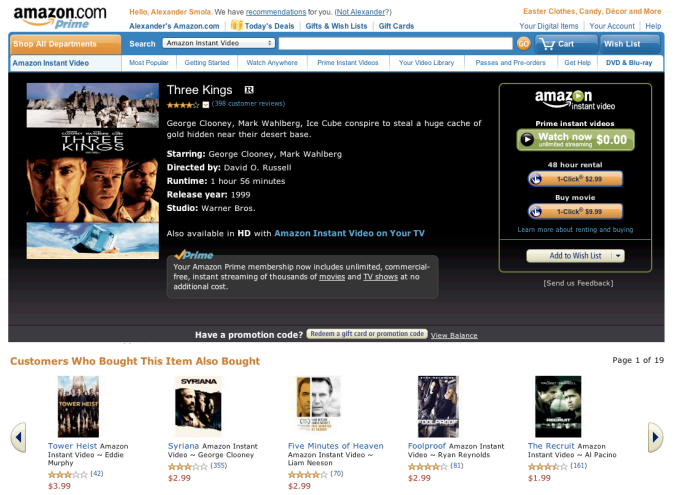
**Recommendation Systems (Tavsiye Sistemleri) Nedir?**

Günümüzde kullanıcılara sunulan ürün ve hizmetler çok yüksek boyutara ulaştığından dolayı kullanıcıların beğenecekleri ürünleri klasik yöntemlerle arayıp bulmaları giderek zorlaşmaktadır. Bu noktada tavsiye sistemleri devreye girerek devasa bilgi ve ürün yığınları arasından kullanıcılara ilgilerini çekebilecek veya beğenecekleri ürünleri bulmalarında yardım ederler.  Kullanıcıların ürünlere verdikleri oylar, ziyaret ettikleri sayfalar, veya arkadaşlık ilişkileri gibi veriler kullanılarak kullanıcı modelleri oluşturulur, bu modeller kullanılarak kullanıcıların sistemde varolan ürünleri ne kadar beğenebilecekleri tahmin edilir ve yüksek beğeni tahmini yapılanlar kullanıcıya tavsiye edilir. Tavsiye sistemleri film tavsiyesi, kitap tavsiyesi, restoran tavsiyesi, otel tavsiyesi, haber tavsiyesi ve müzik tavsiyesi gibi bir çok alanda kullanılabilir. Her kullanıcının istekleri farklıdır. Bu yüzden farklı ürünlere ilgi duyabilirler. Kullanıcılar ilgi alanlarına tam olarak uyan ürünlere daha iyi tepki verirler.

Ticari kullanımda olan tavsiye sistemleri :

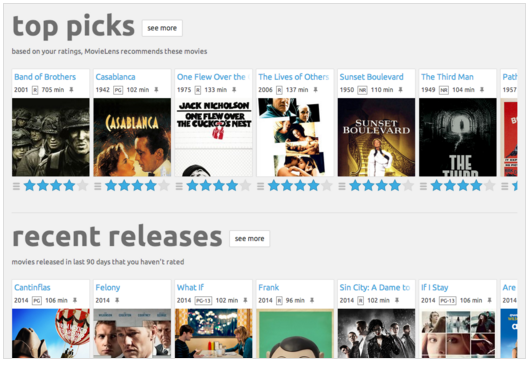
* **Amazon :** Dünyanın en büyük internet alışveriş sitesi olan Amazon, kullanıcının incelemekte olduğu öğe ile benzer olan diğer öğeleri tavsiye etmektedir. Çok bilinen “Bu ürünü alan şunları da aldı.” şeklinde sunulan öneriler çoğunluktadır.



