

Türkçe Müşteri Yorumları Üzerinden Duygu Analizi: Metin Madenciliği ve Makine Öğrenimi Yaklaşımı

Furkan SİDAL^{1*}, Doç. Dr. Yener ALTUN²

¹İstatistik/ Fen Bilimleri Enstitüsü, Yüzüncü Yıl Üniversitesi, Türkiye

²İstatistik/ Fen Bilimleri Enstitüsü, Yüzüncü Yıl Üniversitesi, Türkiye

*(furkansidal@gmail.com) Başlıca yazarın mail adresi

(Received: 03 December 2024, Accepted: 06 December 2024)

(3rd International Conference on Recent Academic Studies ICRAS 2024, December 03-04, 2024)

ATIF/REFERENCE: Sidal, F. & Altun, Y. (2024). Türkçe Müşteri Yorumları Üzerinden Duygu Analizi: Metin Madenciliği ve Makine Öğrenimi Yaklaşımı. *International Journal of Advanced Natural Sciences and Engineering Researches*, 8(11), 238-244.

Özet – Bu çalışma, Türkçe müşteri yorumlarının duygu analizi için bir çerçeve sunmaktadır. Hepsiburada platformundan alınan veri seti üzerinde metin madenciliği ve doğal dil işleme teknikleri uygulanmıştır. Metinler temizlenmiş, kelime köklerine indirgenmiş ve Terim Frekans-Ters Doküman Frekans (TF-IDF) yöntemiyle sayısal formata dönüştürülmüştür. Analiz için Multinomial Naive Bayes (MNB) algoritması kullanılmış ve %83.5 doğruluk oranına ulaşılmıştır. Elde edilen sonuçlar, Türkçe doğal dil işleme ve metin analizi alanında önemli katkılar sağlamaktadır.

Anahtar Kelimeler-Duygu Analizi, Türkçe Doğal Dil İşleme, TF-IDF, Multinomial Naive Bayes, Metin Madenciliği, Müşteri Yorumları.

I. GİRİŞ

Dijital çağda kullanıcı yorumları hem bireysel hem de kurumsal karar verme süreçlerinde önemli bir veri kaynağı haline gelmiştir. E-ticaret platformları, sosyal medya ve kullanıcı odaklı hizmet sunan dijital ortamlar, kullanıcıların ürün ve hizmetlerle ilgili düşüncelerini paylaşmalarına olanak tanır. Bu yorumların duygu analizi yoluyla incelenmesi, işletmelerin müşteri memnuniyetini artırmak, ürün geliştirme stratejileri oluşturmak ve rekabet avantajı sağlamak için değerli bilgiler sunar.

Ancak, duygu analizi süreci, dilin yapısına bağlı olarak çeşitli zorluklarla karşılaşabilir. Türkçe, zengin morfolojik yapısı ve eklemeli dil özellikleri nedeniyle metin işleme teknikleri açısından ek zorluklar barındırır. Kelimelerin farklı bağlamlarda farklı anlamlar taşıması, özellikle duygu analizi için önemli bir engel teşkil eder. Örneğin, "hızlı" kelimesi, olumlu bir bağlamda olumlu bir anlam taşıırken, gecikmeyle ilgili bir yorumda tarafsız veya olumsuz olarak algılanabilir.

Bu çalışmada, Türkçe müşteri yorumlarının duygu analizi için özel bir çerçeve geliştirilmiştir. Çalışmanın amacı, Türkçe'nin dil özelliklerine uygun bir metin işleme ve sınıflandırma süreci tasarlayarak, duygu analizi sonuçlarının doğruluğunu artırmaktır. Bunun için, Hepsiburada platformundan alınan gerçek kullanıcı yorumları üzerinde metin madenciliği teknikleri ve makine öğrenimi algoritmaları uygulanmıştır.

