

# BERT Modeli'nin Sınıflandırma Doğruluğunun Sıfır-Atış Öğrenmesi ile Artırılması

## Increasing the Classification Accuracy of BERT Model with Zero-Shot Learning

Şükrü OZAN

AdresGezgini A.Ş.,  
Ar-Ge Departmanı, İzmir,  
Türkiye  
sukruozan@adresgezgini.com  
ORCID: 0000-0002-3227-348X

Umut ÖZDİL

AdresGezgini A.Ş.,  
Ar-Ge Departmanı, İzmir,  
Türkiye  
umutozdil@adresgezgini.com  
ORCID: 0000-0002-6909-1727

Davut Emre TAŞAR

Dokuz Eylül Üniversitesi  
YL. Öğr., İzmir,  
Türkiye  
davutemre.tasar@ogr.deu.tr  
ORCID: 0000-0002-7788-0478

Büşra ARSLAN

AdresGezgini A.Ş.,  
Ar-Ge Departmanı, İzmir,  
Türkiye  
busraarslan@adresgezgini.com  
ORCID: 0000-0002-6155-7967

Gökçe POLAT

AdresGezgini A.Ş.,  
Ar-Ge Departmanı, İzmir,  
Türkiye  
gokcepolat@adresgezgini.com  
ORCID: 0000-0003-1657-6824

### Öz

Bu çalışmada, çevrim içi reklam platformlarında oluşturulan reklam metinlerinin sektöre göre otomatik olarak sınıflandırılması için Çift Yönlü Kodlayıcı Gösterimleri (BERT) modeli kullanılmıştır. Eğitim veri setimiz 44 farklı sektöre ait yaklaşık 101.000 adet, ilgili sektör isimleri ile etiketlenmiş reklam metninden oluşmaktadır. Sınıflandırma başarımının belirli sektörler için düşük olduğu gözlemlenmiştir. Bunun, ilgili eğitim veri setindeki aykırı bazı reklam metinlerinden kaynaklandığı anlaşılmıştır. Bu sebeple sektör ve reklam metni uygunluğunu otomatik olarak tespit etmek ve veri seti için bir ön işleme gerçekleştirmek amacıyla doğal dil işleme (NLP) alanında güncel bir yöntem olan Sıfır-Atış Öğrenmesi yöntemi kullanılmıştır. Bu ön işleme çalışmasından sonra temizlenen veri seti ile yeniden eğitilen BERT modelinin sınıflandırma başarımının önemli ölçüde arttığı gözlemlenmiştir.

Gönderme ve kabul tarihi: 05.11.2021 - 02.12.2021

Makale türü: Araştırma

**Anahtar Sözcükler:** Dijital Pazarlama, Reklam Metni, NLP, Metin Sınıflandırma, BERT, Sıfır-Atış Öğrenmesi

### Abstract

This study used the BERT model to automatically classify ad texts created on online advertising platforms according to the industry. Our training dataset consists of approximately 101,000 advertisement texts from 44 different sectors labeled with the relevant sector names. It has been observed that the classification performance is low for specific sectors. It has been understood that this is due to some outlier advertising texts in the relevant training dataset. For this reason, the Zero-Shot Learning method, which is a current method in natural language processing (NLP), was used to automatically determine the suitability of the sector and advertisement text and perform a preprocessing for the data set. After this preprocessing study, it was observed that the classification performance of the BERT model, which was retrained with the cleaned data set, increased significantly.

**Keywords:** Digital Marketing, Ad Text, NLP, Text Classification, BERT, Zero-Shot Learning

## 1. Giriş

Geçtiğimiz on yılda internet kullanımının dünya genelinde yaygınlaşması ve özellikle son iki yıldır yaşanan COVID-19 pandemisi ile birlikte internet reklamlarının kullanımı oldukça yaygınlaşmıştır. Bu yaygınlaşmanın sonucunda geleneksel reklamlara göre daha pratik olan internet reklamları, geleneksel reklam mecralarından daha büyük bir pazar payına sahip olmuştur [1].

Dijital pazarlama, ürün ve hizmetlerini internet üzerinden tüketicilere sunmak isteyen işletmeler için önemli bir araç haline gelmiştir. Çevrim içi reklamlar hedef odaklı olmasının yanı sıra reklam performansının takibine ve analizine olanak sağlamaktadır [2]. Bunun paralelinde de çevrim içi reklamları, yatırım ve dönüşüm ilişkisinin incelendiği etkili reklam çalışmaları üzerine birçok araştırma yapılmaya başlanmıştır. Bu araştırmalar kapsamında reklamların yazılı, görsel ve işitsel içerikleri, kullanıcı hareketleri, cihaz ve konum bilgisi veya demografi gibi öznitelikler üzerinden anlamlı sonuçlar çıkarmak için makine öğrenimi algoritmaları ve NLP teknikleri kullanılmaktadır [3].

Özellikle Google gibi arama motorlarında uzun süredir kullanılmakta olan arama ağı reklamları kullanıcılara büyük oranda yazı tabanlı gösterilmektedir. Kaliteli bir arama ağı reklam çalışmasının en önemli özelliği, tüketiciye sunulan içeriğin alakalı ve etkili olmasıdır. Çevrim içi reklamlık platformlarında reklam kalitesi, reklam metninin aramalarla ne kadar alakalı olduğu, tüketicilerin reklama tıklama olasılığı ve reklam tıklandıktan sonra görüntülenen sayfada (landing page) yaşanan kullanıcı deneyimi gibi birçok farklı faktöre bağlıdır. Yüksek reklam kalitesi daha iyi reklam konumu ve daha düşük maliyeti beraberinde getirmektedir [4]. Bu noktada tüketici ihtiyaçlarına ve aramalarına cevap veren reklam metninin hazırlanması, reklam verenin sunduğu ürünün/hizmetin potansiyel müşterilere etkili bir şekilde aktarılmasında önemli bir rol oynamaktadır.

