



Global Business Research Congress (GBRC), May 26-27, 2016, Istanbul, Turkey.

DETERMINATION OF TWITTER USERS SENTIMENT POLARITY TOWARD AIRLINE MARKET IN TURKEY: A CASE OF OPINION MINING

DOI: 10.17261/Pressacademia.2016118690

Bahri Baran Kocak¹, Inci Polat², Cem Burak Kocak³

¹ Anadolu Üniversitesi, bbkocak@anadolu.edu.tr

² Gaziantep Üniversitesi, ipolat@gantep.edu.tr

³ Anadolu Üniversitesi, c_b_k@anadolu.edu.tr

ABSTRACT

The identification of actual and potential customers opinions and sentiments before and after purchase shapes the services offered by airlines in the airline sector as well as in every sector. In this paper, a sentiment analysis is made by compiling Twitter users' comments related to air transport. Comments of users collected from API (Application Programming Interfaces) service provided by Twitter as with many social media applications and were taken from a Java based program on a regular basis between April-May 2016. Obtained 8672 user comments were decomposed as positive, neutr and negative tags. Tags are collected in a tag cloud and results are analysed with Machine Learning Method and standardized and normalized Kernel Polinoms in SMO algorithm.

Keywords: Air Transportation, Twitter, Opinion Mining, Machine Larning Methods, Kernel Polinoms

JEL Codes: L93, M15, M31

TWITTER KULLANICILARININ HAVAYOLU PAZARINA YÖNELİK DUYGU KUTUPLARININ BELİRLENMESİ: BİR FİKİR MADENCİLİĞİ ÖRNEĞİ

ÖZET

Her sektörde olduğu gibi havayolu sektöründe de halihazır ve potansiyel müşterilerin satın alma öncesi ve sonrası fikir ve duygularının tespit edilmesi, havayolu firmalarının gelecekte sunacağı hizmetleri de şekillendirmektedir. Bu çalışmada, Twitter kullanıcılarının havayolu ulaşımı ile ilgili yorumları derlenerek duygu analizi çalışması yapılmıştır. Kullanıcı yorumları, birçok sosyal medya uygulamasında olduğu gibi Twitter' ın da sunmuş olduğu API (Application Programing Interfaces-Uygulama Programlama Arayüzleri) hizmeti vasıtasıyla java tabanlı program kullanılarak Nisan-Mayıs 2016 tarihleri arasında alınmıştır. Elde edilen 8672 kullanıcı yorumu olumlu, nötr ve olumsuz etiketlerle ayrıştırılmıştır. Elde edilen etiketler etiket bulutunda toplanmış ve sonuçlar Makine Öğrenmesi Yöntemi ve SMO sınıflandırmasında standart ve normalize Kernel Polinomları ile analiz edilmiştir.

Anahtar Kelimeler: Hava taşımacılığı, twitter, fikir madenciliği, makine öğrenmesi yöntemleri, Kernel polinomları

JEL Kodları: L93, M15, M31

1. GİRİŞ

Küresel bir hamle olarak ortaya çıkan sosyal medya (Hutton ve Fosdick, 2011) bugün milyarlarca kullanıcının birçok konuda fikirlerini ve şikayetlerini beyan ettiği bir platform haline gelmiştir. İnsanların duygularını gerçek zamanlı ifade ettiği mikroblog siteleri ise gerek firmaların gerekse de kişilerin ürün ve hizmetler hakkındaki duyguları anlaması, tepki ve cevapların verilebilmesi (Agarwal vd., 2011) için geliştirilmiş sosyal medya uygulamalarıdır. Bu mikroblog uygulamalarından biri olan Twitter, toplumun gerçek zamanlı iletişim kurması ve bireylerin 140 karakterle kendilerini ifade edebilmesi için oluşturulmuş küresel bir platformdur. Son yıllarda 35 dilin üstünde hizmet veren bu platformda 288 milyonun üzerinde aylık aktif kullanıcı sayısı (Twitter Annual Report-2014) bulunmaktadır.

Kullanıcı sayısının fazlalığı, fikirlerin doğrudan ve gerçek zamanlı belirtilmesi; bu çalışmada Twitter gönderilerinin duygu sınıflandırılması problemi için kullanılmasına olanak sağlamıştır. Örneklem olarak Twitter kullanıcılarının alınmasının bir diğer sebebi ise bu sosyal ağ servisinin API' ler vasıtasıyla şirketlere ve özel kişilere veriye kolay bir biçimde ulaşma imkanı tanınmasıdır.

