

Çağrı Merkezi Metin Madenciliği Yaklaşımı

Call Center Text Mining Approach

İbrahim Onuralp Yiğit, Ahmet Feyzi Ateş
Türk Telekom Labs
Türk Telekom
İstanbul, Türkiye

Mehmet Güvercin, Hakan Ferhatosmanoğlu, Buğra Gedik
Bilgisayar Mühendisliği Bölümü
İhsan Doğramacı Bilkent Üniversitesi
Ankara, Türkiye

Özetçe —Günümüzde çağrı merkezlerindeki görüşme kayıtlarının sestem metne dönüştürülebilmesi görüşme kaydı metinleri üzerinde metin madenciliği yöntemlerinin uygulanmasını mümkün kılmaktadır. Bu çalışma kapsamında görüşme kaydı metinleri kullanarak görüşmenin içeriğinin duygu yönünden (olumlu/olumsuz) değerlendirilmesi, müşteri memnuniyetinin ve müşteri temsilcisi performansının ölçülmesi amaçlanmaktadır. Yapılan çalışmada görüşme kaydı metinlerinden metin madenciliği yöntemleri ile yeni özellikler çıkarılmıştır. Metinlerden elde edilen özelliklerden yararlanılarak sınıflandırma ve regresyon yöntemleriyle görüşme kayıtlarının içeriklerinin değerlendirilmesini sağlayacak tahmin modelleri oluşturulmuştur. Bu çalışma sonucunda ortaya çıkarılan tahmin modellerinin Türk Telekom bünyesindeki çağrı merkezlerinde kullanılması hedeflenmektedir.

Anahtar Kelimeler—Metin Madenciliği, Makine Öğrenmesi, Gözetimli Makine Öğrenmesi, Tahmin, Sınıflandırma, Regresyon

Abstract—Nowadays, the ability to convert call records from voice to text makes it possible to apply text mining methods to extract information from calls. In this study, it is aimed not only to evaluate the sentiment (positive/negative) of the calls in general, but also to measure the customer satisfaction and representative's performance by using call record texts. New features have been extracted from texts using text mining methods. Using the features extracted, prediction models were developed to evaluate the contents of call records by classification and regression methods. As a result of this study, it is planned to utilize the prediction models developed in Turk Telekom's call centers.

Keywords—Text Mining, Machine Learning, Supervised Machine Learning, Prediction, Classification, Regression

I. GİRİŞ

Şirketler çağrı merkezleri vasıtasıyla müşterilerin şikâyetleri, görüşleri ve isteklerini öğrenip bu doğrultuda ürünlerini ve/veya hizmetlerini geliştirmektedirler. Bu nedenle çağrı merkezleri görüşme kayıtlarından çıkarılan tespitler şirketlerin karar mekanizmaları için büyük önem taşımaktadır. Günümüzde çağrı merkezlerindeki görüşme kayıtlarının sestem metne dönüştürülüp saklanabilmesi müşterilerin yaşadıkları sorunlarla ilgili bilgilere daha kolay erişilebilmesini sağlamaktadır. Bu çalışma kapsamında müşteri memnuniyeti ve hizmet kalitesinin artırılması için metin madenciliği teknikleri kullanılarak Çağrı Merkezi Metin Madenciliği Yaklaşımı geliştirilmiştir.

Geliştirilen yaklaşım kullanılarak çağrı merkezi görüşme kaydı metinleri üzerinden görüşmenin içeriğinin duygu yönünden (olumlu/olumsuz) değerlendirilmesi, müşteri memnuniyetinin ve müşteri temsilcisinin performansının ölçülmesi amaçlanmaktadır.

Bildirinin bundan sonraki bölümleri şu şekilde düzenlenmiştir: İkinci bölümde yapılan literatür araştırması hakkında bilgi verilmiştir. Üçüncü bölümde geliştirilen Çağrı Merkezi Metin Madenciliği Yaklaşımı ile ilgili detaylar paylaşılmıştır. Dördüncü bölümde geliştirilen yaklaşım kullanılarak yapılan deneylere ve bu deneylerin sonuçlarına yer verilmiştir. Son bölümde yapılan çalışma sonucunda gelinen noktanın değerlendirilmesi yapılmış ve gelecek dönemde yapılacak çalışmalardan bahsedilmiştir.