Çalışmamızda, verilen bir reklam metninin hangi sektör ile daha alakalı olabileceğini tespit etmek amacıyla BERT mimarisini [5] kullanan bir sınıflandırma yöntemi önerilmiştir. Yapılan araştırmalarda [6]-[8] çalışmaları incelenmiş ve BERT mimarisinin benzer metin sınıflandırma problemlerinde Word2vec, LSTM gibi modellerden daha iyi sonuç verdiği tespit edilmiştir. Bu sebeple

gerçekleştirilen çalışmada BERT modelinin kullanılmasına karar verilmiştir.

Bu çalışmada, 44 kategori için BERT yöntemi ile gerçekleştirilen sınıflandırma işleminin ortalama doğruluk oranının, gürlütlü olarak tabir edilebilecek çok sayıda aykırı veri sebebiyle görece düşük olduğu gözlemlenmiştir. NLP alanında son dönemde sıklıkla kullanılmaya başlanan Sıfır-Atış Öğrenmesi (SAÖ) sayesinde aykırı örneklerin başarı ile tespit edilebildiği gözlemlenmiştir. Aykırı örneklerin de dışarıda bırakılmasının ardından temizlenmiş veri seti ile tekrardan eğitilen BERT modeli ile gerçekleştirdiğimiz sınıflandırma işleminin ortalama doğruluğunun önemli ölçüde arttığı gözlemlenmiştir.

Bu makalenin devam eden bölümlerinde gerçekleştirilen çalışmaların ve deneylerin ayrıntılarına ve elde edilen sonuçlara yer verilmiştir.

## 2. İlgili Çalışmalar

Son yıllarda yapay zekâ alanındaki hızlı gelişmelere paralel olarak, NLP alanında da birçok yeni mimari yapı ve yöntem geliştirilmiştir. Bir cümle içerisindeki kelimeler arasındaki semantik bağlamı bulmak için uzun kısa süreli bellek (LSTM) [9], dikkat mekanizması [10], yinelemeli sinir ağları [11], kapalı sinir ağları [12] ve evrişimli sinir ağları [13] gibi birçok mimari denenmiştir.

Nergiz vd. [14] çalışmalarında, özyineli yapay sinir ağı (RNN) mimarisi olan LSTM ile fastText, Word2vec ve Doc2vec modellerini birlikte kullanarak, Türkçe haber sitelerinde yer alan farklı kategorilere ait haber metinlerini sınıflandırmıştır. Bu şekilde tasarlanan modellerin sınıflandırma başarı oranlarını karşılaştırılmış ve en başarılı sonucun fastText modeli ile elde edildiğini gösterilmiştir.

Doğru vd. [15] çalışmasında, Türkçe ve İngilizce haber metinleri üzerinde Doc2vec kelime gömme yöntemi ile eğitim modeli oluşturulmuş, derin öğrenme sınıflandırma yöntemi olarak CNN, makine öğrenmesi sınıflandırma yöntemleri olarak Gauss Naive Bayes (GNB), Random Forest (RF), Naive Bayes (NB) ve Support Vector Machine (SVM) modelleri kullanılmıştır. Bu modellerin doğruluk oranlarını karşılaştırarak en yüksek sonucun CNN ile yapılan sınıflandırmada elde edildiği gösterilmiştir.

González-Carvajal ve Garrido-Merchán [16] çalışmasında, sosyal medyada paylaşılan iletler, film-dizi eleştirileri, haber içerikleri gibi veri setleri üzerinde BERT modeli ve geleneksel doğal dil işleme

yaklaşımlarının karşılaştırmalı analizlerini gerçekleştirmiş ve sonuçlarını sunmuştur.

Zhengjie vd. [17] çalışmasında duygu analizine bağlı sınıflandırma yapmak için BERT modelinin hedefe bağlı üç varyasyonu kullanılmıştır. Üç veri kümesi üzerinde yapılan deneyler sonucunda modelin cümle düzeyinde duyarlılık sınıflandırması dâhil olmak üzere birçok NLP probleminde mevcut yöntemlere göre daha iyi sonuçlar verdiği gösterilmiştir.

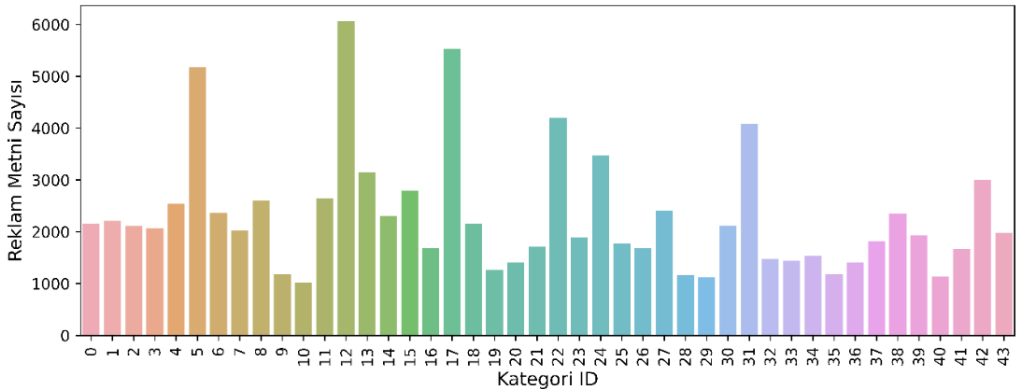
Bu çalışmanın ön çalışması niteliindeki [18]'de gösterildiği üzere, verilen bir reklam metnine ait kategorinin belirlenmesinde BERT mimarisini başarılı bir şekilde kullanılabilmektedir. [18] çalışması özelinde seçilmiş olan 11 adet kategori için, veri setinin geneline göre daha az sayıda aykırı örnek içeren kategorilere ait ortalama F1 skoru metriği ile ölçülen sınıflandırma doğruluğunun %91 dolayında olduğu raporlanmıştır. Bu çalışmada ise eldeki veri setinin büyük bir kısmını kapsayan toplam 44 adet kategori için aynı şekilde gerçekleştirilen sınıflandırma işlemi sonucunda daha düşük bir sınıflandırma doğruluğu elde edilmiştir.

