Karaelmas Fen ve Mühendislik Dergisi
Dergi web sayfası: https://dergipark.org.tr/tr/pub/karaelmasfen
DOI: 10.7212/karaelmasfen.1516627

Araştırma Makalesi
Geliş tarihi / Received : 15.07.2024
Kabul tarihi / Accepted : 15.09.2024

Çeşitli Tekrarlayan Sinir Ağları Kullanarak Siber Saldırı Tespiti

Cyber-Attack Detection Using Various Recurrent Neural Networks

Nesibe Yalçın[1]*, Semih Çakır[2], Sibel Ünaldı[3]

[1]Erciyes Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Kayseri, Türkiye
[2]Zonguldak Bülent Ecevit Üniversitesi, Kdz. Ereğli Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Zonguldak, Türkiye
[3]Bilecik Şeyh Edebali Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, Bilecik, Türkiye

Bu makale, 10-11 Mayıs 2024 tarihleri arasında Karaelmas International Science and Engineering Symposium (KISES 2024) isimli sempozyumda sözlü olarak sunulan ve özet kitabında yayınlanan çalışmanın genişletilmiş versiyonudur (Sempozyum bildirisinin başlığı: "Cyber-Attack Detection Using Various Recurrent Neural Networks").

Öz

Siber saldırıların erken tespiti ve tanımlanması; saldırıların etkisinin doğru bir şekilde değerlendirilmesi, bunlara karşı hızlı ve etkili önlemlerin alınması, veri ve sistemlerin korunması, iş sürekliliğin sağlanması, kurumsal itibarın korunması, yasal ve düzenleyici standartlara uyumun sağlanması açısından hayati önem taşımaktadır. Bu çalışmada, bir siber saldırı tespit sistemi önerilmiştir. Saldırı tespiti için çeşitli Tekrarlayan Sinir Ağı (Recurrent Neural Network, RNN) derin öğrenme yöntemlerinin yanı sıra K En Yakın Komşu (K Nearest Neighbour, KNN), Karar Ağacı ve Rastgele Orman makine öğrenme algoritmaları uygulanmıştır. Sistemin saygınlığı, KDD'99 veri setinin %10'u üzerinden değerlendirilmiş ve tartışılmıştır. Öğrenme modellerinin başarımını karşılaştırmak için çeşitli değerlendirme metrikleri kullanılmıştır. Aynı veri setini kullanan çalışmalarla karşılaştırıldığında önerilen Çift Yönlü Uzun Kısa Süreli Bellek, daha yüksek başarıma sahip RNN modeli olarak öne çıkmaktadır. Ayrıca KNN de yüksek bir test doğruluğu (%99.92) ve duyarlılık (%99.94) sunmuştur. Önerilen modeller, siber saldırıların erken aşamada tespit edilmesini kolaylaştırabilir.

Anahtar Kelimeler: Derin öğrenme, KDD'99, RNN, siber güvenlik, saldırı tespiti.

Abstract

Early detection and identification of cyber-attacks is vital to accurately assess their impact, take swift and effective countermeasures against them, protect data and systems, maintain operational continuity, preserve organizational reputation, and ensure compliance with legal and regulatory standards. A cyber-attack detection system was proposed in this study. K Nearest Neighbour (KNN), Decision Tree, and Random Forest machine learning algorithms, and also various Recurrent Neural Network (RNN) deep learning methods were applied for attack detection. The reputability of the system was evaluated and discussed using 10% of the publicly available KDD'99 dataset. Various evaluation metrics were used to compare the performance of these learning models. In comparison to the studies using the same dataset, the proposed Bidirectional Long Short-Term Memory stands out with its higher performance as an RNN model. KNN also presented a higher test accuracy (99.92%) and recall (99.94%). The proposed models may facilitate the detection of cyber-attacks at an early stage.

Keywords: Deep learning, KDD'99, RNN, cyber security, attack detection.

