DOI:10.25092/baunfbed.876338 *J. BAUN Inst. Sci. Technol., 23(2), 782-791, (2021)*

Siber saldırı tespiti için makine öğrenmesi yöntemlerinin performanslarının incelenmesi

Fatih DEMİR*

Fırat Üniversitesi Teknik Bilimler MYO. Elektrik ve Otomasyon. Bölümü, Rektörlük Kampüsü, Elazığ

Geliş Tarihi (Received Date): 07.02.2021
Kabul Tarihi (Accepted Date): 04.05.2021

Öz

İnternet tabanlı cihazların kullanımının artması, siber ortamda güvenlik sorunlarına yol açmaktadır. Kötü amaçlı yazılımlar, sistemlerin işleyişini bozabilir ve sistemlerdeki güvenlik açıkları nedeniyle veri gizliliğini tehlikeye atabilir. Siber saldırıları belirlemek ve sınıflandırmak için Saldırı Tespit Sistemleri (STS) geliştirilmektedir. Yapay zeka tabanlı yöntemler, STS sistemlerini iyileştirmek için daha sık kullanılmaktadır. Bu çalışmada, STS sistemlerinin geliştirilmesinde yaygın olarak kullanılan ISCX-2012 veri setinin kullanıldığı literatür çalışmaları gözden geçirilmiştir. Ayrıca bu veri seti kullanılarak makine öğrenmesi tabanlı güçlü bir yaklaşım ile siber saldırılar %100 doğrulukla tespit edilmiştir. Önerilen yöntemin sınıflandırma doğruluğu performansını artırmak için öznitelik ve hiperparametre seçme algoritmaları kullanılmıştır. Önerilen yaklaşımın yapay zeka temelli STS sistemleri geliştirmek için faydalı olacağı düşünülmektedir.

Anahtar kelimeler: *Siber saldırı, saldırı tespit sistemleri, makine öğrenmesi.*

Investigation of performance of machine learning methods for cyber-attack detection

Abstract

The increase in internet based devices gets security problems in cyber environment. Malwares can disturb the functioning of systems and compromise data privacy due to vulnerabilities in systems. Intrusion Detection Systems (STS) are improved to determine and classify attacks. Artificial intelligence-based methods are used more frequently to improve STS systems. In this study, literature studies using ISCX-2012 data set, which was widely used in the development of STS systems, were reviewed. Besides, by using this

* Fatih DEMİR, fatihdemir@firat.edu.tr, http://orcid.org/ 0000-0003-3210-3664

data set, cyber-attacks were detected with 100% accuracy with a powerful machine learning-based approach. Feature and hyperparameter selection algorithms are used to increase the classification accuracy performance of the proposed method. These machine learning approaches are thought to be useful for developing artificial intelligence-based STS systems.

Keywords: *Cyber-attack, intrusion detection systems, machine learning.*

1. Giriş

Son yıllarda özellikle Nesnelerin İnterneti teknolojilerindeki gelişmelerle birlikte internet kullanan kişi ve uygulama sayısı sürekli artmaktadır. Dünyada internet kullanımı hakkında bilgi veren DataReportal verilerine göre internet kullanımı giderek artmaktadır. Ayrıca DataReportal verilerine göre her gün 1 milyon internet kullanıcısı eklenmektedir.

İnternet kullanımının artması birçok güvenlik açığını da beraberinde getirmiştir. Bu güvenlik açıklarını önlemek için güvenlik duvarı, veri şifreleme, kullanıcı kimlik doğrulama gibi birçok teknoloji kullanılmaktadır. Bu güvenlik mekanizmaları birçok saldırı türünü engellemektedir. Ancak, bu güvenlik teknolojileri derinlemesine paket analizi yapamaz. Bu nedenle istenilen saldırı tespit seviyesine ulaşamazlar. Saldırı Önleme Sistemi (SÖS) ve STS sistemleri, bu güvenlik mekanizmalarının eksikliklerini tamamlamak için geliştirilmiştir. Bu sistemler, makine öğrenmesi ve derin öğrenmeyi kapsayan yapay zeka tabanlı algoritmalar sayesinde diğer güvenlik sistemlerine göre daha derin veri analizleri yapabilmektedir. SÖS sistemleri hem saldırı algılama hem de önleme mekanizmaları olarak çalışırken, STS sistemleri yalnızca saldırı tespiti ve analizi için kullanılmaktadır [2-4]. Bu çalışmada STS sistemlerine odaklanmıştır.

