



Kahramanmaraş Sutcu Imam University Journal of Engineering Sciences



Geliş Tarihi : 23.01.2023
Kabul Tarihi : 15.05.2023

Received Date : 23.01.2023
Accepted Date : 15.05.2023

DERİN TRANSFORMATÖRLERDEN ÇİFT YÖNLÜ KODLAYICI TEMSİLLERİ VE DESTEK VEKTÖR MAKİNELERİ İLE TÜRKÇE FİLM YORUMLARI ÜZERİNE DUYGU ANALİZİ

SENTIMENT ANALYSIS ON TURKISH MOVIE REVIEWS WITH DEEP BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS (BERT) AND SUPPORT VECTOR MACHINES

Hakan GÜNDÜZ^{1*} (ORCID: 0000-0003-2152-5490)

¹ Kocaeli Üniversitesi, Yazılım Mühendisliği Bölümü, Kocaeli, Türkiye

*Sorumlu Yazar / Corresponding Author: Hakan GÜNDÜZ, hakan.gunduz@kocaeli.edu.tr

ÖZET

Görüş madenciliği olarak da bilinen duygu analizi bir dizi kelimenin ardındaki görüşü belirlemenin yoludur. Duygu analizi, metinsel bir ifadeyle iletilen algıyı, düşünceleri ve duyguları daha iyi anlamak için kullanılır. Bu çalışmada Türkçe film sitesi beyazperde.com'dan derlenen film yorumları üzerinde duygu analizi yapılmıştır. Önerilen yöntem ön eğitilmiş BERTurk modelini temel almıştır. Yapılan ilk deneyde BERTurk modelinin sondan bir önceki dönüştürücü katmanından derin temsiller çıkarılmış ve bu temsiller Destek Vektör Makineleri (DVM) modeline girdi olarak verilmiştir. İkinci deneyde BERTurk üzerinde ince ayarlama yapılarak sınıflandırma gerçekleştirilirken, son deneyde ince ayarlı BERTurk modelinden ilk deneyde olduğu gibi derin temsiller çıkarılmış ve DVM ile sınıflandırma yapılmıştır. Yapılan deneylerde en yüksek doğruluk oranına 0.984 ile ince ayarlı BERTurk temsilleriyle ulaşılmıştır. İnce ayar işlemi sonunda elde edilen temsiller doğruluk oranında yaklaşık %10'luk artışa neden olurken, sınıflandırmada direkt olarak BERTurk yerine BERTurk'ten elde edilen temsiller ile DVM'nin birleşiminin kullanılması yaklaşık %5'lik doğruluk artışıyla sonuçlanmıştır.

Anahtar Kelimeler: Duygu analizi, film yorumları, BERT, BERTurk, derin temsiller

ABSTRACT

Sentiment analysis, also known as opinion mining, identifies the opinion behind a series of words. Sentiment analysis is used to understand better the perception, thoughts, and feelings conveyed in a textual expression. In this study, sentiment analysis was made on the movie reviews collected from the Turkish movie site beyazperde.com. The proposed method is based on the pre-trained BERTurk model. In the first experiment, deep representations were extracted from the penultimate transformer layer of the BERTurk model and given as input to the Support Vector Machines (SVM). In the second experiment, the classification was carried out by fine-tuning on BERTurk. In the last experiment, deep representations were extracted from the fine-tuned BERTurk model as in the first experiment and the classification with SVM was completed. Experiments have shown that fine-tuned BERTurk representations reach the highest accuracy with a rate of 0.984. While the representations obtained at the end of the fine-tuning process caused an increase of about 10% in the accuracy rate, the combination of the representations obtained from BERTurk with the SVM instead of using BERTurk directly in the classification resulted in an accuracy increase of about 5%.

Keywords: Sentiment analysis, movie reviews, BERT, BERTurk, deep representations

ToCite: GÜNDÜZ, H., (2023). DERİN TRANSFORMATÖRLERDEN ÇİFT YÖNLÜ KODLAYICI TEMSİLLERİ VE DESTEK VEKTÖR MAKİNELERİ İLE TÜRKÇE FİLM YORUMLARI ÜZERİNE DUYGU ANALİZİ. *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi*, 26(2), 542-549.

