

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/304460824>

# Development of content based book recommendation system using genetic algorithm

Conference Paper · May 2016

DOI: 10.1109/SIU.2016.7495917

---

CITATIONS

4

---

READS

440

2 authors, including:



[Oktay Yildiz](#)

Gazi University

36 PUBLICATIONS 123 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Rule-Based Mobility Management Routing for Mobile Ad Hoc Networks. [View project](#)

# Genetik Algoritma Kullanılarak İçerik Tabanlı Kitap Tavsiye Sistemi Geliştirilmesi Development Of Content Based Book Recommendation System Using Genetic Algorithm

Celalettin Aygün<sup>1</sup>, Oktay Yıldız<sup>1</sup>

<sup>1</sup> Bilgisayar Mühendisliği Bölümü, Gazi Üniversitesi, Ankara, Türkiye  
{celalettin.aygun, oyildiz}@gazi.edu.tr

**Özetçe**—İnternet kullanıcılarının ihtiyaç duydukları içeriklere doğru ve hızlı bir şekilde ulaşması günümüzde kritik bir konu haline gelmiştir. Bu amaçla tavsiye sistemleri, müzik, kitap, film, turistik gezi veya planlama, e-ticaret, eğitim ve daha pek çok alanda yaygın bir şekilde kullanılmaktadır. Tavsiye sistemleri yaklaşımı, kullanıcının geçmişteki seçim, beğeni ve değerlendirmelerinin yorumlanarak, ilgisi doğrultusunda sonraki seçimlerini öngörme temeline dayanır. Gerçekleştirilen bu çalışmada, yüksek oranda tatmin edici sonuçlar üreten, yeni bir kitap tavsiye sistemi önerilmektedir. Sistem, kullanıcının okumuş olduğu kitapları ve değerlendirmeleri dikkate alarak, kullanıcının ilgi duyacağı yeni kitapları okuyucuyla buluşturmaktadır.

**Anahtar Kelimeler** — tavsiye sistemleri; içerik tabanlı filtreleme; genetik algoritma.

**Abstract**— Trend of internet, makes presentation of right content to right user will inevitably. For this purpose, recommendation systems are widely used for the areas of music, book, movie, touristic travel planning, e-commerce, education and many more. The approach of recommender systems are based on the ground of users history of choices, likings and reviews, each of which is interpreted in order to project the future choices of the user. In this study, a novel and strong recommender system for the books is proposed. A content based book recommendation application was developed which makes recommendations according to user's taste and choices.

**Keywords** — recommendation systems; content based filtering; genetic algorithm.

## I. GİRİŞ

İnternet, gerek barındırdığı içeriğin, gerekse çeşitlilik gösteren kullanıcıların giderek arttığı bir ortam haline gelmektedir. Bu durum da, doğru kullanıcılarla doğru içeriklerin olabildiğince kısa sürede karşılaştırılması ihtiyacını kaçınılmaz kılmaktadır. Başta ticari siteler

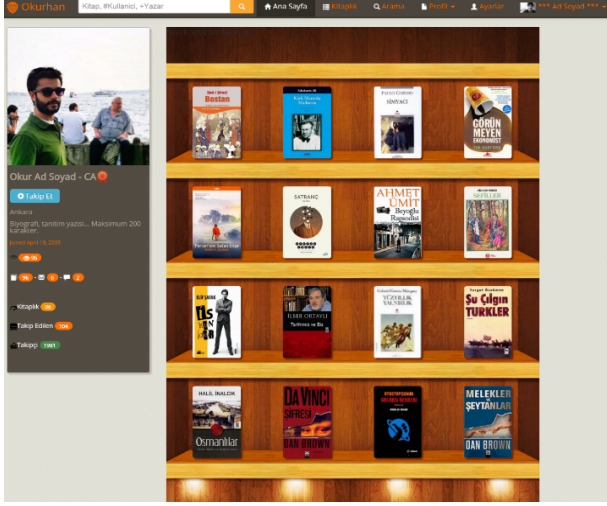
olmak üzere pek çok internet uygulamasında, kullanıcıya ilgisi doğrultusunda içeriklerin sunulması, popüler içeriklerin ötesinde, gün yüzüne çıkmayan özel içeriklerin kendileri için daha anlamlı olan kullanıcılarla buluşturulması büyük önem taşımaktadır [1-3]. Bu noktada devreye Tavsiye Sistemleri (TS) girmektedir.