Bilindiği üzere sosyal ağlar büyük küçük tüm şirketlerin bütünleşik bir parçası haline gelmiş ve başarılarının bir anahtarı olmuştur. Müşteri hizmetlerinden satışa, insan kaynaklarından bilgi teknolojilerine dek işletmelerin her departmanında yaşayan, nefes alan ve her alana dokunan sosyal ağlar, ürünlerin pazarda tanıtımı, bilinirliğine ağızdan ağıza yayılımı ve bu yolla ekonomiye katkı için oldukça önemli araçlar olarak hayatımızda yerini almıştır (Qualman, 2010). Dolayısıyla sosyal ağlardan biri olan Twitter, pazarda yer alan ürünlerle ilgili kullanıcıların yazdığı yorumlara bakılarak ürüne karşı duyguların anlaşılmasında önemli ipuçları sunmaktadır. Ne var ki elde edilen verilerin (yorumların) işlenmemiş olması, büyüklüğü, hatalı yazım, kısaltma ve günlük konuşma dili gibi nedenlerle sosyal ağlar üzerinde çalışmak oldukça zordur. Binlerce hatta on binlerce verinin ham haliyle tek tek incelenmesi ve insan algısıyla anlaşılması güçtür (Meral ve Diri, 2014). Bu nedenle verilerin çeşitli programlar vasıtasıyla API sağlayıcılarından çekilmesi, doğal dil işleme yöntemleri ile işlenmesi ve analizinin doğru bir biçimde yapılması gerekir.

Bir mikroblog uygulaması olan Twitter' da yayınlanan mesajların duygu analizi kapsamında kullanımı üzerine yapılan çalışmalar bu bildirinin literatür kısmında özetlenecek; sonraki bölümde yöntem ve bulgular ortaya konularak sonuçlar değerlendirilecektir. Bildirinin son bölümünde ise gelecek çalışmalar için önerilerde bulunulacaktır.

2. LİTERATÜR TARAMASI

Literatürde birbirlerinin yerine kullanılabilen fikir, duygu ve öznellik kavramlarının bilgisayar tabanlı uygulamalarla çalışılması sonucu bu kavramların fikir madenciliği, duygu analizi ya da öznellik analizi şeklinde değiştiği görülmektedir (Pang ve Lee, 2008: 5). Fikir madenciliği konusu incelendiğinde bu konulara yönelik çalışmaların yakın bir geçmişe sahip olduğu söylenebilir. Fikir madenciliği konusuna öncülük eden ilk çalışma 1994 yılında Wiebe (1994) tarafından kaleme alınmıştır. "Hikayedeki Bakış Açısının İzlenmesi" başlıklı bu makalede; düşünce, algı ve duyguların anlaşılabilmesi için bir algoritma geliştirilmiştir. Fikir madenciliği kavramının öykü ya da roman gibi düz metinlerden ayrılıp Web ortamında kullanılması ise 2003 yılına dayanmaktadır. Dave ve ekibinin (2003) Budapeşte' de düzenlenen WWW (World Wide Web) konferansında ürünlere yönelik kullanıcı yorumlarının çeşitli sınıflandırma yöntemleriyle özellik çıkarımı ve skorlama yaptığı ve sonuçların karşılaştırıldığı çalışma, fikir madenciliğinin ilk örneklerinden biri olarak karşımıza çıkmaktadır.

Metinlerde duyguların tahmin edilmesi, geleceğe dönük yargılara ulaşılması ve Pazar duygusunun elde edilmesi amacıyla yapılan duygu analizi çalışmaları ise Das ve Chen (2001) ve Tong (2001) tarafından literatüre kazandırılmıştır (Pang ve Lee, 2008: 6).

Günümüzde önemi gitgide artan müşterinin anlaşılması, müşteri davranışı altında yatan psikolojik etkenlerin ortaya çıkarılması ve kitlelerin genel duygu yapılarının tahmin edilmesi amacıyla yapılan duygu analizi çalışmaları ülkemizde de popüler hale gelmektedir. Son dönemde ülkemizde internet siteleri, çeşitli bloglar ve sosyal ağlardaki veri setlerinin elde edilmesiyle yapılan bir kısım Türkçe duygu analizi çalışmaları ise Tablo 1' de gösterilmiştir.

