

Sosyal Medyada Rekabet Analizi İçin Karşılaştırma Görevine Yönelik Fikir Madenciliği Modeli

Çiğdem Aytekin, Marmara Üniversitesi, Endüstri Mühendisliği Bölümü,
İstanbul, Türkiye ciğdem.aytekin@marmara.edu.tr

İslam Mayda, Marmara Üniversitesi, Bilgisayar Mühendisliği Bölümü,
İstanbul, Türkiye islam.mayda@stu.khas.edu.tr +905388259530

Kabul tarihi 14 Aralık 2013

ÖZET

Günümüz internet kullanıcıları bir yandan yeni nesil teknolojilerin getirdiği kolaylıklar sayesinde ağ üzerinde sosyalleşirken, diğer yandan da birçok iş alanında dönüşümlere neden olmaktadır. Artık işletmelerin sosyal medya olarak adlandırdığımız bu ortamlara eğilmeleri bir zorunluluk haline almıştır. Zira bu ortamlar kullanıcıyı pasif konumdan içerik üreten aktif kullanıcı haline getirmiş ve bünyesinde ürün ve hizmetlere ilişkin barındırdığı milyonlarca yorum ile işletmelere rekabet analizi, itibar yönetimi, kriz yönetimi, halkla ilişkiler gibi alanlarda yeni rekabet yöntemlerine olanak sağlamıştır. Bu doğrultuda çalışmada sosyal medyada rekabet analizi için karşılaştırma görevine yönelik bir fikir madenciliği modeli geliştirilmiştir. Bu amaçla karşılaştırma siteleri, YouTube ve teknoloji forumlarından iz sürme tekniği ile karşılaştırma ifadesi içeren 100 yorum manuel olarak derlenmiş ve bu yolla bir test veri tabanı oluşturulmuştur. Elde edilen bulgular araştırmanın sınırlılığı kapsamında ele alınmış ve modelin başarısına ilişkin ölçümler duyarlılık ölçütü ile değerlendirilmiştir. Başarı oranları modelin geleceği ile ilgili ümit vermekte olup sonuç kısmında önerilere de ayrıca yer verilmiştir.

Anahtar kelimeler: Sosyal Medya; Rekabet Analizi; Fikir Madenciliği; Karşılaştırma Görevi.

Competitive Analysis on Social Media Comparison Task Oriented: Opinion Mining Model

ABSTRACT

Since social media environments can turn users from a passive position into active users who self-generate contents, in their millions of comments, relating to products and services thus providing enterprises the opportunities for new competition methods in areas such as competitive analysis, reputation management, crisis management, and public relations; an opinion mining model is used in the study for analysis of social media aimed at comparison purposes. One hundred comments containing comparison expressions were manually collected with the chasing technique from the comparison sites, YouTube and technology forums, and by this way a test database was created. The obtained findings were examined within the constraints of the study and computations relating to the performance of the model and were evaluated with the sensitivity criterion. Because the success rates of the model are promising for the future, suggestions are provided in the concluding part of the study.

Keywords: Social Media; Competition Analysis; Opinion Mining; Comparison Task.

GİRİŞ

Günümüz bilgi ve iletişim teknolojileri birçok alanda günlük hayatı ve iş akışlarını dönüştürmekte, birçok yeni eğilimin ortaya çıkmasına neden olmaktadır. Bu eğilimler işletmelerin işini belli yönlerde kolaylaştırırken, belli yönlerde de onları zorlamakta ve küreselleşen dünyaya uyum sağlamak zorunda bırakmaktadır. Artık rakipler ile rekabet edebilmenin yolları da dönüşüm sürecine girmiştir.

Değişim hızının arttığı ve niteliğinin değiştiği yeni rekabet koşullarında, müşteriler de bu değişimden etkilenmektedirler. Müşteri ihtiyaçları ve öncelikleri tüm sektörlerde çok hızlı ve çok dinamik bir şekilde sürekli değişmektedir. Değişim hızının yüksek olduğu günümüzde işletmelerin ürün merkezli stratejilerle ayakta kalabilmeleri artık mümkün olamamaktadır. Çünkü yeni rekabet koşulları, artık ürün performansına dayalı stratejilerin geçerliliğini büyük ölçüde yitirmiş olduğu bir dönemdir ve müşteri merkezli düşünmeyi gerektirmektedir (Kırım, 2003: 74, 78). Bu doğrultuda küresel rekabetin getirdiği müşteri merkezli yaklaşım işletmelerin yönetim anlayışlarını değiştirmiş ve onları müşterileriyle uzun ömürlü ilişkiler sağlamaya ve geliştirmeye zorlamıştır. Artık işletmeler, değişen müşteri gereksinimlerini karşılamak, onların önceliklerini belirlemek ve sorunlara onların bakış açısı ile bakmak durumundadırlar. Ancak bunları sağlayabilmek ve verimliliklerini arttırabilmek için onlar hakkında daha fazla bilgiye sahip olunması gerekmektedir. Bu da dış dünyadaki algılanışın takibi ile mümkün olabilir.

Sosyal medya bu noktada işletmelere geniş bir veri akışı sağlar ve bir dinleme aracı olarak karşımıza çıkar. Zira görece özgür olarak kabul edilen bu ortamlarda müşteriler ürün ve hizmetler hakkında görüşlerini paylaşabilmekte, tavsiyelerde veya şikâyetlerde bulunabilmekte ve deneyimlerini aktarabilmektedirler. Kullanıcı görüşlerinin takip edilmesi yoluyla bu ortamlardan elde edilen bilgiler, rekabet ortamında işletmelerin ilgili kararlarına ışık tutacaktır. Böylece verimliliklerini arttırabilir ve rakipleri karşısında avantajlı duruma geçebilirler, ilgili stratejileri belirlemede ve karar süreçlerinde referans olarak kullanabilirler.