Kategoriler genel olarak incelendiğinde özellikle düşük sınıflandırma performansı gözlemlenen kategorilerin çok sayıda aykırı örnek içerdiği

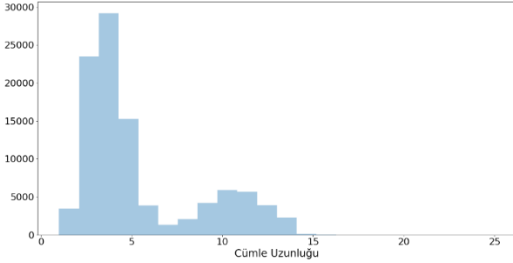
gözlemlenmiştir. Bu sebeple, NLP alanında özellikle dönüştürücü tabanlı dil modellerinin gelişmesi ile birlikte son yıllarda sıklıkla kullanılmaya başlanan SAÖ yöntemi kullanılmıştır.

Bir makine öğrenmesi problemi olan SAÖ'de, test aşamasındaki sistem, eğitim aşamasında hiç görmediği kategorilere ait örneklerin kategorilerini doğru şekilde tahmin edebilmektedir. Temel olarak bu yöntem, örneklerin gözlemlenen ayırt edici özelliklerini kodlayan yardımcı bilgi türleri aracılığıyla, gözlemlenen ve gözlemlenmeyen kategorileri birleştirerek çalışmaktadır.

Bu yöntemin NLP alanında kullanılması ilk olarak 2008 tarihindeki [19] çalışması ile gerçekleştirilmiştir. Metin sınıflandırma için SAÖ yönteminin kullanıldığı [20] gibi çalışmalar da vardır. Bu çalışmada yazarlar yaklaşık dört milyon internet sitesini tarayarak haber başlıkları ile bu haber başlıklarının bulunduğu sayfalarda yer alan arama motoru optimizasyonu (SEO) etiketlerini kullanmışlardır. Word2vec kelime gömme vektörleri [21] ile eğitilen SAÖ modelinin, test verisi üzerinde %94 üzerinde bir doğrulukla etiket tahmini gerçekleştirebildiği gözlemlenmiştir.



**Şekil-1:** Çalışma kapsamında değerlendirilen 44 kategoriye ait veri sayısının dağılımı. Kategorilere göre veri dağılımının dengesiz olduğu gözlemlenmektedir.



**Şekil-2:** Veri setinde yer alan reklam metnlerinin içerdiği kelime sayısının histogram grafiği. Kelime sayısı dağılımının iki farklı ortalama değer etrafında kümelenmiş normal dağılım şeklinde olduğu gözlemlenmektedir.

Bu çalışmada, SAÖ modelinin bu performansını kullanarak, ilgili bir kategoride verilen reklam metninin aşıldığı o kategori ile ilgili olup olmama durumu tespit edilmiştir. İlgili olmama durumunda, söz konusu veri aykırı olarak işaretlenerek bir sonraki aşamada gerçekleştirilecek olan BERT modeli eğitiminde kullanılacak eğitim verisi setinin dışında tutulmuştur. Bu sayede BERT modelinin sınıflandırma doğruluğunun önemli ölçüde arttırıldığı gözlemlenmiştir.

### 3. Tasarım ve Yöntem

#### 3.1 Veri Seti

Kullanılan veri seti, 44 farklı kategoriye ait yaklaşık 101.000 reklam metninden oluşmaktadır. Ön işleme aşamasında veri seti öncelikle tekrar eden verilerden, noktalama işaretleri ve sembollerden arındırılmıştır. 44 kategori arasından rastgele seçilmiş olan 11 kategori için yine rastgele seçilmiş üç reklam metni, veri setinin daha iyi anlaşılabilmesi açısından Çizelge-1'de verilmiştir. Şekil 1'de gösterilen dağılımdan anlaşılacağı üzere, her bir kategoride çok farklı sayıda örnek bulunmaktadır ve dolayısı ile veri seti dengesiz (imbalanced) bir dağılıma sahiptir. Veri setindeki reklamların kelime sayısı histogramını ise Şekil 2'de görülmektedir. Bu dağılımdan da anlaşılacağı gibi kelime sayısı dağılımı iki farklı ortalama değer etrafında kümelenmiş normal dağılım şeklindedir.

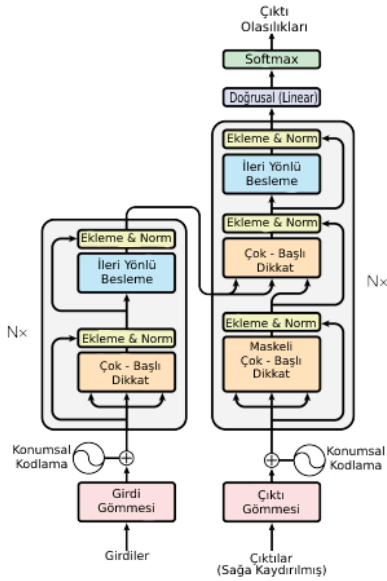
**Çizelge-1:** Rastgele Seçilmiş Kategorilere Ait Veri Örnekleri