GİRİŞ

Web 2.0'ın ortaya çıkmasıyla birlikte Facebook, Twitter, LinkedIn, Instagram gibi çeşitli medyalar kişilerin farklı konularda yorumlarını, görüşlerini, duygularını ve yargılarını paylaşmalarına olanak tanımıştır. Bu platformlar twitler, bloglar ve gönderiler biçiminde büyük miktarda metinsel veri içermektedir. Yayımlanan bu metinsel veriler gerçekler ve görüşler olmak üzere iki grupta ele alınmaktadır. Gerçekler, dünyadaki varlıklar ve olaylar hakkında kesin ifadelerdir. Görüşler ise öznel olmakla beraber kişilerin varlıklar ve olaylar hakkındaki duyguları veya algılarıdır (Yadollahi vd., 2017). Mevcut doğal dil işleme (ddi) araştırmalarının bir kısmı, gerçekleri analiz etmeye ve çıkarmaya (örneğin arama motorları) odaklanırken, web üzerinde yayınlanmış görüşlerin tespiti ve analizi konusunda da ilerleyen literatür mevcuttur (Ain vd., 2017). Görüş madenciliği olarak da adlandırılan duygu analizi bir dizi kelimenin ardındaki görüşü belirlemenin yoludur. Bu analiz, çevrimiçi bir ifadeyle iletilen algıyı, düşünceleri ve duyguları daha iyi anlamak için kullanılır. Bu görev, belirli konular hakkında kamuoyunun bir özetini verdiği için sosyal medyada son derece yararlıdır.

Bilgi teknolojisindeki gelişmeler ve geniş bant internet bağlantılarının yaygınlaşması dâhilinde insanlar çeşitli sosyal medya kaynaklarında ürünler, filmler ve daha pek çok şey hakkındaki fikir ve düşüncelerini paylaşabilmektedir. Duygu analizi, bireylerin ürün ve faaliyetler gibi çeşitli konular hakkındaki görüş, değerlendirme ve duygularını belirleme sürecidir. Duygu analizi, duygu durumunu ve kullanıcı tarafından oluşturulan verilerde bir duygunun varlığını belirler. Duygu analizi genellikle metinsel verilerde yapılırken, konuşma veya video kaydı ile ilgili analiz çalışmaları da vardır (Poria vd., 2017).

Bu çalışmada film incelemeleri üzerine bir duygu analizi çatısı oluşturulmuştur. Önerilen çatı "Bidirectional Encoder Representations from Transformers" olarak kısaltılan BERT modeline dayalıdır. BERT modeli olarak huggingface platformundan indirilen ön eğitilmiş "bert-base-turkish-cased (BERTurk)" seçilmiştir. Çalışmada kullanılan film incelemeleri "beyazperde.com" web sitesinden derlenmiştir ve yaklaşık 54000 adet film incelemesi metnini ve bu incelemelere ait değerlendirme puanlarını içermektedir. Değerlendirme puanları incelemelere atanan sınıf etiketlerini belirlemektedir. Yapılan ilk deneyde ön eğitilmiş BERTurk'un sondan bir önceki dönüştürücü katmanından çıkarılan temsiller Destek Vektör Makinesi (DVM) sınıflandırıcısına girdi olarak verilerek sınıflandırma gerçekleştirilmiştir. İkinci deneyde ön eğitilmiş BERTurk modeli veri örnekleriyle eğitilerek ince ayarlama işlemi yapılmış ve model direkt olarak sınıflandırmada kullanılmıştır. Son deneyde ise ince ayarlanmış BERTurk modelinin sondan bir önceki katmanından derin temsiller elde edilmiş ve bu temsiller ilk deneyde olduğu gibi DVM modeline girdi olarak verilmiştir. Eğitilen modellerin performans değerlendirmelerinde ise doğruluk, kesinlik, anma ve F-Ölçütü metrikleri kullanılmıştır.

Çalışmamızın geri kalan kısmı şu şekildedir: Bölüm 2, Türkçe film incelemeleri üzerine gerçekleştirilen duygu analizi çalışmalarını özetlemektedir. Kullanılan Yöntemler Bölüm 3'te ayrıntılı olarak açıklanmıştır. Bölüm 4'te ise alınan deneysel sonuçlara ve değerlendirmelere yer verilmiştir.

İLGİLİ ÇALIŞMALAR

Duygu analizi alanındaki literatür çalışmaları ağırlıklı olarak İngilizce için olmakla birlikte, Türkçe duygu analizi çalışmaları son yıllarda artış göstermiştir (Gezici ve Yanıkoğlu, 2018; Çatal ve Nangir, 2017; Türkmenoğlu ve Tantuğ, 2014). Bu çalışmalarda sözlük tabanlı ve makine öğrenmesine tabanlı olmak üzere iki farklı yaklaşım kullanılmıştır (Türkmenoğlu ve Tantuğ, 2014). Örneğin Türkçe için makine öğrenmesi yöntemlerini temel alan ilk duygu analizi çalışması Eroğul tarafından gerçekleştirilmiştir (Eroğul, 2009). Bu çalışmada Beyazperde.com web sitesinden toplanan film yorumları üzerine doğal dil işleme teknikleri uygulanmış ve Destek Vektör Makineleri (DVM) ile %85'lik doğruluk oranına erişilmiştir.