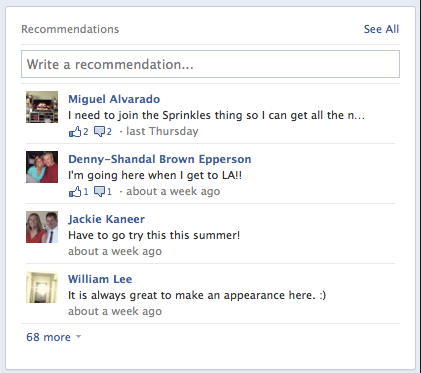
* **Netflix :** Üyelere online film kiralama servisi sunan Netflix, öneri sistemleri konusunda başlattığı 1 milyon dolar ödüllü yarışma ile ses getirmiştir. Netflix, kullanıcılara film önerisinde bulunan bir öneri sistemine sahiptir.



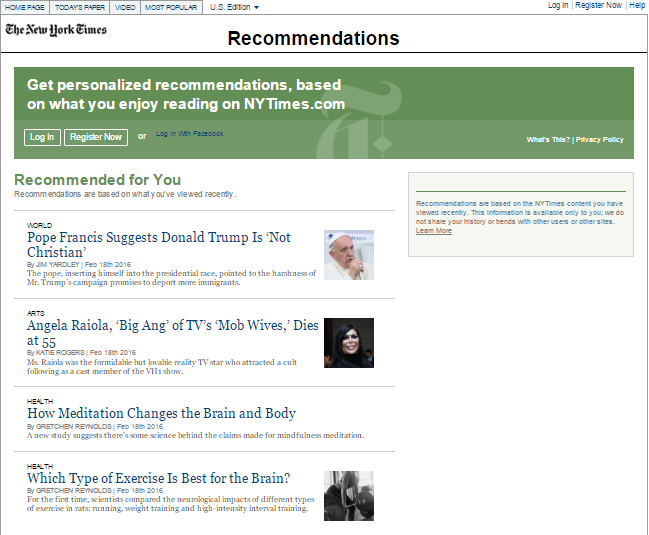
* MovieLens



* **Facebook :** Dünyanın en büyük sosyal ağlarından birisi olan Facebook kullanıcıların tanıyor olabilecekleri kişileri önermektedir.



* **NewYork Times :** ABD'nin ünlü NewYork Times gazetesi üye olan kullanıcılarına haber öneri servisi sunmaktadır.



**Tavsiye Sistemleri Yaklaşımları**

Tavsiye sistemlerinin geliştirilmesi için kullanılan farklı yaklaşımlar bulunmaktadır. Kimi tavsiye sistemleri önerilerde bulunmak kullanıcılar arasındaki benzerlikleri göz önüne alırken, diğerleri içerik analizi ve içerik benzerliklerini kullanmaktadır. Literatürde bu yaklaşımlar genellikle işbirlikçi filtreleme, içerik tabanlı fitreleme ve hibrid olarak üç kısma ayrılmaktadır. Ancak tavsiye yaklaşımlarını daha farklı sınıflandırmak da mümkündür. Örneğin; beş farklı kısımda sınıflandırdığımız da işbirlikçi filtreleme, içerik tabanlı filtreleme, demografik, yardımcı tabanlı ve bilgi tabanlı olarak adlandırılmaktadır.

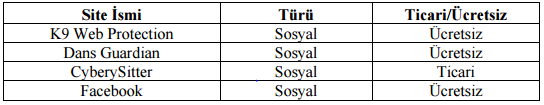
**Content-based Recommendation (İçerik Tabanlı Filtreleme)**

İçerik tabanlı filtreleme sistemi bir kullanıcıya, kullanıcının ilgisine ait belgiye veya öğenin tanımına dayalı bir öğe önerir. İçerik tabanlı filtreleme sistemleri web sayfaları, haber makaleleri, restoran, televizyon öneri programları ve çevrimiçi satış siteleri gibi çeşitli alanlarda kullanılabilir. Her ne kadar sistemler birbirinden farklı olsa da, içerik tabanlı sistemler önerilmiş öğenin tanımlanması, kullanıcı belgisinin oluşturulması ve de kullanıcı belgisindeki öğelerin karşılaştırılması için ortak bir araçtır. Kullanıcı belgisi genellikle kullanıcıya sunulmuş öğelere ait isteklerin geri cevaplarına göre otomatik olarak oluşturulur ve güncellenir. İçerik tabanlı öneri sistemi öğeleri kullanıcı tercihlerine ve öğelerin içerikleri arasındaki ilintiye göre seçim yapar.