Çalışmanın amacı, Türkçe duygu analizi üzerine yapılan literatüre önemli katkılar sağlamaktır. Ayrıca, müşteri yorumlarının etkili bir şekilde analiz edilmesi, işletmelerin daha hızlı ve doğru kararlar almasına olanak tanımaktadır. Gelecekte, daha geniş veri setleri ve gelişmiş modellerle bu alanda yapılan çalışmaların genişletilmesi hedeflenmektedir.

Duygu analizi, metin madenciliği ve doğal dil işleme (NLP) alanında hızla gelişen bir araştırma konusu olmuştur. Kullanıcı yorumlarının otomatik olarak analizi, müşteri davranışlarını anlamak ve ürün geliştirme süreçlerini optimize etmek için önemli bir araçtır ([1]-[3]).

Duygu analizi, metinlerin duygusal tonlarını belirlemek amacıyla sınıflandırılmasını hedefler ve bu bağlamda çeşitli uygulama alanlarına sahiptir [1]. Liu, duygu analizini sistematik bir çerçeve içinde ele alarak, metinlerin olumlu, olumsuz veya tarafsız olarak nasıl sınıflandırılacağını açıklamıştır [2]. Ayrıca Pang ve Lee, özellikle film incelemeleri gibi yapılandırılmamış metinlerin duygu analizine odaklanmıştır [3].

Türkçe, eklemeli yapısı ve morfolojik zenginliği nedeniyle doğal dil işleme süreçlerinde farklı zorluklar yaratır [4]. Çöltekin, Türkçe'nin dil yapısından kaynaklanan bu zorluklara dikkat çekmiş ve morfolojik analiz araçlarının önemini vurgulamıştır [5]. Türkçe tweetlerin duygu analizi üzerine yapılan bir başka çalışmada, TF-IDF ve Naive Bayes algoritması birleştirilerek %82 doğruluk oranına ulaşılmıştır [6]. Bu çalışmalar, Türkçe metinlerde duygu analizi için temel bir çerçeve sunmuş ancak daha gelişmiş yöntemlerin gerekliliğini de ortaya koymuştur [7].

Makine öğrenimi, duygu analizi çalışmalarında sıklıkla kullanılan yöntemlerin başında gelir. Multinomial Naive Bayes (MNB) gibi klasik algoritmalar, küçük veri setlerinde etkili çözümler sunarken, derin öğrenme tabanlı yöntemler, büyük veri setlerinde daha yüksek başarı oranlarına ulaşmaktadır [8]. BERT (Bidirectional Encoder Representations from Transformers), bağlama dayalı anlam çözümlemede önemli bir ilerleme sağlamış ve özellikle Türkçe gibi eklemeli dillerde güçlü bir performans sergilemiştir [9].

II. MATERYAL VE YÖNTEM

A. Veri Seti

Bu çalışmada kullanılan veri seti, Hepsiburada platformundan alınmış ve toplamda 11.426 kullanıcı yorumundan oluşmaktadır. Yorumlar, olumlu, olumsuz ve tarafsız olmak üzere üç farklı kategoride etiketlenmiştir. Veri seti, kullanıcıların çevrim içi alışveriş deneyimlerini yansıtan gerçek yorumları içermekte ve duygu analizi için ideal bir yapı sunmaktadır.

Tablo 1. Veri Seti Özellikleri

Özellik	Değer
Toplam Yorum Sayısı	11.426
Olumlu Yorum Sayısı	4.252
Olumsuz Yorum Sayısı	4.237
Tarafsız Yorum Sayısı	2.937
Eksik Veriler	3 (Temizlendi)

B. Veri Ön İşleme

Veri ön işleme, NLP projelerinde metin verilerini analiz edilebilir ve modellenen bir forma dönüştürmek için kullanılan kritik bir adımdır. Bu süreç, verinin kalitesini artırarak modelin doğruluğunu ve genelleme yeteneğini geliştirir. Çalışmada kullanılan veri ön işleme adımları aşağıda detaylandırılmış ve açıklanmıştır.

i. Küçük Harfe Dönüştürme

Metinlerin analizinde, kelimelerin büyük-küçük harf farklılıkları dikkate alınmadan işlenmesi, dilin doğasına uygun bir genelleme sağlar. Bu nedenle tüm metinler küçük harfe dönüştürülmüştür (ör., "Kedi" "kedi"). Bu işlem, metin madenciliği süreçlerinde standart bir adımdır [1].

ii. Noktalama İşaretleri ve Sayıların Kaldırılması

Metinlerdeki noktalama işaretleri ve sayılar, duygu analizinde genellikle anlam taşımadıkları için kaldırılmıştır.

