

BERT Modeli ile Türkçe Duygu Analizi

Turkish Sentiment Analysis Using BERT

Utku Umur Acikalin, Benan Bardak ve Mucahid Kutlu
Bilgisayar Mühendisliği Bölümü
TOBB Ekonomi ve Teknoloji Üniversitesi, Ankara, Türkiye
{u.acikalin,benan.bardak,m.kutlu@etu.edu.tr}

Özetçe —Duygu analizi, oldukça popüler bir araştırma alanı olmakla birlikte, araştırmaların büyük bir kısmı İngilizce üzerine yapılmış olup, Türkçe üzerine yapılan çalışmalar kısıtlı sayıdadır. Türkçe için diğer dillere nazaran daha kısıtlı olan doğal dil işleme kaynakları da bu konuda yapılan araştırmalar için büyük bir engel teşkil etmektedir. Bu kısıtlamaları aşmak için, bu çalışmada iki türlü yaklaşım önerilmiştir: 1) BERT’in çok dilli modelinin Türkçe duygu analizi için uyarlanması ve 2) Türkçe metinlerin önce İngilizceye otomatik tercümesi yapıp BERT’in ana modelinin kullanılması. Film ve otel yorumlarından oluşan, pozitif ve negatif olmak üzere 2 etikete sahip Türkçe veri kümeleri üzerinde yapılan deneylerde yüksek başarımlar elde edilmiştir. Film veri kümesinde varolan çalışmalardan daha yüksek başarımlar elde edilmiştir.

Anahtar Kelimeler—BERT, Türkçe Duygu Analizi

Abstract—While sentiment analysis is a popular research area, most of the research has been conducted for English and the number of studies for Turkish are rather limited. Limited resources for Turkish natural language processing (NLP) is one of the major challenges for Turkish NLP research. In order to overcome these limitations, we propose two approaches for Turkish sentiment analysis: 1) fine tuning multilingual model of BERT 2) using main model of BERT after machine translation of Turkish texts into English. We conducted experiments on Turkish movie and hotel review datasets where each review is labeled either positive or negative. Our methods achieve high accuracy scores such that in the movie dataset, our BERT models outperform existing methods.

Keywords—BERT, Turkish Sentiment Analysis

I. GİRİŞ

Son yıllarda bilgi teknolojilerinin hızlı gelişmesi ile beraber bireylerin internet kullanımı artmıştır. Böylelikle insanlar internet aracılığıyla uygulamalar üzerinden çeşitli konular hakkındaki fikirlerini daha kolay belirtme imkanı bulmuşlardır. Örneğin, internette alışveriş yaparken almayı düşündüğümüz bir ürün ya da izlemek istediğimiz bir film hakkında diğer insanların yorumları bizim için çok önemli bir bilgi kaynağıdır. Bu kullanıcı yorumları pazarlama, hizmet ve reklamcılık firmaları açısından da son derece önemli olup, firmaların satış politikalarını etkileme potansiyeline sahiptir. Ayrıca insanların sosyal medya üzerinden kurumlar ya da kişiler hakkındaki yorumları da toplumsal olaylar üzerinde araştırma yapanlar için çok önemli bir veri kaynağıdır. İnternet kullanıcılarının çeşitli konular hakkındaki yorumlarının pozitif mi negatif mi olduğunun otomatik tespit edilebilmesi bu önemli veri kaynağının etkin bir şekilde kullanılması için kilit bir rol oynamaktadır.

Duygu analizi alanında gerçekleştirilen çalışmalar çoğunlukla İngilizce metinler üzerinde gerçekleştirilmiştir. Fakat, önerilen yöntemlerin büyük çoğunluğu belirli bir dile özgü özellikler kullanılarak geliştirildiği için, bu yöntemlerin başka bir dil için uyarlanması oldukça zordur. Türkçe duygu analizi konusunda yapılan çalışmalar İngilizceye nazaran oldukça kısıtlı olup, bu konuda araştırma yapmak isteyen araştırmacılar için birçok zorluk bulunmaktadır. Örneğin, Türkçenin zengin morfolojik yapısı ve cümlelerdeki serbest kelime sırası metinlerin anlamsal analizini zorlaştırmaktadır. Ayrıca bir metnin analizi için gereken doğal dil işleme kütüphanelerinin diğer dillere göre Türkçede çok daha az olması araştırmacılar önündeki büyük engellerden biridir.