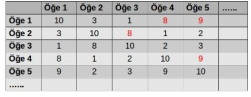
İçerik tabanlı filtreleme yaklaşımında ürünlerin ve kullanıcı tercih bilgilerinin içeriğinden elde edilen özellik vektörleri kullanılmaktadır. Ürün ve kullanıcı tercih vektörleri arasındaki benzerlikler bulunarak önerilecek ürünler belirlenmektedir.

İçerik tabanlı öneri sistemlerinde sınıflandırma algoritmaları anahtar bileşenlerdir. Çünkü her kullanıcının ilgi alanlarını modelleyen işlevleri öğrenirler. Yeni bir öğe ve kullanıcı modeli verildiğinde işlev kullanıcın öğe ile ilgilenip ilgilenmeyeceğini öngörür.

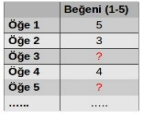
İçerik tabanlı öneri sistemlerinin kullandığı ticari ve ticari olmayan popüler siteler :



İçerik tabanlı öneri sistemlerinde bir öneride bulunmak için öğelerin özellikleri kullanılmaktadır. Bu tür sistemlerde kullanıcının geçmişte tercih ettiği öğelerle ortak özellikleri bulunan yeni öğeler kullanıcıya tavsiye edilir. Şekil 1'de öğeler arasındaki benzerlik oranlarının tutulduğu matrise bir örnek verilmiştir. Buna göre, örneğin öğe 2 ve öğe 3 arasında yüksek oranda benzerlik bulunmaktadır. Şekil 2'de ise bir kullanıcının hangi öğeleri daha çok beğendiğine dair bir matris verilmiştir. Bu örneğe göre, eğer kullanıcı 2 numaralı öğeyi beğenmişse, 2 ve 3 numaralı öğeler benzer olduğu için bu kullanıcıya 3 numaralı öğeyi önermek mantıklı olacaktır. Öğeler arasındaki benzerlikler hesaplanırken çeşitli yöntemler ve parametreler kullanılabilir. Metin analizi, anahtar kelime benzerlikleri gibi içerikteki her türlü benzerliğin tespit edilmesi ile öneri sistemine dahil edilmesi içerik tabanlı filtreleme yöntemine girmektedir. İçerik tabanlı filtreleme yönteminin kökenleri bilgi elde etme yöntemlerine dayanmaktadır . Bazı çalışmalarda içerik tabanlı filtreleme yerine bilgi filtreleme (information filtering) terimi kullanılmıştır ancak tamamen aynı yöntemden söz edilmektedir.

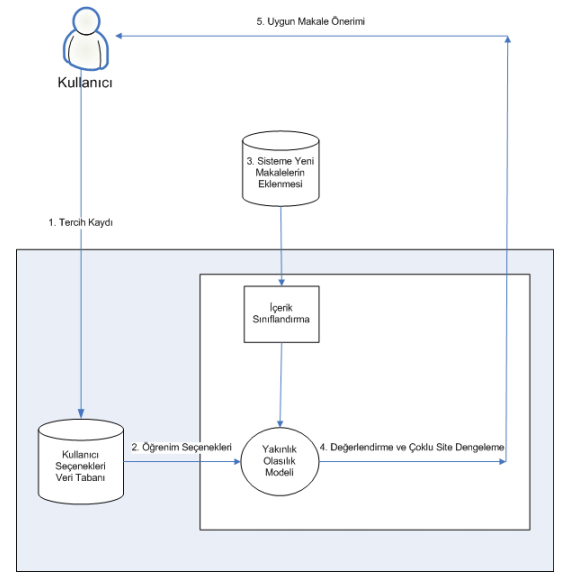


**Şekil 1 -** Öğeler arası benzerlik matrisi



**Şekil 2 –** Bir kullanıcının öğeleri beğenme oranı gösterme matrisi

5 adımdan oluşan içerik tabanlı tavsiye sistemi şeması :



1) Kullanıcı tercih kaydı yapılır ve tercihler veri tabanına kayıt edilir.