Türkmenoğlu ve Tantuğ (2014) çalışmalarında iki farklı veri türüne hem sözlük tabanlı hem makine öğrenimi yöntemlerini uygulamıştır. Daha kısa ve yapılandırılmamış metinlere sahip Twitter verisi üzerindeki duygu analizi sözlük tabanlı yaklaşımla %75'lik doğruluk oranıyla sonuçlanmıştır. Aynı veri üzerinde DVM ile yapılan sınıflandırma sonucunda başarı oranı %85'e yükselmiştir. Kullanılan diğer veri türü olan film yorumları veri kümesi Twitter verilerine nazaran daha uzun ve nispeten yapılandırılmış metinlerden oluştuğundan DVM ile %89 doğruluk oranına erişilmiştir.

Nizam ve Akın (2014) çalışmalarında farklı gıda firmalarına ait Twitter verileri üzerinde duygu analizi gerçekleştirmiştir. Çalışmada veri örneklerine pozitif, negatif ve nötr olmak üzere çoklu sınıf etiketi ataması yapılmıştır. Ön işlenmiş twitler farklı yapay öğrenme modellerine girdi olarak verilirken, en yüksek sınıflandırma başarısına DVM'nin bir türevi olan Farklı Sıralı Minimal Optimizasyon (SMO) modeli ile erişilmiştir. Twitter verisinin kullanıldığı diğer bir duygu analizi çalışmasında ise sözlük tabanlı bir yöntem önerilmiştir. Çalışmada performans değerlendirmesi sınıf bazında yapılırken pozitif, negatif ve nötr etiketli sınıflar sırasıyla %72.7, % 70.31 ve % 73.16 doğruluk oranlarına sahip olmuştur.

Diğer bir duygu analizi çalışması Türkçe film yorumları üzerine Akgül vd. (2016) tarafından yapılmıştır. Çalışmada sözlüğe dayalı yaklaşım ve yapay öğrenme yöntemleri kullanılırken, model performansları F-Ölçütü metriği ile değerlendirilmiştir. Çalışmada yapay öğrenme yöntemlerinin sözlüğe dayalı yaklaşıma göre daha yüksek F-Ölçütü oranlarına sahip olduğu keşfedilirken, en yüksek performans sergileyen modelin %83'lük F-Ölçütü oranıyla DVM olduğu gözlemlenmiştir.

Son yıllarda literatürde BERT modelinin kullanıldığı Türkçe duygu analizi çalışmaları da bulunmaktadır. Açıkalin (2020) tarafından gerçekleştirilen çalışmada Türkçe film incelemeleri ve otel yorumları BERT modeli kullanılarak sınıflandırılmıştır. Kilimci (2020) ise çalışmasında derin öğrenme temelli finansal duygu tahmin sistemi önermiştir. Tahmin işleminde veri olarak bankacılık hisseleri ile ilgili twitler kullanılırken, derin öğrenme mimarileri olarak Evrişimsel Sinir Ağları, Uzun-Kısa Süreli Bellek, Özyineli Sinir Ağları ve BERT seçilmiştir. Elde edilen sınıflandırma performansları incelendiğinde, hisselerin büyük çoğunluğunda en yüksek doğruluk oranlarına BERT temsilleri ile erişildiği görülmüştür. BERT modelinin kullanıldığı diğer bir Türkçe duygu analizi çalışmasında ise veri kümesi olarak Türkçe twitler kullanılmıştır. Çalışmada twitlerdeki duyguların tespiti için BERT modeliyle farklı yapay öğrenme modelleri kıyaslanmış ve en yüksek doğruluğa %98,75 ile BERT'in eriştiği görülmüştür (Güven, 2021). Covid-19 pandemisi ile ilgili Instagram gönderileri üzerinde yapılan duygu analizi çalışmasında ise ön-eğitilmiş BERT modellerinin duyguların sınıflandırılmasında en yüksek F-ölçütü oranlarına sahip olduğu belirtilmiştir (Karayığit vd., 2022).

YÖNTEMLER

Bu bölümde çalışmada kullanılan BERT dil modeli ve Destek Vektör Makineleri sınıflandırıcısı hakkında özet bilgiler verilmiştir. Ayrıcı model performanslarının değerlendirilmesinde kullanılan ölçütlerin detayları açıklanmıştır.