Kategori	Reklam Metni
Evcil Hayvan Ürünleri	Hayvan Dostunuz için En İyi Si Köpek Maması ve Vitaminleri Türkiye'nin Rakipsiz Petshop'u
Mücevher, Takı & Aksesuar	Özel Taşlı Bileklikler Size En Yakışan Saat Burada! En Ucuz & En Yeni Kolye Modelleri.
Nakliyat, Kargo	Nakliye ve Depolamada Öncü Firma Güvenilir ve Hızlı Kargo Paketleyerek Profesyonel Taşıma
Oto Aksesuar & Yedek Parça	Son Teknoloji Hidrolik Yağlar Araç Koltuk Temizliği Aradığınız Tüm Jant Çeşitleri
Psikolojik Danışmanlık	İlaçsız Terapi Yöntemi Öfke Kontrolü Yöntem & Tedavi Kişiyeye Özel Klinik Terapi Yöntemleri
Temizlik & Halı Yıkama	Profesyonel Yıkama Hizmetleri Temizlik Malzemeleri ve Aparatlar. Uygun Fiyatlarla Mükemmel Temizlik!
Tur Acenteleri	Yurtiçi & Yurtdışı Tur Fırsatı Avantajlı Otel & Tatil Fırsatı Aradığınız Konaklama Fırsatları
Vize İşlemleri	Hızlı & Kolay Vize İşlemleri Evraksız Rusya Vizesi Vize İşlemlerinin Doğru Adresi
Yabancı Dil Eğitimi	Müfredata Uygun Eğlenerek Öğrenme Erken Yaşta İngilizce Dil Eğitimi Konuşma Odaklı Drama ile İngilizce
Yurtlar	Ücretsiz wireless & Çalışma Odaları Evinizdeki Konfor için Bize Ulaşın! 7/24 Sıcak Su & Güvenlik Hizmeti.

#### 3.2 BERT Modeli ile Sınıflandırma

BERT modeli [5], etiketlenmemiş metinden çift yönlü gösterimleri önceden eğtmek ve sonrasında farklı NLP görevleri için etiketli metinler kullanılarak ince ayar yapmak için tasarlanmış ve Google tarafından 2018 yılında sunulmuş bir dönüştürücü modelidir. BERT ile birlikte Google Translate uygulamasının çeviri başarısı da önemli ölçüde artmıştır ve artık bu

uygulama, çok uzun cümleleri bile, anlam kaybı olmadan büyük oranda doğru bir şekilde farklı diller arasında çevirebilmektedir [23][24]. Dönüştürücüler, kuyruk yapısıyla oluşturulmuş öz dikkat mekanizması ile çalışan bir yapıya sahiptir (bkz. Şekil 3).



Şekil-3: Dönüştürücü Model Mimarisi [22]

BERT modeli, bir sorguyu ve bir dizi anahtar-değer çiftini bir çıktıya eşlemektedir. Burada sorgu, anahtarlar, değerler ve çıktının kendi aralarındaki korelasyonu ifade edecek vektörler oluşmaktadır. Çıktı, değerlerin ağırlıklı toplamı ile hesaplanmaktadır. Bir değere atanan ağırlık ise sorguya karşılık gelen anahtarla uyumluluk oranı ile hesaplanmaktadır [22]. BERT modeli, bir metni hem sağdan sola hem de soldan sağa işlemektedir, bu sayede metin içerisindeki öğeler arasındaki ilişkileri öğrenebilmektedir. Eğitim aşamasında, MLM (Masked Language Modeling) ve NSP (Next Sentence Prediction) teknikleri kullanılmaktadır. MLM tekniğinde, maskelenen kelimeler, açık (maskelenmeyen) kelimeler kullanılarak tahmin edilmeye çalışılmaktadır. Bu teknik ile cümle içindeki kelimeler üzerinden inceleme ve tahminleme yapılmaktadır. NSP tekniğinde, bir cümlenin kendisinden sonra gelen cümle ile ilişkisine bakılarak

cümlelerin birbirleri ile ilişkisi incelenmektedir. BERT modeli ile kurulan yapılarda önceden eğitilmiş bir model gereksinimi vardır. Bu sebeple çalışmamızda, Loodos takımı tarafından Türkçe dili için e-kitaplar, haber yazıları, çevrim içi blog yazıları ve Wikipedia gibi kaynaklardan elde edilmiş 200GB büyüklüğündeki veri seti ile MLM tekniği kullanılarak önceden eğitilmiş bir BERT modeli olan BERT-BASE-TURKISH-UNCASED [25] modeli tercih edilmiştir. Bu modelin tercih edilmesinin esas nedeni ise daha önce yapılan çalışmalarda benzer sınıflandırma problemleri içerisinde en yüksek doğruluk oranlarına bu model ile ulaşılmış olmasından kaynaklanmaktadır [26]. Kullanılan modelin ön eğitim parametreleri Çizelge-2’de verilmiştir. Model parametreleri ön eğitilmiş BERT modeline ait parametreler ile aynı olacak şekilde seçilmiştir. Bu modelde kullanılan Gizli Katman Boyutu, Maksimum Sekans Uzunluğu, Dikkat Ana Başlığı Sayısı, Gizli Katman Sayısı, Mimari parametreleri ön eğitilmiş modeller için standart kullanılan parametrelerden oluşmaktadır. Bu model üzerinde ince ayar işlemi yapılarak kullanıma hazır hale getirilmiştir.

