

Düzce Üniversitesi Bilim ve Teknoloji Dergisi

Araştırma Makalesi

Gezi Öneri Sistemine Kullanıcı Özelliklerinin Etkisi

Gizem Zeynep PARİM [a,*], Nevcihan DURU [b]

[a] *Bilgisayar Mühendisliği Bölümü, Fen Bilimleri Enstitüsü, Kocaeli Üniversitesi, Kocaeli, TÜRKİYE*

[b] *Bilgisayar Mühendisliği Bölümü, Mühendislik Fakültesi, Kocaeli Üniversitesi, Kocaeli, TÜRKİYE*

** Sorumlu yazarın e-posta adresi: 145112010@kocaeli.edu.tr*

ÖZET

Günümüzde öneri(tavsiye) sistemlerinin kullanıldığı alanlar giderek artmaktadır. Öneri sistemi kullanıcılarına uygun önerilerde bulunmak da gün geçtikçe önem kazanmaktadır. Bu çalışmada, birçok alanda karşılaştığımız öneri sistemlerine, kullanıcıların kişisel özelliklerinin etkisi araştırılmıştır. Yapılan çalışmada çeşitli illerdeki katılımcılardan internet aracılığı ile İstanbul ili içerisinde yer alan 50 turistik yerden ziyaret etmiş oldukları turistik yerleri oylamaları istenmiştir. Ayrıca katılımcılardan bazı demografik verilerini(yaş, cinsiyet, medeni hal vb.) paylaşmaları istenmiştir. Elde edilen veriler ile öneri sistemlerinde kullanıcı oylarının yanı sıra bazı kullanıcı özelliklerinin de sisteme etkisi araştırılmak istenmiştir. Elde edilen veriler Azure Machine Learning üzerinde kullanılabilecek formata dönüştürülüp sisteme yüklenmiştir. İlk olarak sadece kullanıcı oyları kullanılarak işbirlikçi filtreleme, ardından kullanıcılara ait demografik veriler de sisteme eklenerek içerik tabanlı filtreleme ve işbirlikçi filtrelemenin birleştirildiği hibrit yaklaşım gerçekleştirilmiştir. Eğitim ve test işlemleri sonrasında kullanıcıların hangi özelliklerinin sisteme nasıl etki ettiği gözlemlenmiştir.

Anahtar Kelimeler: *İşbirlikçi Filtreleme, İçerik Tabanlı Filtreleme, Hibrit Yaklaşım, Demografik Veri, Azure Machine Learning*

The Impact of User Features on Trip Recommendation Systems

ABSTRACT

Nowadays, the areas that recommendation systems are used are on the increase. It has become considerably important to make appropriate recommendations to users of the recommendation system. In this study, the influence of the personal characteristics of the users on the recommender systems we encountered in many areas has been investigated. In the study conducted, participants from various cities were asked to vote for the touristic places they visited from 50 touristic places in Istanbul province via the internet. Participants were also asked to share some demographic data (age, gender, marital status, etc.). It is desired to investigate the system effects of some user characteristics as well as user votes in the obtained data. The obtained data is converted into a form that can be used on Azure Machine Learning. Then the data loaded into the system. First, collaborative filtering was performed by using only the user votes, then hybrid approach that collaborative filtering and content based filtering

Geliş: 4/12/2017, Düzeltme: 25/12/2017, Kabul: 12/28/2017

are combined was performed by using user votes and demographic data. After training and testing, it was observed which features of the users affected the recommendation system.

Keywords: *Collaborative Filtering, Content Based Filtering, Hybrid Approach, Demographic Data, Azure Machine Learning*

I. GİRİŞ

Son yıllarda erişmenin kolaylığı ile internet, herhangi bir soruya cevap aranırken faydalanılan ilk araç haline gelmiştir. Kişiler ne izlemeli, ne okumalı gibi sorularına yanıt ararken büyük ölçüde interneti kullanmaktadır. Kişilerin ilgilenebilecekleri ürünleri sunan bu sistemlere öneri(tavsiye) sistemi denmektedir. Bilinen ilk öneri sistemi, Grundy adı verilen bir kütüphane görevlisi oluşturma sistemiydi. Bu yazılımda kullanıcılara bazı sorular sorulup verecekleri cevaplara göre erkek, sporcu, feminist gibi önceden hazırlanan bir kalıba dâhil ederek öneride bulunma amaçlanmıştır [1]. Ekstrand, Riedl, Konstan [2], Grundy sisteminin günümüz öneri sistemleri ile karşılaştırıldığında oldukça ilkel olduğunu fakat öneri sistemlerinin temelini oluşturması açısından oldukça önemli olduğunu vurgulamışlardır.