İnternet kullanımının ve veri aktarım hızlarının artması da birçok anormalliğe neden olmuştur [5]. Sonuç olarak, internete yapılan saldırılar sürekli artmaktadır. Skybox Security tarafından 2020 yılında yayınlanan zafiyetler ve iş parçacıkları raporuna göre 2019 yılında 17220 yeni zafiyet tespit edilmiş olup, bir önceki yıla göre % 3,8'lik bir artış söz konusu olmuştur [6]. Bu yüzden, kurum ve kuruluşlar, kullanıcılarına güvenli ve istikrarlı bir hizmet sunabilmek için siber güvenlik teknolojileri harcamalarını sürekli artırmaktadır. Crystal Market Research (CMR) tarafından yayınlanan rapora göre 2012 yılında yaklaşık 58,13 milyar ABD doları değerinde olan Siber Güvenlik Piyasasının 2022 yılına kadar 173,57 milyar ABD dolarına ulaşması beklenmektedir. Yine bu rapora göre bulutun gelişmesi depolama ve Nesnelerin İnterneti gibi teknolojiler veri ihlali riskini artırmaktadır [7].

İnternet kullanımının artması, siber güvenlik şirketlerini geleneksel güvenlik yöntemlerinin yanı sıra daha hassas sistemler üretmeye zorlamıştır. Sonuç olarak, ağ davranış analizi, makine öğrenmesi, tehdit analizi gibi proaktif siber güvenlik sistemleri geliştirilmiştir. Günümüzde STS sistemleri, siber tehditlere daha duyarlı hale gelmek için sıklıkla kullanılmaktadır.

STS sistemleri geliştirmek için literatürde birçok makine öğrenmesi ile ilgili çalışmalar yapılmıştır. Önerilen yöntemin performansı ISCX 2012 veri seti üzerinde değerlendirildiği için bu veri setini kullanan çalışmalara yer verilmiştir.

Awad vd. [8], katmanlı örnekleme yöntemini ve farklı maliyet fonksiyonu şemalarını Ağırlıklandırılmış Aşırı Öğrenme Makinesi ile birleştirmiştir. Önerilen yöntem ISCX2012 veri kümesine uygulanmıştır. Bouteraa vd. [9] ISCX 2012 veri setini kullanarak izinsiz giriş tespiti için veri madenciliği tekniklerinin karşılaştırmalı bir çalışmasını gerçekleştirmiştir. Injadat vd. [10], anormallik tespiti için etkili bir çerçeve önermiştir. Rastgele Değişken (RD), DVM ve K-EYK makine öğrenmesi yöntemleri sınıflandırma için kullanılmıştır. DVM, RD ve K-EYK parametrelerini ayarlamak için Bayes Optimizasyon tekniği kullanılmıştır. Yassin vd. [11] K-Merkezli Kümeleme ve Saf Bayes sınıflandırıcılarını entegre bir şekilde kullanan hibrit bir algoritma önermiştir. Hassan vd. [12], yüksek doğruluk değerlerine sahip ağ saldırılarını tespit etmek için hibrit bir derin öğrenme modeli oluşturmuştur. Bu amaçla CNN ve ağırlığı azaltılmış Uzun Kısa Vadeli Bellek (UKVB) içeren bir derin öğrenme yöntemi kullanılmıştır.

Bu çalışmada literatürde sık kullanılan ISCX 2012 veri seti ile makine öğrenmesi tabanlı STS uygulaması gerçekleştirilmiştir. Ki-kare öznitelik seçme algoritması ile yüksek seviyeli öznitelikler seçilmiştir. Böylece hesaplama maliyeti düşürülmüştür. KA algoritması ile sınıflandırma işlemi yapılmıştır. KA sınıflandırıcısının hiperparametreleri Bayes optimizasyon yöntemiyle otomatik olarak seçilmiştir. Böylelikle hiperparametre seçimindeki pratik deneyim gereksinimi ortadan kaldırılmıştır.

2. Sınıflandırma teknikleri

Makine öğrenmesi teknikleri uygulanırken sistemin performansını etkileyen en önemli aşamalardan birisi sınıflandırmadır. Çıkartılan özniteliklere göre her sınıfa ait örnekler belli bir bölgeye konumlandırılır. Sınıflar arası ayrım her sınıfa ait örneklerin bu konumlarına göre yorumlanmalıdır ve elde edilen bulgulara göre uygun sınıflandırıcı veya sınıflandırıcılar seçilmelidir.