Günümüzde sosyal medyada rekabet analizi hizmeti “Dijital Ajans”, “Dijital Medya Ajansı” veya “Sosyal Medya Ajansı” olarak tanımlanan kuruluşlarca yapılmaktadır. Zira artık markaların karşısında sosyal ağlarda etkileşim halinde olan yeni bir hedef kitle bulunmaktadır. Kullanıcılar markanın kendisine, ürünlerine, hizmetlerine olan yaklaşımlarını bu ortamlarda karşısındaki ile konuşur gibi informal bir dille ortaya koymaktadırlar. Ajanslar ise bir yandan markalara sosyal ağlarda doğru bir şekilde konumlanmanın yollarını sunarken, diğer yandan da rakipleri ile nasıl mücadele edebilecekleri konusunda onlara özel uygulamalar geliştirmektedirler. Bazı ajanslar tarafından “rakip analizi” adıyla da anılan bu uygulamalar, daha çok fırsatların ve tehditlerin belirlenmesine yönelik olarak geliştirilen swot analizlerine dayanmaktadır. Yani, markanın potansiyel rakipleri karşısındaki güçlü ve zayıf yönleri ortaya konmakta, analizler bu doğrultuda yapılmaktadır.

Bu çalışmada sosyal medyada yer alan verilerin rekabet analizi açısından çok önemli olduğu öngörüsünden hareketle, bu verileri analiz eden bir model geliştirilmiştir. Modelin veri mimarisi, bu ortamlarda yer alan “marka üstünlüğü bildiren yorum” temeli üzerine kuruludur. Ancak belirtmek gerekir ki, bu çalışmada geliştirilen model Türkiye’deki sosyal medya ajanslarının “rekabet analizi” kapsamında verdiği hizmetlerden farklı bir yaklaşıma sahiptir. Modelde

analitik ölçümler yerine yorumlar üzerine odaklanılmış ve bu çerçevede fikir madenciliği yaklaşımı söz konusu edilmiştir. Bu çerçevede önce fikir madenciliği, görevleri ve özellikle karşılaştırma görevi ele alınacak ve ardından modelin detayları açıklanacaktır.

FİKİR MADENCİLİĞİ VE KARŞILAŞTIRMA GÖREVİNE YÖNELİK ÇALIŞMALAR

Fikir Madenciliği, sosyal medyada yer alan verinin katlanarak çoğalması ile birlikte, bu verilerin analiz ihtiyacı çerçevesinde Metin Madenciliği'nin bir uygulama alanı olarak ortaya çıkmış bir kavramdır. Falcon'a göre fikir madenciliği, verilen bir konu üzerine görüş sahibinin fikrinin sınıflandırılması veya tanımlanması için istatistiksel model ve yazılımların kullanılmasıdır (Falcon, 2010).

Fikir madenciliği ile yapılabilecek işler literatürde "fikir madenciliği görevleri" olarak tanımlanmaktadır. Bu çerçevede farklı yazarlar tarafından farklı görevlerden söz edilmektedir. Esuli ve Sebastiani (Esuli ve Sebastiani, 2006: 417), Liu (Liu, 2007: 411), Abbasi ve diğerleri (Abbasi ve diğerleri, 2008: 14) çalışmalarında "metnin pozitif-negatif kutbunu belirleme" görevinden söz ederler. Bu görev, fikir ifade eden bir metnin pozitif-negatif kutuplardan hangisine dahil olacağını belirlemeye yönelik olarak geliştirilir. Ayrıca üzerinde en çok uygulama geliştirilen görev ünvanına sahiptir. Esuli ve Sebastiani aynı çalışmalarında "metnin pozitif-negatif kutup şiddetini belirleme" görevinden de bahsederler. Bu görevde, pozitif-negatif kutuplara atanmış metin tabanlı fikir verileri belli seviyelerde derecelendirilir. Levene (Levene, 2010: 416), Verbeke ve Eynde (Verbeke ve Eynde, 2013: 9), Qiu ve diğerleri (Qiu ve diğerleri, 2010: 2) ise ilgili çalışmalarında başka bir görev olan "özellik tabanlı madencilik" görevine yer vermişlerdir. Bu görevde ürün veya hizmetin herhangi bir özelliği veya bazı özelliklerinin nasıl algılandığına ilişkin bir değerlendirme yapılır. Yine Verbeke ve Eynde aynı çalışmalarında "sınıflandırma problemi olarak sorun konu etme" görevinden söz etmişlerdir. Burada da fikir madenciliğine serbest olarak istedik yönde bir problem konu etmek esastır.

Fikir madenciliğine ilişkin bir diğer görev ise bu çalışmaya konu olan "karşılaştırmalı cümle" görevidir. Bu görevde, iki ürün veya hizmetin karşılaştırılmasına ilişkin fikir ifadelerinden otomatik olarak bilgi çıkarımı esastır. Bir markanın diğer markaya üstünlüğü bu görevle tespit edilebilir.

Fikir madenciliğinin karşılaştırma görevine yönelik literatür incelendiğinde çalışmaların son yedi yıla yayıldığı görülmektedir. Bugüne kadar bu göreve ilişkin olarak İngilizce, Korece, Arapça ve Çince dillerinde çalışmalar yapıldığı gözlenmiştir. Aşağıda bu çalışmalardan bazı örnekler verilmiştir.

