

Yıldız Teknik Üniversitesi Bilgisayar Mühendisliği Bölümü

<u>DERS</u> BLM3760 – UZMAN SISTEMLERE GİRİŞ

<u>KONU</u> Kitap Öneri Sistemi

DERSİN YÜRÜTÜCÜSÜ M. Fatih AMASYALI

ÖĞRENCİ NUMARA: 13011704 AD: Abdullah SOYAD: KAYAHAN

TANIM

Projem Kitap Öneri Sistemidir. Bu sistem, kullanıcıdan alınan maksimum 10 kitaba karşılık, kullanıcının daha önceden okuduğu kitapları tahmin etme ve/veya beğeneceği kitapları önermek üzerine tasarlanan bir sistemdir.

Sistem Visual Studio 2015 üzerinde C# Dilinde masaüstü uygulama şeklinde tasarlanmıştır. Veritabanı olarak MSAccess kullanılmışıdır.

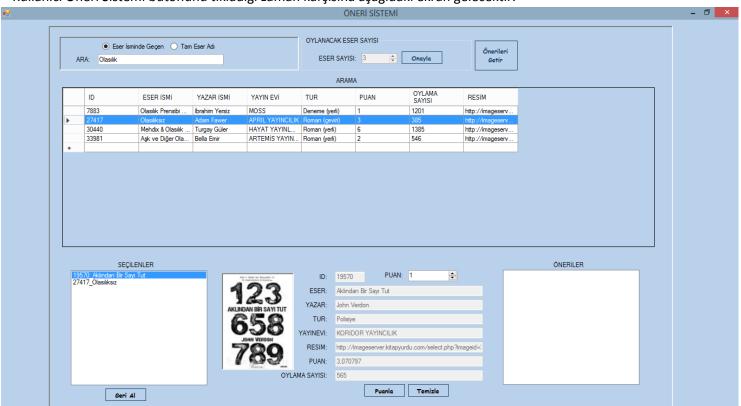
SİSTEMİN ÇALIŞMASI

Sistem açıldığında kullanıcı bir ana sayfa ekrani ile karşılaşmaktadır.



Bu ana ekrandan hem internetten veri çekme kısmına hemde öneri sistemine ulaşılabilir. Programa bir kere veri çekmek yeterli olduğundan "veri al" kısmı şimdilik anlatılmayacaktır.

Kullanıcı Öneri Sistemi butonunu tıkladığı zaman karşısına aşağıdaki ekran gelecektir.



Bu ekranda kullanıcının sırası ile şunları yapması beklenir;

Kullanıcı, sisteme girdi olacak kaç kaç kitap vereceği bilgisini girer bu sayı 1 ile 10 arasında sınırlandırılmıştır.

Kitap sayısını girdikten sonra kullanıcı kitabı arar gelen listeden seçerek ekleme işlemi yapar. Bunu girdiği sayı kadar tekrar eder. Sayıya ulaştığı zaman önerileri getir butonu aktif olur.

Kullanıcı yanlış bir kitap eklerse geri al butonunu kullanarak listeden kaldırabilir yerine yeni kitap girebilir.

Girdiği kitaplardan emin olduktan sonra ve gireceği kitap sayısını tamamladıktan sonra aktif olan önerileri getir butonuna basarak bir sonraki adıma geçer.

Ayrıca kullanıcı öneri getirmeksizin seçtiği kitabı sol alt tarafta ki listeden seçerek puanlayabilir.

Önerileri Getir Butonuna basıldığında gelen ekran aşağıdaki gibidir.