Bu çalışmada, Türkçe dilinde yazılmış farklı alanlara ait kullanıcı yorumlarını içeren veriler üzerinde duygu analizi çalışması gerçekleştirilmiştir. Çalışmamızda Google tarafından geliştirilmiş ve birçok doğal dil işleme alanında en yüksek başarımların elde edilmesini sağlayan dönüştürücü (transformer) mimarisini kullanan BERT [5] modeli tabanlı iki yöntem önerilmiştir. İlk modelimizde BERT’in çok dilli versiyonu [14] uyarlanmıştır. İkinci yöntemimizde ise, Türkçe metinler makine-tercümesi kütüphaneleri kullanılarak önce İngilizceye çevrilmiş sonra da BERT’in sadece İngilizce için geliştirilen ana modeli duygu analizi için uyarlanmıştır. Otel ve film yorumlarını içeren veri kümesi üzerindeki deneylerimizde tercüme edilmiş veriyle eğitilen BERT’in İngilizce versiyonunun, yeterli eğitim kümesi sağlandığında çok yüksek başarımlı çalıştığı gözlemlenmiştir.

Bu çalışmanın katkıları aşağıdaki gibidir.

- Bildiğimiz kadarı ile, BERT Türkçe duygu analizi için daha önce kullanılmamıştır. Çalışmamızda BERT’in 3 farklı versiyonunun Türkçe duygu analizi için performansı incelenmiştir.
- Çok dilli modeller yerine otomatik tercüme sistemlerinin entegre edilip dil spesifik modellerin kullanımı incelenmiştir. Otomatik tercümenin genelde daha iyi sonuçlar verdiği gözlemlenmiştir.
- Önerdiğimiz modeller farklı veri kümelerinde eğitilip, eğitim ve test kümesinin farklı temalarda olduğu durumlarda modellerin başarımları incelenmiştir. BERT’in öğrendiği bilgiyi farklı alanlara genellemekte fastText’e göre daha başarılı olduğu gözlemlenmiştir.

II. LİTERATÜR

Türkçe duygu analizi çalışmalarının birçoğunda pozitif ya da negatif olmak üzere ikili bir sınıflandırma yapılmıştır.

Eroğul [6] film yorumları üzerine yaptığı duygu analizi çalışmasında kullandığı öznelilikler arasında n-gram ve kelime türleri bulunmaktadır. Vural vd. [16] yine film yorumları üzerine yoğunlaşmış ve İngilizce için geliştirilmiş gözetimsiz bir yöntem olan ve metin parçalarını pozitif veya negatif olarak puanlandıran SentiStrength yöntemini, Türkçeye uyarlamışlardır. Kaya vd. [9] siyasal içerikli Türkçe haberleri üzerinde ikili sınıflandırma yapmışlardır. Bu çalışmada n-gram'ın yanı sıra sıfatlar ve etkin kelimeler adını verdikleri seçilmiş kelimelerin kullanılmasının etkilerini, Destek Vektör Makinesi (Support Vector Machine), Naif Bayes gibi farklı makine öğrenmesi modelleri kullanarak analiz etmişlerdir.

Ucan vd. [15] duygu sözlüğü oluşturmak için bir yöntem önermişler ve İngilizce için geliştirilmiş olan SentiWordNet (SWN) kullanılarak bir Türkçe sözlük oluşturmuşlardır. Geliştirdikleri sözlük tabanlı modeli pozitif veya negatif etiketlere sahip film ve otel yorumları veri kümelerinde Destek Vektör Makinesi (DVM) modeliyle karşılaştırmışlardır. Sözlük tabanlı modelin, daha az veriden oluşan otel verisinde DVM'den daha iyi çalıştığını, fakat çok sayıda veriden oluşan film yorumları veri kümesinde daha kötü çalıştığını belirtmişlerdir. Erşahin vd. [7] SentiTurkNet duygu sözlüğünü, Automated Synonym Dictionary [1] kullanarak genişletip, sözlük tabanlı yöntemlerle makine öğrenmesi yöntemlerini birleştirerek hibrit bir çözüm önermişlerdir. Genişletilmiş sözlük kullandıklarında elde ettikleri başarımlarda film ve otel verisetlerindeki sırasıyla, %53 ve %88'lik bir artış sağlamışlardır.