Tablo 1: Türkiye' de son dönemde yapılmış duygu analizi çalışmaları

Yazar		
Çetin ve Amasyalı (2013)	Türkçe twitter gönderilerinden oluşan 2 veri kümesi üzerinde Terimlerin sınıf dağılımlarını da hesaba katan eğitici yöntemler çeşitli boyutlarda karşılaştırılmıştır.	Bu çalışmada metin sınıflandırma da kullanılan eğitici ve eğitici olmayan terim ağırlıklandırma yöntemlerinin iki adet Türkçe twitter gönderimi veri kümesi üzerinde karşılaştırmalı analizi yapılmıştır. Analiz sonucunda metin temsilinde karakter ngramları kelime köklerine göre daha başarılıdır. Terim ağırlıklandırma da eğitici yöntemlerin geleneksel eğitici yöntemlerden daha başarılı olduğu özellik sayısı sebebiyle daha hızlı ve kolay modellenebildiği görülmektedir.
Meral ve Diri (2014)	Algı analizi, kelime tabanlı yöntem, N-Gram Naive Bayes, Rastgele Orman ve Destek Vektör Makinesi gibi makine öğrenme yöntemleri.	8.321 adet twit etiketlenmiş, kelime tabanlı yöntem kullanıldığında %89,5'lik başarı alınmış N-gram yöntemi ile alan bağımlı verilerde %90 başarı vermesine rağmen, alan bağımsız verilerde aynı başarıyı gösterememiştir. Ayrıca bu çalışmada kelime tabanlı doğal dil işleme sürecinden geçirilen twitter verileri ile sınıflandırıcıları eğitilmiş %90 başarılı sınıflandırma performansı gösteren bir sistem geliştirilmiştir.
Nizam ve Akin (2014)	Naive Bayes (NB), Random Forest (RF), Sequential Minimal Optimization (SMO), Decision Tree (J48) ve 1- Nearest Neighbors (IB1)	İlki 2000, ikincisi 824 olmak üzere toplam 2824 veri seti kullanılmıştır. Analiz sonucunda SMO yönteminin %72.33 ile metin sınıflandırmasında en yüksek başarıyı elde ettiği gözlemlenmiştir.

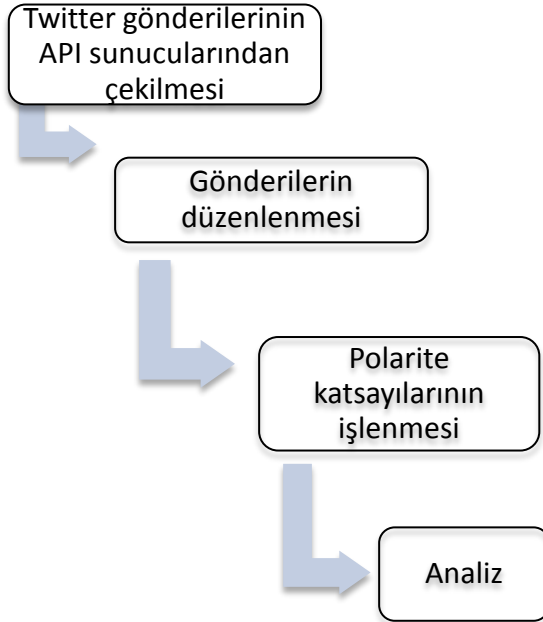
Tablo 1 (dvm) : Türkiye' de son dönemde yapılmış duygu analizi çalışmaları

Çoban vd. (2015)	SVM, Naive Bayes ve Multinom Naive Bayes, KNN	14777 adet mesajdan oluşan veri setinden elde edilen sonuçlara göre twitter mesajlarının makine öğrenmesi yöntemleriyle sınıflandırılabilceği tezi doğrulanmıştır. Veri setinin rastgele ve sadece his simgeleriyle oluşturulması, eğitim aşamasında bu simgelerin elenmesi sınıflandırma başarısını düşürmüştür. Ayrıca mesajların konu gözetilmeksizin rastgele çekilmesi, sınıflandırıcının genelleme yapma yeteneğini de sınırlamıştır.
Eliaçık ve Erdoğan (2015)	Finans topluluğuna ait duygu polarite değerleri ile Borsa İstanbul 100 endeks hareketleri arasında haftalık bazda korelasyon analizi, unigram , bigram ve PMI yöntemi ile özellik çıkarımı, veri kümesinin oluşturulmasında uzman görüşü ve destekçi vektör makinelerinin duygu sınıflandırmada kullanımı.	Haftalık borsa fiyat değişimleri ile haftalık finansal sosyal topluluk duygu polaritesi değişimleri arasında önceki yaklaşımlara göre daha yakın bir doğrusal ilişki elde edilmiştir. Duygu sınıflandırmasında başarı oranı % 73,63 bulunmuştur.

