

Yapay Sinir Ağı Tabanlı Sınıflandırıcılarda Nitelik Seçiminin Siber Zorbalık Tespitine Etkileri

The Effects of Attribute Selection in Artificial Neural Network Based Classifiers on Cyberbullying Detection

Eren Çürük

Dept. of Electrical-Electronics Engineering
Mersin University
Mersin, Turkey
erencuruk@gmail.com

Çiğdem Acı

Dept. of Computer Engineering
Mersin University
Mersin, Turkey
caci@mersin.edu.tr

Esra Saraç Eşsiz

Dept. of Computer Engineering
Adana Science and Technology University
Adana, Turkey
esarac@adanabtu.edu.tr

Öz— Son zamanlarda, bilgi iletişim teknolojilerinin hızla artması sonucunda akıllı telefonlar, tabletler ve dizüstü bilgisayarların kullanımı yaygınlaşmıştır. Özellikle gençler arasında sosyal ağların günlük yaşamın bir parçası haline gelmesi ve siber uzayda kişilerin kimlik bilgilerini gizleyerek her ortama erişebilmeleri sonucunda siber zorbalık problemi ortaya çıkmıştır. Bu çalışmada sosyal ağlarda meydana gelen siber zorbalığın tespit edilmesine yönelik analizler gerçekleştirilmiştir. Youtube yorumlarından elde edilmiş toplam 3469 yorum siber zorbalık barındırıp barındırmamasına göre pozitif ve negatif olarak etiketlenmiştir. Analizlerde nitelik seçme algoritmalarından maksimum uygunluk minimum fazlalık (MRMR), ReliefF ve yinelemeli özellik eliminasyonu (RFE) algoritmaları kullanılarak veri kümesinin nitelik sayısı 10, 50, 100, 250 ve 500 olarak azaltılmıştır. Sınıflandırma algoritmaları olarak destek vektör makinaları (DVM), stokastik dereceli azalma (SGD), radyal tabanlı fonksiyon (RBF) ve lojistik regresyon (LR) tercih edilmiştir. Deneysel çalışmalar sonucunda RFE nitelik seçme algoritmasıyla birlikte SGD sınıflandırıcısının kullanılması sonucunda 0.943'lük bir F-ölçütü değeri elde edilmiştir. Diğer nitelik seçme algoritmaları da RFE kadar yüksek değerler oluşturamasa da yaklaşık 0.76 ila 0.84 arası F-ölçütü değerleri elde edilmiştir.

Anahtar Sözcükler—Siber Zorbalık, Sınıflandırma, Metin Madenciliği, MRMR, ReliefF, RFE, DVM, RBF, SGD, Lojistik Regresyon, Sosyal Medya

Abstract— Recently, as a result of the rapid increase of information communication technologies, the use of smartphones, tablets and laptop computers has become widespread. Especially among young people, social networks have become a part of everyday life and the cyberbullying problem has arisen as a result of hiding the credentials of people in cyberspace and reaching every level. In this study was carried out assays for the detection of cyberbullying. A total of 3469 reviews from Youtube were tagged as positive and negative, depending on whether have contained bullying. In the analyzes, the number of features of the data set was reduced to 10, 50, 100, 250 and 500 using the minimum redundancy and maximum relevance (MRMR), ReliefF and recursive feature elimination (RFE) algorithms as

feature selection algorithms. support vector machines (SVM), stochastic gradient descent (SGD), radial basis function (RBF) and logistic regression (LR) have been preferred as classification algorithms. As a result of the experimental studies, the use of the SGD classifier together with the RFE attribute selection algorithm resulted 0.943 F-measure value. Other quality selection algorithms did not produce as high values as RFE, but F-measure values of about 0.76 to 0.84 were obtained.

Keywords—Cyberbullying, Classification, Text Mining, MRMR, ReliefF, RFE, DVM, RBF, SGD, Logistic Regression, Social Media

I. GİRİŞ

Bilgi ve iletişim teknolojilerinin artmasıyla birlikte bireylerin sosyal platformlara erişimi, telefon, tablet bilgisayar gibi iletişim araçlarını kullanarak her geçen gün kolaylaşmaktadır. Bu araçların kullanım amaçları her ne olursa olsun, kullanım oranları incelendiğinde şaşırtıcı sonuçlar görülmektedir. İngiltere'de 2016 yılında yapılan araştırmada, 3 ila 4 yaş arasındaki çocukların %55'inin bir tablet ve %24'ünün masaüstü veya dizüstü bilgisayar kullandığı sonucuna ulaşılmıştır. 5 ila 7 yaşındakiler arasında tablet kullanımı %67, masaüstü veya dizüstü bilgisayar kullanımı %49 ve akıllı telefon kullanımı %2 olarak değerlendirilmiştir. 8 ila 11 yaşındakiler arasında ise, tablet kullanımı % 80, masaüstü veya dizüstü bilgisayar kullanımı % 66 ve akıllı telefon kullanımı % 32 olarak rapor edilmiştir [1]. Amerika Birleşik Devletleri'nde, ilkökul öğrencileri arasında da bilgi iletişim teknolojilerini kullanım istatistikleri benzerdir. 2015 yılında ABD'de yapılan araştırmada, ilkökul çağındaki gençlerin %66'sı dizüstü bilgisayar, %53'ü akıllı telefon ve %78'i tablet kullanmaktadır. Ayrıca, bu gençlerin %35'i bir dizüstü bilgisayara sahip olduğunu, %36'sı bir akıllı telefona sahip olduğunu ve %69'u ise bir tablete sahip olduğunu söylemektedir [2]. Yapılan bu çalışmalar teknolojik iletişim araçlarının gençler arasında ne kadar yaygın olarak kullanıldığını göstermektedir. Bu artışla birlikte siber zorbalığın gençler arasında artışı paralellik

