



Türkçe Duygu Kütüphanesi Geliştirme: Sosyal Medya Verileriyle Duygu Analizi Çalışması

Semra Yıldırım¹, Yücel Batu Salman², Serkan Ayvaz^{2*}

¹ Bahçeşehir Üniversitesi, Fen Bilimleri Enstitüsü, Mühendislik Yönetimi Bölümü, İstanbul, Türkiye

² Bahçeşehir Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Yazılım Mühendisliği Bölümü, İstanbul, Türkiye

(İlk Geliş Tarihi 7 Mart 2019 ve Kabul Tarihi 17 Nisan 2019)

(DOI: 10.31590/ejosat.537085)

ATIF/REFERENCE: Yıldırım, S., Salman, Y. B. & Ayvaz, S. (2019). Türkçe Duygu Kütüphanesi Geliştirme: Sosyal Medya Verileriyle Duygu Analizi Çalışması. *Avrupa Bilim ve Teknoloji Dergisi*, (16), 51-60.

Öz

Bu çalışmanın amacı, Türkçe için kapsamlı yeni bir duygu kütüphanesi geliştirmektir. Bu yeni kütüphane ile Türkçe sosyal medya paylaşımlarında etkili duygu analizi çalışmalarının yapılmasına katkı sunmak hedeflenmektedir. Bu çalışmada, varolan diğer kütüphanelerden bazıları incelenmiş olup bunları genişleten kapsamlı bir kütüphane oluşturulmuştur. Daha önce yapılmış çalışmalara ek olarak kütüphaneye basit emoji karakterler ve puanlama altyapısı eklenmiştir. Geliştirilen Türkçe Duygu Kütüphanesi'nin verimliliğini ölçmek için Twitter'da belli etiketlerle oluşturulmuş paylaşımlar toplanmıştır ve bunların üzerinden duygu analizi çalışmaları yapılmıştır. Analiz çalışmaları, birbirinden bağımsız iki farklı konu başlığında gerçekleştirilmiştir. Bu analiz kapsamında yapılan çalışmalardan ilki hava değişikliklerinin kullanıcılar üzerindeki etkisini ölçmek adına yapılmıştır. Bu çalışma kapsamında yaz aylarında ve kış aylarında toplanan veri kümeleri incelenmiştir. Türkçe paylaşım yapan kullanıcıların kış aylarında daha olumsuz paylaşımlar yaptıkları, bahar ve yaz aylarında ise kış aylarına oranla daha olumlu paylaşımlar yaptıkları gözlemlenmiştir. Diğer analiz çalışması ise Türk televizyonlarında belirli bir süre boyunca yayımlanan, günlerce popüler konu listesinin en üst sırasında yer alan Survivor adlı yarışma programı izleyicilerinin Türkçe paylaşımlarını konu almıştır. Analiz sonucunda kullanıcıların takip etmekte oldukları bu programla ilgili genelde olumsuz paylaşımlar yaptığı gözlemlenmiştir. Ayrıca, paylaşımlarda bozuk Türkçe kullanım alışkanlıkları tespit edilmiştir.

Anahtar Kelimeler: Duygu Analizi, Sentimental Analiz, Türkçe Duygu Kütüphanesi, Sosyal Medya, Twitter

Developing Turkish Sentiment Lexicon: A Sentiment Analysis Study using Social Media Data

Abstract

The objective of this study was to develop a comprehensive Sentiment Analysis Lexicon for Turkish language. The aim of this new lexicon is to facilitate effective sentiment analysis on Turkish social media posts. Previously developed lexicons have been examined and a comprehensive lexicon extending previous work has been developed. As an extension to the earlier work, the lexicon consist simple emoji characters and scoring infrastructure. To evaluate the performance of the Turkish Sentiment Analysis Lexicon, the tweets with certain hashtags were collected and analyzed. Analysis studies were carried out on two different topics. The first of these studies was done to measure the effect of weather changes on users. The data collected during the summer and winter were examined. The users who share Turkish post negative tweets in winter and appeared to share more positive posts in spring and summer compared to winter. The other analysis was conducted on Survivor TV Show, which has been aired on Turkish TVs, and placed at the top of the Trend Topic list for days. We observed that the users share negative posts about this program. Additionally, the habit of improper use of Turkish language on social media was noted in their posts.

Keywords: Emotion Analysis, Sentimental Analysis, Sentimental Lexicon for Turkish, Social Media, Twitter

* Sorumlu Yazar: Bahçeşehir Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Yazılım Mühendisliği Bölümü, İstanbul, Türkiye, ORCID:0000-0003-2016-4443, serkan.ayvaz@eng.bau.edu.tr

1. Giriş

Son yıllarda yaygın internet kullanımı ile devasa boyutlarda veriler oluşmaya başlamıştır. Bu toplanan büyük verilerin yönetilmesi, analiz edilmesi, metinlerin içeriklerinde duygu ve düşünce yoğunluklarının belirlenmesi ve duygu yüklerine göre gruplanması birçok bilimsel çalışmaların ana konusunu oluşturmaya başlamıştır. Yapılan bu çalışmalar ile birlikte duygu analizi bir bilimsel çalışma alanı olarak ortaya çıkmıştır (Liu, 2012; Pang, Lee, & others, 2008).

Duygu analizi büyük verinin analiz edilip, kullanılabilir verilerinin içinden anlamlı veriyi çıkartmak için oluşturulan yöntemlerdendir. “Duygu Analizi (Sentiment Analysis)” ya da “Düşünce Madenciliği (Opinion Mining)” gibi birçok isimlendirmeler kullanılmaktadır (Nasukawa & Yi, 2003).

İnternetin yaygın kullanılması nedeniyle artık geleneksel metotlar ile el yordamıyla yapılabilecek olan takipler söz konusu olmaktan çıkmıştır. Örneğin bir kurumla ilgili çıkan tüm haberleri takip etmek ya da bu iş için birkaç kişilik ekipler kurmak daha önce mümkün iken günümüzde bu söz konusu bile değildir. Her geçen gün yeni bloglar, haber siteleri ve paylaşım siteleri ortaya çıkarak bu dünyayı zenginleştirmektedir. Bu büyük verileri saklayabilmek ve ihtiyaç dahilinde kullanabilmek ve ayrıca sonuçlar çıkartarak yeni yollar bulmak günümüzün en popüler konularından biridir. Sınırsız bilgi, yorum akışının olduğu bir dünyada, verileri sınıflandırma işini otomasyona bırakmaktan başka bir yol bulunmamaktadır (Ayvaz & Shiha, 2018).