Jindal ve Liu 2006 yılında yayımladıkları çalışmalarında, o tarihe kadar karşılaştırma görevine yönelik herhangi bir çalışma bulamadıklarını belirtmişlerdir. Onlar bu çalışmada forum, blog ve haber sitelerindeki yorumların içinde geçen karşılaştırma cümlelerini tespit etmeye ve bu cümlelerdeki karşılaştırma ilişkilerini sınıflandırmaya çalışmışlardır. İlk adımda, İngilizce dilindeki karşılaştırma cümlelerinde geçen 83 son ek ve anahtar kelimeyi kullanarak tüm karşılaştırma cümlelerinin %81'ini %82'lik bir doğruluk oranıyla tespit edebilmişlerdir. İkinci adımda ise, bu cümleleri karşılaştırma ilişkilerine göre eşitsizlik, eşitlik ve en üstünlük olmak üzere üç ayrı kategoride sınıflandırmışlardır. Jindal ve Liu bu çalışmalarında karşılaştırma cümleleri üzerinde

fikir madenciliği yapmamışlardır. Bu çalışma, karşılaştırma cümleleri üzerinde yapılan ilk doğrudan çalışma olması nedeniyle sonraki araştırmalara temel teşkil etmesi bakımından önemlidir (Jindal ve Liu, 2006: 1331, 1334, 1336).

Benzer bir çalışmayı 2009'da Yang ve Ko Kore dili için yapmışlardır. Bu çalışmada da sadece karşılaştırma cümleleri tespit edilmeye çalışılmış, yine bu cümleler üzerinde fikir madenciliği yapılmamıştır. Onlar Kore dilinin yapısının İngilizce'den farklı olduğunu, bu nedenle Kore dilinde yazılmış yorumlardaki karşılaştırma cümlelerini tespit edebilmenin İngilizce dilinde yazılan yorumlara göre daha zor olduğunu belirtmişlerdir. Çalışmada ilk adımda, Kore dilinde karşılaştırma cümlelerinde kullanılan anahtar kelimeler kullanılarak yorumlar üzerinde bir tespit yapılmış, daha sonra Naïve Bayes ve Maksimum Entropi Modelleri ile bu yorumlarda karşılaştırma cümlesi olarak yer almayan yorumlar elenmiştir. Yang ve Ko, önerdikleri metod ile %88,5 oranında karşılaştırma cümlelerini doğru olarak tespit edebilmişlerdir. Bu çalışma, Kore dilinde karşılaştırma görevine yönelik ilk fikir madenciliği çalışması olma özelliğine sahiptir (Yang ve Ko, 2009: 153-156).

Arapça dilinde karşılaştırma cümlelerini tespit üzerine ilk çalışma ise, 2012 yılında El-Halees tarafından yapılmıştır. Halees çalışmasının başında Arapça'nın şekilsel bakımdan İngilizce'ye göre çok daha kompleks bir dil olduğunu ifade etmiştir. Bu çalışmada eğitim, teknoloji ve spor olmak üzere üç farklı alanda internet kullanıcıları tarafından yapılan 1048 yorum toplanmış, bunların 435'inin karşılaştırma ifadesi içerdiği manuel olarak belirlenmiştir. Karşılaştırma ifadeleri içeren bu yorumların tespit edilebilmesi için farklı yaklaşımlar ortaya konmuştur. İlk yaklaşım olarak, sıfat ve zarf sözcük türleri üzerinden tespitler yapılmış ve %63,7 oranında bir başarı elde edilmiştir. Daha sonra üç farklı makine öğrenmesi metodu denenmiştir. Naïve Bayes, K-En Yakın Komşu ve Destek Vektör Makinesi ile yapılan tespit çalışmasında en yüksek başarı oranı K-En Yakın Komşu metodu ile %86,6 olarak elde edilmiştir. Son olarak sözcük türleri ve makine öğrenmesi metodları bir arada kullanılarak bir test yapılmıştır. Bu çalışma sonucunda Destek Vektör Makinesi ve sözcük makineleri birlikte kullanıldığında %88,8 ile en yüksek başarı oranına ulaşılmıştır. El-Halees bu çalışmasında sadece karşılaştırma cümlelerini tespit etmeyi sağlayan bir yaklaşım ortaya koymuş, ancak bu cümleler üzerinde fikir madenciliği uygulaması gerçekleştirilmemiş, sonraki çalışmalarına bırakmıştır (El-Halees, 2012: 265, 266, 268, 269).

Ganapathibhotla ve Liu 2008'de yayımladıkları çalışmalarında, karşılaştırma cümleleri üzerinde yapılan ilk fikir madenciliği çalışmasını sunmuşlardır. Onlar karşılaştırma cümlelerinde hangi ürün veya hizmetin diğerine göre tercih edildiğini bulmaya çalışmışlar ve başarı oranı yüksek bir metod ortaya koymuşlardır. Öncelikle sıfat ve zarfları manuel olarak pozitif veya negatif sınıfta kategorize etmişler ve cümle içinde geçen bu sıfat ve zarflara göre o cümleyi pozitif veya negatif bir cümle olarak tanımlamışlardır. Örnek veri seti olarak Jindal ve Liu'nun 2006 yılında kullandıkları yorumlara ilave olarak e-ticaret ve teknoloji forumlarındaki laptop, kamera ve cep telefonları hakkındaki yorumları da kullanmışlardır. Binlerce yorum ve forum gönderisi içinden sadece 837 tanesinin (418 tanesi Jindal ve Liu'nun çalışmasından, 419 tanesi diğer e-ticaret ve teknoloji forumlarından) bu çalışmada kullanılabilir veri olduğunu görmüşler ve sosyal medyada yer alan yorumların sadece %10'unun karşılaştırma ifadesi içerdiğini tespit etmişlerdir. Değerlendirme sonunda cümlede ilk geçen ürün veya hizmetin %84 oranında ikinci ürün veya hizmete göre tercih edildiği şeklinde ilginç bir asimetric dağılım gözlemlenmiştir. Yani, verilerin %84'ünde ilk geçen ürün veya hizmet hakkında pozitif bir yorum yapılmaktadır. Dolayısıyla,