II. LİTERATÜR ARAŞTIRMASI

Duygu analizi, metin içerisinden görüş, duygu ve tutum gibi öznel bilgilerinin çıkarılması için doğal dil işleme, metin madenciliği gibi yapay zekâ tekniklerinin kullanıldığı bir araştırma alanıdır [1]. Duygu analizi alanında makine öğrenmesi ve sözlük tabanlı yöntemlerle birçok akademik çalışma yapılmıştır. Makine öğrenmesi tabanlı yöntemler kullanılarak duygu analizi gerçekleştirilirken, duygu yönünden olumlu/olumsuz olarak etiketlenmiş veri kümesi üzerinde makine öğrenmesi algoritmaları uygulanarak sınıflandırma modeli oluşturulmaktadır. Ardından oluşturulan model yeni örneklerin sınıflandırılmasında kullanılmaktadır. Sözlük tabanlı yöntemler ise daha önceden oluşturulmuş duygu sözcükleri içeren bir duygu sözlüğüne dayalı olarak duygu analizini gerçekleştirmektedir. Metinde geçen sözcük ve cümlelerin anlamsal yönelimlerine dayalı bir hesaplama gerçekleştirilerek duygusal sınıflandırma yapılmaktadır.

Duygu analizi alanında makine öğrenmesine dayalı yöntemler kullanılarak yapılan temel çalışmalardan biri Pang ve diğerleri [2] tarafından gerçekleştirilmiştir. Bu çalışmada, çeşitli makine öğrenmesi yöntemleri (SVM, Naïve Bayes ve maksimum entropi) IMDB platformundaki film yorumlarını içeren veri kümesine uygulanarak film yorumları olumlu-olumsuz olarak sınıflandırılmıştır. Uygulanan yöntemlerin başarımları ve performansları kıyaslanmıştır. En yüksek doğru sınıflandırma oranının %82,9 ile destek vektör makinesi (SVM) yöntemi uygulandığında elde edildiği görülmüştür. Türkçe için benzer yüksek lisans tez çalışmaları gerçekleştirilmiştir. Eroğlu çalışmasında destek vektör makineleri yöntemini ve N-gram modelini kullanarak %85 başarı ile Beyazperde platformundaki film yorumları olumlu-olumsuz olarak sınıflandırılmıştır [3].

Akbaş ise çalışmasında Türkçe tweetler üzerinde kişiler ve başlıklar hakkındaki alt konuları çıkarıp bu konuları duygu kutbuna göre gruplayabilmiştir [8].

Vural ve diğerleri [4] çalışmalarında Türkçe film yorumlarını sözlük tabanlı yöntemle duygu analizi yapmışlardır. SentiStrength Kütüphanesini Türkçe'ye çevirerek film yorumlarını olumlu-olumsuz olarak sınıflandırmışlardır. Eroğlu'nun çalışmasında kullandığı Beyazperde platformundan alınan veri kümesine kendi yöntemlerini uygulayarak %76 başarı elde etmişlerdir.

Ergün ve arkadaşları [5] yaptıkları çalışmada, sözcük tabanlı yöntemlerden yararlanarak internetteki tüketici yorumları üzerinden duygu analizi yapmışlardır. Tüketici yorumlarının metinleri içerisinde ürün özelliğini gösteren kelimeler ve olumlu-olumsuz anlamlı sıfatlar belirlenmiştir. Ürünü niteleyen sıfatlar temsil ettikleri memnuniyet derecelerine göre puanlanmıştır. Niteleyen ve nitelenen kelimelerin tespiti için açık kaynak kodlu Zemberek Kütüphanesi [6] kullanılmış ve Türkçe dilbilgisine göre ağaç yapısı oluşturulmuştur. Ağaç yapısı üzerinde Derinlik Öncelikli Arama algoritması kullanılarak ürünün herhangi bir özelliğine ilişkin memnuniyet derecesini ifade eden sayısal bilgiler hesaplanmıştır.

