



Duygu Analizi İle Kişiyeye Özel İçerik Önerme

Beyzanur BOSTANCI^{1*}, Ahmet ALBAYRAK²

¹Düzce Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği A.B.D, Düzce

²Düzce Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği, Düzce

Özet

Bu çalışmada, günümüzde yoğun biçimde kullanılan Facebook ve Twitter sosyal medya platformlarında paylaşılan kullanıcı yorumları duygu analizi teknikleri ile değerlendirilmiştir. Kullanıcı yorumları olarak üniversite tercih döneminde olan kişilerin paylaşımlarını içeren kısıtlı bir veri seti üzerinde çalışılmıştır. Bu veri setinin seçilmesinin nedeni günümüzde üniversite tercih dönemlerinde özellikle TV ve gazeteler aracılığıyla özel ve vakıf üniversitelerinin verdikleri yoğun reklamların sosyal medya kullanıcı profilleri analiz edilerek kişiyeye özel içerik oluşturma amacıyla kullanılmak istenmesidir. Sosyal medya üzerinden verilecek reklam içeriklerinin TV'lerde verilen reklamlardan maliyet olarak daha uygun olacağı açıktır. Özellikle kamu üniversitelerinin sosyal medya üzerinden reklam verebilmesi daha nitelikli ve bilinçli öğrencileri çekebileceği düşünülmektedir. Burada iki aşamalı bir yaklaşım önerilmektedir. İlk aşamada sosyal medya ortamından toplanan veriler analiz edilmiştir. İkinci aşamada önceden belirlenen sınıfların TF-IDF tekniği puanına göre iyimser, karamsar, mizahi, üretken ve dışa dönük olmak üzere kategorilere ayrılmaktadır. Ayrılan bu kategorilere göre de kullanıcıların profiline uygun reklamlar içerikleri sunulmaktadır.

Anahtar Kelimeler: Duygu Analizi, Eğitimde Reklam, Sosyal Medya Analizi, TF-IDF

Makale Bilgisi

Başvuru:
06/08/2020
Kabul:
11/10/2020

Suggesting Individual Content with Sentiment Analysis

Abstract

In this study, user comment shared on Facebook and Twitter social media platform, which are used extensively today, were evaluated using emotion analysis techniques. As user comments, a limited set of data was included, which includes the sharing of people who are in university preference period. The reason for choosing this data set is that today, during the university preference periods, the intensive advertisements given by private and foundation universities, especially through TV and newspapers, are asked to be used to create personalized content by analyzing the social media user profiles. It is clear that advertising content on social media will be more affordable as a cost than ads on TVs. It is thought that public universities, in particular, can attract more qualified and informed students to advertise through social media. A two-step approach is proposed here. In the first stage, the data collected from the social media environment were analyzed. In the second stage, according to the TF-IDF technique scores of the predetermined classes, they are divided into categories as optimistic, moody, humorous, productive and extroverted.

* İletişim e-posta: beyzanurbstnci@gmail.com

According to these allocated categories, ads content that is appropriate to the users ' profile is offered.

Keywords: Sentiment Analysis, Advertising in Education, Social Media Analysis, TF-IDF

1 Giriş

Üniversite eğitiminin gelişmesi ve modernleşmesi, gelişmekte olan ülkeler için önemli bir rol üstlenmektedir. Günümüzde üniversiteler, toplumların kültürel, sosyal, ekonomik, bilimsel ve teknolojik gelişmelerde kendilerine yer aramaktadır. Daha kaliteli eğitim çıktıları elde etmek daha kaliteli öğrenci topluluğu ile olabilmektedir. Bunun için birçok vakıf üniversitesi; öğrenci adayları için TV ve sosyal medya platformlarında reklamlar vermektedir. Bu reklamlar genel olarak öğrencilerin yapacakları tercihleri yönlendirmek için vaatler içermektedir.

2011 yılında yapılan bir çalışmada; gazetelerden ve dergilerden alınan 385 farklı reklam ilanına ulaşılmıştır. Çalışmaya göre üniversite kavramının farklı topluluklar da nasıl algılandığı açıklanmıştır. Araştırma sonucunda öğrencilerin kariyer vaadi üzerinde yoğunlaştığı görülmektedir [1].