Tablo 2. Noktalama İşaretleri ve Sayıların Kaldırılması Örneği

Orijinal Metin	Düzenlenmiş Metin
Bu ürün harika!!! 5 yıldız.	Bu ürün harika

Bu işlem, modelin yalnızca anlamlı özelliklere odaklanmasını sağlar ve metinleri daha yalın hale getirir [4].

iii. Durak Kelimelerin (Stopwords) Çıkarılması

Türkçe metinlerde sıkça kullanılan "ve", "bir", "ama", "de" gibi durak kelimeler, genellikle duygu analizi açısından anlam taşımaz. Bu kelimelerin çıkarılması, modelin önemli kelimelere odaklanmasına yardımcı olur [5].

Tablo 3. Durak Kelimelerin (Stopwords) Çıkarılması Örneği

Orijinal Metin	Düzenlenmiş Metin
Bu ürünü çok beğendim ve tavsiye ederim.	Bu ürünü beğendim tavsiye ederim.

iv. Kelimelerin Köklerine İndirgenmesi (Stemming ve Lemmatization)

Türkçe'nin eklemeli yapısı nedeniyle kelimeler, ekler aracılığıyla bağlama bağlı olarak farklı anlamlar taşıyabilir. Bu nedenle kelimeler, köklerine indirgenmiştir. Bu işlem için Zemberek-NLP gibi Türkçe doğal dil işleme araçları kullanılmıştır.

Tablo 4. Kelimelerin Köklerine İndirgenmesi (Stemming ve Lemmatization) Örneği

Orijinal Metin	İndirgenmiş Metin
"Beğeniyorum", "Beğenmiş", "Beğenecek"	Beğen

Tablo 5. Kelimelerin Köklerine İndirgenmesi (Stemming ve Lemmatization) Örneği

Adım	Açıklama	Sonuç
Orijinal Metin	İşlenmemiş haliyle kullanıcı yorumu.	"Bu ürün gerçekten mükemmel! Fiyatı uygun ama kargo çok geç geldi."
Küçük Harfe Dönüştürme	Metindeki tüm harfler küçük harfe dönüştürülür.	"bu ürün gerçekten mükemmel! fiyatı uygun ama kargo çok geç geldi."
Noktalama İşaretlerinin Kaldırılması	Noktalama işaretleri metinden çıkarılır.	"bu ürün gerçekten mükemmel fiyatı uygun ama kargo çok geç geldi"
Durak Kelimelerin Çıkarılması	"bu", "gerçekten", "ama" gibi anlam taşımayan durak kelimeler çıkarılır.	"ürün mükemmel fiyatı uygun kargo geç geldi"
Kelimelerin Köklerine İndirgenmesi	"geçti", "fiyatı" gibi kelimeler köklerine indirgenir.	"ürün mükemmel fiyat uygun kargo geç geldi"

