



Kullanıcı Tabanlı ve Öge Tabanlı İşbirlikçi Filtreleme ile Kümeleme Algoritmalarının Değerlendirilmesi

Mustafa Özgür Cingiz^{1*}, Kadriye Marangoz²

^{1*} Bursa Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, Bursa, Türkiye (ORCID:0000-0003-4469-1440),

mustafa.cingiz@btu.edu.tr

² Bursa Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Bölümü, Bursa, Türkiye (ORCID:0000-0003-3667-9150),

kadriyemarangoz@hotmail.com

(1st International Conference on Applied Engineering and Natural Sciences ICAENS 2021, November 1-3, 2021)

(DOI: 10.31590/ejosat.1005391)

ATIF/REFERENCE: Cingiz, M.Ö. & Marangoz, K. (2021). Kullanıcı Tabanlı ve Öge Tabanlı İşbirlikçi Filtreleme ile Kümeleme Algoritmalarının Değerlendirilmesi. *Avrupa Bilim ve Teknoloji Dergisi*, (28), 453-458.

Öz

Öneri sistemleri, kullanıcıların memnuniyetini ve bağlılığını artırıp, kullanıcılara kişiselleştirilmiş sistem deneyimini yaşatabilmek için geliştirilmiştir. Öneri sistemleri sayesinde kullanıcılar tercihlerine en uygun olan sonucu en az çaba göstererek bulabilmektedirler. Kullanıcıya özel öneri sistemlerinin önemi son yıllarda giderek artmakta ve filmler, şarkılar, haberler başta olmak üzere çeşitli alanlarda uygulanmaktadır. Öneri sistemleri hafıza tabanlı ve model tabanlı olmak üzere ikiye ayrılmaktadır. Model tabanlı filtreleme yaklaşımlarından olan işbirlikçi filtreleme yöntemleri, öneri sistemlerinde yaygın olarak kullanılmaktadır. Bu çalışmada Jester veri seti içerisinde bulunan şakalar kullanıcı tabanlı ve öge tabanlı işbirlikçi filtreleme yöntemleri ile kümelendi. Sonuçlar Davies–Bouldin İndeksi, Dunn İndeksi ve Silhouette Katsayısı değerlerine göre karşılaştırılmıştır. Karşılaştırmaların sonuçlarına göre öge tabanlı işbirlikçi filtreleme yönteminin kullanıcı tabanlı işbirlikçi filtreleme yöntemine göre daha iyi bir doğruluk sağladığı görülmüştür.

Anahtar Kelimeler: Öneri Sistemleri, İşbirlikçi Filtreleme, K-ortalama, Hiyerarşik Kümeleme, Spektral Çift Kümeleme.

Evaluation of Clustering Algorithms with User Based and Item Based Collaborative Filtering

Abstract

Recommendation systems have been developed to increase the satisfaction and loyalty of the users and to provide the users with a personalized system experience. Because of recommendation systems, users can find most suitable result for their preferences with least effort. The importance of user-specific recommendation systems has been increasing in recent years and has been applied in various fields, especially movies, songs, and news. Suggestion systems are divided into memory-based and model-based. Collaborative filtering methods, which are model-based filtering approaches, are widely used in recommendation systems. In this study, jokes in jester dataset were clustered with user-based and item-based collaborative filtering methods. The results were compared according to the davies–bouldin index, dunn index and silhouette score. According to the general result of the comparisons, it has been seen that item-based collaborative filtering method provides better accuracy than the user-based collaborative filtering method.

Keywords: Recommender Systems, Collaborative Filtering, K-means, Hierarchical Clustering, Spectral Bi-clustering.

* Mustafa Özgür CİNGİZ: Mustafa Özgür CİNGİZ: Bursa Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, Bursa, Türkiye (ORCID:0000-0003-4469-1440) mustafa.cingiz@btu.edu.tr

1. Giriş

Web 2.0 ile birlikte artan verinin oluşturduğu büyük veri havuzu içerisinde, kullanıcıların tercihlerine uygun bir sonucu elle bulması günden güne zorlaşmaktadır. Müzik (Subramaniyaswamy, Logesh, Chandrashekar, Challa, & Vijayakumar, 2017), film (Sánchez-Moreno, ve diğerleri, 2016), seyahat (Krishna, ve diğerleri) gibi birçok alanda kullanılan öneri sistemleri sayesinde, kullanıcılar kendilerine uygun sonuçları otomatik olarak bulabilmektedir.