hiçbir değerlendirme yapılmadan cümlede ilk geçen ürün veya hizmet hakkında pozitif bir yorum yapıldığı varsayılırsa, %84 oranında bir doğruluk söz konusu olabilecektir. Ancak, Ganapathibhotla ve Liu sundukları metot ile %94,4 oranında bir doğruluk oranı elde etmeyi başarmışlardır. Diğer ilginç bir tespit de, cümlede geçen ikinci ürün veya hizmetin tercih edildiği yorumların, cümlede geçen ilk ürün veya hizmetin tercih edildiği yorumlara göre genelde çok daha kompleks olmasıdır (Ganapathibhotla ve Liu, 2008: 241, 245, 247, 248).

Görüldüğü gibi, karşılaştırma görevine yönelik fikir madenciliği çalışmaları nicelik, kapsam ve uygulandığı dil bakımından henüz çok sınırlı bir düzeydedir. Kaldı ki, mevcut çalışmaların bir kısmı da fikir madenciliğinde karşılaştırma yapmak üzere veri toplamaya yönelik olarak geliştirilmiştir. Bu çalışmada ise detayları aşağıda ayrıntılı olarak belirtileceği gibi veri toplamaya yönelik herhangi bir uygulama geliştirilmemiş, model üzerine yoğunlaşmıştır.

ARAŞTIRMANIN AMACI VE ÖNEMİ

Bu çalışmada sosyal medyada rekabet analizi için karşılaştırma görevine yönelik bir fikir madenciliği modeli geliştirilmiştir. Sosyal medyada yer alan kullanıcı görüşleri işletmeler için paha biçilmez bir değerdedir. Zira bu ortamlar tüketime sunulan ürün ve hizmetlere ilişkin önemli ölçüde kullanıcı yorumlarını içermektedir. Kullanıcılar bu ortamlarda kendi istekleriyle yer aldıklarından başka yollarla zor elde edilebilecek görüşlere bu yolla kolayca ulaşılabilirler. Benzer konularda ilgisi olan bireyleri de bir araya getiren sosyal medya, ilgili konularda milyonlarca sayıda yorum içermekte ve bu sayı her geçen gün katlanarak çoğalmaktadır. Bu nedenle sosyal medyada yer alan metin yığınları çok önemli veriler barındırmaktadır. Dolayısıyla işletmeler, sundukları ürün ve hizmetlerin tüketiciler tarafından nasıl karşılandığını öğrenmek amacıyla bu ortamları takip etmeli ve rekabet analizi açısından etkin bir değerlendirme yapabilecekleri mecralar olarak görmelidirler.

VERİNİN HAZIRLANMASI

Sosyal medya özelinde karşılaştırma siteleri, YouTube ve teknoloji haber sitelerinin forum sayfaları incelenmiş ve tercih doğrultusunda belirlenen “cep telefonu” ürününe yönelik karşılaştırma metinleri araştırmaya konu edilmiştir.

Karşılaştırma siteleri günümüzde e-ticaret alanında müşterilere sağladığı kolaylıklar açısından ön plana çıkmışlardır. Özellikle en iyi teknoloji ürünlerini en ucuzda nerede bulabileceği arayışında olan tüketiciler çoğunlukla bu tür sitelere başvurumaktadırlar. Site yapısı bir “karşılaştırma motoru” mimarisine dayanır ve kullanıcı dostu bir arayüze sahiptir. Karşılaştırılmak istenen ürünler karşılaştırma motoruna girilerek veya popüler karşılaştırma alternatiflerinden seçilerek incelenebilir ve eğer istenirse yorumlanabilir.

YouTube 2005 yılında kurulmuş kullanıcılara video içerikleri sunan bir web sitesidir. Kullanıcıların arama motoru vasıtasıyla girdikleri konuya ilişkin videolar izleyiciye sunulur ve eğer istenirse hemen altında yorumlara fırsat tanınır.

Teknoloji haber siteleri ise, kullanıcılara belli kategorilerdeki teknolojilere ilişkin haberleri sunar ve forum sayfaları ile kullanıcılara etkileşim imkânı sağlar. Forumlar çoğu zaman üyelik

gerektirmekle birlikte ziyaretçi olarak mesaj girilmesine de olanak verilmektedir.

Nitel araştırma yöntemlerinden iz sürme yöntemi (tracer), araştırmalarda veri toplamayı ve örnekleme için daha sistematik yapan etkili anahtar bilgi kaynaklarına ulaşmaya yardım eder (Altunışık ve diğerleri, 2010: 305). Araştırma kapsamında yukarıdaki web sitelerinden elde edilen kullanıcı yorumları, iz sürme yöntemi ile elde edilmiş ve bu yolla “karşılaştırma veri tabanı” oluşturulmuştur. Şöyle ki;

- Web sitelerindeki yorumlarda yapılan incelemede, iki ürün birbirleri ile karşılaştırılırken zarf/sıfat/edat tabanlı şu ifadelerin yoğunlukla kullanıldığı tespit edilmiştir: Daha, gibi, farklı (olarak), göre, den/dan, oranla, kadar, da/de, benzer olarak, biraz daha, en az, ise, fazla, fazlaca, azıcık, epey, çokça, çok (çok), pek çok, çok az, gayet, aşağı yukarı, şöyle böyle, oldukça, bayağı, bile.
- Yorumlarda bu temel ifadeler ek olarak sıfat kökenli şu kelimelere de büyük ölçüde yer verilmiştir: Hızlı, dayanıklı, iyi, güçlü, üstün, uzun, stabil, büyük, pahalı, önde, güvenli, gelişmiş, güzel, performanslı, sağlam, ucuz, kötü, ağır, düşük, zor, farksız, zayıf, berbat, kalitesiz.