Türkçe duygu analizi çalışmalarında pozitif, negatif veya nötr olmak üzere çoklu sınıflandırma yapan çalışmalar da bulunmaktadır. Yıldırım vd. [18] kelimelerin ters döküman frekansını (inverse document frequency) dikkate alarak öznelilikleri belirlemişler ve detaylı bir ön-işleme aşamasından sonra Türkçe tweetlerin duygu analizi için DVM modeli kullanmışlardır. Çoklu sınıflandırma gerçekleştirilen başka bir çalışmada [4] duygu analizini cümle seviyesine ek olarak döküman ve taraf (aspect) açısından ele almışlar ve bunlar arasında elde ettikleri sonuçları karşılaştırarak döküman seviyesindeki analizlerin daha başarılı olduğunu belirtmişlerdir.

Klasik doğal dil işleme yöntemlerinin dışında son yıllarda popülerleşen derin öğrenme yöntemleri de Türkçe duygu analizi için kullanılmıştır. Ayata vd. [2] Word2Vec kelime temsili modelini DVM ve Rassal Orman (Random Forest) algoritmalarıyla kullanarak banka, futbol, telekom ve perakende sektörü ile ilgili tweetler üzerinde duygu analizi gerçekleştirmişlerdir. Word2Vec [10] modeli ile elde edilen vektörlerin yine konvolüsyonel yapay sinir ağını eğitmek için kullanıldığı bir başka çalışmada [13], ürün yorumlarının pozitif veya negatif olmak üzere ikili sınıflandırılması amaçlanmıştır. Word2vec + CNN modelinin klasik modellerden daha iyi çalıştığı belirtilmiştir. Panthathi vd. [11], ürün yorumları üzerine derin öğrenme tabanlı bir yöntem kullanan çalışmada, kelime vektör modeli yerine genel amaçlı bir dil modeli olan ULMFit [8] algoritmasını Türkçe Wikipedia içerikleriyle gözetimsiz bir şekilde ön-eğitime tabi tutmuşur. Daha sonra model, restoran ve ürün yorumları için gözetimli bir veri seti ile ince-ayarlama (fine-tune) işlemine tabii tutulmuştur. Çalışmalarında, etiketlerin dağılımının dengesiz olduğu bir veri kümesinin eğitimde kullanılmasının negatif sonuçları olduğu belirtilmiştir.

Bildiğimiz kadar ile, BERT modelleri Türkçe duygu analizi için daha önce hiç uygulanmamış olup çalışmamız BERT'in

farklı boyut ve dil versiyonlarını kullanıp başarımını detaylı bir şekilde incelemiştir.

III. KULLANILAN MODELLER

Türkçe metinler üzerinde, geçmişte yapılan duygu analizi çalışmaları incelendiğinde, genellikle ilk olarak ön işlem adımlarının gerçekleştirildiği, daha sonra makine öğrenmesi veya derin öğrenme bazlı kelime temsil modellerinin uygulandığı görülmüştür ([16], [9], [18], [4], [2]). Modern kelime temsil modellerinden biri olan fastText'in, Türkçe Vikipedi verisi ile eğitilmiş bir modeli, fastText'in kendi sitesinde¹ mevcuttur. Çalışmada karşılaştırma yapmak için bu model tercih edilmiştir. Son yıllarda İngilizce üzerine yapılan duygu analizi çalışmalarında en iyi sonuçların, kelime temsil modellerinden genel amaçlı dil modelleriyle alındığı görülmüştür². 2018 yılında Google tarafından yayınlanan BERT, birçok doğal dil işleme probleminde en başarılı sonuçları elde etmiştir [5]. Bu yüzden biz de bu çalışmada BERT modelinin çok dilli versiyonu ve İngilizce versiyonlarını Türkçe duygu analizine uyarladık.