2.1. Destek vektör makineleri (DVM)

DVM, hem sınıflandırma hem de regresyon problemlerini çözmek için kullanılabilen gözetimli bir makine öğrenme algoritmasıdır. Bununla birlikte, en çok sınıflandırma problemlerinin çözümünde kullanılır. DVM algoritmasında, her veri çifti ve her bir öznitelik değeri özel bir koordinat düzleminde olacak şekilde n boyutlu uzayda bir nokta olarak temsil edilebilir. Ardından, iki sınıflı bir sınıflandırma problemini çözmek için bir hiper düzlem bulunur ve sınıflandırma yapılır [13]. Hiper düzlemdeki doğrusal denklemler ve sınıflar arasındaki mesafenin Lagrange çarpanları ile optimize edilmesi gerekir. Eğer sınıflar doğrusal olarak ayrılamaz ise Gauss, Dairesel Taban ve Polinominal gibi çekirdek fonksiyonları kullanarak problem başka uzaya taşınır. Böyle bir durumdaki DVM sınıflandırıcısının karar fonksiyonu Denklem (1)'de verilmiştir.

$$f(x) = \text{sgn}\left(\sum_{i}^{n} \alpha_i y_i K(x_i, x) + b \right) \tag{1}$$

Burada, "α_i" Lagrange çarpanlarından elde edilen ağırlık vektörü, "y_i" sınıf etiketi ,"x_i" giriş değeri ve "b" bias vektörüdür.

2.2. K-en yakın komşu (K-EYK)

K-EYK algoritması, en basit ve en yaygın kullanılan sınıflandırma algoritmalarından biridir. K-EYK, parametrik olmayan tembel bir öğrenme algoritmasıdır. İstekli öğrenmenin aksine tembel öğrenmede bir eğitim aşaması yoktur. K-EYK, eğitim verilerini öğrenmez; bunun yerine eğitim veri kümesini ezberler. Bir tahmin yapmak istenildiğinde, tüm veri setinde en yakın komşuları aranır. Algoritma çalışmasında bir k değeri belirlenir. Bu k değerinin anlamı, bakılacak eleman sayısıdır. Bir değer geldiğinde, en yakın k elemanı alınarak gelen değer arasındaki mesafe hesaplanır. Öklid işlevi genellikle mesafe hesaplamasında kullanılır. Öklid işlevine alternatif olarak City Block, Manhattan, Minkowski ve Chebyshev işlevleri de kullanılabilir [14]. Mesafe hesaplandıktan sonra en yakın komşular sıralanır ve gelen değer uygun sınıfa atanır.

2.3. Karar ağacı (KA)

KA sınıflandırıcısı, çoğunlukla sınıflandırma problemlerini basit yoldan çözmek için kullanılmaktadır. KA algoritmasında, sınıflandırma problemini çözmek için en doğru yol belirli yöntemlerle aranır. KA sınıflandırıcısında, yukarıdan aşağıya inen kökler, yapraklar ve dallardan oluşan bir yapı bulunmaktadır. KA sınıflandırıcısında ayırma kriteri olarak Gini katsayısı, Büyüyen KA veya Sapmalar metotları uygulanmaktadır [15].

2.4. Lineer ayırtaç analizi (LAA)

LAA algoritması, sınıflandırma ve öznitelik seçimi için sıklıkla kullanılmaktadır. LAA algoritması, sınıf içinde eşit olmayan frekanslarda ve rastgele oluşturulan test verilerinde kolayca uygulanır. LAA algoritmasının temel çalışma prensibi, belirli bir veri setindeki sınıflar arasındaki varyans oranını sınıflar içindeki varyans oranına maksimize ederek en iyi ayrıştırmayı bulmaya dayanmaktadır [16]. Genel bir LAA algoritmasında, dağılım matrisleri sınıflar arası dağılım matrisleri hesaplanır. LAA fonksiyonu Denklem (2)'de verilmiştir

$$d(x) = w^T\left(x - \frac{m}{w}\right) \tag{2}$$

.

Burada, "w" dağılım matrislerine göre oluşturulan ağırlık vektörü, "x" giriş örneklerini ve "m" sınıf etiketlerini temsil etmektedir.