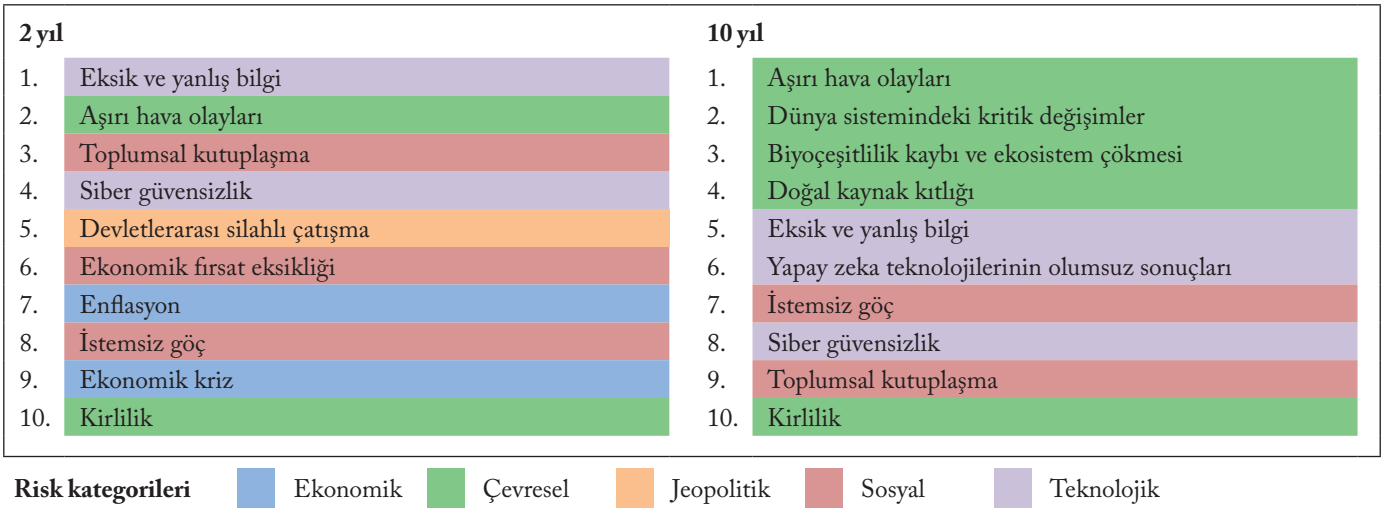
*Sorumlu yazarın e-posta adresi: nesibeyalcin@erciyes.edu.tr

Nesibe Yalçın orcid.org/0000-0003-0324-9111
Semih Çakır orcid.org/0000-0003-3072-9532
Sibel Ünaldı orcid.org/0000-0001-9948-4284

1. Giriş

Dünya Ekonomik Forumu (World Economic Forum, WEF)'nun 2024 Küresel Riskler Raporu'nda (WEF 2024) siber güvensizlik, en büyük 10 küresel risk arasında gösterilmiş, kısa vadede en ciddi 4. ve uzun vadede 8. küresel risk olarak sıralanmıştır (Şekil 1). Ayrıca siber saldırılar, küresel ölçekte maddi bir krize neden olma ihtimali en yüksek olduğuna inanılan ve en acil çözüm gerektiren 5 küresel risk arasındadır. Siber saldırıların önümüzdeki yıllarda sürekli ve önemli bir endişe kaynağı olmaya devam edeceği açıktır. Bununla birlikte siber güvenlik saldırılarını tespit etmek ve etkilerini azaltmak için çok çeşitli güvenlik araçları ve sistemleri (güvenlik duvarları, saldırı tespit sistemleri, bulut güvenlik yazılımları gibi) mevcuttur.