A. Çok Dilli BERT Modeli

Google tarafından tanıtılan önceden eğitilmiş BERT (Bidirectional Encoder Representations for Transformers) girdileri ardışık olarak okuyan genel amaçlı dil modellerinden farklı olarak, kelime gruplarını tek seferde okumaktadır. Yönsüz olarak adlandırılan bu yöntemle, kelimelerin, etrafındaki diğer kelimelerle olan ilişkisi öğrenilmektedir. BERT, öncelikle etiketsiz bir veri üstünde iki aşamalı bir ön-eğitime tabii tutulmaktadır. İlk aşamada cümledeki kelimelerin %15'i rastgele seçilmekte ve maskelenmektedir. Daha sonra iki yönlü dönüştürücü kodlayıcıya (bidirectional transformer encoder) verilerek maskelenen kelimeler tahmin edilmeye çalışılmaktadır. İkinci aşamada, cümleler arası ilişkilerin yakalanabilmesi için, BERT bir sonraki cümleyi tahmin ettirilerek eğitilmektedir. BERT'in iki farklı parametre sayısına sahip, ön-eğitimi tamamlanmış hazır 2 modeli bulunmaktadır. Bu modellerden orta ölçeklisi, 12 katman, 768 gizli boyut ve 12 dikkat noktasından (attention head) oluşmaktadır ve toplamda yaklaşık 110 milyon parametreden oluşmaktadır. Büyük ölçekli model ise, 24 katman, 1024 gizli boyut ve 16 dikkat noktasından oluşup, toplamda yaklaşık 340 milyon parametreye sahiptir. Bu devasa boyutlardaki modellerin eğitilmesi çok vakit aldığı için, modellerin belirli bir doğal dil işleme görevi için direkt eğitilmesi yerine, önceden eğitilmiş modellere ekstra katman eklenerek bu katman, amaçlanan görev için etiketli veriyle eğitilmektedir.

Google, farklı veri kümeleri üzerinde ön eğitilmiş olan BERT modellerini paylaşmıştır. Bu modeller İngilizce, Çince ve çok-dilli versiyonlarıdır. Biz bu çalışmada çok-dilli BERT modelini kullandığımız Türkçe veri kümeleri üzerinde ince-ayarlama (fine tuning) işlemine tutup Türkçe duygu analizine uyarladık.

B. İngilizce BERT Modeli

Çok dilli BERT modeli farklı diller üzerinden eğitildiğinden belirli bir dil için başarımı düşebilmektedir. Cümle

¹<https://fasttext.cc/>

²https://nlpprogress.com/english/sentiment_analysis.html

içindeki farklı kelime türü sıralamaları olan diller için de başarımının düştüğü gözlemlenmiştir [14]. Önerdiğimiz bu modelde, Türkçe metinleri önce makine tercüme yöntemleri ile İngilizceye çevirip daha sonra BERT'in sadece İngilizce için oluşturulan modeli ince-ayarlıma işlemine tabii tutulup modelimiz eğitilmiştir.

IV. DENEYSEL SONUÇLAR VE TARTIŞMA

A. Veri Kümesi

Geliştirilen modellerin başarımalarının ölçülmesinde, literatürde başka çalışmalar tarafından da kullanılmış (Ör., [7], [15]), film ve otel yorumlarını içeren iki farklı veri seti³ kullanılmıştır.

Film: Film yorumları beyazperde.com sitesinden kişilerin çeşitli filmler hakkındaki yorumlarını içermektedir. Film yorumları yazarları tarafından 1 ile 5 arasında puanlanmıştır. Veri kümesinin oluşturucuları tarafından, 1 ve 2 puan verilmiş yorumların negatif, 4 ve 5 puan verilen yorumların pozitif etikete sahip olduğu kabul edilmiştir. 26700 pozitif ve 26700 negatif olmak üzere toplam 53,400 örnekten oluşmaktadır.

Otel: Bu veri kümesi otelpuan.com sitesindeki kullanıcıların kaldıkları otellere dair yorumlarını içermektedir. Otel yorumları da yazarları tarafından 0 ile 100 arasında puanlandırılmıştır. Veri kümesinin oluşturucuları tarafından, 0-40 arası puan alan yorumlar negatif, 80-100 arası puan alan yorumlar pozitif olarak kabul edilmiştir. 5800 pozitif, 5800 negatif olmak üzere toplam 11600 yorumdan oluşmaktadır.