Bu doğrultuda yapılan incelemelerde, anlamsal bakımdan karşılaştırma ifade eden en etkili anlatım biçiminin yukarıdaki listede yer alan “den/dan” tabanlı ifadeler olduğuna karar verilmiş ve diğer karşılaştırma biçimleri çalışma kapsamı dışında tutulmuştur. Bu durum çalışmanın bir sınırlılığı olarak değerlendirilebilir. Tablo 1’de bu çerçevede ele alınabilecek olumlu ve olumsuz yorum tiplerine örnekler verilmiştir. Olumlu yorumlar 1. markanın (X) üstünlüğünü ortaya koyarken, olumsuz yorumlarda durum tam tersidir. Yorumlarda değişebilir ifadeler parantez içinde verilmiştir.

Tablo 1: Kategorilere İlişkin Yorum Tipleri

Kategori (Olumlu/Olumsuz)	Yorum Tipi
Olumlu	“... X, Y’den (çok) daha (hızlı) ...”
	“... X, Y’den oldukça (iyi) ...”
	“... X, Y’den bile daha (üstün) ...”
	“... X, Y’den iki kat (güçlü) ...”
	“... X, Y’den bayağı (üstün) ...”
	“... X, Y’den (ucuz) ...”
	“... X, Y’den bir adım önde ...”
Olumsuz	“... X, Y’den daha (kötü) ...”
	“... X, Y’den daha (iyi) değil ...”
	“... X, Y’den (kötü) ...”
	“... X, Y’den (çok) daha (kötü) ...”

Ürünlerin karşılaştırma sırasında kullanıcılar tarafından esas alınan kriterleri ise Tablo 2’de verilmiş olup bu kriterler genel çerçeveyi temsil etmektedir.

Tablo 2: Yorumlarda Yer Alan Karşılaştırma İfadelerinin Kategori ve Kriterleri

Karşılaştırma Kategorisi	Kriterler
Ürünün Kendisine Yönelik Karşılaştırmalar	Hız performansı
	Pil ömrü
	Kamera özellikleri
	Hafıza büyüklüğü
	İşlemci özellikleri
	Ekran boyutu
	Sağlamlık
Ürünün İşletim Sistemine Yönelik Karşılaştırmalar	iOS
	Android
Ürünün Fiyatına Yönelik Karşılaştırmalar	Ucuz
	Pahalı

Yukarıda açıklanan iz sürme yöntemi sayesinde anahtar bilgilere açıklanan biçimde ulaşılmış ve yorumlar bir nitel araştırma tekniği olan içerik analizinin “analiz üniteleri” olarak değerlendirilmiştir. Bu doğrultuda “karşılaştırma veri tabanı” adını verdiğimiz bir test veri tabanı oluşturulmuştur. Karşılaştırma siteleri, YouTube ve teknoloji haber sitelerinin forum sayfalarındaki karşılaştırma ifadelerinden elde edilen bu veri tabanı 100 adet yorumun oluşturduğu örneklem grubudur. Bu veri tabanında sadece cep telefonu ürünü ile ilgili yorumlara yer verilmiş, ancak belli iki markayı karşılaştırma sınırlandırmasına gidilmemiştir.

Karşılaştırma Modeli

Sosyal Medya özelinde karşılaştırma siteleri, YouTube ve teknoloji haber sitelerinin forum sayfalarında yer alan kullanıcı yorumlarının otomatik analizi için Aytekin (Aytekin, 2013: 184) tarafından geliştirilen bir programdan yararlanılmış ve yorumlar “marka üstünlüğü” esasına göre analiz edilmiştir. Program 4745 adet sıfat/zarf temelli kelimedenden oluşan bir sözlük yapısına dayanmaktadır. Sözlükteki her bir kelime ait olduğu duygunun rengi çerçevesinde bir olumlu/olumsuz olasılık değerine sahiptir. Yorumları olumlu/olumsuz sınıflara ayırma, sözlükte yer alan kelimelerin olasılık değerleri ölçüsünde Naïve Bayes Bit Ağırlıklandırma Algoritması kurallarına göre gerçekleşmektedir. Ayrıca program yarı denetimli öğrenme tekniğine göre çalışmaktadır.

Bu noktada şunu da belirtmek gerekir ki, “kullanıcı dostu”, “uzun ömürlü”, “renksiz”, “mat” gibi cep telefonlarının olumlu/olumsuz özelliklerini belirten birtakım kelimeler de tespit edilmiş ve 4745 adetlik sözlüğe öngörülen olasılık değerleri ile eklenerek sözlük genişletilmiştir.

Öncelikle karşılaştırma veri tabanında yer alan 100 yorum manuel olarak olumlu/olumsuz yorum kapsamında işaretlenmiş ve bunun sonucunda 80 olumlu, 20 olumsuz yorum elde edilmiştir. Diğer yandan olumsuz sınıfta değerlendirilen yorum sayısının az olması doğaldır. Zira iki ürünün karşılaştırması yapılırken genellikle, birinin diğerinden üstün yanlarını ifade etme yoluna gidilir. Tablo 3’de manuel olarak olumlu ve olumsuz sınıfta işaretlenen örnek iki yorumun program tarafından hesaplanan değerleri verilmiştir.