Bidirectional Encoder Representations from Transformers (BERT)

Dönüştürücüden Çift Yönlü Kodlayıcı Temsilleri (Bidirectional Encoder Representations from Transformers) olarak kısaltılan BERT, dönüştürücü (transformer) mimarisine dayalı bir dil modelidir. Dil modelleme, basitçe bir metindeki eksik kelimeyi tahmin etme görevi olarak tanımlanır. Geleneksel dil modelleri bir sonraki kelimeyi kendinden önce gelen kelimeleri kullanarak (soldan sağa yaklaşımla) tahmin ederken, BERT mimarisi bu yaklaşımın aksine tahminde önceki ve sonraki kelimelerden yararlanır. Soldan sağa yaklaşım, özellikle metin üretilmesi gibi kelimelerin tek tek tahmin edildiği görevlerde başarıyla uygulanır. Ancak, cümle içinde rastgele bir kelime tahmin edildiğinde, BERT'i klasik yaklaşımlardan ayıran özellik rastgele kelimeyi maskeleyerek tahmin etmeye çalışmasıdır (Devlin vd., 2018).

BERT modelinden önce geliştirilen Word2Vec ve GloVe gibi kelime gömülmesi modelleri büyük ve etiketlenmemiş derlem ile eğitilmesine rağmen bu modeller birleşik kelimeleri ve olumsuzluk anlamlarını kotarma yeteneğine sahip değildir. Modellerin diğer bir eksikliği ise kelimelerin geçtiği bağlamı göz önünde bulundurmamasıdır (Jin vd., 2020). BERT modeli Word2Vec ve GloVe modellerinin bu dezavantajlarını derin öğrenme mimarilerini kullanarak giderir. BERT'in üzerine inşa edildiği dönüştürücü mimarisine kelimenin bağlamındaki diğer kelimeler göz önünde bulundurulur ve bağlamdaki kelimelere ağırlık mekanizması uygulanarak kelime temsilleri oluşturulur. Dönüştürücülerin uzun vadeli kelime bağımlılıklarını yüksek performansla tespit etme yeteneği, BERT modelinin dönüştürücü mimarisi üzerine inşa edilmesinin temel nedenidir. BERT, her biri 1024 gizli birimden oluşan 24 dönüştürücü bloğundan oluşur ve yaklaşık 340 milyon parametreye sahiptir.

Destek Vektör Makineleri (DVM)

Destek Vektör Makineleri (DVM), Vapnik tarafından geliştirilen, sınıflandırma ve regresyon problemlerinde kullanılan bir yapay öğrenme yöntemidir (Joachims, 1999). DVM, diğer yapay öğrenme yöntemlerine kıyasla doğrusal olmayan veri kümelerini modelleme başarısıyla ön plana çıkmaktadır. İkili sınıflandırma problemlerinde doğrusal yapıya sahip olmayan eğitim örnekleri, mevcut giriş uzayında bir doğruyla iki sınıfa ayrılamaz. Bu gibi durumlarda, DVM, mevcut uzayı doğrusal olmayan bir haritalama fonksiyonuyla eğitim örneklerinin doğrusal olarak ayrılabilceği yüksek boyutlu bir öznelik uzayına taşır. Haritalama işleminde kullanılan fonksiyonlar genellikle ikinci dereceden, kübik ve trigonometrik terimler içermektedir. Çekirdek fonksiyonları olarak adlandırılan bu fonksiyonlar ile ilk aşamada tüm eğitim örnekleri için fonksiyon değerleri hesaplanır. Sonraki aşamada ise yeni örnekler ile eğitim örnekleri arasındaki benzerlik değerleri hesaplanan çekirdek fonksiyon değerleri üzerinden kolayca bulunur. DVM'de yaygın olarak kullanılan çekirdekler doğrusal, polinom, sigmoid ve radyal tabanlı fonksiyonlardır (Gunduz, 2021).