Günümüzde büyük bir tehdit haline gelen siber saldırılara karşı otomatik/erken tespit ve önlem alma, önemli bir savunma stratejisidir. Yapay zeka ve makine öğrenmesi gibi teknolojilerin kullanılması, anomalileri ve potansiyel saldırı işaretlerini tanımlamada yardımcı olabilmektedir. Literatürde bilgisayar ağlarında saldırı tespiti için çeşitli yapay zeka yöntemlerini kullanan birçok model önerilmiş ve yüksek başarım elde edilmiştir. 2020 yılında Iwendi ve diğerleri, yaptıkları çalışmada KDD'99 (KDD Cup 1999 Data) ve NSL-KDD veri setlerini makine öğrenmesi tabanlı saldırı tespitinde kullanılmışlardır (Iwendi vd. 2020). 5 ana saldırı türüne ait sınıfları ele alarak korelasyon tabanlı öznitelik seçimi + topluluk sınıflandırıcıları (Bagging ve Adaboost) yaklaşımlarını önermişler ve yüksek doğruluk ile düşük yanlış alarm oranına sahip saldırı takip sistemi geliştirmişlerdir. Önerdikleri yaklaşım, KDD'99 veri setinde 0 yanlış alarm oranı ve %99.90 tespit oranı, NSL-KDD veri setinde %0.5 yanlış alarm oranı ve %98.60 tespit oranı başarımına sahiptir. Yapılan bir başka çalışmada (Gao vd. 2021) ise SCADA ağları için saldırılar, ilk kez zamansal ilişkili olma ve olmama durumları için kategorize edilmiş ve derin öğrenmeye dayalı çok yönlü bir saldırı tespit sistemi üzerine çalışılmıştır. Çalışmada %10 KDD'99 veri seti; 1) normal ve ilişkisiz saldırılar, 2) normal ve ilişkili saldırılar ve 3) hem ilişkili hem de ilişkisiz saldırıları içeren 3 farklı veri seti şeklinde düzenlenmiştir. Çalışmadan elde edilen bulgular zamansal ilişkili saldırıları tespit etmede Uzun Kısa Süreli Bellek (Long Short-Term Memory, LSTM)'nin daha iyi performans gösterdiğini, zamansal ilişkili olmayan saldırıların tespitinde ise İleri Beslemeli Sinir Ağı (Feedforward Neural Network, FNN)'nın daha avantajlı olduğunu ortaya koymuştur. Saldırı tespitinde Zamansal Evrişimli Ağlar (Temporal Convolutional Networks, TCN)'ın başarımının incelendiği bir çalışmada (Çakır ve Angin 2021), KDD'99 veri seti kullanılmıştır. Araştırmada Tam Evrişimli Ağlar (Fully Convolutional Networks, FCN) ile LSTM ve TCN tabanlı otomatik kodlayıcılar geliştirilerek başarımları karşılaştırılmış ve ikili sınıflandırmada TCN'nin en az LSTM kadar başarılı olduğu raporlanmıştır. Laghrissi ve diğerleri (2021), siber saldırıları tespit etmek için LSTM tabanlı bir derin öğrenme yaklaşımı sunmuşlardır. Boyut azaltmanın uygulandığı bu yaklaşım, KDD'99 veri seti üzerinde test edilmiş ve %99.49 doğruluk - %99.15 duyarlılık sağlamıştır. LSTM Tekrarlayan Sinir Ağı (Recurrent Neural Network, RNN) uygulanan bir saldırı tespit modelinde ise değerlendirme UNSW-NB15 veri seti üzerinde yapılmış ve Basit RNN modeli ile karşılaştırıldığında %99'un üzerinde doğ-



Şekil 1. Önem derecesine göre küresel riskler (WEF 2024).

