

Yapay Zeka Desteđi İle Parfüm Öneri Sistemi

Yusuf Kocaman¹, Asli Uyar¹

¹ Okan Üniversitesi, Bilgisayar Mühendisliđi Bölümü, İstanbul

yusuf.kocaman@msn.com, asli.uyar@okan.edu.tr

Özet: Önceki kullandığı parfümlere olan tepkisi bilinen bir kullanıcının, parfümlerinin koku tonları sınıflandırılıp, yeni belirlenen parfüm grubuna karşı tepkisinin tahmin edilebilmesi amaçlanmaktadır.

Anahtar Sözcükler: Parfüm öneri, kokuların sınıflandırılması, yapay öğrenme.

Abstract: This study aims at estimating people's response to a new parfüm by classifying the smell tonnes of his/her previous perfume choices.

1. Giriş

Parfüm öneri algoritmalarında popüler olarak şu 3 algoritma kullanılır;

1. Kullanılan markanın diđer parfümlerinin sunulması
2. Kullanıcı beğeni popülaritesi yüksek olan diđer parfümlerin sunulması
3. Genel kullanıcıların baktığı parfümlerin gezinme ağacını çıkarıp, aynı parfüm sırasından başlayan kullanıcıya diđer parfümlerin sunulması

Oysa bireysel olarak renk beğenisinde de-facto ayrımlar olduğu gibi, koku beğenisinde de durum aynıdır. [1]

Parfüm satın alırken genellikle diđer tanıdığımız kişilerin tavsiyeleri yada önceki kullanılan bir parfüm markası baz alınıp aynı markanın diđer modelleri arasında seçim yapılıyor.

Aslında bu iki tercih yöntemi de size en uygun parfümü bulmanızı sağlamaz. Her parfüm için içerisinde bulunan aroma çeşitleri ve yayılma süresine göre yaydığı bir koku notası vardır. [2]

Bu çalışmada amaç yapay zeka destekli olarak kullanıcının sahip olduğu parfümlerin türüne göre sınıflandırılıp, satın alma ihtimali yüksek olan yeni parfümlerin tespiti için güvenilir karar ağacı oluşturmaktır.

2. Uygulama

Problem Tanımı

Genel olarak satılan popüler parfümler, kullanıcıya özel filtreleme yapılmadan sunulmaktadır ve bu da son kullanıcı tercihlerinde kararsızlık oluşmasına sebep olmaktadır. Son kullanıcı için öneri kirliliğine engel olmak için temel filtreleme oluşturulması amaçlanmıştır. Parfümlerin kokularına göre kategorileri bulunmaktadır ve bu kategorilerine göre yapay zeka desteđi ile kullanıcıya özel karar ağacı oluşturulup en tutarlı önerileri sunabilmek için çalışılmıştır.

Müşterilerin satın alma eğilimleri için gerçekleştirilen tüm bu uygulamalarda müşterilerin tahsil durumu, gelir durumu, medeni durumu, mesleđi, yaşı, ikamet adresi gibi demografik özellikleri dikkate alınmamaktadır.

Uygulamada Veri Madenciliđi Süreci

Veri Temizleme: Müşteri analizi yapabilmek için müşterilere ait harcama tutarı, satın alma periyotları ve müşterilerin yaşları, ikametleri

gibi verilere ihtiyaç vardır. Daha doğru bir şekilde müşteri analizi yapabilmek için bu veriler gereklidir. Ancak biz verileri mağazaya ait internet sitesi üzerinden aldığımız için yukarıda sayılan verilere sahip değiliz. Bu nedenle müşteri verilerini işlemek (satın alma eğilimlerini ortaya çıkarmak) için sadece müşterilerin ürüne sahip olup-olmaması ve ürünü alma istekleri değerlendirilmiştir. [3]

Eksik olan kısımlar için o kategoriye ait ortalama değer kullanılsaydı müşteriler doğru analiz edilemeyecekti. Bu nedenlerden dolayı orijinal veri tablosu üzerinde eksik veri içeren kısımlar sıfır "0" olarak alınmıştır.