Değerlendirme Ölçütleri

Çalışmada model performanslarını doğruluk, kesinlik, anma ve F-Ölçütü metrikleriyle değerlendirilmiştir. Sınıflandırma problemlerinde doğruluk oranından yaygın olarak yararlanır. Ancak dengesiz sınıf dağılımına sahip veri kümelerinin performans değerlendirmesinde doğruluk oranı yanıltıcı sonuçlar doğurmaktadır. Bu durumun önüne geçebilmek için sınıf ayrılabilirliğinin modeller tarafından ne kadar iyi gerçekleştirilebildiğini gösteren F-Ölçütü'nden yararlanır (Gulsen vd., 2015). F-Ölçütü, karışıklık matrisi elemanları kullanılarak hesaplanır. Karışıklık matrisi ikili sınıflandırma problemlerinde sınıf bazında doğru ve yanlış sınıflandırılan örneklerin sayısını tutar ve doğru pozitif (dp), yanlış pozitif (yp), yanlış negatif (yn) ve doğru negatif (dn) olmak üzere 4 elemana sahiptir. Karışıklık matrisi temel alınarak kesinlik, anma ve F-Ölçütü metrikleri aşağıdaki şekilde hesaplanır (Catal vd., 2021):

$$\text{Kesinlik (K)} = \frac{dp}{dp + yp} \quad (1)$$

$$\text{Anma (A)} = \frac{dp}{dp + yn} \quad (2)$$

$$\text{F-Ölçütü} = \frac{2 * K * A}{(K + A)} \quad (3)$$

Eşitliklerde belirtildiği üzere, F-Ölçütü; anma ve kesinlik ölçütlerinin harmonik ortalaması alınarak hesaplanır.

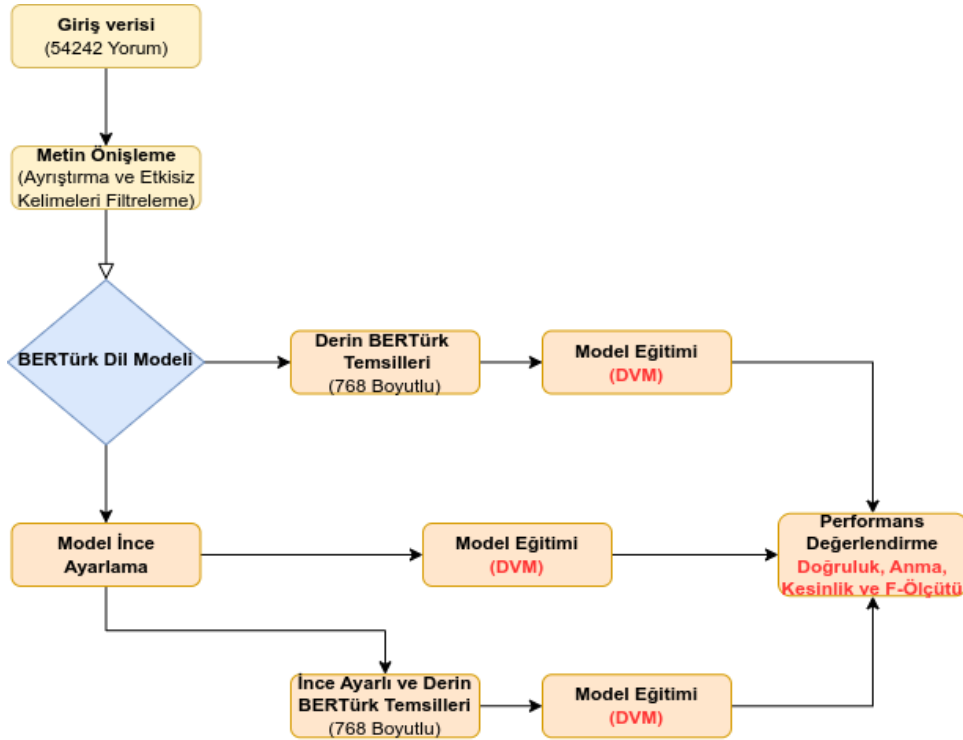
DENEYSEL SONUÇLAR

Çalışmada kullanılan veri kümesi beyazperde.com isimli web sitesinden derlenmiştir. İlk olarak Türkmenoğlu ve Tantuğ (2014) 'un çalışmasında kullanılan veri kümesi 54242 adet örnekten oluşmaktadır. Örnekler yorum yapılan film adlarını, kullanıcıların filmlere yaptıkları yorumları ve filmlere verilen beğeni puanlarını içermektedir. Atanan değerlendirme puanları ölçeği 1 ile 5 arasındadır. Çalışmada filmlerin beğeni durumu tahmin edileceğinden filmlere atanan puanlar 2 sınıfa ayrılmıştır. İzleyici tarafından filme verilen puan 4'ün altında ise sınıf etiketi olarak 0 (beğenilmedi), 4 ve üzerinde ise etiket olarak 1 (beğenildi) atanmıştır. Etiket atama aşamasından sonra film yorumları üzerinde ön işleme gerçekleştirilmiştir. Ön işlemede karakter katarı olarak bulunan metinler kelimelere ayrıştırılmış (Tokenization) ve etkisiz kelimeler ihmal edilmiştir.