Eliaçık ve Erdoğan [7] çalışmalarında makine öğrenmesi tabanlı bir yöntemle Twitter'daki Türkçe finans iletilerinin duygu kutbunu (pozitif, negatif) belirlemiştir. Çalışma kapsamında finans konusunda uzman kişiler tarafından 1501 negatif, 907 pozitif Türkçe kısa iletinin bulunduğu veri kümesi oluşturulmuştur. Özellik çıkartmak için unigram ve bigram, özellik seçmek için PMI (Point-wise Mutual Information) yöntemi kullanılmıştır. Özellik çıkarma ve seçme işlemlerinden sonra 35030 özellikten oluşan özellik kümesi elde edilmiştir. Bu çalışmada, destek vektör makinesi tabanlı bir duygu sınıflandırma yöntemi kullanılarak %73,63 başarı oranı ile duygu kutbu belirlenmiştir.

Literatürde yapılan benzer çalışmalar duygu kutbuna göre kategorize etme ve sınıflandırmayla ilgilidir. Bu çalışmada çağrı merkezlerindeki görüşmelerin duygu yönünden (olumlu/olumsuz) değerlendirilmesi yanında müşteri memnuniyeti ve temsilci performansının ölçülmesi de hedeflenmiştir.

III. ÇAĞRI MERKEZİ METİN MADENCİLİĞİ YAKLAŞIMI

Bu çalışma kapsamında ses kayıtlarına ilişkin sesten dönüştürülmüş metinlerin sınıflandırılması ve içeriklerinin değerlendirilmesi için Çağrı Merkezi Metin Madenciliği Yaklaşımı geliştirilmiştir¹. Bu yaklaşım kapsamında ses kaydı metinlerinin üzerinde metin madenciliği teknikleri uygulanarak metinlere ilişkin yeni özellikler elde edilmektedir. Metinlerden çıkarılan yeni özelliklerle beraber gözetimli makine öğrenmesi algoritmaları kullanılarak görüşmenin havasının, müşteri memnuniyetinin ve temsilci performansının tahmin edilmesi için sınıflandırma ve regresyon modelleri oluşturulmuştur.

Çağrı Merkezi Metin Madenciliği Yaklaşımı ile sınıflandırma modelleri oluşturulurken sırasıyla aşağıdaki adımlar takip edilmiştir.

¹Bu çalışmada kullanılan sesten metne dönüştürülmüş veriler Türk Telekom Grubu şirketlerinden AssisTT A.Ş. tarafından sağlanmıştır.

- 1) Ses kaydı metinlerinde bulunan ve doğrudan tahmin modellerinde kullanılacak meta özellikler belirlenmiştir. Tahmin modellerine girmeye uygun halde olmayan özellikler ön işlemden geçirilerek sayısal hale getirilmiştir.
- 2) Ses kaydı metinlerinden metin madenciliği yöntemleri yardımıyla işlenerek saklı olan metin tabanlı özellikler açığa çıkarılmıştır.
- 3) Metin tabanlı özellikler meta özelliklerle birleştirilerek kullanılacak özellik kümesi belirlenmiş ve nihai veri kümesi oluşturulmuştur.
- 4) Veri kümesi üzerinde sınıflandırma ve regresyon algoritmaları kullanılarak deneyler yapılmıştır.
- 5) Sınıflandırma algoritmaları karşılaştırılarak en başarılı sınıflandırma algoritması belirlenmiş ve görüşmenin havasının tahmini için kullanılacak sınıflandırma modeli oluşturulmuştur. Lineer regresyon algoritmasından yararlanarak müşteri memnuniyetinin ve temsilci performansının tahmin edilmesi için regresyon modeli oluşturulmuştur.

Bu bölümün alt başlıklarında sınıflandırma modellerine eklenen özelliklere ilişkin ayrıntılı bilgilere yer verilmiştir.

A. Ses Kayıtlarının Meta Özelliklerinin Eklenmesi

Veri kümesinde yer alan ses kayıtlarının meta özellikleri müşteriyle, müşteri temsilcisiyle veya doğrudan görüşmenin kendisiyle ilgili tanımlayıcı istatistiksel özelliklerdir. Ses kayıtlarının meta özellikleri sesin analizi sonucunda ortaya çıkmıştır. Çalışma kapsamında meta özelliklerin tamamı gözden geçirilmiş ve tahmin modellerinde kullanılmaya uygun olanlar belirlenmiştir. Görüşmenin analizi sonucunda çıkarılan meta özelliklere konuşma süresi, konuşmaların üst üste gelme süreleri, monotonluk, sınırlılık yerleri ve sınırlılık yüzdeleri örnek olarak verilebilir.