Müzik [4-7], kitap [2, 6, 8], film [6, 9], turistik gezi planlama [10], web içeriği [3], e-ticaret [2, 6, 11], eğitim [12] gibi pek çok alanda TS kullanımına yönelim görülmektedir. Çalışmaların tümündeki ortak nokta, kullanıcının geçmişteki seçim ve değerlendirmeleri yorumlanarak, ilgisi doğrultusunda sonraki seçimlerini öngörebilmektir ki TS yapısı da bunun üzerine kuruludur. Yine benzer olarak TS'lerin amacı, kullanım kayıtlarından, kullanıcının aklındaki örtülü kalmış nesneleri keşfetmek, ortaya çıkarmak olarak ifade edilmektedir [1].

Geliştirilen kitap tavsiye sistemi, kullanıcının okumuş olduğu kitaplar ve değerlendirmeleri gözetilerek, kendisinde ilgi uyandıracağı ya da beğenebileceği düşünülen kitapları okuyucuyla buluşturmaya hedeflemektedir. Uygulama bu yönüyle kişiye özgü, birebir tavsiye çıktıları üretmektedir. Böylelikle, kullanıcı karşısına yalnızca popüler olan ya da belli kategorideki kitaplar değil, okuma profiliyle örtüşen, ilgi duyduğu yelpazeden kitapların bulunduğu bir öneri şablonu çıkacaktır.

## II. GENETİK ALGORİTMA TABANLI KİTAP TAVSİYE SİSTEMİ

Kitap tavsiye sistemi, kitap temalı sosyal ağ modeliyle geliştirilmesi hedeflenen web tabanlı bir uygulama için düşünülmüştür. Söz konusu sistemde, okuyucular ve kitaplar yer almaktadır. Her bir kullanıcının, okumuş olduğu kitaplardan oluşan sanal kitaplığı bulunmaktadır.



Şekil 1. Sanal kitaplık örnek ekran görüntüsü

Şekil 1’de uygulamadaki örnek kullanıcı kitaplığı paylaşılmıştır. Kullanıcılar kitap okudukça sanal kitaplıklarını güncelleyerek, kitaplara ilişkin yorum ve değerlendirmelerini site üzerinden girebilmektedir.

#### A. Kullanılan Veri Kümesi ve Yapısı

Tamamı Türkçe kitaptan oluşan veri tabanında 133.386 adet kitap bulunmaktadır. Her bir kullanıcının okumuş olduğu kitaplar ve her bir kitap için yapmış olduğu değerlendirmeler uygulama kayıtlarında tutulmaktadır. Her bir kitap için veri tabanında bulunan alanlar ve bu alanların örnek bir kitap üzerinde gösterimine Tablo 1’de yer verilmiştir.

|                         |  |
|-------------------------|--|
| <b>Kitap Adı</b>        | Şu Çılgın Türkler                                    |
| <b>Yazar</b>            | Turgut Özakman                                       |
| <b>Yayınevi</b>         | Bilgi Yayınevi                                       |
| <b>Kategorileri</b>     | Edebiyat, Tarih                                      |
| <b>Alt kategorileri</b> | Tarihi Roman, Türk Edebiyatı, Siyasi ve Askeri Tarih |
| <b>Dizi</b>             | Roman Dizisi   |
| <b>Etiket</b>           | Atatürk, Kurtuluş Savaşı                             |

Tablo 1. Özniteliklerin örnek bir kitap üzerinde gösterilmesi

#### B. Uygulama Gerçekleştirme Adımları

1) Okuyucu Profili Çıkarılması: Okuyucu profili, tavsiyede bulunulacak kullanıcının okuma geçmişi taranarak, okuduğu kitapların yazarlarının, kategorilerinin, alt kategorilerinin, etiketlerinin, dahil oldukları seri ve dizilerin ayrı ayrı listelerinin çıkarılmasıyla oluşturulmaktadır.