2.5. Ki-Kare Algoritması

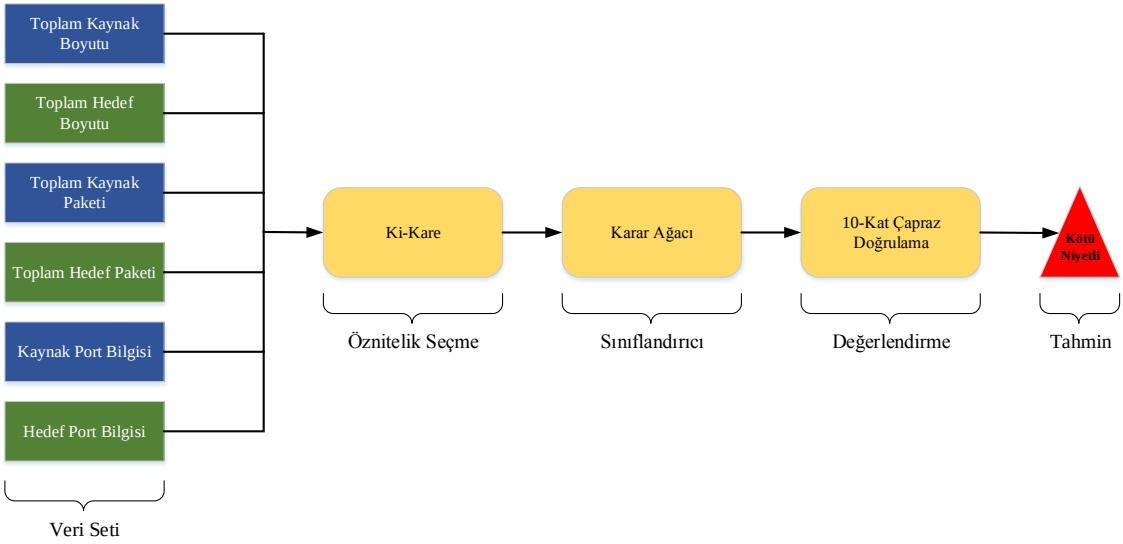
Ki-Kare yönteminde [17] r_i öznitelik setindeki seçim, bir S_j sınıfı ile korelasyonu temel alınarak yapılır ve r_i öznitelik setinde bulunan S_j sınıfının ayırt etme yeteneği Denklem (3)'teki gibi hesaplanır:

$$x^2\left(r_i, S_j\right) = \frac{L \times (m_{ij}c_{ij} - n_{ij}d_{ij})^2}{(m_{ij}c_{ij}) \times (m_{ij}d_{ij}) \times (n_{ij}c_{ij}) \times (n_{ij}d_{ij})} \tag{3}$$

L, toplam giriş sayısını ifade etmektedir. m_{ij}, r_i özniteliğinde bulunan S_j sınıfına ait örnek sayısını ve n_{ij}, S_j kategorisindeki r_i özniteliğine ait olmayan örneklerin sayısını karşılık gelmektedir. d_{ij}, r_i özniteliğine ait fakat S_j kategorisini içermeyen örneklerin sayısına karşılık gelmektedir. c_{ij}, S_j kategorisinde bulunmayan ve r_i özniteliğine ait olmayan örneklerin sayısına karşılık gelmektedir.

2.6. Materyal metot

Makine öğrenme tabanlı STS sistemlerinin uygulamasını gerçekleştirmek için ISCX 2012 veri kullanılmıştır. Bu veri seti, yedi günlük ağ verilerinden oluşturulmuştur. Veri seti, normal ve kötü niyetli ağ trafiğini içermektedir. Kötü amaçlı ağ trafiği, ağa içeriden sızma, HTTP Hizmet Reddi, Dağıtılmış Hizmet Reddi ve Brute Force SSH saldırılarını içermektedir. Makine öğrenme uygulamasını gerçekleştirmek için 2516 normal ve 2515 kötü niyetli sınıf örneği olacak şekilde veri seti azaltılmıştır. [18]. Veri setinde yer alan toplam kaynak boyutu (bayt), toplam hedef boyutu (bayt), toplam kaynak paketi (adet), toplam hedef paketi (adet), kaynak portu ve hedef portu öznitelik olarak kullanılmıştır. Çıkarılan özniteliklerin içinden yüksek seviyeli olanlar için Ki-Kare algoritması ile seçilmiştir ve öznitelik sayısı 6'dan 3'e düşürülmüştür. Seçilmiş öznitelikler en iyi sınıflandırma sonucunu verdiği için KA sınıflandırıcısına iletilmiştir ve kötü niyetli örnekler tahmin edilmiştir. Önerilen metodun temsili grafiği Şekil 1'de verilmiştir. Sonuçları değerlendirmek için 10-kat çapraz doğrulama metodu kullanılmıştır.



Şekil 1. Önerilen metodun temsili grafiği.

2.7. Performans metrikleri

Bu çalışmada önerilen metodun performansını ölçmek için doğruluk (Do) ana kriter olmak üzere duyarlılık (Du), özgüllük (Öz), kesinlik (Ks) ve *F-skor* değerleri kullanılmıştır. Bu ölçütleri hesaplamak içim karmaşıklık matrisinde yer alan doğru pozitif (DP), doğru negatif (DN), yanlış pozitif (YP), yanlış negatif (YN) değerleri kullanılmıştır. Bu performans metrikleri aşağıdaki denklemler ile elde edilmiştir.