Veri kümesindeki bütün özellikler tahmin modellerinde kullanılması için uygun değildir. Özelliklerin bir kısmı sayısal değerler içermektedir ve doğrudan sınıflandırma modeline girecek durumdadırlar. Sayısal olmayan özellikler ön işleme geçirilerek sayısal hale getirilmiştir. Böylece sayısal olmayan özellikler de tahmin modellerine girebilecek duruma getirilmiştir.

B. Ses Kayıtlarının Metin Tabanlı Özelliklerinin Eklenmesi

Çalışma kapsamında veri kümesindeki meta özelliklerin yanı sıra konuşmanın müşteri ve temsilci metinlerinden çıkarılan özelliklerle birlikte özellik kümesi zenginleştirilmiştir ve tahmin doğruluğu artırılmaya çalışılmıştır. Metin tabanlı özellikler çıkarılırken üç farklı metin madenciliği tekniği uygulanmıştır. Uygulanan metin madenciliği tekniklerine ilişkin ayrıntılar aşağıda verilmiştir. Bu başlık altında verilen metin tabanlı özellikler örnek olmakla birlikte eldeki veriye göre daha farklı özellikler eklenebilecektir veya bu özelliklerden bazıları kullanılamayabilmektedir.

1) *Duygu analizi ile yeni özellikler elde etme:* Duygu analizi modelleri serbest metin verileri üzerinde oluşturulmaktadır ve metinlerin duygu yönünü tahmin etmede kullanılmaktadır. Bu modeller kullanılarak bir cümlenin duygu yönü pozitif/negatif ekseninde tahmin edilebilmekte ve bu tahminin kuvvet derecesini gösteren skorlar elde edilebilmektedir. Duygu

analizi modelleri gözetimli [9] [10] [11] veya gözetimsiz [11] [12] [13] sınıflandırma modelleri olarak oluşturulabilmektedir. Duygu analizi modelleri oluşturulduktan sonra ses kaydı metinlerindeki cümlelerin duygu yönü ölçülerek her ses kaydına ait çeşitli istatistikler çıkarılmaktadır. Çıkarılan bu istatistikler elde edilmeye çalışılan özellik kümesinde birer özellik olarak değerlendirilmektedir.

Ses kaydı metinlerindeki bütün cümleler duygu analizi modelinden geçirilmektedir ve bu modelin verdiği skorlar (0 ile 1 arasında) kullanılarak yeni özellikler hesaplanmaktadır. Yeni özellikler hem müşteri hem de temsilci metinleri için ayrı ayrı çıkarılmaktadır. Aşağıda duygu analizi ile elde edilen özellikler açıklamalarıyla birlikte verilmektedir.

- **Negatif/Pozitif yüzdesi:** Negatif/Pozitif olarak tahmin edilen cümlelerin sayısının metin içinde geçen toplam cümlelerin sayısına oranıyla hesaplanmaktadır.
- **Ortalama negatif/pozitif skoru:** Negatif/Pozitif olarak tahmin edilen cümlelerin skorlarının metin içinde geçen tüm cümleler bazında ortalaması alınarak hesaplanmaktadır.
- **Toplam negatif/pozitif skoru:** Negatif/Pozitif olarak tahmin edilen cümlelerin skorlarının toplamı alınarak hesaplanmaktadır.

2) Ayırt edici kelimeler ile yeni özellikler elde etme:

Ses kaydı metinleri için özellik kümesi oluşturma sürecinde kullanılan diğer bir metod ayırt edici kelimeleri tespit etme yöntemidir. Bu metodun uygulanabilmesi için ses kaydı metinlerinin olumlu/olumsuz olarak etiketlenmiş olması gerekmektedir. Etiketlenmiş ses kaydı metinleri üzerinde WLLR (Weighted Log Likelihood Ratio) [14] [15] tekniği uygulanarak her bir etiket sınıfı için ayırt edici kelimeler tespit edilebilir ve daha sonra bu kelimeler her sınıf için oluşturulacak tahmin modelleri için birer özellik olarak kullanılabilir.