göstermekte ve her geçen gün siber zorbalığın olumsuz etkileri artarak devam etmektedir.

Siber zorbalık eylemleri farklı şekillerde gerçekleştirilebilmektedir. Siber zorbalıklar bazen mağdurların kişisel verilerini (kişisel verilerimiz ev ya da işyeri adresleri, gerçek isimler ve eğitim yerleri gibi bilgileri kapsar) kamuya açık olarak toplamak ve dağıtmak için forumları ya da web sitelerini kullanabilmektedir [3]. Bu bilgilerle gerçek olmayan hesaplar yaratarak yorumlar yapabilmekte ya da web siteleri oluşturabilmektedirler. Bu eylemlerle hedeflerini olumsuz yönde göstermekte veya insanların mağdurlara karşı saygılarını kaybetmelerine neden olmaktadır. Bunun yanı sıra siber zorbalık mesajları ve e-posta gibi yöntemlerle de gerçekleştirilebilmekte ve mağdurlara tehdit, taciz veya dedikodu gibi konularla zarar verilebilmektedir.

Siber zorbalık tespiti gibi metin analizine dayalı çalışmalar incelendiğinde en önemli problemlerden biri de veri kümesinin büyük olmasından dolayı nitelik alanının yüksek boyutlara sahip olmasıdır [4]. Bu sebepten dolayı, nitelik seçimi algoritmaları veri kümelerinin boyutunu azaltmak için yaygın olarak kullanılmaktadır. Bazı durumlarda tam veri kümesi ile çalışmak imkânsız olmakta ve nitelik alanının yüksek boyutluluğu, metin sınıflandırmasının temel sorunu haline gelmektedir. Nitelik seçimi algoritmaları ile en iyi nitelik alt kümesi, en yüksek sınıflandırma doğruluğu ve verimliliğe sahip en az sayıda nitelik içerecek şekilde uygulanmaktadır. Bu çalışmada nitelik seçiminin siber zorbalık tespitinde üzerine etkilerinin araştırılması amacıyla farklı nitelik seçimi algoritmaları kullanılmıştır. Yapılan çalışmalar incelendiğinde ki-kare testi (Chi-2) ve bilgi kazancı (IG) yöntemleri metin sınıflandırma problemlerinde en sık kullanılan nitelik seçimi algoritmaları olarak görülmektedir [4], [5], [6], [7]. Yaptığımız bu çalışmada ise ReliefF, maksimum uygunluk minimum fazlalık (MRMR) ve yinelemeli özellik eliminasyonu (RFE) nitelik seçimi algoritmaları kullanılarak, yapay sinir ağı (YSA) tabanlı, radyal tabanlı fonksiyon (RBF), stokastik dereceli azalma (SGD), lojistik regresyon (LR), destek vektör makinaları (DVM) sınıflandırıcıları ile başarımlar analizleri gerçekleştirilmiştir.

Bu çalışmanın ilerleyen bölümleri aşağıdaki şekilde düzenlenmiştir: ikinci bölümde konuyla ilgili literatürde yapılmış çalışmaların kısa birer özeti sunulmuştur. Üçüncü kısımda, kullanılan veri kümesi ve uygulanan yöntemler anlatılmıştır. Dördüncü bölümde ise deneysel sonuçlar ve tartışma olarak çalışmamız tamamlanmıştır.

II. LİTERATÜR ÖZETİ

Siber zorbalık tespiti üzerine yapılan çalışmalar incelendiğinde J48 karar ağacı, naive bayes (NB), DVM, ve K en yakın komşuluğu (K-NN) sınıflandırıcıları gibi sınıflandırıcılar kullanılmıştır [5], [8]–[10]. Ancak, çevrimiçi algılamaya çok az dikkat edilmiştir. [11]–[13]'daki çalışmalar, siber zorbalığı otomatik olarak algılamak için iyi bilinen metin madenciliği tekniklerini uygulayan başlıca örneklerdir. Siber zorbalığı otomatik olarak tespit etme, ön işleme, nitelik çıkarma, nitelik seçimi ve sınıflandırma gibi metin madenciliği yöntemlerini içerir.