İnternetin yaygınlaşmış kullanılmaya başlanmasıyla beraber insanlar diğer insanlar, kurumlar, ürünler gibi birçok konu hakkındaki görüşlerini açık bir şekilde ifade etmeye başlamışlardır. Günümüzde bu veriler üzerinden birçok analiz çalışması yapılabilmektedir (Dokoohaki, Zikou, Gillblad, & Matskin, 2015; Pak & Paroubek, 2010; Tang & Liu, 2010). Duygu analiz modelleri Twitter üzerinde satış performans tahminleri, ürünleri kademelendirme, kamuoyu yoklamaları, oylama sonucu tahminleri, politik bakış açılarının keşfi, filmlerin gişe tahminleri gibi birçok alanda kullanılmaktadır (Dokoohaki et al., 2015; Duncombe, 2011; Java, Song, Finin, & Tseng, 2007; Ozturk & Ayvaz, 2017).

Son yıllarda artan sosyal medya kullanımı nedeniyle sosyal medyada duygu analizleri çok önem kazanmıştır. İnsanların sosyal medya üzerinden paylaştıkları metinlerden kişiler hakkında olumlu, olumsuz gibi analizler yapmak mümkün olmaktadır. Günümüzde kullanılan sosyal medya uygulamaları arasından en fazla duygu yüklü kelime içeren platformlardan biri Twitter’dir (Aylin & Elif, 2017; Nasukawa & Yi, 2003; Shiha & Ayvaz, 2017; Tang & Liu, 2010). Twitter’da kullanıcılar dünyada daha önceden gerçekleşmiş ya da şu anda gerçekleşmekte olan herhangi bir konu hakkında anında paylaşım yapmaktadır. Bu paylaşımlar sayesinde kitleler olumlu ya da olumsuz olarak etkilenmektedir. Bu etkilenmeleri ölçümlendirmek ve sonuçlar üretebilmek bu çalışmada için sürekli veri akışı olan Twitter platformu kullanılmıştır (Ayvaz & Shiha, 2018).

Bu araştırma kapsamında Twitter’da yer alan Türkçe veriler üzerinden duygu analizi çalışmaları yapılmıştır. Türkçe Twitter kullanıcılarının belli etiketlerle paylaştıkları veriler incelenmiştir. Etiketlerin belirlenebilmesi için Twitter’ın popüler konu listeleri ve çok takipçiye hitap eden kullanıcıların paylaşımları incelenmiş ve tüm bu çalışmaların ardından etiket seçimleri yapılmıştır. İncelenen bu veriler toplanarak veri tabanında saklanmıştır. Verilerde kullanıcı bilgisi, paylaşım metni, tarih, aldığı beğeni sayısı gibi detaylı bilgiler yer almaktadır. Bu bilgiler sayesinde farklı etiketlerle (hava durumu, yazgelsin, ilkbahar, sonbahar, kış, survivor) ve farklı başlıklarda analiz çalışmaları yapılmıştır. Tüm etiketler için tek bir duygu kütüphanesi kullanılmıştır.

2. Literatür Taraması

Günümüzde duygu analizi politik ve siyasi konular içinde kullanılmaya başlanmıştır (Dokoohaki et al., 2015). Vatandaşlar siyasi parti adaylarıyla ilgili olarak diğer insanların görüşlerini bilmek istemektedirler. Geçmişte bu gibi düşünce ihtiyaçları aile, arkadaş ve yakın çevreyle sınırlıydı. Fakat şimdilerde web üzerinden her türlü bilgi kolayca elde edilebilmektedir. Bu gibi amaçlarla tüketicilerin düşüncelerini öğrenmek için anketler, kamuoyu yoklamaları yapılmaktadır. Varolan bütün verileri bir araya getirip bu veriler üzerinden anlamlı sonuçları görüntüleyebilmek için duygu analizi çalışmaları yapılmaktadır (Parycek, 2012). Bunun en büyük örneklerinden Amerika Birleşik Devletleri ve Almanya seçimleri gösterilmektedir (Java et al., 2007). Örneğin, 2008 yılında Barack Obama’nın kampanya ekibi tarafından Twitter, Facebook, MySpace gibi birçok sosyal medya platformundan veriler toplanmış ve bunlar üzerinden seçim sonuçları tahmin edebilmek için analiz çalışmaları yapılmıştır (Pak & Paroubek, 2010; Dokoohaki, Zikou, Gillblad, & Matskin, 2015; Tang & Liu, 2010). Obama ve ekibi bu başarılı sosyal medya analizi sayesinde 500 milyon dolar bağış elde etmiştir (Cihan, Mustafa, Evren, H. Kaan, & Duygu, n.d.; Vargas, 2008).

Son yıllarda, sosyal medya paylaşımlarının organizasyonları tekrar şekillendirdiğine, toplumsal düşünce ve duyguları sarstığına şahit olunmuştur (Akgul, 2011; Aylin & Elif, 2017). Örneğin, 17 Aralık 2010’da bazı Arap ülkelerinde sosyal medyanın katkılarıyla politik değişiklikler olduğu görülmüştür (Duncombe, 2011). Arap ülkelerindeki sosyal medya kullanımının artması isyanların büyümesine ve halkın örgütlenmesine katkı sağlamıştır. Bu nedenle sosyal medya devrimi olarak adlandırılmıştır. Bu durum sonucunda da artık sosyal medya üzerindeki bu fikirleri içeren paylaşımları toplama ve analiz etmenin önemi artmıştır.

Pang ve Lee, destekçi vektör makinası algoritmaları ile yıldız (*) ve sembollerle etiketlenmiş verilerde gruplama yapma problemini araştırmışlardır ve oluşturdukları çoklu sınıf destekçi vektör makinası regresyon algoritmaları ile yüzde 60 gibi bir başarı elde etmişlerdir (Pang & Lee, 2005). Daha sonra ise veriden öznel ifadeleri çıkartarak tekrardan hesaplamalar yapmış ve ortalama yüzde 85 olarak başarı oranlarını güncellemişlerdir. Turney (2002), web üzerinden topladığı 410 yorumu cümlelerin polaritesine PMI-IR algoritmasını kullanarak sınıflandırmıştır. Turney, bu araştırmasında 4 farklı konu üzerine yoğunlaşmıştır. Çalışma sonucunda

yüzde 74 gibi bir başarı elde etmiştir (Turney, 2002). Arapça Sentimental Analiz (ASA) çalışmalarının popülerliği hızla artmaktadır. Bu nedenle de Duwairi çalışmasında diyalektik kelimeler kullanarak Arapça paylaşımlar üzerinden duygu analiz çalışmaları gerçekleştirmiştir. Makine öğrenmesi gruplama yöntemi ile analiz çalışmaları yapmıştır (Duwairi, 2015). Ayrıca Arapça için Doğal Dil İşleme (DDİ) ile de birçok çalışmalar yapılmıştır (Nassar & Khamayseh, 2015). Shankar Setty Facebook'da bulunan Haber Akışı (New Feed) üzerinde Gmail'in "Gelen Kutusu"nda kullanıldığı gibi gruplama önerisinde bulunmuştur. Önerdiği bu yapı haber akışından gönderilen fikirlerin önem durumlarına göre sınıflanması üzerinedir (Setty, Jadi, Shaikh, Mattikalli, & Mudenagudi, 2014). Wöllmer, Wenninger, Sun ve ekibi tarafından (2015) YouTube'da özellikle film incelemeleri içeren videolar üzerine çalışmalar gerçekleştirmiştir. Yaptığı çalışma videoda bulunan konuşmayı alıp bu konuşma üzerinden duygu analiz çalışmaları yapmayı hedeflemiştir (Wollmer et al., 2013).