Çizelge-2: Kullanılan Ön Eğitilmiş BERT Modelinin Parametreleri

Model	Gizli Katman Boyutu	Maks. Sekans Uzunluğu	Dikkat Ana Başlığı Sayısı	Gizli Katman Sayısı	Mimarî	Sözlük Boyutu
(Loodos 2020)	768	512	12	12	BERT for Masked LM	32000

Modelin başarımını doğru bir şekilde değerlendirmek için, doğruluk değerinin yanı sıra, F1 skoru olarak tabir edilen değer de takip edilmiştir. F1 skorunun hesaplanmasında True Positive (TP), False Positive (FP), True Negative (TN) ve False Negative (FN) değerleri kullanılmaktadır. TP, modelin tahmini ve gerçek değerlerin her ikisinde de olumlu sonuç vermesi; TN, modelin tahmini ve gerçek değerlerin her ikisinde de olumsuz sonuç vermesi; FP, model tahmini olumlu iken gerçek değer olumsuz sonuç vermesi; FN, modelin tahmini olumsuz iken gerçek değer olumlu sonuç vermesi şeklinde açıklanmaktadır. Bu durumda TP ve TN doğru sonuç, FP ve FN ise yanlış sonuç kabul edilmektedir. Doğruluk değeri, modelin doğru tahmin ettiği TP ve TN değerlerinin, tahmin edilen tüm TP, TN, FP, FN değerlerine oranı ile hesaplanmaktadır (Denklem (1)). Kesinlik (precision) değeri, modelin tahmin ettiği TP

değerlerin sayısının, modelin ürettiği tüm olumlu sonuçlar olan TP ve FP değerlerinin sayısına oranıdır (Denklem (2)). Duyarlılık değeri ise modelin tahmin ettiği TP değerlerinin sayısının, modelin üretmesi gereken tüm olumlu sonuçlar olan TP ve FN sayılarına oranı ile bulunabilir (Denklem (3)). F1 skoru ise kesinlik ve duyarlılık değerlerinin harmonik ortalaması olarak tanımlanabilir (Denklem (4)).

$$\text{Doğruluk} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Kesinlik} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{F1 Skoru} = 2 \left( \frac{\text{Kesinlik} \times \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \right) \quad (4)$$

### 3.3 Sıfır-Atış Öğrenmesi ile Veri Setinin Aykırı Verilerden Temizlenmesi

SAÖ yöntemi hem çoklu sınıflandırma özelliği ile hem de çoklu sınıflandırma özelliği olmadan kullanılmıştır. Her iki değerlendirme kapsamında, elde edilen sonuçta, en yüksek olasılığa ait sınıflandırma değeri eldeki veri setinin önceden atanmış olan etiketi ile aynı değilse ilgili veri metni aykırı bir metin olarak işaretlenerek eğitimin dışında tutulmuştur.

Aykırı metinlerin daha iyi anlaşılabilmesi için Çizelge-3'de verilen örnekler incelenebilir. Görüleceği üzere sol sütunda gösterilen aykırı metinlerin, sağ sütunda belirtilen kategoriler ile anlamsal bir bağının kurulması pek mümkün görülmemektedir. Bu sebeple bu metinler BERT modeli ile tasarladığımız sınıflandırıcının doğruluğunun görece düşük seviyelerde seyretmesine sebep olmaktadır.

**Çizelge-3: Aykırı Metin Örnekleri**

Aykırı Metin	Bulunduğu Kategori
9.90 TL Fiyat Avantajını Kaçırmanın	Anne & Bebek Ürünleri
Nefes Almanın En Sağlıklı Yolu	Elektronik & Bilgisayar
Taksitle Ödeme İmkânı	Emlakçılar
Cebe Dost Fiyatlar Burada	Güvenlik & Alarm Sistemleri
En Uygun Fiyata Hizmet Garantisi.	Marka Tescil & Belgelendirme
Bilgi için Şimdi Arayın	Yazılım Hizmetleri

Uzman Ekip ve Uygun Maliyet.	Enerji
Proje, Mühendislik, Kurulum Hizmeti	Ambalajlama

Bu şekilde aykırı verilerden ayrıştırılmış veri setinden, eğitimde kullanılmaya uygun yaklaşık 20.000 adet veri elde edilmiştir. Şekil 1'e benzer şekilde kategorilere göre veri sayısı dağılımı çizdirildiğinde Şekil 4'teki dağılım elde edilmiştir. Yine benzer şekilde reklam metinlerindeki kelime sayısının histogram grafiği de Şekil 5'teki gibi olmaktadır.

## 4. Tartışmalar ve Sonuçlar

Çevrim içi reklam platformlarında oluşturulan reklam metinlerinin sektöre göre otomatik olarak sınıflandırılması için gerçekleştirdiğimiz bu doğal dil işleme uygulamasında BERT modeli ile bir sınıflandırma gerçekleştirilmiştir. Veri setindeki yaklaşık 101.000 adet verinin 80.000'lik kısmı ile sistemi eğitimi gerçekleştirilmiştir. Veri setimizdeki 44 farklı sektöre ait yaklaşık 20.000 adet etiketli test verisini %79'luk doğruluk ile sınıflandırabilmiştir. Gerçekleştirilen sınıflandırma işlemine ait kategori bazında hesaplanmış doğruluk ve F1 skoru değerleri Çizelge-4'te görülebilir. Doğruluk değerlerine göre sıralanan bu çizelgede görüleceği üzere en yüksek doğruluk ve F1 skor değerleri sırasıyla %93 ve %94 olmaktadır. 44 kategoriye ait doğruluk ve F1 skoru değerlerinin, kategorilerdeki veri sayısı göz önünde bulundurularak hesaplanmış olan ağırlıklı ortalama değerleri ise %79 değeri ile sınırlı kalmaktadır.