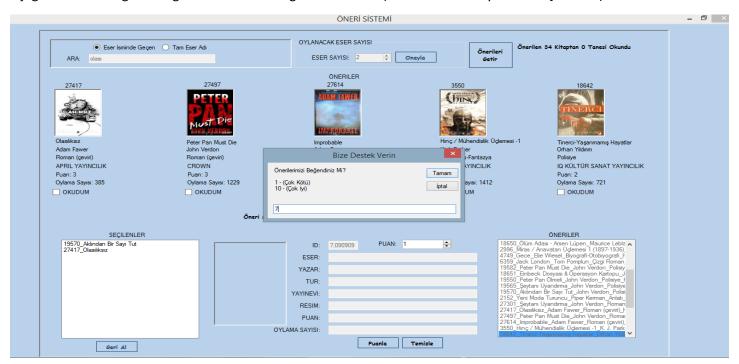


Bu ekranda karşımıza 5 adet resim kutusu, ileri-geri butonları ve her resim kutusunun altında okuduğumuzu işaretleyebilmemiz için birer onay kutusu çıkmaktadır. Bu ekran oluştururken kullanıcıdan alınan kitaplar referans alınarak veri tabanından girilen kitaplara benzer (tür, yazar, yayın evi bazında benzerlikler) kitaplar kullanıcıya getirilmektedir. Kullanıcıya öneri olarak getirilen kitapların sayısı sisteme girdi olarak verilen kitapların sayısı ile doğru orantılı olarak artmaktadır. Getirilen öneriler 5 e bölünerek sayfalar halinde kullanıcıya sunulmaktadır. Kullanıcı ileri-geri butonları ile sayfalar arasında gezinebilmektedir.

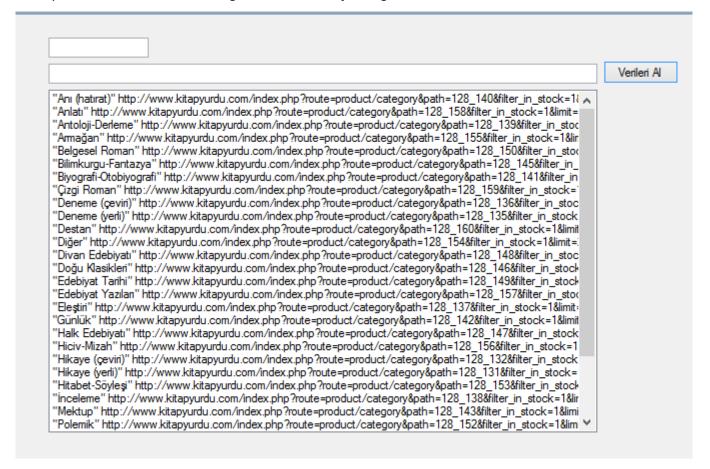
Kullanıcıdan beklentiler, bu sayfa için kullanıcıdan sistem başarısını ölçmek adına daha önce okuduğu kitapları işaretlemesi beklenmektedir. Ayrıca kullanıcı bu sayfadaki resimlere tıklayarak kitaplara puan verebilir.

Kullanıcıdan sistemden çıkarken önerilen kitapları değerlendirmesi istenilir. Kullanıcı 1 çok kötü 10 çok iyi olmak üzere bir değer girerek yine sistem başarısını ölçmek adına geri dönüş bilgisi vermesi beklenmektedir.

Aşağıdaki resimde genel değerlendirme ekranı görünmektedir. (Öneri Sistemi Sayfasından Çıkarken)



Ana sayfada Veri Al butonuna basıldığında bu ekran karşımıza gelmektedir.



Bu ekranda orta kısımda görülen listBox'ta http://www.kitapyurdu.com/kategori/kitap-edebiyat/128.html adresindeki alt kategorilerinin linklerinin düzenlenmiş hali bulunmaktadır. Aşağıda listbox içinde bulunan bir link örneği verilmiştir.

"http://www.kitapyurdu.com/index.php?route=product/category&path=128 144&filter in stock=1&limit=2&page="

Yukarıdaki linkte fontu kalın olarak gösterilen yerler özellikle ayarlanmalıdır. Limit değeri bir sayfa içinde kaç kitabın geleceğini, page değeri ise o an hangi sayfada olacağımızı göstermektedir. Link listbox'a eklenirken bir sayfa numarası verilmemiştir. Sayfa numarası program içerisinde döngü ile verilip her sayfayı gezmesi sağlanmıştır.

Her sayfada iki kitap getirilmesi: C# ile web üzerinden veri çekerken kullanılan fonksiyonlara pek aşına olmadığım için her sayfada iki kitap getirip kullandığım fonksiyonda ki "IndexOf" ve "LastIndexOf" methodlarını kullanarak her seferinde iki kayıt almaya odaklandığım için her sayfada iki kitap gerilmektedir.