2.1. Türkçe Tweetler Üzerinden Yapılan Duygu Analizi Çalışmaları

Türkçe Twitter paylaşımları üzerinden birçok deney ve araştırmalar yapılmıştır. Fakat bu çalışmalar esnasında birçok engelle karşılaşmıştır. Bu engeller genel olarak iki grup altında incelenebilir: (i) dil yapısı; (ii) kullanıcı alışkanlıklarıdır. Türkçe'nin diğer dillerden daha farklı yapıya sahip olması, daha esnek olması ve kullanıcıların yaygın olarak İngilizce klavye kullanma alışkanlıkları nedeniyle Türkçe duygu analiz yapma çalışmaları çok zaman almaktadır. Ayrıca kullanıcıların Twitter'da yer alan karakter kısıtlaması nedeniyle kelimelerin bazı harflerini eksik yazma alışkanlıkları bulunmaktadır. Bunlar gibi engellere rağmen Türkçe kelimeler üzerinden duygu analizi konusunda birçok çalışma yapılmıştır. Bu konuda firmalar uzmanlaşmak için bölümler kurmuştur. İnternette yer alan Türkçe verinin analiz edilmesi ve böylelikle yeni hizmetlerin ve iş dallarının oluşması araştırmacılar ve çeşitli sektörler için faydalar içermektedir.

Yapılan bu çalışmalarda birçok gruplama teknikleri, analiz seviyeleri ve teknolojiler kullanılmıştır. Çalışmalar farklı teknikler ve yöntemlerde gerçekleştirilmiştir. Çalışmalarda çeşitli veri tabanları kullanılmıştır. Bazıları ilişkisel veri tabanları kullanmıştır bazıları ise ilişkisel olmayan veri tabanları kullanmıştır. Bu durum tamamen analiz edilmesi planlanan veri büyüklüğüne ve teknik ekibin bilgisine göre belirlenmektedir. Yapılan bazı çalışma örnekleri aşağıda yer almaktadır: Erkut Evirgen tez çalışmasında Türkçe tweetlerde duygu analizi konusunda araştırmalar yapmıştır. Tez çalışması kapsamında R programlama dilini kullanmıştır. Tweetlerde ki düzgün formatlı olmayan kelimeleri düzenleyip işlediğinden bahsetmiştir. Ayrıca yaygın makine öğrenme metotları olan destekli vektör makinası, rasgele orman karar ağaçları, artırılma algoritmaları (boosting), maksimum entropi, yapay sinir ağları karşılaştırmalarını yapmıştır (Evirgen, 2016). Dilek Kayahan, Aslı Sergen ve Banu Diri yaptıkları duygu analizi çalışmalarıyla Twitter üzerinden Türkiye'deki televizyon programlarının reytinglerinin tahminlemesini gerçekleştirmişlerdir. Bu çalışma için kendi belirledikleri etiketler sınıflandırılmış ve sınıflandırmalar sonrasında da duygu kütüphanesinden kontroller yapılmıştır (Kayahan, Sergin, & Banu, n.d.). B. İbrahim Sevindi, Türkçe film yorumlarının duygu kutuplarını çeşitli yöntemler kullanılarak belirlemeye çalışmıştır. Makine öğrenmesi ve sözlük tabanlı yaklaşımlar kullanılarak karşılaştırmalar yapmıştır. Türkçe film yorumlarının duygu analizinde makine öğrenmesi yöntemleri başarılı sonuçlar üretmiştir (Sevindi, 2013). Sinem Demirci tarafından yapılan araştırmalar esnasında "finer-grained emotion analysis/parçalı duygu analizi" yönteminin özellikle Türkçe kaynaklı yapılan analiz çalışmalarının da etkili olduğu gözlemlenmiştir. Türkçe için hali hazırda bulunabilecek bir veri kümesi kullanılması, duygu analizi işleminde yetersiz kalmaktadır. Bu durumun önüne geçebilmek amacıyla, kullanılan kelimeler duygulara göre sınıflandırılmış ve bu veri kümeleri yardımıyla tweet gönderen kişinin duygu analizi gerçekleştirilebilmiştir (Demirci, 2014). Eyüp S. Akgül, Caner Ertona ve Banu Diri taraflarınca yapılan araştırmalarda Twitter verileri kullanılmıştır. Bu veriler analiz edilip gruplanmıştır. Gruplamalar "olumlu", "olumsuz" ve "nötr" olarak yapılmıştır. Böylelikle bir duygu kütüphanesi kullanılmıştır. Bu makale kapsamında geliştirilen duygu kütüphanesini zenginleştirme aşamasında Sercan ve diğerleri tarafından oluşturulan kütüphaneden de yararlanılmıştır (Sercan, Caner, & Banu, 2017). Nazan Öztürk ve Serkan Ayvaz tarafından gerçekleştirilen sentimental analiz çalışmalarında Suriyeli mültecilerle ilgili olarak İngilizce ve Türkçe Twitter paylaşımları R programlama dili yardımıyla toplanmıştır. Toplanan verilerde yapılan analiz çalışmalarında Türkçe paylaşımlarda pozitif duygulara daha fazla yer verildiği gözlemlenmiştir. İngilizce paylaşımların ise genelde nötr ve olumsuz oldukları sonucuna varmışlardır. Bu araştırma kapsamında da bir duygu kütüphanesi kullanılmıştır (Ozturk & Ayvaz, 2017).

3. Veri ve Yöntem

Türkçe'nin esnek bir dil olması ve Twitter'da çok fazla kirli verilerin bulunması nedeniyle temiz veriler elde edebilmek için yaklaşık bir yıl boyunca farklı farklı etiketlerle paylaşımlar toplanmıştır. Toplanan kayıtlar üzerinden incelemeler yapıp kütüphanede zenginleştirme ve iyileştirmeler yapılmıştır. Bu incelemeler esnasında Türkçe paylaşımlar yapan kullanıcıların genelde İngilizce klavye kullanmalarından dolayı oluşan aksaklıklarla karşılaşmıştır. Bu durumları yönetebilmek adına algoritmalar üretilmiştir. Örneğin kelimelerin tamamı küçük harfe dönüştürülmüş ve ardından kütüphaneden kontrol edilmiştir. Boşluklar ve noktalama işaretleri gibi kontroller yapılmıştır. Bu düzenlemelerin ardından bazı kelimelerin büyük-küçük harfle yazımından oluşan hatalar giderilmiştir. Buna ek olarak paylaşım metnindeki kelimeler hem Türkçe alfabeyle hem de İngilizce alfabeyle çevrilmiştir. Bu dönüştürmelerin ardından duygu kütüphanesinden eşleşmeler kontrol edilmiştir. Sadece bir eşleme üzerinden puanlama yapılmıştır. Diğer bir ifadeyle, hem İngilizce metne hem Türkçe metne puanlama yapılmamıştır. Ayrıca, Türkçe'de yer alan ikilemelerden dolayı oluşan hesaplama hatalarını yönetebilmek için de ek algoritmalar geliştirilmiştir. Her bir kelime grubu sadece tekil puanlanacak şekilde düzenlemeler gerçekleştirilmiştir. Bu çalışmalar sonrasında yapılan deneylerden anlamlı birçok sonuç elde edildiği gözlemlenmiştir. Örneğin; kış aylarında kişilerin daha umutsuz ve karamsar oldukları sonucu elde edilmişken, beklenildiği gibi yaz aylarının gelmesiyle kişilerin daha umutlu ve baharı heyecanla karşılayan duygulara sahip oldukları paylaşımlarda görülmüştür.

