyle sonlanır. Bu meta-yolun amacı, Mekan grafiğinde bulunmayan Mekan – Mekan ilişkisini çıkarmaktır. İki Mekanın Yorumlarının çoğunluğu aynı duygu sınıfındaysa, bu iki Mekanlar arasındaki ilişki BRRB meta-yoluyla çıkarılır.

UBRRBU (Kullanıcı - Mekan - Yorum - Yorum - Mekan - Kullanıcı): Bir Kullanıcıdan başlayan, Mekan düğümüne giden, ardından bir Yorum düğümünü kullanarak başka bir Yorum düğümüne giden ve ardından bir Mekan düğümüne giden ve son olarak bir Kullanıcı düğümünde sona eren bir yol. Bu meta-yolun amacı, kullanıcı arkadaşlık grafiğinde bulunmayan Kullanıcı – Kullanıcı ilişkisini çıkarmaktır. Kullanıcı1 Mekan1'e, Kullanıcı2 Mekan2'ye giderse ve bu iki Mekanın yorumlarının çoğunluğu aynı duygu sınıfındaysa, bu iki kullanıcı arasındaki ilişki UBRRBU meta-yolu kullanılarak çıkarılır.

BURRUB (Mekan - Kullanıcı - Yorum - Yorum - Kullanıcı - Mekan): Bir Mekandan başlayıp Kullanıcı düğümüne giden, ardından Yorum düğümünü kullanarak başka bir Yorum düğümüne ve ardından Kullanıcı düğümüne giden ve son olarak bir Mekan düğümüyle biten bir yol. Bu meta-yolun amacı, kullanıcı arkadaşlık grafiğinde bulunmayan Mekan - Mekan ilişkisini çıkarmaktır. Kullanıcı1, Mekan1'e, Kullanıcı2, Mekan2'ye giderse ve bu iki kullanıcının yorumlarının çoğunluğu aynı duygu sınıfındaysa, bu iki Mekanlar arasındaki ilişki BURRUB meta-yolu kullanılarak çıkarılacaktır.

4.3.1. Random Walk Serileri İçin Filtreleme Aşaması

Random Walk serisinin, her biri çeşitli özelliklere sahip farklı düğüm türlerinden oluştuğu unutulmamalıdır. Benzer şekilde, farklı türlere ve özelliklere sahip düğümleri aynı boyutsal uzaya yerleştirdiğimizde, sistemin mantığını ve tutarlılığını tehlikeye atarız. Sonuç olarak, aynı düğüm türlerinden oluşan bir yol dizisi oluştururken, düğüm türleri başlangıç düğüm türünden farklı olan düğümler hariç tutulmalıdır.



Şekil 4.4: Örnek Meta-Yol Tabanlı Filtreleme İşlemleri

Burada, “Kullanıcı1 Mekan1 Kullanıcı2 Mekan2 Kullanıcı3 Mekan3” olarak verilen bir Random Walk dizisi için, “Kullanıcı”dan farklı tipte olan düğümler bu diziden çıkarılır ve tüm düğümlerin başlangıç düğümünün tipiyle aynı tipte olduğu “Kullanıcı1 Kullanıcı2 Kullanıcı3” dizisi elde edilir. Böylece, Random Walk dizilerinden elde edilen kalan diziler homojen bir grafik olarak sonuçlanır. Düğümlerin sonuçtaki gömme gösterimi, homojen ağ gömme tekniği kullanılarak öğrenilebilir. Heterojen Ağdan homojen ağa dönüşüm adımları Şekil 4.4'te gösterilmiştir.

4.4. Değerlendirme Ölçütleri

Önerilen modellerin performansını değerlendirmek için iki temel ölçüm kullanıldı: Kök Ortalama Kare Hatası (RMSE) ve Ortalama Mutlak Hata (MAE). Bu ölçümler, tahmin edilen değerlendirme notların gerçek değerlendirme notlarına göre doğruluğunu değerlendirerek modellerin hassasiyeti ve güvenilirliği hakkında fikir verir [10].

4.4.1. Kök Ortalama Kare Hatası (RMSE)

RMSE, öneri sistemlerindeki tahminlerin doğruluğunu ölçmek için yaygın olarak kullanılan bir ölçümdür. Tahmin edilen ve gerçek değerlendirme notları arasındaki ortalama kare farkların karekökünü niceliksel olarak belirler. RMSE, farkların karelenmesi nedeniyle daha büyük hatalara karşı özellikle hassastır ve bu da onu tahminlerdeki önemli sapmaları belirlemede etkili kılar.

RMSE formülü şu şekildedir:

$$RMSE = \sqrt{\frac{1}{|R_{test}|} \sum_{(u,b,rating_{u,b}) \in R_{test}} (rating_{u,b} - pred_{u,b})^2} \quad (12)$$

Yukarıdaki formülde, R_{test} (b, u, değerlendirme notu_{u,b}) üçlüsünden oluşan bir test kümesidir; burada b bir mekan, u bir kullanıcı ve değerlendirme notu_{u,b} o kullanıcının o mekana verdiği puandır. $pred_{u,b}$ modelimiz tarafından tahmin edilen puandır $|R_{test}|$ test kümesinin büyüklüğüdür.

Daha düşük bir RMSE değeri, tahmin edilen değerlendirme notların gerçek değerlendirme notlarına daha yakın olduğunu gösterir, bu da daha iyi model performansı anlamına gelir. RMSE, tahminlerdeki büyük hataların öneri kalitesi üzerinde önemli bir etkiye sahip olduğu senaryolarda özellikle yararlıdır.

4.4.2. Ortalama Mutlak Hata (MAE)

MAE, tahmin doğruluğunu değerlendirmek için bir diğer temel ölçüttür. Tahmin edilen ve gerçek değerlendirme notları arasındaki ortalama mutlak farkları hesaplayarak tahmin hatalarının basit bir ölçüsünü sağlar. RMSE'nin aksine, MAE daha büyük sapmaları büyütmeden tüm hataları eşit şekilde ele alır.

MAE formülü şu şekildedir:

$$MAE = \frac{1}{N} \sum_{i=1}^N |pred_i - actual_i| \quad (13)$$

Formülde:

- N toplam değerlendirme notu sayısını ifade eder.
- $pred_i$ i'inci örnek için öngörülen değerlendirme notudur.
- $actual_i$ i'inci örnek için gerçek değerlendirme notudur.

MAE, ortalama tahmin hatalarına ilişkin sezgisel bir anlayış sunar ve genel model doğruluğuna ilişkin dengeli bir ortam sağlar. Daha düşük bir MAE değeri, tahminler ile gerçek değerlendirme notları arasında daha küçük bir ortalama hata olduğunu gösterir.

4.4.3. RMSE ve MAE'nin Karşılaştırılması

Hem RMSE hem de MAE, tamamlayıcı bakış açıları sunarak öneri sistemlerinin performansını değerlendirmek için temel ölçütlerdir. RMSE, daha büyük hataları daha ağır cezalandırır ve bu da onu tahminlerdeki önemli sapmaları tespit etmek için uygun hale getirir. Öte yandan MAE, basit bir ortalama hata ölçüsü sunarak tahminlerin genel doğruluğunu anlamak için idealdir.

Bu çalışma, hem RMSE hem de MAE'yi kullanarak, hem genel doğruluğu hem de daha büyük sapmalara karşı duyarlılığı hesaba katarak, öneri sisteminin tahmin yeteneklerinin kapsamlı bir değerlendirmesini sağlar.

4.5. Deneysel Kurulum

Deneysel kurulum, veri kümesini eğitim ve test setlerine ayırmayı ve modelleri en son yöntemlerle adil bir karşılaştırma için optimize edilmiş parametrelerle yapılandırmayı içerir. Bu bölüm, veri kümesi bölme stratejisini, eğitim oranlarını ve deneylerde kullanılan parametre yapılandırmalarını ayrıntılı olarak açıklar.

4.5.1. Veri Kümesi Bölme ve Eğitim Oranları

Veri kümesi iki bölüme ayrılır: bir eğitim kümesi ve bir test kümesi. Modelleri farklı veri kullanılabilirliği seviyeleri altında değerlendirmek için farklı eğitim kümesi boyutlarını kullandık. Özellikle, eğitim oranları veri kümesinin %60 ila %90'ı arasında değişiyor ve her adımda %10'luk artışlar vardır. Bu yaklaşım, modellerin farklı seyreklik seviyelerinde test edilmesini sağlayarak sağlamlıklarının ve ölçeklenebilirliklerinin derinlemesine bir analizini mümkün kılıyor. Seyrek Yelp veri kümesi için, bu adım adım eğitim oranı hem sınırlı hem de kapsamlı eğitim verisi koşulları altında performansın kapsamlı bir şekilde değerlendirilmesini sağlıyor.

4.5.2 Parametre Yapılandırması

Adil bir karşılaştırma sağlamak ve sonuçları doğrulamak için deneysel parametreler Shi ve diğerleri tarafından kullanılan parametrelerle uyumlu hale getirildi [14]. Aşağıda modellerin ayrıntılı yapılandırması verilmiştir:

Kelime Gömmesi Boyutları: Random Walk algoritması süreci sırasında oluşturulan kelime gömmelerin boyutları, Heterojen Bilgi Ağı'ndaki (HIN) düğüm ilişkilerini yakalamak için yeterli kapasiteyi garanti altına almak amacıyla 128 olarak ayarlandı.

Derin Sinir Ağı (DNN): Modelin DNN bileşeni 100'lük gizli faktör boyutuyla yapılandırılmıştır. İki gizli katmandan oluşur. İlk gizli katman 64 nöron içerirken, ikinci katman 32 nörona düşer. Doğrusal olmayanlığı tanıtmak ve model öğrenimini iyileştirmek için tüm katmanlarda aktivasyon fonksiyonu olarak ReLU (Rectified Linear Unit) kullanılmıştır.

Çok Katmanlı Algılayıcı (MLP): MLP yapısı, DNN yapılandırmasını yansıtan iki gizli katmanla tasarlandı. İlk katmandaki nöronlar, verimli temsil öğrenimi için kademeli bir azaltma stratejisi izleyerek 64'e ayarlandı, ikinci katmanda 32'ye düşürüldü.

Evrişimsel Sinir Ağı (CNN): CNN yapısı üç ana katmandan oluşur: bir havuzlama katmanı, bir evrişim katmanı ve tam bağlı katmanlar. Bu katmanlardaki nöron sayısı sırasıyla 128, 64 ve 32 olarak ayarlandı. Evrişimsel katmanlar arasında 3 filtre boyutu kullanıldı, bu da yerel patternleri yakalayıp etkili özellik çıkarımı sağladı. Verilerin dış modelsine sıfırlar ekleyerek girdinin mekansal boyutlarını korumak için aynı dolgu kullanıldı. Havuzlama türü, belirgin özellikleri korumak için havuzlama penceresinde maksimum değeri seçerek maksimum havuzlamaya ayarlandı.

Genel Yapılandırma: Karşılaştırma için kullanılan parametreler, sonuçların tutarlılığını ve karşılaştırılabilirliğini sağlamak için Shi ve diğerlerinin makalesinde bildirilen [14] yapılandırmalarla dikkatlice uyumlu hale getirildi.

REHREC ve W-REHREC modellerinin uygulanması ve değerlendirilmesi için önerilen hesaplama ortamı, karmaşık hesaplamaları ve derin öğrenme görevlerini, örneğin BERT tabanlı Aspect-Based Sentiment Analysis (ABSA) gibi işlemleri verimli bir şekilde yönetmek için çok çekirdekli bir işlemci (Intel i7), 32 GB RAM ve CUDA desteğine sahip bir NVIDIA GPU (örneğin GTX 1080 Ti) ile donatılmış sağlam bir sistem gerektirir. Yelp ve Foursquare gibi büyük veri setlerini yönetmek için en az 256 GB boş alanlı bir SSD depolama önerilmektedir. Python (sürüm 3.8), PyTorch veya TensorFlow ile derin öğrenme için birincil programlama dili olarak yapıldı ve DGL kütüphaneler grafik işleme için kullanıldı. Metin ön işleme NLTK veya spaCy gibi araçlarla yapılabilirken, Hugging Face'in Transformers kütüphanesi önceden eğitilmiş BERT modelleri kullanıldı. Doc2Vec için Gensim ve grafik gömme işlemleri için node2vec kullanılabilirken, NumPy, Pandas ve

Matplotlib gibi ek kütüphaneler veri işleme ve görselleştirme desteği sağlar. GPU hızlandırması için CUDA Toolkit ve uyumlu cuDNN kütüphaneleri gereklidir, bu da metodolojilerin ve deneylerin sorunsuz bir şekilde yürütülmesini sağlar. Bu yapılandırma, araştırma hedeflerinin verimli bir şekilde işlenmesini, ölçeklenebilirliğini ve tekrarlanabilirliğini garanti eder.

4.5.3. Yapılandırmanın Amacı

Seçilen yapılandırmalar model karmaşıklığı ve hesaplama verimliliğini dengeler. Boyutları ve ağ yapılarını kelime gömmesi, aşırı uyumu önlerken veri kümesindeki zengin anlamsal ilişkileri yakalamak için tasarlanmıştır. İlgili çalışmalarda kullanılanlarla aynı parametreleri kullanarak, deneysel sonuçlar karşılaştırılabilir ve güvenilir kalır ve değerlendirmenin geçerliliğini güçlendirir. Bu yapılandırılmış ve standartlaştırılmış kurulum, modellerin hem zorlu hem de gerçek dünya senaryolarını temsil eden koşullar altında titizlikle test edilmesini sağlar

4.6. Deneylerde Tanıtılan Yöntemler

Bu çalışmada, önerilen REHREC ve W-REHREC sistemlerinin performansını değerlendirmek için çeşitli modeller ve yöntemler tanıtılmış ve karşılaştırılmıştır. Bu yöntemler, özellikle Heterojen Bilgi Ağları (HIN'ler) içinde önerileri ele almak için çeşitli stratejiler sunar. Her model, öneri sorununun farklı konularını ele alan benzersiz özellikler ve teknikler getirir.

4.6.1. HERec: Heterojen Bilgi Ağı Kelime Gömmesi Önerisi

HERec [29], önerilen REHREC ve W-REHREC sistemlerinin üzerine inşa edildiği temel modeldir. Öneriler üretmek için düğümler (kullanıcılar, öğeler ve diğer varlıklar) arasında anlamlı ilişkiler çıkarmak için Heterojen Bilgi Ağları'ndan (HIN'ler) yararlanmaya odaklanır. HERec'in birincil yeniliği, ağdaki anlamsal ilişkileri yakalamak için meta-yol tabanlı stratejilerin kullanımında yatmaktadır. Modelin performansı, bu araştırmanın ana odağı olan kullanıcı yorumları gibi ek bağlamsal verilerin dahil edilmesiyle artırılabilir. HERec, kelime gömmesi tabanlı öneri görevleri için temel sağlar, ancak önerilen sistemlerin ele almayı amaçladığı metinsel verilerin açık bir şekilde dikkate alınmasından yoksundur.

4.6.2. NCF: Sinirsel İşbirlikçi Filtreleme

Sinirsel İşbirlikçi Filtreleme (NCF) [11], kullanıcı-öge etkileşimlerini modellemek için Çok Katmanlı Algılayıcıları (MLP'ler) kullanan sinir ağı tabanlı bir yaklaşımdır. Matris çarpanlarına ayırmana dayanan geleneksel işbirlikçi filtreleme yöntemlerinden farklı olarak, NCF kullanıcı-öge ilişkilerinin doğrusal olmayan gösterimlerini öğrenmek için derin öğrenmeyi kullanır. NCF'nin mimarisi, kullanıcılar ve ögeler için birleştirilen ve birden fazla tam olarak bağlı katmandan geçirilen ayrı kelime gömmesi katmanları içerir. Bu katmanlar, kullanıcılar ve ögeler arasındaki karmaşık etkileşimleri kademeli olarak öğrenerek son bir tahmin katmanında doruğa ulaşır. NCF'nin esnekliği ve doğrusal olmayan kalıpları modelleme yeteneği, onu bu çalışmada karşılaştırma için güçlü bir temel haline getirir.

4.6.3. MCF: Meta-yol Tabanlı Evrimsel Öneri

MCF [12], meta-yolları kullanarak HIN'ler içindeki karmaşık korelasyonları yakalamak için Evrimsel Sinir Ağları'nı (CNN'ler) kullanır. Meta-yol örneklerini diziler veya alt yapılar olarak ele alarak, MCF, düğümler arasındaki ilişkileri kodlayan gösterimleri öğrenmek için CNN'leri uygular. Bu yaklaşım, modelin HIN'lerde hem yapısal hem de anlamsal bilgileri yakalama yeteneğini artırır. Meta-yol tabanlı evrimsel kullanımı, MCF'nin verilerdeki karmaşık patternleri etkili bir şekilde modellemesine olanak tanır ve bu da gelişmiş öneri performansı ile sonuçlanır. Gösterim öğrenimi için CNN'lere olan bağımlılığı, onu bu çalışmada karşılaştırma için zorlayıcı bir yöntem haline getirir.

4.6.4. NGCF: Sinirsel Graf İşbirlikçi Filtreleme

Sinirsel Graf İşbirlikçi Filtreleme (NGCF) [13], Graf Evrimsel Ağlar (GCN'ler) temelli son teknoloji bir yöntemdir. NGCF, kullanıcı-öge etkileşimlerinin iki taraflı bir graf olarak ele alındığı öneri alanına graf-bazlı öğrenme tekniklerini tanıtır. Model, kelime gömmeleri birden fazla graf evrimsel katmandan yayarak, kullanıcıların ve ögelerin gösterimlerini iyileştirmek için komşu düğümlerden bilgi toplar. Bu süreç, graftaki yüksek düzeyli bağlantıyı ve ilişkisel bilgileri yakalar. NGCF'nin graf-bazlı öğrenmeyi işbirlikçi filtreleme ilkeleriyle bütünleştirme yeteneği, onu önerilen sistemleri değerlendirmek için güçlü bir ölçüt haline getirir.