Öneri sistemlerinin temel amacı kullanıcılara uygun bir şekilde öneride bulunmaktır. Netflix kullanıcılarına sevebileceği bir diziyi veya filmi önermeyi amaçlarken Amazon satın alabileceği bir eşya; İnstagram, Facebook ve Twitter gibi sosyal medya siteleri ise kullanıcılarına arkadaş listelerine ekleyebileceği bir kişi veya beğenebilecekleri bir sayfa önermeyi amaçlamaktadır. Netflix, öneri sistemlerindeki doğruluk oranını arttırma amacıyla 2006 yılında bir yarışma başlattığını duyurmuş, kazanana 1 milyon dolar ödül verileceğini açıklamıştır. Ödülü, 2009 yılında Netflix sisteminin doğruluğunu %10.06 arttıran BellKor's Pragmatic Chaos takımı almaya hak kazanmıştır [8]. Öneri sisteminin temel amacı kullanıcıya doğru öneride bulunmak olsa da yanlış kurulan bir sistem ile kullanıcılara oldukça yanlış öneriler de sunulabilir. Amazon sitesi, sistemlerinden kaynaklanan bir sorun sebebi ile bir süre boyunca müşterilerine bomba yapım malzemeleri; çocuk sahibi olmayan müşterilerine ise çocukları için alabilecekleri bebek bezi önerilerinde bulunmuştur [9].

Literatürde öneri sistemlerinde elde edilen doğruluk oranını arttırmak için yapılan farklı çalışmalar mevcuttur. Bulut ve Milli [3] öneri sistemlerindeki doğruluk oranını arttırmak için iki farklı yöntem önermiştir. Bu yöntemlerden ilkinde, kullanıcıların vermiş olduğu oyların tamamı kullanılarak yapılan standart sapma hesabı ile kullanıcıların negatif ve pozitif standart sapma değerlerine göre sistemde birbirlerine yakınlık dereceleri bulunarak öneri sisteminde daha doğru sonuçlar elde etmeyi amaçlamışlardır. Önerdikleri bir diğer yöntemde ise kullanıcıların, Pearson korelasyon ve kosinüs benzerlikleri ile pozitif ve negatif benzerlikleri hesaplanarak kullanıcıların ağırlıklandırılmış ortalamasının kullanılmasını önermişlerdir. Bakır [4] ise kişilerin tercihlerinin zamana göre değişebileceğini göz önünde bulundurmuş, puan verilen tarihe göre eğer tarih eski ise zayıflatılmasını, yeni ise puanın kuvvetlendirilmesi gerektiğini öne sürmüştür. Taşçı [5], içerik tabanlı filtreleme kullanarak gerçekleştirdiği haber tavsiye sistemindeki doğruluk oranını arttırmak için okunma zamanını kullanmıştır. Kullanıcı, haberi yakın bir zaman diliminde okuduysa kelime ağırlığını kuvvetlendirmiş; fakat eski bir zaman dilinde okuduysa herhangi bir zayıflatma işlemi yapmamıştır. Lvn, Wang ve Raj [6] haber metinlerinde öneride bulunmak için işbirlikçi filtreleme ile kosinüs benzerliğini kullandıklarında daha doğru sonuçlar elde ettiklerini belirtmişlerdir. Srinivas, Balaji, Saravanan [7], kullanıcıların cinsiyet, yaş bilgisinin yanı sıra sosyal ağlardaki ortak arkadaş sayılarını da sisteme ekleyerek ürün tavsiye sistemi geliştirmişlerdir. Netflix'in düzenlediği yarışmada kazanan takım olan BellKor's Pragmatic Chaos ise tek algoritma kullanmak yerine farklı algoritmaları birleştirmiştir [10].

Bu çalışmada ülke genelinde internet üzerinden gerçekleştirilen bir anket ile katılımcılardan İstanbul ilinde ziyaret ettikleri yerlere puan vermeleri istenmiştir. Katılımcılardan ayrıca cinsiyet, yaş, eğitim düzeyi gibi veriler de istenmiş olup bu özelliklerinin öneri sistemine etkisi Azure Machine Learning [11] üzerinde incelenmiştir.