Not: Bu işlem veri alma sürecini biraz uzatmaktadır.

	kitap1							
4	id ▼	Eser →	Yazar 🕶	YayinEvi -	Tur →	Resim -	Puan 🔻	OylamaSayi: 🕶
	1	Life With a Vie	Toni Sepeda	ÇİTLEMBİK YAY	Anı (hatırat)	http://imagese	7,986238	436
	2	Öyle Bir Aşk	Yusuf Ziya Bah	İMGE KİTABEVİ	Anı (hatırat)	http://imagese	6	540
	3	Çocukluğum /	Lev N. Tolstoy	ANTİK YAYINLA	Anı (hatırat)	http://imagese	1	81
	4	Kendim İçin Ar	Domenica Ruta	PALOMA YAYIN	Anı (hatırat)	http://imagese	4	27
	5	Anılar	Doris Lessing	KIRMIZI KEDÎ Y.	Anı (hatırat)	http://imagese	6	1349
	6	Arabanın Dili C	Burçin Yıldırım	FERFİR YAYINL	Anı (hatırat)	http://imagese	7	372
	7	Ben Nojoud 10	Nojoud Ali	MARTI YAYINL	Anı (hatırat)	http://imagese	9	598
	8	Kemal Hadi Ge	Gül Sunal	DOĞAN KİTAP	Anı (hatırat)	http://imagese	6	621

Veri tabanından bir kesit.

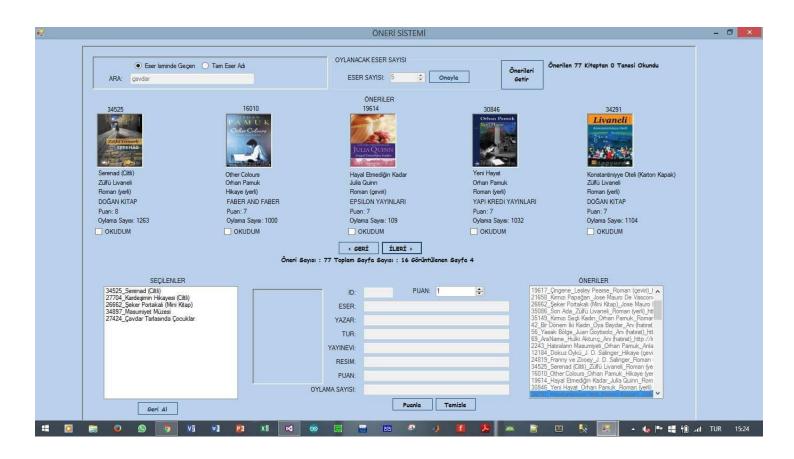
GELİŞTİRME SÜRECİNDE YAŞANANLAR

1- Veri tabanının oluşturulması;

- a. Veri tabanı http://www.kitapyurdu.com/kategori/kitap-edebiyat/128.html adresindeki kitaplardan oluşturulmuştur. Site içerisinde edebiyat kategorisine ait 32 adet alt kategori bulunmaktadır. Bu alt kategorilerin hepsinin içerisindeki kitapları veri tabanına çekmek için 32 adet linkin hazırlanıp programın veri al kısmına girilmesi gerekti.
- b. Veri al kısmında her sayfada 2 kitap getirilecek şekilde linkler düzenlenerek bir döngü yardımı ile 32 alt kategorinin tamamındaki veriler alındı.
- c. Proje önerimde veri tabanımım 500 kitaptan oluşacağını söylemiştim lakin veri almanın kolaylığından kaynaklı veri tabanımda 42000 den fazla kitap yer almaktadır.