3. VERİ VE YÖNTEM

Türkiye' de havayolu ulaştırması ile ilgili Twitter üzerinden paylaşılan gönderilerin toplanma, işleme ve analiz süreci modeli Şekil 1' de gösterilmiştir.

Şekil 1: Twitter üzerinden paylaşılan gönderilerin toplanma, işleme ve analiz süreci modeli



•Havayolu ulaştırması ile ilgili gönderiler Twitter' ın API sunucularından java tabanlı program ile çekilmiştir.

•Tekrar eden ve havayolu ulaştırması ile ilgili alakasız gönderiler Şekil 2' de görüldüğü üzere temizlenmiş, imla ve yazım kurallarına göre yeniden düzenlenmiş ve argo ifadelerden arındırılmıştır.

•Gönderilerdeki olumlu, nötr ve olumsuz duygular Şekil 3' teki gibi 0, 1, 2 olarak ARFF dosyasına işlenmiştir.

•İşlenen metin ve duygu polariteleri Weka programında sınıflandırma algoritmaları arasında yer alan SMO ile analiz edilmiştir.

Şekil 2: Gönderilerde üzeri işaretlenmiş ifadelerin temizlenerek düzenlenmesi

@korkakkurbaga Karar verdim A firmasının hostesleri tirt 1.10 hostes mi olur lan!
RT @korkakkurbaga Karar verdim firmanın hostesleri tirt 1.10 hostes mi olur lan!

Şekil 3: Twitter gönderilerinin ARFF şablonunda gösterimi

```

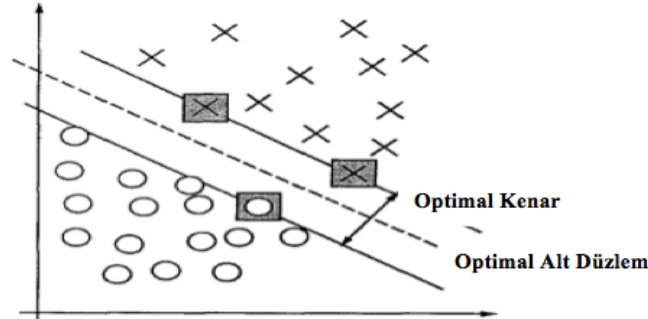
@relation Twitter_analizi
@attribute Tweetler string
@attribute Polarite {0,1,2}
@data
'Uçakta bekletildik. Konu size de ulaşmıştır. Şirketinizin hizmeti de personel de berbat. İlk ve son uçuşum oldu',2
'Yanıtınız için teşekkür ederim. İlk uçuşum 6 Mart tarihinde olacak. Bilet alındı sadece.',0
'Benim o tarihten 3 gün sonra İtalya uçuşum var. Bir de vize işi sorun, çok geç belli oluyor ve çıkmama ihtimali çok yüksek.',1
  
```

3.1. Destekçi Vektör Makineleri

İlk kez Cortes ve Vapnik (1995) tarafından öğrenen makinelerin yeni bir şekli olarak geliştirilen Destek Vektör Ağları, iki sınıflı problemlerde gruplar arasında sınır çizerek grupların özellik çıkarımı doğrultusunda bir alanda birbirlerinden ayrılmasını sağlamaktadır. Bu alanda ağın yüksek genelleme yeteneğine bağlı olan belirli özellikler

Şekil 4' te görüldüğü üzere doğrusal bir karar vermek için bir araya getirilir. Şekildeki her bir nokta ve çarpı işareti sistem girdilerinden hareketle çıkarılan özellikleri ifade etmektedir.