2016 yılında yapılan başka bir çalışma da eğitim üzerinden yapılan reklamların metalaşmasından bahsedilmektedir. Çalışma için birçok gazetenin eğitim haberlerine ilişkin sayfalarındaki reklamların sayısal bilgileri incelenmiştir. Reklamlar özel ve devlet okulları olarak ikiye ayrılmıştır. Araştırmaya göre özel okulların özellikle de üniversitelerin olumlu veya olumsuz haber sayısının çok fazla olduğu gözlenmiştir. Bunun yanında devlet okullarının/üniversitelerinin özel okullar ile eşitlenebilmesi için eğitim haberciliğinin gelişmesi önerilmiştir [2].

Teknolojinin gelişmesiyle birlikte reklamlar gazete ve dergilerin yanı sıra internet üzerinden de yapılmaya başlanmıştır. İnsanlar bu reklamlara olumlu veya olumsuz tepkilerini, sosyal medya üzerinden paylaşmaktadır. Bunun sonucunda sosyal medya önemli bir bilgi kaynağı haline gelmiştir. İnsanların sosyal medya üzerinden paylaştığı bir tweet in duygu ve düşüncelerini belirlemek için duygu analizi yapılmaktadır. Duygu analizi (Görüş madenciliği), doğal dil işleme, istatistik, bilgisayar bilimleri gibi alanlardan yöntem ve tekniklerin kullanılması ile görüş sahibinin metin içerisinde belirttiği duygu, görüş, tutum gibi öznel bilgilerin belirlenmesini amaçlayan güncel bir araştırma alanıdır [3].

Duygu analizi yapılırken metnin öznel içeriğine göre sınıflandırma işlemleri yapılır. Sınıflandırma işlemlerini yapabilmek için sınıflandırma algoritmaları kullanılır. Duygu analizi amacıyla literatürde en çok tercih edilen sınıflandırma algoritmaları Naive Bayes, en yakın komşu, Karar ağaçları, Destek Vektör Makineleri (SVM) ve Random Forest algoritmalarıdır [4].

Günümüzde sosyal medyadan toplanan veriler ile ilgili birçok duygu analizi çalışması bulunmaktadır. Ortigosa ve arkadaşlarının 2014 yılında yapmış olduğu çalışmada Facebook kullanıcıları tarafından yazılan mesajların duygu polariteleri (pozitif, nötr veya negatif) ve duygu analizi sonucunda e-öğrenme üzerinde etkisi incelenmiştir. Mesajları alması ve alınan mesajlar üzerinden duygu analizine göre sınıflandırma yapmak için Facebook'un uygulaması olan SentBuk kullanılmıştır. SentBuk, sözcük tabanlı ve makine öğrenmesi sınıflandırma tekniklerinin birleşimidir. E-öğrenme üzerindeki etkisi için bir kurs örneği verilmiştir. Bu kursa giden öğrencilerin duygu ve düşüncelerini sosyal medya üzerinden alınan mesajlar üzerinden duygu analizi yapılarak özellikle yüz yüze olmayan kurslar için geri bildirim olarak kullanılmıştır[5].

Akın ve Şimşek tarafından yapılan çalışmada, televizyon kanallarındaki ürünlerin müşteri üzerindeki etkisini anlayabilmek ve gelecek ile ilgili yorum yapabilmek için sosyal medya üzerinden toplanan veriler ile duygu analizi yapılmıştır. Haziran - Kasım 2017 tarihleri arasındaki Twitter kullanıcılarının televizyon yayınları hakkındaki duygu ve düşüncelerinin pozitif, negatif veya nötr olma durumu incelenmiştir. Araştırmanın sonucunda televizyon kanallarındaki programların reyting bazlı değil, izleyici görüşüne bağlı olarak değiştiği gözlemlenmiştir. Bunun yanında televizyon kanallarına gelecek olan reklamların fiyat çizelgesi önceden belirlenebilir[6].