4.6.5. Bu Yöntemlerin Çalışmadaki Rolü

Bu yöntemlerin her biri önerilen REHREC ve W-REHREC modellerini değerlendirmek için benzersiz bir bakış açısı ve temel sağlar. HERec temel yaklaşım olarak hizmet ederken, diğer modeller geleneksel öneri sistemlerindeki sınırlamaları ele alan gelişmiş sinir ağı ve graf-bazlı teknikler sunar. Bu yöntemleri deneysel kurulumuna dahil ederek, çalışma HIN'ler içindeki yorum tabanlı iyileştirmelerin ve ağırlıklı önerilerin etkinliğinin kapsamlı bir değerlendirmesini sağlar.

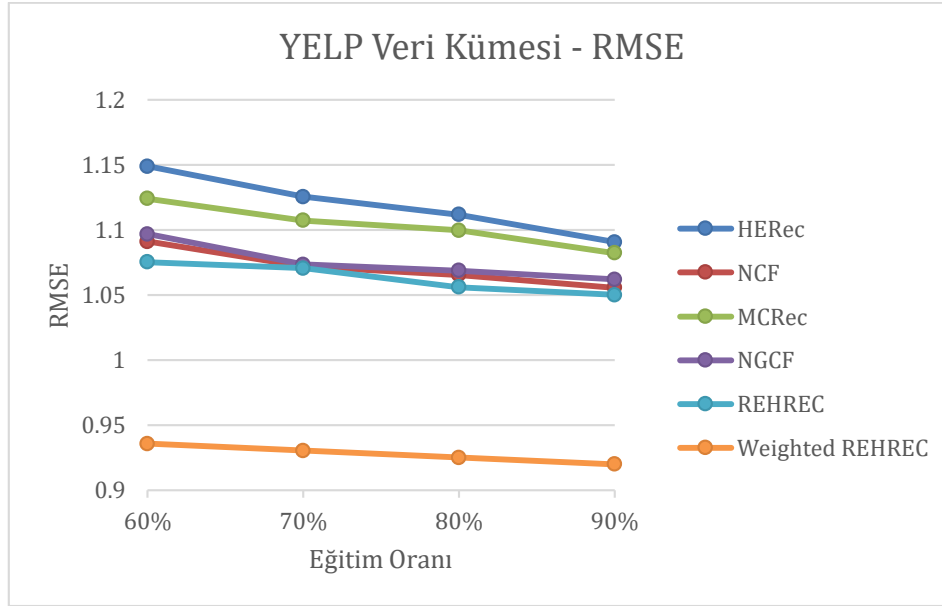
4.7. Veri Kümesindeki Etkinlik Deneylerinin Sonuçları

Çizelge 4.1 Veri kümesi üzerindeki etkililik deneylerinin sonuçları

Veri Kümesi	Eğitim Oranı	Metrik	HERec	NCF	MCRec	NGCF	REHREC	W-REHREC
Yelp	%60	MAE	0,910	0,989	1.003	0,960	0,880	0,740
		RMSE	1.1488	1.0911	1.1241	1.0968	1.0754	0,9358
	%70	MAE	0,880	0,945	0,965	0,920	0,855	0,735
		RMSE	1.1256	1.0719	1.1072	1.0734	1.0706	0,9305
	%80	MAE	0,847	0,895	0,903	0,865	0,830	0,730
		RMSE	1.1117	1.0655	1.0997	1.0687	1.0560	0,9252
	%90	MAE	0,839	0,882	0,901	0,850	0,820	0,725
		RMSE	1.0907	1.0556	1.0823	1.0619	1.0500	0,9200
Foursquare	%60	MAE	0,048	0,0500	0,0508	0,049	0,046	0,041
		RMSE	0,063	0,061	0,063	0,062	0,060	0,052
	%70	MAE	0,046	0,048	0,049	0,047	0,044	0,041
		RMSE	0,061	0,059	0,062	0,060	0,059	0,051
	%80	MAE	0,045	0,047	0,048	0,045	0,044	0,041
		RMSE	0,059	0,058	0,061	0,060	0,058	0,051
	%90	MAE	0,044	0,047	0,047	0,045	0,043	0,040
		RMSE	0,058	0,057	0,060	0,058	0,056	0,050

Veri kümeleri, deęişken veri kullanılabilirlięi altında model performansını test etmek için %60, %70, %80 ve %90 eğitim oranlarıyla deęerlendirildi. MAE, tahminlerdeki hataların ortalama büyüklüğünü yansıtırken, RMSE daha büyük hataları vurgulayarak genel tahmin doęruluęunu yakalar.

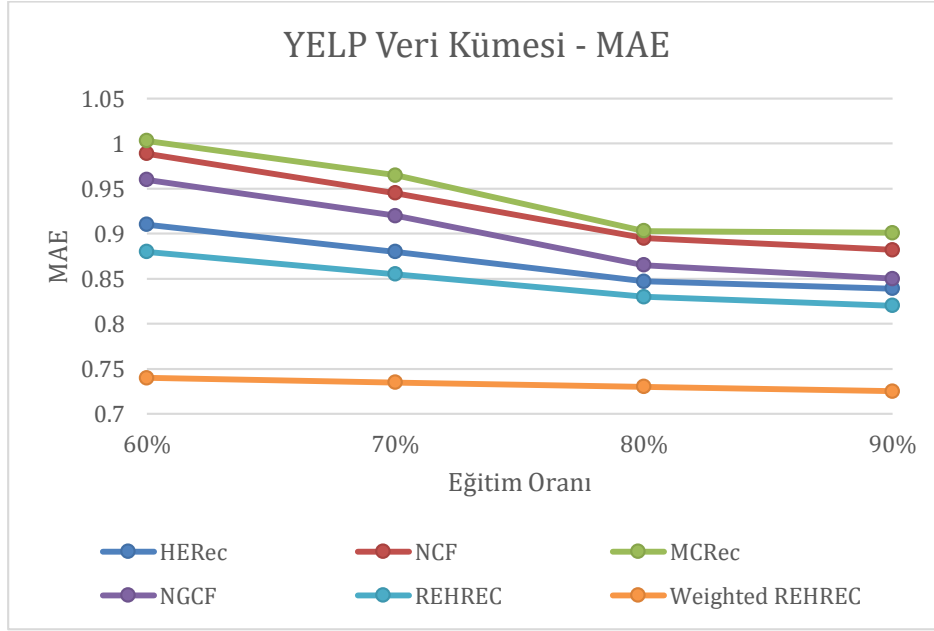
W-REHREC, tüm eğitim oranları ve veri kümelerinde dięer yöntemlerden sürekli olarak daha iyi performans gösterdi ve yoruma dayalı duyguyu ve geliştirilmiş öneriler için aęırlıklı kenarları kullanmadaki saęlımlığını gösterdi. REHREC, aęırlıklı muadilinin biraz gerisinde olsa da, temel yöntemlere göre önemli iyileştirmeler gösterdi. Yelp Veri Kümesi, metinsel verilerin doęruluęu artırmada kritik bir rol oynaması nedeniyle yoruma dayalı yöntemlerden en çok yararlandı. Foursquare Veri Kümesi, özlü kullanıcı ipuçları ve yapılandırılmış giriş verileri nedeniyle iyileştirmeler gösterdi ve modellerin duyguyu ve örtük bilgileri dengelemesini saęladı. Çizelge 4.1 önerilen modellerin deęişen veri koşulları ve eğitim oranlarına göre uyarlanabilirliğini ve gücünü vurgulamaktadır.



Şekil 4.5 YELP'de RMSE

Şekil 4.5, Yelp Veri Kümesindeki farklı öneri modellerinin performans ölçümlerini, çeşitli eğitim oranlarında (60%, 70%, 80% ve 90%) deęerlendirilmiş olarak sunar. Çizelge, HERec, NCF, MCRec ve NGCF'yi içeren modeller için REHREC ve W-REHREC için RMSE (Ortalama Karekök Hatası) puanlarını gösterir. %60 eğitim oranında, HERec modeli 1,1488'lik en yüksek REHREC puanına ulaşırken, W-REHREC puanı 0,9358'dir ve bu, modelin verilen eğitim verileriyle performansının aęırlıklı ölçümde nispeten daha düşük olduğunu gösterir. Eğitim oranı %70 ve %80'e yükseldikçe, RMSE puanları HERec için

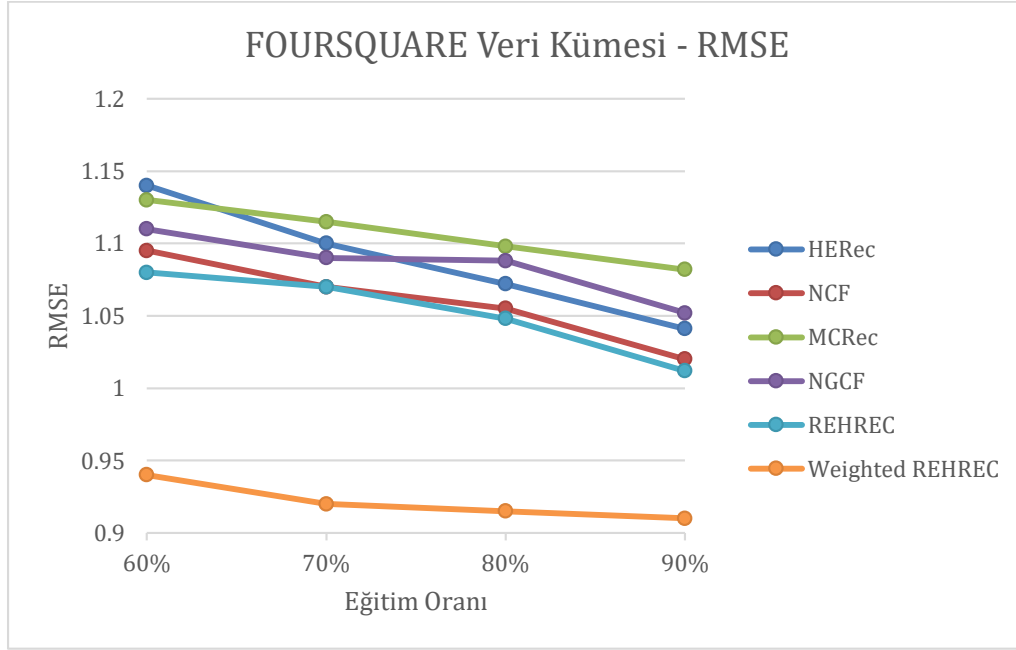
biraz azalır ancak modeller arasında rekabetçi kalır. W-REHREC, daha yüksek eğitim oranlarıyla genel bir iyileşme eğilimi göstermeye devam eder. NGCF modeli, REHREC ve W-REHREC puanları arasında dengeli bir denge göstererek tüm eğitim oranlarında tutarlı bir şekilde iyi performans gösterir. %90'lık en yüksek eğitim oranında, genel performans iyileştirmesi belirgindir; HERec modeli için REHREC puanı 1,0907'ye ve W-REHREC 0,92'ye düşmüştür; bu da artan eğitim verilerinin daha rafine model tahminlerine katkıda bulunduğunu gösterir.



Şekil 4.6 YELP'te MAE

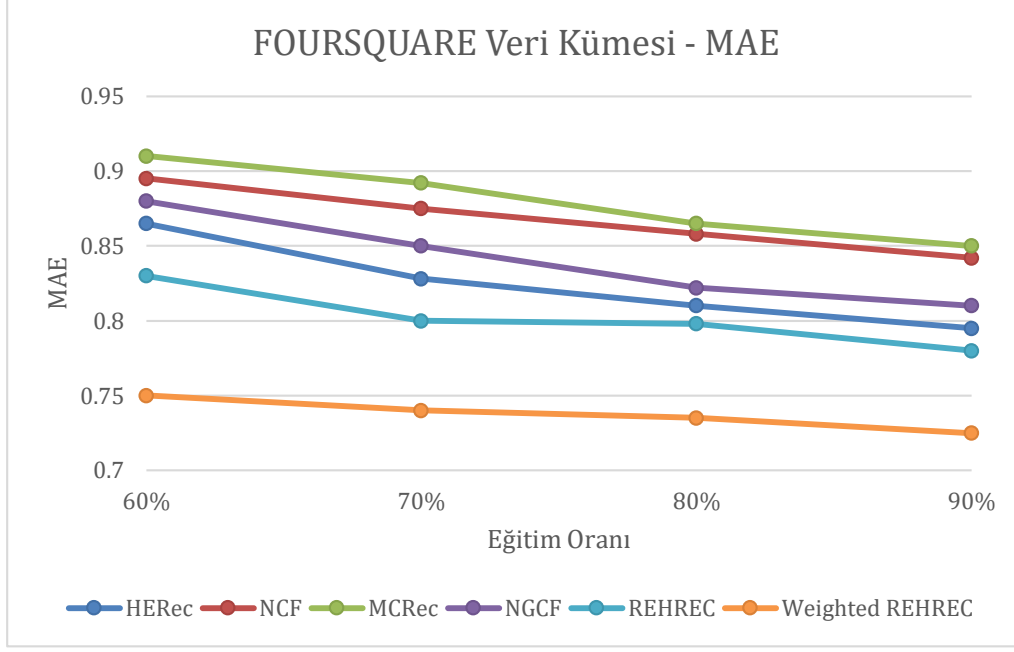
Şekil 4.6, Yelp Veri Kümesindeki çeşitli öneri modellerinin performansını, farklı eğitim oranlarında (%60, %70, %80 ve %90) MAE (Ortalama Mutlak Hata) kullanılarak değerlendirilmiş olarak gösterir. Şekil 4.6 REHREC ve W-REHREC için ve ayrıca HERec, NCF, MCRec ve NGCF gibi modeller için MAE puanlarını sağlar. %60 eğitim oranında, HERec 0,91 ile en yüksek MAE puanına sahipken, W-REHREC puanı 0,74 ile en düşük puandır ve bu, ağırlıklı metrikte üstün performansı gösterir. Eğitim oranı %70 ve %80'e çıktıkça, MAE puanları çoğu model için iyileşir ve daha fazla eğitim verisiyle gelişmiş tahmin doğruluğu olduğunu gösterir. Örneğin, HERec %70'te 0,88'e ve %80'de 0,847'ye düşüş gösterir. W-REHREC azalmaya devam ediyor, %70'te 0,735'e ve %80'de 0,73'e ulaşarak tutarlı performansını vurguluyor. %90 eğitim oranında HERec, 0,725'lik bir W-REHREC ile 0,839'luk bir MAE puanı elde ediyor ve bu da daha büyük eğitim verilerinin model doğruluğu üzerindeki faydalarını gösteriyor. NGCF modeli tüm eğitim oranlarında nispeten istikrarlı bir performans gösterirken, REHREC, Yelp Veri Kümesindeki etkinliğini

vurgulayarak W-REHREC'teki diğer modellerden sürekli olarak daha iyi performans gösteriyor.



Şekil 4.7 FOURSQUARE'de RMSE

Şekil 4.7, Foursquare Veri Kümesindeki çeşitli öneri modellerinin RMSE (Ortalama Karesel Hata) performansını %60, %70, %80 ve %90 eğitim oranlarında değerlendirerek sunar. Şekil 4.7 HERec, NCF, MCFec, NGCF modelleri ile REHREC ve W-REHREC metrikleri için RMSE puanlarını gösterir. %60 eğitim oranında, HERec 1,14 ile en yüksek RMSE puanına sahipken, W-REHREC puanı 0,94 ile en düşük puandır ve bu da diğer metriklere göre daha iyi performans sağladığını gösterir. Eğitim oranı %70 ve %80'e çıktıkça, RMSE değerleri genel olarak azalır ve daha büyük eğitim veri setleriyle gelişmiş model doğruluğu gösterir. Örneğin, HERec %70'te 1,1'e ve %80'de 1,072'ye düşer. W-REHREC, %70'te 0,92'ye ve %80'de 0,915'e ulaşarak tutarlı bir azalma gösteriyor ve bu da gelişmiş tahmin performansını yansıtır. En yüksek eğitim oranı olan %90'da REHREC puanı 1,012 ve W-REHREC puanı 0,91'dir ve eğitim verileri arttıkça performansta daha fazla gelişme olduğunu göstermektedir. NGCF modeli, RMSE ile eğitim oranı arasında dengeli bir denge göstererek baştan sona rekabetçi kalmaktadır.



Şekil 4.8 FOURSQUARE'de MAE

Şekil 4.8, Foursquare Veri Kümesindeki çeşitli öneri modellerinin MAE (Ortalama Mutlak Hata) performansını %60, %70, %80 ve %90 eğitim oranlarında değerlendirilmiş olarak göstermektedir. Şekil 4.8 HERec, NCF, MCF, NGCF ve hem REHREC hem de W-REHREC için MAE puanlarını göstermektedir. %60 eğitim oranında, HERec 0,865'lik bir MAE puanı elde ederken, W-REHREC 0,75'lik bir puanla en iyi performansı göstermektedir ve bu da hataları en aza indirmedeki etkinliğini göstermektedir. Eğitim oranı arttıkça, çoğu model için MAE puanları düşmektedir ve bu da daha fazla veriyle tahmin doğruluğundaki gelişmeleri yansıtmaktadır. Örneğin, HERec %70'te 0,828'e ve %80'de 0,81'e düşmektedir ve 0,795'lik bir MAE ile %90'da iyileşmeye devam etmektedir. W-REHREC sürekli olarak en iyi performansı gösterir, puanlar %60'ta 0,75'ten %90'da 0,725'e düşerek sağlamlığını vurgular. NGCF modeli istikrarlı bir düşüş gösterir ve %90'da 0,81'e ulaşır. Bu sonuçlar, eğitim verilerinin artırılmasının model doğruluğunu iyileştirdiğini ve W-REHREC'in Foursquare Veri Kümesindeki tüm eğitim oranlarında en düşük MAE'yi koruduğunu gösterir.