Tablo 5, bir metin üzerinde gerçekleştirilen veri ön işleme adımlarını ve bu adımların metne etkisini detaylı bir şekilde göstermektedir. İşlenmemiş orijinal metin, duygu analizi için uygun bir formatta değildir ve analiz sırasında gereksiz bilgilere yol açabilir. Veri ön işleme işlemleri, metni sadeleştirerek anlamlı özelliklere odaklanır. Küçük harfe dönüştürme adımında, kelimelerin "Kargo" ve "kargo" gibi varyasyonlarını tek bir nesne olarak algılanmasını sağlar. Örneğin, farklı büyük/küçük harf kullanımları model performansını olumsuz etkileyebilir. Noktalama işaretleri genellikle kelime tabanlı analizde önemli değildir. Ancak, duygu analizi için kullanılan bazı modeller (ör. sinir ağları), noktalama işaretlerini

anlamı belirlemek için kullanabilir. Bu durum, analizin bağlamına bağlıdır. Durak kelimelerinin çıkarılması adımı, "Bu", "çok", "ama" gibi kelimeler duygu taşımadığı için çıkarılır. Ancak, bağlam analizi gereken durumlarda durak kelimeler önem taşıyabilir. Kelimelerin köklerine indirgenmesi işlemi, Türkçe gibi eklemeli dillerde özellikle önemlidir. Örneğin, "kargo" kelimesinin "kargolar", "kargonun", "kargoya" gibi varyasyonları, tek bir kök kelime "kargo" olarak işlenir. Bu, veri boyutunu küçültür ve modelin genelleme yeteneğini artırır.

C. Özellik Çıkarımı

Metinlerin sayısal verilere dönüştürülmesi, metin madenciliği ve NLP süreçlerinin temel adımlarından biridir. Bu çalışmada, metin verilerini modellemeye uygun hale getirmek için TF-IDF (Term Frequency-Inverse Document Frequency) yöntemi kullanılmıştır. TF-IDF, metindeki kelimelerin önemini belirlemek için kullanılan yaygın ve etkili bir ağırlıklandırma yöntemidir [4].

TF, bir kelimenin bir belgede ne kadar sık geçtiğini ölçer. Kelimenin bir belgedeki toplam kelime sayısına oranıyla hesaplanır.

$$TF(t) = \frac{\text{Bir belgedeki kelimesinin sayısı}}{\text{Belgedeki toplam kelime sayısı}} \quad (1)$$

Formül 1, kelimenin o belgede ne kadar önemli olduğunu değerlendirirken, sık kullanılan kelimelere daha fazla ağırlık verir [8].

$$IDF(t) = \log \left(\frac{\text{Toplambelge sayısı}}{\text{Kelimeyi içeren belge sayısı} + 1} \right) \quad (2)$$

Formül 2'de, kelimenin genel veri setinde ne kadar bilgi taşıdığını ölçülmektedir [10].

TF-IDF, TF ve IDF'nin çarpımı ile hesaplanır ve bir kelimenin hem belgede hem de tüm veri setinde ne kadar önemli olduğunu değerlendirir.

Bu çalışmada TF-IDF, metinlerin sayısal verilere dönüştürülmesinde kullanılmıştır. Örneğin, "güzel" kelimesi, hem sık kullanıldığı belgelerde hem de nadir bulunduğu belgelerde yüksek bir TF-IDF skoruna sahip olmuştur. Bu da, kelimenin duygu analizi açısından bilgi taşıdığını göstermektedir.

D. Sınıflandırma Algoritması

Bu çalışmada, metin sınıflandırma için MNB algoritması kullanılmıştır. Bu algoritma, özellikle metin sınıflandırma ve duygu analizi gibi NLP problemlerinde yaygın olarak tercih edilen bir yöntemdir. Naive Bayes algoritması, olasılık tabanlı bir model olup, bir sınıfa ait olma olasılığını hesaplamak için Bayes Teoremi'ni kullanır [11].

MNB, metin sınıflandırma problemlerinde kelime sayım verilerini modellemek için geliştirilmiştir. Bu algoritma, her kelimenin bir belgeye ait olma olasılığını değerlendirir ve sınıflar arasında maksimum olasılığı seçerek sınıflandırma yapar. MNB, metindeki kelimelerin frekanslarını göz önüne alır ve bu frekansların sınıf üzerindeki etkilerini hesaplar. Örneğin, "memnun" kelimesi olumlu yorumlarda sıkça geçiyorsa, bu kelimenin ağırlığı sınıflandırmada artırılır [12].