Aydın 2018 yılında yapmış olduğu çalışmada itfaiye istasyonu ihtiyaçlarını belirlemek için sınıflandırma işlemi yapmıştır. Bunun için 2015-2017 yılları arasında yangın kayıtlarını veri seti olarak kullanmıştır. Sınıflandırma işlemleri için makine öğrenmesi algoritmaları kullanılmaktadır.

Analiz sonucunda bu çalışma için en verimli algoritma Random Forest olduğu gözlemlenmiştir[7]

Troussas ve arkadaşlarının yapmış olduğu çalışmada, Facebook kullanıcıları tarafından oluşturulan paylaşımların Naive Bayes sınıflandırma tekniği kullanılarak duygu analizinin dil öğrenmesinde nasıl yardımcı olacağı sunulmuştur. Yapılan çalışma sonucunda, Naive Bayes sınıflandırıcısı Facebook kullanıcılarının duygusal durumu analiz etme doğruluğunun yüksek olduğu belirtilmiştir[8].

Maite Giménez ve arkadaşlarının yapmış olduğu çalışmada, doğal dil işleme performansını arttırabilmek için evrimsel sinir ağlarında (CNN) anlamsal tabanlı dolgu önerilmiştir ve bir duygu analizi örneği yapılmıştır. Yapılan çalışmada birçok metinden örnekler alınarak anlamsal tabanlı dolgu uygulanmış ve duygu analizi sonucunda performansın daha iyi olduğu kanıtlanmıştır[9].

Ayan ve arkadaşlarının 2019 yılında yaptığı çalışmada 162.000 Tweet incelenmiş ve bunların islamofobik olup olmadığının belirlenmesi için duygu analizi yapılmıştır. Duygu analizi için Naive Bayes sınıflandırma algoritması ve Ridge regresyonu kullanılmıştır. Bu çalışmada Ridge ile birlikte Naive Bayes daha verimli çalışırken, bunun yanında sınıflandırma işlemi daha kısa sürmüştür[10].

Joeran Beel1 ve arkadaşlarının yapmış olduğu çalışmada, araştırma - kağıt tavsiye sistemleri üzerinden bir literatür araştırması yapılmıştır. Bu çalışma için 16 yıllık 200'den fazla makale bulunmuştur. Bulunan makaleleri tavsiye yaklaşımının yarısından fazla olanlar, işbirlikçilik, grafik tabanlı olmak üzere çeşitli filtrelemeler yapılmıştır. Her filtreye model önermek için Terim Frekansı-Ters Belge Frekansı (TF-IDF) ve n gram teknikleri kullanılmıştır[11].

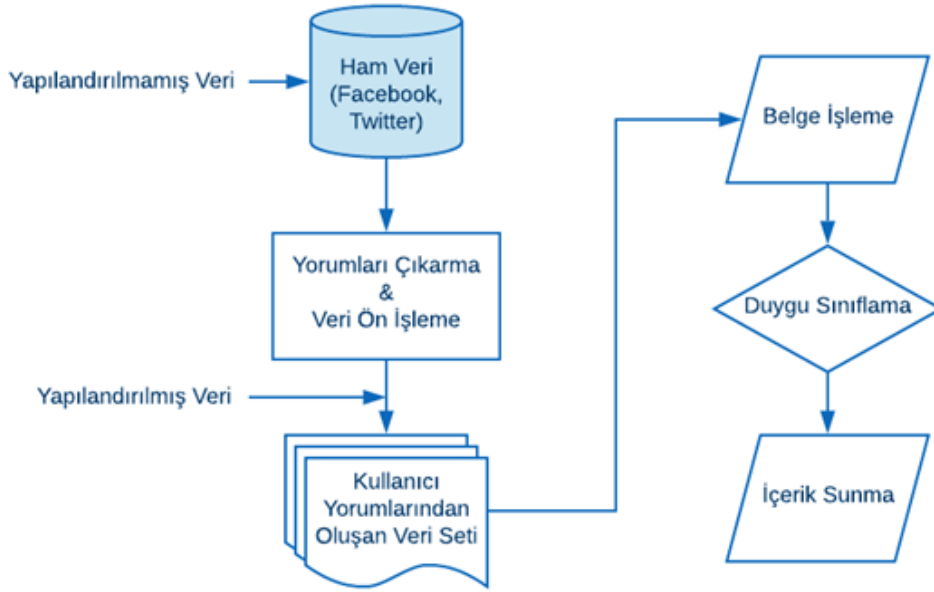
Mirjana Pejic-Bach ve arkadaşlarının 2020 yılında yaptığı çalışmada endüstri 4.0 kapsamında iş özelliklerinin hızlı değişiminden dolayı iş reklamlarının analizini hızlandırmak amacıyla metin madenciliği kullanılmıştır. Veri seti oluştururken endüstri 4.0 için verilen iş ilanları dikkate alınmıştır. Metin madenciliği analizi sonucunda iki grup iş profili oluşmuştur. İlk iş profili grubu sadece Endüstri 4.0 ile ilgili bilgilere odaklanmış: siberfizik sistemler ve robotlu üretim için nesnelerin interneti ve akıllı üretim tasarımı ve üretim kontrolüdür. İkinci iş profili grubu, Endüstri 4.0'a uyarlanan daha genel bilgi alanlarına odaklanmaktadır: tedarik değişim