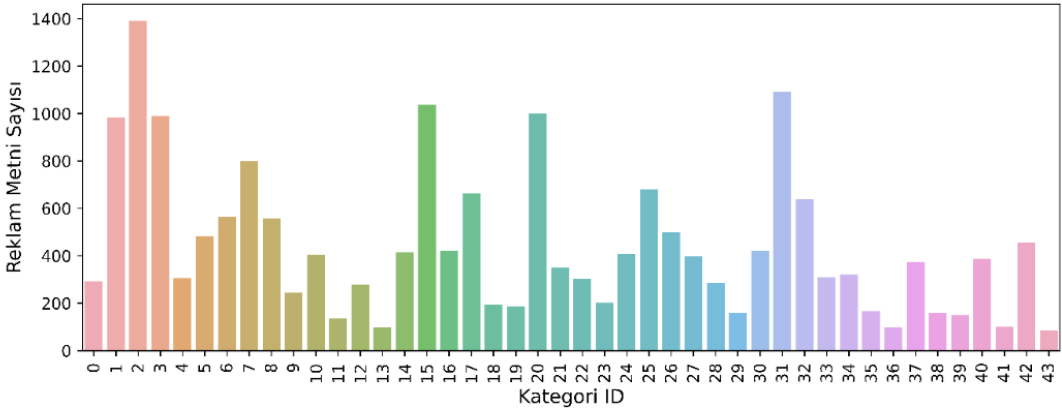
BERT modeli ile 44 kategori için elde edilen sonucun düşük olmasından dolayı veri odaklı yapay zekâ yaklaşımı ile modelden ziyade eldeki veri setine odaklanarak modelin doğruluk değerinin artırılması hedeflenmiştir. Bunun manuel olarak gerçekleştirilmesinin zorluğundan dolayı bu işlemi otonom olarak gerçekleştirmek için güncel yöntemlerden birisi olan SAÖ yöntemi kullanılmıştır. SAÖ yöntemi ile gerçekleştirilen ayrıştırma işleminden sonra yaklaşık 20.000 adet aykırı örnekten temizlenmiş veri seti elde edilmiştir. Bu işlemin yapılmasındaki temel amaç BERT modeline ince ayar yapılarak gerçekleştirilen öğrenme aktarması ile elde edilen başarımın SAÖ ile elde edilen başarı ile birleştirilmesidir. Bu işlemi gerçekleştirmek için SAÖ sayesinde elde edilen veri seti, tercih edilen BERT modeline girdi olarak verilerek modelin tahmin ettiği değerlere göre sınıflandırma işlemi gerçekleştirilmiştir. Veri setinin yaklaşık 16.000'lik

kısmı ile BERT modelinin eğitimi yeniden gerçekleştirilmiş, kalan yaklaşık 4.000 adet test verisi ile doğruluk ve F1 skoru değerleri yeniden hesaplanarak Çizelge-5'te gösterilen sonuçlar elde edilmiştir. Yine doğruluk değerlerine göre sıralanmış olan bu çizelgede görüleceği üzere birçok kategorinin doğruluk ve F1 skoru %100 değerlerine ulaşabilmiştir.

**Çizelge-4: Eldeki Veri Seti ile Doğrudan Eğitilen BERT Modeli ile Elde Edilen Kategorilere Göre Sınıflandırma Doğruluk Değerleri Çizelgesi**

I D	Kesinlik	Duyarlılık	Doğruluk	F1 Skoru	Örnek Sayısı
41	0,94	0,93	0,93	0,94	363
35	0,89	0,92	0,92	0,91	424
31	0,85	0,92	0,92	0,88	474
28	0,88	0,91	0,91	0,89	507
27	0,85	0,90	0,90	0,87	694
39	0,81	0,89	0,89	0,84	840

12	0,86	0,88	0,88	0,87	519
4	0,92	0,88	0,88	0,9	353
2	0,89	0,86	0,86	0,88	288
40	0,83	0,86	0,86	0,84	560
⋮	⋮	⋮	⋮	⋮	⋮
15	0,72	0,70	0,70	0,71	816
11	0,67	0,69	0,69	0,68	252
3	0,78	0,69	0,69	0,73	342
34	0,80	0,68	0,68	0,74	282
29	0,79	0,68	0,68	0,73	442
9	0,67	0,68	0,68	0,67	385
0	0,76	0,67	0,67	0,71	236
23	0,71	0,67	0,67	0,69	338
16	0,73	0,65	0,65	0,69	462
38	0,79	0,62	0,62	0,7	306
	<b>0,79</b>	<b>0,79</b>	<b>0,79</b>	<b>0,79</b>	<b>20203</b>
	<b>(Ağ. Ort.)</b>	<b>(Ağ. Ort.)</b>	<b>(Ağ. Ort.)</b>	<b>(Ağ. Ort.)</b>	<b>(Toplam)</b>

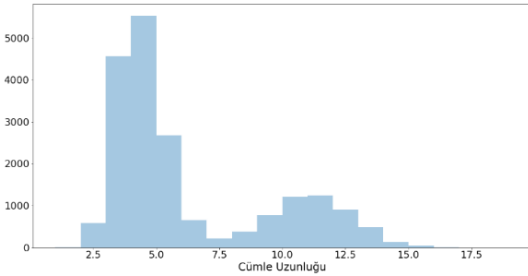


**Şekil-4:** Şekil 1'de görülen 44 kategoriye ait veri sayısı dağılımının, veri setinin Sıfır-Atış Öğrenmesi yöntemi ile aykırı verilerden arındırıldıktan sonraki dağılım grafiği.