Şekil 4: İki boyutlu uzayda gri kutularla gösterilen destek vektörlerinin optimum düzeyde doğrularla iki sınıfa ayrımı



Sistem girdilerinin optimum düzeyde birbirinden ayrılabilmesi için destek vektörler adı verilen eğitim verilerinin öncelikle sisteme girilmesi gerekmektedir. Bu şekilde optimal kenarlar belirlenebilecektir. Dolayısıyla yapmış olduğumuz çalışmada sınıflandırmanın sağlıklı gerçekleştirilebilmesi için tvtillerin bir kısmı eğitim verisi olarak ayrılmıştır. Eğitim vektörlerinin hata payı ise şu şekilde hesaplanmaktadır:

$$E[\text{Pr}(\text{hata})] \leq E[\text{destek vektör sayısı}] / \text{eğitim vektör sayısı}$$

Doğrusal olarak ayrılabilen bir sınıflandırma probleminde ise eğitim verileri k sayıda örnek için $\{x_i, y_i\}$ $i=1, \dots, k$ kabul edilirse, optimum hiper-düzleme ait eşitsizlikler aşağıdaki gibi olur:

$$w \cdot x_i + b \geq +1 \text{ her } y = +1 \text{ için}$$

$$w \cdot x_i + b \leq -1 \text{ her } y = -1 \text{ için}$$

Burada x , N boyutlu bir uzaydaki tüm reel sayı kümelerinin bir elemanı, y ise $\{-1, +1\}$ kümesinin elemanıdır. Denklemden W , ağırlık vektörünü, b ise eğilim değerini göstermektedir (Osuna vd., 1997). Denklemden kümeler arasına çekilecek iki hiper-düzlemin belirlenmesi gerekmektedir. Bu düzlem ya da doğru $w \cdot x_i + b = \pm 1$ şeklinde ifade edilir.

3.2. Sequential Minimal Optimization (SMO)

Destekçi vektör makinelerini eğitmek için Platt (1998) tarafından oluşturulmuş SMO, ikinci dereceden programlama ile büyük problemleri hızlı ve analitik bir çözümle küçültmekte kullanılan bir algoritmadır. SMO için ön tanımlı değerlerde olan Kernel Polinomu'nda polinom derecesi d olan denklem ise;

$$K(x, y) = ((x \cdot y) + 1)^d$$

şeklinde dir. Bu doğrultuda normalleştirilmiş Polinom Kerneli;

$$K(x, y) = \frac{((x \cdot y) + 1)^d}{\sqrt{((x \cdot x) + 1)^d ((y \cdot y) + 1)^d}}$$

Şeklinde ifade edilmektedir (Çölkesen, 2010).