3) Metin üzerinden önceden belirlenmiş kurallar ile yeni özellikler elde etme: Ses kaydı metinleri üzerinden önceden belirlenmiş kurallara göre yeni özellikler çıkarılmaktadır. Aşağıda tahmin modellerinde bulunan özellikler açıklamalarıyla birlikte verilmektedir.

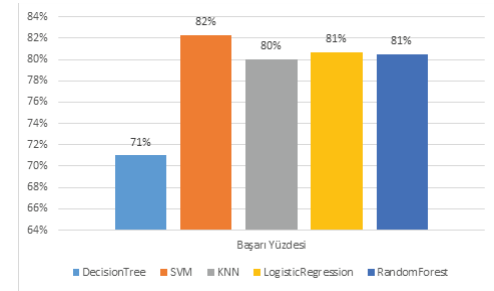
- **Kelime sayısı:** Ses dosyasındaki toplam kelime sayısını yansıtmaktadır.
- **Temsilcinin görüşmeyi olumlu sonlandırması:** Temsilcinin görüşmenin sonuna doğru müşterinin sorununa yardımcı olacak veya yönlendirici şekilde davranıp davranmadığına, son cümlelerde bazı kelimelerin söylenip söylenmediğine bakılır.
- **Müşterinin görüşmeyi olumlu sonlandırması:** Konuşmanın sonuna doğru müşterinin olumlu cümleler söyleyip söylemediğinin tespiti için kullanılır. Müşteri konuşmayı kapatırken naziklik içeren bazı kelimeleri söyleyip söylemediğine bakılır.
- **Görüşmenin olumlu sonlandırması:** Yukarıdaki iki özellik dikkate alınarak konuşmanın sonuna doğru hem müşterinin hem de temsilcinin davranışı ölçülür.
- **Negatiflik/Pozitiflik:** Konuşmadaki olumsuzluk/olumluluk sayısını belirten özelliktir. Önceden

belirlenen kötü/iyi kelime listesi kullanılarak belirlenir. Görüşmedeki herhangi bir cümlede kötü/iyi kelimelerden birisi geçiyorsa bu özelliğin değeri bir artırılır.

- **Kibarlık:** Temsilcinin müşteriye karşı ne kadar kibar davrandığına, kibarlık ifade eden kelimeleri kullanıp kullanmadığına bakılır.
- **Telefonun kapanması:** Müşterinin konuşmayı olağan bir biçimde sonlandırıp sonlandırmadığına bakılır.

IV. DENEYLER

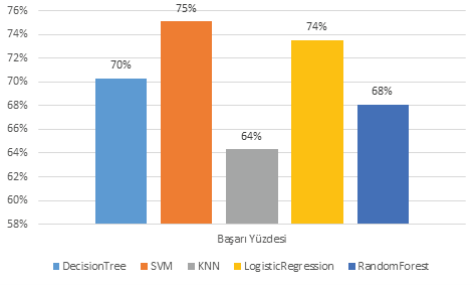
Tahmin modellerinin oluşturulması için ilk olarak ses kayıt dosyalarının etiketlemesi ve metin madenciliği teknikleri kullanılarak özellik vektörlerine dönüştürülmesi işlemleri tamamlanmıştır. Özellik vektörlerine dönüştürme sürecinde bir önceki bölümde anlatılan özellik çıkarma yöntemleri kullanılmıştır. Etiketleme işlemi ise görüşmenin havası olumlu/olumsuz, müşteri memnuniyeti ve temsilci performansı ise 1 ile 5 arasında olacak şekilde derecelendirerek tamamlanmıştır. Etiketleme çalışmaları sonucunda 400 tane ses kaydı olumlu/olumsuz olarak etiketlenmiştir. Bu örneklerden 92 tanesi olumsuz, kalan 308 tanesi ise olumlu olarak tespit edilmiştir. Bu aşamada iki farklı deney gerçekleştirilmiştir. Birinci deneyde Karar Ağacı (Decision Tree), Destek Vektör Makineleri (SVM), K En Yakın Komşu (KNN), Lojistik Regresyon (Logistic Regression), Rasgele Orman (Random Forest) algoritmaları denenerek görüşmenin havasını tespit etmek için en başarılı sınıflandırma algoritması belirlenmeye çalışılmıştır. Aşağıdaki şekilde denenen algoritmaların sınıflandırma başarılarıyla ilgili sonuçlar gösterilmektedir. Bu sonuçlara göre %82 doğruluk ile az bir farkla en başarılı sınıflandırma SVM algoritması olmuştur.