**Çizelge-5: Sıfır-Atış Öğrenmesi Yöntemi ile Aykırı Verilerden Arındırılmış Veri Seti İle Eğitilen BERT Modeli ile Elde Edilen Kategorilere göre Sınıflandırma Doğruluk Çizelgesi**

ID	Kesinlik	Duyarlılık	Doğruluk	F1 Skoru	Örnek Sayısı
0	1,00	1,00	1,00	1,00	75
2	0,99	1,00	1,00	1,00	100
4	1,00	1,00	1,00	1,00	133
5	0,90	1,00	1,00	0,95	19
10	1,00	1,00	1,00	1,00	27
11	0,99	1,00	1,00	0,99	81
12	0,98	1,00	1,00	0,99	62
24	0,99	1,00	1,00	1,00	113
25	1,00	1,00	1,00	1,00	196
29	1,00	1,00	1,00	1,00	37
⋮	⋮	⋮	⋮	⋮	⋮
32	1,00	0,97	0,97	0,98	30
16	0,92	0,96	0,96	0,94	49
36	0,91	0,96	0,96	0,93	91
13	1,00	0,95	0,95	0,97	61
21	0,95	0,95	0,95	0,95	20
27	1,00	0,95	0,95	0,97	20
42	0,97	0,93	0,93	0,95	80
6	0,88	0,88	0,88	0,88	17
38	0,96	0,82	0,82	0,89	33
9	0,88	0,69	0,69	0,77	32
	<b>0,98</b>	<b>0,98</b>	<b>0,98</b>	<b>0,98</b>	<b>3897</b>
	<b>(Ağ. Ort.)</b>	<b>(Ağ. Ort.)</b>	<b>(Ağ. Ort.)</b>	<b>(Ağ. Ort.)</b>	<b>(Toplam)</b>



**Şekil-5:** Şekil 2'de görülen veri setinde yer alan reklam metinlerinin içerdiği kelime sayısının histogram grafiğinin, Sıfır-Atış Öğrenmesi yöntemi ile aykırı verilerden arındırıldıktan sonraki durumu şekildeki gibi olmaktadır.

44 kategoriye ait doğruluk ve F1 skoru değerlerinin, kategorilerdeki veri sayısı göz önünde bulundurulacak hesaplanmış olan ağırlıklı ortalamaları ise %98

değerine ulaşmıştır. Çizelge-4 ve Çizelge-5'in genişletilmiş hali GitHub platformunda paylaşılmıştır [27].

Manuel etiketlenerek kategorilendirilen veri setinin NLP algoritmasının çalışma prensibine uygun olmayan bir yapıda girdi içermesi sorunu, SAÖ yöntemi ile giderilerek bu değerlere ulaşılmıştır. Bu sayede öğrenme aktarması ile yapılan ilk denemede düşük başarımlarına sahip olan 0, 23, 16 gibi kategorilerde distilasyon sonrasında yüksek başarı elde edilebilmiştir. Bunun gerçekleşmesini sağlayan durum ise Çizelge-3'te görülen aykırı metinlerin SAÖ yöntemi ile öğrenme aktarması gerçekleşmeden önce elemine edilmiş olmasından kaynaklanmaktadır. Ayrıca yine Çizelge-5'te görüldüğü gibi düzgün şekilde ayıklanmış bir veri seti ile uygulanan öğrenme aktarması işleminden sonra örnek sayısı 50'nin altında olsa dahi yüksek doğruluk oranlarına sahip sonuçlar elde edilebilmektedir. Bu nedenden ötürü SAÖ yönteminin, kategorik sınıflandırma problemlerinde uygulanmasının modelin başarımını olumlu etkilediği kanaatine ulaşılmıştır.



Bu çalışmada, çevrim içi reklam platformlarında oluşturulan reklam metinlerinin kategorilere göre sınıflandırılması amacı ile BERT modeli kullanılmıştır. 44 adet kategoriye dengesiz bir dağılıma sahip veri setinin yaklaşık %80'lik kısmı ile gerçekleştirilen eğitimin ardından kalan %20'lik kısmı ile yapılan doğruluk değeri hesaplamalarında modelin ağırlıklı ortalama doğruluğunun ve F1 skoru değerlerinin %80'in altında kaldığı görülmüştür. Bununla birlikte modelin eğitimi esnasında, çapraz-doğrulama 5 defa gerçekleştirilmiştir. Çizelge-6'da sunulan deneysel sonuçlar incelendiğinde çapraz-doğrulama işleminin modelin başarımlarına olan etkisinin dikkate alınmayacak düzeyde düşük olduğu tespit edilmiştir. Modelin eğitildiği veri seti incelendiğinde SAÖ modeli ile distile edilmiş veri setinden reklam metinlerinin kategorik olarak homojen bir yapıya sahip olması bunun başlıca nedeni olarak görülmektedir. En gelişmiş ve güncel modellerden birisi olan BERT'in sınıflandırma başarısındaki bu düşüklüğün sebebinin modelden ziyade veri setinden kaynaklanabileceği düşüncesiyle veri seti daha dikkatle analiz edilmiştir. Sonuçta veri setinden çok miktarda aykırı metin örnekleri bulunduğu görülmüştür.

**Çizelge-6: Tekli ve 5 Katlı Çapraz-Doğrulama Kullanan Modellerin Test Verisi Üzerindeki Performansı**

Model	BERT		BERT + Sıfır-Atış	
	Tek (Ağ. Ort.)	5-Kat Çapraz Doğrulama (Ağ. Ort.)	Tek (Ağ. Ort.)	5-Kat Çapraz Doğrulama (Ağ. Ort.)
Kesinlik	0,79	0,80	0,98	0,97
Duyarlılık	0,79	0,79	0,98	0,97
Doğruluk	0,79	0,79	0,98	0,97
F1 Skoru	0,79	0,79	0,98	0,97

Aykırı metinlerin ayıklanması için yine çok güncel bir makine öğrenmesi yöntemi olan SAÖ yöntemi kullanılmış ve veri setinde önceden etiketlenmiş olan verilerin etiketlerinin tutarlılığı analiz edilmiştir. Bu analiz sonucunda temizlenen veri seti ile aynı eğitim ve test prosedürü gerçekleştirildiğinde BERT modelinin ağırlıklı ortalama sınıflandırma doğruluğunun ve F1 skoru değerlerinin %98 seviyesine yükseldiği gözlemlenmiştir.