2) Aday Kitap Havuzu Oluşturulması: Aday kitap havuzu, okuyucu profili çıkartma adımıyla oluşturulan listelerdeki her bir öznitelikle ilişkili kitapların bütünüdür kapsar.

Yalnızca kullanıcıya önerilebilecek kitaplar arasından değerlendirme yapabilmek için, bütün veri tabanı içerisinde okuyucunun kitaplığında yer alan kitaplarla ortak özniteliklere sahip kitaplar (aynı yazar, aynı alt kategori, etiket gibi) aday kitap havuzunu oluşturur.

3) Değerlendirme İşlemi: Her bir aday öneri setindeki kitaplar, tek tek okuyucu profili ve kullanıcının puanlaması göz önüne alınarak değerlendirilir. Değerlendirme işlemi, okuyucu profilindeki kategori, alt kategori, etiket, yazar dağılımına ve ne oranda örtüştüğüne, okuyucunun kitap puanlamasına bakılarak her bir kitap için puanlama yapılır. Öneri setindeki her bir kitabın puanlarının toplamı, o öneri setinin puanını oluşturacaktır.

#### C. Genetik Kodlama

1) Kromozom: Her bir öneri sonucunda okuyucuya 12 adet kitap sunulmaktadır. Dolayısıyla her bir aday çözüm 12’lik kitap setinden oluşmaktadır.

Her bir kitaba ait kategori, alt kategori, yazar, yayın evi, dizi, etiket öznitelikleri kromozomdaki genleri temsil etmektedir. Ayrıca kullanıcının kitap değerlendirmesi de (1-5 aralığında puanlama olarak), genlere ağırlıklandırılarak yansıtılmıştır.

2) Başlangıç Popülasyonu: Başlangıç popülasyonu için kitaplar, oluşturulan aday kitap havuzundan seçilir. PB popülasyon boyutu ve AKHB aday kitap havuzu boyutu olmak üzere,

$$PB = AHKB/12 \quad (1)$$

popülasyon boyutu (1)’deki gibi hesaplanmaktadır ve her kullanıcı için değişkendir.

Popülasyon boyutunca aday kitap havuzundan rassal olarak seçilen her 12’lik kitap, bir aday çözümü temsil etmektedir.

3) Uygunluk Fonksiyonu: Öneri setindeki her bir kitap, kullanıcı tarafından ağırlıklandırılabilen katsayılarla uygunluk fonksiyonuna tabi tutulur. Değerlendirilen öznitelikler, kullanıcı okuma profiliyle uyuşan yazar, kategori, alt kategori, etiketler şeklindedir. Şekil 2’de uygunluk fonksiyonuna ait kaba kod gösterimine yer verilmiştir.

Değerlendirme adımıyla her bir öznitelik, her kullanıcı için sonuca farklı ağırlıklarda etki etmektedir. Dolayısıyla öznitelikler dinamik olarak ağırlık kazanmaktadır.

Popülasyondaki her bir aday çözüm seti için;  
Çözüm setindeki her bir kitap için;  
Kitaptaki her bir nitelik için;

$$\text{Birey Puanı} += (\text{Öznitelik katsayısı}) \times (\text{Okuyucu profilinde yer alma oranı})$$

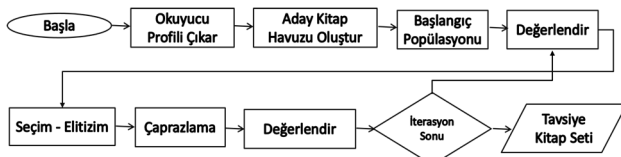
Şekil 2. Uygunluk fonksiyonu kaba kod gösterimi

4) Çaprazlama ve Seçim Yöntemi: Uygulamada tek noktali çaprazlama kullanılmış ve ebeveyn seçim tekniği olarak rulet tekerleği yöntemi seçilmiştir. Ebeveyn iki birey, her iterasyonda rastgele belirlenen 4 ardışık noktadan (kromozom boyutunun 1/3'ü kadar) çaprazlanırlar.