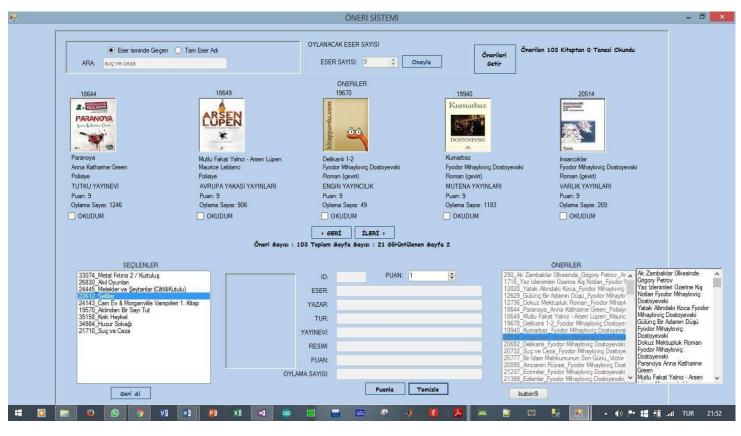
2- Sorgulama sorunu;

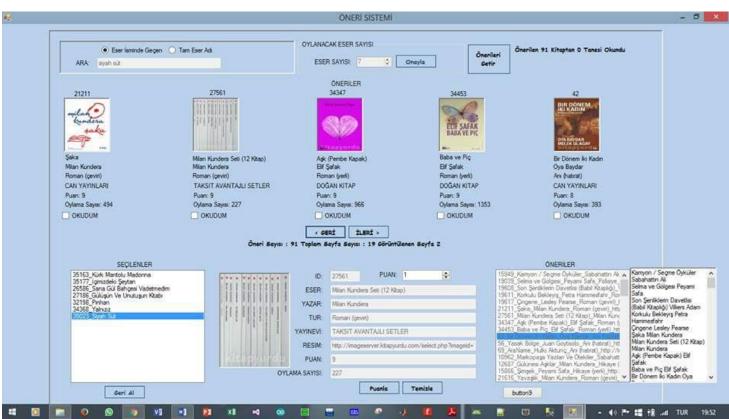
- a. Veri tabanına kitaplar yazılırken hiçbir kısıtlama getirilmemiş sitede nasıl yazılmışsa aynen veri tabanına aktarılmıştır. Bu verileri sorgularken (tam kitap ismine göre) bazen kitaplar veri tabanında olmasına rağmen bulunamamaktadır. Bunun nedeni ise kitap isminin yanında kitabın cilt özelliğinin yazılmasıdır. Örneğin: Serenad (ciltli) ve Serenad (karton kapaklı) gibi.
- b. Tekrarlı veri sorunu yaşanmaktadır. Bir kitabı özellikle dünya klasiklerini birden fazla yayın evi bastığı için sorgulama esnasında aranan kitabın birden fazla çıktığı gözlenmiştir. Her ne kadar sorguda tekrarsız verilerin getirilmesi istense de ID bazında sınırlandırıldığı için bu işlem işe yaramamıştır.
- c. Kitap türü problemi, bu problemde ise bir kitap hem polisiye hem Roman (Yerli) veya Roman (Çeviri) kategorisinde yer alabilmektedir. Örneğin "Aklından Bir Sayı Tut" kitabı Polisiye türünde bir kitaptır ancak isim olarak aratıldığında kitaptan 2 adet çıkmaktadır diğer gelen kayıt Roman (Çeviri) türündedir. Kullanıcının burada ki seçimi önerilecek olan kitapları etkilemektedir.
- d. Box Set kitaplar. Öneriler içerisinde gelen set halinde satılan kitaplar mevcut. Örneğin John Verdon Polisiye Roman Seti gibi. Kullanıcılar setin tamamını okumasa da set içindeki bazı kitapları okumuş oluyor ve bunlar tek tek gelmediği zaman(Okudum Olarak İşaretleyemeyince) sistem başarısını olumsuz yönde etkiliyor.





Aşağıdaki iki ekran görüntüsü proje testi sırasında çekilmiştir. Kullanıcılar programı uzaktan kullanamadıkları için önerileri sağ altta görülen bir richTextBox'a yazdırarak kendilerine gönderdim ve geri dönüşler üzerine programın ortalama başarısını hesaplamaya çalıştım.





KULLANICI	Girilen Kitap Sayısı	Önerilen Kitap Sayısı	Önerilenler İçinde Okunan Kitap Sayısı	Başarı Yüzdesi	Önerileri Beğenme
1	5	77	5	6,49%	70%
2	5	96	18	18,75%	80%
3	10	113	16	14,16%	75%
4	4	57	9	15,79%	70%
5	10	112	11	9,82%	60%
6	10	101	14	13,86%	75%
7	7	126	7	5,56%	65%
8	9	103	12	11,65%	85%
9	4	109	10	9,17%	50%
10	7	91	15	16,48%	80%
ORTALAMA	7,1	98,5	11,7	12,17%	0,71

Yandaki tabloda 10 kullanıcıya ait bilgiler bulunmaktadır. Bu bilgiler kullanıcılardan kitap isimleri isteyip onlara gelen önerileri göndererek elde edilmiştir.