Tablo 3: Olumlu ve Olumsuz Örnek İki Yoruma İlişkin Değerlendirmeler

Yorum	Manuel Değerlendirme	Program Tarafından Yapılan Değerlendirme	
“... X, Y'den daha iyi ...”	Olumlu	Yorumun olumlu sınıfta olma olasılığı: %39,83	39,83>10,16 (Olumlu)
		Yorumun olumsuz sınıfta olma olasılığı: %10,16	
“... X, Y'den daha ağır ...”	Olumsuz	Yorumun olumlu sınıfta olma olasılığı: %0,0014	9,08 > 0,0014 (Olumsuz)
		Yorumun olumsuz sınıfta olma olasılığı: %9,08	

Tablo 3'e göre, manuel olarak olumlu sınıfta değerlendirilen “... X, Y'den daha iyi ...” yorumu program tarafından da olumlu sınıfa atanmış olup X markasının Y markasına olan üstünlüğünü ifade eder. Benzer şekilde, manuel olarak olumsuz sınıfta değerlendirilen “... X, Y'den daha ağır ...” yorumu program tarafından da olumsuz sınıfa atanmış olup birini durumun tersine Y markasının X markasına olan üstünlüğünü ifade eder.

Diğer yandan, bazı yorumlar program tarafından olumlu veya olumsuz olarak sınıflandırılmamaktadır. Örneğin, bu durum “X, Y'den daha iyidir.” yorumu için geçerlidir. Çünkü bu yorumda geçip de program sözlüğünde yer alan herhangi bir sözcük veya sözcük grubu bulunmamaktadır. Dolayısı ile program olumlu veya olumsuz sınıflara ilişkin olasılık değerlerini hesaplayamamakta, bu da algoritmanın her iki sınıf için %50 olasılık değerini öngörmesine sebep olmaktadır. Eğer bu yorumda geçen “daha iyidir” ifadesi “daha iyi” biçiminde yer almış olsaydı, sorun olmayacaktı. Bu durumda bir metin madenciliği görevi olan gövde analizi görevi söz konusu olmaktadır. Yani gövde analizi ile “daha iyidir” ifadesi “-dir” ekinden ayrılarak “daha iyi” biçimine getirilebilir. Bununla birlikte, “X, Y'den daha iyidir.” yorumu eylem bakımından isim cümlesidir. Oysaki program sıfat/zarf temelli sözcüklerle çalışmaktadır. Bu nedenle “daha iyidir” ifadesinin gövde analizi ile ayrıştırılması yoluna gidilmemiştir.

BULGULAR VE DEĞERLENDİRME

Karşılaştırma modelinin başarısı, bir model başarı ölçütü olan duyarlılık ölçütü ile değerlendirilmiştir. Bu amaçla, 80 olumlu ve 20 olumsuz örneklerden oluşan karşılaştırma veri tabanının manuel olarak ve program tarafından atanan değerleri karşılaştırılmıştır. Buna göre;

Olumlu Duyarlılık Ölçütü=Doğru Olarak Sınıflandırılmış Olumlu Örnek Sayısı/Toplam Olumlu Örnek Sayısı

$$\text{Olumlu Duyarlılık Ölçütü} = 61/80 = \%76,25$$

Olumsuz Duyarlılık Ölçütü=Doğru Olarak Sınıflandırılmış Olumsuz Örnek Sayısı/Toplam Olumsuz Örnek Sayısı

$$\text{Olumsuz Duyarlılık Ölçütü} = 13/20 = \%65,00$$

Başka bir deyişle; öngörülen karşılaştırma modeli %76,25 oranla 1. markanın 2. markaya göre üstünlüğünü ve %65,00 oranla 2. markanın 1. markaya göre üstünlüğünü doğru olarak hesaplayabilmektedir.

422 Journal Academic Marketing Mysticism Online

Ayrıca belirtmek gerekir ki, olumlu sınıfta program tarafından sınıflandırılmayan yorum sayısı 5, olumsuz sınıfta sınıflandırılmayan yorum sayısı ise 1 olup, yukarıdaki değerler hesaplanırken bu yorumlar “yanlış sınıflandırılmış yorum” kapsamında değerlendirilmiştir.

Yukarıdaki değerlendirmelere ilave olarak, literatür çalışmaları ile benzerlik gösterdiğinden aşağıdaki tespitlerden de söz etmek yerinde olacaktır:

- Özellikle YouTube ve teknoloji haber sitelerinin forum sayfalarında yer alan yorumların içinde karşılaştırma ifadeleri daha az yer almıştır.
- Karşılaştırma ifadelerinde 1. olarak söz edilen marka, diğerine göre yaklaşık %75 oranında daha tercih edilebilirdir.
- İfadede geçen ilk marka hakkındaki olumsuz olan yorumlar, diğerine göre daha karmaşık bir yapıdadır. Bu durum, olumsuz duyarlılık ölçütünün daha düşük çıkması sonucu ile paralellik göstermektedir.

Ayrıca çalışma kapsamında, olumlu ve olumsuz yorumların kelime sayıları program veri tabanına kaydedildiğinden doğru/yanlış/sınıflandırılmamış yorum-kelime sayısı ilişkisi de incelenmiş ve bu doğrultuda Tablo 4 oluşturulmuştur.