## Kaynakça

- [1] Emarketer., “Advertisers will spend nearly \$600 billion worldwide in 2015,” 2014.
- [2] Google., “1.1 Çevrimci reklamcılık ve Google Ads’in avantajları - Google Ads Yardım,” <https://support.google.com/google-ads/answer/6123875?hl=tr>, May 2021.
- [3] Z, A. and Adali, E., “Opinion mining and sentiment analysis for context-tual online-advertisement,” in 2016 IEEE 10th International Conference on Application of Information and Communication Technologies (AICT). IEEE, 2016, pp. 1–3.
- [4] Google., “Reklam kalitesi hakkında - Google Ads Yardım,” <https://support.google.com/google-ads/answer/156066?hl=tr&reftopic=10549746>, May2021.
- [5] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K., “Bert: Pre-training of deep bidirectional transformers for language understanding,” arXiv preprint arXiv:1810.04805, 2018.
- [6] Ozan, S., and Tasar, D. E., “Auto-tagging of short conversational sentences using natural language processing methods,” in 2021 29th Signal Processing and Communications Applications Conference (SIU), 2021, pp. 1–4
- [7] Zulfat Miftahutdinov, Ilseyar Alimova, E. T., “Kfu nlp team at smm4h2019 tasks: Want to extract adverse drugs reactions from tweets? bertto the rescue.,” ACL 2019, pp. 52–57, 2019
- [8] Rønningstad, E., “Targeted sentiment analysis for norwegian text,” 2020.
- [9] Hochreiter, S. and Schmidhuber, J., “Long short-term memory,” Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997.
- [10] Chen, P., Sun, Z., Bing, L., and Yang, W., “Recurrent attention network on memory for aspect sentiment analysis,” in Proceedings of the 2017 conference on empirical methods in natural language processing, 2017, pp. 452–461.
- [11] Dong, L., Wei, F., Tan, C., Tang, D., Zhou, M., and Xu, K., “Adaptive recursive neural network for target-dependent twitter sentiment classification,” in Proceedings of the 52nd annual meeting of the association for computational linguistics (volume 2: Short papers), 2014, pp. 49–54.
- [12] Xue, W. and Li, T., “Aspect based sentiment analysis with gated convolutional networks,” arXiv preprint arXiv:1805.07043, 2018.
- [13] Huang, B. and Carley, K. M., “Parameterized convolutional neural networks for aspect level sentiment classification,” arXiv preprint arXiv:1909.06276, 2019.

- [14] Nergiz, G., Safali, Y., Avarođlu, E., and Erdođan, S., "Classification of turkish news content by deep learning based lstm using fasttext model," in 2019 International Artificial Intelligence and Data Processing Symposium (IDAP). IEEE, 2019, pp. 1–6.
- [15] Dogru, H. B., Tilki, S., Jamil, A., and Hameed, A. A., "Deep learning-based classification of news texts using doc2vec model," in 2021 1st International Conference on Artificial Intelligence and Data Analytics(CAIDA). IEEE, 2021, pp. 91–96.
- [16] Gonz'alez-Carvajal, S. and Garrido-Merch'an, E. C., "Comparing bertagainst traditional machine learning text classification," arXiv preprintarXiv:2005.13012, 2020.
- [17] Gao, Z., Feng, A., Song, X., and Wu, X., "Target-dependent sentimentclassification with bert," IEEE Access, vol. 7, pp. 154290–154299, 2019.
- [18] Ozdil, U., Arslan, B., Tařar, D. E., Polat, G., and Ozan, ř., "Adtext classification with bidirectional encoder representations," in 2021 6th International Conference on Computer Science and Engineering (UBMK), 2021, pp. 1–6.
- [19] Chang, M., Ratinov, L., Roth, D., and Srikumar, V., "Importance of semantic representation: Dataless classification," in AAAI-08/IAAI-08 Proceedings - 23rd AAAI Conference on Artificial Intelligence and the 20th Innovative Applications of Artificial Intelligence Conference, 2008, vol. 2, pp. 830–835, 23rd AAAI Conference on Artificial Intelligence and the 20th Innovative Applications of Artificial Intelligence Conference, AAAI-08/IAAI-08 ; Conference date: 13-07-2008 Through 17-07-2008.
- [20] Pushp, P. K. and Srivastava, M. M., "Train once, test anywhere: Zero-shot learning for text classification," CoRR, vol. Abs/1712.05972, 2017.
- [21] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J., "Distributed representations of words and phrases and their compositionality," in Advances in Neural Information Processing Systems, Burges, C. J. C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K. Q., Eds. 2013, vol. 26, Curran Associates, Inc.
- [22] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I., "Attention is all you need," arXivpreprint arXiv:1706.03762, 2017.
- [23] Gupta, Shashij, et al. "Machine translation testing via pathological invariance." Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering. 2020.
- [24] Do, Quang-Minh, Kungan Zeng, and Incheon Paik. "Resolving Lexical Ambiguity in English-Japanese Neural Machine Translation." 2020 3rd Artificial Intelligence and Cloud Computing Conference. 2020.
- [25] Loodos., "loodos/bert-base-turkish-uncased · hugging face," <https://github.com/Loodos/turkish-language-models>, Aug. 2020.
- [26] Tařar, D. Emre, et al. "Auto-tagging of Short Conversational Sentences using Transformer Methods." arXiv preprint arXiv:2106.01735 (2021).
- [27] Adresgezgini. (n.d.). Adresgezgini/ICABMWZSL. GitHub. Retrieved November 11, 2021, from <https://github.com/adresgezgini/ICABMWZSL>.