3.1. Duygu Kütüphanesi

Duygu analizi bir cümlenin ya da bir metnin duygu içerip içermediğini ve bu duygunun durumunun saptanması süreci olarak da adlandırılır (Liu, 2010). Duygu barındıran metinler herhangi bir konu, şahıs, marka ya da siyasi görüş hakkında olabilir. Toplumsal ve sosyal araştırmalar kapsamında yapılan anketler ve incelemelere bakıldığında duygu analizinin sosyoloji bilim dalını da yakından ilgilendirdiği görülmektedir. Örneğin; toplumsal tepkilerin ya da ayaklanmaların yoğun olduğu dönemlerde genellikle ilgili konularda olumsuz bir dil kullanıldığı görülmektedir. Bu tepkiyi doğuran kişi, kurum ya da kuruluşlar hakkında birçok duygu barındıran metin sosyal medya üzerinden keşfedilebilir ve analiz edilebilir. Duygu barındıran kelimeler yönelimlerine uygun olarak sınıflandırılıp taşıdıkları duygu yoğunluklarına uygun olarak puanlandırılabilir. Örneğin; “güzel”, “harika”, “mükemmel” olumlu duygu içeren kelimeler olup “Kötü”, “zayıf”, “korkunç” ise olumsuz duygu içeren kelimeler olarak değerlendirilebilir. Bunun gibi kelimelerin ve ifadelerin liste olarak saklanması sonrasında duygu kütüphaneleri oluşturulur.

Geçen yıllar boyunca araştırmacılar duygu kütüphaneleri derlemek için birçok algoritmalar tasarlamışlardır (Claster, Dinh, & Cooper, 2010). Duygu kelimeleri iki türde olurlar: baz (base) kelimeler ve karşılaştırmalı (comparative) kelimeler. Örneğin: “iyi”, “kötü” baz kelimedir. Öte yandan “daha iyi”, “daha kötü” karşılaştırmalı kelimedir (Ganapathibhotla & Liu, n.d.). Bu araştırma kapsamında yeni bir duygu kütüphanesi oluşturulmuştur. Bu çalışma kapsamında oluşturulan Türkçe duygu kütüphanesi ilgilene araştırmacılar için araştırma çalışmalarında kullanılabilir hale getirildi². Kütüphanenin zenginleştirme süreci ise aşağıdaki gibidir:

- i. Türk Dil Kurumu (TDK) sözlüğü: Kütüphanenin zenginleştirme çalışmalarında TDK sözlüğü üzerinden birçok kelime incelenmiş ve duygu analiz çalışmalarına katkı sağlayacağı öngörülenler ayıklanarak oluşturulan duygu kütüphanesine eklenmiştir.
- ii. Daha önceden oluşturulan kütüphaneler: Eyüp Akgül, Caner Ertona, Banu Diri (2010) tarafından geliştirilmiş kütüphaneden bazı kelimeler, oluşturulan duygu kütüphanesine eklenmiştir (Eyüp Sercan et al., 2017).
- iii. Emoji Listesi: Saif Mohammad ve Peter Turney (2011) tarafından oluşturulan liste duygu kütüphanesine eklenmiştir (Turney & Mohammad, 2011). Emojiler listesinde düzenlemeler yapılmıştır. Daha detaylı belirtmek gerekirse, ilk olarak her bir emoji için olumlu, olumsuz veya nötr olarak gruplama yapılmıştır. Ardından emojilerde kelimelere benzer şekilde duygu değerleri verilerek puanlama yapılmıştır. Değerler -5 ile +5 arasındadır.
- iv. Toplanan Twitter verileri incelenerek kütüphanede bulunmayan duygu yüklü kelimeler tespit edilmiş ve duygu kütüphanesine eklenmiştir.
- v. Türkçe karakter kullanamamaktan kaynaklanan sorunları göz ardı etmemek için bazı yeni küçük algoritmalar yazılmıştır. Başka bir deyişle, kelimelerde yer alan Türkçe harflerin tamamı İngilizce harflerle değiştirilerek yeni bir kolona kaydedilmiştir. Her bir kelime için duygu kütüphanesinden kontrol yapılırken öncelikli olarak Kelimeler tablosunda kontrol yapılır. Eğer metin kolonunda kelime bulunamazsa eşleniği kolonundan kontrol yapılır. Kelimelerin eşleşmesi durumunda kelimeye ilgili puan verilir.

Ayrıca bazı özel kelime grupları için farklı puanlama algoritması geliştirilmiştir. Bu kelime gruplarında duygu yoğunluğu daha fazla olduğu için puanlama sistemi de farklı olarak oluşturulmuştur. Olumlu kelime grupları +6 ile +11 arasında değer verilmiştir. Olumsuz kelime gruplarına ise -6 ile -11 arasında değer verilmiştir. Örneğin: “burnu büyük” kelime grubu olumsuz olarak kategorize edilmiştir ve duygu yoğunluk puanı -8 olarak belirlenmiştir. Argo/kötü söz içeriklerine ise -11 puan verilmiştir. Bu puanlama skalasının seçiminde amaç argo kelimeler ve mükemmel olumluluğu ifade eden kelime gruplarının kullanılması durumunda genel duygu değerlerinin yansıtılmasında daha belirleyici bir etki oluşturulmasının sağlanmasıdır. Bu yaklaşımın daha gerçekçi sonuçlar elde ettiği gözlemlenmiştir. Objektif bir puan elde etmek için, sözlükteki terimler anadili Türkçe olan 2 farklı yüksek lisans öğrencisi tarafından elle puanlanmıştır. Çıkan sonuçlar araştırma ekibi tarafından kontrol amaçlı tekrar incelenmiştir. Aşağıdaki Tablo 1’de yer alan duygu kütüphanesi veri kümesi tablosunda bazı örnek kelimelerin olumlu (1), olumsuz (2) durumları, bu kelimelerin İngilizce alfabedeki karşılıkları ve kelimelerin duygu değerleri görüntülenmektedir.