Sınıflandırma işleminde doğal dil işleme deposu **huggingface'den** indirilen "**bert-base-turkish-cased (BERTurk)**" modeli kullanılmıştır (Wolf vd., 2019). BERT modelinin temel versiyonu olan "bert-base-turkish-cased" her biri 768 adet gizli birime sahip 12 adet dönüştürücü bloğu içermektedir. Elde edilen model sınıflandırma görevinde kullanılacağından bu modelin son dönüştürücü katmanından sonra doğrusal bir katman daha eklenmiştir. BERTurk modeli OSCAR derlemi, Wikipedia metinleri ve Kemal Oflazer tarafından sağlanan özel bir derlem ile ön eğitilmiştir (BERTurk, 2019).

Kullanılacak modelin elde edilmesinden sonra veri kümesi örnekleri BERTurk modeli girdisi olarak biçimlendirilmiştir. Örneklerin biçimlendirilmesi için yine huggingface tarafından sağlanan **BertTokenizer** modülü

kullanılmıştır (Wolf vd., 2019). Bu modül ile veri örnekleri en fazla 64 token (kelime) içeren nitelik vektörlerine dönüştürülmüştür. Veri biçimlendirildikten sonra örneklerin %80'lik kısmı eğitim kümesi, geri kalan %20'lik kısmı ise test kümesi olarak belirlenmiştir. Önerilen sınıflandırma modellerine ait şematik gösterim Şekil 1'de verilmiştir.



Şekil 1. Önerilen Sınıflandırma Modellerine ait Şematik Gösterim

Yapılan ilk deneylerde ön eğitilmiş BERTürk modelinden derin öznitelik temsilleri çıkarılmıştır. Modelin son dönüştürücü katmanı sınıflandırma işleminin yapıldığı lineer katmanla benzerliğe sahip olduğundan derin temsiller BERTürk modelinin sondan ikinci (11.) dönüştürücü bloğunun çıkışından elde edilmiştir. Her veri örneği 64 token ile temsil edildiğinden, çıkarılan özniteliklerin boyutu her örnek için 64×768 olmuştur. Örnekleri tek boyutlu vektörlere indirgemek için ortalama havuzlama (average pooling) işlemi gerçekleştirilmiş ve her örnek 768 boyutlu vektörle temsil edilmiştir. Elde edilen vektörler son aşamada DVM modeline girdi olarak verilmiş ve sınıflandırma işlemi gerçekleştirilmiştir. Sınıflandırma sonuçları Tablo 1'de gösterilmektedir.

Tablo 1. Ön-Eğitilmiş BERT Temsilleriyle Alınan Sonuçlar

Nitelik Türü	Doğruluk	Anma	Kesinlik	F-Ölçütü
BERT Temsilleri (Ön eğitilmiş)	0,874	0,926	0,895	0,910

Gerçekleştirilen ikinci deneylerde, veri örnekleri ön eğitilmiş BERTürk modeline girdi olarak verilmiş ve ince ayarlama işlemi gerçekleştirilmiştir. Ayrılan eğitim kümesiyle BERTürk modelinin ince ayarlaması Tablo 2'de listelenen ağ parametreleri kullanılarak yapılmıştır. İnce ayarlama kullanılan parametrelerin belirlenmesinde ızgara arama yönteminden yararlanılmıştır.

Modelin ince ayarlaması tamamlandıktan sonra test kümesi örnekleriyle model başarımları değerlendirilmiştir. Elde edilen sınıflandırma sonuçları Tablo 3'te listelenmiştir.

İnce ayarlı modelle direkt sınıflandırma yapıldıktan sonra, en son aşamada bu modelden elde edilen derin özniteliklerle deneyler yapılmıştır. İlk deneyde olduğu gibi BERTürk'un sondan ikinci dönüştürücü katmanın (11.

dönüştürücü bloğun) çıkışından her örnek için 64*768 boyutlu derin temsiller elde edilmiştir. Ortalama havuzlama işlemi ile örnekler yine 768 boyutlu vektörlerle temsil edilmiştir. Oluşturulan eğitim kümesi örnekleriyle DVM sınıflandırıcısı eğitilirken, test kümesi örnekleriyle model performansı değerlendirilmiştir. Elde edilen sınıflandırma sonuçları Tablo 4'te gösterilmiştir.