E. Performans Değerlendirmesi

Performans değerlendirme, sınıflandırma algoritmalarının başarılarını ölçmek için kullanılan bir dizi formülü içerir. Bu formüller, modelin doğruluğunu, duyarlılığını ve genel etkinliğini değerlendirerek sonuçların güvenilirliğini sağlar [11]. Çalışmada kullanılan temel performans hesaplamaları aşağıda açıklanmıştır.

i. Precision (Kesinlik)

Precision, modelin pozitif olarak tahmin ettiği sonuçların ne kadarının doğru olduğunu ölçer. Yanlış pozitif tahminlerin etkisini değerlendirmek için kritik bir hesaplama aracıdır [12].

ii. Recall (Duyarlılık)

Recall, gerçek pozitif örneklerin model tarafından ne kadarının doğru tahmin edildiğini ölçer. Modelin, tüm pozitif örnekleri ne derece iyi yakaladığını gösterir [4].

iii. F1-Score

F1-Score, Precision ve Recall metriklerinin harmonik ortalamasıdır. Modelin dengeli bir performans sergileyip sergilemediğini değerlendirmek için kullanılır [10].

iv. Doğruluk (Accuracy)

Doğruluk, modelin doğru tahmin ettiği örneklerin toplam örnek sayısına oranıdır [10].

III. BULGULAR

Bu çalışmada, Türkçe müşteri yorumlarının duygu analizi için geliştirilmiş modelin performansı kapsamlı bir şekilde değerlendirilmiştir. Bulgular, hem modelin genel başarısını hem de sınıf bazında analiz sonuçlarını içermektedir.

A. Model Performansı

Tablo 6. Model Performansı (Precision, Recall, F1-Score)

Kategori	Precision	Recall	F1-Score
Olumlu	0.85	0.82	0.84
Olumsuz	0.82	0.85	0.83
Tarafsız	0.78	0.75	0.76

Modelin olumlu, olumsuz ve tarafsız yorumlar için sınıf bazında performans sonuçları Tablo 6'da özetlenmiştir. Olumlu yorumlarda, Precision (0.85) ve Recall (0.82) değerleri, modelin bu sınıfı doğru bir şekilde sınıflandırma becerisini göstermektedir. F1-Score değeri (0.84), olumlu sınıf için dengeli bir performans ifade etmektedir. Olumsuz yorumlarda, Recall (0.85) ve Precision (0.82) değerleri, modelin olumsuz sınıfları doğru bir şekilde ayırt edebildiğini ortaya koymaktadır. Precision (0.78) ve Recall (0.75) değerlerinin diğer sınıflara göre daha düşük olması, modelin tarafsız sınıfta sınıflandırma zorluğu yaşadığını göstermektedir. Bu durum, tarafsız yorumların düşük duygu yoğunluğu içermesi ve bağlam analizinin eksikliğinden kaynaklanabilir.

B. Yanlış Sınıflandırma Analizi

Tablo 7'de, modelin yanlış sınıflandırdığı örnekler ve bu sınıflandırmaların olası nedenleri gösterilmiştir.

Tablo 7. Yanlış Sınıflandırma Analizi Sonuçları

Yanlış Sınıflandırma Türü	Örnek Yorum	Tahmin	Gerçek	Olası Neden
Yanlış Pozitif (False Positive)	"Fiyat uygun ama kargo çok yavaş."	Olumlu	Olumsuz	Yorumda geçen "fiyat uygun" gibi olumlu ifadeler model tarafından daha baskın değerlendirilmiştir.
Yanlış Negatif (False Negative)	"Ürün beklediğimden daha iyi çıktı."	Tarafsız	Olumlu	Düşük duygu yoğunluğuna sahip ifadeler modelin tarafsız sınıfa yönelmesine neden olmuştur.
Yanlış Pozitif (False Positive)	"Hızlı teslimat, fakat ürün kalitesi kötü."	Olumlu	Olumsuz	Teslimatla ilgili olumlu ifade, modelin yorumu olumlu olarak sınıflandırmasına yol açmıştır.
Yanlış Negatif (False Negative)	"Beklentilerimi karşıladı, güzel ürün."	Tarafsız	Olumlu	Tarafsız ifadelerle başlayan yorumlar, modelin sınıflandırmasında kararsızlığa neden olmuştur.
Yanlış Pozitif (False Positive)	"Memnun kalmadım, ama ambalaj sağlamdı."	Olumlu	Olumsuz	"Ambalaj sağlamdı" gibi olumlu bir özellik, modelin bu özelliğe daha fazla önem vermesine neden olmuştur.