Tablo 4: Doğru/Yanlış/Sınıflandırılmamış Yorum-Kelime Sayısı İlişkisi

	Doğru Olarak Sınıflandırılmış Yorumların Ortalama Kelime Sayısı	Yanlış Olarak Sınıflandırılmış Yorumların Ortalama Kelime Sayısı	Sınıflandırılmamış Yorumların Ortalama Kelime Sayısı
Olumlu	34,54	39,92	8,80
Olumsuz	35,15	35,33	6,00

Görüldüğü gibi, her iki yorum tipinde de doğru olarak sınıflandırılmış yorumların ortalama kelime sayısının üstünde olan yorumlar yanlış olarak sınıflandırılmaktadır. Demek ki, karşılaştırma ifadeleri uzunluk bakımından belli bir sınır değerinde olmalıdır.

Ayrıca, her iki yorum tipinde de geçerli olmak üzere ortalama kelime sayısının altındaki yorumlar ise sınıflandırılmamakta, program olasılık değeri atayacak yeterlilikte sözcük bulamamaktadır.

Bulgulara ilişkin son değerlendirme ise yazım denetimi üzerinedir: Toplumda iletişim süreci, kitle iletişimi, örgütsel düzey, gruplar arası düzey, kişilerarası düzey ve kişisel düzey olmak üzere beş düzeyde incelenebilir (Lazar, 2001:13). Sosyal medyada dört düzey iletişim sürecinden de söz etmek mümkündür. Özellikle gruplar arası düzey, karşılıklı olarak birbirlerine mesajlar gönderen belirli sayıda kişiler arasındaki iletişimi olarak sosyal medyada önemli bir yere sahiptir. Kullanıcılar bu ortamlarda etkileşim halindedirler ve bu esnada ilgili konulardaki fikirlerini yazarken konuşma dilini kullanırlar. Dolayısıyla sosyal medya dilinin informal bir yapıda olduğu söylenebilir. Ancak bu yapı, fikir madenciliği çalışmalarında olumsuz bir faktör olarak karşımıza çıkar. Zira kullanıcı yorumlarının yazım bakımından denetimini gerektirir ki, bu çok zor bir iştir

ve literatürde metin madenciliği çalışmaları kapsamında değerlendirilir.

Dünyada farklı diller üzerinde yazım denetimi çalışmaları yapılmaktadır. Ancak örneğin, bir kelimenin yanlış yazılablmesinin birçok alternatifi bulunabildiği gibi, yanlış yazımlar farklı anlamlara gelebilecek yeni kelimeler de oluşturabilirler. Metin madenciliği çalışmalarında bu durumda bazen yanlış yazımların da sözlüğe eklenmesi yoluna gidilebilmektedir. Örneğin, “güzel” sözcüğünün “quzel, güzel, güzelli, ” biçimleri program sözlüklerinde yerini alabilmektedir.

Bu çalışmada yazım denetimi büyük ölçüde, program tarafından Microsoft Word programına yönlendirilerek yapılmıştır. Ancak anlamsal birtakım denetimlerin manuel olarak yapılması da gerekebilmiştir. Örneğin, “..... telefonun yana donuk kullanımı da var” yorumunu ele alalım. Bu yorum Türkçe karakter kullanılmadan yazılmıştır. Ama denetim sırasında “donuk” sözcüğü de ayrı bir kelime olduğundan altı çizilmeyecek, doğru olarak kabul edilerek atlanacaktır. Halbuki “donuk” sözcüğü gibi gerçekte olması gereken “dönük” sözcüğü de sözlükte yer almaktadır. Dolayısı ile bu anlamsal denetim atlanırsa program “dönük” sözcüğü için yanlış bir olasılık değeri kullanacak ve böylece belki de yorum yanlış sınıfa atanabilecektir.

SONUÇ VE ÖNERİLER

Bu çalışma birtakım sınırlılıklara sahip olmakla birlikte, sosyal medyadaki Türkçe yorumların rekabet analizi açısından fikir madenciliği yöntemleri ile değerlendirilmesi konusunda yapılan ilk çalışma olması bakımından önemlidir. Türk insanının sosyal ağ kullanımı konusunda dünya listelerinin ön sıralarını zorladığı düşünüldüğünde bu önem daha da anlamlı hale gelmektedir.

Ayrıca geliştirilen modelin başarısının tespitine yönelik bulgular modelin geleceği için ümit vermekte ve gelecek çalışmalar için bir başlangıç noktası oluşturabileceği düşünülmektedir. Diğer yandan yazım denetimi ve gövde analizi konularının fikir madenciliği çalışmalarının önünde önemli bir engel teşkil ettiği söylenebilir.

Yazım denetimi ile benzer paralellikte karşımıza çıkan ve modelin başarısını olumsuz yönde etkileyebilen bir diğer konu ise, yorumların içerdiği kelime sayısıdır. Kelime sayısı bakımından fakir olan yorumlar sınıflandırma için yeterli sözcük zenginliğini sağlayamazken, fazla uzun yorumlar da başarısız sınıflandırmalara neden olabilmektedir. Bu konunun yorum yazan kullanıcının hazır bulunuşluk durumundan kaynaklandığı söylenebilir.

Bu çalışmada kullanılan karşılaştırma içeriğine sahip yorumlar sosyal medyadan manuel olarak kopyalanmıştır. Halbuki otomatik bir veri toplama, bütünsel bir fikir madenciliği yaklaşımı için çok önemlidir. Bu çerçevede gelecek çalışmalar için bir “veri toplama motoru” konusundan söz edilebilir. Bu nedenle sosyal medyada yer alan karşılaştırma ifadeleri çok iyi analiz edilmeli ve yorumlar farklı kategorilerde indekslenmelidir. Ayrıca her kategoriye ilişkin farklı analiz modelleri geliştirilmelidir.

Gelecek çalışmalara ilişkin bir diğer öngörü, analiz çalışmalarının online olarak gerçekleştirilmesi boyutunda ele alınabilir. Böylelikle işletmeler kendi ürün ve hizmetlerinin karşısında yer alan rakipleri anlık olarak görebilecekleri gibi, belli periyodlarla dönemsel değerlendirme yapma şansına da sahip olabilirler.