II. ÖNERİ SİSTEMLERİ

A. ÖNERİ SİSTEMLERİ

Öneri sistemleri, kullanıcılara, önceki alışkanlıklarından yola çıkarak nelerden hoşlanabileceğini gösteren sistemlerdir. En yaygın öneri sistemleri, kişilere çeşitli müzik, dizi, film, kitap vb. önerilerinde bulunur. Bir öneri sistemi farklı şekillerde oluşturulabilir. Temel olarak üç farklı yöntem ile öneri sistemi oluşturulabilir. Bunlar işbirlikçi filtreleme, içerik tabanlı filtreleme ve hibrit(melez) yaklaşımdır.

İşbirlikçi filtreleme, kullanıcıların verdiği oy değerlerine göre birbirine yakın kullanıcıları belirleyip bu doğrultuda öneride bulunan sistemlerdir. "A ve B kişileri benzer şeylerden hoşlanıyor ise A kişisinin sevdiği bir şeyi B kişisi de sevebilir." fikrini temel alır. Tablo 1'de *k* kullanıcı, *n* nesne değeri olmak üzere kullanıcı ve nesneye bağlı olarak oluşturulabilecek oy değerleri tablosu gösterilmiştir.

Tablo 1. *Kullanıcı ve nesne oy tablosu*

	n_1	n_2	....	n_{i-1}	n_i
k_1					
k_2					
.					
.					
.					
.					
k_{j-1}					
k_j					

Kullanıcıların benzerlikleri Pearson korelasyonu, kosinüs benzerliği gibi yöntemlerle hesaplanır. Elde edilen sonuçlara göre kullanıcıya en yakın olan diğer bir kullanıcı bulunur. Kullanıcıların verdikleri oy değerleri *r*, oy verilen nesneler *I* olmak üzere *u* ve *v* kullanıcıları arasındaki benzerlik Pearson korelasyonu ile hesaplanırken ortak olarak oy verilen nesneler kullanılmalıdır. Pearson korelasyonu, Eşt. 1'de verilmiştir;

$$sim(u,v) = \frac{\sum_{i \in I_u \cap I_v}(r_{u,i} - \bar{r}_u)(r_{v,i} - r_v)}{\sqrt{\sum_{i \in I_u \cap I_v}(r_{u,i} - \bar{r}_u)^2}\sqrt{\sum_{i \in I_u \cap I_v}(r_{v,i} - \bar{r}_v)^2}} \quad (1)$$

Kosinüs benzerliğinde de kullanıcıların benzerliği verdikleri oylara göre hesaplanır. Kullanıcıların verdikleri oy değerleri *r* vektörü olmak üzere *u* ve *v* kullanıcıları arasındaki benzerlik kosinüs benzerliği ile Eşt. 2'deki gibi hesaplanabilir.

$$sim(u,v) = \frac{r_u . r_v}{||r_u||_2 . ||r_v||_2} = \frac{\sum_i r_{u,i} r_{v,i}}{\sqrt{\sum_i r_{u,i}^2}\sqrt{\sum_i r_{v,i}^2}} \quad (2)$$

İçerik tabanlı filtrelemede ise önemli olan içeriklerin benzerliğidir. Metin, kategori vb. benzerlikleri kullanılarak öneri yapılır. Bir film öneri sistemi ele alındığında içerik tabanlı filtreleme kullanılacaksa; oyunculara, yönetmene, film türüne göre benzerlik göz önünde bulundurularak öneride bulunabilir.
İşbirlikçi filtreleme ile içerik tabanlı filtrelemenin birleştirilerek daha doğru sonuçlar elde edileceği düşünülmüştür. Bu sebeple hibrit(melez) yaklaşım ortaya çıkmıştır.