B. Deney Düzeni

Deneilerimizde 3 farklı düzenek kurulmuştur. Otel ve film veri kümeleri, oluşturucuları tarafından yarısı eğitim yarısı test olmak üzere ikiye ayrılmış ve bu şekilde paylaşılmıştır. Eğitim kümeleri $Otel_E$ ve $Film_E$, test kümeleri de $Otel_T$ ve $Film_T$ olarak gösterilecektir. Varolan çalışmalarla adil bir karşılaştırma olması için kendi modellerimizin eğitim ve testinde aynı veri kümeleri kullanılmıştır. Ayrıca farklı konulardaki veri kümelerinde eğitilip test edildiğindeki performanslarını ölçümleyebilmek için, kendi modellerimizi daha çok veriye sahip olan Film veri kümesinde eğitip, Otel veri kümesinin test kısmı üzerinde de değerlendirdik.

C. Gerçekleme Detayları

Türkçe metinlerin İngilizceye makine çevirisi için mT-ranslate kütüphanesi⁴ kullanılmıştır. BERT modelleri transformers kütüphanesinin 2.3.0 versiyonu kullanılarak Nvidia RTX 2080Ti ekran kartı ile eğitilmişlerdir. Çok kaynak isteyen bu modellerin eğitim parametreleri aşağıdaki tablodaki gibidir.

TABLO I: BERT MODEL PARAMETRELERİ

| Model | Çok Dilli Bert | Temel Bert | Büyük Bert |
|------------------|----------------|------------|------------|
| max_seq_length | 256 | 256 | 256 |
| train_batch_size | 4 | 4 | 1 |
| eval_batch_size | 4 | 4 | 1 |
| learning_rate | $2e^{-5}$ | $2e^{-5}$ | $2e^{-5}$ |
| num_train_epochs | 3 | 3 | 3 |

³<http://humirapps.cs.hacettepe.edu.tr/tsad.aspx>

⁴<https://pypi.org/project/mtranslate/>

D. Referans Modeller

Geliştirdiğimiz modelleri 4 farklı referans model ile karşılaştırdık.

fastText-Tr: fastText [3], Facebook tarafından yayınlanmış olup, kelime temsil vektörleri (word embeddings) ve bir metin sınıflandırma yöntemi içeren bir kütüphanedir. Word2Vec [10] ve Glove [12] kelime temsil modellerinden farklı olarak, en küçük birim olarak kelimeleri değil karakter n-gramlarını kullanmaktadır. Bu sayede daha önce hiç görmediği kelimeler için, Word2Vec ve Glove metodlarının aksine, vektör oluşturabilmektedir. Bu özelliği ile, yazım hatalarının veya gündelik dil kullanımının yaygın görüldüğü kullanıcı yorumlarının duygu analizi için önemli bir temel yöntemdir. Bu yöntemde, fastText'in Türkçe için geliştirilmiş olan kelime vektörleri ile kendi sınıflandırıcısı veri kümelerimizde eğitilerek bir duygu analizi modeli oluşturulmuştur. Modeli gerçekleştiren Zemberek'te bulunan versiyonu kullanılmış olup, learningRate ve epochCount parametreleri sırasıyla 0.1 ve 50 olarak seçilmiştir.

fastText-En: Bu referans yöntemde, kendi BERT modelimizde uyguladığımız gibi, Türkçe metinler önce makine-tercüme yöntemleri ile İngilizceye çevrilmiş sonra da fastText'in İngilizce için geliştirdiği kelime vektörleri kullanılarak sınıflandırıcı eğitilmiştir. Modelin gerçekleştirilmesi için fastText'in kendi sitesindeki C++ implementasyonu kullanılmış olup, lr, epoch ve wordNgrams parametreleri sırasıyla, 0.1, 50 ve 2 olarak seçilmiştir.

(Ersahin vd., 2019) [7]: Bölüm II'de daha detaylı anlatıldığı üzere, sözlük tabanlı yöntemlerle makine öğrenmesi yöntemlerini birleştirerek hibrit bir model geliştirmişlerdir.

(Ucan vd., 2016) [15]: Bölüm II'de detaylı anlatıldığı üzere, Ucan vd. İngilizceden uyarlanarak Türkçe için geliştirilen bir sözlüğü kullanarak Destek Vektör Makinelerini eğitmişlerdir.

E. Deneysel Sonuçlar

Tablo II'de kullandığımız tüm modellerin farklı eğitim ve test kümeleri için sonuçları gösterilmiştir. [7] ve [15] çalışmalarının sonuçları ilgili makalelerde raporlanan sonuçlardır. Deneilerimizde aşağıda belirtilen araştırma soruları (AR) incelenmiştir.