yönetimi, müşteri memnuniyeti ve kurumsal yazılımlardır.[12]

Bu çalışmada, özellikle özel/vakıf üniversite ve okullar tarafından verilen reklamların bir kamu üniversitesi için nasıl olabileceği araştırılmıştır. Özel/Vakıf üniversiteleri, tercih edecek öğrenci profilini dikkate almaksızın herkes için geçerli aynı reklamı yapmaktadır. Buna ek olarak özel/vakıf üniversiteleri reklam için yoğun biçimde TV ve gazeteleri tercih etmekte ve büyük sayılabilecek paralar harcamaktadır. Bir kamu üniversitesinin reklam amacıyla yüksek bütçeler ayırması mümkün değildir. Günümüzde üniversite tercih edecek öğrenci profilinin sosyal medyayı sık kullandığı da düşünülünce, kamu üniversitelerinin sosyal medya platformlarında reklam vermesi yerinde olacaktır. Bu amaçla, bu çalışmada iki aşamalı bir yaklaşım önerilmiş ve kısıtlı veri seti ile denenmiştir. İlk aşamada sosyal medya ortamında (Facebook ve Twitter) kullanıcı yorumları duygu analizi ile analiz edilmiştir. Bu işlemin amacı kullanıcı profiline uygun reklamlar üretmek için kullanıcıların duygusal yönlerini tespit etmektir. Bunun için beş farklı kategori belirlenmiştir. Bunlar iyimser, karamsar, mizahi, üretken, dışa dönük şeklindeki kategorilerdir. Sosyal medya kullanıcısının hangi kategoride olduğu sınıflandırıldıktan sonra, kullanıcı profiline uygun hazırlanmış reklamlar kullanıcıya sunulmaktadır. Sınıflandırma ve kişiye uygun reklamın sunulması da çalışmanın ikinci aşamasını oluşturmaktadır. Bu çalışma duygu analizi ve sınıflandırma sürecinin paylaşıldığı materyal ve metot bölümü, sonuçların analiz edildiği bulgular ve ardından sonuç ve öneriler bölümlerinden oluşmaktadır.

2 Materyal ve Yöntem

Bu çalışmada doğal dil işleme teknikleri kullanılarak duygu analizi yapılmış ve kullanıcılara içerik sunulmuştur. Bir prototip olarak geliştirilen sistemin blok şeması Şekil 1'de verilmektedir. Şekil 1'de de verildiği gibi ilk olarak twitter ve facebook platformlarından kullanıcı yorumları alınmıştır. Alınan bu veriler ham veri formatında olup, yapılandırılmamış veri tipindedir. Yapılandırılmamış veri ön işleme tabii tutularak içinde anlam ifade etmeyen, duygu analizi için kullanılmayan ifadeler/sözcükler/karakterler çıkarılmış ve yorumlar elde edilmiştir. Her yoruma benzersiz bir nitelik alanı eklenerek, csv uzantılı dosyaya kaydedilmiştir. Bu aşamada veri seti yapılandırılmış veri formuna dönüşmüştür.