B. AZURE MACHİNE LEARNİNG ÖNERİ SİSTEMİ

Azure Machine Learning, Microsoft tarafından oluşturulan ve birçok algoritma ve örnek veri setini içinde barındıran bir makine öğrenmesi sistemidir. Sınıflandırma, kümeleme gibi birçok makine öğrenmesi işlemini yapabilmektedir. Azure Machine Learning gibi öneri sistemlerini çalıştırabilen çeşitli veri bilimi platformları mevcuttur. Dataiku [12] da üzerinde öneri sistemi oluşturulabilen bu platformlardan biridir. Masaüstü ve çevrimiçi olarak çalışan iki farklı kullanımı mevcuttur. Masaüstü kullanımı ücretsizdir fakat içerisindeki modüller erişim açısından kısıtlıdır. Çevrimiçi kullanımda ise ücretsiz kullanım süresi 2 haftadır. Deneme süresinden sonra çevrimiçi kullanmak ücretlidir. Azure Machine Learning sistemi masaüstü olarak kullanılamamaktadır. Çalışılan bilgisayara herhangi bir program yüklemeye gerek duymadan tarayıcı üzerinden çevrimiçi olarak çalışır. Azure Machine Learning sistemine kredi kartı bilgisi verilmeden, okul elektronik postası(edu uzantılı) ile giriş yapıldığında 1 yıl boyunca ücretsiz kullanım imkânı sunuluyor. İçerisinde yer alan modüller herhangi bir kod yazımına gerek duymadan sürükle bırak yöntemi ile çalışır. Bu özelliği ile diğer sistemlerden farklı olarak kod yazma bilgisi gerektirmeden herkesin makine öğrenmesi alanında proje gerçekleştirmesine olanak sağlar. İçerisinde yer alan örnek veri setlerinden birini kullanmak yerine kendi verilerinizi kullanmak istediğinizde bu veriyi sisteme bir kez yüklemeniz gereklidir. Ücretsiz sürümde verileri saklamak için 10 GB alan ayrılmıştır. Azure Machine Learning ile web servisleri de oluşturulabilir; fakat ücretsiz sürümde web servis oluşturma sayısı ve bu servisler kullanılarak yapılabilecek aylık işlem sayısı sınırlıdır. Kullanıcı görüşleri ile oluşturulan en iyi veri bilimi platformları sırasında Azure Machine Learning 9. sırada yer alırken Dataiku 12. sırada yer almaktadır [13]. Bu çalışmada kullanım kolaylığı, 10 GB boyutuna kadar olan verileri ücretsiz bir şekilde depolayabilme ve herhangi bir kredi kartı bilgisi gerektirmeden üye olmayı sağladığı için Azure Machine Learning sistemi tercih edilmiştir.

Azure Machine Learning üzerinde yer alan öneri sistemi ile işbirlikçi filtreleme, içerik tabanlı filtreleme ve hibrit yaklaşım oluşturulabilir. Burada oluşturulan sistemde kullanıcıların ve nesnelerin tanımlayıcıları(id değerleri) ile sadece oy değerleri kullanıldığında işbirlikçi filtreleme; kullanıcı ve nesnelerin benzerlikleri için özellik değerleri kullanıldığında içerik tabanlı filtreleme oluşturulabilir. Azure Machine Learning içerisinde yer alan sistem içerik tabanlı filtreleme ile işbirlikçi filtrelemeyi birleştirebilen bir yapıya sahip olduğu için hibrit yaklaşımda da bulunabilir [14].
Yapılan çalışmada sisteme ilk olarak kullanıcı, nesne ve oy değerleri verisi yüklenmiştir. Bu işbirlikçi filtreleme olarak çalışır. Kullanıcı özelliklerinin eklenmesi ile sistem, hibrit yaklaşımda bulunur.

III. YÖNTEM

A. ÇALIŞMA GRUBU

Araştırma için İstanbul'da ziyaret edilebilecek 50 yer belirlenmiştir. Kullanıcı özellikleri için ise cinsiyet, medeni hal, eğitim düzeyi, gelir düzeyi, yaş grubu, meslek ve yaşadıkları il soruları kullanıcılara yöneltilmek için hazırlanmıştır. Anket 11.11.17 tarihinde sosyal medya ve forumlarda paylaşılmıştır. Kullanıcılardan, daha önce ziyaret ettikleri yerlere puan vermeleri ve ayrıca kişisel sorulara da cevap vermeleri istenmiştir. Gerçekleştirilen anket çalışması ile 209 adet farklı kullanıcıdan yanıt toplanmıştır. Kullanıcıların daha önce ziyaret etmedikleri yerler toplanan yanıtlardan çıkarılmış ve toplamda 209 kullanıcıya ait listede yer alan 50 turistik yer için 7652 oy verisi elde edilmiştir.

B. VERİ TOPLAMA

Katılımcılara Tablo 2'de verilen yerleri daha önce ziyaret etmemişlerse 0(sıfır); ziyaret etmişlerse en düşük 1 en yüksek 10 olmak üzere bir puan vermeleri istenmiştir.