İşbirlikçi filtreleme sistemleri, içerik tabanlı ve bu iki yapının hibritlenmesiyle oluşmuş hibrit sistemler öneri sistemlerinin temel alanlarıdır. İşbirlikçi filtreleme sistemleri model tabanlı ve hafıza tabanlı olmak üzere ikiye ayrılmaktadır (Fathan, Adji, & Ferdiana, 2018). Model tabanlı filtreleme yaklaşımı, geçmiş derecelendirmeye dayalı olarak derecelendirme yapmamış kullanıcıların verecekleri dereceleri tahmin etmektedir. Bu yaklaşımdaki ana yöntemlerden biri gizli faktör modelidir ve bu model sayesinde, kullanıcıların derecelendirdikleri öğelerin gizli özellikleri bulunmaktadır (Koren, 2008). Hafıza tabanlı filtreleme yaklaşımı, öğelerin geçmiş davranışlarından ve kullanıcılardan derecelendirmeye yönelik benzerliklerini keşfederek çalışmaktadır. Bu yaklaşım içeriğe ihtiyaç duymadığından karmaşık öğeler üzerinde doğru bir şekilde çalışabilmektedir (Ozkaya & Seyfi, 2021).

Kullanıcı tabanlı ve öge tabanlı işbirlikçi filtreleme yöntemleri hafıza tabanlı filtreleme yaklaşımının alt başlıklarıdır. Kullanıcı tabanlı işbirlikçi filtreleme yönteminde birbirine benzeyen kullanıcılar ilişkilendirilip kümelenmektedir. Kümedeki kullanıcılara birbirlerinin derecelendirmelerine dayanarak öneride bulunmaktadır. Öge tabanlı işbirlikçi filtreleme yönteminde öge benzerliği baz alınarak bir ilişkilendirme yapılıp kümelenmektedir. Kümedeki öğelerin derecelendirilmelerine dayanarak öneride bulunmaktadır (Sarwar, Karypsis, & Konstan, 2001). Araştırmalar sonucunda öge tabanlı işbirlikçi filtrelemenin, kullanıcı tabanlı işbirlikçi filtrelemeden önemli ölçüde daha yüksek bir performans gösterdiği görülmektedir (Raghuwanshi & Pateriya, 2019), (Ponnam, Punyasamudram, Nallagulla, & Yellamati, 2016).

Bu çalışmada hafıza tabanlı filtreleme yöntemlerinin alt başlığı olan kullanıcı tabanlı ve öge tabanlı işbirlikçi filtreleme yöntemleri kullanılmıştır. Çalışmanın amacı, işbirlikçi filtreleme yöntemleri kullanılarak üç farklı kümeleme algoritmasının beş farklı parametre değerlerine göre karşılaştırmalarının yapılmasıdır.

2. Materyal ve Metot

2.1. Veri Kümesi

Çalışmamızda Jester veri kümesi (The Goldberg Berkeley, 2021) kullanılmıştır. Jester veri kümesi eksik veriler içeren bir veri kümesi olmakla birlikte eksik verilerin oranına göre üç alt versiyonu bulunmaktadır. Veri setleri toplamda 73.421 kullanıcıdan alınan 100 adet şakayla ilgili anonim

değerlendirmeyi içermektedir. Değerlendirmeler [-10,10] arasındadır. 99 olarak belirtilen değerler eksik veri olduğu anlamına gelmektedir.

2.2. Kullanılan Kümeleme Algoritmaları

2.2.1. K-ortalama

K-ortalama algoritması k değeri küme sayısını veren denetimsiz öğrenme algoritmasıdır K değeri belirlendikten sonra algoritma içerisinde k değeri kadar merkez nokta seçilmektedir. Tek tek veriler ile merkez noktalar arasındaki mesafe hesaplanarak veri için yeni bir en yakın merkez nokta belirlenmekte ve buna göre bir kümeye atanmaktadır. Bu yeni merkez nokta belirleme işlemi sistem kararlı duruma gelene kadar devam etmektedir (Hatipoğlu, 2021).

2.2.2. Hiyerarşik Kümeleme

Hiyerarşik kümeleme bir kümeleme algoritmasıdır. Parçadan bütüne (agglomerative) ve bütünden parçaya (divisive) olarak iki alt yaklaşımı vardır. Parçadan bütüne yaklaşımında tüm parçalar tek bir küme halindedir. Benzer özelliklere sahip olan kümeler birleşerek yeni bir küme oluşturmaktadır. Süreç sonunda tek bir küme olması amaçlanmaktadır. Bütünden parçaya yaklaşımında başlangıçta tüm veriler tek bir küme halindedir. Uzaklık/benzerlik matrisine göre veriler ana kümeden ayrılıp alt kümelere bölünmektedir. Süreç sonunda her bir veri için bir küme olması amaçlanmaktadır (Veri Bilimi Okulu, 2021).