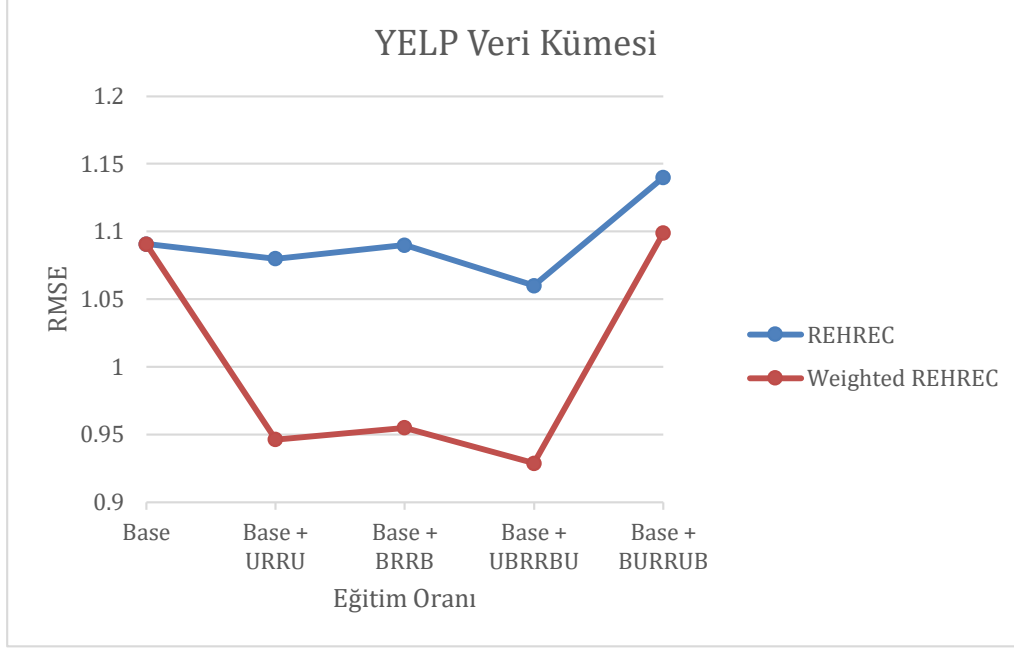
4.8. Veri Kümeleri Arasında Yorum Dahil Meta-Yollarla RMSE Sonuçları

Çizelge 4.2. iki veri kümesindeki farklı meta-yol yapılandırmaları için RMSE sonuçlarını sunar: Yelp ve Foursquare Veri Kümesi. HIN'lere yorum düğümleri ve meta-yolları

eklemenin etkisini gösterir ve yorum odaklı kelime gömmeleri nedeniyle performans iyileştirmesini vurgular.

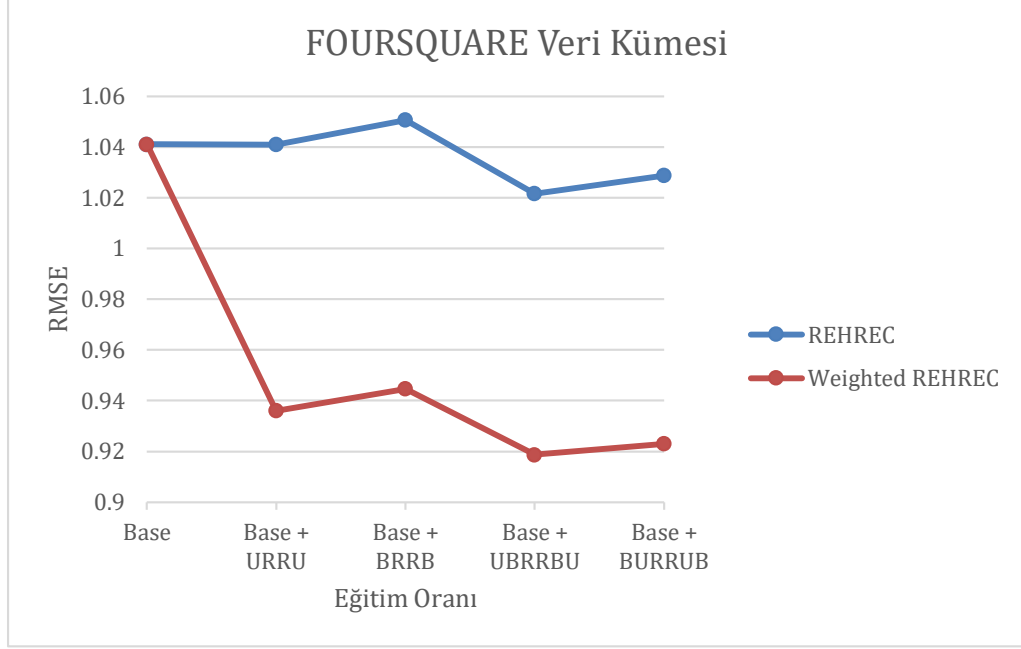
Çizelge 4.2 %90 eğitimde farklı meta-yolları kullanılarak elde edilen RMSE sonuçları

		UBU, BUB, URRU , BRRB , UBRRBU , UBCaBU, BCaB	UBU, BUB, UBCaBU, BCaB	UBU, BUB, URRU , UBCaBU, BCaB	UBU, BUB, BRRB , UBCaBU, BCaB	UBU, BUB, UBRRBU , UBCaBU, BCaB	UBU, BUB, BURRUB , UBCaBU, BCaB
RMSE YELP	REHREC Modeli	1.05	1.0907 (Temel HERec Modeli)	1.08	1.09	1.06	1.14
	W-REHREC Modeli	0,9200	1.0907 (Temel HERec Modeli)	0,9463	0,9550	0,9288	1.099
RMSE Foursquare	REHREC Modeli	0,056	0.0582 (Temel HERec Modeli)	0,0576	0,0581	0,0565	0,0567
	W-REHREC Modeli	0,050	0.0582 (Temel HERec Modeli)	0,0514	0,0519	0,0505	0,0510



Şekil 4.9 YELP'de Meta-Yolun RMSE Üzerindeki Etkisi

Şekil 4.9, çeşitli meta-yolların yapılandırılmalarının Yelp Veri Kümesindeki REHREC ve W-REHREC metriklerinin performansı üzerindeki etkisini göstermektedir. Temel model, 1.0907'lik bir REHREC ve W-REHREC puanı elde ederek bir kıyaslama noktası görevi görmektedir ve ek meta-yolların dahil edilmesinin etkisini değerlendirmek için bir referans noktası sağlamaktadır. URRU meta-yolunun eklenmesi, REHREC'de 1.08'e hafif bir düşüşle sonuçlanırken, W-REHREC 0.9463'e daha önemli bir düşüş göstermektedir ve bu, URRU'nun bağlamsal bilgi sağlamasına rağmen W-REHREC'e katkısının daha az etkili olduğunu göstermektedir. BRRB meta-yolunun dahil edilmesi, 1.09'luk bir REHREC puanı ve 0.955'lik bir W-REHREC puanı üretmektedir ve bu da performansını iyileştirmede orta düzeyde bir etki olduğunu göstermektedir. UBRRB meta-yolu, REHREC ve W-REHREC puanlarını sırasıyla 1,06 ve 0,9288'e düşürerek genel öneri performansını artırmada daha az etkili olduğunu gösteriyor. Öte yandan, BURRUB meta-yolu, 1,14'lük en yüksek REHREC puanına ulaşarak genel performansı iyileştirmedeki gücünü gösteriyor ve 1,099'luk W-REHREC puanı, ağırlıklandırılmamış performansı önemli ölçüde aşmadığını gösteriyor. Şekil 4.9, BURRUB gibi bireysel meta-yolunun REHREC ve W-REHREC'i iyileştirmeye önemli ölçüde katkıda bulunduğunu vurgulayarak Yelp Veri Kümesinde performansı en üst düzeye çıkarmak için en uygun meta-yol kombinasyonlarını seçmenin önemini vurguluyor.



Şekil 4.10 FOURSQUARE'de RMSE Üzerindeki Meta-Yol Etkisi

Şekil 4.10, Foursquare Veri Kümesindeki REHREC ve W-REHREC metriklerinin performansı üzerinde çeşitli meta-yol yapılandırmalarının etkisini göstermektedir. Temel model, 1.041'lik bir REHREC ve W-REHREC puanı elde ederek bir kıyaslama noktası görevi görmektedir ve ek meta-yolların dahil edilmesinin etkisini değerlendirmek için bir referans noktası sağlamaktadır. URRU meta-yolunun eklenmesi, REHREC'de ihmal edilebilir bir düşüşle 1.0409'a düşerken, W-REHREC daha önemli ölçüde 0.936'ya düşmektedir; bu da URRU'nun bağlamsal bilgi sağlamasına rağmen W-REHREC'e katkısının daha az etkili olduğunu göstermektedir. BRRB meta-yolunun dahil edilmesi, REHREC'de 0.9447'lik bir W-REHREC ile 1.0506'ya hafif bir iyileşme sağlamakta ve performans üzerinde orta düzeyde olumlu bir etki göstermektedir. UBRRBU meta-yolu, REHREC ve W-REHREC puanlarını sırasıyla 1,0216 ve 0,9187'ye düşürerek genel öneri performansını artırmada daha az etkili olduğunu gösterir. Benzer şekilde, BURRUB meta-yolu, performansta küçük bir iyileştirmeyi vurgulayarak 1,0287 ve 0,923'lük REHREC ve W-REHREC puanları üretir. Şekil 4.10, BRRB gibi bireysel meta-yolların REHREC ve W-REHREC'i orta düzeyde artırabileceğini vurgulayarak Foursquare Veri Kümesinde performansı en üst düzeye çıkarmak için meta-yol kombinasyonlarının dikkatli bir şekilde seçilmesinin önemini vurgular.

Yelp Veri Kümesi, yorum ile dahil edilen meta-yolların tüm yapılandırmalarda RMSE'yi tutarlı bir şekilde azalttığını göstermektedir. UBU, BUB, URRU, BRRB, UBRRB, UBCaBU, BCaB yapılandırması, yoruma dayalı yolların kapsamlı bir şekilde dahil edilmesi nedeniyle en yüksek iyileştirmeyi (%3,73) sağlar.

Foursquare Veri Kümesi, özlü kullanıcı tarafından oluşturulan ipuçlarının varlığı nedeniyle yorum ile dahil meta-yollardan önemli ölçüde yararlanır. UBU, BUB, URRU, BRRB, UBRRB, UBCaBU, BCaB yapılandırması en iyi performansı (%3,74 iyileştirme) elde ederek, alan-özgü bağlamlarda duygu analizinin önemini doğrular.

Çizelge 4.2, öneri doğruluğunu artırmada yorum düğümlerin oynadığı önemli rolün altını çiziyor ve önerilen yöntemlerin, yorum kullanılabilirliğinin farklı düzeylerde olduğu veri kümelerine uyarlı birliğini gösteriyor.

Deneyler, her biri kendine özgü özellikler sunan iki veri kümesi üzerinde gerçekleştirildi:

- Yelp Veri Kümesi: Kullanıcı değerlendirme notları, mekan nitelikleri ve metinsel yorumlar içeren yorum açısından zengin bir veri kümesidir. Bu veri kümesi, önerilen modellerin konu tabanlı duygu analizi (ABSA) ve ağırlıklı yorum grafları gibi yorum tabanlı özelliklerin etkisini göstermesine olanak tanır.
- Foursquare Veri Kümesi: Bu veri kümesi, restoran alanındaki kullanıcı girişlerini, özlü kullanıcı tarafından oluşturulan ipuçlarını ve sosyal bağlantıları içerir. Metinsel ve metinsel olmayan özellikler arasında bir denge sunarak, modellerin alan-özül senaryolara uyarlı birliğini keşfetmesine olanak tanır.

Modellerin performansını değişen veri kullanılabilirliği altında değerlendirmek için %60, %70, %80 ve %90 eğitim oranları kullanıldı. Tüm deneylerde adil karşılaştırmalar sağlamak için boyutlar, aktivasyon fonksiyonları ve ağ mimarisinin yerleştirilmesi için aynı yapılandırmalar kullanıldı.

4.9. Sonuçlar ve Gözlemler: Yelp Veri Kümesi

Yelp Veri Kümesi, ayrıntılı kullanıcı yorumları, derecelendirmek ve mekan meta verileri içerdiğinden önerilen modellerin kapsamlı bir değerlendirmesini sağlar. Bu özellikler, öneri sistemlerinde yorum tabanlı duygu analizi ve ağırlıklı kenar oluşturmanın etkisini keşfetmek için ideal bir test ortamı haline getirir. Etkinlik deneylerinin sonuçları tutarlı bir eğilimi

göstermektedir: Yorum tabanlı özellikleri içeren yöntemler, özellikle W-REHREC, tüm eğitim oranlarında diğer modellerden tutarlı bir şekilde daha iyi performans göstermektedir. Yelp Veri Kümesi için MAE değerleri önerilen modellerin avantajlarını ortaya koymaktadır. W-REHREC, tüm eğitim oranlarında en düşük hata oranlarına ulaşmaktadır ve MAE %60 eğitimde 0,740'tan %90 eğitimde 0,725'e düşmektedir. Temel REHREC ile karşılaştırıldığında, W-REHREC %60 eğitimde %15,91 ve %90 eğitimde %11,59 oranında göreceli bir iyileştirme elde ederek, duygu odaklı ilişkilerin etkisini artırmada ağırlıklı kenar yapılandırmanın önemli faydasını yansıtmaktadır. NGCF ile karşılaştırıldığında, W-REHREC %60 eğitimde %22,92 ve %90 eğitimde %14,71 oranında daha iyi performans göstererek, metinsel verileri etkili bir şekilde entegre etme yeteneğini göstermektedir. HERec ile karşılaştırıldığında, W-REHREC %60 eğitimde %18,68'lik ve %90 eğitimde %13,61'lik bir iyileştirme elde ederek, Konu-Bazlı Duygu Analizini dahil etmenin kritik rolünü vurgulamaktadır.

REHREC de iyi performans gösteriyor ve MAE değerleri %60 eğitimde 0,880'den %90 eğitimde 0,820'ye kadar değişiyor. NGCF ile karşılaştırıldığında REHREC, %60 eğitimde %8,33 ve %90 eğitimde %3,53 oranında göreceli bir iyileştirme gösteriyor ve yorum entegrasyonunun katma değerini vurguluyor. HERec'e göre iyileştirme %60 eğitimde %3,30 ve %90 eğitimde %2,27 olup, öneri doğruluğunu artırmada duygu analizinin önemini vurguluyor.

Yoruma dayalı olmayan yöntemler arasında NGCF, %60 eğitimde 0,960'tan %90'da 0,850'ye kadar değişen MAE değerleriyle en iyi performansı gösterir. Graf evrişim tabanlı yaklaşımı, HIN'deki yüksek dereceli ilişkileri yakalar ancak yorum analizi tarafından sağlanan anlamsal derinlikten yoksundur. MCF, %60 eğitimde 1,003 ve %90 eğitimde 0,901 MAE değerleriyle yakından takip eder. Metinsel verilerin eksikliği, kullanıcı-öge ilişkilerini REHREC ve W-REHREC kadar etkili bir şekilde modelleme yeteneğini sınırlar. NCF, MAE'nin %60 eğitimde 0,989'dan %90 eğitimde 0,882'ye düşmesiyle MCF'ten biraz daha iyi performans gösterir. Doğrusal olmayan etkileşimleri modellerken, metinsel verileri kullanamaması daha yüksek hatalara neden olur. Temel HIN tabanlı model olan HERec, %60 eğitimde 0,910'dan %90'da 0,839'a kadar değişen en yüksek MAE değerlerini kaydeder. Duygu analizinin olmaması, kullanıcı tercihlerinin karmaşıklığını yakalama yeteneğini sınırlar ve bu da daha az doğru tahminlere yol açar.

Yelp Veri Kümesi için RMSE sonuçları, modellerin etkinliği hakkında daha fazla içgörü sağlar. W-REHREC, %60 eğitimde 0,9358'den %90 eğitimde 0,9200'e kadar değişen değerlerle tüm eğitim oranlarında en düşük RMSE'yi elde eder. REHREC ile karşılaştırıldığında, W-REHREC, RMSE'yi %60 eğitimde %12,98 ve %90 eğitimde %12,38 oranında azaltır ve tahminlerdeki büyük sapmaları azaltmada ağırlıklı kenarların önemini vurgular. NGCF ile karşılaştırıldığında, W-REHREC, %60 eğitimde %14,68 ve %90 eğitimde %13,36 oranında bir iyileştirme elde eder ve yorum odaklı özelliklerin avantajlarını daha da vurgular. HERec ile karşılaştırıldığında, W-REHREC %60 eğitimde %18,54, %90 eğitimde ise %15,63 daha iyi performans göstererek, öneri kalitesini önemli ölçüde artırma yeteneğini ortaya koymaktadır.