5) Elitizm: İterasyon süresince uygun bireylerin yitirilmesini önlemek amacıyla, popülasyon boyutunun %10'u kadar birey doğrudan bir sonraki nesle aktarılmıştır.

6) Sonlandırma: Algoritma, belirlenen sayıda iterasyon gerçekleştiğinde sonlandırılmaktadır.

Şekil 3'te uygulamanın akış diyagramı görülmektedir.



Şekil 3. Uygulama akış diyagramı

### III. DENEYSEL ÇALIŞMA

Literatürde tavsiye sistemleri değerlendirmeleri geçmiş kullanıcı tercihleri baz alınarak yapıldığı gibi, kullanıcı geri bildirimleri alınarak yapılan değerlendirme yaklaşımları bulunmaktadır. Geliştirilen tavsiye sistemi, farklı sayıda kitap okuyan 15 kişinin kayıtları sisteme girilerek, üretilen tavsiye sonuçları üzerinden iki yaklaşımla değerlendirilmiştir.

İlk yaklaşımda, üretilen tavsiye sonuçları kullanıcı seçimleriyle karşılaştırılarak değerlendirme yapılmaktadır. İkinci yaklaşımda ise, tavsiye sonuçları kullanıcılara sunulurken, önerilen her bir kitap için değerlendirmede bulunmaları istenilmiştir ve sistem değerlendirmesi kullanıcıların geri bildirimleri üzerinden yapılmaktadır.

|                            | 1. Yaklaşım   | 2. Yaklaşım  |
|----------------------------|---|--|
| <b>Doğru Pozitif (DP)</b>  | Kullanıcının pozitif değerlendirdiği ve sistem tarafından önerilen kitaplar       | Sistem tarafından önerilen ve sonrasında kullanıcı tarafından onaylanan kitaplar       |
| <b>Yanlış Negatif (YN)</b> | Kullanıcının pozitif değerlendirdiği ancak sistem tarafından önerilmeyen kitaplar | Kullanıcının pozitif değerlendirdiği ancak sistem tarafından önerilmeyen kitaplar      |
| <b>Doğru Negatif (DN)</b>  | Kullanıcının negatif değerlendirdiği ve sistem tarafından önerilmeyen kitaplar    | Kullanıcının negatif değerlendirdiği ve sistem tarafından önerilmeyen kitaplar         |
| <b>Yanlış Pozitif (YP)</b> | Kullanıcının negatif değerlendirdiği ancak sistem tarafından önerilen kitaplar    | Sistem tarafından önerilen ancak sonrasında kullanıcı tarafından onaylanmayan kitaplar |

Tablo 2. Değerlendirme yaklaşımlarında kullanılan sınıflandırıcı etiketleri

Değerlendirme yapılırken gözetilen sınıflandırma yaklaşımı Tablo 2'deki gibidir.

Literatürde tavsiye sistemi değerlendirmelerinde en yaygın kullanılan metrikler kesinlik ve geri çağırma metrikleridir [8, 13]. Buradan hareketle sistem değerlendirmesinde kullanılan kesinlik, geri çağırma, doğruluk ve F-Ölçütü değerleri 2-5 eşitliklerinde gösterildiği şekilde hesaplanmıştır.

$$\text{Kesinlik} - P (\text{Precision}) = DP / (DP + YP) \quad (2)$$

$$\text{Geri çağırma} - R (\text{Recall}) = DP / (DP + YN) \quad (3)$$

$$\text{Doğruluk} (\text{Accuracy}) = (DP + DN) / (DP + DN + YP + YN) \quad (4)$$

$$\text{F-Ölçütü} (\text{F-Measure}) = 2PR / (P + R) \quad (5)$$

#### A. Yaklaşım 1

Geliştirilen içerik tabanlı TS uygulamasını değerlendirebilmek için kitap okuyan kullanıcılara ve kullanıcıların okudukları kitaplara ihtiyaç vardı. Bu amaçla farklı sayılarda kitap okumuş 15 adet kullanıcının okuduğu kitaplar ve bu kitaplar için 1-5 aralığındaki puanlamaları sisteme girilerek tavsiye sistemi 5 kez işletildi. Bu yaklaşımda tavsiye sonuçları kullanıcı geri bildirimlerine başvurulmadan değerlendirilmiştir.