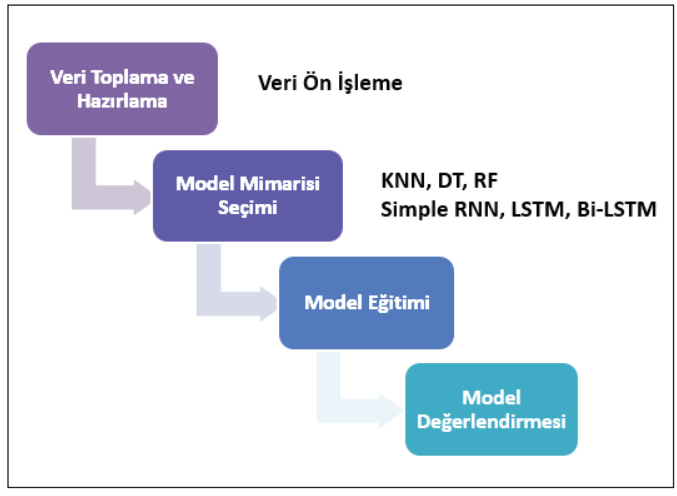
ruluk ile daha iyi başarım göstermiştir (Thant vd. 2023). Kasongo tarafından Basit RNN, LSTM ve Geçitli Tekrarlayan Birim (Gated Recurrent Unit, GRU) yöntemlerini kullanan bir saldırı tespit sistemi önerilmiştir. Sistemin performansını değerlendirmek için UNSW-NB15 ve NSL-KDD veri setleri dikkate alınmış, ayrıca özellik seçimi yapılmıştır. LSTM, NSL-KDD veri seti üzerinde %88.13 test doğruluğu ile en iyi başarımı göstermiştir. UNSW-NB15 veri seti üzerinde Basit RNN, %87.07 doğruluk oranı ile en etkili model olmuştur (Kasongo 2023). Ağ saldırı tespit sistemi için Çift Yönlü LSTM (Bidirectional LSTM, Bi-LSTM) kullanan sinir ağlarının değerlendirildiği bir çalışmada (Pooja ve Shrinivasacharya 2021) ise KDD'99 ve UNSW-NB15 veri setleri üzerinde deneyler gerçekleştirilmiştir. Bi-LSTM modeli, her iki veri seti için %99 doğrulukla yüksek sonuçlar vermiştir. KDD'99 veri seti üzerinde yapılan bir başka çalışmada (Liu ve Zhang 2020) ise Evrişimli Sinir Ağları kullanılarak %98.02 doğruluk elde edilmiştir. Alenazi ve Mishra (2024), saldırı tespit için Ekstrem Gradyan Artırma (Extreme Gradient Boosting, XGBoost) modeli önermişler, KDD'99 veri setinde %99.98 ve NSL-KDD veri setinde %99.97 ise doğruluk elde etmişlerdir. Ayrıca çalışmalarında, Gaussian Naïve Bayes (NB) ile XGBoost modelini karşılaştırmışlar ve her iki veri setinde de XGBoost modelinin üstün yanlarını vurgulamışlardır. İlgün ve Samet, 2024 yılında yaptıkları çalışmada ise NSL-KDD veri setine ön işlem uygulanmadan ve sırasıyla kategorik veri kodlama, ölçeklendirme, hibrit öznitelik seçimi ön işlemlerini hem ayrı ayrı hem de beraber uygulayarak 5 farklı senaryo oluşturmuş ve böylelikle 5 farklı veri seti elde etmişlerdir. Çalışmanın bir sonraki adımında bu veri setleri üzerinde çeşitli makine öğrenmesi yöntemlerini kullanmışlardır. En son adımda ise en başarılı sonuçları veren yöntemler ele alınarak hiperparametre optimizasyonu ile modellerin performansları iyileştirilmiştir. Çalışmada optimizasyon uygulanmasının, genel olarak saldırı tespit başarısını arttırsa da eğitim ve test süresini uzattığı görülmüştür (İlgün ve Samet 2024).

Bu çalışmada, Basit RNN, LSTM ve Bi-LSTM olmak üzere farklı türde RNN modelleri kullanılarak siber saldırı tespiti gerçekleştirilmiştir. Ayrıca Karar Ağacı (Decision Tree, DT), Rastgele Orman (Random Forest, RF) ve K En Yakın Komşu (K Nearest Neighbour, KNN) modelleri geliştirilmiş ve karşılaştırmalı analizi yapılmıştır. Yapılan analizlere ilişkin sonuçlar incelendiğinde, bütün öğrenme modelleri için %97.76 ve üzeri sınıflandırma doğruluğu elde edilmiştir. En yüksek duyarlılığa sahip modeller ise %99.94 ile KNN ve Bi-LSTM olmuştur. Ayrıca, elde edilen sonuçlar aynı veri seti kullanılarak saldırı tespit sistemi geliştirilmesi amaçlı literatürdeki çalışmalar ile karşılaştırıldığında daha yüksek başarım sağlanmıştır.

KDD'99 veri seti kullanılarak siber saldırı tespitinin araştırıldığı bu çalışmanın geri kalan bölümleri şu şekilde organize edilmiştir: 2. Bölümde, veri seti ve çalışmada ele alınan yöntemler açıklanmıştır. Ardından, 3. Bölümde, çalışma kapsamında elde edilen bulgular sunulmuş ve tartışılmıştır. Son bölümde ise sonuçlar değerlendirilmiştir.

2. Gereç ve Yöntemler

Saldırı tespiti için farklı yapay zeka yöntemleri kullanılarak modeller geliştirilmiş ve modellerin performansları çeşitli metrikler açısından değerlendirilmiştir. Çalışmanın ana hatları Şekil 2'de gösterilmiştir.