KAYNAKLAR

Abbasi, A., Chen, H. ve Salem, A. (2008). Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums. *ACM Trans. Information Systems*. 26 (3), p12-34.

http://ai.arizona.edu/intranet/papers/AhmedAbbasi_SentimentTOIS.pdf [VIEW ITEM](#)

Altunışık, R., Coşkun, R., Bayraktaroğlu, S. ve Yıldırım, E. (2010). *Sosyal Bilimlerde Araştırma Yöntemleri*. Sakarya: Sakarya Yayıncılık.

<http://www.idifix.com/kitap/sosyal-bilimlerde-arastirma-yontemleri-spss-uygulamali-engin-yildirim/tanim.asp?sid=ICBYVJ1SDB0B4KABRME4> [VIEW ITEM](#)

Aytekin, Ç. (2013). An Opinion Mining Task in Turkish Language: A Model for Assigning Opinions in Turkish Blogs to the Polarities. *Journalism and Mass Communication*. 3 (3), p179-198.

<http://www.davidpublishing.com/show.html?12556> [VIEW ITEM](#)

<http://mimoza.marmara.edu.tr/~cigdem.aytekin/> [VIEW ITEM](#)

El-Halees, A. (2012). Opinion Mining From Arabic Comparative Sentences. *The 13th International Arab Conference on Information Technology ACIT'2012*. p265-271.

<http://www.acit2k.org/ACIT/2012Proceedings/6025.pdf> [VIEW ITEM](#)

Esuli, A. ve Sebastiani, F. (2006). Sentiwordnet: A Publicly Available Lexical Resource For Opinion Mining. *Language Resources and Evaluation (LREC)*. p417-422.

http://gandalf.aksis.uib.no/lrec2006/pdf/384_pdf.pdf [VIEW ITEM](#)

Falcon, J. (2010). Opinion Mining in eDiscovery. *Opinion Mining Blog*. Retrieved December 8, 2013, from <http://jedefalconit.com/opinion-mining/opinion-mining-in-ediscovery>

<http://jedefalconit.com/opinion-mining/opinion-mining-in-ediscovery> [VIEW ITEM](#)

Ganapathibhotla, M. ve Liu, B. (2008). Mining Opinions in Comparative Sentences. *Proceedings of the 22nd International Conference on Computational Linguistics (Coling, 2008)*. p241–248. Manchester.

http://delivery.acm.org/10.1145/1600000/1599112/p241-ganapathibhotla.pdf?ip=78.173.245.224&id=1599112&acc=OPEN&key=BF13D071DEA4D3F3B0AA4BA89B4BCA5B&CFID=393339847&CFTOKEN=85780157&acm_=1388330283_8cf9f44d44b43af521c87c11a96db32a [VIEW ITEM](#)

Jindal, N. ve Liu, B. (2006). Mining Comparative Sentences and Relations. *Proceedings of the 21st National Conference on Artificial Intelligence*. 2, p1331-1336.

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.80.1219&rep=rep1&type=pdf> [VIEW ITEM](#)

Kırım, A. (2003). *Strateji ve Bire-Bir Pazarlama CRM*. İstanbul: Sistem Yayıncılık.

<http://www.sistem.com.tr/tanim.asp?sid=GZQTHFDRN73XCBEI0OJ4> [VIEW ITEM](#)

Lazar, J. (2001). *İletişim Bilimi*. (Çev: C. Anık). Ankara: Vadi Yayınları.

http://www.iku.edu.tr/EN/97-2-51674-64-1259-598-1-1-1/ders_detay [VIEW ITEM](#)

http://books.google.com.tr/books/about/%C4%B0leti%C5%9Fim_bilimi.html?id=5V4cAAAACAAJ&redir_esc=y [GS SEARCH](#)

Levene, M. (2010). *An Introduction to Search Engines and Web Navigation*. 2nd ed. New Jersey: Wiley Publisher. p 416.

http://books.google.com.tr/books?id=mDI72_9-bw0C&printsec=frontcover#v=onepage&q&f=false

[GS SEARCH](#)

Liu, B. (2007). *Web Data Mining*. Chicago: Springer.

http://neerci.ist.utl.pt/neerci_shelf/MERC/AEP_Gestao%20das%20Redes,%20da%20Informacao%20e%20dos%20Servicos/2%20Semestre/Extracao%20e%20Analise%20de%20Dados%20Web/Books/web_data_mining.pdf [VIEW ITEM](#)

Qiu, G., Liu, B., Bu, J. ve Chen, C. (2010). Opinion Word Expansion and Target Extraction Through Double Propagation. *Computational Linguistics*. 37 (1), p9-27.

https://dspace-prod-lib.cc.uic.edu/bitstream/handle/10027/8193/coli_a_00034.pdf?sequence=1
[VIEW ITEM](#)

Verbeke, M. ve Eynde, W. V. (...). Opinion Mining. Retrieved December 8, 2013, from <http://people.cs.kuleuven.be/~bettina.berendt/WebMining10/L3.pdf>

<http://people.cs.kuleuven.be/~bettina.berendt/WebMining10/L3.pdf> [VIEW ITEM](#)

Yang, S. ve Ko, Y. (2009). Extracting Comparative Sentences from Korean Text Documents Using Comparative Lexical Patterns and Machine Learning Techniques. *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*. p153-156. Suntec, Singapore. 4 August 2009

<http://wmmks.csie.ncku.edu.tw/ACL-IJCNLP-2009/Short/pdf/Short039.pdf> [VIEW ITEM](#)