REHREC ayrıca, %60 eğitimde 1.0754'ten %90'da 1.0500'e düşen değerlerle güçlü bir RMSE performansı göstermektedir. NGCF ile karşılaştırıldığında, REHREC %60 eğitimde %1.95 ve %90 eğitimde %0.99'lük bir göreceli iyileştirme elde ederek duygu analizini dahil etmenin faydalarını vurgulamaktadır. HERec ile karşılaştırıldığında, REHREC %60 eğitimde %6.38 ve %90 eğitimde %3.73'lük bir iyileştirme göstererek öneri sistemlerinde metinsel özelliklerin önemini doğrulamaktadır.

Yoruma dayalı olmayan yöntemler makul derecede iyi performans gösterir ancak yorum odaklı modellerin sürekli gerisinde kalır. NGCF, kullanıcı-öge ilişkilerini etkili bir şekilde yaymak için graf yapılarını kullanarak %60 eğitimde 1,0968'den %90'da 1,0619'a kadar değişen RMSE değerleri elde eder. MCRec, %60 eğitimde 1,1241'den %90'da 1,0823'e kadar değişen RMSE değerleri kaydeder. Yapısal ilişkileri yakalarken, metinsel verileri analiz edememesi nispeten daha yüksek hatalara yol açar. NCF, %60 eğitimde 1,0911'den %90'da 1,0556'ya kadar değişen RMSE değerleri sergiler. Derin öğrenmeye dayalı yaklaşımı, yorum entegrasyonunun eksikliğini kısmen telafi etse de REHREC ve W-REHREC'ten daha az etkili olmaya devam eder. HERec, temel model olarak, yorum verilerinin entegrasyonundan yoksun olduğu için %60 eğitimde 1,1488'den %90 eğitimde 1,0907'ye kadar değişen en yüksek RMSE değerlerini kaydeder.

Yelp Veri Kümesindeki sonuçlar, özellikle W-REHREC olmak üzere yorum odaklı modellerin doğru ve kişiselleştirilmiş öneriler sağlamada üstün performansını açıkça göstermektedir. Konu-Bazlı Duygu Analizi ve ağırlıklı kenar inşasını entegre ederek, W-REHREC hem REHREC hem de yorum tabanlı olmayan yöntemlere kıyasla sürekli olarak

daha düşük MAE ve RMSE elde eder. REHREC ayrıca yorum duygu verilerini dahil etmenin deęerini doęrularak önemli iyileştirmeler sergiler.

Bu bulgular, önerilen modellerin, özellikle W-REHREC'in, deęişen eğitim oranlarına ve veri kullanılabilirliğine ölçeklenebilirliğini ve uyarılana bilirliğini vurgular. Temel ve son teknoloji yöntemlere göre iyileştirmeler, yorum verilerinin ve gelişmiş HIN kelime gömmesi tekniklerinin modern öneri sistemlerine entegre edilmesinin önemini vurgular.

4.10. Sonuçlar ve Gözlemler: Foursquare Veri Kümesi

Foursquare Veri Kümesi, kullanıcı girişlerine ve özlü kullanıcı tarafından oluşturulan ipuçlarına odaklanan alana-özgü bir bağlamda önerilen modelleri değerlendirmek için benzersiz bir fırsat sunar. Bu özellikler, yapısal ve metinsel olmayan verilerin etkisini dengeleyerek duyguya duyarlı önerilerin etkisinin araştırılmasına olanak tanır. Etkililik deneylerinin sonuçları, bu veri kümesinde yorum odaklı tekniklerin, özellikle W-REHREC'in entegre edilmesinin avantajlarını vurgular.

Foursquare Veri Kümesi için MAE deęerleri, modeller arasında net bir performans hiyerarşisi gösteriyor ve W-REHREC tüm eğitim oranlarında en düşük hata oranlarına ulaşıyor. %60 eğitimde, W-REHREC 0,041'lik bir MAE elde ediyor ve bu %90'da 0,040'a çıkıyor. Temel REHREC ile karşılaştırıldığında, W-REHREC %60 eğitimde yaklaşık %10,87 ve %90 eğitimde %6,98'lik bir iyileştirme gösteriyor ve duygu odaklı bağlantıların etkisini artırmada ağırlıklı kenar yapısının deęerini gösteriyor. NGCF ile karşılaştırıldığında, W-REHREC %60 eğitimde %16,33 ve %90 eğitimde %11,11'lik bir iyileştirme elde ediyor ve bu da hem yapısal hem de metinsel özellikleri entegre etme konusundaki üstün yeteneğini yansıtıyor. HERec ile karşılaştırıldığında, W-REHREC %60 eğitimde %14,58 ve %90 eğitimde %9,09 gibi önemli bir oranda daha iyi performans gösteriyor ve bu da duygu analizi ve ağırlıklı avantajların dahil edilmesinin önemini vurguluyor.

REHREC ayrıca MAE deęerlerinin %60 eğitimde 0,046'dan %90'da 0,043'e düşmesiyle güçlü bir performans sergiliyor. NGCF ile karşılaştırıldığında, REHREC %60 eğitimde %6,12 ve %90 eğitimde %4,44'lük göreceli bir iyileştirme elde ederek duygu bütünleşmesinin deęerini vurguluyor. HERec'e göre iyileştirme %60 eğitimde %4,17 ve %90 eğitimde %2,27'dir ve bu da ağırlıklı kenarlar olmadan bile duygu analizinin tahmin doęruluęunu önemli ölçüde artırdığını göstermektedir.

Yoruma dayalı olmayan yöntemler arasında NGCF, %60 eğitimde 0,049'dan %90'da 0,045'e kadar değişen MAE değerleriyle en iyi performansı gösteren yöntem olarak öne çıkıyor. Graf teorisine dayalı yaklaşımı, Yorum odaklı ve geleneksel yöntemler arasındaki boşluğu kısmen kapatıyor. Ancak, duygu analizi eksikliği nedeniyle W-REHREC ve REHREC'in gerisinde kalıyor. HERec ile karşılaştırıldığında NGCF, %60 eğitimde %2,08 ve %90 eğitimde %2,27'lik bir iyileşme göstererek, graf-bazlı yayılma mekanizmasının etkinliğini kanıtıyor. MCREc, %60 eğitimde 0,0508 ve %90 eğitimde 0,047'lik MAE değerleriyle onu yakından takip ediyor. HERec ile karşılaştırıldığında MCREc, %60 eğitimde %4,17 ve %90 eğitimde %4,55'lik daha küçük bir iyileşme gösteriyor ve bu iyileşme metinsel verileri değerlendirmedeki yetersizliğiyle sınırlı. NCF, MAE değerlerinin %60 eğitimde 0,0500'den %90'da 0,047'ye düşmesiyle MCREc'e benzer şekilde performans gösterir. HERec ile karşılaştırıldığında NCF, %60 eğitimde %4,17 ve %90 eğitimde %4,55'lik bir iyileşme gösterir; bu, öncelikle doğrusal olmayan kullanıcı-öge etkileşimlerini modelleme becerisinden kaynaklanır. Temel HIN tabanlı model olan HERec, kullanıcı tarafından oluşturulan ipuçlarının entegrasyonundan yoksun olduğu için %60 eğitimde 0,048'den %90 eğitimde 0,044'e kadar değişen en yüksek MAE değerlerini sergiler.

Foursquare Veri Kümesi için RMSE sonuçları MAE'de gözlemlenen eğilimleri destekler. W-REHREC, %60 eğitimde 0,052'den %90 eğitimde 0,050'ye kadar değişen tüm eğitim oranlarında en düşük RMSE değerlerine ulaşır. REHREC ile karşılaştırıldığında, W-REHREC, RMSE'yi %60 eğitimde %13,33 ve %90 eğitimde %10,71 oranında iyileştirerek, ağırlıklı kenarların büyük sapmaları azaltmadaki etkisini gösterir. NGCF ile karşılaştırıldığında, W-REHREC, %60 eğitimde %16,13 ve %90 eğitimde %13,79 oranında iyileştirerek, hem yapısal hem de metinsel özellikleri entegre etmedeki üstün yeteneğini yansıtır. HERec ile karşılaştırıldığında, W-REHREC %60 eğitimde %17,46, %90 eğitimde ise %13,79 daha iyi performans gösteriyor ve bu da duyguya duyarlı uç yapının önemini daha da vurguluyor.

REHREC ayrıca RMSE değerlerinin %60 eğitimde 0,060'tan %90'da 0,056'ya kadar değişmesiyle iyi performans göstermektedir. NGCF ile karşılaştırıldığında, REHREC %60 eğitimde %3,23 ve %90 eğitimde %3,45 oranında göreceli bir iyileştirme elde ederek duygu bütünleşmesinin değerini vurgulamaktadır. HERec ile karşılaştırıldığında, REHREC %60 eğitimde %4,76 ve %90 eğitimde %7,14 oranında bir iyileştirme göstererek HIN'lere metinsel özelliklerin dahil edilmesinin önemini doğrulamaktadır.

Foursquare Veri Kümesi'ndeki sonuçlar birkaç önemli içgörüyü vurgulamaktadır. W-REHREC, hem MAE hem de RMSE'de diğer tüm yöntemlerden sürekli olarak daha iyi performans göstererek, Konu-Bazlı Duygu Analizi ve ağırlıklı kenar oluşturmayı birleştirmenin önemli faydalarını göstermektedir. Kullanıcı tarafından oluşturulan ipuçlarından yararlanarak, W-REHREC kullanıcı tercihlerinin nüanslarını yakalar ve bu da gelişmiş tahmin doğruluğu ile sonuçlanır. REHREC ayrıca güçlü bir performans göstererek, HIN tabanlı öneri sistemlerine duygu bilgilerini eklemenin değerini vurgular. Ancak, yorum odaklı kenarları ağırlıklandıramaması, W-REHREC'e kıyasla genel etkinliğini sınırlar.

Yoruma dayalı olmayan yöntemler rekabetçi olsa da, yorum odaklı modellerin elde ettiği performansın sürekli gerisinde kalmaktadır. NGCF ve MCF, HIN içindeki yapısal ilişkileri yakalama yetenekleri nedeniyle geleneksel işbirlikçi filtreleme yaklaşımlarından daha iyi performans göstermektedir. Ancak, duygu analizinin olmaması, kullanıcı tercihlerini tam olarak yakalama yeteneklerini sınırlamaktadır. Temel model olarak HERec, metinsel verilerin entegre edilmesinin önemini göstermektedir, çünkü hariç tutulması sürekli olarak daha yüksek hatalara yol açmaktadır.

Bulgular, kullanıcı ipuçları gibi alana-öзgü verileri kullanarak öneri doğruluğunu artırmada önerilen modellerin etkinliğini doğrulamaktadır. W-REHREC'in tüm eğitim oranlarındaki tutarlı performansı, sınırlı metinsel bilgiye sahip veri kümelerine ölçeklenebilirliğini ve uyarlanabilirliğini sergilemektedir. Bu sonuçlar, özellikle yapısal ve metinsel özellikleri dengeleyen Foursquare gibi veri kümeleri için öneri sistemlerine duygu analizi ve ağırlıklı uç oluşturmayı dahil etmenin önemini vurgulamaktadır.

4.11. Sonuçlar ve Gözlemler: Meta-Yollar

REHREC ve W-REHREC'in meta-yol odaklı modellerini tanıttıktan sonra, şimdi sonuçların ayrıntılı bir analizine ve tartışmasına geçiyoruz. Bu bölüm, önerilen modellerin genel performansı üzerinde bireysel incelemeye dahil edilen meta-yolların etkisini, özellikle de öneri doğruluğuna ve iş akışı etkinliğine katkılarına odaklanarak inceler. Her meta-yol, anlamlı ilişkiler çıkarmak için kullanıcıları, işletmeleri ve incelemeleri benzersiz şekillerde birbirine bağlayan Heterojen Bilgi Ağı (HIN) içinde farklı bir bakış açısını temsil eder.

URRU, BRRB, UBRRBU ve BURRUB dahil olmak üzere bu bölümde analiz edilen meta-yollar, yerleştirme sürecini yönlendiren ve öneri çıktılarını etkileyen temel bileşenler olarak hizmet eder. Ağdaki belirli rotaları geçerek, bu meta-yollar kullanıcılar arasındaki paylaşılan

incelemeler veya işletmeler arasındaki benzer nitelikler gibi hem doğrudan hem de dolaylı etkileşimleri ortaya çıkarır. Her meta-yol, paylaşılan duygular, kategori tabanlı benzerlikler veya işbirlikçi davranışlar gibi kullanıcı-öge ilişkisinin farklı yönlerini vurgulamak için dikkatlice tasarlanmıştır. Bu analiz, her meta-yolun öneri sürecine nasıl katkıda bulunduğu dair kapsamlı bir anlayış sağlamayı ve anlamsal, yapısal veya örtük ilişkileri yakalama becerisine odaklanmayı amaçlamaktadır. Bu meta-yolların veri kümeleri (Yelp ve Foursquare) genelinde sunduğu performans iyileştirmelerini değerlendireceğiz ve hem REHREC hem de W-REHREC modellerini geliştirmedeki rollerini vurgulayacağız. Bu meta-yolların, inceleme açısından zengin Yelp veri kümesi gibi değişken özelliklere sahip veri kümeleri genelinde modellerin genelleme yapma becerisini nasıl etkilediğine özel dikkat gösterilmektedir.

Aşağıdaki bölümler, genel model ve öneri doğruluğu üzerindeki belirli etkilerini analiz etmek için "URRU" meta-yoluyla başlayarak bireysel meta-yolların ayrıntılı tartışmasıyla başlar. Bu sistematik yaklaşım, her meta-yolun güçlü ve zayıf yönlerine ilişkin içgörüler sağlar ve modern öneri sistemlerindeki önemlerini belirler.

4.11.1. "URRU" Meta-Yolunun Veri Kümeleri Üzerindeki Etkisi

"URRU" meta yolu (Kullanıcı-Yorum-Yorum-Kullanıcı), Heterojen Bilgi Ağı (HIN) içindeki paylaşılan yorumlar veya etkileşimler aracılığıyla kullanıcıdan kullanıcıya ilişkileri yakalar. Bu meta yol, özellikle paylaşılan görüşlere ve deneyimlere dayalı kullanıcı benzerliğini belirlemede etkilidir ve bu da onu Yelp gibi veri kümelerinde değerli bir katkı haline getirir. Bu bölüm, "URRU" meta yolunun hem REHREC hem de W-REHREC modelleri üzerindeki etkisini, HERec makalesindeki temel modele kıyasla performanslarına odaklanarak Yelp ve Foursquare veri kümeleri genelinde değerlendirir.

Yelp Veri Kümesi Analizi:

Ayrıntılı kullanıcı yorumlarıyla zenginleştirilmiş Yelp Veri Kümesi, "URRU" gibi yorum tabanlı meta-yolların potansiyelini keşfetmek için oldukça uygundur. Kullanıcıları paylaşılan yorumlar aracılığıyla birbirine bağlayan bu meta-yol, modellerin ayrıntılı kullanıcı tercihlerini yakalamasını ve öneri doğruluğunu iyileştirmesini sağlar.

REHREC modelinde, "URRU" meta-yoluyla RMSE, temel HERec modelindeki 1.0907'ye kıyasla 1.080'dir. Bu, %0.89'luk bir iyileştirmeyi temsil eder ve yorum aracılı kullanıcı

bağlantılarını dahil etmenin tahmin performansını önemli ölçüde artırdığını gösterir. "URRU" meta-yolunun eklenmesi, REHREC'in kullanıcılar arasındaki anlamsal ilişkilerden yararlanmasını sağlayarak temel modele kıyasla tahmin hatalarını azaltır.

W-REHREC'te RMSE, temel HERec modelindeki 1,0907'ye kıyasla 0,9463'e ulaşarak önemli ölçüde iyileşir. %13,26'lık bu iyileşme, daha güçlü yorum tabanlı ilişkilerin etkisini artıran ağırlıklı kenar yapısının önemini vurgular. Konu-Bazlı Duygu Analizini ve ağırlıklandırılmış yorum bağlantılarını entegre ederek, W-REHREC, temel çizgiyi önemli ölçüde geride bırakarak, kritik yorum odaklı bilgileri çıkarma ve kullanma yeteneğini gösterir.

Hem REHREC hem de W-REHREC, HERec temel çizgisine göre önemli iyileştirmeler göstermektedir. W-REHREC'teki REHREC'e kıyasla ek iyileştirme, bol miktarda metinsel bilgiye sahip veri kümelerinde kritik olan daha etkili yorumlara öncelik vermede ağırlıklı kenarların değerini vurgular.

Foursquare Veri Kümesi Analizi:

Açık yorumlardan yoksun olan Foursquare Veri Kümesi, "URRU" meta-yolunun genelleme yeteneğini test eder. Bu bağlamda, meta-yol, yorumlar yerine girişler veya ipuçları gibi paylaşılan etkileşimler aracılığıyla kullanıcıdan kullanıcıya ilişkileri yakalar. Metinsel verilerin yokluğuna rağmen, "URRU" meta-yolu, öneri doğruluğunu iyileştirmeye anlamlı bir şekilde katkıda bulunur.

REHREC'te, "URRU" meta-yoluyla RMSE, temel HERec modelindeki 0,0582'ye kıyasla 0,0576'dır. Bu, %1,03'lük bir iyileştirmeyi temsil eder ve modelin yorumdan esinlenen meta-yolu metin dışı veri kümelerine genelleştirme yeteneğini gösterir. REHREC, örtüşen giriş konumları veya zamansal veriler gibi paylaşılan davranışlar aracılığıyla kullanıcı benzerliğini modelleyerek, temel çizgiye göre tahmin hatalarını azaltır.