2) Tercihler olasılık modeli ile eğitilir.

3) Sisteme kayıt edilen makaleler anahtar kelimelere yakınlıklarına göre sınıflandırılır.

4) Kestirim, eğitilmiş model kullanılarak yüklenmiş makalelerin değerlendirilmesi ile yapılır ve çoklu site dengelemesi yapılır.

5) Son olarak uygun makale önerilir.

**İçerik Tabanlı Filtreleme Sisteminde Olasılık Modelleme**

İçerik tabanlı filtreleme sistemlerinde olasılık modellemesini anlamak için örnek bir uygulama olarak kitap öneri sistemi ele alalım. Bu kitap öneri sisteminde, seçilmiş anahtar kelimeler ve tercih edilmiş kitaplardaki değerler yeni kitapları değerlendirmek üzere olasılık modelinde kullanılır. Tercih edilen kitaplar sadece tercihlerin öğrenilmesinde kullanılır. Bu öğrenmede kullanıcı sadece tercihlerini vermek zorundadır. Bu tercih edilen kitapların elde edilmesinde yararlı bir yöntemdir. Aşağıda öğrenme sürecindeki olasılık modellemesinin detayları verilmiştir.

İçerik filtreleme gösteriminde d kitap, z demetlemeye karşılık olarak gizli bir değişken, s alan ve ya başlık veya özet, w bir kelime, ns,d(w) ise d kitabında s alanında ki w kelimesinin sayısı olarak gösterilmektedir. Kitabın olasılık yapısı aşağıdaki sonlu karışım modeli ile modellenmiştir.



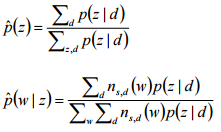
Daha sonra tercih edilmiş kitaplardan olasılığı olan p(z) ve p (w z) s | eğitilir. Bu genellikle demetleme için kullanılan bir sözde karışım modelidir. Bu modeli normal bir karışım modelinden farklı kılan önemli bir özellik ise bu modelde alanlar ayrılmaktır yani her alandaki tercihler birbirinden bağımsız olarak yakalanabilir. Aslında, burada iki alan kullanılmaktadır: başlık ve özet.

Olasılık değerleri bir EM tarafından (Beklenti-Maksimizasyon) algoritması ile eğitilir ve bu algoritmada aşağıdaki E- ve M-adımları durum memnun edici olana kadar dönüşümlü olarak devam eder.

**E-Adımı:**

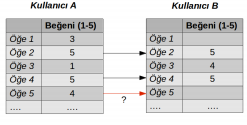


**M- Adımı:**



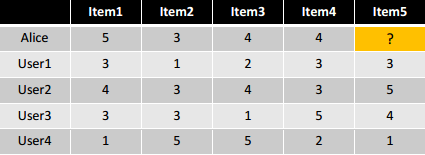
**Collaborative Recommendation (İşbirlikçi Filtreleme)**

Bu yöntemde öneriler kullanıcın diğer kişilerle olan benzer tercihleri kullanılarak yapılmaktadır. Örneğin; kullanıcının geçmişte izlediği ve beğendiği filmleri izleyip beğenen başka bir kullanıcı varsa bu iki kullanıcı arasında bağlantı kurulur, birisinin izleyip beğendiği filmler diğerine de önerilir. Şekil 3'te iki kullanıcının aynı öğeleri ne kadar beğendiğine dair tablolar verilmiştir. Buna göre, her iki kullanıcı da 2 ve 4 numaralı öğeleri aynı derecede beğenmiştir. A kullanıcısı 5 numaralı öğeyi de beğenmiştir. Ancak B kullanıcısının henüz bu öğeye dair bir geribildirimi olmamıştır. Bu durumda A ve B kullanıcısının geçmişteki benzerlikleri göz önüne alınarak, bu iki kullanıcının benzer beğenileri olduğu çıkarsanıp, 5 numaralı öğe B kullanıcısına da tavsiye edilebilir. İşbirlikçi filtreleme yönteminde kullanıcı benzerliklerinin hesaplanmasında çeşitli yöntemler kullanılmaktadır. Bayes ağları (Bayesian network) ve kümeleme (clustering) yöntemleri bu konuda en çok kullanılan yöntemlerdendir.