Şekil 1. Önerilen modelin blok şeması

Şekil 1’de kullanıcı yorumlarından oluşan veri seti üzerinde belge işleme aşamasında belirlenmiş anahtar kelimeler taranmıştır. Kelimelerin dokümanda/yorumda geçme sıklığına bağlı olarak duygu analizi yapılmıştır. Kullanıcının yorumlarındaki duygu durumu tespit edildikten sonra, önceden hazırlanmış reklam içerikleri kullanıcıya sunulmaktadır.

2.1 Veri Setinin Elde Edilmesi

Doğal dil işleme tabanlı duyarlılık analizi Twitter ürün/hizmetleri etrafında odaklanmış olsa da, Facebook durum mesajlarında kişilerin duygularını doğası gereği daha doğru bir şekilde sınıflandırmanın mümkün olduğu açıktır. Zira Facebook durum mesajları sınıflandırılması tweetlerden daha kolaydır, çünkü daha fazla sözcük içermektedir. Bu sayede daha iyi biçimde duyguların tasvir edilmesi sağlanır.

Veri seti, oluşturmak için Twitter ve Facebook’ta “geliştirici hesabı- developer hesabı” açılmıştır. Geliştirici hesabı ile API üzerinden veriler çekilmiştir. Burada hesabı gizli olmayan ve en az 10 yorum yazmış kullanıcılar seçilmiştir. Geliştirici hesabı ile çekilen yorumlar JSON dosya formatındadır. Veri setinde Twitter’dan 82 kullanıcı ve Facebook’tan ise 65 kullanıcı yorumu bulunmaktadır. Twitter ve Facebook’tan çekilen kullanıcı yorumlarının aynı kullanıcılara ait olmamasına dikkat edilmiştir.

2.2 Veri Ön İşleme

Veri temizleme işlemi istenmeyen kelime ve karakterlerin ayrıştırılmasında yardımcı olur. Veri temizleme işlemi yapılmazsa gürültülü ve tutarsız veriler ile doğru bir sonuca ulaşılması olası değildir. Bu adımın amacı, noktalama, özel karakter, sayılar ve istenilen metin dışındaki kelimeleri temizleyerek gürültüsüz veri seti elde etmektir. Geliştirilen uygulama Jupiter ortamında kodlanmıştır. Veri temizleme işlemi için yazılan fonksiyon aşağıda verilmektedir.

#sözcük temizleme

```
def gereksiz_sil(metin, silinecek_desen):
    r = re.findall(silinecek_desen, metin)
    for i in r:
        metin = re.sub(i, "", metin)
```

return metin

Fonksiyon içinden gereksiz karakter ve sözcüklerin temizleneceği, yorum parametre olarak almaktadır. Fonksiyon bir diğer parametre olarak ise, ne tür sözcük ve karakterin kaldırılmasını tutan deseni almaktadır. Veri temizleme adımından sonra her bir yoruma benzersiz nitelik verilerek (Id) csv uzantılı bir dosyaya alınmıştır. Bu aşamada veri seti yapılandırılmış verilerden oluşmaktadır.

2.3 Belge İşleme

Çalışmanın bu aşamasında, ilk olarak sınıflandırma için anahtar kelimeler belirlenmiştir. Belirlenen bu anahtar kelimeler iyimser, karamsar, mizahi, üretken ve dışadönük olarak beş farklı kategoriye göre gruplandırılmıştır. Tablo 1’de hazırlanmış anahtar kelimelerden bir kesit verilmektedir.

Tablo 1. Duygu analizi için belirlenen kategoriler

Kategori	İyimser	Karamsar	Mizahi	Üretken	Dışadönük
1	Hayırlısı	Kapkara	Zeka	Tarz	Hissetme
2	Olumlu	Olumsuz	Komik	Üretmek	Duygu
3	Pozitif	Negatif	Gülmek	Plan	Değişim
4	Doğru	Yanlış	Şaka	Proje	İlişki
5	Saygı	Nefret	Espri	Çalışmak	Arkadaş
6	Sevgi	Çirkin	Eğlenmek	İş	Pratik
7	Fırsat	Stres	Düşünmek	Deneyim	Görev
8	İyi	Kötü	Eleştiri	Başarı	Hoşgörü
9	Güzel	Umutsuz	Dalga geçmek	Mantık	Sosyal
10	Şükür	Dert	Karikatür	Neden	Çevre