2.2.3. Spektral Çift Kümeleme

Bu algoritma, girdi veri matrisini gizli bir dama tahtası yapısına sahip olduğunu varsaymaktadır. Bu yapıya sahip bir matrisin satırlarını ve sütunlarını böler, böylece karşılık gelen blok sal sabit dama tahtası matrisi, orijinal matrise iyi bir yaklaşım sağlar. Veriler üretilmekte, ardından karıştırılmakta ve spektral çift kümeleme algoritmasına aktarılmaktadır. Karıştırılmış matrisin satırları ve sütunları, algoritma tarafından bulunan çift kümeleri göstermek için yeniden düzenlenmektedir. Satır ve sütun etiket vektörlerinin dış çarpımı, dama tahtası yapısının bir temsilini göstermektedir (Scikit-Learn, 2021).

2.3. Değerlendirme Yöntemleri

2.3.1. Davies-Bouldin İndeksi

Davies-bouldin indeksi, her kümenin ortalama benzerlik ölçüsü olarak tanımlanmaktadır. Burada benzerlik, küme içi mesafelerin küme arası mesafelere oranıdır. Böylece, birbirinden daha uzak ve daha az dağınık, grup içi varyansı düşük verilerden oluşan, olan kümeler daha iyi bir skorla sonuçlanacaktır. Minimum puan sıfırdır ve daha düşük değerler daha iyi kümelenmeyi göstermektedir (Scikit-Learn, 2021).

2.3.2. Silhouette Katsayısı

Etiket değeri bilinmeyen modellerin değerlendirilmesi için modelin kendisi kullanılmaktadır. Silhouette katsayı puanının

yüksek olması modelin iyi tanımlanmış kümelerle sahip olduğunu göstermektedir. Siluet katsayısı her örnek için tanımlanmaktadır. Her bir örnek ait olduğu kümedeki diğer örneklerle olan uzaklıkları ve diğer kümelerdeki örneklerle olan uzaklıklarına hesaplanarak Silhouette katsayısı değeri belirlenir. En iyi değer 1 ve en kötü değer -1'dir. 0'a yakın değerler, örtüşen kümeleri göstermektedir. Negatif değerler genellikle, bir örneğin yanlış kümeye atandığını göstermektedir (Scikit-Learn, 2021).

2.3.2. Dunn İndeksi

Dunn indeksi, kümelenecek verinin kendisine dayandığı bir iç değerlendirme şemasıdır. Amaç, kümenin üyeleri arasında varyansın düşük olup ve kümeler arası mesafenin fazla olduğu kümelemeleri belirlemektir. Dunn indeksi kümeler arası uzaklığın kümelerin kendi örnekleri arasındaki uzaklığa oranını verdiği için yüksek bir Dunn indeksi, daha iyi kümelenebilirliği göstermektedir (Geekforgeeks, 2021).

3. Araştırma Sonuçları ve Tartışma

Uygulama, öge tabanlı işbirlikçi filtreleme yöntemi ve kullanıcı tabanlı işbirlikçi filtreleme yöntemi olmak üzere iki farklı yaklaşım kullanılarak oluşturulmuştur. Kullanılan Jester veri seti üç versiyonludur. Versiyonlar arasındaki fark veri setinin içerdiği eksik veri miktarıdır.