Gönderilen kitaplar arasından kullanıcının okuduğu kitap sayısını ve gönderilen önerileri genel olarak yüzde kaç oranında beğendiği sorulmuştur. Alınan cevaplar ışığında tablo doldurulmuştur.

Ortalama Başarı: %12,17 Ortalama Beğenilme: %71,00

- 1- Sistem Başarısı: sistem başarısı iki farklı şekilde ölçülmektedir bunlar;
 - a. Önerilen kitaplar içinden daha önce okuduğu kitap yüzdesi.
 - i. Kullanıcıya girdiği kitaplara göre alternatif kitaplar önerilmektedir. Bu kitaplar içinde kullanıcının daha önce okuduğu kitapları işaretlemesi istenir. (İşaretlenen Kitap Sayısı * 100)/ (Önerilen Kitap Sayısı) Formülü ile kullanıcının önerilen kitaplar içinden daha önce okuduğu kitapların yüzde oranı ortaya çıkarılmaktadır. Bu daha önce okuduğu kitapları tahmin etme başarısıdır.
 - b. Önerilen kitapları beğenme yüzdesi.
 - i. Kullanıcıya program çıkışında sorulan önerilen kitapları beğenme yüzdesidir. Program çıkışında kullanıcının 1 ile 10 arasında önerilen kitapları beğenip beğenmediğini girmesi istenilmektedir. Bu da önerdiğimiz kitapların başarısını göstermektedir.

2- Çıkarım:

- a. Okuduğu kitap sayısı az olan kişilerde programın başarı yüzdesi oldukça düşük çıkmıştır (Bknz 1. Ve 7. Kullanıcı). Bunun sebebi okuduğu tüm kitapları girdi olarak vermektir. Sonuç olarak önerilen kitaplar içinde sadece okuduğu kitaplar olmaktadır. Bazen okunan ancak puanı düşük olan kitaplar listede yer alamamıştır. Bu da programın başarısını düşüren faktörlerden biridir.
- b. Puan bazlı sıralama yapıldığından eğer kişinin okuduğu kitaplar düşük puanlı ise önerilen kitaplar arasında çıkmama ihtimali olmaktadır.

3- **Yorum:**

- a. Programa kendi okuduğum kitapları girdiğim zaman çok kitap okuyan biri olmadığım için programın önerdiği bazı kitapların isimlerini bile yeni duydum. Bunun bir başka sebebinin de veri tabanındaki kitapların sayısının fazla olması olduğunu düşünüyorum.
- b. Gelen geri dönüşlerde çoğunda kullanıcılar genel olarak yazar benzerliklerinin daha ön planda olduğunu ancak türün ön planda olması gerektiğini söylediler. Lakin programdaki SQL sorgum içinde bunu nasıl sağlayacağımı çözümleyemedim.
- c. 10 kişi üzerinden yaptığım testler sonucunda sistemin başarısını 1. kritere göre %12,17, 2. kritere göre %71 buldum. Ancak bu sonuçların gerçek başarının üstünde olduğunu düşünmekteyim. Programı yazan ve test eden kişi olarak düşündüğüm zaman sistemin başarısı 1. kriterde %10, 2. kriterde %60 civarında olduğunu düşünmekteyim.

KAYNAKÇA

- 1- http://www.kitapyurdu.com/kategori/kitap-edebiyat/128.html
- 2- http://www.gencayyildiz.com/blog/c-ile-bot-yapimi-web-sitelerinden-icerik-cekme/
- 3- http://www.kursadaltan.com/c-ile-bot-yapimiweb-sitelerinden-veri-cekme.html
- 4- https://msdn.microsoft.com/tr-tr/library/456dfw4f(v=vs.110).aspx
- 5- Lise ve Önlisans Proje Arşivim