² <https://github.com/sayvaz/turkish-lexicon/>

Tablo 1. Duygu Kütüphanesi Örnek küme

İzlenim	Metin	Değer	Eşlenik
2	minimum	-3	minimum
2	alkolik	-5	alkolik
2	az	-3	az
2	açgözlü	-5	acgozlu
2	cüzi	-2	cuzi
2	düztaban	-4	duztaban
2	haram	-5	haram
2	hödük	-5	hoduk
1	enerjik	5	enerjik
1	enfes	5	enfes
2	kahpe	-5	kahpe
2	serseri	-3	serseri
2	kokoş	-2	kokos
1	bedava	2	bedava
1	ağırbaşlı	1	agirbasli
2	burnu büyük	-8	burnu buyuk
2	bölük pörçük	-7	boluk porcuk
2	eli sopalı	-8	eli sopali
2	kuş beyinli	-8	kus beyinli
1	en güzel	11	en guzel

4. Analiz ve Bulgular

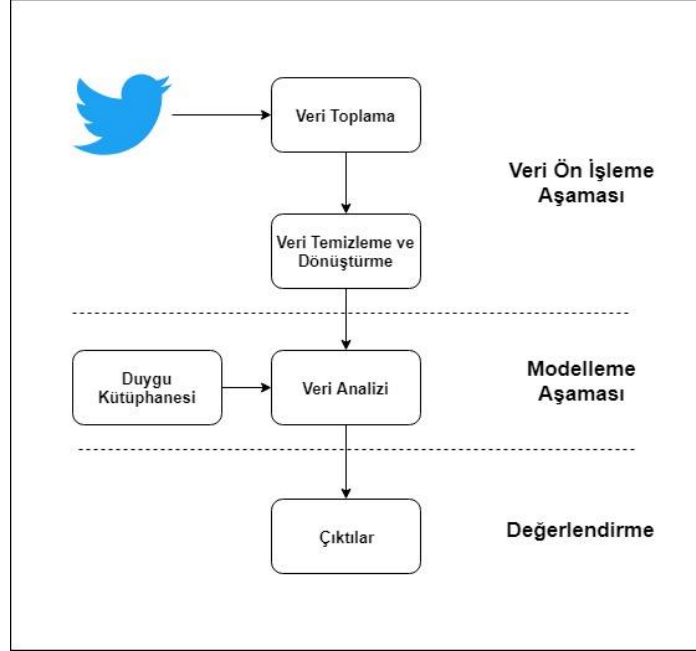
Bu araştırmada yapılan analiz çalışmalarında Twitter verileri incelenmiştir. Bu incelemelerde popüler konu listesi değerlendirilip bazı etiketler belirlenmiştir. Böylelikle kararlaştırılan etiketler kümeler halinde Twitter'dan oluşturulan veritabanına aktarılmıştır. Şekil 1'de duygu analizi algoritmasının çalışma prensibi ve işleyişi görselleştirilmiştir. Verileri çekme, ayıklama ve düzenleme işlemleri için C#.NET ve Python yazılım dilleri kullanılmıştır. Elle veri üzerinde gözlemler yapılmış ve bu incelemeler sayesinde duygu kütüphanesine eklemeler gerçekleştirilmiştir. Her yeni eklenen kelime için olumlu-olumsuz değerleri ve duygu yoğunlukları düzenlenmiştir. Duygu kütüphanesine ayrıca basit emojiler de eklenmiştir. Analiz aşamasında her bir paylaşım içinde yer alan kelimeler ayrıştırılarak duygu değeri saptanmıştır. Geliştirilen duygu kütüphanesinin verimli çalıştığını sınavabilmek için iki farklı analiz çalışmaları yapılmıştır. Ek olarak, veri sayısına göre analiz sonuçlarında farklılık olup olmadığına dair incelemeler yapılmıştır. Analiz çalışmalarında iki ayrı veri kümesi kullanılmış olup bu veri kümeleri hiçbir şekilde birleştirilmemiştir. Hava durumu için ayrı (hava durumu ve mevsimler), ve survivor televizyon programı için ayrı veri kümesi kullanılmıştır. Araştırma kapsamında, 2016 Eylül-2017 Kasım ayları arasında veri toplama işlemleri yapılmıştır.

4.1. Hava Değişikliği ve Mevsimlerin Kullanıcılar Üzerindeki Etkisinin Analizi

Hava durumunun mevsimsel olarak Twitter kullanıcıları üzerindeki etkilerini ölçmek için analiz çalışmaları yapılmıştır. Bu çalışmalar yaklaşık 1 yıllık zaman dilimi içinde toplanan veriler üzerinde yapılmıştır. Yapılan analiz çalışmalarında kullanılan etiketler: "ilkbahar, yaz, sonbahar, kış, spring, summer, fall, winter" dir. Hava değişikliklerinin Türkçe tweet paylaşan Twitter kullanıcıları üzerindeki etkisini gözlemleyebilmek için toplamda 8788 adet paylaşım içeren veri kümesi üzerinden 4 mevsim için 4 deney yapılmıştır.

- "Yaz" ve "summer" etiketleriyle toplanan verilerde (2017 Ağustos – Eylül - Ekim) yapılan çalışmalarda kullanıcıların çok fazla resim paylaşması nedeniyle genel olarak nötr oldukları görülmektedir. Fakat olumsuz olma oranı olumlu oranına göre düşüktür.
- "Sonbahar" ve "fall" etiketleriyle toplanan verilerde (2017 Ağustos – Eylül - Ekim) yapılan çalışmalarda kullanıcıların olumlu duygu içeren paylaşımlar yaptıkları görülmektedir.
- "İlkbahar" ve "spring" etiketleriyle toplanan verilerde (2017 Temmuz-Ağustos-Eylül) yapılan çalışmalarda kullanıcıların daha çok olumlu ve nötr paylaşımlar yaptıkları görülmüştür.
- "Kış" ve "winter" etiketleriyle toplanan verilerde (2017 Eylül-Ekim-Kasım) yapılan çalışmalarda kullanıcıların daha çok nötr ve olumsuz paylaşımlar yaptıkları gözlemlenmiştir.

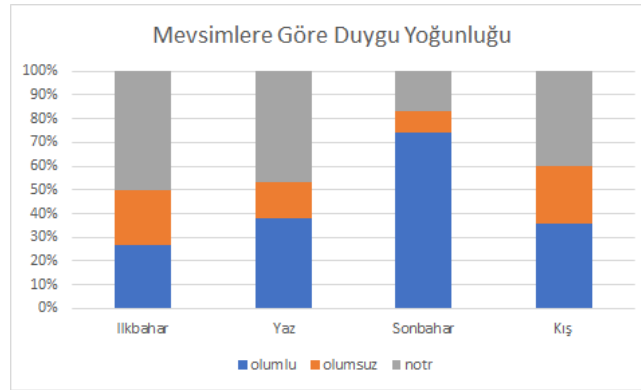
Bu dört deney sonrasında ise kullanıcıların genel olarak kış mevsiminde paylaşımlarının daha karamsar oldukları tespit edilmiştir. Fakat yaz aylarının gelmesiyle kullanıcıların paylaşımlarının daha olumlu olduğu Şekil 1’de gösterildiği gibi gözlemlenmiştir. Ayrıca sonbahar aylarında ise kullanıcıların halen olumlu olarak paylaşım yaptıkları tespit edilmiştir. Kullanıcılar yazın özlemlerini sıklıkla bahsetmektedir. Kullanıcılar bu özlemlerini belirtirken olumlu kelimeler seçmişlerdir. Bu nedenle toplanan verilerden olumlu ve nötr ağırlıklı bir duygu analizi sonucu çıkmıştır. Duygu analizi sonrasında karşılaşılan sonuçların beklenen değerlerle uyumlu olduğu gözlemlenmiştir.