Tablo 2. *Ziyaret edilebilecek yerlerin listesi*

Sıra No	**Yerin Adı**
1	Topkapı Sarayı
2	Dolmabahçe Sarayı
3	Çırağan Sarayı
4	Galata Kulesi
5	Kız Kulesi
6	Eyüp Sultan Cami
7	Ortaköy Cami
8	Mihrimah Sultan Cami
9	Süleymaniye Cami
10	Sultanahmet Cami
11	Ayasofya Müzesi
12	İstanbul Arkeoloji Müzesi
13	Rahmi Koç Müzesi
14	Masumiyet Müzesi
15	Pera Müzesi
16	Aşiyan Müzesi
17	Oyuncak Müzesi
18	Panoroma 1453
19	Miniatürk
20	Yerebatan Sarnıcı
21	Yedikule Surları
22	Cağaloğlu Hamamı
23	Kapalı Çarşı
24	Mısır Çarşısı
25	Kuzguncuk Evleri
26	Pierre Loti Tepesi
27	Çamlıca Tepesi
28	Otağtepe
29	Atatürk Arboretumu
30	Emirgan Korusu
31	Fethi Paşa Korusu
32	Mihrabat Korusu
33	Cemile Sultan Korusu
34	Gülhane Parkı
35	Gezi Parkı
36	Yoğurtçu Parkı
37	Yıldız Parkı
38	Rumeli Hisarı
39	Anadolu Hisarı
40	Haydarpaşa Garı
41	Taksim Meydanı
42	Büyükada
43	Heybeli Ada
44	Kınalı Ada
45	Burgazada
46	Bebek Sahili
47	Boğaz Turu
48	Eminönü Balıkçıları
49	İstanbul Akvaryum
50	Vialand

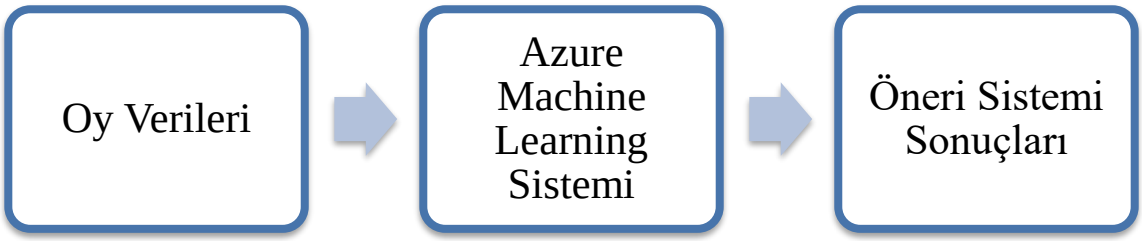
Katılımcıların kişisel özelliklerine ait verileri elde etmek için Tablo 3'te yer alan ifadelere yanıt vermesi istenmiştir. Ankete katılan doktora mezunu olmadığı için tabloda gösterilmemiştir. Tablo 3'teki sorulara ek olarak katılımcılara meslekleri ve yaşadıkları şehir de sorulmuştur. Elde edilen verilerde 28 farklı meslek ve 16 farklı şehir bilgisi yer almaktadır.

Tablo 3. *Kullanıcıdan istenen kişisel özellik soruları*

Nitelik	**1**	**2**	**3**	**4**	**5**	**6**
Cinsiyet	Erkek	Kadın				
Medeni Hal	Evli	Bekâr				
Eğitim Düzeyi	İlkokul	Ortaokul	Lise	Ön lisans	Lisans	Yüksek Lisans
Gelir Düzeyi	Geliri yok.	1000TL altı	1000TL-2500TL	2500-5000TL	5000TL ve üstü	
Yaş	18 yaş altı	18-25 yaş grubu	26-40 yaş grubu	41 ve üstü		