$$Do = \frac{DP + DN}{DP + DN + YP + YN} \tag{4}$$

$$Du = \frac{DP}{DP + YN} \tag{5}$$

$$Öz = \frac{DN}{DN + YP} \tag{6}$$

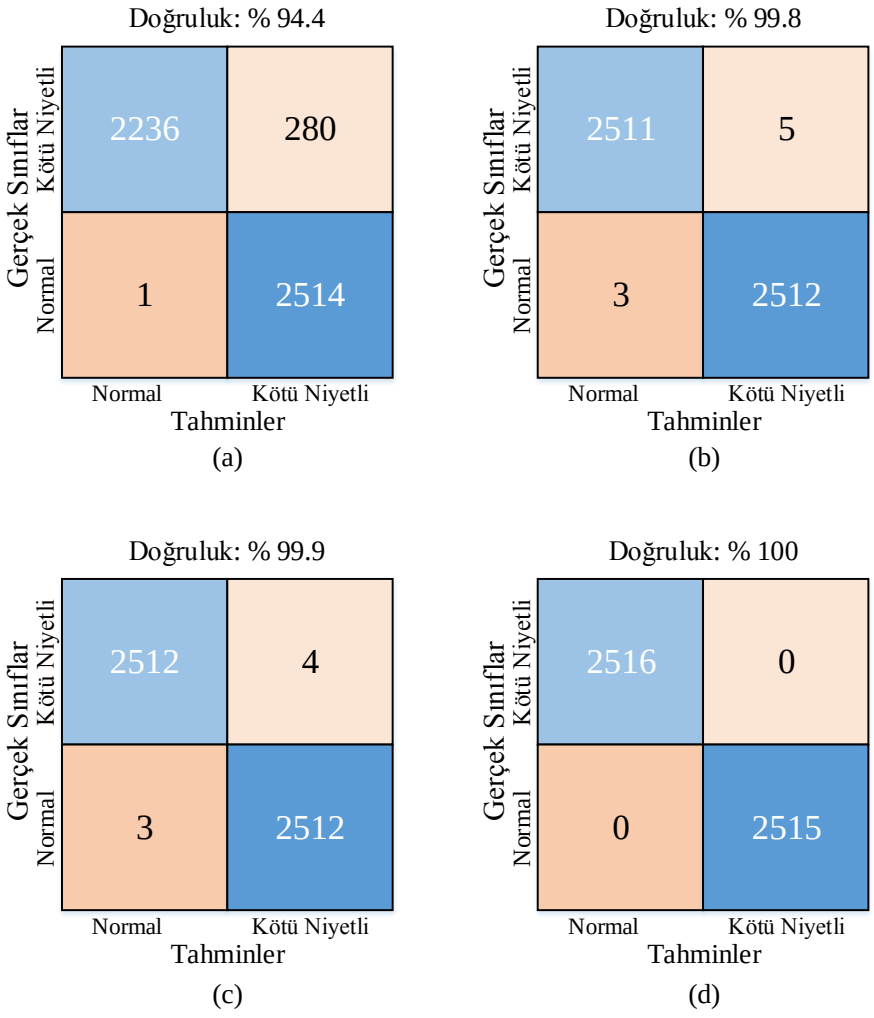
$$Ks = \frac{DP}{DP + YP} \tag{7}$$

$$F - skor = 2 \times \frac{Ks \times Du}{Ks + Du} \tag{8}$$

Ayrıca bir performans ölçütü olarak ROC eğrisi ve AUC değeri de verilmiştir. ROC eğrisi DP oranının YP oranına göre değişimi kullanılarak çizdirilir. AUC değeri ise ROC eğrisi altında kalan alandır.

3. Sonuçlar ve tartışma

Bu çalışmada Ki-kare öznitelik seçimi haricindeki tüm kodlamalar Matlab ortamında gerçekleştirilmiştir. Öznitelik seçimi için Python programlama dili kullanılmıştır. Intel® Core™ i7-5500 CPU işlemcisi, 8GB DRR3 bellek ve 2GB grafik kartına sahip bir bilgisayar donanım ile çalışılmıştır. DVM sınıflandırıcısı için Gauss çekirdek fonksiyonu tercih edilmiştir. Çekirdek ölçeği, "0.64" ve kutu kısıtlaması "1" seçilmiştir. K-EYK sınıflandırıcında uzaklık denklemi "kosinüs" ve en yakın komşu sayısı "10" seçilmiştir. LAA sınıflandırıcısındaki hiperparametre gama, "0" olarak girilmiştir. 10-Kat Çapraz değerlendirme sonunda kullanılan sınıflandırıcıların karmaşıklık matrisi sonuçları Şekil 2'de verilmiştir.