Tablo 1’de hazırlanmış anahtar kelimelerin bir kısmı verilmektedir. Her bir kategori için belirlenen 25 anahtar kelime çalışma kapsamında kullanılmaktadır. Belirlenen anahtar kelimeler ardından, veri seti içinde anahtar kelimeler aranmış ve kelime karşılıkları bulunduğu durumda ilgili kategorinin puanı bir artırılmaktadır. Bu amaçla kullanılan fonksiyon aşağıda verilmektedir.

Bu fonksiyonda ilk olarak ifade eğer büyük harfli ise küçük harfe çevrilmiş ve ardından sözcüklere ayrılmıştır. Aynı kökte başka sözcüklerin olup olmadığını tespit etmek için bir sözcüğün çekimsel biçimleri değerlendirilmiş ve ortak bir temel biçime indirgeme işlemi yapılmıştır. Bu işlem adımının amacı kullanıcıların sözcükleri farklı kullanması durumunda sözcüğün kökü üzerinden işlem yapmanın daha doğru olacağıdır.

#sözcük hesapla

```
def metin(sozcuk):
    lower_ifade = ifade.lower()
    lemmatizer = WordNetLemmatizer()
    lem_words = filtered_data.map(metin)
    tfidf = TfIdfVectorizer(sozcuk, min_df=2,
max_features=None, stop_words='english')
    tfidf = tfidf_vectorizer.fit_transform(sozcuk)
    return lemmatizer.lemmatize(tfidf)
```

Köklerin de bulunması ile artık tüm veri seti atomik sözcüklerden oluşmaktadır. Veri kümesi içinde her bir sözcüğün tekrarının ne kadar olduğunun bulunması dördüncü işlem adıdır. Burada yöntem olarak TF-IDF yöntemi seçilmiştir. TF-IDF, bir kelimenin belge içerisinde kullanıma sıklığını gösteren istatistiki yöntem ile bulunabilen ağırlık faktörüdür. TF-IDF değeri, bir sözcüğün belgede kaç kez görüldüğüyle orantılı olarak artar ve bazı sözcüklerin genel olarak daha sık görüldüğü gerçeğinin belirlenmesine yardımcı olan bir tekniktir. TF-IDF günümüzde en popüler

terim ağırlık faktörü yaklaşımlarındandır. Bu çalışmada belirlenen anahtar kelimelerin belge içerisindeki ağırlık fonksiyonu değerini bulabilmek amacıyla TF-IDF yöntemi seçilmiştir. 2015 yılında yapılan bir araştırmada, dijital kütüphanelerdeki metin tabanlı tavsiye sistemlerinin %83’ünün TF-IDF kullandığını göstermiştir [13].

2.4 Duygu Sınıflandırma

Bu çalışmada duygu sınıflama adımında, kullanıcı yorumları TF-IDF ile belirlenen anahtar kelimeler ile sayılmış en yüksek puanı alan kategori kullanıcının duygusunu vermektedir. Doğal dil işleme kütüphanesindeki varsayılan teknikler duygu sınıflamasını olumlu, olumsuz ve nötr gibi en fazla üç kategoriye ayırmaktadır. Bu çalışmada ise kategori sayısı artırılmıştır. Sınıflama kategorilerinin artırılması sonucunda duygu analizi teknikleri/fonksiyonları farklı amaçla kullanılmaktadır. Bu neden çalışmanın bu adımı literatürdeki duygu analizi gibi kısa süreli sürmemiştir. Bu çalışmada referans anahtar kelimelerinin veri setinde ne sıklıkta olduğunun tespiti makine öğrenmesi kapsamındaki danışmanlı öğrenme kavramı kapsamındadır. Duygu analizinde sınıflandırıcıları değerlendirirken, değerlendirme için farklı yaklaşımlar bulunmaktadır. Bu çalışmada her bir kategori/sınıf için karmaşıklık matrisi oluşturulmuş. Oluşturulan bu matris algoritmanın performansını test etmek için kullanılmıştır ve başarıyı değerlendirilmiştir. İyimser kullanıcı kategorisi Tablo 2’de verilmektedir.