Tablo 2. BERTurk Modelinin Ağ Parametreleri

Parametre	Değer
Yığın boyutu	32
Optimizasyon fonksiyonu	Adam
Öğrenme oranı	2e-5
İterasyon sayısı	20

Tablo 3. İnce Ayarlanmış BERT Temsilleriyle Alınan Sonuçlar

Nitelik Türü	Doğruluk	Anma	Kesinlik	F-Ölçütü
BERT Temsilleri (İnce ayarlı)	0,931	0,960	0,938	0,949

Tablo 4. İnce Ayarlanmış BERT Temsilleriyle Alınan Sonuçlar

Nitelik Türü	Doğruluk	Anma	Kesinlik	F-Ölçütü
BERT Temsilleri (İnce ayarlı ve derin öznitelikler)	0,984	0,989	0,981	0,985

Tablo 5. İnce Ayarlanmış BERT Temsili Modeline ait Karmaşıklık Matrisi

Tahmin/ Gerçek	0	1
0	25941	486
1	303	27512

Tüm deneylerin sonuçları incelendiğinde en yüksek sınıflandırma sonucuna ince ayarlı ve derin özniteliklerle ulaşıldığı görülmüştür. Bu modele ait karmaşıklık matrisi Tablo 5'te gösterilmiştir. İnce ayarlı derin temsillerle eğitilen modelle 0.984'lük doğruluk ve 0.981'lik F-Ölçütü oranlarına erişilmiştir.

SONUÇLAR VE DEĞERLENDİRME

Görüş madenciliği olarak da nitelendirilen duygu analizi bir dizi kelimenin ardındaki görüşü belirlemenin temel yoludur. Bu çalışmada ön eğitilmiş BERTurk modeli temel alınarak Türkçe film yorumları üzerinde duygu analizi yapılmıştır. En yüksek model performansına (0.984'lük doğruluk ve 0.981'lik F-Ölçütü) ince ayarlı derin temsillerle ulaşılrken, ince ayarlanmış bu temsiller sadece ön eğitilmiş BERT temsillerine göre doğruluk oranında yaklaşık %10'luk, F-Ölçütü oranında ise yaklaşık %7'lik performans artışına sebep olmuştur. İnce ayarlı BERT modeliyle doğrudan sınıflandırma yapıldığında ise doğruluk oranı 0.931 olmuştur. BERT'in dönüştürücü katmanlarından sonra lineer bir sınıflandırma katmanı kullanması doğrudan BERT temsilleriyle eğitilen modelin başarısının düşük kalmasında önemli rol oynamıştır. DVM'nin doğrusal olmayan verileri modellemedeki kabiliyeti, derin temsiller ile DVM modeli kombinasyonunun BERT'in doğrudan kullanımına göre daha yüksek başarı oranına sahip olmasının temel nedenidir.

Önceden belirtildiği üzere bu çalışmanın veri kümesinin ilk kullanıldığı çalışma Türkmenoğlu ve Tantuğ (2014)' tür. Bu çalışmada en yüksek doğruluk oranı TF-IDF (Term Frequency-Inverse Document Frequency) öznelikleri ve DVM modeliyle 0.89 olarak gerçekleşmiştir. Bu çalışmaya kıyasla doğruluk oranında yaklaşık %9'luk artış meydana gelmiştir. Artışın ön önemli sebebi BERT'in cümlelerdeki kelime bağlamlarını göz önünde bulundurarak derin temsiller oluşturmaktır. Çalışmada alınan sonuçlar yakın dönem Türkçe duygu analizi çalışmaları ile de kıyaslanmış ve deney sonuçlarının bu çalışmalardan daha yüksek doğruluk ve F-Ölçütü oranlarına eriştiği görülmüştür (Tablo 6).

Tablo 6. Yakın Dönem Türkçe Duygu Analizi Çalışmaları

Nitelik Türü	Doğruluk	F-Ölçütü
Polat ve Ağca (2022)	0,870	0,917
Şahinaslan vd. (2022)	0,656	0,654
Uca vd. (2022)	0,940	-
Sağbaş (2023)	0,943	0,942
*Bu Çalışma	0,984	0,981

İleri dönem çalışmalarında bu çalışmaya ek olarak elde edilen derin temsillerin topluluk öğrenme modellerine girdi olarak verilmesi ve bu modellerden nitelik önem puanlarının öğrenilmesi planlanmaktadır. Elde edilen nitelik puanları ile düşük puanlı nitelikler elimine edilecek ve sınıflandırma daha az boyutlu nitelik temsilleriyle yapılacaktır.