Şekil 1: Dengesiz veri kümesi başarı yüzdeleri

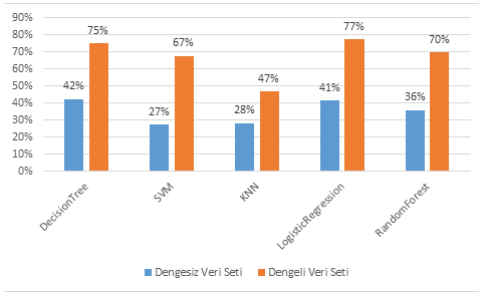
İkinci deneyde olumlu ve olumsuz sayısının dengeli/eşit olacak şekilde veri kümesi düzenlenmiştir. Veri kümesi düzenlenikten sonra sınıflandırma algoritmaları tekrar karşılaştırılmıştır. Karşılaştırma sonuçları incelendiğinde SVM ve Lojistik Regresyon metodlarının diğerlerine göre daha başarılı sonuçlar elde ettiği gözlemlenmiştir.

Şekil 3'deki grafikte algoritmaların olumsuz görüşmelerin tahminine ilişkin dengeli ve dengesiz veri kümeleriyle yapılan deneylerdeki başarı yüzdeleri yer almaktadır. Dengeli veri kümesi kullanılarak oluşturulan sınıflandırma modellerinde genel tahmin başarısı açısından düşüş görülsede bile olumsuz görüşmelerin tahmin etmedeki başarı yüzdesinde kayda değer artış olduğu görülmektedir. Olumsuz görüşmeleri sınıflandırmada



Şekil 2: Dengeli veri kümesi başarı yüzdeleri

en başarılı algoritmaların Karar Ağacı ve Lojistik Regresyon olduğu tespit edilmiştir. Ayrıca, tüm sınıflandırma algoritmalarının dengeli veri kümeleriyle beraber kullanıldığında olumsuz görüşmeleri tahmin etmedeki başarı yüzdeleri dengesiz veri kümelerinin kullanıma göre ciddi oranda artmıştır. Örneğin, genel başarı açısından en iyi algoritma olan SVM sonuçları incelendiğinde dengesiz veri kümesi kullanıldığı zaman olumsuz görüşmeleri doğru tahmin etme başarısı %27 iken dengeli veri kümesi ile yapılan deneyde başarının %67'e çıktığı gözlemlenmiştir.



Şekil 3: Olumsuz görüşme tahmininde başarı yüzdeleri

Etiketleme sürecinde konuşmanın içeriği dikkate alınarak 1 ile 5 arasında derecelendirilen müşteri memnuniyeti ve temsil performansı değişkenlerinin tahmini için Lineer Regresyon algoritması kullanılarak regresyon modelleri oluşturulmuştur. Müşteri memnuniyeti ve temsilci performansı değişkenlerinin bağımlı, özellik kümesinde yer alan diğer özelliklerin bağımsız değişken olduğu regresyon modellerinde ilk 320 örnek öğrenim kümesi için son 80 örnek ise test kümesi için kullanılmıştır. Müşteri memnuniyeti ve temsilci performansı değişkenleri için bu deneyle ilgili sonuçlar aşağıdaki tabloda verilmiştir.

TABLO I: REGRESYON MODELİ SONUÇLARI

Değişken	Varyans	Ort. Mutlak Hata	Ort. Karesel Hata
Müşteri Memnuniyeti	-0,06	0,68	0,66
Temsilci Performansı	-0,01	1,12	1,81

Regresyon modelinin sonuçları incelendiğinde müşteri memnuniyetinin ve temsilci performansının düşük varyans değeri ile tahmin edilebildiği görülmektedir. Ortalama mutlak hata ve ortalama karesel hata metriklerine bakıldığında regresyon modelinin müşteri memnuniyetini temsilci performansına göre daha başarılı bir şekilde ölçebildiği gözlemlenmektedir.