**Şekil 3 –** İki kullanıcı arasındaki beğeni benzerlikleri

**Örnek :** Alice ve diğer bazı kullanıcılar ve geçerli kullanıcılarının yorumu içeren bir veri tabanı verilmiştir.



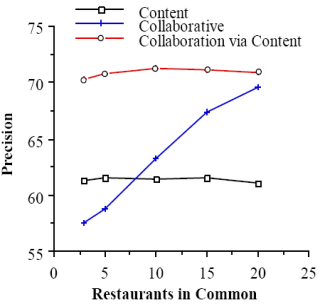
Alice item5’ i daha değerlendirmedi ya da görmedi amaç Alice’nin item5 ‘i sevip sevmeyeceğini belirlemek

**Hibrid Recommendation (Karma yaklaşım)**

Tavsiye sistemlerinin kalitesini arttırmak için bu yaklaşımda, içerik tabanlı ve işbirlikçi filtreleme yöntemlerinin birlikte kullanılır. Bu yaklaşımın amacı, tek bir yöntemin sahip olduğu dezavantajlardan mümkün olduğunca kurtulmak ve yöntemlerin avantajlarını birleştirmektir. İçerik tabanlı ve iş birlikçi filtreleme yöntemleri bir arada farklı şekillerde kullanılabilir. Karma yaklaşım öneri sistemleri 3 ayrı kümede toplanır.

İlk kümedeki sistem, içerik tabanlı ve işbirlikçi süzgeçler ile gerçekleştirilmiş öneri sonuçlarını basitçe birleştirmeye çalışır. İkinci kümedeki sistem, iki yaklaşımı ürün bilgisinin ve işlem bilgisinin bir arada olduğu tek sunum katında birleştirir. Bu sistemlerin bazıları işbirlikçi filtrelemeye dayanır ve kullanıcı gösteriminde öğelerin içerik bilgilerini ekler veya öğelerin içeriklerine bağlı olarak oyları ekler. Üçüncü kümedeki ise farklı bilgi kaynakları ile birleşmiş kapsamlı modellerden oluşmuş sistemleri içermektedir. Kapsamlı karma yaklaşım algoritmasının önemli bir sınıfı da demetleme modellerinden oluşturulmuştur. Kullanıcı ve öğe özellikleri ve aynı zamanda buna uygun ek varlıklar ve bunların özellikleri daha yüksek kalitede öneri sağlayabilmek için demetleme modelleri ile birlikte çalışırlar. Bu alandaki Hoffman’ın çalışmaları bulunmaktadır. Şekil 4 de de görüldüğü gibi işbirlikçi filtreleme içerik tabanlı filtrelemeye göre daha kesin sonuçlar verebilmektedir. Karma sistemler ise veri ve bilginin az olduğu tasarımlarda başarı göstermekte iken örnek sayısı çoğaldığında işbirlikçi sistemler karma sistemler kadar başarılı olabilmektedirler.

Şekil 4’de görüldüğü üzere içerik filtreleme öğe sayısı bilgisi artsa da aynı şekilde başarım gösteriyor. İşbirlikçi filtreleme genelde az veri olduğu durumlarda düşük başarım gösterir. İki filtreleme yönteminin beraber kullanıldığı karma yaklaşım yöntemi ile ise en yüksek başarım sağlanır.



**Şekil 4 –** İçerik ve işbirlikçi filtrelemenin beraber uygulandığı çalışma