Veri dönüştürme: Veri temizlemeden sonraki adım veri dönüştürmedir. Veri dönüşümü elde edilen bilginin maksimize edilmesini sağlar.

1. Parfümlerin kullanımının analizinde, kullanan (değerlendiren) müşteriler için "1", kullanmayan (değerlendirmeyen) müşteriler için ise "0" değeri alındı.
2. Parfümler içeriklerine göre (Citrus baharat), Sweet, Fresh Spicy (taze baharat),...) 1000 tane parfüm çeşidi içerisinde toplam 70 farklı accord olduğu belirlenmiştir.
3. Müşterilerin parfüme sahip olup-olmaması veya ürünü alma istekleri dikkate alınmıştır.

Modelleme: Veri temizleme ve dönüşümünden sonraki adım modelleme adımdır. Bu adımda farklı modeller veri kümesi üzerinde deneyerek doğruluğu en yüksek olan model seçilir.

Bu uygulamada model kurma aşamasında WEKA paket programı kullanılmıştır. Modellemenin ilk aşamasında; fragrantica.com internet sitesi üzerinden parfüm ve kozmetik ürünleri alan müşterilerin sitedeki parfümlere sahip olup olmadıklarının analizi, ikinci aşamasında; fragrantica.com internet sitesi üzerinden parfüm alışverişini yapan müşterilerin farklı marka veya çeşit

parfüm alma isteklerinin analiz ederek müşterilerin satın alma eğilimlerinin ortaya çıkarmak ve onlara uygun pazarlama politikaları geliştirmek amacıyla müşteri segmentasyonu gerçekleştirilmiştir.

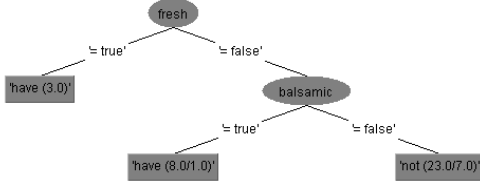
Eldeki ham veriler kullanılarak müşteriler analizi edildiğinde sistem doğru/mantıklı sonuçlar vermemektedir. Bu durumu engellemek (müşterilerin satın alma eğilimlerini doğru bir şekilde ortaya çıkarabilmek) için altı adet müşteri profili kullanılmıştır. Müşteri profilleri, kullanıcı verileri ile aynı sayıda rastgele seçilen parfümler arasında oluşturulan veri dosyalarından oluşmaktadır ve iki adet arff dosyası elde edilmektedir.

- a. Müşterilerin mevcut kullandıkları parfümleri gösteren "have.arff" dosyası,
- b. Müşterilerin satın almak istedikleri parfümleri gösteren "want.arff" dosyasıdır.

Müşteri profilleri; sitede alveriş yapan, sisteme sahip olduğu ve almak istediği parfümlerin listesini ve özelliklerini ekleyen kişilerin bilgileri kullanılarak bilgisayar programı aracılığıyla sistemden verilerin alınması ile elde edilmiştir.

J.48 algoritması; C4.5 karar ağacı algoritmasının WEKA'ya uyarlanmış versiyonudur. Her satır, ağaçtaki bir düğümü; alt satırlar, ilk satırın çocuk düğümlerini; düğümlerde parantezin içindeki ilk sayı veri kümesindeki kaç vakanın bu düğüm için doğru olarak sınıflandırıldığını; eğer varsa, parantezin içindeki ikinci sayı, düğüm tarafından yanlış olarak sınıflandırılan vakaların sayısını gösterir.

J.48 algoritmasının örnek bir müşteri profili için çalıştırılması sonucunda elde edilen çıktı da; WEKA'da "classifier output" panelinde metin formatında sunulan karar ağacını "visualize tree" seçeneğini kullanarak görsel olarak ifade etmek mümkündür.