KAYNAKLAR

- Ain, Q. T., Ali, M., Riaz, A., Noureen, A., Kamran, M., Hayat, B., & Rehman, A. (2017). Sentiment analysis using deep learning techniques: a review. *International Journal of Advanced Computer Science and Applications*, 8(6).
- Akgül, E. S., Ertano, C., & Banu, D. İ. R. İ. (2016). Twitter verileri ile duygu analizi. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 22(2), 106-110.
- Catal, C., Gunduz, H., & Ozcan, A. (2021). Malware detection based on graph attention networks for intelligent transportation systems. *Electronics*, 10(20), 2534.
- Catal, C., & Nangir, M. (2017). A sentiment classification model based on multiple classifiers. *Applied Soft Computing*, 50, 135-141.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Eroğul, U. (2009). Sentiment analysis in Turkish (Master's thesis). *Middle East Technical University, Ankara*.
- Gezici, G., & Yanıkoğlu, B. (2018). Sentiment analysis in Turkish. In *Turkish natural language processing* (pp. 255-271). Springer, Cham.
- Gulsen, E., Gunduz, H., Cataltepe, Z., & Serinol, L. (2015, May). Big data feature selection and projection for gender prediction based on user web behaviour. In *2015 23rd Signal Processing and Communications Applications Conference (SIU)* (pp. 1545-1548). IEEE.
- Gunduz, H. (2021). An efficient stock market prediction model using hybrid feature reduction method based on variational autoencoders and recursive feature elimination. *Financial Innovation*, 7(1), 1-24.
- Guyen, Z. A. (2021, September). Comparison of BERT models and machine learning methods for sentiment analysis on Turkish tweets. In *2021 6th International Conference on Computer Science and Engineering (UBMK)* (pp. 98-101). IEEE.
- Jin, D., Jin, Z., Zhou, J. T., & Szolovits, P. (2020, April). Is bert really robust? a strong baseline for natural language attack on text classification and entailment. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 34, No. 05, pp. 8018-8025).
- Joachims, T. (1999). Making large-scale svm learning. *Practical Advances in Kernel Methods-Support Vector Learning*.

- Karayığit, H., Akdagli, A., & Acı, Ç. İ. (2022). BERT-based Transfer Learning Model for COVID-19 Sentiment Analysis on Turkish Instagram Comments. *Information Technology and Control*, 51(3), 409-428.
- Kilimci, Z. H. (2020). Financial sentiment analysis with Deep Ensemble Models (DEMs) for stock market prediction. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 35(2), 635-650.
- Nizam, H., & Akın, S. S. (2014). Sosyal medyada makine öğrenmesi ile duygu analizinde dengeli ve dengesiz veri setlerinin performanslarının karşılaştırılması. *XIX. Türkiye'de İnternet Konferansı*, 1(6).
- Polat, H., & Ağca, P. (2022). Tripadvisor kullanıcılarının türkçe ve ingilizce yorumları kapsamında duygu analizi yöntemlerinin karşılaştırmalı analizi. *Abant Sosyal Bilimler Dergisi*, 22(2), 9
- Poria, S., Cambria, E., Hazarika, D., Majumder, N., Zadeh, A., & Morency, L. P. (2017, July). Context-dependent sentiment analysis in user-generated videos. In *Proceedings of the 55th annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 873-883).
- Sağbaş, E. A. (2023). Filtre Tabanlı Öznitelik Seçim Yöntemleri Kullanılarak Metinlerde Duygu Sınıflandırması Üzerine Karşılaştırmalı Bir Çalışma. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 35(1), 239-250.
- Şahinaslan, Ö., Dalyan, H., & Şahinaslan, E. (2022). Naive bayes sınıflandırıcısı kullanılarak youtube verileri üzerinden çok dilli duygu analizi. *Bilişim Teknolojileri Dergisi*, 15(2), 221-22
- Türkmenoglu, C., & Tantug, A. C. (2014, June). Sentiment analysis in Turkish media. In *International Conference on Machine Learning (ICML)*.
- Uca, E., Yilmazer, S., Kizilhan, H., & Karaköse, M. (2022). Sağlık Alanındaki Etkileşimlerde Duygu Analizi Yaklaşımları ve Analizi. *Fırat Üniversitesi Uzay ve Savunma Teknolojileri Dergisi* 1(1), 465-470,
- Yadollahi, A., Shahraki, A. G., & Zaiane, O. R. (2017). Current state of text sentiment analysis from opinion to emotion mining. *ACM Computing Surveys (CSUR)*, 50(2), 1-33.
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... & Rush, A. M. (2019). Huggingface's transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*.
- BERTurk (2019). <https://huggingface.co/dbmdz/bert-base-turkish-cased>, Erişilme tarihi: 19.10.2022.