Yanlış pozitifler, genellikle olumsuz yorumlarda yer alan olumlu ifadelerin, model tarafından yanlış bir şekilde olumlu olarak sınıflandırılmasına neden olmuştur. Yanlış negatifler ise tarafsız olarak tahmin edilen olumlu yorumlarda, düşük duygu yoğunluğu veya bağlama dayalı ifadeler modelin performansını sınırlamıştır.

C. Confusion Matrix Analizi

Modelin doğru ve yanlış sınıflandırmaları, Confusion Matrix ile değerlendirilerek sonuçlar Tablo 8'de gösterilmiştir. Olumlu ve olumsuz sınıflarda doğru tahmin oranı yüksek bulunmuş, bu sınıflarda düşük hata oranları gözlenmiştir. Tarafsız sınıfta, modelin yanlış sınıflandırma oranı diğer sınıflara kıyasla daha yüksektir. Bu durum, tarafsız yorumların daha düşük bir duygu yoğunluğuna sahip olması ile ilişkilendirilebilir.

Tablo 8. Confusion Matrix Analizi Sonuçları

Gerçek Sınıf	Olumlu (Tahmin)	Olumsuz (Tahmin)	Tarafsız (Tahmin)
Olumlu	3.520	102	630
Olumsuz	95	3.615	527
Tarafsız	637	425	1.875

IV. TARTIŞMA

Bu çalışmada, Türkçe müşteri yorumlarının duygu analizi için TF-IDF ve MNB algoritmalarının kullanıldığı bir sınıflandırma modeli geliştirilmiştir. Çalışma kapsamında elde edilen sonuçlar, literatürdeki diğer çalışmalarla karşılaştırılarak değerlendirilmiş, güçlü ve zayıf yönler ele alınmış ve gelecekteki çalışmalara yönelik öneriler sunulmuştur.

A. Çalışmanın Güçlü Yönleri

Çalışmanın en önemli güçlü yönlerinden biri, Türkçe'nin eklemeli yapısına uygun bir veri ön işleme süreci tasarlanmasıdır. Metinlerin köklerine indirgenmesi ve TF-IDF yöntemi ile sayısal verilere dönüştürülmesi, sınıflandırma modelinin genelleme yeteneğini artırmıştır. Literatürde, bu tür ön işleme adımlarının metin madenciliği projelerinde başarı oranını artırdığı belirtilmiştir. Ayrıca, MNB algoritmasının hesaplama açısından verimli olması, modelin hızlı bir şekilde eğitilmesini ve test edilmesini sağlamıştır.

B. Performans Değerlendirmesi

Model, genel olarak olumlu ve olumsuz yorumların sınıflandırılmasında başarılı sonuçlar elde etmiştir. Doğruluk oranı (%83.5), literatürde benzer yöntemlerle yapılan çalışmalara kıyasla rekabetçi bir performans sergilemiştir. Precision, Recall ve F1-Score değerleri, olumlu ve olumsuz sınıflarda dengeli bir performansı ortaya koymuştur. Ancak tarafsız sınıfta, düşük Precision (0.78) ve Recall (0.75) değerleri, tarafsız yorumların sınıflandırılmasının diğer sınıflara göre daha zor olduğunu göstermektedir.