Şekil 2. Sınıflandırma sonuçları: (a) LAA, (b) DVM (c) K-EYK, (d) KA

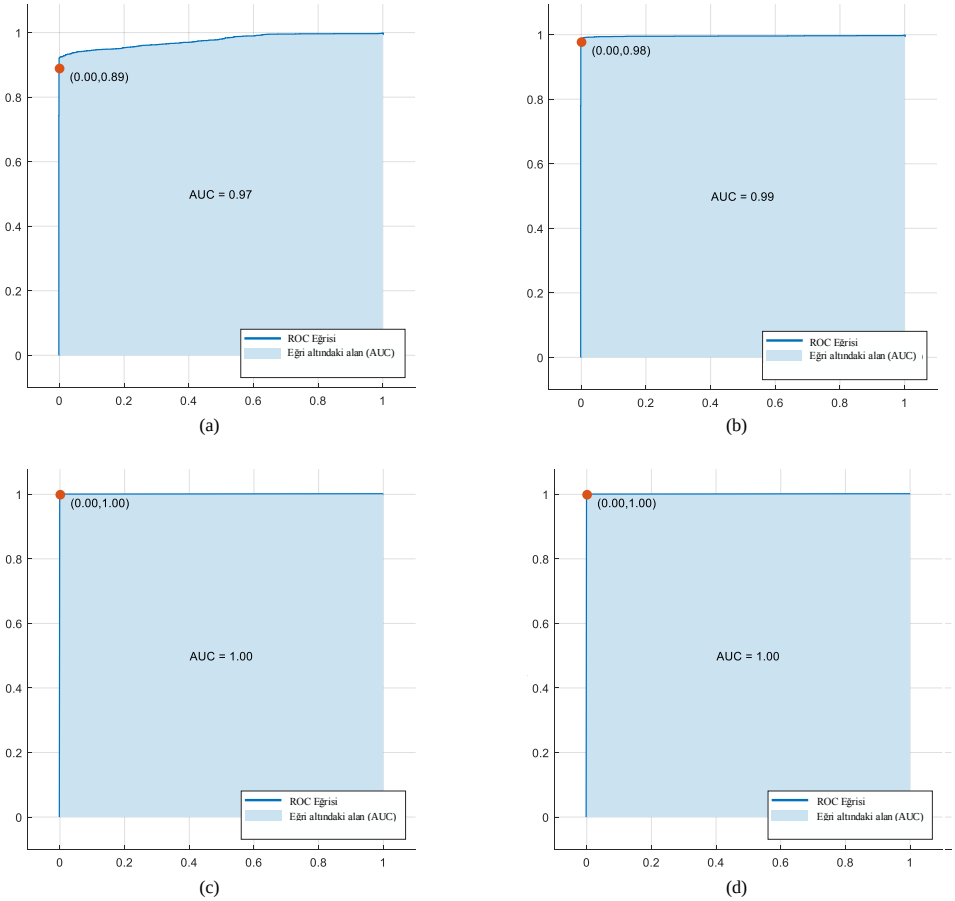
Şekil 2'de verilen sonuçlara göre en iyi sınıflandırma sonucuna %100 doğruluk ile KA sınıflandırıcısıyla ulaşılırken, en kötü sonuca %94.11 doğruluk ile LAA sınıflandırıcısıyla ulaşılmıştır. Şekil 2'deki karmaşıklık matrisleri kullanılarak elde edilen Du, Öz, Ks ve F-skor değerleri Tablo 1'de verilmiştir. En iyi sonuçlara yine KA sınıflandırıcısıyla ulaşılırken en kötü sonuçlara LAA sınıflandırıcısıyla ulaşılmıştır.

Tablo. 1 Sınıflandırıcıların Du, Öz, Ks ve F-skor sonuçları.