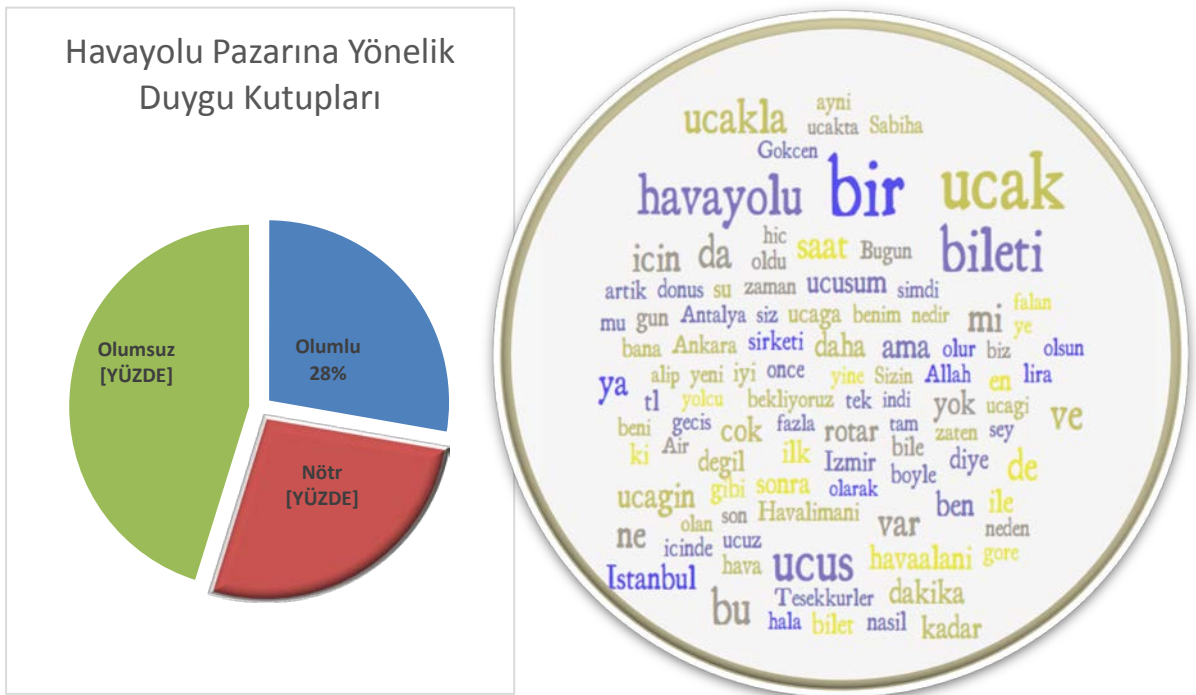
3.3. Veri Seti ve Özellikleri

Çalışmada Türkiye havayolu pazarına yönelik Twitter' da Nisan 2016-Mayıs 2016 tarihleri arasında paylaşılan 8672 gönderi veri seti olarak kullanılmış olup 6803 adet gönderi haber ve tekrarlardan oluştuğu için sınıflandırmaya dâhil edilmemiştir. Kalan 1879 adet veri seti el yordamıyla olumlu, olumsuz ve nötr olarak ayrıştırılmıştır. Ayrıştırma sonucunda elde edilen duygu kutupları ve etiket bulutu Şekil 5' te gösterilmiştir.

Ayrıştırma işlemi esnasında metinler Weka yazılımının gerektirdiği şekilde Türkçe karakterlerden arındırılmış ve ARFF dosyasına bu şekilde kaydedilmiştir. Veri setine dahil 5207 kelime özellik olarak kullanılmış ve analiz bu şekilde yapılmıştır.

Veri setinin %25' i eğitim, kalanı ise test verisi olacak şekilde makineye öğretilmiştir. Yapılan SMO analizi esnasında uygulanan standart ve normalize edilmiş Kernel Polinomu sonuçlarının karşılaştırıldığı Tablo 2 aşağıda gösterilmiştir. Analizlerden elde edilen sınıflandırma başarıları veri setindeki doğruluk payının toplam örnek sayısına oranı iken (Leo, 2001; Nizam ve Akın, 2014) Kappa İstatistiği ise +1 ile -1 arasında değişmekle birlikte sınıflar arasındaki gözlenen uyum ile şansa bağlı uyum arasındaki ilişkiyi vermektedir. Kappa İstatistiği' nin 1' e eşit olduğu durumda tam uyum söz konusu iken, 0' dan büyük olması durumunda gözlenen uyumun şansa bağlı uyumdan büyük ya da eşit olduğunu ifade etmektedir. Kappa İstatistiği' nin 0' dan küçük olması durumunda ise sınıflandırmanın güvenilir olmadığı anlaşılmaktadır (Aha ve Kibler, 1991; Nizam ve Akın, 2014).

Şekil 5: Türkiye' de havayolu pazarına yönelik değerlendirmelerin duyu kutupları ve bir etiket bulutu



Tablo 2: SMO sınıflandırma yönteminde standart ve normalize Kernel Polinomları' nın başarı ölçütlerinin karşılaştırılması

SMO Yöntemi	Sınıflandırma	Toplam Veri Sayısı	Sınıflandırılmış Veri Sayısı	Sınıflandırma Başarısı	Kappa İstatistiği
Standart Polinomu	Kernel	1406	812	%58	0,32
Normalize Polinomu	Kernel	1406	777	%55	0,23

4. SONUÇ

Ülkemizde havayolu ulaştırmasına yönelik hâlihazır ve potansiyel yolcuların fikirlerinin öğrenilmesi ve incelenmesi; havayolu pazarlamacıları açısından gün geçtikçe daha önemli bir hal almaktadır. Özellikle

internette havayolu firmalarına yönelik akan büyük bilginin veri ve metin madenciliği yöntemleriyle işlenmesi, modellenmesi ve analiz edilmesi içinde yaşadığımız bilgi çağında önem kazanmaktadır.

Türkiye’ de havayolu pazarına yönelik genel duygu haritasının bir kısmının çıkarılarak sınıflandırma algoritmalarından SMO yöntemi ile analiz edildiği bu bildiri standart ve normalize edilmiş Kernel Polinomları’ nın sınıflandırma başarıları karşılaştırılmıştır.