W-REHREC, temel HERec modelindeki 0,0582'ye kıyasla "URRU" meta-yoluyla 0,0514'lük bir RMSE elde eder. %11,68'lik bu iyileştirme, modelin açık yorum verileri olmadan bile daha güçlü kullanıcı bağlantılarını önceliklendirme yeteneğini vurgular. Ağırlıklı kenar yapısı, W-REHREC'in giriş bilgilerinde sık sık birlikte bulunma veya paylaşılan konum tercihleri gibi örtük sinyalleri etkili bir şekilde kullanmasına olanak tanır.

Hem REHREC hem de W-REHREC, HERec temel çizgisine göre belirgin iyileştirmeler göstermektedir. "URRU" meta-yolu, yorumların yokluğunda bile kullanıcı benzerliklerinin etkili bir şekilde modellenmesine katkıda bulunur. W-REHREC, en alakalı kullanıcı

etkileşimlerini vurgulayan ağırlıklı bağlantıları dahil ederek REHREC'i sürekli olarak geride bırakır ve bu da RMSE'de önemli bir azalmayla sonuçlanır.

Veri Kümeleri Arası Gözlemler:

Yelp ve Foursquare veri kümelerindeki "URRU" meta-yolunun sonuçları, uyarılana bilirliliğini ve etkinliğini vurgular. Yorumların bol olduğu Yelp veri kümesinde, meta-yol, paylaşılan yorum içeriği aracılığıyla kullanıcılar arasındaki anlamsal ilişkileri yakalayarak öneri doğruluğunu artırır. Temel HREc modeliyle karşılaştırıldığında, REHREC ve W-REHREC sırasıyla %0,89 ve %13,26 oranında iyileştirmeler elde ederek yorum odaklı bağlantıların etkisini vurgular. Özellikle W-REHREC, daha güçlü ilişkilere öncelik vermek için ağırlıklı kenarları kullanır ve bu da önemli doğruluk kazanımlarıyla sonuçlanır.

Yorumların bulunmadığı Foursquare veri kümesinde, "URRU" meta yolu, paylaşılan etkileşimler aracılığıyla örtük kullanıcı ilişkilerini yakalamak için geliştirilir. Temel HREc modeliyle karşılaştırıldığında, REHREC ve W-REHREC sırasıyla %1,03 ve %11,68 oranında iyileştirmeler elde ederek, yorumdan esinlenen teknikleri metin dışı bağlamlara uyarılama becerilerini gösterir. W-REHREC'in üstün performansı, en etkili kullanıcı davranışlarını vurgulamada ağırlıklı bağlantıların önemini vurgular.

"URRU" meta-yolu, hem yorum açısından zengin hem de yorum açısından eksik veri kümelerinde öneri doğruluğunu iyileştirmede oldukça etkilidir. Yelp veri kümesinde, anlamsal bağlantıları ortaya çıkarmak için yorum içeriğinden yararlanır ve bu da temel HREc modeline kıyasla önemli RMSE azalmalarıyla sonuçlanır. Foursquare veri kümesinde, örtük kullanıcı ilişkilerini modellemek için etkili bir şekilde geliştirir ve sağlamlığını ve çok yönlülüğünü sergiler. W-REHREC tarafından elde edilen önemli iyileştirmeler, özellikle temel çizgiye kıyasla, modern öneri sistemlerinde ağırlıklı kenarların ve gelişmiş meta-yol odaklı kelime gömmelerin önemini doğrular.

4.11.2. "BRRB" Meta-Yolunun Veri Kümeleri Üzerindeki Etkisi

"BRRB" meta-yolu (Mekan-Yorum-Yorum-Mekan), Heterojen Bilgi Ağı'ndaki (HIN) paylaşılan yorumlar aracılığıyla mekanlar arasındaki ilişkileri yakalamada önemli bir rol oynar. Bu meta-yol, öneri modellerinin, Yelp gibi metinsel yorumlarla zenginleştirilmiş veri kümelerinde özellikle yararlı olan, ortak kullanıcı görüşleri ve geri bildirimlerine dayalı olarak mekanlar arasında benzerlikler oluşturmasına olanak tanır. Bu bölüm, "BRRB" meta-

yolunun hem REHREC hem de W-REHREC modelleri üzerindeki etkisini Yelp ve Foursquare veri kümeleri genelinde değerlendirir ve HERec makalesindeki temel modele göre performanslarına odaklanır.

Yelp Veri Kümesi Analizi:

Yelp Veri Kümesi, kullanıcılar tarafından oluşturulan çok sayıda yorumla, "BRRB" gibi yorum tabanlı meta-yolların potansiyelini keşfetmek için mükemmel bir ortam sağlar. Bu meta-yol, mekanları paylaşılan yorumlar aracılığıyla birbirine bağlayarak, modellerin gizli ilişkileri ortaya çıkarma ve öneri doğruluğunu iyileştirme yeteneğini artırır.

REHREC modelinde, "BRRB" meta-yoluyla RMSE, temel HERec modelindeki 1.0907'ye kıyasla 1.090'dır. Bu, %0,06'lık marjinal bir iyileştirmeyi temsil eder ve "BRRB" meta-yolu değer katarken, REHREC'teki ağırlıklı kenarların eksikliğinin genel etkisini sınırladığını gösterir. İşletmelerin paylaşılan yorumlar aracılığıyla bağlanması, REHREC'in bazı anlamsal benzerliklerden yararlanmasına olanak tanır, ancak etkisi diğer meta-yollara kıyasla daha az belirgindir.

W-REHREC'te RMSE, temel HERec modelindeki 1,0907'ye kıyasla 0,9550'ye ulaşarak önemli ölçüde iyileşir. %12,47'lik bu iyileştirme, daha güçlü yorum bağlantılarının etkisini artırmada ağırlıklı kenar yapısının önemini vurgular. W-REHREC, daha etkili yorumlara etkili bir şekilde öncelik vererek mekanlar arasındaki ilişkilerin daha derin bir şekilde anlaşılmasını sağlar ve bu da önemli ölçüde daha iyi öneri performansına yol açar.

Hem REHREC hem de W-REHREC, HERec temel çizgisinden daha iyi performans gösterir, ancak aralarındaki fark, ağırlıklı kenarları dahil etmenin avantajını vurgular. W-REHREC'in kritik yorumları yükseltme yeteneği, önemli ölçüde iyileştirilmiş doğrulukla sonuçlanır ve bu da onu Yelp gibi yorum açısından zengin veri kümelerinde "BRRB" meta-yolunu kullanmak için üstün bir seçim haline getirir.

Foursquare Veri Kümesi Analizi:

Açık yorumlardan yoksun olan Foursquare Veri Kümesi, modelleri "BRRB" gibi yorumdan esinlenen meta-yolları metinsel olmayan verilere genelleştirmeye zorlar. Bu bağlamda, "BRRB" meta-yolu, örtüşen girişler veya benzer etkinlik kalıpları gibi paylaşılan kullanıcı etkileşimlerine dayalı olarak mekanları birbirine bağlar.

REHREC'te, "BRRB" meta-yoluyla RMSE, temel HERec modelindeki 0,0582'ye kıyasla 0,0581'dir. Bu, %0,17'lik mütevazı bir iyileştirmeyi temsil eder ve modelin meta-yolu bazı

örtük iş ilişkilerini yakalamak için genelleştirebileceğini gösterir. Bu veri kümesindeki "BRRB" meta-yolu, paylaşılan kullanıcı davranışına dayalı olarak mekanlar arasındaki ortak noktaları vurgular, ancak açık yorumların eksikliği potansiyelini sınırlar.

W-REHREC, "BRRB" meta-yoluyla 0,0519'luk bir RMSE elde ederken, temel HERec modelindeki 0,0582'ye kıyasla. %10,82'lik bu iyileştirme, modelin ağırlıklı yorumdan esinlenen stratejileri metinsel verilerden yoksun veri kümelerine uyarlama yeteneğini vurgular. Daha güçlü bağlantıları vurgulayarak, W-REHREC, yorumun olmadığı bir senaryoda bile anlamlı iş ilişkilerini etkili bir şekilde yakalar.

Hem REHREC hem de W-REHREC, HERec temel çizgisine göre iyileştirmeler gösterirken, W-REHREC daha belirgin bir avantaj sergiliyor. W-REHREC'in "BRRB" meta-yolunu örtük etkileşimlere genelleştirme yeteneği, sağlamlığını ve uyarlanabilirliğini vurgular.

Veri Kümeleri Arası Gözlemler:

Yelp ve Foursquare veri kümelerindeki "BRRB" meta-yolunun sonuçları, öneri sistemlerindeki önemini ve çok yönlülüğünü göstermektedir. Yelp veri kümesinde, meta-yol, mekanları paylaşılan yorumlar aracılığıyla birbirine bağlayarak, öneri doğruluğunu artıran anlamsal ilişkileri ortaya çıkarır. Temel HERec modeliyle karşılaştırıldığında, REHREC ve W-REHREC sırasıyla %0,06 ve %12,47 oranında RMSE iyileştirmeleri elde ederek, yorum odaklı bağlantıların etkisini artırmada ağırlıklı kenarların kritik rolünü göstermektedir.

Foursquare veri kümesinde, "BRRB" meta-yolu, paylaşılan kullanıcı etkinliğine dayalı olarak mekanlar arasındaki örtük ilişkileri yakalamak için genelleştirilir. Temel HERec modeliyle karşılaştırıldığında, REHREC ve W-REHREC sırasıyla %0,17 ve %10,82 oranında RMSE iyileştirmeleri elde eder. W-REHREC'in üstün performansı, açık yorumların olmadığı durumlarda bile daha alakalı bağlantılara öncelik vermenin önemini vurgular.

"BRRB" meta-yolu, hem yorum açısından zengin hem de yorum açısından yoksun veri kümelerinde öneri doğruluğunu iyileştirmede etkilidir. Yelp veri kümesinde, mekanlar arasında anlamsal ilişkiler kurmak için yorum içeriğinden yararlanır ve bu da temel HERec modeline kıyasla W-REHREC'te önemli iyileştirmeler sağlar. Foursquare veri kümesinde, örtük iş ilişkilerini yakalamak için etkili bir şekilde genelleştirir ve metin dışı bağlamlarda sağlamlığını gösterir. W-REHREC tarafından elde edilen önemli iyileştirmeler, ağırlıklı

kenarların ve gelişmiş meta-yol odaklı kelime gömmelerin önemini doğrular ve modern öneri sistemlerindeki rolünü sağlamlaştırır.

4.11.3. "UBRRBU" Meta-Yolunun Veri Kümeleri Üzerindeki Etkisi

"UBRRBU" meta-yolu (Kullanıcı-Mekan-Yorum-Yorum-Mekan-Kullanıcı), hem doğrudan kullanıcı-mekan bağlantılarını hem de paylaşılan yorumları geçerek kullanıcılar ve mekanlar arasındaki karmaşık ilişkileri yakalar. Bu meta-yol, paylaşılan kullanıcı deneyimleri ve görüşleri aracılığıyla kullanıcı ve mekan varlıklarını birbirine bağlamada özellikle etkilidir. Öneri modellerinin karmaşık ilişkisel kalıpları ortaya çıkarmasına olanak tanır ve bu da onu Yelp gibi yorum açısından zengin veri kümelerinde değerli bir katkı haline getirir. Bu bölüm, "UBRRBU" meta-yolunun hem REHREC hem de W-REHREC modelleri üzerindeki etkisini Yelp ve Foursquare veri kümelerinde değerlendirir ve bunların performansını HERec makalesindeki temel modelle karşılaştırır.

Yelp Veri Kümesi Analizi:

Ayrıntılı yorumlarla zenginleştirilmiş Yelp Veri Kümesi, "UBRRBU" gibi yorum merkezli meta-yolların potansiyelinden yararlanmak için ideal senaryoyu sunar. Bu meta-yol, yalnızca doğrudan kullanıcı-mekan etkileşimlerini değil, aynı zamanda paylaşılan yorum içeriği aracılığıyla oluşan dolaylı ilişkileri de yakalayarak, gelişmiş öneri doğruluğuna yol açar.

REHREC modelinde, "UBRRBU" meta-yoluyla RMSE, temel HERec modelindeki 1.0907'ye kıyasla 1.060'tır. Bu, %2.81'lik bir iyileştirmeyi temsil eder ve bu karmaşık yorum aracılı meta-yolun dahil edilmesinin REHREC'in kullanıcılar ve mekanlar arasındaki nüanslı ilişkileri daha iyi yakalamasını sağladığını gösterir. "UBRRBU" meta-yolunu dahil ederek, REHREC dolaylı bağlantılardan ve paylaşılan yorum bağlamlarından yararlanır ve temel modele kıyasla tahmin hatalarını azaltır.

W-REHREC'te RMSE, temel HERec modelindeki 1,0907'ye kıyasla 0,9288'e daha da iyileşiyor. %14,86'lık bu önemli iyileştirme, daha alakalı yorumları ve etkileşimleri vurgulayan ağırlıklı kenar yapısının katma değerini vurgular. W-REHREC, önemli bağlantıları önceliklendirmek için Konu-Bazlı Duygu Analizi ve ağırlıklı yorumlardan yararlanarak karmaşık kullanıcı-öge ilişkilerini yakalama yeteneğini artırır.

Hem REHREC hem de W-REHREC, HERec temel çizgisini geride bırakıyor ve W-REHREC, ağırlıklı kenar yapısı nedeniyle dikkate değer bir avantaj gösteriyor. Her iki modele de "UBRRBU" meta yolunun dahil edilmesi, öneri doğruluğunu iyileştirmek için ayrıntılı yorum verilerinden yararlanmadaki etkinliğini vurguluyor.

Foursquare Veri Kümesi Analizi:

Açık yorum verilerinden yoksun olan Foursquare Veri Kümesi, "UBRRBU" gibi yorumdan esinlenen meta-yolların örtük kullanıcı-öge etkileşimlerinin hakim olduğu veri kümelerine uyarlanabilirliğini test eder. Bu bağlamda, meta yol kullanıcıları ve mekanları Yorum içeriği yerine girişler veya eş zamanlı oluşumlar gibi paylaşılan etkinlik kalıpları aracılığıyla birbirine bağlar.

REHREC'te, "UBRRBU" meta-yoluyla RMSE, temel HERec modelindeki 0,0582'ye kıyasla 0,0565'tir. Bu, %2,93'lük bir iyileştirmeyi temsil eder ve modelin, paylaşılan kullanıcı-öge ilişkilerini etkili bir şekilde belirlemek için yorumdan esinlenen meta-yolu genelleştirebileceğini gösterir. Metinsel veriler olmadan bile, "UBRRBU" meta-yolu, kullanıcıları ve mekanları örtük davranış kalıpları aracılığıyla birbirine bağlayarak iyileştirilmiş doğruluğa katkıda bulunur.

W-REHREC, "UBRRBU" meta-yoluyla 0,0505'lik bir RMSE elde ederken, temel HERec modelindeki 0,0582'lik değere kıyasla. %13,22'lik bu iyileştirme, modelin ağırlıklı yorumdan esinlenen teknikleri açık metinsel bilgi içermeyen veri kümelerine uyarlama yeteneğini vurgular. Daha güçlü kullanıcı-öge bağlantıları vurgulayarak, W-REHREC, "UBRRBU" meta-yolu tarafından yakalanan örtük sinyalleri etkili bir şekilde kullanır.

Hem REHREC hem de W-REHREC, temel HERec modeline göre iyileştirmeler gösterirken, W-REHREC önemli bir avantaj göstermektedir. "UBRRBU" meta yolunu yorum dışı bağlamlara genelleştirme yeteneği, önerilen modellerin sağlamlığını ve çok yönlülüğünü vurgulamaktadır.

Veri Kümeleri Arası Gözlemler:

Yelp ve Foursquare veri kümelerindeki "UBRRBU" meta-yolunun sonuçları, öneri doğruluğunu artırmadaki önemini vurgular. Yelp veri kümesinde, meta-yol, kullanıcıları ve mekanları paylaşılan yorumlar aracılığıyla birbirine bağlayarak doğruluğu önemli ölçüde artıran karmaşık ilişkileri ortaya çıkarır. Temel HERec modeliyle karşılaştırıldığında, REHREC ve W-REHREC sırasıyla %2,81 ve %14,86 oranında RMSE iyileştirmeleri elde

ederek, ağırlıklı kenarların yorum odaklı bağlantıların etkisini artırmadaki kritik rolünü gösterir.

Yorumların bulunmadığı Foursquare veri kümesinde, "UBRRBU" meta-yolu, paylaşılan etkinlik kalıpları aracılığıyla örtük kullanıcı-öge ilişkilerini yakalamak için etkili bir şekilde genelleştirilir. Temel HERec modeliyle karşılaştırıldığında, REHREC ve W-REHREC sırasıyla %2,93 ve %13,22 oranında RMSE iyileştirmeleri elde eder. W-REHREC'in üstün performansı, açık yorumların olmadığı durumlarda bile en etkili etkileşimleri vurgulamada ağırlıklı bağlantıların önemini vurgular.

"UBRRBU" meta-yolu, çeşitli veri kümeleri arasında öneri doğruluğunu iyileştirmek için güçlü bir araçtır. Yelp veri kümesinde, karmaşık kullanıcı-öge ilişkilerini yakalamak için ayrıntılı yorum içeriğinden yararlanır ve bu da temel HERec modeline kıyasla önemli iyileştirmeler sağlar. Foursquare veri kümesinde, örtük bağlantıları tanımlamak için etkili bir şekilde genelleştirir ve metin dışı bağlamlarda sağlamlığını gösterir. W-REHREC tarafından elde edilen önemli iyileştirmeler, özellikle temele kıyasla, modern öneri sistemlerinde ağırlıklı kenarların ve gelişmiş meta-yol odaklı kelime gömmelerin önemini doğrular.