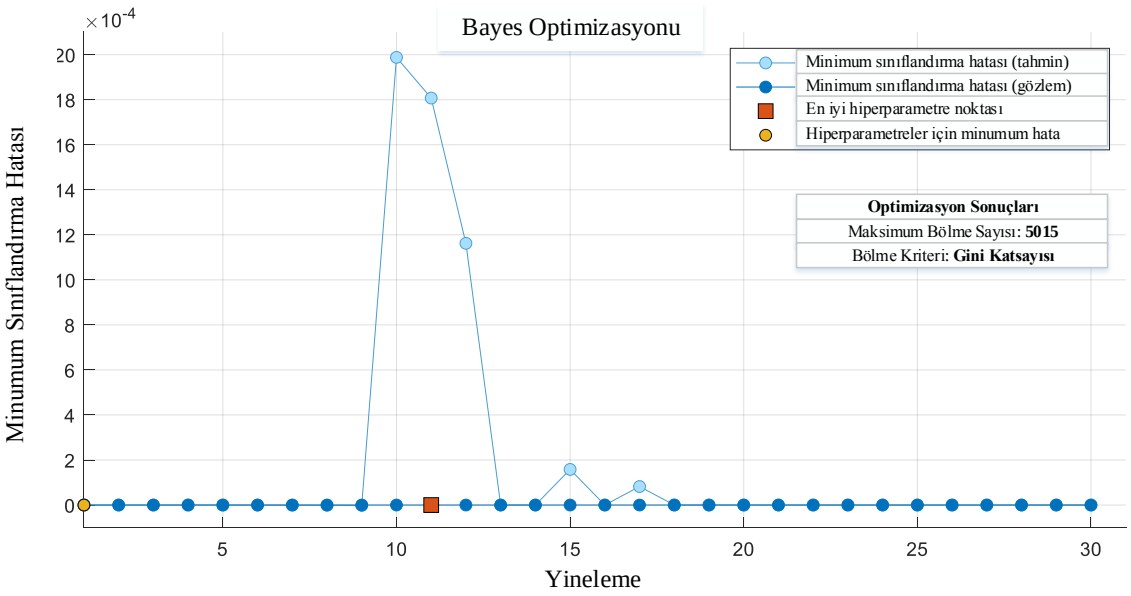
Sınıflandırıcı	Sınıf	Du	Öz	Ks	F-skor
LAA	Normal	0,88	0,99	0,99	0,94
	Kötü Niyetli	0,99	0,88	0,89	0,94
DVM	Normal	0,99	0,99	0,99	0,99
	Kötü Niyetli	0,99	0,99	0,99	0,99
K-EYK	Normal	0,99	0,99	0,99	0,99
	Kötü Niyetli	0,99	0,99	0,99	0,99
KA	Normal	1	1	1	1
	Kötü Niyetli	1	1	1	1

Tablo 1'e göre en iyi sonuçlar KA sınıflandırıcısıyla alınırken, en kötü sonuçlar LAA sınıflandırıcısıyla alınmıştır.

Tüm sınıflandırıcılar için ROC eğrileri ve AUC değerleri Şekil 3'te verilmiştir. Şekil 3'e göre en kötü AUC değeri LAA sınıflandırıcısıyla elde edilirken, en iyi AUC değerine DVM ve KA sınıflandırıcısıyla ulaşılmıştır.



Şekil 3. Sınıflandırıcıların ROC eğrileri: (a) LAA, (b) DVM (c) K-EYK, (d) KA



Şekil 4. KA sınıflandırıcısının hiperparametre seçimi

Şekil 4'te KA sınıflandırıcısının hiperparametre seçimi için Bayes optimizasyon algoritması uygulanmıştır. Bu optimizasyon tekniğinin uygulanışı ile ayrıntılı bilgi [20]'de yer almaktadır. Bayes algoritması en iyi sınıflandırma sonucunu verecek hiperparametreleri 30 yineleme sonucunda otomatik olarak bulmuştur. Hiperparametrelerden biri olan maksimum bölme sayısı 1-5030 aralığında aranmıştır ve en iyi maksimum bölme sayısı 5015 olarak bulunmuştur. Bölme kriteri olarak Gini katsayısı ve maksimum sapma azaltma kullanılmıştır ve Gini katsayısı kriterinin daha iyi sonuç verdiği görülmüştür.

Tablo 2'de ISCX 2012 setini kullanan mevcut çalışmaların performans metrikleri verilmiştir. Tablo 2'ye göre tüm performans metrikleri birbirine çok yakın olmakla beraber en iyi sonuçlar önerilen metotta kullanılan KA sınıflandırıcısıyla elde edilmiştir.

Tablo. 2 Metotların karşılaştırılması

Referans	Sınıflandırıcı	Do	Du	Ks	F-skor
Injadet vd. [10]	DVM	0.9984	1.00	0.9998	-
	K-EYK	0.9993	1.00	0.9999	-
	RD	0.9992	1.00	0.9999	-
Kılıncer vd. [19]	KA	1.00	0.9994	0.9994	0.9994
	K-EYK	0.9980	0.9978	0.9978	0.9979
	DVM	0.9980	0.9978	0.9978	0.9979
Önerilen Metot	KA	1.00	1.00	1.00	1.00
	K-EYK	0.9990	0.9986	0.9986	0.9986
	DVM	0.9980	0.9984	0.9984	0.9984

Elde edilen sonuçlar STS için kullanılan makine öğrenmesi sınıflandırıcılarının üstün performans sağladığını göstermiştir. Aynı veri setini kullanan diğer çalışmalara göre daha iyi sınıflandırma doğrulukları elde edilmiştir. Dahası, bu sınıflandırma doğrulukları daha az öznitelikle (3 adet öznitelik) elde edilmiştir ve hesaplama maliyeti azaltılmıştır. Bu

çalışmada, sınıflandırma performansı için sınıflandırıcı ve hiperparametre seçiminin önemli olduğu görülmüştür. En iyi sınıflandırma sonucunu elde KA sınıflandırıcısı ile elde edilirken hiperparametre seçimi Bayes algoritması ile otomatik olarak yapılmıştır. Kötü niyetli yazılımları tespit etmek için geliştirilen bu sistem yüksek başarım sağladığından STS sistemlerinde kullanılabilir. Ayrıca gelecekteki çalışmalarda daha çok sınıfı yüksek başarımla otomatik olarak tanıyan Derin Öğrenme Tabanlı STS sistemlerini geliştirmek de önem arz etmektedir.