3.1. Öge Tabanlı İşbirlikçi Filtreleme Yöntemi

Veri setindeki şakalar öge tabanlı işbirlikçi filtreleme yöntemine göre öge olarak kabul edilmiştir ve okunan eksik verilerin giderilmesi amacıyla şakalar arasında öklid uzaklığı hesaplanmıştır. Bunu yapabilmek için şaka tabanlı bir süreç olması adına okunan veri setinin transpozunu alınmıştır. Bu sayede satırlarda şakalar sütunlarda kullanıcılar olmuştur. 100x100 boyutlu bir matris otomatik olarak oluşturulup bu matriste her şakanın birbirleriyle olan öklid uzaklığı hesaplanmıştır. Her bir eksik veri için bulunduğu sütundaki en küçük 21 verinin öklid uzaklığının ortalaması alınıp eksik verinin yerine doldurulmuştur. En yakın komşu sayısının 21 seçilmesinin nedeni daha küçük değer seçildiği zaman eksik verileri dolduramıyor olmasındandır. K-ortalama, hiyerarşik kümeleme ve spektral çift kümeleme algoritmalarıyla veri seti kümelenebilir. Kümeleme işleminde seçilen kümeleme sayısı 10,20,30,40,50 olacak şekilde beş farklı değer baz alınmıştır. Kümelemenin değerlendirme parametreleri olarak Davies-Bouldin indeksi, Silhouette katsayısı ve Dunn indeksi metrikleri kullanılmıştır.

Aynı işlemler veri setinin ikinci ve üçüncü versiyonları için de yapılmaktadır. Üç veri seti için de uygulanan bu adımların sonucunda alınan değerler aşağıda Tablo 1, Tablo 2 ve Tablo 3 de gösterilmiştir.

Tablo 1: Jester veri seti-1 için öge tabanlı işbirlikçi filtreleme yöntemi sonuçları

Küme sayısı	K-Ortalama			Hiyerarşik Kümeleme			Spektral Çift Kümeleme		
	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi
10	2.33717	0.05119	0.82257	2.38367	0.05999	0.93525	2.87887	0.01310	0.73863
20	1.69847	0.03490	0.79316	1.94184	0.04781	0.93662	2.31734	0.00591	0.61890
30	1.63778	0.02293	0.77198	1.61773	0.04741	0.86372	1.79727	-0.00950	0.56439
40	1.2179	0.02630	0.78287	1.08513	0.04036	0.92065	1.34198	-0.01475	0.56439
50	1.07239	0.79363	0.79363	2.87887	0.01310	0.73863	1.34198	-0.01475	0.56439

Tablo 2: Jester veri seti-2 için öge tabanlı işbirlikçi filtreleme yöntemi sonuçları

Küme sayısı	K-Ortalama			Hiyerarşik Kümeleme			Spektral Çift Kümeleme		
	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi
10	2.64432	0.06853	0.83733	2.62967	0.02934	0.73136	2.74053	0.04474	0.87629
20	1.7758	0.01970	0.86776	1.88854	0.03693	0.86306	2.28793	0.00408	0.84134
30	1.5753	0.02014	0.86790	1.49331	0.03882	0.90653	1.93962	-0.00264	0.64633
40	1.27942	0.01546	0.83194	1.2116	0.03973	0.97273	1.45671	-0.00798	0.80920

50	1.05227	0.01406	0.73536	1.06806	0.03020	0.86228	1.30228	-0.00612	0.68227
----	----------------	---------	---------	----------------	---------	---------	----------------	----------	---------

Tablo 3: Jester veri seti-3 için öge tabanlı işbirlikçi filtreleme yöntemi sonuçları

Küme sayısı	K-Ortalama			Hiyerarşik Kümeleme			Spektral Çift Kümeleme		
	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi
10	0.96040	0.34067	0.82197	1.24758	0.29731	0.90155	1.34634	0.25201	0.74970
20	0.78314	-0.00754	0.63070	0.90240	0.12665	0.83868	1.64119	-0.02300	0.51877
30	0.50017	0.08078	0.98932	0.81551	0.03201	0.74258	1.41302	-0.02103	0.48855
40	0.56069	0.00154	0.84532	0.77384	0.03377	1.00288	1.07396	-0.02028	0.45399
50	0.65489	0.02463	0.91028	0.70664	0.04527	0.98743	1.02661	-0.05022	0.36105

3.2. Kullanıcı Tabanlı İşbirlikçi Filtreleme Yöntemi

Veri setinden sırayla okunan değerlerde eksik veriler vardır ve eksik verilerin giderilmesi amacıyla kullanıcılar arasında öklid uzaklığı hesaplanmıştır. Yaklaşık 25000x25000 boyutlu bir matris otomatik olarak oluşturulup bu matriste her kullanıcının birbirleriyle olan öklid uzaklığı hesaplanmıştır. Her bir eksik veri için bulunduğu sütundaki en küçük 5001 verinin öklid uzaklığının ortalaması alınıp eksik verinin yerine doldurulmuştur. En yakın komşu sayısının 5001 seçilmesinin nedeni daha küçük değer seçildiği zaman eksik verileri

dolduramıyor olmasındandır. K-ortalama, hiyerarşik kümeleme ve spektral çift kümeleme algoritmalarıyla veri seti kümelenmiştir. Kümeleme işleminde, kümeleme sayısı 10,20,30,40,50 olacak şekilde beş farklı kümeleme sayısı baz alınmıştır. Kümelemenin değerlendirme parametreleri olarak Davies-Bouldin indeksi, Silhouette katsayısı ve Dunn indeksi yöntemleri kullanılmıştır.