Tablo 2. Karmaşıklık matrisi

		Tahmin Edilen		Hassasiyet	Doğruluk
		Pozitif	Negatif		
İyimser Kullanıcı	Pozitif	826	644	55,89	56,2
	Negatif	163	1307	87,47	
Karamsar Kullanıcı	Pozitif	1012	458	68,1	68,8
	Negatif	298	972	79,1	
Mizahi Kullanıcı	Pozitif	785	685	53	53,4
	Negatif	621	849	58,2	
Üretken Kullanıcı	Pozitif	812	658	55,15	55,23
	Negatif	214	1256	84,8	
Dışadönük Kullanıcı	Pozitif	1144	326	77,34	77,82
	Negatif	462	1008	68,79	

Tablo 2’de kullanıcı yorumları incelendiğinde en yüksek doğruluğun dışadönük kategorisi için elde edildiği görülmektedir. Bunun sebeplerinden birinin bu kategori için belirlenen anahtar kelimelerin genel kullanımda dolaşan kelimeler ağırlıklı olması olduğu düşünülmektedir.

2.5 İçerik Sunma

Bu çalışmada Facebook ve Twitter yorumları alınan kullanıcıların doğal dil işleme teknikleri ile duyguları analiz edildikten sonra, kendilerine reklam içeriği sunulmaktadır. Bu çalışma Düzce üniversitesi bünyesinde gerçekleştirildiğinden reklam içerikleri de Düzce üniversitesi afişlerinden oluşmaktadır. Hazırlanan reklam içerikleri şekil 2’de verilmektedir.

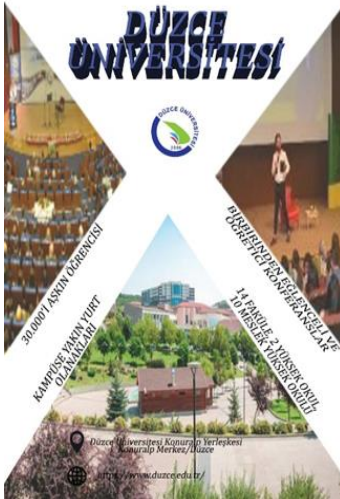
İyimser Reklam Afışı



Karamsar Reklam Afışı



Mizahi Reklam Afışı



Üretken Reklam Afışı



Dışadönük Reklam Afışı



Şekil 2 Afişler

Şekil 2'de afişler; iyimser, karamsar, mizahi, üretken ve dışadönük olmak üzere beş sınıfa göre kategorilenir. Afişlerin üzerlerinde sınıfları destekleyen tasarımlar verilmiştir. Bu tasarımlar sosyal medya üzerinden öğrencilerin dikkatini

çekmek ve Düzce Üniversitesinin imkanlarını öğrencilere anlatmak amaçlı oluşturulmuştur. Afişlerin arka planları, oluşturulan kategoriye uygun olmasına dikkat edilerek üniversitenin fotoğrafları kullanılmıştır.

3 Sonuçlar

Günümüzde reklam sektörü televizyon ve gazetelerden sosyal medya platformlarına doğru yönelmektedir. Özellikle Z kuşağı diye tabii edilen günümüz üniversite öğrencileri ve tercih durumunda olan öğrenciler sosyal medya platformlarını bağımlılık derecesinde yoğun biçimde kullanmaktadır. Üniversite tercih dönemlerinde büyük maddi kaynaklar ayrılarak yapılan reklamların kamu üniversiteleri tarafından yapılması mümkün değildir. Böylesi dezavantajlı bir durumda daha düşük maliyetle nitelikli ve bilinçli öğrenci çekmek son derece önemli hale gelmektedir. Bu amaçla bu çalışmada bir yaklaşım önerilmiştir. Bunun için 65 Facebook kullanıcısı ve 82 Twitter kullanıcısı profilinden alınan veriler ilk olarak veri ön işleme sürecine tabi tutulmuştur. Ardından gürültüsüz olan verilerin, belirlenen kategorilere göre sınıflandırılması yapılmıştır. Sınıflandırma yapılırken her kategori için 25 tane anahtar kelime belirlenmiştir. Seçilen bu kelimelerin tekrarlanma sıklığını ölçmek için TF – IDF tekniği kullanılmıştır. Bu teknik sonucunda %77.82 ile dışadönük kategorisinin diğer kategorilere göre doğruluk oranı daha fazla olduğu gözlemlenmiş ve bu duygu analizinin sonucunda, reklam amaçlı Düzce Üniversitesi için beş farklı afiş oluşturulmuştur.