#### B. Yaklaşım 2

Kitap tavsiye sisteminde temel amaç, okuyucuyla ilgili duyacağı düşünülen kitapları karşılaştırmaktır. Dolayısıyla kendi seçimleri üzerinden üretilmiş tavsiye sonuçlarına kullanıcının geri dönüşü alınmaksızın yapılan bir değerlendirmenin eksik olacağı düşünülmüştür. Bu nedenle Yaklaşım 1'den farklı olarak, her kullanıcıdan kendileri için önerilen tavsiye çıktılarını incelemeleri ve her bir kitap için "Onaylama" ve "Reddetme" şeklinde değerlendirmede bulunmaları istenmiştir.

#### C. Değerlendirme

Yapılan değerlendirmeler neticesinde; kesinlik, geri çağırma, doğruluk ve F-ölçütü değerleri Tablo 3'teki gibi elde edilmiştir.

|                            | Yaklaşım 1 | Yaklaşım 2 |
|----------------------------|------------|------------|
| <b>Kesinlik</b>            | 0,9257     | 0,8383     |
| <b>Geri Çağırma</b>        | 0,2367     | 0,4219     |
| <b>Doğruluk</b>            | 0,5024     | 0,5629     |
| <b>F-Ölçütü</b>            | 0,3698     | 0,5569     |
| <b>Doğru Pozitif (DP)</b>  | 1,9866     | 4,64       |
| <b>Yanlış Pozitif (YP)</b> | 0,1733     | 0,9067     |

Tablo 3. Değerlendirme sonuçları

Sonuçlar kıyaslandığında, 2. yaklaşımda ortalama kesinlik değerlerinde düşüş, geri çağırma, doğruluk ve F-Ölçütü değerlerinde ise artış görülmektedir. Bu değişimlerdeki temel etken, okuyucu değerlendirmesi sonrası yeniden şekillenen DP ve YP sınıflandırıcılarıdır. DP, tavsiye çıktılarının kalitesini belirleyen sınıflandırıcıdır. Tavsiye çıktılarının ne oranda kabul gördüğünü doğrudan ortaya koymaktadır.

2. Yaklaşımda, sistem tarafından önerilen ve kullanıcı tarafından onaylanan kitaplar olarak sınıflandırılmış DP sonuçlarında, 1. Yaklaşım göre 2,3 kat artış görülmektedir. Okuyucular her bir öneri setinden ortalama 4,6 kitabı dikkat çekici ve okumaya değer olarak nitelendirmişlerdir. Öneri setinin 12 kitaptan oluştuğu göz önüne alınırsa, okuyucular açısından bu oranlar, “Her öneri setinde ortalama 4,6 ilgi duyacağım kitapla karşılaştım.” anlamına gelmektedir.

#### IV. SONUÇ

Çalışmada, kitap temalı bir sosyal ağ düşünülerek, çok sayıda kitap alternatifi arasından okuyuculara ilgi alanları doğrultusunda kitapların önerilmesi amacıyla, tavsiye sistemi geliştirilmiştir. Böylece kullanıcılara pek çok kitap alternatifi arasından zevkleri doğrultusunda önerilerde bulunmak, ilgi duyabileceği kitaplardan haberdar kılmak amaçlanmıştır. Gerçekleştirilen uygulama ile yalnızca popüler olan içerikler değil, bilinirliği az olan ancak kullanıcı profiliyle uyuşan içerikler de gün yüzüne çıkarılarak okuyucuyla buluşturulmaktadır.