Şekil 2. Çalışmanın iş akışı.

2.1. Veri Seti

Çalışmada %10 KDD'99 veri seti kullanılmıştır (UCI KDD Archive 1999). Yaklaşık 25 yıllık bir veri seti olmasına rağmen, zaman yayılımı, içerdiği saldırı çeşitliliği ve toplam veri sayısının büyüklüğü gibi nedenlerle saldırı tespit sistemlerinin başarımının değerlendirilmesinde halen yaygın olarak kullanılmaktadır. Veri setinde, bağlantının başlangıcından bitişine kadar geçen süre (duration), bağlantının durumu (flag), kullanılan iletişim protokolü (protocol_type), ağ hizmeti (service), kaynağa/hedefe iletilen toplam veri miktarı (src_bytes, dst_bytes), geçersiz giriş denemesi sayısı (num_failed_logins), sunucu hata oranı (dst_host_srv_rerror_rate) gibi 41 özellik yer almaktadır. 97,277 normal ve 396,744 saldırı içeren trafik verisi olmak üzere toplamda 494,021 kayıt bulunmaktadır.

2.2. RNN

RNN, düğümler arasında bağlantıların yer aldığı ve yönlendirilmiş döngü oluşturduğu bir sinir ağı türüdür (Şeker vd. 2017, Mikolov vd. 2010). İleri beslemeli bir sinir ağında, tüm girdiler ile çıktıların birbirinden bağımsız olduğunu varsayılırken RNN'de, önceki adımlarda elde edilen çıktılar bir sonraki adımda girdi olarak kullanılır ve böylece zamanla (veya döngüyle) geri besleme sağlanır. İleri beslemeleri ağların aksine RNN'lerin kısa süreli hafıza oluşturma olasılığı vardır (Mikolov vd. 2010) ve geçmiş girdilerden gelen bilgileri kullanarak durumu koruyan bağlantılara sahiptir. Başka bir ifade ile sabit sayıda giriş vektörü kullanmak yerine, tahmin için mevcut bir zaman dilimine kadar tüm giriş bilgilerinden yararlanabilir. Diziye yakın veri noktaları arasındaki korelasyonları içeren (zaman) sıralı verilerle baş etmek için bir çözüm sunar (Schuster ve Paliwal 1997). Basit RNN, model olarak bir girdi katmanı (input layer), bir gizli (hidden/context) katman ve bir çıkış (output) katmanına sahiptir.

LSTM ağları, dizi verilerinin zaman adımları arasındaki uzun süreli bağımlılıkları öğrenebilen bir RNN türüdür (Cakir vd. 2020, Bouktif vd. 2020). Geniş bir kullanım alanına sahip olan LSTM, basit yapılı RNN'lerin eğitim zamanında meydana gelen kaybolan (vanishing) gradyan sorununu çözmek amaçlı Hochreiter ve Schmidhuber (1997) tarafından tasarlanmıştır. LSTM'nin temel yapısını 3 ana kapı oluşturur: 1) giriş kapısı (input gate), 2) çıkış kapısı (output gate) ve 3) unutma kapısı (forget gate). Bu üç kapı, LSTM hücresinin hücre durumunu ve gizlilik durumunu kontrol ederek elde ettiği bilginin depolanması, güncelleme zamanı ve çıkış zamanı konularında karar vermektedir.

Graves ve Schmidhuber tarafından 2005 yılında ilk kez LSTM ağına çift yönlü eğitim uygulanmış ve Bi-LSTM, LSTM'nin gelişmiş bir türü olarak önerilmiştir. Dizi içerisinde yer alan bağımlılıkları iki yönlü olarak ileri ve geri yönde öğrenebilmeyi sağlamaktadır. Hem önceki hem de sonraki bağlamı her bir zaman adımında dikkate alması, modelin daha doğru tahmin gerçekleştirmesine imkan tanımaktadır. Bi-LSTM modeli yapısında çift ayrı LSTM katmanı bulundurur. Katman çiftlerinden biri girdiyi zaman adımı sırası ile ileri yönde işlerken, diğeri ise girdiyi zaman adımı ters sırası ile geri yönde işlemektedir. Elde edilen çıktılar ise daha sonra birleştirilmektedir. Bu çift yönlü yapısı ile Bi-LSTM birçok uygulamada yüksek doğruluk ile kullanılabilmektedir.