C. VERİ ANALİZİ

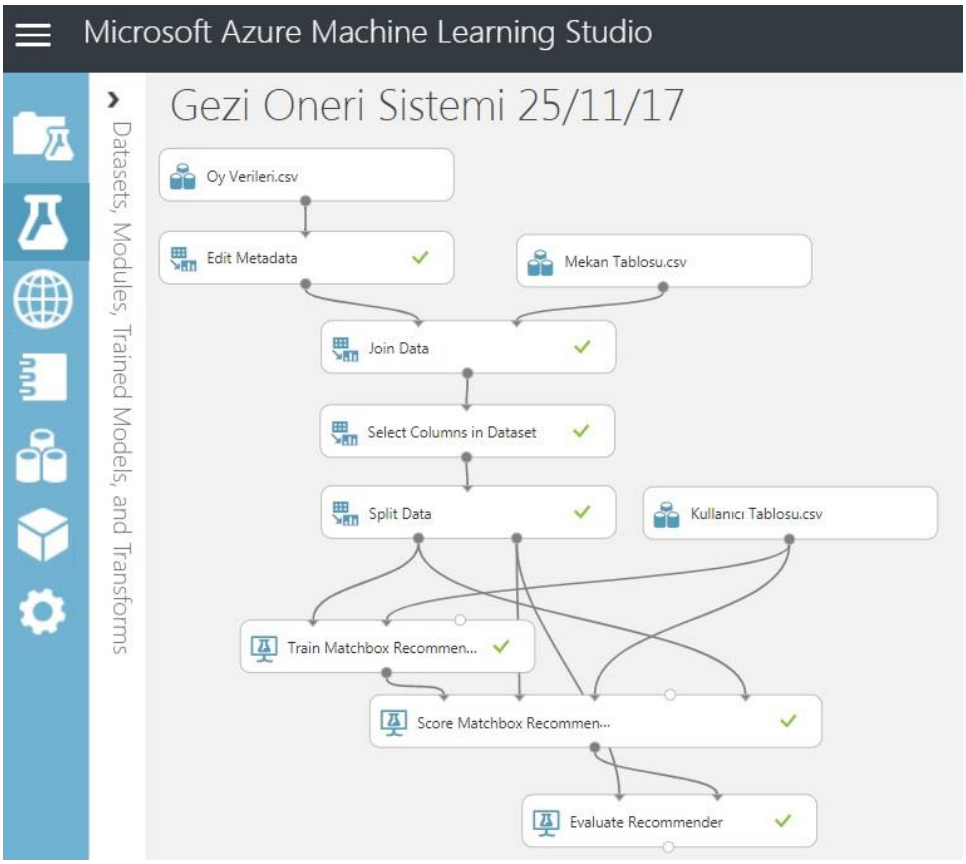
Araştırma kapsamında elde edilen veriler Azure Machine Learning [11] sistemine yüklenmiş ve eğitim aşamasından geçirilmiştir. Eğitim verisi 0,75, test verisi 0,25 olarak ayrılmıştır. İterasyon sayısı 10 olarak seçilmiştir. İlk olarak kullanıcılara ait hiçbir veri sisteme yüklenmeden doğruluk oranı ölçülmüştür. Bu işlemden sonra sırayla kişisel özelliklerden sadece biri eklenerek ölçüm yapılmıştır. Bir sonraki adımda ise kullanıcılardan elde edilen tüm veriler sisteme eklenerek ölçüm yapılmıştır. Son aşamada ise tek tek eklendiğinde en yüksek sonucu veren özelliğe diğer özellikler sırayla doğruluk oranına göre eklenmiş ve ikili, üçlü, dörtlü, beşli, altılı gözlem yapılmıştır. Şekil 1'de Azure Machine Learning sisteminin temel çalışma prensibi verilmiştir.



Şekil 1. *Azure Machine Learning sisteminin çalışma prensibi*

Şekil 2'de Azure Machine Learning üzerinde oluşturulan ve kullanıcılara ait tüm özelliklerin eklendiği öneri sistemine ait görüntü verilmiştir. Kullanıcıların mekânlara verdikleri oylar, *Oy Verileri* dosyasında "kullaniciID, mekanID, raiting" olarak; Tablo 2'de belirtilen yer isimleri *Mekan Tablosu* dosyasında "mekanID, mekanAdi" olarak yer almaktadır. Bu dosyalar Azure Machine Learning üzerinde *Join Data* modülü ile birleştirilmiştir. *Oy Verileri* dosyasında yer alan verilerden oy değerinin hangi sütun başlığı altında yer aldığı *Edit Metadata* bölümünde belirtilmiştir. Join işlemi sonrasında birleştirilen verilerden hangi sütunların eğitim ve test işlemlerinde kullanılacağı *Select Columns in Dataset* modülü ile belirlenmiştir. Veriler *Split Data* modülü kullanılarak eğitim ve test verisi olmak üzere ikiye bölünmüştür. *Train Matchbox Recommender* ile eğitim işlemi gerçekleştirilmiş; *Score Matchbox Recommender* modülü ile de eğitim işlemi sonrası test verileri için doğruluk kontrolü

gerçekleştirilmiştir. Son basamakta *Evaluate Recommender* ile kurulan sistemin doğruluk oranı hesaplanmıştır.