V. SONUÇ

Ses kaydı verisi günümüzde birçok kurum tarafından elde edilmekte ve sesten metne dönüştürülerek metin formatında saklanmaktadır. Bu çalışma kapsamında geliştirilen Çağrı Merkezi Metin Madenciliği Yaklaşımı ile görüşme kaydı metinlerinden özellik kümesi oluşturulmuş ve makine öğrenmesi algoritmaları kullanılarak sınıflandırma ve regresyon modelleri geliştirilmiştir. Bu çalışma sonucunda ortaya çıkarılan tahmin modellerini kullanarak Türk Telekom Çağrı Merkezine gelen çağrıların görüşmenin havasının olumlu/olumsuz olması, müşteri memnuniyeti ve temsilci performansı açısından değerlendirilmesi hedeflenmiştir. Yaptığımız araştırmalar sonucunda çağrı merkezlerindeki görüşme kaydı metinlerin işlenmesi için piyasada kullanılan bir ürünün varlığı tespit edilememiştir. Çalışmalar neticesinde Türk Telekom bünyesinde geliştirilecek Ar-Ge proje sonunda kapsamlı bir ürün çıkarılması planlanmaktadır. Ayrıca çağrı merkezi alanında geliştirilecek ürünün farklı alanlarda uygulamaları da değerlendirilecektir.

KAYNAKLAR

- [1] A. Onan, S. Korukoğlu, "Makine öğrenmesi yöntemlerinin görüş madenciliğinde kullanılması üzerine bir literatür araştırması", Pamukkale Univ Muh Bilim Derg, cilt 22, no. 2, pp. 111-122, 2016.
- [2] B. Pang, L. Lee, S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques", Proceedings of EMNLP, 2002.
- [3] U. Eroğlu, "Sentiment Analysis in Turkish", The Graduate School of Natural and Applied Sciences of Middle East Technical University, Ankara, 2009.
- [4] A. G. Vural, B. B. Cambazoglu, P. Senkul, Z. O. Tokgoz, "A Framework for Sentiment Analysis in Turkish: Application to Polarity Detection of Movie Reviews in Turkish", Computer and Information Sciences III, Springer, , pp. 437-445, 2012.
- [5] K. Ergün, C. Kubat, G. Çağıl, R. Cesur, "İnternet ortamındaki tüketici yorumlarından özet bilgi çıkarımı", SAÜ. Fen Bil. Der., cilt 17, no. 1, pp. 33-40, 2013.
- [6] M. D. Akın, A. A. Akın, "Türk Dilleri İçin Açık Kaynaklı Doğal Dil İşleme Kütüphanesi: Zemberek," Elektrik mühendisliği, cilt 431, pp. 38-44, 2007.
- [7] A. B. Elicaçık, N. Erdoğan, "Mikro Bloglardaki Finans Toplulukları için Kullanıcı Ağırlıklandırılmış Duygu Analizi Yöntemi", UYMS, İzmir, 2015.
- [8] E. Akbaş, "Aspect Based Opinion Mining on Turkish Tweets", The Graduate School of Engineering and Science of Bilkent University, Ankara, 2012.
- [9] A. Go, R. Bhayani, L. Huang, "Twitter sentiment classification using distant supervision", CS224N Project Report, Stanford, 2009.
- [10] A. L. Maas, "Learning word vectors for sentiment analysis", Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1, Association for Computational Linguistics, 2011.
- [11] B. Pang, L. Lee, "Opinion mining and sentiment analysis", Foundations and trends in information retrieval 2.1-2, 1-135, 2008.
- [12] X. Hu, "Unsupervised sentiment analysis with emotional signals", Proceedings of the 22nd international conference on World Wide Web, International World Wide Web Conferences Steering Committee, 2013.
- [13] P. D. Turney, "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews.", Proceedings of the 40th annual meeting on association for computational linguistics, Association for Computational Linguistics, 2002.
- [14] V. Ng, D. Sajib, S. M. Arifin, "Examining the role of linguistic knowledge sources in the automatic identification and classification of reviews", Proceedings of the COLING/ACL on Main conference poster sessions, Association for Computational Linguistics, 2006.
- [15] K. Nigam, "Text classification from labeled and unlabeled documents using EM", Machine learning 39:2-3, 103-134, 2000.