Denetimli bir makine öğrenme yaklaşımının uygulandığı bir çalışmada, Youtube yorumları toplanmış, el ile etiketlenmiş ve cinsellik, fiziksel görünüm, zeka ve algılama olmak üzere üç farklı konu üzerine ikili ve çok sınıflı sınıflamalar uygulanmıştır. Yapılan test sonuçlarında ikili sınıflandırmalarda %80,2 doğruluk elde edilirken çok sınıflı testlerde %66,7 doğruluk elde edilmiştir [9].

Massachusetts Teknoloji Enstitüsünde, Youtube video yorumlarından metinsel bağlamda sanal zorbalığın tespitini gerçekleştiren bir sistem geliştirilmiştir. Sistem, yorumları cinsellik, kültür, zeka ve fiziksel özellikler gibi hassas konulardaki sınıflar dizisi olarak tanımlamaktadır ve hangi yorumun hangi sınıfa ait olduğunu tespit etmektedir [10]. Bad-of-Word yaklaşımını kullanarak, temel metin madenciliği sisteminin incelendiği bir çalışmada, duygular ve bağlamsal özellikler geliştirilerek oluşturulan bir modelde, % 61,9'luk bir doğruluk oranı elde edilmiştir [6]. Başka bir çalışmada, zorbalık izleri çeşitli doğal dil işleme teknikleri kullanılarak tanımlanmıştır. Çevrimiçi ve çevrimdışı zorbalık örnekleri incelenmiş ve zorbalığı türünü belirlemek için duygu analiz sistemi ve gizli dirichlet ayırımı (LDA) yöntemleri kullanılmıştır. Bu yöntemde, zorbalık örnekleri doğru bir şekilde saptanmamıştır [14].

Siber zorbalık tespitine yönelik başka bir çalışmada, Türkçe yazılmış sosyal medya mesajlarında siber zorbalığı saptamak amaçlanmıştır. Türkçe metinler üzerinde siber zorbalık tespiti yapılan çok az çalışmadan biri olarak literatürde yer almaktadır. Türkçe yazılan Instagram ve Twitter mesajlarından bir veri kümesi hazırlanmış ve daha sonra DVM, C4.5 Karar Ağacı, NB ve K-NN siber zorbalık tespiti için uygulanmıştır. Ayrıca, sınıflandırıcıların doğruluğunu artırmak için IG ve Chi-2 nitelik seçim yöntemleri de kullanılmıştır. Metin mesajlarındaki hem sözcükler hem de ifadeler özellik olarak dikkate alındığında, siber zorbalık tespitinin geliştiği gözlemlenmiştir. Sınıflandırıcılar arasında, NB hem sınıflandırma doğruluğu hem de çalışma süresi açısından en başarılı model olarak sonuç vermiş. Özellik seçimi uygulandığında, sınıflandırma doğruluğu kullanılan veri kümesi için %84'e kadar artmıştır [4].

Bu alanda yapılan ve Yapay Zeka Algoritmalarından Karınca Koloni Algoritması ile nitelik alt kümelerinin seçilmesine yönelik bir tez çalışmasında Formspring.me, Youtube, MySpace ve Twitter veri kümeleri üzerinde Karınca Koloni Algoritması ve Chi-2 algoritması birleştirilerek yapılan nitelik seçimi yönteminde sınıflandırma başarımlarının arttığı gözlemlenmiştir. Yapılan bu tez çalışması Yapay Zeka algoritmalarının IG ve Chi-2 gibi genel nitelik seçimi yöntemlerinden daha iyi sonuçlar verdiğini göstermiştir [5].

Literatürde yapılan çalışmalar incelendiğinde genellikle IG ve Chi-2 algoritmaları nitelik seçiminde kullanıldığı görülmektedir. Siber zorbalık tespiti üzerine yapılan çalışmalar incelendiğinde ReliefF, MRMR ve RFE nitelik seçimi algoritmaları ile yapılmış bir siber zorbalık tespiti analizine bulunmadığı görülmektedir. Bu nedenle ReliefF, MRMR ve RFE algoritmalarının, RBF, SGD, LR ve DVM sınıflandırıcılarının üzerine etkileri bu çalışmada analiz edilmiştir.

III. MATERYAL VE YÖNTEM

Bu çalışmada Youtube web sitesindeki yorumlardan elde edilmiş veri kümesinin siber zorbalık sınıflandırması YSA tabanlı DVM, LR, RBF ve SGD sınıflandırıcıları ile gerçekleştirilmiş ve nitelik seçiminde Relief, MRMR ve RFE nitelik seçimi algoritmaları kullanılmıştır.