Şekil 2. *Kullanıcıya ait verilerin tamamı kullanılarak oluşturulan öneri sistemi*

IV. BULGULAR ve TARTIŞMA

Veriler ilk olarak herhangi bir kullanıcı verisi eklenmeden Azure Machine Learning sistemine yüklenmiştir. Eğitim ve test aşamaları sonrasında doğruluk oranının 0,845738 olduğu gözlemlenmiştir. Oluşturulan sisteme cinsiyet, medeni hal, eğitim düzeyi, gelir düzeyi, yaş, meslek ve yaşanılan şehir verileri teker teker eklenmiş ve her biri eklendiğinde elde edilen sonuçlar Tablo 4'te gösterilmiştir.

Tablo 4. *Kullanıcı verileri eklendiğinde sistemin doğruluk değeri*

Nitelik	**Kullanıcı Verisi Olmadan**	**Kullanıcı Verisi Eklendiğinde**
Cinsiyet	0,845738	0, 896298
Medeni Hal	0,845738	0,895957
Eğitim Düzeyi	0,845738	0,895377
Gelir Düzeyi	0,845738	0,898373

Yaş	0,845738	0,899404
Meslek	0,845738	0,895513
Şehir	0,845738	0,89463
Kullanıcı Verilerinin Tümü	0,845738	0,891131

Tablo 4'te yer alan verilere göre en yüksek doğruluk değerinden en düşük değere doğru olan sıralama yaş, gelir düzeyi, cinsiyet, medeni hal, meslek, eğitim düzeyi, şehir şeklindedir. En yüksek doğruluk değerinden itibaren sırayla sisteme ekleme yapılarak elde edilen ikili, üçlü, dörtlü, beşli, altılı ve yedili özelliklere göre oluşan sonuçlar Tablo 5'te gösterilmiştir.

Tablo 5. *En yüksek doğruluk değerinden itibaren ekleme yapıldığında elde edilen sonuçlar*

Nitelik	**Kullanıcı Verisi Eklendiğinde**
Yaş-Gelir Düzeyi	0,8971
Yaş-Gelir Düzeyi-Cinsiyet	0,897746
Yaş-Gelir Düzeyi-Cinsiyet-Medeni Hal	0,894812
Yaş-Gelir Düzeyi-Cinsiyet-Medeni Hal-Meslek	0,892571
Yaş-Gelir Düzeyi-Cinsiyet-Medeni Hal-Meslek-Eğitim Düzeyi	0,893062
Yaş-Gelir Düzeyi-Cinsiyet-Medeni Hal-Meslek-Eğitim Düzeyi-Şehir	0,891131

Tablo 5'teki değerlere göre en doğru sonuç "Yaş-Gelir Düzeyi-Cinsiyet" ile 0,897746 olarak elde edilmiştir. Tablo 4 ve Tablo 5 birlikte incelendiğinde en yüksek değerin Tablo 4'te yer alan yaş verisi ile elde edildiği görülebilir. Ayrıca kullanıcılardan elde edilen veriler doğrultusunda, ziyaret edilen yerlere ait ortalama oy değerleri hesaplanmıştır. En yüksek ortalama oya sahip ilk yirmi yer Tablo 6'da gösterilmiştir.

Tablo 6. *Oy ortalamalarına göre en çok beğenilen ilk 20 yer*

Sıra No	**Yerin Adı**	**Oy Ortalaması**
1	Boğaz Turu	8,9362
2	Ayasofya Müzesi	8,8506
3	Dolmabahçe Sarayı	8,7219
4	Galata Kulesi	8,6080
5	Topkapı Sarayı	8,5698
6	Sultanahmet Cami	8,5323
7	Yerebatan Sarnıcı	8,4233
8	Çamlıca Tepesi	8,4222
9	Büyükada	8,4205
10	Atatürk Arboretumu	8,3889
11	Kapalı Çarşı	8,3819
12	Kız Kulesi	8,3270
13	Pierre Loti Tepesi	8,3257
14	Mısır Çarşısı	8,2053
15	Heybeli Ada	8,1987

16	Eyüp Sultan Cami	8,1742
17	Miniatürk	8,1515
18	Haydarpaşa Garı	8,0492
19	Anadolu Hisarı	8,0452
20	Emirgan Korusu	8,0422