Bir sonraki çalışmada, sosyal medya platformundan alınan kullanıcı veri sayısını arttırılarak farklı makine öğrenmesi yöntemleri denenecektir. Bunun yanında reklamları sosyal medya platformları üzerinden yayınlamak kullanıcıların geri dönüşleri incelenecektir.

Teşekkür

Bu çalışmada Bulut Bilişim, Veri Bilimi ve Uygulamaları dersleri kapsamında üretilmiştir. Çalışmaya katkıda bulunan Necmiye ARABACI 'ya ve Düzce Üniversitesine teşekkür ederim.

Kaynaklar

- [1] Yıldız S. "Üniversiler Neler Vaat Ediyor? Türkiye'deki Üniversitelerin Basın İlanları Üzerine Bir İçerik Analizi". Anadolu Üniversitesi Sosyal Bilimler Dergisi, 14(2), 155-170, 2014.
- [2] Dağtaş E, Ünal MA. "Eğitim Haberlerinin Metalaşmasını Reklam ve Reklam Haberler Üzerinden Okumak". Maltepe Üniversitesi İletişim Fakültesi Dergisi, 6(1), 38-61, 2019.
- [3] Onan A. "SENTIMENT ANALYSIS ON TWITTER MESSAGES BASED ON MACHINE LEARNING METHODS". Yönetim Bilişim Sistemleri Dergisi, 3(2), 1-14, 2017.

- [4] Zhang Y, Chen M, Liu L. "A Review on Text Mining". 2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS), 23-25 September 2015.
- [5] Ortigosa A, Martin JM, Carro RM. "Sentiment analysis in Facebook and its application to e-learning". Computers in Human Behavior, 31, 527-541, 2014.
- [6] Akın B, Gürsoy Şimşek UT. "SOSYAL MEDYA ANALİTİĞİ İLE DEĞER YARATMA: DUYGU ANALİZİ İLE GELECEĞE YÖNELİM". Mehmet Akif Ersoy Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 5(3), 797-811, 2018.
- [7] Aydın C. "Makine Öğrenmesi Algoritmaları Kullanılarak İtfaiye İstasyonu İhtiyacının Sınıflandırılması". Avrupa Bilim ve Teknoloji Dergisi, 14, 169-175, Aralık 2018.
- [8] Troussas C, Virvou M, Espinosa KJ, Llaguno K, Caro J. "Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning". Information, Intelligence, Systems and Applications (IISA), 2013 Fourth International Conference on, July 2013.
- [9] Giméne M, Palanca J, Botti V. "Semantic-based padding in convolutional neural networks for improving the performance in natural language processing. A case of study in sentiment analysis". Neurocomputing, 378, 315-323, 2020.
- [10] Ayan B, Kuyumcu B, Ciylan B. "Detection of Islamophobic Tweets on Twitter Using Sentiment Analysis". Gazi Üniversitesi Fen Bilimleri Dergisi, 7(2), 495-502, 2019.
- [11] Beel J, Gipp B, Langer S, Breitingner C. "Research-paper recommender systems: a literature survey". International Journal on Digital Libraries, 17, 305-338, 2016.
- [12] Pejic-Bacha M, Bertonce T, Meškob M, Krstić Z. "Text mining of industry 4.0 job advertisements". International Journal of Information Management, 50, 416-431, 2020.
- [13] Salo F, Injadat M, Nassif AB, Shami A, Essex A. "Data Mining Techniques in Intrusion Detection Systems: A Systematic Literature Review". IEEE Access, September 2018.