Aynı işlemler veri setinin ikinci ve üçüncü versiyonları için de yapılmaktadır. Üç veri seti için de uygulanan bu adımların sonucunda alınan değerler aşağıda Tablo 4, Tablo 5 ve Tablo 6 de gösterilmiştir.

Tablo 4: Jester veri seti-1 için kullanıcı tabanlı işbirlikçi filtreleme yöntemi sonuçları

Küme sayısı	K-Ortalama			Hiyerarşik Kümeleme			Spektral Çift Kümeleme		
	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi
10	3.73459	0.01365	0.83163	4.83862	-0.00727	0.80621	6.49305	-0.03479	0.77610
20	4.11136	-0.00890	0.85825	5.04393	-0.03009	0.80874	7.61757	-0.06367	0.67440
30	4.20946	-0.01113	0.71779	5.20977	-0.03993	0.70955	8.4288	-0.06902	0.71722
40	4.20524	-0.02596	0.77966	5.31518	-0.04486	0.74801	8.81453	-0.07191	0.62088
50	4.2611	-0.03890	0.73774	5.44109	-0.05458	0.73961	9.01735	-0.07771	0.63158

Tablo 5: Jester veri seti-2 için kullanıcı tabanlı işbirlikçi filtreleme yöntemi sonuçları

Küme sayısı	K-Ortalama			Hiyerarşik Kümeleme			Spektral Çift Kümeleme		
	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi
10	3.70781	0.02033	0.85424	4.65457	0.00894	0.84597	6.63666	-0.03668	0.73766
20	4.06868	-0.00107	0.83905	5.53483	-0.02151	0.80183	6.76342	-0.06170	0.70024
30	4.21953	-0.01489	0.77388	5.51942	-0.04186	0.77027	8.57485	-0.06856	0.68955
40	4.13174	-0.03056	0.68810	5.32962	-0.05104	0.69953	8.75609	-0.07594	0.65748

50	4.20895	-0.03419	0.70188	5.31809	-0.06058	0.61307	9.37358	-0.07246	0.64144
----	---------	----------	---------	---------	----------	---------	---------	----------	---------

Tablo 6: Jester veri seti-3 için kullanıcı tabanlı işbirlikçi filtreleme yöntemi sonuçları

Küme sayısı	K-Ortalama			Hiyerarşik Kümeleme			Spektral Çift Kümeleme		
	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi	Davies-Bouldin İndeksi	Silhouette Katsayısı	Dunn İndeksi
10	2.60751	0.03093	0.72434	3.65457	-0.00694	0.64578	4.38003	-0.01415	0.85595
20	3.06938	-0.02163	0.73936	4.53483	-0.00131	0.60145	5.31581	-0.03068	0.82216
30	3.20693	-0.02489	0.67968	4.51942	-0.02186	0.57085	6.5759	-0.03874	0.75411
40	3.15694	-0.03063	0.58690	4.32962	-0.03254	0.59945	6.12925	-0.04254	0.76712
50	3.36995	-0.03489	0.60228	4.31809	-0.04087	0.51378	6.53782	-0.04850	0.75550

Araştırma sonuçlarına bakıldığında en iyi performansın Jester-3 veri seti üzerinde alındığı görülmektedir. Bunun nedeni Jester-3 veri setinde olan eksik verilerin diğer versiyonlara oranla fazla olmasından kaynaklanmaktadır. Eksik veriler en yakın komşuların öklid uzaklığı ortalaması baz alınarak doldurulduğunda performans üzerindeki etkisini arttırmaktadır. Yani eksik veri miktarı model performansını etkilemektedir. Öge tabanlı işbirlikçi filtreleme yöntemine göre seçilen kümeleme algoritmalarından k-ortalama kümeleme algoritması hiyerarşik kümeleme ve spektral çift kümeleme algoritmalarına göre daha iyi bir performans göstermiştir. Ayrıca kümeleme algoritmalarındaki küme sayısını veren parametre değerlerinin artması Davies-Bouldin indeksi ve Dunn indeksine göre performansı artırırken Silhouette katsayısına göre azaltmıştır. Kullanıcı tabanlı işbirlikçi filtreleme yöntemine göre seçilen kümeleme algoritmalarından yine k-ortalama kümeleme algoritması hiyerarşik kümeleme ve spektral çift kümeleme algoritmalarına göre daha iyi bir performans göstermiştir. Ayrıca kümeleme algoritmalarındaki küme sayısı değerlerinin artması tüm değerlendirme sonuçlarına göre performansı azaltmaktadır. Elde edilen bulgular ışığında literatür incelemelerinde olduğu gibi öge tabanlı işbirlikçi filtreleme yönteminin kullanıcı tabanlı işbirlikçi filtreleme yöntemine göre daha iyi performans elde ettiği görülmektedir.