Türkçe duygu analizinde, otomatik tercüme ile İngilizceye çevrilen metinler üzerinde İngilizce için geliştirilen modellerin başarımı nedir? fastText-TR tüm durumlarda fastText-EN'den daha yüksek başarımlar sağlamıştır. Aynı temadaki veriler ile eğitilip test yapıldığı durumlarda (Örneğin, $Film_E + Film_T$) tercüme metinler üzerinden çalışan BERT modelleri çok dilli BERT modelinden daha iyi çalışmaktadır. Ancak farklı temalı veri kümeleri kullanıldığında (Örneğin, $Otel_E + Film_T$) çok dilli BERT modeli de diğer BERT modellerine göre Film veri kümesinde daha iyi, Otel veri kümesinde ise daha kötü çalışmaktadır.

Eğitim kümesinin performans üzerindeki etkisi nedir? fastText-TR ve fastText-EN her iki veri kümesi üzerinde de, test kümesinden farklı temada bir veri kümesini eğitimde kullandığımız durumlarda, aynı temadaki veriler ile eğitilmiş modellere göre daha düşük performans elde etmiştir. Bu da fastText modellerinin temalar arası kullanımını zorlaştırmaktadır. Ancak her 3 BERT modeli için de, otel yorumları yerine

TABLE II: DENEY SONUÇLARI.

| Eğitim Kümesi | Test Kümesi | fastText-TR | fastText-EN | Çok Dilli Bert-TR | Temel Bert | Büyük BERT | [7] | [15] |
|---------------|-------------|--------------|-------------|-------------------|------------|--------------|--------------|-------|
| $Film_E$ | $Film_T$ | 91,15 | 74,10 | 92,79 | 92,91 | 93,32 | 88,93 | 84,60 |
| $Otel_E$ | $Film_T$ | 70,55 | 58,00 | 69,80 | 66,35 | 65,35 | - | - |
| $Otel_E$ | $Otel_T$ | 80,18 | 80,80 | 66,03 | 74,07 | 73,04 | 91,96 | 80,70 |
| $Film_E$ | $Otel_T$ | 79,96 | 79,40 | 85,17 | 87,48 | 91,34 | - | - |

film yorumları kullandığımızda her iki veri kümesinde de performansı artmaktadır. Bunun sebebi otel yorumlarının film yorumlarından daha az olması olabilir. Deneylerimiz yeterli veri elde edildiğinde BERT'in öğrendiği bilgiyi başka temalara genelleştirebildiğini göstermektedir.

Geliştirdiğimiz modellerin başarımı varolan modellerin başarımlarına kıyasla nasıldır? Film veri kümesinde tüm BERT modelleri var olan çalışmalardan daha yüksek başarımlar elde etmiştir. Ancak Otel veri kümesinde tersi bir durum söz konusudur. Bu sorunun sebebinin, otel yorumlarına ait veri kümesinin boyutunun BERT'in iyi çalışması için yetersiz olmasından kaynaklandığını düşünmekteyiz. Otel veri kümesi için modellerimizi film yorumları ile eğittiğimizde Büyük BERT modeli en yüksek başarımlı olan [7] çalışmasının başarımlarından az bir farkla geride kalmıştır. Adil bir karşılaştırma olması için, varolan çalışmalar da film veri kümesi ile eğitilmesi gerekmektedir. Ancak sonuçlarımız BERT modellerine farklı temalarda bile olsa yeterli veri sağlandığında yüksek başarımlar elde edilebileceğini göstermektedir.

V. SONUÇ

Gerçekleştirilen bu çalışmada, Türkçe duygu analizi için BERT bazlı 3 farklı model geliştirilmiştir. Bu modeller, çok dilli modelin Türkçe metinler ile eğitilmesi ve Türkçe metinlerin önce otomatik tercüme ile İngilizceye çevrilip BERT'in temel ve büyük modellerinin eğitilmesi ile oluşturulmuştur. Otel ve film yorumlarını içeren iki farklı veri kümesinde gerçekleştirdiğimiz deneylerde elde ettiğimiz sonuçlar şöyledir: 1) BERT modeli fastText'e göre farklı temalardaki metinlerin duygusunu analiz etmede daha başarılıdır. 2) Yeterli miktarda eğitim verisi olduğunda, çok dilli BERT modeli yerine, otomatik tercüme ile metinlerin İngilizce'ye çevrilip sadece İngilizce için geliştirilmiş olan BERT modellerini kullanmak daha iyi sonuç vermektedir. 3) BERT modelleri film veri kümesinde varolan çalışmalardan daha yüksek başarımlar elde etmiştir. Otel veri kümesinde de, film yorumları ile eğitildiğinde varolan en iyi sistemden az bir farkla daha düşük bir başarımlar elde etmiştir.