Duygu kutuplarına bakıldığında halihazır ve potansiyel yolcuların daha çok olumsuz yargılarda bulunması *ucuz-değil-hiç-rötar* kelimelerinin etiket bulutunda ön plana çıkmasıyla açıklanabilmektedir.

Analiz sonucu başarı ölçütleri incelendiğinde ise Standart Kernel Polinomu’ nun SMO algoritması içinde daha iyi bir performans gösterdiği; ancak genel olarak sınıflandırma başarısının düşük çıktığı görülmektedir. Bunun sebebi sınıflandırma verilerinin dengesiz dağılımı olması ve kelimelerin tamamının analizde kullanılmasıdır. Sınıflandırmanın yargısal yapılması bazı kelimelerin tüm sınıflarda geçmesine yani ayırt edici olmamasına sebep olmuştur. Ayrıca veri setinin dar olması da sınıflandırma başarısını olumsuz etkilemektedir. Kappa sayısının 0’ in üzerinde çıkması ise yapılan sınıflandırma gözleminin şansa dayalı olmadığını göstermektedir.

Gelecek çalışmaların diğer sektörlerde ve farklı sosyal medya mecralarında yapılması fikir madenciliği çalışmalarına katkıda bulunacağı aşikârdır.

KAYNAKLAR

- Agarwal, A., Xie, B., Vovsha, I., Rambow, O., Passonneau, R. Sentiment analysis of twitter data. In Proceedings of the ACL 2011 Workshop on Languages in Social Media (2011), pp. 30–38.
- Aha, D. ve Kibler, D. (1991), "Instance-based learning algorithms", Machine Learning, vol. 6, Issue no. 1, January 1991
- Cortes, C., Vapnik, V. (1995). Support-vector networks. Machine learning,20(3), 273-297.
- Çetin, M., Amasyalı, M. F. (2013). Eğitici ve Geleneksel Terim Ağırlıklandırma Yöntemleriyle Duygu Analizi. In *Proceedings of Signal Processing and Communications Applications Conference (SIU)*.
- Coban, O., Ozyer, B., Ozyer, G. T. (2015, May). Sentiment analysis for Turkish Twitter feeds. In *Signal Processing and Communications Applications Conference (SIU), 2015 23th* (pp. 2388-2391). IEEE.
- Çölkesen, İ. (2010). Destek Vektör Makineleri ile Uydu Görüntülerinin Sınıflandırılmasında Kernel Fonksiyonlarının Etkilerinin İncelenmesi. *Harita Dergisi*, 144, 73-82.
- Das, S., Chen, M. (2001, July). Yahoo! for Amazon: Extracting market sentiment from stock message boards. In Proceedings of the Asia Pacific finance association annual conference (APFA) (Vol. 35, p. 43).
- Dave, K., Lawrence, S., & Pennock, D. M. (2003, May). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In *Proceedings of the 12th international conference on World Wide Web* (pp. 519-528). ACM.
- Eliaçık, A. B., Erdoğan, N. (2015). Mikro Bloglardaki Finans Toplulukları için Kullanıcı Ağırlıklandırılmış Duygu Analizi Yöntemi.
- Hutton, G., & Fosdick, M. (2011). The globalization of social media. *Journal of Advertising Research*, 51, 564–570.
- Leo Breiman, *Machine Learning*, 45, 5–32, 2001: Random Forests
- Meral, M., Diri, B. (2014). Twitter Üzerinde Duygu Analizi. IEEE 22nd Signal Processing and Communications Applications Conference (SIU). 690-693.
- Nizam, H., Akin, S. S. (2014). Sosyal Medyada Makine Öğrenmesi ile Duygu Analizinde Dengeli ve Dengesiz Veri Setlerinin Performanslarının Karşılaştırılması. XIX. Türkiye'de İnternet Konferansı.
- Osuna, E., Freund, R., & Girosi, F. (1997). Support vector machines: Training and applications.
- Pang, B., Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, 2(1-2), 1-135.
- Platt, J. (1998). Sequential minimal optimization: A fast algorithm for training support vector machines.
- Qualman, E. (2010). *Socialnomics: How social media transforms the way we live and do business*. John Wiley & Sons.
- Tong, R. M. (2001). An operational system for detecting and tracking opinions in on-line discussion. In *Working Notes of the ACM SIGIR 2001 Workshop on Operational Text Classification* (Vol. 1, p. 6).
- Wiebe, J. M. (1994). Tracking point of view in narrative. *Computational Linguistics*, 20(2), 233-287.