Şekil 1. J.48 Karar Ağacı

Şekil 1.'de ilk dallanma örnek müşteri profilinin sahip olduğu parfümlerin “fresh” accorduna (taze koku kategorisi); ikinci dallanma “balsamic” accorduna (yumuşatıcı/ dinlendirici koku kategorisi) sahip olup olmadığını göstermektedir.

Yukarıda görülen J.48 karar ağacı algoritmasının sonuçlarını aşağıdaki şekilde değerlendirmek mümkündür:

İlk dallanmada parfüm içeriğinde fresh kokular olması durumunda müşterinin bu ürünleri tercih etme eğiliminde olacağı yani müşterinin fresh kokuya sahip olan bütün parfüm çeşitlerini alabileceği görülmektedir. Eğer parfüm içeriğinde fresh koku yoksa müşterinin balsamic kokusuna sahip olan parfüm çeşitlerine karşı eğilim gösterecektir.

Algoritma	J.48
Doğru olarak sınıflandırılan birimlerin yüzdesi	74,29
Yanlış olarak sınıflandırılan birimlerin yüzdesi	25,71

Tablo 1. Modellemede karar ağacının ortalama doğruluk oranı

Yukarıdaki altı son kullanıcı verisine istinaden hazırlanan profiller değerlendirildiğinde;

No	Kaynak Uye ID[5]	Doğruluk Yüzdesi (%)
1	19936	80,71
2	166650	65,02
3	869627	80
4	564421	75
5	804919	70
6	483398	75

Tablo 2. Son kullanıcı verilerine göre doğruluk oranı

Yukarıda verilen algoritmalar en yüksek başarı seviyesini vermektedir ve algoritmaların başarı oranı ortalama % 74,29'dir.

İnternet sitesi üzerinden alınan bu altı son kullanıcı verisi için oluşturulan, sistemdeki parfümlere sahip olup olamaları ve parfümleri alma istekleri değerlendirildiğinde; WEKA üzerinde kurulan model parfüm alma isteğine göre ortalama %74,29 oranında doğru tahminler yapmaktadır.

Şuanda sadece altı tane müşterinin verileri kullanılarak elde edilen veriler kullanılarak program üzerinde model çalıştırıldığında yukarıda verilen yüzdelik değerler elde edilmiştir. Site üzerinden alışveriş yapan daha fazla müşteri profilleri (verileri) incelenerek daha fazla tahmin yapabilmek mümkündür.

Sonuç ve Öneriler

Veri madenciliğinin en yaygın uygulama alanlarından biri pazarlamadır. Şirketler, veri madenciliğini kampanya yönetimi, müşteri segmentasyonu, müşteri edinme, mevcut müşterileri elde tutma, müşteri yaşam ömrü analizi, müşteri ilişkileri yönetimi, pazar sepeti analizleri ve kredi derecelendirme gibi pek çok alanda kullanmaktadır.

Bu çalışmada, veri madenciliği teknikleri kullanılarak internet üzerinden parfüm ürünleri satan bir sitenin müşterilerini analiz ederek; onların parfüm satın alma eğilimlerini (tercihlerini) belirlemek; müşterilere göre onların sevebileceği içeriklere sahip ürünler eklendiğinde onlara e-posta gibi iletişim araçlarıyla ürünlerden haberdar olmaları amaçlanmıştır.

Kaynaklar

[1] Perfume Reviews and Online Community, fragrantica.com, FRAGRANTICA

[2] Parfüm Özel, parfumozel.com,
Bitenekadar Bilgi Teknolojileri ve Ticaret
A.Ş.

[3] Çetin Elmas (2012), Yapay Zeka
Uygulamaları, Yapay Sinir Ağları – Bulanık
Mantık– Genetik Algoritma, Ankara: Seçkin
Yayınevi ISBN 9789750216961

[4] Şadi Evren ŞEKER (2013) (Türkçe). İş
Zekası ve Veri Madenciliği (Weka ile) ISBN
9786051276717. Cinius.

[5] Perfume Wardrobe and Profile,
fragrantica.com/member/ID,
FRAGRANTICA