C. Sınırlamalar

Çalışmanın bazı sınırlamaları, modelin geliştirilmesi gereken yönlerini ortaya koymaktadır. Naive Bayes algoritması, kelimelerin bağımsız olduğunu varsayar. Ancak Türkçe gibi bağlam odaklı dillerde bu varsayım, modelin performansını sınırlayabilir. Veri setindeki sınıflar arasında küçük bir dengesizlik bulunmakta olup, bu durum bazı sınıfların performansını olumsuz yönde etkileyebilir. Tarafsız yorumlar genellikle düşük duygu yoğunluğuna sahip olduğundan, modelin bu sınıfta düşük performans göstermesine neden olmuştur. Yanlış sınıflandırmalar, tarafsız yorumların genellikle olumlu veya olumsuz ifadeler içermesiyle ilişkilidir.

V. SONUÇLAR

Bu çalışma, Türkçe müşteri yorumlarının duygu analizi için TF-IDF özellik çıkarımı ve MNB algoritmasını kullanarak bir sınıflandırma modeli geliştirmiştir. Çalışma kapsamında elde edilen bulgular ve analizler, modelin genel başarısını ve sınıf bazındaki performansını değerlendirme açısından önemli bilgiler sunmaktadır.

Veri ön işleme sürecinde kullanılan kelime köklerine indirgeme ve TF-IDF yöntemleri, Türkçe'nin eklemeli yapısına uygun şekilde modellenmiştir. Bu, sınıflandırma başarısına önemli bir katkı sağlamıştır. Model, %83.5 doğruluk oranı ile olumlu, olumsuz ve tarafsız yorumları etkili bir şekilde sınıflandırmıştır. Precision, Recall ve F1-Score gibi metrikler, modelin duygu analizi görevini başarıyla gerçekleştirdiğini göstermektedir.

Çalışmanın sonuçları, Türkçe metinlerde yapılan diğer çalışmalarla uyumlu bulunmuş ve TF-IDF ile MNB algoritmalarının basit ama etkili bir yöntem olduğunu ortaya koymuştur.

Bu sonuçlar, Türkçe doğal dil işleme alanında hem akademik hem de pratik açıdan değerli bir katkı sağlamaktadır.

KAYNAKLAR

- [1] B. Liu, *Sentiment Analysis and Opinion Mining*, 1st ed., Morgan & Claypool, 2012.
- [2] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis", *Foundations and Trends in Information Retrieval*, vol. 2, pp. 1–135, 2008.
- [3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proc. NAACL-HLT*, 2018, pp. 4171–4186.
- [4] C. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*, 1st ed., Cambridge University Press, 2008.
- [5] Ç. Çöltekin, "Morphological Challenges in Turkish NLP," *Computational Linguistics*, vol. 47, no. 1, pp. 1–34, 2021.
- [6] İ. Bozkurt, et al., "Türkçe Tweet Duygu Analizi: TF-IDF ve Naive Bayes Kullanımı," in *Proc. NLP-Turkey'20*, 2020, pp. 89–97.
- [7] O. Eryılmaz, "Türkçe Metinlerde Duygu Analizi ve Performans Kıyaslaması", *Applied NLP Journal*, vol. 5, pp. 25–42, 2021.
- [8] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Information Processing & Management*, vol. 24, no. 5, pp. 513–523, 1988.
- [9] P. Domingos and M. Pazzani, "On the Optimality of the Simple Bayesian Classifier under Zero-One Loss," *Machine Learning*, vol. 29, pp. 103–130, 1997.
- [10] J. Ramos, "Using TF-IDF to Determine Word Relevance in Document Queries," in *Proc. First International Conference on Machine Learning*, 2003, pp. 133–142.
- [11] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, 1st ed., MIT Press, 2012.
- [12] A. McCallum and K. Nigam, "A Comparison of Event Models for Naive Bayes Text Classification," in *Proc. AAAI Workshop on Learning for Text Categorization*, 1998.