4.11.4. "BURRUB" Meta-Yolunun Veri Kümeleri Üzerindeki Etkisi

"BURRUB" meta-yolu (Mekan-Kullanıcı-Yorum-Yorum-Kullanıcı-Mekan), hem doğrudan mekan-kullanıcı bağlantılarını hem de paylaşılan yorumları geçerek mekanlar ve kullanıcılar arasındaki karmaşık ve çok düzeyli ilişkileri yakalar. Bu meta-yolu, karşılıklı yorumlar ve paylaşılan deneyimler aracılığıyla kullanıcılar ve mekanlar arasındaki etkileşimin kapsamlı bir temsilini sağlar. Yelp gibi yorum açısından zengin veri kümelerinde özellikle değerlidir, çünkü modellerin nüanslı ilişkileri ortaya çıkarmasına ve öneri doğruluğunu iyileştirmesine olanak tanır. Bu bölüm, "BURRUB" meta yolunun hem REHREC hem de W-REHREC modelleri üzerindeki etkisini Yelp ve Foursquare veri kümelerinde değerlendirir ve HERec makalesindeki temel modele kıyasla performanslarına odaklanır.

Yelp Veri Kümesi Analizi:

Kapsamlı kullanıcı yorumlarıyla zenginleştirilmiş Yelp Veri Kümesi, "BURRUB" meta-yolunun potansiyelini keşfetmek için ideal bir test ortamı sağlar. İşletmeleri ve kullanıcıları

paylaşılan yorumların birden fazla katmanıyla birbirine bağlayarak, bu meta-yol modellerin karmaşık ilişkileri daha iyi yakalamasını ve tahmin doğruluğunu iyileştirmesini sağlar.

REHREC modelinde, "BURRUB" meta-yoluyla RMSE, temel HERec modelindeki 1.0907'ye kıyasla 1.140'tır. Şaşırtıcı bir şekilde, bu RMSE'de marjinal bir artışa neden olur ve bu da REHREC'in ağırlıklı kenarların desteği olmadan bu meta-yolun getirdiği ek karmaşıklıkta zorlandığını gösterir. REHREC'te ağırlıklı yorum bağlantılarının olmaması, kritik ilişkilerin alaka düzeyini zayıflatabilir ve bu da doğruluğun azalmasına yol açabilir.

W-REHREC'te RMSE, temel HERec modelindeki 1.0907'ye kıyasla 1.099'a ulaşarak önemli ölçüde iyileşir. %0,08'lik iyileşme mütevazı olsa da, daha karmaşık meta-yolların oluşturduğu zorlukları hafifletmede ağırlıklı uç yapısının önemini vurgular. W-REHREC, daha güçlü anlamsal veya duygu bağlantıları olan yorumlara öncelik vererek, "BURRUB" meta-yolunu REHREC'ten daha etkili bir şekilde kullanmasına olanak tanır.

Karşılaştırmalı analiz, W-REHREC'in etkili yorumlara daha fazla önem verme yeteneği sayesinde "BURRUB" meta-yolunun karmaşıklığını ele almak için daha donanımlı olduğunu ortaya koyuyor. Performans kazanımı daha basit meta-yollara kıyasla daha az belirgin olsa da, sonuçlar bu meta yolun potansiyelini tam olarak gerçekleştirmede ağırlıklı kenarların gerekliliğini vurguluyor.

Foursquare Veri Kümesi Analizi:

Açık yorumlardan yoksun olan Foursquare Veri Kümesi, "BURRUB" meta-yolunun yorum açısından zengin ortamların ötesinde genelleme yapma yeteneğini test eder. Bu veri kümesinde, meta-yol mekanları ve kullanıcıları yorumlar yerine paylaşılan girişler veya benzer etkinlik kalıpları aracılığıyla birbirine bağlar.

REHREC'te, "BURRUB" meta-yoluyla RMSE, HERec temel modelindeki 0,0582'ye kıyasla 0,0567'dir. Bu, %2,58'lik bir iyileştirmeyi temsil eder ve modelin örtük kullanıcı-öge ilişkilerini etkili bir şekilde belirlemek için yorumdan esinlenen "BURRUB" meta-yolunu genelleştirebileceğini gösterir. İşletmeleri ve kullanıcıları paylaşılan etkileşimler aracılığıyla birbirine bağlayarak, meta-yol metinsel veriler olmadan bile tahmin hatalarını azaltır.

W-REHREC, "BURRUB" meta-yoluyla 0,0510'luk bir RMSE elde ederken, temel HERec modelindeki 0,0582'lik değere kıyasla bunu başarıyor. %12,37'lik bu iyileştirme, modelin yapısal ilişkilerle dolu veri kümelerine yorumdan esinlenen teknikleri uyarlama yeteneğini

vurguluyor. W-REHREC, "BURRUB" meta-yolundaki daha etkili bağlantılara öncelik vererek, yorumların olmamasına rağmen yüksek doğruluğu korumasını sağlıyor.

REHREC ile W-REHREC arasındaki karşılaştırma, ikincisinin "BURRUB" meta-yolunun önemini artırmak için ağırlıklı kenarları kullanma yeteneğini vurgular ve bu da Foursquare veri kümesinde RMSE'de önemli bir azalmaya yol açar.

Veri Kümeleri Arası Gözlemler:

Yelp ve Foursquare veri kümelerindeki "BURRUB" meta-yolunun sonuçları, potansiyelini ve zorluklarını vurgulamaktadır. Yelp veri kümesinde, meta-yol, mekanları ve kullanıcıları paylaşılan yorumlar aracılığıyla birbirine bağlayarak zengin bir anlamsal ilişki kaynağı sağlar. Ancak, REHREC, ağırlıklı kenarlar olmadan bu meta-yolu etkili bir şekilde kullanmakta zorlanır ve bu da HERec modelinin başlangıç düzeyine kıyasla RMSE'de hafif bir artışla sonuçlanır. Buna karşılık, W-REHREC, %0,08'lik bir iyileştirme göstererek, karmaşık meta-yollardan anlamlı içgörüler çıkarmada ağırlıklı kenarların önemini gösterir.

Yorumların bulunmadığı Foursquare veri kümesinde, "BURRUB" meta-yolu, paylaşılan etkileşimler aracılığıyla örtük kullanıcı-öge ilişkilerini yakalamak için etkili bir şekilde geliştirilir. REHREC ve W-REHREC, temel HERec modeline kıyasla sırasıyla %2,58 ve %12,37 oranında RMSE iyileştirmeleri elde eder. W-REHREC'in üstün performansı, metinsel veri içermeyen veri kümelerinde bile en etkili kullanıcı-öge etkileşimlerini vurgulamada ağırlıklı bağlantıların önemini vurgular.

"BURRUB" meta-yolu, hem yorum açısından zengin hem de yorum açısından yoksun veri kümelerinde çok düzeyli kullanıcı-öge ilişkilerini yakalamak için değerli bir araçtır. Yelp veri kümesinde, ağırlıklı kenarların olmaması nedeniyle REHREC'te etkinliği sınırlıdır; oysa W-REHREC, etkili yorumlara öncelik vererek mütevazı iyileştirmeler göstermektedir. Foursquare veri kümesinde, meta-yol örtük etkileşimlere iyi bir şekilde geliştirilerek her iki modelin de doğruluğu artırmasını sağlar. W-REHREC tarafından elde edilen önemli iyileştirme, özellikle temel HERec modeliyle karşılaştırıldığında, ağırlıklı kenarların ve gelişmiş meta-yol odaklı kelime gömmelerin modern öneri sistemlerinde, özellikle "BURRUB" gibi karmaşık meta-yollar için gerekliliğini doğrulamaktadır.

4.11.5. Veri Kümeleri Arasında Birleştirilmiş Meta-Yolların Etkisi

Birleşik meta-yollar (UBU, BUB, URRU, BRRB, UBRRBU, UBCaBU, BCaB) çeşitli ilişki kümelerini kapsar ve yorumlar ve etkileşimler aracılığıyla kullanıcı-kullanıcı, kullanıcı-mekan ve mekan-mekan bağlantılarını yakalar. Bu kombinasyon, öneri modellerinin kullanıcı-öge ilişkilerinin birden fazla bakış açısını entegre etmesine olanak tanır ve altta yatan Heterojen Bilgi Ağı'nın (HIN) kapsamlı bir şekilde anlaşılmasını sağlar. Bu meta-yolların dahil edilmesi, modellerin anlamsal ve yapısal ilişkileri ortaya çıkarma kapasitesini güçlendirir ve bunları Yelp ve Foursquare gibi veri kümelerindeki öneri görevleri için özellikle etkili hale getirir. Bu bölüm, bu birleşik meta-yolların Yelp ve Foursquare veri kümelerindeki hem REHREC hem de W-REHREC modelleri üzerindeki etkisini değerlendirir ve temel HERec modeline göre performanslarına odaklanır.

Yelp Veri Kümesi Analizi:

Kapsamlı yorumlarla zenginleştirilmiş Yelp Veri Kümesi, birleşik meta-yolların tüm gücünden yararlanmak için ideal bir senaryo sunar. Çeşitli ilişki düzeylerini yakalayarak, meta-yollar kullanıcı görüşlerini, mekan özelliklerini ve etkileşimlerini HIN'e entegre ederek öneri performansını önemli ölçüde artırır.

REHREC modelinde, birleşik meta-yollarla RMSE, temel HERec modelindeki 1.0907'ye kıyasla 1.050'dir. Bu, kullanıcı-öge ilişkilerinin daha bütünsel bir konu ünü yakalamak için çeşitli meta-yolları birleştirmenin faydasını gösteren %3,73'lük bir iyileştirmeyi temsil eder. REHREC, URRU ve UBRRBU gibi meta-yolları dahil ederek, kullanıcılar ve mekanlar arasındaki anlamsal bağlantıları ve paylaşılan yorumlarla aracılık edilen dolaylı ilişkileri dahil edebilir. UBCaBU ve BCaB'nin dahil edilmesi, modelin iş niteliklerini ve kategorilerini de hesaba katabilmesini sağlayarak önerilerine derinlik katar.

W-REHREC, temel HERec modelindeki 1,0907'ye kıyasla 0,9200'lük bir RMSE ile önemli bir iyileştirme elde ediyor. %15,65'lik bu iyileştirme, yüksek etkili yorumların ve ilişkilerin etkisini artıran ağırlıklı kenar yapısının gücünü vurguluyor. Model, paylaşılan duygular ve son derece alakalı yorumlar gibi temel bağlantıları önceliklendirerek üstün doğruluk sağlıyor. Tüm meta-yollardan yararlanarak, W-REHREC çeşitli ilişkiyel sinyalleri entegre etmede mükemmellik göstererek kullanıcı ve mekan tercihlerinin çok yönlü bir temsili elde ediyor.

REHREC ile karşılaştırıldığında W-REHREC, %12,38'lik bir ek iyileştirme sağlayarak, HIN'den en alakalı bilgilerin çıkarılması ve kullanılmasında ağırlıklı kenarların kritik rolünü vurgulamaktadır.

Foursquare Veri Kümesi Analizi:

Açık yorumlardan yoksun olan Foursquare Veri Kümesi, örtük kullanıcı-öge etkileşimlerinin baskın olduğu senaryolarda birleşik meta-yolların genelleme yeteneğini test eder. Bu meta-yollar, kullanıcıları ve mekanları paylaşılan girişler, kategoriler ve sosyal ilişkiler aracılığıyla birbirine bağlayarak modellerin tercihleri ve benzerlikleri etkili bir şekilde çıkarmasına olanak tanır.

REHREC'te, birleşik meta-yollarla RMSE, temel HERec modelindeki 0,0582'ye kıyasla 0,056'dır. Bu, %3,79'luk bir iyileştirmeyi temsil eder ve meta-yolların örtük ilişkileri etkili bir şekilde yakalamak için genelleştirilebileceğini gösterir. UBCaBU ve BCaB gibi meta-yollar, paylaşılan kategoriler aracılığıyla mekanları birbirine bağlamada önemli bir rol oynarken, UBU ve BUB yapısal kullanıcı-mekan ilişkilerini yakalayıp tahmin hatalarını azaltır.

W-REHREC, temel HERec modelindeki 0,0582'ye kıyasla 0,050'lik bir RMSE elde eder. %14,11'lik bu önemli iyileştirme, modelin ağırlıklı yorumdan esinlenen stratejileri metinsel veriler olmadan veri kümelerine uyarlama yeteneğini vurgular. Önemli kullanıcı-öge etkileşimlerini vurgulayarak ve meta-yolların çeşitliliğinden yararlanarak, W-REHREC sağlamlığını ve çok yönlülüğünü gösterir.

REHREC ile W-REHREC arasındaki karşılaştırma, W-REHREC'te %10,71'lik bir iyileştirme ortaya koyuyor ve örtük yapısal sinyallerin hakim olduğu veri kümelerinde ağırlıklı bağlantıların önemini vurguluyor.

Veri Kümeleri Arası Gözlemler:

Yelp ve Foursquare veri kümelerindeki birleşik meta-yolların sonuçları, öneri sistemlerindeki çok yönlülüklerini ve etkinliklerini vurgulamaktadır. Yelp veri kümesinde, tüm meta-yolların dahil edilmesi, modellerin kullanıcılar ve mekanlar arasındaki karmaşık ilişkileri yakalama yeteneğini önemli ölçüde artırır. Temel HERec modeliyle karşılaştırıldığında, REHREC ve W-REHREC sırasıyla %3,73 ve %15,65 oranında RMSE iyileştirmeleri elde ederek çeşitli anlamsal ve yapısal bağlantıları entegre etmenin değerini göstermektedir. Özellikle W-REHREC, yüksek etkili yorumlara ve ilişkilere öncelik vererek önemli doğruluk kazanımları elde ederek öne çıkmaktadır.

Foursquare veri kümesinde, birleştirilmiş meta-yollar kullanıcılar ve mekanlar arasındaki örtük ilişkileri belirlemek için etkili bir şekilde genelleştirilir. REHREC ve W-REHREC, temel HERec modeline kıyasla sırasıyla %3,79 ve %14,11 oranında RMSE iyileştirmeleri elde eder. W-REHREC'in üstün performansı, açık metinsel verilerden yoksun veri kümelerine yorumdan esinlenen teknikleri uyarlama yeteneğini vurgular. UBCaBU ve BCaB gibi meta-yolların kullanımı, iş kategorisi ve yapısal bilgileri kullanarak etkinliğini daha da artırır.

Birleşik meta-yollar (UBU, BUB, URRU, BRRB, UBRRBU, UBCaBU, BCaB) hem yorum açısından zengin hem de yorum açısından yoksun veri kümelerinde öneri doğruluğunu iyileştirmede etkilidir. Yelp veri kümesinde, bu meta-yollar modellerin karmaşık anlamsal ve yapısal ilişkileri yakalamasını sağlayarak hem REHREC hem de W-REHREC'te önemli iyileştirmeler sağlar. Ağırlıklı kenar yapısıyla W-REHREC, etkili bağlantıları önceliklendirmenin kritik rolünü göstererek dikkate değer bir performans elde eder.

Foursquare veri kümesinde, birleştirilmiş meta-yollar örtük ilişkileri yakalamak için etkili bir şekilde genelleştirilir ve modellerin metin dışı bağlamlara uyum sağlamasına olanak tanır. W-REHREC tarafından elde edilen önemli iyileştirmeler, özellikle temel HERec modeliyle karşılaştırıldığında, modern öneri sistemlerinde çeşitli meta-yolları ve ağırlıklı kenarları entegre etmenin önemini doğrular. Bu kapsamlı yaklaşım, çeşitli veri ortamlarında sağlamlık, ölçeklenebilirlik ve üstün performans sağlar.

5. TARTIŞMA, SONUÇ VE GELECEĞE YÖNELİK DÜŞÜNCELER

5.1. Tartışma

Bilgi bolluğunun olduğu modern çağda, öneri sistemleri dijital platformların ayrılmaz bir parçası haline geldi ve kullanıcıları büyük miktarda veri arasında ilgili içeriklere yönlendirdi. İşletmeler ve kullanıcılar karar alma süreçlerinde kişiselleştirilmiş önerilere giderek daha fazla güvendikçe, ayrıntılı kullanıcı-öge etkileşimlerini ve tercihlerini yakalama yeteneği önemli hale geliyor. Bu tez, Heterojen Bilgi Ağları (HIN'ler) içindeki Yorum verilerinden yararlanarak öneri doğruluğunu iyileştirmek için tasarlanmış iki gelişmiş model önererek bu alana katkıda bulunmaktadır: REHREC ve W-REHREC.

HIN tabanlı öneri sistemleri, kullanıcılar, mekanlar ve yorumlar gibi varlıklar arasındaki karmaşık ilişkileri modellemenin benzersiz avantajına sahiptir. Bu tezde, kullanıcı tercihlerinin temsilini geliştirmek için Konu-Bazlı Duygu Analizi (ABSA) ve meta-yol odaklı kelime gömmeleri entegre ederek bu yeteneği genişletiyoruz. Yaklaşımımız, paylaşılan yorumlar, mekan kategorileri ve duygu puanları aracılığıyla karmaşık ilişkileri yakalayan URRU, UBRRBU ve BCaB gibi meta-yolların rolünü vurgular. W-REHREC modeli, ağırlıklı kenarları tanıtarak, yüksek etkili yorumlara ve etkileşimlere öncelik vererek ve çeşitli veri kümeleri arasında üstün performansla sonuçlanarak bu bağlantıları daha da iyileştirir.