V. SONUÇ VE GELECEK ÇALIŞMALAR

Bu araştırma kapsamında internet aracılığı ile Türkiye'deki çeşitli illerde ikamet eden katılımcıların İstanbul'da daha önce ziyaret ettikleri turistik mekânlara verdikleri oylar ile demografik verileri birleştirilerek öneri sistemlerinde doğruluk oranı arttırılmak istenmiştir. Gerçekleştirilen bu çalışma ile demografik verilerin öneri sistemindeki doğruluk oranını arttırdığı tespit edilmiştir. Katılımcılara ait hiçbir bilgi olmadığında sistemdeki 0,845738 olan doğruluk oranının katılımcılara ait özelliklerin sisteme eklenmesiyle arttığı gözlemleniştir. Katılımcıların sadece yaş bilgisinin sisteme eklenmesiyle en yüksek doğruluk oranı 0,899404 olarak elde edilmiş ve 0,002 değerindeki fark ile bir sonraki yüksek değerin kişilerin yaş, cinsiyet ve gelir düzeylerinin birlikte kullanıldığında elde edildiği gözlemlenmiştir. Yapılan bu çalışma ile öneri sistemlerinde kullanıcılara ait çeşitli demografik verilerin kullanılması ile sistemdeki doğruluk oranının arttığı belirlenmiştir.
Gelecek çalışmalarda kullanıcı verilerinin yanı sıra ziyaretin yapıldığı gün, ay, yıl bilgisi, ulaşım şekli(özel araç ile ulaşım, otobüs ile ulaşım vb.), müze vb. yerler için giriş ücretleri, gezinin toplam maliyeti gibi seçenekler de eklenerek öneri sistemindeki doğruluk oranına etkisi araştırılabilir.

VI. KAYNAKLAR

[1] E. Rich, "User Modeling Via Stereotypes," *Cognitive science* vol. 3, no.4, pp. 329-354, 1979.

[2] M. D. Ekstrand, J. T. Riedl, and A. J. Konstan. "Collaborative filtering recommender systems," Foundations and Trends® in Human–Computer Interaction vol. 4, no.2, pp. 81-173, 2011.

[3] H. Bulut, M. Milli, "İşbirlikçi Filtreleme İçin Yeni Tahminleme Yöntemleri," *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, c. 22, s. 2, ss.123-128, 2016.

[4] Ç. Bakır, "Zamana bağlı işbirlikçi filtreleme," Yüksek lisans tezi, Bilgisayar Mühendisliği Bölümü, Yıldız Teknik Üniversitesi, İstanbul, Türkiye, 2014.

[5] S. Taşçı, "İçerik Bazlı Medya Takip ve Haber Tavsiye Sistemi," Yüksek lisans tezi, Bilgisayar Mühendisliği Bölümü, Hacettepe Üniversitesi, Ankara, Türkiye, 2015.

[6] R. Lvn, Q. Wang, J.D. Raj, "Recommending news articles using cosine similarity function," *SAS Global Forum 2014,* Washington, USA, 2014, pp 1-8.

[7] R.S. Srinivas, C.S.A. Balaji, P. Saravanan, " Online Product Recommendation using Relationships and Demographic Data on Social Networks," *Indian Journal of Science and Technology*, vol. 9, no. 44, 2016.

[8] Anonim, (26 Kasım 2017). *[Online]*. Access: https://www.netflixprize.com/community/topic_1537.html.

[9] S. Kennedy, (2017, 28 Kasım). *Potentially Deadly Bomb Ingredients-On-Amazon.* Access: https://www.channel4.com/news/potentially-deadly-bomb-ingredients-on-amazon.

[10] R.M. Bell, Y. Koren, C. Volinsky. (2017, 1 Aralık). *Netflixprize.* Access*:* https://www.netflixprize.com/assets/ProgressPrize2008_BellKor.pdf.

[11] Anonim, (25 Kasım 2017). *[Online]*. Access: https://studio.azureml.net

[12] Anonim, (26 Kasım 2017). *[Online]*. Access: https://www.dataiku.com

[13] Anonim, (26 Kasım 2017). *[Online]*. Access: https://www.itcentralstation.com/categories/data-science-platforms#top_rated.aspx.

[14] Anonim, (26 Kasım 2017). *[Online]* https://msdn.microsoft.com/en-us/library/azure/dn905987.aspx.