Kitap tavsiyeleri, haber, internet içerikleri, müzik gibi tavsiyelerin aksine çabuk tüketilebilen tavsiyeler olmayıp, kullanıcıların daha uzun zaman dilimlerinde gündemlerine alarak değerlendirecekleri tavsiyelerdendir. Yine önerilen bir kitap, kullanıcının ilgisini hemen çekebilme ancak kullanıcı okuma safhasına aylar sonra geçebilmektedir. Bu sebeple yalnızca kullanıcıların geçmiş seçimleri gözetilerek yapılan sistem değerlendirmelerinin yetersiz olduğu düşünülmekte, tavsiyelerden sonraki kullanıcı davranışları ya da geri bildirimlerinin değerlendirmeye katılması önemli bulunmaktadır.

Kitap öneri sistemi genetik algoritma tabanlıdır. Geliştirilen tavsiye modeli içerik tabanlı filtrelemeye dayalı olup, genetik algoritmanın yapısı gereği, elde kullanıcı verileri biriktikçe kolaylıkla ortak filtreleme yaklaşımıyla birlikte hibrid yapıda tavsiye yaklaşımını da destekleyebilecektir. Bu adımla daha etkili tavsiye sonuçlarının elde edilebileceği düşünülmektedir.

#### KAYNAKÇA

[1] Su, J.-H., Wang, B.-W., Hsiao, C.-Y. ve Tseng, V. S., “Personalized Rough-Set-Based Recommendation By Integrating Multiple Contents And Collaborative Information”, Information Sciences, Cilt 180, No 1, 113–131, 2010.

[2] Linden, G., Smith, B. ve York, J., “Amazon.com Recommendations Item to Item Collaborative Filtering”, IEEE Internet Computing, Cilt 7, No 1, 76-80, 2003.

[3] Göksedef, M. ve Gündüz-Öğüdücü, Ş., “Combination of Web page recommender systems”, Expert Systems with Applications, Cilt 37, No 4, 2911–2922, 2010.

[4] Kim, H.-T., Kim, E., Lee, J.-H. ve Ahn, C. W., “A recommender system based on genetic algorithm for music data”, Computer Engineering and Technology (ICCET), 2010 2nd International Conference, Chengdu, 414-417, 16-18 Nisan 2010.

[5] Park, D. H., Kim, H. K., Choi, I. Y. ve Kim, J. K., “A literature review and classification of recommender systems research”, Expert Systems with Applications, Cilt 39, No 11, 10059–10072, 2012.

[6] Lee, C., Kim, Y. ve Rhee, P., “Web personalization expert with combining collaborative Filtering and association rule mining technique”, Expert Systems with Applications, Cilt 21, No 3, 131–137, 2001.

[7] Liu, N.-H., Hsieh, S.-J. ve Tsai, C.-F., “An intelligent music playlist generator based on the time parameter with artificial neural networks”, Expert Systems with Applications, Cilt 37, No 4, 2815–2825, 2010.

[8] Düzgün, S. S., “BORGO: A Book Recommender For Reading Groups”, Yüksek Lisans Tezi, Orta Doğu Teknik Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 2012.

[9] Fong, S., Ho, Y. ve Hang, Y., “Using Genetic Algorithm for Hybrid Modes of Collaborative Filtering in Online Recommenders,” Hybrid Intelligent Systems, HIS '08. Eighth International Conference, Barselona, 174-179, 10-12 Eylül 2008.

[10] Bobadilla, J., Ortega, F., Hernando, A., ve Alcalá, J., “Improving collaborative filtering recommender system results and performance using genetic algorithms”, Knowledge-Based Systems, Cilt 24, No 8, 1310–1316, 2011.

[11] Sarwar, B., Karypis, G., Konstan, J. ve Riedl, J., “Analysis of recommendation algorithms for e-commerce”, In Proceedings of the 2nd ACM conference on Electronic commerce (EC '00), ACM, New York, 158-167, 17-20 Ekim 2000.

[12] Chang, Y., Li J. ve Ku, H., “A Genetic Algorithm (GA) - Based Personalized Learning Service in Cloud Learning Environments”, IEEE Technology and Engineering Education (ITEE), Cilt 7, No 2, 28–32, 2012.

[13] Bobadilla, J., Ortega, F., Hernando, A., Gutiérrez, A., “Recommender systems survey”, Knowledge-Based Systems, Cilt 46, 109-132, 2013.