A. Veri Kümesi

Yapılan çalışmalarda kullanılan Youtube veri kümesi 2013 yılında yapılmış bir çalışmada hazırlanmıştır [15]. Veri kümesi içerisinde yer alan Youtube yorumları siber zorbalık içerip içermemesi durumuna göre pozitif ve negatif olarak etiketlenmiştir. Tablo 1'de veri kümesi için siber zorbalık içeren ve içermeyen mesaj sayıları gösterilmiştir. Bu veri kümesi 3469 mesajdan oluşmaktadır. Deneysel çalışmalar sırasında veri kümesinde yer alan pozitif ve negatif mesajların dengelenmesi amacıyla pozitif örnekler dengeli bir veri kümesi oluşacak şekilde çoğaltılmıştır. Veri kümesini dengelemek amacıyla pozitif örnek sayısı negatif örnek sayısına yaklaşık olarak eşit oluncaya kadar pozitif örnekler tekrarlanarak veri kümesine eklenmiştir. Bu işlem için her bir pozitif örnek 8 defa tekrarlanmıştır. Veri kümesi içerisindeki mesajların %75'i eğitim verisi %25'i ise test verisi olarak ayrılmıştır.

TABLO I. VERİ KÜMESİ POZİTİF/NEGATİF MESAJ DAĞILIMI

Veri Kümesi	Pozitif Mesajlar	Negatif Mesajlar
Youtube	422	3047

B. Önışleme Ve Nitelik Çıkarımı

Bu çalışmada önışleme adımında nitelik çıkartımı için dizgilerin oluşturulması sağlanmıştır. Bu dizgiler veri kümesi içerisinde yer alan kelimeler, emoji olarak adlandırılan ifadeler ve diğer anlamlı ifadeleri belirtmektedir. Sınıflandırma başarımının artırılması amacıyla önışleme adımında ayrılan dizgiler içerisinde yer alan etkisiz kelimeler (stop-words) ayıklanmaktadır. Ayrıca çıkartılan niteliklerden doküman frekansı 0.001'in altında olanlar da ayıklanmış ve bu şekilde sınıflandırma başarımına etkisi çok düşük olan niteliklerden kurtularak sınıflandırma süresinin artırılması amaçlanmıştır. Ön işleme ve nitelik çıkarımı adımlarının ardından elde edilmiş nitelik sayısı 14215 nitelikten oluşmaktadır.

C. Nitelik Seçimi

Nitelik Seçimi, veriyi anlamada, hesaplama gereksinimini azaltmada, boyutsallık etkisini azaltmada ve tahmin edicinin başarımını geliştirmede yardımcı olur [16]. Nitelik seçiminin odak noktası, giriş verisini etkili bir şekilde gürültü ve alakasız değişkenleri azaltarak tanımlayabilen ve iyi tahmin sonuçları sağlayan, girdi değişkenlerinden oluşan bir alt küme seçmektir [17]. Nitelik seçimi yaklaşımları kabaca, filtre tabanlı yöntemler [18], sarmalayıcı yöntemler [19] ve gömülü yöntemler [20] olarak kategorize edilebilir. Filtre tabanlı yöntemler, değişken sıralama tekniklerini, değişken seçimi için temel kriter olarak kullanır. Sıralama yöntemleri sade bir yapıya sahiptir dolayısıyla pratik uygulamalar için başarılı bir yöntem olarak rapor edilmektedir. Değişkenleri derecelendirmek için uygun bir sıralama kriteri kullanılabilir ve bir

eşik değeri kullanarak, bu eşikğin altındaki değişkenleri kaldırır [21]. Sarmalayıcı yöntemler, algoritmaların yinelemeli doğası nedeniyle sıralı olarak adlandırılır. Sıralı Özellik Seçimi (SFS) algoritması boş bir setle başlar ve ilk adım için objektif fonksiyon için en yüksek değeri veren bir özellik ekler. İkinci adımdan itibaren kalan özellikler mevcut alt kümeye ayrı ayrı eklenir ve yeni alt kümeler değerlendirilir. Maksimum sınıflandırma doğruluğunu verirse, bireysel özellik, alt kümeye kalıcı olarak dahil edilir. Gerekli özellik sayısı eklenene kadar işlem tekrarlanır [22]. Gömülü yöntemler sarmalayıcı yöntemlerde yapılan farklı alt kümeleri yeniden sınıflandırmak için alınan hesaplama süresini azaltmak ister. Temel yaklaşım, özellik seçiminin eğitim sürecinin bir parçası olarak dahil edilmesidir [23].

MRMR algoritması, iki koşullu olarak çalışan, yani sınıf etiketleriyle maksimum ilişkiyi ve minimum fazlalığı (artıklığı) birleştiren filtre tabanlı bir özellik seçme yöntemidir [24]. Bu koşulları, uygunluk ve artıklık değerini elde etmek için karşılıklı bilgileri hesaplayarak, artımlı seçim şeması kullanılarak birleştirilir. MRMR, tüm özellik alanını en iyi temsil eden özelliklerin bir alt kümesini seçen bir diskriminant analiz yöntemidir. Bu yöntem, aralarındaki benzerlik düzeyini yansıtan özellik çiftleri arasındaki karşılıklı bilgilere dayanmaktadır. Karşılıklı bilgi, iki özelliğin paylaştığı bilgi miktarını ölçmeyi amaçlamaktadır. Bu nedenle, eğer iki özellik yüksek bir bilgi miktarına sahipse, bunların yüksek korelasyonlu olduğunu ve sonuç olarak bir diğerinin minimum bilgi kaybıyla yerini alabileceğini söylemektedir [25].