4. Sonuçlar ve Öneriler

Bu çalışmada, verilerin artmasıyla birlikte hayatımızda önemli bir yer tutan öneri sistemlerinin farklı kümeleme algoritmalarıyla değerlendirilmesi üzerine çalışılmıştır. Kullanıcı tabanlı işbirlikçi filtreleme yöntemi ve öge tabanlı işbirlikçi filtreleme yöntemi olmak üzere iki tabanlı bir çalışma yapılmıştır. Kümeleme algoritmaları olarak k-ortalama, hiyerarşik kümeleme ve spektral çift kümeleme seçilmiştir. Kümeleme algoritmalarının sonuçları kümeleme değerlendirme metriklerine göre karşılaştırılmıştır. Öge tabanlı işbirlikçi filtreleme yöntemlerinin kullanıcı tabanlı işbirlikçi filtreleme yöntemlerinden daha iyi sonuç verdiği görülmüştür. Gelecekteki çalışmada mevcut çalışmaya ek olarak matris hesaplamalarını daha basit hale getirip performansı arttırmak için tekil değer ayrışımı (SVD) ile boyut azaltma uygulanacaktır.

Kaynakça

- (2021). Geeksforgeeks: <https://www.geeksforgeeks.org/dunn-index-and-db-index-cluster-validity-indices-set-1/>
- (2021). Scikit-Learn: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.silhouette_score.html
- (2021). Scikit-Learn: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.davies_bouldin_score.html
- (2021). Scikit-Learn: <https://scikit-learn.org/stable/modules/biclustering.html>
- (2021). Veri Bilimi Okulu: <https://www.veribilimiokulu.com/hiyerarşik-kümeleme/>
- (2021). The Goldberg Berkeley: <https://goldberg.berkeley.edu/jester-data/>
- Fathan, G., Adji, T. B., & Ferdiana, R. (2018). Impact of Matrix Factorization and Regularization Hyperparameter on a Recommender System for Movies. *2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*.
- Hatipoğlu, E. (2021). The Medium website: <https://medium.com/@ekrem.hatipoglu/machine-learning-clustering-k%C3%BCmeleme-k-means-algorithm-part-13-be33aeef4fc8>
- Koren, Y. (2008). Factorization meets the neighborhood: a multifaceted collaborative filtering model. *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*.
- Krishna, M. G., Haseeb, M., Siyad, B. M., Zameel, P., Budget, & Raj, S. . Experience Based Travel Planner Using Collaborative Filtering. *2021 1st Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON)*.
- Ponnam, L. T., Punyasamudram, S. D., Nallagulla, S. N., & Yellamati, S. (2016). Movie recommender system using item based collaborative filtering technique. *2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS) (pp. 1-5). IEEE*
- Raghuwanshi, S., & Pateriya, R. (2019). Collaborative Filtering Techniques in Recommendation Systems. *In Data, Engineering and Applications. (pp 11-21). Springer, Singapore*

- Sánchez-Moreno, D., González, A. B., Dolores, M., Vicente, M., Batista, V. F., N., M., & García, M. (2016). A collaborative filtering method for music recommendation using playing coefficients for artists and users. *Expert Systems with Applications*, 66, 234-244
- Sarwar, B., Karypis, G., & Konstan, J. (2001). Item-based Collaborative Filtering Recommendation Algorithms. In *Proceedings of the 10th International World Wide Web Conference on World Wide Web* (pp. 285-295)
- Subramaniaswamy, V., Logesh, R., Chandrashekhar, M., Challa, A., & Vijayakumar, V. (2017). A personalised movie recommendation system based on collaborative filtering. *International Journal of High Performance Computing and Networking*, 10(1-2), 54-6