Bu çalışmanın devamında, farklı veri kümeleri üzerinde XLNet [17] gibi diğer genel amaçlı metin gösterim modellerinin uyarlanması ve BERT ile sonuçların birleştirilmesi düşünülmektedir. Bunların yanı sıra, geliştirdiğimiz modellerin farklı veri kümeleri üzerinde de değerlendirilmesi gerekmektedir. BERT'in çok dilli versiyonu yerine sadece Türkçe için bir model geliştirip duygu analizi için kullanılması gelecekte yapılabilecek bir diğer çalışmadır.

KAYNAKLAR

- [1] Aktas, Ö., Birant, Ç. C., Aksu, B., and Çebi, Y. Automated synonym dictionary generation tool for turkish (asdict). *Bilig* 65 (2013), 47.
- [2] Ayata, D., Saraçlar, M., and Özgür, A. Turkish tweet sentiment analysis with word embedding and machine learning. In *2017 25th Signal Processing and Communications Applications Conference (SIU)* (2017), IEEE, pp. 1-4.

- [3] Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T. Enriching word vectors with subword information. *arXiv preprint arXiv:1607.04606* (2016).
- [4] Dehkharghani, R., Yanikoglu, B., Saygin, Y., and Oflazer, K. Sentiment analysis in turkish: Towards a complete framework. *Natural Language Engineering* 1 (2015).
- [5] Devlin, J., Chang, M., Lee, K., and Toutanova, K. BERT: pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)* (2019), pp. 4171-4186.
- [6] Eroglu, U. Sentiment analysis in turkish. Master's thesis, 2009.
- [7] ERŞAHİN, B., AKTAŞ, Ö., Kilinc, D., and ERŞAHİN, M. A hybrid sentiment analysis method for turkish. *Turkish Journal of Electrical Engineering & Computer Sciences* 27, 3 (2019), 1780-1793.
- [8] Howard, J., and Ruder, S. Universal language model fine-tuning for text classification. *arXiv preprint arXiv:1801.06146* (2018).
- [9] Kaya, M., Fidan, G., and Toroslu, I. H. Sentiment analysis of turkish political news. In *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 01* (2012), pp. 174-180.
- [10] Mikolov, T., Chen, K., Corrado, G. S., and Dean, J. Efficient estimation of word representations in vector space, 2013.
- [11] Panthathi, J., Bhaskar, J., Ranga, T. K., and Challa, M. R. Sentiment analysis of product reviews using deep learning. In *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (2018), IEEE, pp. 2408-2414.
- [12] Pennington, J., Socher, R., and Manning, C. Glove: Global vectors for word representation. In *Proc. of the 2014 conference on empirical methods in natural language processing* (2014), pp. 1532-1543.
- [13] Pervan, N., Keleş, Y., et al. *Derin öğrenme yaklaşımları kullanarak Türkçe metinlerden anlamsal çıkarım yapma*. PhD thesis.
- [14] Pires, T., Schlinger, E., and Garrette, D. How multilingual is multilingual bert? In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers* (2019), pp. 4996-5001.
- [15] Ucan, A., Naderalvojud, B., Sezer, E. A., and Sever, H. Sentiwordnet for new language: Automatic translation approach. In *2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)* (2016), IEEE, pp. 308-315.
- [16] Vural, A. G., Cambazoglu, B. B., Senkul, P., and Tokgoz, Z. O. A framework for sentiment analysis in turkish: Application to polarity detection of movie reviews in turkish. In *Computer and Information Sciences III*. Springer, 2013, pp. 437-445.
- [17] Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., and Le, Q. V. Xlnet: Generalized autoregressive pretraining for language understanding. *arXiv preprint arXiv:1906.08237* (2019).
- [18] Yıldırım, E., Çetin, F. S., Eryiğit, G., and Temel, T. The impact of nlp on turkish sentiment analysis. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi* 7, 1 (2015), 43-51.