Relief algoritması yüksek verimliliğe sahiptir ve veri tiplerinin özelliklerini sınırlamamaktadır. Relief veri kümelerinin ayrık veya sürekli olarak analiz edilmesini sağlayabilir. Çok sınıflı problemlerle uğraşırken, Relief algoritması, farklı kategorilerdeki örneklerin her birinden en yakın komşu örneklerini seçer. Özellik ağırlıklandırma vektörü ile komşu örneklerinden sınıf içi mesafeyi ve sınıflar arası mesafeyi karşılaştırarak özelliklere daha fazla ağırlık vererek vektörü düzenler. Her bir özellik için aynı prosedür tekrarlanır, sonunda her bir özelliğin ağırlık değeri oluşturulmuş olur [26].

RFE algoritması, spesifik öğrenme modeline bağlı olan özellik seçimi için etkili bir algoritmadır [27]. [28]'de kanser sınıflandırması için RFE nitelik seçme algoritması kullanılarak DVM sınıflandırıcısı ile analizler gerçekleştirilmiştir. RFE, bir DVM modeli oluşturmak için tüm özellikleri kullanmakta ve daha sonra sıralı bir özellik listesi oluşturan DVM modelindeki her bir özelliğin katkısını sıralamaktadır. Son olarak da DVM modeline daha az katkıda bulunan anlamsız özellikleri ortadan kaldırarak nitelik seçme işlemini gerçekleştirmektedir.

Nitelik seçimi adımı tamamlandıktan sonra her bir niteliğin veri kümesi içerisindeki ağırlıklarının hesaplanması için terim frekansı (tf) ve ters belge frekansı (idf) kullanılarak Denklem 2'de verilen TF-IDF ağırlıklandırma yöntemi kullanılmıştır.

$$tf_idf_{t,d} = tf_{t,d} \times idf_{t,d} \quad (1)$$

TF-IDF ağırlıklandırma yöntemi kullanılarak elde edilen nitelik ağırlıkları ve veri kümesi içerisinde yer alan sınıf etiketinde belirtilen siber zorbalık barındırma durumu bilgileri ile birlikte bir arff dosyası halinde kaydedilmektedir. Bu dosyada yer alan

nitelik ağırlıkları ve sınıf türü sınıflandırılmada kullanılmak üzere WEKA veri madenciliği aracına aktarılmaktadır.

D. Sınıflandırma Ölçütü

Bu çalışmada siber zorbalık tespiti için WEKA veri madenciliği aracının YSA tabanlı sınıflandırıcılarından DVM, SGD, LR ve RBF sınıflandırıcıları kullanılmıştır.

Sınıflandırma başarımı Denklem 2’de verilen F-ölçütü değeri ile ölçülür.

$$F_ölçütü = \frac{2 \times \text{hassasiyet} \times \text{doğruluk}}{\text{hassasiyet} + \text{doğruluk}} \quad (2)$$

Doğruluk, sınıflandırma algoritmasının kesinliğidir ve pozitif sınıfa ait olarak doğru şekilde etiketlenmiş örneklerin sayısının, pozitif olarak etiketlenmiş toplam örnek sayısına oranıdır. Hassasiyet, sınıflandırma algoritmasının bütünlüğüdür ve gerçek pozitiflerin sayısının, aslında pozitif sınıfa ait olan elementlerin toplam sayısına oranıdır. Çalışmalarda kullanılan veri kümesi dengeli hale gerildiğinden, her iki sınıf için de F-ölçütü değerleri hesaplanmaktadır (yani, siber zorbalık varken ve siber zorbalık olmadan). Sınıflandırma başarımı gösterilirken bu iki F-ölçütü değerinin ortalaması alınarak gösterilmektedir.

IV. SONUÇLAR VE TARTIŞMA

Youtube veri kümesi üzerinde yapılan çalışmada, MRMR, ReliefF ve RFE nitelik seçimi algoritmaları kullanılarak, nitelik sayısı 10, 50, 100, 250 ve 500 olacak şekilde ayrı ayrı seçilme işlemleri gerçekleştirilmiştir. Tablo II, III ve IV incelendiğinde uygulanan nitelik seçimi algoritmalarının test verisi sonuçlarına göre başarımları karşılaştırmaları görülmektedir. Bu adımda gerçekleştirilen karşılaştırmalarda birden fazla sınıflandırma algoritması kullanılarak nitelik seçiminin etkilerinin daha net bir şekilde anlaşılması amaçlanmıştır. DVM, SGD, RBF ve LR sınıflandırıcılarının kullanıldığı çalışmada WEKA veri madenciliği aracı kullanılarak veri kümesinin eğitim ve test adımları sonucunda elde edilen F-ölçütü değerleri tablolarda gösterilmektedir.