Kaynaklar

[1] Kemp, S., Digital 2019: Global digital overview, **Retrieved from Datareportal**, (2019).

[2] Hajisalem, V. ve Babaie, S., A hybrid intrusion detection system based on ABC-AFS algorithm for misuse and anomaly detection, **Computer Networks**, 136, 37-50, (2018).

[3] Inayat, Z., Gani, A., Anuar, N. B., Khan, M. K. ve Anwar, S., Intrusion response systems: Foundations, design, and challenges, **Journal of Network and Computer Applications**, 62, 53-74, (2016).

[4] Ashoor, A. S. ve Gore, S., Difference between intrusion detection system (IDS) and intrusion prevention system (IPS), **In International Conference on Network Security and Applications**, 497-501, Berlin, Heidelberg, (2011).

[5] Jabez, J. ve Muthukumar, B., Intrusion detection system (IDS): anomaly detection using outlier detection approach, **Procedia Computer Science**, 48, 338-346, (2015).

[6] Quepons, I., Vulnerability and Trust, **PhaenEx**, 13, 2, 1-10, (2020).

[7] Garcia-Teodoro, P., Diaz-Verdejo, J., Maciá-Fernández, G. ve Vázquez, E., Anomaly-based network intrusion detection: Techniques, systems and challenges, **Computers and Security**, 28,1-2, 18-28, (2009).

[8] Awad, M., ve Alabdallah, A., Addressing Imbalanced Classes Problem of Intrusion Detection System Using Weighted Extreme Learning Machine, **International Journal of Computer Networks & Communications (IJCNC)**, 1-11, Toronto, Canada, (2019).

[9] Bouteraa, I., Derdour, M. ve Ahmim, A., Intrusion Detection using Data Mining: A contemporary comparative study, **In 2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS)**, 1-8, Algeria (2018).

[10] Injadat, M., Salo, F., Nassif, A. B., Essex, A. ve Shami, A., Bayesian optimization with machine learning algorithms towards anomaly detection, **In 2018 IEEE global communications conference (GLOBECOM)**, 1-6, Abu Dhabi, (2018).

[11] Yassin, W., Udzir, N. I., Muda, Z. ve Sulaiman, M. N., Anomaly-based intrusion detection through k-means clustering and naives bayes classification, **In Proc. 4th Int. Conf. Comput. Informatics (ICOCI)**, 298-303, Kuching, Sarawak, (2013).

[12] Hassan, M. M., Gumaei, A., Alsanad, A., Alrubaian, M. ve Fortino, G., A hybrid deep learning model for efficient intrusion detection in big data environment, **Information Sciences**, 513, 386-396, (2020).

[13] Cortes, C. ve Vapnik, V. Support Vector Machine**, Machine Learning**, 20, 3, 273-297, (1995).

[14] Akbulut, Y., Sengur, A. ve Guo, Y., Smarandache, F., NS-k-NN: Neutrosophic set-based k-nearest neighbors classifier, **Symmetry**, *9*, 9, 179, (2017).

[15] Safavian, S. R. ve Landgrebe, D., A survey of decision tree classifier methodology**, IEEE transactions on systems, man, and cybernetics**, 21, 3, 660-674, (1991).

[16] Altay, O. ve Ulas, M., Prediction of the autism spectrum disorder diagnosis with linear discriminant analysis classifier and K-nearest neighbor in children, **In 2018 6th International Symposium on Digital Forensic and Security (ISDFS)**, Antalya, Turkey, 1-4, (2018).

[17] Kira, K. ve Rendell, L. A., A practical approach to feature selection, **In Machine learning proceedings 1992**, 249-256, Morgan Kaufmann, (1992).

[18] Shiravi, A., Shiravi, H., Tavallaee, M. ve Ghorbani, A. A., Toward developing a systematic approach to generate benchmark datasets for intrusion detection, **Computers and Security**, 31, 3, 357-374, (2012).

[19] Kilincer, I. F., Ertam, F. ve Sengur, A., Machine Learning Methods for Cyber Security Intrusion Detection: Datasets and Comparative Study, **Computer Networks**, 107840, (2021).

[20] Ucar, F. ve Korkmaz, D, COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images, **Medical Hypotheses**, 140, 109761, (2020).