Bu modellerin iki veri kümesinde (Yelp ve Foursquare) değerlendirilmesi, sağlamlıklarını ve uyarlanabilirliklerini göstermektedir. Örneğin, zengin yorum içeriğine sahip Yelp veri kümesi, önerilen modellerin duygu odaklı geliştirmelerinden tam olarak yararlanmamızı sağlamıştır. REHREC, temel modellere kıyasla RMSE'de önemli iyileştirmeler elde ederken, W-REHREC, RMSE'yi %15,65 oranında daha da azaltarak ağırlıklı yorum graflarının değerini göstermiştir. Açık yorumlardan yoksun olan Foursquare veri kümesinde, modellerimiz örtük kullanıcı-mekan etkileşimlerini kullanarak genelleme yapma yeteneklerini göstermiş ve %14,11'e kadar RMSE azaltmaları elde etmiştir.

Önceki araştırmalar öneri sistemlerinde kayda değer ilerlemeler kaydetmiş olsa da, bu tez birkaç nedenden ötürü öne çıkıyor. Birincisi, ABSA'nın HIN kelime gömmeleriyle bütünleştirilmesi, kullanıcı tercihleri hakkında daha rafine bir anlayış sağlıyor. İnce ayrıntılı duyguları çıkararak, modeller yiyecek kalitesi, servis ve ambiyans gibi birden fazla açıdan kullanıcı görüşlerini farklılaştırabiliyor. Bu nüanslı yaklaşım, önerilerin kişiselleştirilmesini önemli ölçüde artırıyor. İkincisi, ağırlıklı yorum graflarının dahil edilmesi, kritik

bağlantıların önceliklendirilmesini sağlayarak yüksek değerli yorumların öneri süreci üzerinde daha güçlü bir etki yaratmasını sağlıyor. Bu yenilik, geleneksel olarak öneri sistemlerini etkileyen veri seyrekliği ve Cold-Start sorunlarının zorluklarını doğrudan ele alıyor.

Bu çalışmanın bir diğer ayırt edici özelliği, çapraz veri kümesi genellemesine yaptığı vurgudur. Birçok çalışma tek bir veri kümesindeki performansı optimize etmeye odaklanırken, birden fazla farklı veri kümesindeki değerlendirmemiz önerilen modellerin sağlamlığını ve çok yönlülüğünü vurgular. REHREC ve W-REHREC'in yorum odaklı teknikleri Foursquare gibi veri kümelerine uyarlama yeteneği, gerçek dünya senaryolarında daha geniş uygulanabilirlik potansiyellerini vurgular.

Son olarak, bu tezin en önemli katkılarından biri karmaşık ilişkileri modellemek için meta-yolların kapsamlı kullanımınıdır. UBRBU ve BCaB gibi meta-yollar, anlamsal ve yapısal bilgileri dahil ederek basit kullanıcı-öge etkileşimlerinin ötesine geçer ve modellerin aksi takdirde gizli kalacak olan kalıpları ortaya çıkarmasını sağlar. W-REHREC'te tanıtılan ağırlıklı uzantılar, bu bağlantıların önemini daha da artırarak tüm veri kümelerinde tutarlı iyileştirmeler sağlar.

Gelişmiş HIN tabanlı modelleme, duygu entegrasyonu, ağırlıklı graf oluşturma ve çapraz veri kümesi doğrulamasının birleşimi, bu tezi öneri sistemleri alanına önemli bir katkı haline getirir. Veri seyrekliği, Cold-Start sorunları ve genelleme ihtiyacı gibi temel zorlukları ele alarak, bu çalışma kişiselleştirilmiş ve bağlam farkında önerilerde gelecekteki araştırmalar için güçlü bir temel sağlar.

İleriye bakıldığında, bu araştırma keşif için birkaç heyecan verici konu açıyor. Kullanıcı tercihlerindeki zamansal değişiklikleri yakalamak için dinamik duygu analizini dahil etmek, öneri sistemlerinin uyarlanabilirliğini daha da artırabilir. Modelleri çok dilli yorumları destekleyecek şekilde genişletmek, bunları küresel pazarlara uygulanabilir hale getirirken, modelleri gerçek zamanlı dağıtım için optimize etmek, pratik faydalarını doğrulayabilir. Ek olarak, belirli meta-yolların ve duygu puanlarının etkisini vurgulayan açıklanabilir öneri modelleri geliştirmek, kullanıcı güvenini ve sistem şeffaflığını iyileştirebilir.

Özetle, bu tezde geliştirilen metodolojiler, HIN'ler içindeki yorum verilerinden yararlanmak için sağlam, ölçeklenebilir ve uyarlanabilir bir yaklaşım sunarak öneri sistemlerinde önemli bir ilerlemeyi temsil etmektedir. Burada sunulan yenilikler, daha kişiselleştirilmiş, verimli

ve bağlamsal olarak farkında önerilere giden yolu açarak, akıllı sistemlerin gelişen manzarasına anlamlı bir şekilde katkıda bulunmaktadır.

Etkinlik deneyleri, önerilen modellerin (REHREC ve W-REHREC) tahmin doğruluğunu ve öneri kalitesini, çeşitli veri kümeleri genelindeki en son teknoloji yöntemlerine göre değerlendirmeyi amaçlamaktadır. Bu deneyler, modellerin doğru ve kişiselleştirilmiş öneriler sağlama yeteneğini değerlendirmek için temel değerlendirme ölçütleri olarak Kök Ortalama Kare Hatası (RMSE) ve Ortalama Mutlak Hata (MAE) kullanır.

5.2. Sonuç ve Geleceğe Yönelik Düşünceler

Bu tez, iki yenilikçi öneri modelini tanıtıyor ve titizlikle değerlendiriyor: REHREC (Yorumdan Etkilenen Heterojen Bilgi Ağı Tabanlı Öneri Sistemi) ve ağırlıklı yorum analizini entegre eden geliştirilmiş versiyonu W-REHREC. Her iki model de Heterojen Bilgi Ağlarının zengin ilişkisel yapısından yararlanarak ve yorum tabanlı duygu analizini dahil ederek geleneksel öneri sistemlerindeki kritik zorlukları ele alıyor. Bu modeller birlikte, kullanıcı-öge ilişkilerini yakalama, öneri doğruluğunu iyileştirme ve çeşitli veri kümeleri arasında genelleme yapma yeteneğini önemli ölçüde ilerletiyor.

REHREC modeli, HIN tabanlı öneri sistemlerini yorum verileriyle zenginleştirmede temel bir adımı temsil eder. Yorum bilgilerini yerleştirerek ve çeşitli meta-yolları kullanarak REHREC, kullanıcılar, mekanlar ve yorumlar arasındaki anlamsal ve yapısal ilişkileri yakalar. Bu model, yorum odaklı iç görüşleri geleneksel HIN kelime gömmeleriyle bütünleştirerek öneri doğruluğunu önemli ölçüde iyileştirir. W-REHREC, duygu puanlarına dayalı ağırlıklı kenarlar sunarak, etkili ilişkilere öncelik vererek ve kritik yorumların etkisini artırarak bu temelin üzerine inşa edilir. Bu geliştirme, hem REHREC hem de temel modellere kıyasla üstün doğruluk ve kişiselleştirmeye yol açar.

Bu başarılarla rağmen, bu modellerin geliştirilmesi ve değerlendirilmesi zorluklardan uzak değildi. Birincil zorluklardan biri, özellikle birçok kullanıcı ve mekanın sınırlı etkileşimlere sahip olduğu Yelp ve Foursquare gibi veri kümelerinde veri seyrekliğiyle başa çıkmaktı. Seyrek etkileşimler, özellikle açık yorum verilerinin bulunmadığı senaryolarda, işbirlikçi filtreleme ve anlamlı meta-yolların oluşturulması için engeller oluşturuyordu. Ayrıca, Yelp veri kümesinde belirli mekan kategorilerinin aşırı temsil edilmesi gibi veri dağıtımındaki dengesizlikler, öneri modellerinin iyi temsil edilen gruplara doğru önyargılı olma riski

taşıyordu. Benzer sorunlar, popüler konuların orantısız şekilde daha fazla giriş aldığı Foursquare'de de gözlemlendi.

Bir diğer önemli zorluk, Konu-Bazlı Duygu Analizinin (ABSA) HIN'e entegre edilmesiydi. Yorumlardan duygu puanlarını çıkarmak ve analiz etmek, kullanıcı yorumlarındaki gayriresmi dil, argo ve alakasız içeriklerin işlenmesi de dahil olmak üzere kapsamlı ön işleme gerektiriyordu. Duygu puanlarının ağ yapısına orantılı olarak katkıda bulunduğundan emin olmak, ağırlıklandırma mekanizmalarının yinelemeli olarak ayarlanmasını gerektiriyordu. Dahası, meta-yolların tasarımı ve seçimi başka bir zorluk oluşturuyordu. UBU, BUB, URRU, BRRB, UBRRBU, UBCaBU, BCaB gibi meta-yollar, çeşitli ilişkileri yakalamada etkiliydi, ancak dahil edilecek en uygun yol kümesini belirlemek kapsamlı deneyler gerektiriyordu. Her meta-yol, hesaplama yükü ekliyordu ve aşırı uyum sağlamadan performansı en üst düzeye çıkaran yolları seçmek karmaşık bir görevdi.

Hesaplama karmaşıklığı da zorluklar ortaya çıkardı, özellikle de büyük ölçekli HIN'leri Random Walk algoritması gibi yöntemlerle yerleştirirken. Yelp gibi veri kümeleri için bu süreçleri kelime gömmesi kalitesinden ödün vermeden konu etmek önemli hesaplama kaynakları ve optimizasyon çabaları gerektirdi.

Bu zorluklara rağmen, REHREC ve W-REHREC tarafından veri kümeleri arasında elde edilen iyileştirmeler, önerilen metodolojilerin etkinliğini vurgulamaktadır. Yelp veri kümesinde, REHREC, HERec temel düzeyine kıyasla RMSE'yi %3,73 oranında azaltarak yorum odaklı meta-yolların entegre edilmesinin faydasını göstermiştir. W-REHREC, temel düzeye kıyasla RMSE'de %15,65 oranında dikkate değer bir iyileştirme elde ederek, ağırlıklı kenarların yüksek değerli yorumların etkisini artırmadaki kritik rolünü vurgulamıştır. Foursquare veri kümesinde, açık yorum verilerinin olmamasına rağmen REHREC, RMSE'yi %3,79 oranında azaltarak, yorumdan ilham alan teknikleri örtük kullanıcı-öge etkileşimlerine genelleştirme yeteneğini göstermiştir. W-REHREC, yapısal verilerin hakim olduğu veri kümelerinde uyarlanabilirliğini ve sağlamlığını doğrulayarak %14,11 oranında önemli bir RMSE iyileştirmesi elde etmiştir.

REHREC ve W-REHREC'in birleşik güçleri, çeşitli meta-yolları ve duygu odaklı içgörülerini entegre etme yeteneklerinde yatmaktadır. UBU, BUB, URRU, BRRB, UBRRBU, UBCaBU ve BCaB dahil olmak üzere kullanılan meta-yollar, modellerin kullanıcı-öge ilişkilerinin bütünsel bir yönünü yakalamasını sağlar. W-REHREC'e ağırlıklı kenarların dahil edilmesi,

anlamli etkileşimleri önceliklendirme yeteneğini artırarak, temel ve son teknoloji modellere kıyasla sürekli olarak daha iyi performansla sonuçlanır.

Bu tez ayrıca bu modellerin değişken özelliklere sahip veri kümeleri arasında ölçeklenebilirliğini ve çok yönlülüğünü göstermektedir. Metinsel yorumlarla zengin olan Yelp veri kümesi, modellerin anlamsal içgörüler çıkarma ve bunlardan yararlanma yeteneğini sergilemiştir. Foursquare veri kümesi, yorumdan esinlenen metodolojilerin genelleme potansiyelini vurgulamış ve modellerin seyrek, sosyal graf odaklı ortamlarda ölçeklenebilirliğini doğrulamıştır.

Bu çalışmanın benzersizliği, duygu analizi, meta-yol odaklı kelime gömmeleri ve ağırlıklı kenar inşasını sorunsuz bir şekilde entegre ederek öneri sistemlerini iyileştirmeye yönelik bütünsel yaklaşımında yatmaktadır. Modeller, metinsel ve yapısal veriler arasındaki boşluğu kapatarak, nüanslı kullanıcı-öge ilişkilerini yakalamalarını sağlar. Ağırlıklı kenarların tanıtımı, W-REHREC'in etkili yorumların önemini artırmasına, seyrek etkileşimler ve Cold-Start sorunları gibi zorlukları ele almasına olanak tanır. Modellerin farklı veri bağlamlarına uyarlanabilirliği ve ölçeklenebilir, doğru ve kişiselleştirilmiş öneriler sağlama yetenekleri, öneri sistemleri alanına yaptıkları önemli katkıların altını çizmektedir.

Bu tezde elde edilen sonuçlar gelecekteki araştırmalar için de çeşitli yollar açmaktadır. Dinamik duygu analizi, kullanıcı tercihlerindeki zamansal değişiklikleri dahil ederek modelleri daha da iyileştirebilir. ABSA'yı çok dilli yorumları ele alacak şekilde genişletmek, modelleri küresel pazarlarda uygulanabilir hale getirecektir. REHREC ve W-REHREC'i gerçek zamanlı dağıtım için optimize etmek, pratik faydalarını doğrulayabilirken, meta-yol ve duygu katkılarının yorumlanabilirliğini artırmak şeffaflığı ve kullanıcı güvenini iyileştirebilir.

Bu tez, REHREC ve W-REHREC modellerini geliştirerek yorumla zenginleştirilmiş öneri sistemleri için sağlam bir temel oluşturur. Geleneksel sistemlerdeki kritik zorlukları ele alarak ve önemli performans iyileştirmeleri elde ederek, bu modeller HIN tabanlı öneri araştırmasında yeni bir ölçüt belirler. Burada sunulan yenilikler, kişiselleştirilmiş, bağlam farkında ve ölçeklenebilir öneri sistemlerinde daha fazla ilerlemenin yolunu açar.

KAYNAKLAR

- [1] Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. *The Adaptive Web*, 4321, 291–324. https://doi.org/10.1007/978-3-540-72079-9_9
- [2] Su, X., & Khoshgoftaar, T. M. (2009). A survey of collaborative filtering techniques. *Advances in Artificial Intelligence*, 2009, 421425. <https://doi.org/10.1155/2009/421425>
- [3] Lops, P., De Gemmis, M., & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender Systems Handbook* (pp. 73–105). Springer.
- [4] Kabbur, S., Ning, X., & Karypis, G. (2013). FISM: Factored item similarity models for top-N recommender systems. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 659–667). <https://doi.org/10.1145/2487575.2487589>
- [5] Musto, C., Semeraro, G., de Gemmis, M., & Lops, P. (2016). Learning word embeddings from Wikipedia for content-based recommender systems. In *Proceedings of the 9th International Conference on Knowledge Discovery and Information Retrieval* (pp. 729–734). https://doi.org/10.1007/978-3-319-30671-1_60
- [6] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. In *Proceedings of ICLR*.
- [7] Konstan, J. A., & Riedl, J. (2012). Recommender systems: From algorithms to Kullancı experience. *Kullancı Modeling and Kullancı-Adapted Interaction*, 22(1-2), 101–123. <https://doi.org/10.1007/s11257-011-9112-x>
- [8] Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th International Conference on World Wide Web* (pp. 285–295). <https://doi.org/10.1145/371920.372071>
- [9] Koren, Y. (2008). Factorization meets the neighborhood: A multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD*

International Conference on Knowledge Discovery and Data Mining (pp. 426–434). <https://doi.org/10.1145/1401890.1401944>

- [10] McNee, S. M., Riedl, J., & Konstan, J. A. (2006). Being accurate is not enough: How accuracy metrics have hurt recommender systems. In Proceedings of CHI '06 Extended Abstracts on Human Factors in Computing Systems (pp. 1097–1101).
- [11] He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T.-S. (2017). Neural collaborative filtering. Proceedings of the 26th International World Wide Web Conference (WWW 2017), 173–182. <https://doi.org/10.1145/3038912.3052569>
- [12] Hu, B., Shi, C., Zhao, W. X., & Yu, P. S. (2018). Leveraging meta-path-based context for top-N recommendation with a neural co-attention model. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, United Kingdom. <https://doi.org/10.1145/3219819.3219965>
- [13] Wang, X., He, X., Wang, M., Feng, F., & Chua, T. (2019). Neural graph collaborative filtering. Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, 165–174.
- [14] Wang, H., Zhang, P., Lu, T., Gu, H., & Gu, N. (2017). Hybrid recommendation model based on incremental collaborative filtering and content-based algorithms. In 2017 IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD) (pp. 337–342). <https://doi.org/10.1109/CSCWD.2017.8066717>
- [15] Huang, X., Wang, J., & Cui, J. (2024). A personalized collaborative filtering recommendation system based on bi-graph embedding and causal reasoning. *Entropy*, 26(5), 371. <https://doi.org/10.3390/e26050371>
- [16] Jafri, S. I. H., Ghazali, R., Javid, I., Mahmood, Z., & Hassan, A. A. A. (2022). Deep transfer learning with multimodal embedding to tackle cold-start and sparsity issues in recommendation system. *PLoS ONE*, 17(8), e0273486. <https://doi.org/10.1371/journal.pone.0273486>
- [17] Thorat, S. A., Ashwini, G., & Seema, M. (2023). Survey on collaborative and content-based recommendation systems. In 2023 5th International Conference on