TABLO II. MRMR TEST VERİSİ F-ÖLÇÜTÜ SINIFLANDIRMA BAŞARIMI

Nitelik Sayısı	DVM	SGD	RBF	LR
10	0.569	0.566	0.602	0.568
50	0.630	0.635	0.690	0.662
100	0.701	0.702	0.668	0.688
250	0.742	0.766	0.786	0.746
500	0.797	0.830	0.744	0.822

Elde edilen F-ölçütü değerleri incelendiğinde bütün nitelik seçimi algoritmaları için SGD sınıflandırıcısı en iyi sonucu vermiştir. ReliefF ve MRMR nitelik seçimi algoritmaları için

en yüksek F-ölçütü değerleri sırasıyla 0,830 ve 0.777 olurken RFE nitelik seçimi algoritmasında 0.943 olarak elde edilmiştir.

TABLO III. RELIEFF TEST VERİSİ F-ÖLÇÜTÜ SINIFLANDIRMA BAŞARIMI

Nitelik Sayısı	DVM	SGD	RBF	LR
10	0.355	0.369	0.558	0.488
50	0.553	0.561	0.642	0.586
100	0.624	0.627	0.689	0.638
250	0.705	0.719	0.737	0.712
500	0.755	0.777	0.776	0.768

TABLO IV. RFE TEST VERİSİ F-ÖLÇÜTÜ SINIFLANDIRMA BAŞARIMI

Nitelik Sayısı	DVM	SGD	RBF	LR
10	0.614	0.632	0.631	0.650
50	0.720	0.724	0.697	0.729
100	0.751	0.749	0.716	0.764
250	0.855	0.867	0.719	0.879
500	0.923	0.943	0.734	0.941

3469 Youtube yorumundan oluşan veri kümesi için uygulanan nitelik seçme algoritmalarının test verisi F-ölçütü değerlerine bakıldığında, en iyi sonuçları RFE nitelik seçme algoritmasında elde edildiğini görülmektedir. Uygulanan diğer nitelik seçme algoritmaları için sınıflandırma başarım ölçütümüz sırasıyla MRMR için ~0.8 ve üzeri, ReliefF için ~0.77 olarak hesaplandığı görülmektedir. Sınıflandırıcılar dikkate alınarak incelendiğinde RBF sınıflandırıcısı dışındaki sınıflandırıcılar genel olarak benzer sonuçlar vermiştir. RBF sınıflandırıcısı ise bu çalışma için diğerlerine göre daha düşük seviyelerde kaldığı görülmektedir.

TABLO V. incelediğinde en iyi sonucu veren RFE algoritması için 500 nitelik seçilerek yapılan sınıflandırmaların test verisi detaylı başarımları analizi görülmektedir. 500 nitelik için DVM, SGD, RBF ve LR sınıflandırıcılarının kesinlik, hassasiyet, gerçek pozitif, yanlış pozitif ve F-ölçütü değerleri gösterilmiştir. Tablo incelendiğinde pozitif ve negatif mesajlar olarak etiketlenmiş olan veri kümemizde her iki sınıf içinde bütün sınıflandırıcılarda birbirine yakın oranda gerçek pozitif, yanlış pozitif, hassasiyet ve F-ölçütü değerleri yakalandığı görülmektedir. Örneğin SGD sınıflandırıcısında negatif mesajlar yani siber zorbalık barındırmayan mesajlar için gerçek pozitif değeri 0.900 olarak hesaplanırken, pozitif mesajlar için 0.984, ağırlıklı ortalamaları ise 0.944 olarak hesaplanmıştır. Bu sonuçlar sınıflandırıcıların tek bir etiket üzerine ağırlıklı bir sınıflandırma gerçekleştirmediklerini, dengeli bir sınıflandırma algoritmasının yakalandığını göstermektedir.

TABLO VI'da ise RFE nitelik seçme algoritması ile seçilmiş 10, 50, 100, 250 ve 500 nitelik için yapılan sınıflandırmaların süre analizleri saniye cinsinden gösterilmektedir. Nitelik sayısının artması sınıflandırma süresinin de artmasına neden olduğu tablo üzerinden anlaşılabilir. Bu sonuçlar incelendiğinde sırasıyla DVM, SGD, RBF ve LR için en iyi sonuçların alındığı 500 nitelik için eğitim ve test toplam süreleri:

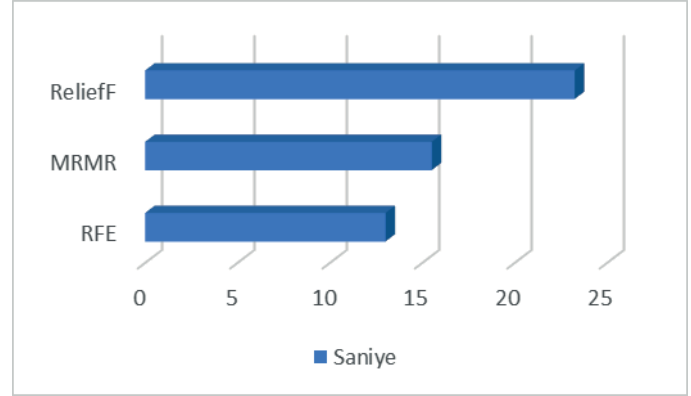
TABLO V. RFE_500 TEST VERİSİ DETAYLI SINIFLANDIRMA BAŞARIMI

	Sınıf	Kesinlik	Hassasiyet	GP	YP	F-Ölçütü
DVM	Negatif	0.974	0.864	0.864	0.022	0.916
	Pozitif	0.885	0.978	0.978	0.136	0.929
	Ağırlıklı Ortalama	0.928	0.923	0.923	0.081	0.923
SGD	Negatif	0.982	0.900	0.900	0.016	0.939
	Pozitif	0.913	0.984	0.984	0.100	0.947
	Ağırlıklı Ortalama	0.946	0.944	0.944	0.059	0.943
RBF	Negatif	0.715	0.749	0.749	0.281	0.732
	Pozitif	0.753	0.719	0.719	0.251	0.736
	Ağırlıklı Ortalama	0.735	0.734	0.734	0.265	0.734
LR	Negatif	0.991	0.886	0.886	0.007	0.936
	Pozitif	0.903	0.993	0.993	0.114	0.946
	Ağırlıklı Ortalama	0.946	0.941	0.941	0.062	0.941

69.52sn, 13sn, 5.44sn ve 159.37sn olarak ölçülmüştür. Bu çalışmalarda ölçümler için Intel(R) Core (TM) i5-4210U CPU @ 1.70GHz-2.40GHz özelliklerinde işlemciye sahip bir bilgisayar kullanılmıştır. Sınıflandırıcılar arasında en hızlı çözüm süresi RBF sınıflandırıcısına aitken hem süre hem de doğruluk başarımı olarak dikkate alındığında SGD sınıflandırıcısının en uygun sonucu verdiği görülmektedir. Şekil I'de ise SGD sınıflandırıcısı ile yapılan sınıflandırmalarda ReliefF, MRMR ve RFE nitelik seçme algoritmalarının sınıflandırma süresine etkilerinin gösterilmesi amacıyla sınıflandırma süresi başarımları çizdirilmiştir. Nitelik seçme algoritmaları ile elde edilen farklı nitelikler sınıflandırıcının doğruluk başarımı etkilediği gibi süre başarımı da etkilemektedir. Grafikten de anlaşılacağı üzere sınıflandırma algoritması ve nitelik sayısı aynı olmasına rağmen seçilen nitelikler farklı olduğundan dolayı sınıflandırma süreleri de farklılık göstermektedir.

TABLO VI. RFE NİTELİK SAYILARINA GÖRE SINIFLANDIRICILARIN SÜRE ANALİZİ (SN)

Nitelik Sayısı	DVM	SGD	RBF	LR
10	0.60	0.46	0.55	0.16
50	1.25	1.63	1.31	0.94
100	1.65	1.76	0.92	1.85
250	22.88	10.07	4.64	17.85
500	69.52	13.00	5.44	159.37



ŞEKİL I. 500 NİTELİK İÇİN NİTELİK SEÇME ALGORİTMALARININ SGD SINIFLANDIRICISINDAKİ SÜRE ANALİZİ

Yapılan bu çalışma siber zorbalık tespiti için MRMR, ReliefF ve RFE nitelik seçme algoritmaları ile nitelik sayısının azaltılması sonucunda siber zorbalık tespitinin başarımları analizleri gerçekleştirilmiştir. Youtube yorumlarının derlenmesi sonucuna elde edilen veri kümesi ilk olarak TF-IDF ağırlıklandırma yöntemi kullanılarak oluşturulan niteliklerden etkisiz kelimeler ve doküman frekansı 0.001'in altında olan nitelikler ayıklanmıştır. Gerçekleştirilen bu ön işlem adından sonra veri kümesini temsil eden 14215 nitelik çıkartılmıştır. Elde edilen nitelikler MRMR, ReliefF ve RFE nitelik seçme algoritmalarına alınarak 10, 50, 100, 250 ve 500 nitelik olarak ayrı ayrı hesaplanmış ve YSA tabanlı DVM, SGD, RBF ve LR sınıflandırıcıları ile sınıflandırılmıştır. Deneysel çalışmalar sonucunda RFE nitelik seçme algoritmasının SGD sınıflandırıcısı ile birlikte kullanılmasıyla optimal sonuçlar elde edilmiştir.