Şekil 1 Duygu Analizi Algoritma İşleyişi

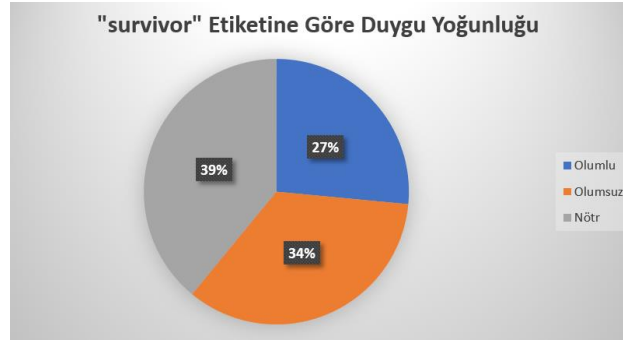
4.2. Survivor İsimli TV Programı ile İlgili Paylaşım Yapan Kullanıcıların Duygu Analizi

Bu program Türk televizyonlarında yer alan dünyaca ünlü yarışma programıdır. Popüler konu listesinde günlerce ilk sırada yer almıştır. Bu nedenle “#survivor” etiketiyle yapılan Twitter paylaşımları arşivlenmiştir. Daha sonrasında ise veriler üzerinde elle kontroller yapılmıştır. Bozuk Türkçe kullanım alışkanlıklarına uygun örnek olması nedeniyle bu araştırma kapsamına alınmıştır. Duygu analiz çalışması sonrasında survivor yarışmasıyla ilgili Twitter üzerinde Türkçe yapılan paylaşımların %27’sinin olumlu, %34’ünün olumsuz olduğu görülmüştür. Ayrıca, geriye kalan %39’luk kısım içerisinde bulunan paylaşımlar büyük oranda bozuk Türkçe kullanımı ve görsel paylaşım içermeleri nedeniyle nötr olarak kaydedilmiştir.



Şekil 2. Mevsimlere Göre Duygu Yoğunluğu

Yapılan bu analiz bulguları Şekil 2’de yer almaktadır. Bu analiz için toplam 766 paylaşım incelenmiştir. Sonuç olarak Türkçe paylaşım yapan kullanıcıların Survivor gibi bir eğlence programı için bile genel olarak olumsuz duyguları daha çok ifade ettiği gözlemlenmiştir. Beklenen sonuç ise çok sayıda kişinin takip ettiği bir eğlence programının olumlu duygu değerine sahip sonuçlar üretmesiydi.



Şekil 3. #survivor Etiketine Göre Yapılan Duygu Yoğunluk Dağılımı

5. Değerlendirme

Bu çalışmada duygu kütüphanesi kullanılarak sözlüksel benzeşim gruplama yöntemleri aracılığıyla duygu analizi yapılmıştır. Duygu analizleri, kelime seviyesinde ve cümle seviyesinde yapılmıştır. Ayrıca, bu çalışma kapsamında yeni bir Türkçe duygu kütüphanesi oluşturulmuştur. Duygu kütüphanesi birçok kelime, kelime grubu ve basit emoji karakterlerinden oluşmaktadır. Bunun yanında, her kelimenin İngilizce klavye kullanılması durumunda oluşacak versiyonları da kütüphaneye eklenmiştir. Türkçe paylaşım yapan kullanıcıların İngilizce klavye kullanma alışkanlıklarına sahip oldukları gözlemlenmiştir. Bu nedenle, İngilizce klavye ile Türkçe kelimelerin yazım şekilleri düzenlenmiştir.

Ayrıca, paylaşımlar üzerinden yapılan genel puanlama esnasında argo ya da küfür içeren kelime kullanımları puanlama sisteminde yüksek negatif (-), mükemmel olumlu değerler içeren kelime kullanımı yüksek pozitif (+) etki yapacak şekilde ağırlıklandırılmıştır. Yapılan analizlerde kullanıcıların hem argo ve hem de olumlu kelimeleri bir arada kullanmasından kaynaklanan anlam karmaşası engellenmeye çalışılmıştır. Fakat bazı durumlarda bu durum engellenememiştir. Örneğin; bir Twitter kullanıcısı tarafından paylaşılan “Nisan’ın 8’i olmuş, bahar gelmiş, sevgi, ask, kardeşlik geleceğine yine terör geliyor. Bu dünyayı bu hale getirenler yok olsun.” paylaşımında yer alan duygu içeren birden fazla olumlu ve olumsuz kelime yer almaktadır. Bu kelimeler “ask”, “sevgi”, “terör”dür. Bu örnekte yer alan metin olumsuz anlam ifade etmektedir fakat analiz sonucunda olumlu olarak işaretlenmiştir. Bunun sebebi metnin içinde yer alan “aşk” ve “sevgi” kelimelerinin duygu yoğunluk değerlerinin toplamı “terör” kelimesinin duygu yoğunluk değerinden büyük olmasıdır. Bu gibi durumlar Türkçe’nin esnek kullanım yapısından kaynaklanmaktadır.

Esnek bir dil olan Türkçe’nin duygu analizi çalışmaları sırasında karşılaşılan sorunları yönetebilmek için birçok algoritma geliştirilmiştir. Bu çalışmada genel olarak Türk kullanıcıların oluşturduğu paylaşım analiz çalışmaları esnasında karşılaşılan sorunlar aşağıdaki gibi listelenmektedir:

- Kullanıcıların büyük çoğunluğunun Türkçe klavye kullanmamalarından kaynaklanan birçok anlam bozukluğu ve eksikliğiyle karşılaşılmıştır. Kullanıcıların Türkçe metinler yazdıklarını düşünüp Türkçe olmayan kelimelerle paylaşım yapmaları nedeniyle birçok hibrit kelimelerin oluştuğu tespit edilmiştir. Bu sebeple iyileştirme yapma gereksinimi ortaya çıkmıştır.
- Yapılan araştırmalarda genelde negatif sonuçlar elde edilmiştir.
- Türk kullanıcıların az kelime kullanarak çok fazla konuyu ifade etme durumu ile karşılaşılmıştır.
- Türkçe’de yer alan ikilemelerden dolayı oluşan hesaplama işlemlerindeki hataları yönetebilmek içinde algoritmalar geliştirilmiştir.
- Twitter’da yer alan verilerin kirliliği olduğu ve Türkçe paylaşımlarda yüksek oranda küfür içerikli cümlelerin olduğu gözlemlenmiştir.