- Smart Systems and Inventive Technology (ICSSIT) (pp. 1541–1548). Tirunelveli, India. <https://doi.org/10.1109/ICSSIT55814.2023.10061072>
- [18] Yang, F., Zhao, X., Huang, P., Xiao, W., & de Rijke, M. (2019). M-HIN: Complex embeddings for heterogeneous information networks via metagraphs. In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'19) (pp. 913–916). Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3331184.3331281>
- [19] Shi, C., Li, Y., Zhang, J., Sun, Y., & Yu, P. S. (2017). A survey of heterogeneous information network analysis. *IEEE Transactions on Knowledge and Data Engineering*, 29(1), 17–37. <https://doi.org/10.1109/TKDE.2016.2598561>
- [20] Khalilzadeh, F., & Cicekli, I. (2024). REHREC modeli: Yorum effected heterogeneous information network recommendation system. *IEEE Access*, 12, 123456–123457. <https://doi.org/10.1109/ACCESS.2024.1234567>
- [21] Wang, Q., Jin, E., Zhang, H., Chen, Y., Yue, Y., Dorado, D. B., Hu, Z., & Xu, M. (2024). Enhancing personalized recommendations: A study on the efficacy of multi-task learning and feature integration. *Information*, 15(6), 312. <https://doi.org/10.3390/info15060312>
- [22] Althbiti, A., Alshamrani, R., Alghamdi, T., Lee, S., & Ma, X. (2021). Addressing data sparsity in collaborative filtering-based recommender systems using clustering and artificial neural network. 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), NV, USA, 0218–0227. <https://doi.org/10.1109/CCWC51732.2021.9376008>
- [23] Zhu, Y., Lin, J., He, S., Wang, B., Guan, Z., Liu, H., & Cai, D. (2018). Addressing the item cold-start problem by attribute-driven active learning. arXiv preprint arXiv:1805.09023. Retrieved from <https://arxiv.org/abs/1805.09023>
- [24] Wang, J., Li, Y., Liang, M., & Li, A. (2022). Embedding representation of academic heterogeneous information networks based on federated learning. arXiv preprint arXiv:2210.03290. Retrieved from <https://arxiv.org/abs/2210.03290>

- [25] Li, L., Gui, X., & Lv, R. (2023). Recommendation algorithm based on heterogeneous information network and attention mechanism. *Applied Sciences*, 14(1), 353. <https://doi.org/10.3390/app14010353>
- [26] Troussas, C., Krouska, A., Koliarakis, A., & Sgouropoulou, C., (2023). Harnessing the Power of Kullanıcı-Centric Artificial Intelligence: Customized Recommendations and Personalization in Hybrid Recommender Systems. *Computers*, 12(5), 109. Available at: <https://doi.org/10.3390/computers12050109>
- [27] Tsai, H.C., Wong, C.H., Shen, J., & Yin, G. (2023). Source-Aware Embedding Training on Heterogeneous Information Networks. *Data Intelligence*, 5(3), 611–635. https://doi.org/10.1162/dint_a_00200
- [28] Wei, J., Liu, Y., Huang, X., Zhang, X., Liu, W., & Yan, X. (2024). Self-Supervised Graph Neural Networks for Enhanced Feature Extraction in Heterogeneous Information Networks. arXiv preprint arXiv:2410.17617. Available at: <https://arxiv.org/abs/2410.17617>.
- [29] Shi, C., Hu, B., Zhao, W.X., & Yu, P.S. (2019). Heterogeneous Information Network Embedding for Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 31(2), 357–370. <https://doi.org/10.1109/TKDE.2018.2833443>
- [30] Mao, C., Wu, Z., Liu, Y., & Shi, Z. (2024). Matrix Factorization Recommendation Algorithm Based on Attention Interaction. *Symmetry*, 16(3), 267. <https://doi.org/10.3390/sym16030267>
- [31] Shang, J., Qu, M., Liu, J., Kaplan, L., Han, J., & Peng, J., (2016). Meta-Path Guided Embedding for Similarity Search in Large-Scale Heterogeneous Information Networks. arXiv preprint arXiv:1610.09769.
- [32] Shang, J., Qu, M., Liu, J., Kaplan, L., Han, J., & Peng, J., (2016). Meta-Path Guided Embedding for Similarity Search in Large-Scale Heterogeneous Information Networks. arXiv preprint arXiv:1610.09769. <https://doi.org/10.48550/arXiv.1610.09769>

- [33] Wu, S., Sun, F., Zhang, W., Xie, X., & Cui, B. (2022). Graph neural networks in recommender systems: A survey. *ACM Computing Surveys*, 55(5), Article 97, 37 pages. <https://doi.org/10.1145/3535101>
- [34] Zhong, Z., Li, C.T., & Pang, J. (2022). Personalised meta-path generation for heterogeneous graph neural networks. *Data Mining and Knowledge Discovery*, 36(10), 2299–2333. <https://doi.org/10.1007/s10618-022-00862-z>
- [35] Wang, L., Song, Y., Huang, H., Ye, F., Shi, X., & Jin, H. (2020). Modeling heterogeneous edges to represent networks with graph auto-encoder. In Nah, Y., Cui, B., Lee, S.W., Yu, J.X., Moon, Y.S., & Whang, S.E. (Eds.), *Database Systems for Advanced Applications. DASFAA 2020, Lecture Notes in Computer Science (Vol. 12113)*. Springer, Cham. https://doi.org/10.1007/978-3-030-59416-9_30
- [36] Choudhary, S., Sharma, K., & Bajaj, M. (2023). Social networks analysis and machine learning: An overview of approaches and applications. 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 123–128. <https://doi.org/10.1109/ICSCSS57650.2023.10169353>
- [37] Ben-Eliezer, O., Eden, T., Oren, J., & Fotakis, D. (2021). Sampling multiple nodes in large networks: Beyond random walks. arXiv preprint. <https://arxiv.org/abs/2110.13324>
- [38] Huang, Z., Silva, A., & Singh, A. (2021). A broader picture of random-walk based graph embedding. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD '21)*, 685–695. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3447548.3467300>
- [39] Peng, H., Li, J., Yan, H., & others. (2020). Dynamic network embedding via incremental skip-gram with negative sampling. *Science China Information Sciences*, 63(202103). <https://doi.org/10.1007/s11432-018-9943-9>
- [40] Ren, X., Wei, W., Xia, L., Su, L., Cheng, S., Wang, J., Yin, D., & Huang, C. (2024). Representation learning with large language models for recommendation. *Proceedings of the ACM Web Conference 2024 (WWW '24)*, 3464–3475.

Association for Computing Machinery, New York, NY, USA.
<https://doi.org/10.1145/3589334.3645458>

- [41] Gao, C., Zheng, Y., Li, N., Li, Y., Qin, Y., Piao, J., Quan, Y., Chang, J., Jin, D., He, X., & Li, Y. (2023). A survey of graph neural networks for recommender systems: Challenges, methods, and directions. *ACM Transactions on Recommender Systems*, 1(1), Article 3, 51 pages. <https://doi.org/10.1145/3568022>
- [42] Perozzi, B., Al-Rfou, R., & Skiena, S., (2014). DeepWalk: Online Learning of Social Representations. *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 701-710. <https://doi.org/10.1145/2623330.2623732>
- [43] Ahmed, A., Shervashidze, N., Narayanamurthy, S., Josifovski, V., & Smola, A.J. (2013). Distributed large-scale natural graph factorization. *Proceedings of the 22nd International Conference on World Wide Web (WWW '13)*, 37–48. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/2488388.2488393>
- [44] Zhang, X., Xie, K., Wang, S., & Huang, Z. (2021). Learning based proximity matrix çarpanlarına ayırma for node embedding. *arXiv preprint*. <https://arxiv.org/abs/2106.05476>
- [45] Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., & Manzagol, P.-A. (2010). Stacked denoising otomatik kodlayıcılar (autoencoders): Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research*, 11, 3371–3408.
- [46] Grover, A., & Leskovec, J. (2016). Node2vec: Scalable feature learning for networks. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*, 855–864. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/2939672.2939754>
- [47] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint*. <https://arxiv.org/abs/1301.3781>

- [48] Dong, Y., Chawla, N.V., & Swami, A. (2017). Metapath2vec: Scalable representation learning for heterogeneous networks. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '17), 135–144. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3097983.3098036>
- [49] Sun, Y., Han, J., Yan, X., Yu, P.S., & Wu, T. (2011). PathSim: Meta path-based top-K similarity search in heterogeneous information networks. Proceedings of the VLDB Endowment, 4(11), 992–1003. <https://doi.org/10.14778/3402707.3402736>
- [50] Mei, G., Pan, L., & Liu, S. (2022). Heterogeneous graph embedding by aggregating meta-path and meta-structure through attention mechanism. Neurocomputing, 468(C), 276–285. <https://doi.org/10.1016/j.neucom.2021.10.001>
- [51] Hu, M., & Liu, B. (2004). Mining and summarizing customer Yorum. Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '04), 168–177. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/1014052.1014073>
- [52] Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O., Hoste, V., Apidianaki, M., Tannier, X., Loukachevitch, N., Kotelnikov, E., Bel, N., Jiménez-Zafra, S.M., & Eryiğit, G. (2016). SemEval-2016 Task 5: Aspect Based Sentiment Analysis. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), 19–30. Association for Computational Linguistics, San Diego, California. <https://doi.org/10.18653/v1/S16-1002>
- [53] Belguith, M., Aloulou, C., & Gargouri, B. (2024). Aspect Level Sentiment Analysis Based on Deep Learning and Ontologies. SN Computer Science, 5, 58. <https://doi.org/10.1007/s42979-023-02362-3>
- [54] Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval, 2(1–2), 1–135. <https://doi.org/10.1561/1500000011>

- [55] Arras, L., Montavon, G., Müller, K.-R., & Samek, W. (2017). Explaining Recurrent Neural Network Predictions in Sentiment Analysis. In A. Balahur, S. M. Mohammad, & E. van der Goot (Eds.), *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (pp. 159–168). Association for Computational Linguistics. <https://doi.org/10.18653/v1/W17-5221>
- [56] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805. <https://arxiv.org/abs/1810.04805>
- [57] Liao, W., Zeng, B., Liu, J., Wei, P., Cheng, X., & Zhang, W. (2021). Multi-level graph neural network for text sentiment analysis. *Computers & Electrical Engineering*, 92, 107096. <https://doi.org/10.1016/j.compeleceng.2021.107096>
- [58] Musto, C., Lops, P., de Gemmis, M., & Semeraro, G. (2019). Justifying recommendations through aspect-based sentiment analysis of Kullanıcı Yorumları. In *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '19)* (pp. 4–12). Association for Computing Machinery. <https://doi.org/10.1145/3320435.3320457>
- [59] Liang, B., Du, J., Xu, R., Li, B., & Huang, H. (2019). Context-aware embedding for targeted aspect-based sentiment analysis. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL '19)*, 4678–4683. <https://doi.org/10.18653/v1/P19-1462>
- [60] Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of tricks for efficient text classification. arXiv preprint arXiv:1607.01759. <https://arxiv.org/abs/1607.01759>
- [61] Sharbatian, K., & Moattar, M. H. (2023). Deep aspect extraction and classification for opinion mining in e-commerce applications using convolutional neural network feature extraction followed by long short term memory attention model. *Applied AI Letters*, 4(3), e86. <https://doi.org/10.1002/ail2.86>
- [62] Chen, Y., Li, Y., & Ma, J. (2022). A study of aspect-level sentiment analysis based on deep learning. In *2022 2nd International Symposium on Artificial Intelligence and its Application on Media (ISAIAM)* (pp. 6–9). IEEE. <https://doi.org/10.1109/ISAIAM55748.2022.00009>

- [63] Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. arXiv preprint arXiv:1408.5882. <https://arxiv.org/abs/1408.5882>
- [64] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv preprint arXiv:1907.11692. <https://arxiv.org/abs/1907.11692>
- [65] Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2020). ALBERT: A Lite BERT for Self-supervised Learning of Language Representations. arXiv preprint arXiv:1909.11942. <https://arxiv.org/abs/1909.11942>
- [66] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All You Need. *Advances in Neural Information Processing Systems*, 30. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fb-d053c1c4a845aa-Paper.pdf
- [67] Liang, B., Liu, Q., Xu, J., Zhou, Q., & Zhang, P. (2017). Aspect-Based Sentiment Analysis Based on Multi-Attention CNN. *Jisuanji Yanjiu yu Fazhan/Computer Research and Development*, 54, 1724–1735. <https://doi.org/10.7544/issn1000-1239.2017.20170178>
- [68] Pang, B., & Lee, L. (2005). Seeing Stars: Exploiting Class Relationships for Sentiment Categorization with Respect to Değerlendirme notu Scales. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)* (pp. 115–124). Association for Computational Linguistics. <https://doi.org/10.3115/1219840.1219855>
- [69] Zhao, G., Luo, Y., Chen, Q., & Qian, X. (2023). Aspect-based sentiment analysis via multitask learning for online Yorum. *Knowledge-Based Systems*, 264, 110326. <https://doi.org/10.1016/j.knosys.2023.110326>
- [70] Henzinger, M. (2016). PageRank Algorithm. In M.Y. Kao (Ed.), *Encyclopedia of Algorithms*. Springer, New York, NY. https://doi.org/10.1007/978-1-4939-2864-4_277

- [71] Park, S., Lee, W., Choe, B., & Lee, S.-G. (2019). A Survey on Personalized PageRank Computation Algorithms. *IEEE Access*, 7, 163049-163062. <https://doi.org/10.1109/ACCESS.2019.2952653>
- [72] Eedi, H., Peri, S., Ranabothu, N., & Utkoor, R. (2021). An Efficient Practical Non-Blocking PageRank Algorithm for Large Scale Graphs. In *Proceedings of the 29th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*, 35-43, Valladolid, Spain. <https://doi.org/10.1109/PDP52278.2021.00015>
- [73] Chowdhary, A., & Kumar, A. (2019). Study of Web Page Ranking Algorithms: A Yorum. *Acta Informatica Malaysia*, 3, 01-04. <https://doi.org/10.26480/aim.02.2019.01.04>
- [74] Li, Z., Piao, W., Sun, Z., Wang, L., Wang, X., & Li, W. (2023). Kullanıcı Real-Time Influence Ranking Algorithm of Social Networks Considering Interactivity and Topicality. *Entropy*, 25(6), 926. <https://doi.org/10.3390/e25060926>
- [75] Thaiprayoon, S., & Unger, H. (2024). Enhancing a Kullanıcı Matchmaking Algorithm using Personalized PageRank. In *Proceedings of the 2023 7th International Conference on Natural Language Processing and Information Retrieval (NLPIR '23)* (pp. 319–327). Association for Computing Machinery. <https://doi.org/10.1145/3639233.3639346>
- [76] Lee, J., & Ahn, S. (2024). PageRank Algorithm-Based Recommendation System for Construction Safety Guidelines. *Buildings*, 14(10), 3041. <https://doi.org/10.3390/buildings14103041>
- [77] Sankaranarayanan, B., Shivarajan, S., Kumaran, T., & Saravanasankar, S. (2018). PageRank Algorithm-Based Recommender System Using Uniformly Average Değerlendirme notu Matrix. In *Advances in Information and Communication Technology*, 10.4018/978-1-5225-5445-5.ch006.
- [78] Al-Sultany, G. (2022). Enhancing Recommendation System using Adapted Personalized PageRank Algorithm. *Proceedings of the International Conference on Information and Communication Technology and Applications (IICETA)*, 1-5. <https://doi.org/10.1109/IICETA54559.2022.9888678>.

- [79] Rezvani, M., & Hashemi, S. M. (2012). Enhancing Accuracy of Topic Sensitive PageRank Using Jaccard Index and Cosine Similarity. In 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, 620–624. Macau, China. <https://doi.org/10.1109/WI-IAT.2012.166>.
- [80] Sallinen, S., Luo, J., & Ripeanu, M. (2023). Real-Time PageRank on Dynamic Graphs. In Proceedings of the 32nd International Symposium on High-Performance Parallel and Distributed Computing (HPDC '23), 239–251. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/3588195.3593004>.
- [81] Wu, S., Wu, D., Quan, J., Chan, T. N., & Lu, K. (2024). Efficient and Accurate PageRank Approximation on Large Graphs. *Proc. ACM Manag. Data*, 2(4), Article 196, 26 pages. <https://doi.org/10.1145/3677132>.
- [82] Srivastava, A., Garg, R., & Mishra, P. (2017). Discussion on Damping Factor Value in PageRank Computation. *International Journal of Intelligent Systems and Applications*, 9, 19-28. <https://doi.org/10.5815/ijisa.2017.09.03>.
- [83] Bahrami Bidoni, Z., George, R., & Shujaee, K. (2014). A Generalization of the PageRank Algorithm. <https://doi.org/10.13140/RG.2.1.1018.7361>.
- [84] Gupta, S., & Kumar, A. (2024). Sentiment-aware Enhancements of PageRank-based Citation Metric, Impact Factor, and H-index for Ranking the Authors of Scholarly Articles. arXiv preprint arXiv:2403.08176. <https://arxiv.org/abs/2403.08176>.
- [85] Gasteiger, J., Bojchevski, A., & Günnemann, S. (2022). Predict then Propagate: Graph Neural Networks meet Personalized PageRank. arXiv preprint arXiv:1810.05997. <https://arxiv.org/abs/1810.05997>.
- [86] Yelp Dataset. (2021). Retrieved from <https://www.yelp.com/dataset>
- [87] Foursquare Dataset. (2013). Retrieved from <https://www.foursquare.com>

EKLER

EK 1 - Tezden Türetilmiş Yayınlar

F. Khalilzadeh and I. Cicekli, "REHREC: Review Effected Heterogeneous Information Network Recommendation System," in IEEE Access, vol. 12, pp. 42751-42760, 2024, doi: 10.1109/ACCESS.2024.3379271.

keywords: {Recommender systems;Reviews;Filtering;Collaborative filtering;Data mining;Semantics;Collaboration;Heterogeneous networks;Information systems;Recommender systems;Heterogeneous information networks;recommendation systems;network embedding;meta-path base random walk}