KAYNAKLAR

- [1] "Children and parents: media use and attitudes report 2017 - Ofcom."
- [2] "Pearson Student Mobile Device Survey 2015."
- [3] R. Sugandhi, A. Pande, S. Chawla, A. Agrawal, and H. Bhagat, "Methods for detection of cyberbullying: A survey," in *2015 15th International Conference on Intelligent Systems Design and Applications (ISDA)*, 2015, pp. 173-177.
- [4] S. A. Ozel, E. Sarac, S. Akdemir, and H. Aksu, "Detection of cyberbullying on social media messages in Turkish," in *2017*

- International Conference on Computer Science and Engineering (UBMK)*, 2017, pp. 366–370.
- [5] E. S. Eşsiz, “Selecting Optimum Feature Subsets With Nature Inspired Algorithms for Cyberbully Detection,” Çukurova University, 2016.
- [6] L. E. Dawei Yin, Zhenzhen Xue, Liangjie Hong, Brian D. Davison, April Kontostathis, “Detection of Harassment on Web 2.0,” in *Proceedings of the Content Analysis in the WEB 2.0 (CAW2.0) Workshop at WWW2009*, 2009.
- [7] V. Nahar, S. Unankard, X. Li, and C. Pang, “Sentiment Analysis for Effective Detection of Cyber Bullying,” 2012, pp. 767–774.
- [8] H. Sanchez, “Twitter Bullying Detection,” *UCSC ISM245 Data Min. course Rep.*, 2011.
- [9] H. L. Karthik Dinakar, Roi Reichart, “Modeling the Detection of Textual Cyberbullying | Request PDF,” in *Social Mobile Web Workshop at International Conference on Weblog and Social Media*, 2011.
- [10] M. Dadvar, F. M. G. de Jong, R. J. F. Ordelman, and R. B. Trieschnigg, “Improved cyberbullying detection using gender information.” Ghent University, 23-Feb-2012.
- [11] A. Kontostathis and A. Kontostathis, “ChatCoder: Toward the Tracking and Categorization of Internet Predators,” in *Proc. Text Mining Workshop 2009 Held In Conjunction With The Ninth Siam International Conference On Data Mining (SDM 2009)*. Sparks, Nv. May 2009., 2009.
- [12] M. Dadvar and F. M. G. de Jong, “Improved Cyberbullying Detection Through Personal Profiles.” Université de Bourgogne Observatoire International de la Violence Scolaire, 2012.
- [13] Noviantho, S. M. Isa, and L. Ashianti, “Cyberbullying classification using text mining,” in *2017 1st International Conference on Informatics and Computational Sciences (ICICoS)*, 2017, pp. 241–246.
- [14] I. McGhee, J. Bayzick, A. Kontostathis, L. Edwards, A. McBride, and E. Jakubowski, “Learning to Identify Internet Sexual Predation,” *Int. J. Electron. Commer.*, vol. 15, no. 3, pp. 103–122, Apr. 2011.
- [15] M. Dadvar, R. B. Trieschnigg, and F. M. G. de Jong, “Expert knowledge for automatic detection of bullies in social networks.” TU Delft, 2013.
- [16] G. Chandrashekar and F. Sahin, “A survey on feature selection methods,” *Comput. Electr. Eng.*, vol. 40, no. 1, pp. 16–28, Jan. 2014.
- [17] I. Guyon and A. Elisseeff, “An Introduction to Variable and Feature Selection,” *J. Mach. Learn. Res.*, vol. 3, no. Mar, pp. 1157–1182, 2003.
- [18] H. Hosni and F. Mhamdi, “A Filter Correlation Method for Feature Selection,” in *2014 25th International Workshop on Database and Expert Systems Applications*, 2014, pp. 59–63.
- [19] S. Maldonado and R. Weber, “A wrapper method for feature selection using Support Vector Machines,” *Inf. Sci. (Njy)*, vol. 179, no. 13, pp. 2208–2217, Jun. 2009.
- [20] S. Maldonado and J. López, “An embedded feature selection approach for support vector classification via second-order cone programming,” *Intell. Data Anal.*, vol. 19, no. 6, pp. 1259–1273, Nov. 2015.
- [21] R. Kohavi and G. H. John, “Wrappers for feature subset selection,” *Artif. Intell.*, vol. 97, no. 1–2, pp. 273–324, Dec. 1997.
- [22] J. Reunanen, “Overfitting in making comparisons between variable selection methods,” vol. 3, pp. 1371–1382, Mar. 2003.
- [23] A. L. Blum and P. Langley, “Selection of relevant features and examples in machine learning,” *Artif. Intell.*, vol. 97, no. 1–2, pp. 245–271, Dec. 1997.
- [24] H. Hanchuan Peng, F. Fuhui Long, and C. Ding, “Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.
- [25] H. Oufaida, O. Nouali, and P. Blache, “Minimum redundancy and maximum relevance for single and multi-document Arabic text summarization,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 26, no. 4, pp. 450–461, Dec. 2014.
- [26] Z. Wang *et al.*, “Application of ReliefF algorithm to selecting feature sets for classification of high resolution remote sensing image,” in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 755–758.
- [27] G. Ou, Y. Wang, W. Pang, and G. M. Coghil, “Large margin distribution machine recursive feature elimination,” in *2017 4th International Conference on Systems and Informatics (ICSAI)*, 2017, pp. 1518–1523.
- [28] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, “Gene Selection for Cancer Classification using Support Vector Machines,” *Mach. Learn.*, vol. 46, no. 1/3, pp. 389–422, 2002.