Ayrıca, toplanan verilerin birçoğu duygu içermemekte ya da eksik harflerle paylaşıldığından analiz çalışmalarında anlamlı verinin toplam veriye oranı düşüktür. Bu nedenle veri toplama ve anlamlandırma aşamaları çok uzun zaman almıştır. Hali hazırda Türkçe için istikrarlı olarak DDİ veri tabanı olmaması nedeniyle Türkçe duygu kütüphanesi geliştirilmiştir.

Geliştirilen Türkçe duygu kütüphanesinin performansı analizler için toplanan paylaşımlar üzerinde incelenmiş olup gereksinim durumlarında kütüphaneye yeni eklemeler yapılmıştır. Kütüphanenin kalitesi ve performansını ölçmek amacıyla analiz çalışmaları yapılmıştır. Bu analizlerden birincisi olarak hava durumu ve mevsimlerin geçişlerinin Twitter üzerinden Türkçe paylaşım yapan kullanıcıların paylaşımları üzerinden duygu analizi çalışmaları gerçekleştirilmiştir. Duygu kütüphanesini kullanarak elde ettiğimiz sonuçlar beklendiği üzere kış aylarında insanların hava ile ilgili daha olumsuz tepkiler verdiklerini ve daha karamsar olduklarını göstermiştir. Ayrıca mevsim geçişlerinde kullanıcıların duygularında olan değişiklikleri ölçebilmek için yine aynı duygu kütüphanesi üzerinden çalışmalar yapılmıştır. Bu sayede sonbahar mevsiminde kullanıcıların diğer mevsimlere göre daha fazla duygu aktaran paylaşım yaptıkları görülmüştür. Bu analiz hava durumunun kullanıcıları etkilediği tespit edilip beklentiyle örtüşmüştür. Yapılan incelemeler, geliştirilen Türkçe duygu kütüphanesinde belirtilen kelimelerin duygu değerlerinin gerçek metinlerdeki duygu ifadeleri ile örtüştüğü tespit edilmiştir.

Bu analizlerden ikincisi olarak ise televizyon reytinglerde üst sıralarda olan Survivor adlı program hakkında yapılan Twitter paylaşımları ile ilgili duygu analizi çalışması yapılmıştır. Bu çalışmayla Twitter üzerinden Türkçe paylaşım yapan kişilerin genelde olumsuz yorum yapma ve bozuk Türkçe kullanımı alışkanlıklarını tespit etmek amaçlanmıştır. Analiz sonucunda “olumsuz” ve “nötr” olarak etiketlenen paylaşımların ağırlıklı oldukları bulunmuştur. Sonuçlarda en büyük ağırlığa sahip nötr kategorisindeki paylaşımlar incelendiğinde çok fazla sayıda görsel paylaşıldığı veya Türkçe karşılığı olmayan kelimeler kullanıldığı gözlemlenmiştir.

5.1. Kısıtlamalar ve Eksiklikler

Bu çalışma dâhilinde geliştirilen Türkçe duygu kütüphanesinin iyi sonuçlar verdiği belirlenmiştir. Bununla birlikte, bütün Türkçe kelimeleri içermediği bilinmelidir ve zenginleştirmelere açıktır. Kütüphanenin zenginleşmesi sonuçların daha iyi olmasını sağlayacaktır.

Günümüz teknolojisinde DDİ algoritmalarıyla oluşturulmuş istikrarlı bir Türkçe duygu kütüphanesinin bulunmaması nedeniyle DDİ çalışmaları çok kısıtlı kalmıştır. Gelecekte gerçekleştirilecek kapsamlı ve başarılı bir Türkçe DDİ kütüphanesi, şüphesiz Türkçe metinler üzerinde yapılan metin madenciliği ve duygu analizi çalışmalarına katkı sunacaktır.

6. Sonuç

Bu çalışmada Twitter üzerinden belli etiketler ile toplanan Türkçe veriler üzerinde duygu analizi çalışması yapılmıştır. Duygu analiz yöntemi olarak varlık-görüş seviyesi ve cümle seviyesi kullanılmıştır. Ayrıca duygu kütüphanesi geliştirilmiştir. Duygu kütüphanesine duygu yönelim gruplamaları ve duygu yoğunluğu parametreleri eklenmiştir. Böylelikle, cümle seviyesinde yapılan analiz çalışmalarında oluşan anlam kayıplarından çıkan hatalar en aza indirilmiştir. Ek olarak, duygu kütüphanesine eklenen emojiler de duygu analizini destekleyecek şekilde puanlandırılmıştır. Böylelikle emojilerin duygu etkileri de kelimeler gibi dikkate alınmıştır. Ayrıca Türkçe’de yer alan kelime grupları ve ikilemeler de kapsama dâhil edilmiştir ve duygu yoğunluk değeri olarak daha büyük puanlar verilerek etkime oranı artırılmıştır.

Türkçe paylaşımlarda yaşanan sorunları çözmek için kelimelerden Türkçe klavyede yer alan ve İngilizce klavyede yer almayan harflerin yerine neler kullanılabileceği belirlenerek bu kütüphane zenginleştirilmiştir. Bu sayede Twitter kullanıcılarının Türkçe metinleri İngilizce alfabe kullanarak yazmaları sonrasında oluşan paylaşımları da analiz edebilme kabiliyeti kazandırılmıştır. Ayrıca büyük, küçük harf karmaşasından oluşabilecek sorunlar içinde çözümler üretilmiştir.

Bu çalışmada iki deney yapılmıştır. Deneylerden ilkinde, kullanıcıların paylaşımlarının hava durumuna göre değişen duygularını ölçme amacıyla analiz çalışmaları yapılmıştır. Bu araştırmalar kapsamında kış ve bahar-yaz için veriler toplanmıştır. Bu çalışmalar sonrasında paylaşımlardaki duygu düzeyinin hava durumuna göre değişiklik gösterdiği sonucu ortaya çıkmıştır. Kış aylarında olumsuz, yaz ise pozitif paylaşım oranlarının arttığı gözlemlenmiştir. Mevsimlerin değişmesiyle kullanıcıların duygularının da değiştiği görülmüştür. Bu nedenle en fazla verinin Sonbahar ayında paylaşıldığı görüntülenmiştir. İkinci deneyde ise Türkçe paylaşım yapan Twitter kullanıcılarının bozuk dil kullanma, genelde olumsuz durumları daha çok ifade etme alışkanlıklarını ölçmek adına Survivor etiketi ile yapılan paylaşımlar incelenmiştir. Survivor izleyen kitlenin genel olarak olumsuz içerikli paylaşımlar yaptığı sonucu çıkartılmıştır.

Kaynakça

Akgul, M. (2011, July 19). Internet Yasakları, Bilgi Toplumu ve Demokrasi.

Aylin, Y. A., & Elif, S. (2017). Social Media in Social Organization. *European Scientific Journal*.

Ayvaz, S., & Shiha, M. O. (2018). A Scalable Streaming Big Data Architecture for Real-Time Sentiment Analysis. In *Proceedings of the 2018 2nd International Conference on Cloud and Big Data Computing* (pp. 47–51). ACM.

Christine B., W., & Girish J., G. (2007). Social networks in political campaigns: Facebook and the 2006 midterm elections. *Annual Meeting of the American Political Science Association*.

Christine B., W., & Girish Jeff, G. (2008). *What is a Social Network Worth? Facebook and Vote Share in the 2008 Presidential Primaries*. Boston, MA: Department of International Studies and Government Bentley College.

Cihan, C., Mustafa, E., Evren, K., H. Kaan, T., & Duygu, A. (n.d.). Sosyal Medyanın Politik Katılım ve Hareketlerdeki Rolü. *Bilkent Üniversitesi, Bilgisayar Teknolojisi ve Bilişim Sistemleri Bölümü, Ankara*.

Cluster, W. B., Dinh, H., & Cooper, M. (2010). Na #x00EF;ve Bayes and unsupervised artificial neural nets for Cancun tourism social media data analysis. In *2010 Second World Congress on Nature and Biologically Inspired Computing (NaBIC)* (pp. 158–163). <https://doi.org/10.1109/NABIC.2010.5716370>

- Demirci, S. (2014). *Emotion Analysis on Turkish Tweets*. Middle East Technical University.
- Dokoohaki, N., Zikou, F., Gillblad, D., & Matskin, M. (2015). Predicting Swedish Elections with Twitter: A Case for Stochastic Link Structure Analysis. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015* (pp. 1269–1276). New York, NY, USA: ACM. <https://doi.org/10.1145/2808797.2808915>
- Duncombe, C. (2011). The Twitter revolution? Social media, representation and crisis in Iran and Libya (pp. 1–12). Presented at the Australian Political Science Association Conference (APSA) 2011, Australian National University, School of Politics and International Relations. Retrieved from <https://espace.library.uq.edu.au/view/UQ:269526>
- Duwairi, R. M. (2015). Sentiment analysis for dialectical Arabic. In *2015 6th International Conference on Information and Communication Systems (ICICS)* (pp. 166–170). <https://doi.org/10.1109/IACS.2015.7103221>
- Evirgen, E. (2016). *Türkçe Tweetlerin Duygu Analizi*. Bahcesehir Universitesi.
- Eyup Sercan, A., Caner, E., & Banu, D. (2017, September 26). Twitter verileri ile duygu analiz. Retrieved September 26, 2017, from <http://docplayer.biz.tr/28952591-Pamukkale-universitesi-muhendislik-bilimleri-dergisi-pamukkale-university-journal-of-engineering-sciences.html>
- Ganapathibhotla, M., & Liu, B. (n.d.). Mining Opinions in Comparative Sentences (pp. 241–248). Presented at the Proceedings of the 22nd International Conference on Computational Linguistics - Volume 1 (COLING '08), Vol. 1. Association for Computational Linguistics, Stroudsburg, PA, USA: Department of Computer Science University of Illinois at Chicago.
- Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why We Twitter: Understanding Microblogging Usage and Communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis* (pp. 56–65). New York, NY, USA: ACM. <https://doi.org/10.1145/1348549.1348556>
- Kayahan, D., Sergin, A., & Banu, D. (n.d.). *Twitter ile TV Program Reytinglerinin Belirlenmesi*. Bilgisayar Mühendisliği Bölümü Yıldız Teknik Üniversitesi.
- Liu, B. (2010). Sentiment Analysis and Subjectivity. *Handbook of Natural Language Processing*, 2, 627–666.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1–167.
- Nassar, I. N., & Khamayseh, F. T. (2015). Constructing activity diagrams from Arabic user requirements using Natural Language Processing tool. In *2015 6th International Conference on Information and Communication Systems (ICICS)* (pp. 50–54). <https://doi.org/10.1109/IACS.2015.7103200>
- Nasukawa, T., & Yi, J. (2003). Sentiment Analysis: Capturing Favorability Using Natural Language Processing. In *Proceedings of the 2Nd International Conference on Knowledge Capture* (pp. 70–77). New York, NY, USA: ACM. <https://doi.org/10.1145/945645.945658>
- Ozturk, N., & Ayvaz, S. (2017). Sentiment Analysis on Twitter: A Text Mining Approach to the Syrian Refugee Crisis. *Telematics and Informatics*. *Bahcesehir University*. <https://doi.org/10.1016/j.tele.2017.10.006>
- Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *LREc* (Vol. 10, pp. 1320–1326).
- Pang, B., & Lee, L. (2005). Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales (pp. 115–124). Presented at the Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics, Association for Computational Linguistics. Retrieved from <https://scholar.google.com/citations?user=qCdLtl0AAAAJ>
- Pang, B., Lee, L., & others. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2), 1–135.
- Parycek, P. (2012). *CeDEM 12 Conference for E-Democracy and Open Government 3-4 May 2012 Danube-University Krems, Austria*. Edition-Donau-Univ. Krems.

- Setty, S., Jadi, R., Shaikh, S., Mattikalli, C., & Mudenagudi, U. (2014). Classification of facebook news feeds and sentiment analysis. In *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 18–23). <https://doi.org/10.1109/ICACCI.2014.6968447>
- Sevindi, B. I. (2013). *Türkçe metinlerde denetimli ve sözlük tabanlı duygu analizi yaklaşımlarının karşılaştırılması*. Gazi Üniversitesi Fen Bilimleri Enstitüsü. Retrieved from <http://acikarsiv.gazi.edu.tr/index.php?menu=2&secim=10&YayinBIK=9431>
- Shiha, M., & Ayvaz, S. (2017). The effects of emoji in sentiment analysis. *Int. J. Comput. Electr. Eng. (IJCEE.)*, 9(1), 360–369.
- Tang, L., & Liu, H. (2010). Toward Predicting Collective Behavior via Social Dimension Extraction. *IEEE Intelligent Systems*, 25(4), 19–25. <https://doi.org/10.1109/MIS.2010.36>
- Turney, P. D. (2002). Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. *arXiv:cs/0212032*. Retrieved from <http://arxiv.org/abs/cs/0212032>
- Turney, P. D., & Mohammad, S. (2011). *Crowdsourcing a Word–Emotion Association Lexicon*. Institute for Information Technology, National Research Council Canada. Ottawa, Ontario, Canada, K1A 0R6.
- Vargas, J. (2008, November 20). Obama raised half a billion online. *Washington Post*. Retrieved from [fromvoices.washingtonpost.com/44/2008/11/obama-raised-half-a-billion-on.html](http://voices.washingtonpost.com/44/2008/11/obama-raised-half-a-billion-on.html)
- Wollmer, M., Weninger, F., Knaup, T., Schuller, B., Sun, C., Sagae, K., & Morency, L.-P. (2013). YouTube Movie Reviews: Sentiment Analysis in an Audio-Visual Context. *IEEE Intelligent Systems*, 28(3), 46–53.