2.3. Makine Öğrenmesi Algoritmaları

KNN, en çok kullanılan makine öğrenmesi algoritmalarındandır. Algoritma, kendisine en yakın *k* komşuya göre tahmin yapmaktadır. Komşuların belirlenmesinde yakınlık hesabı için kosinüs ve Öklid mesafesi gibi uzaklık yöntemleri kullanılmaktadır (Dolgun vd. 2009). KNN, sınıflandırma ve regresyon problemleri için etkili yöntemlerden biridir.

DT, ters ağaç yapısına benzer. Verileri dallara ayıran yapısı ile karar verme süreci oluşturur ve bu sürecin sonunda bir tahmin (sınıflandırma/regresyon) gerçekleştirir (Özger 2023). Veriyi bölme kararı, çeşitli ölçütlere (bilgi kazancı, Gini indeksi gibi) göre yapılmaktadır. Örneğin bilgi kazancı en yüksek olan özellik, kök düğümde tutulur ve veriyi bölmek için kullanılan ilk kriterdir. Her bir özelliğin aldığı değerler dalları ifade eder ve her bir dalın sonunda bir karar noktası (iç düğüm) oluşturulur ya da yaprak yer alır. Yapraklar, ağacın en alt seviyesidir ve nihai tahmini sağlar (Ünaldı ve Yalçın 2022).

RF, sınıflandırma ve regresyon amaçlı karar ağaçlarını kullanan bir topluluk (ensemble) yöntemi olarak bilinmektedir. Algoritmanın çalışma mantığı, veri kümesinden rastgele oluşturulan her bir alt küme için DT meydana getirilmesine dayanır. Her bir DT tarafından üretilen tahminlere göre ağırlıklandırma/oylama yapılır ve elde edilen sonuç nihai tahmin olarak sunulur (Özger 2023, Yalçın vd. 2024).

3. Bulgular ve Tartışma

Çalışma kapsamında geliştirilen öğrenme modelleri, Google Colaboratory platformunda Python programla dili ile geliştirilmiştir. Pandas, NumPy, Keras (models, layers, callbacks), Matplotlib, Sklearn (metrics, model_selection, preprocessing, neighbors, tree, ensemble) ve Mlxtend kütüphaneleri kullanılmıştır.

Sinir ağları, sayısal girdiler gerektirdiği için veri setinde yer alan kategorik veriler (protocol_type, flag, service, target) sayısal değerlere dönüştürülmüştür. Daha sonra bütün değerler, 0 ile 1 arasında olacak şekilde normalize edilmiştir. Veri seti, eğitim ve test veri seti olmak üzere 80:20 oranında ikiye ayrılmıştır. Bununla birlikte KNN, DT ve RF modelleri için 5 katlı çapraz doğrulama (cross validation) yöntemi uygulanmıştır. KNN için *k* komşu sayısı, Elbow yöntemi ile 5 olarak elde edilmiştir. Denemeler sonucunda DT ve RF için maksimum derinlik (max_depth) 5 ve RF içinde oluşturulacak ağaç sayısı (n_estimators) 40 olarak belirlenmiştir. Çalışmada Basit RNN, LSTM ve Bi-LSTM yöntemleri kullanılarak 3 farklı RNN derin öğrenme modeli tasarlanmıştır

ve tasarlanan RNN modellerine ilişkin detaylar ise Çizelge 1'de verilmiştir. Her bir RNN modeli, model ile aynı isimli katman ve bir yoğun (dense) katmandan oluşmaktadır. Çizelge 1'de her bir katmandaki nöron, eğitilebilir (trainable) ve toplam (total) parametre sayısı (params) görülebilmektedir. Katmanlarda Adam optimize edici, sigmoid aktivasyon fonksiyonu ve ikili çapraz entropi kaybı (binary crossentropy loss) kullanılmıştır.

Geliştirilen tüm modeller için elde edilen test karmaşıklık matrisleri, Şekil 3'te sunulmuştur. Test veri setinin 79452 adeti saldırı, 19353 adeti ise normal trafik verilerinden oluş-

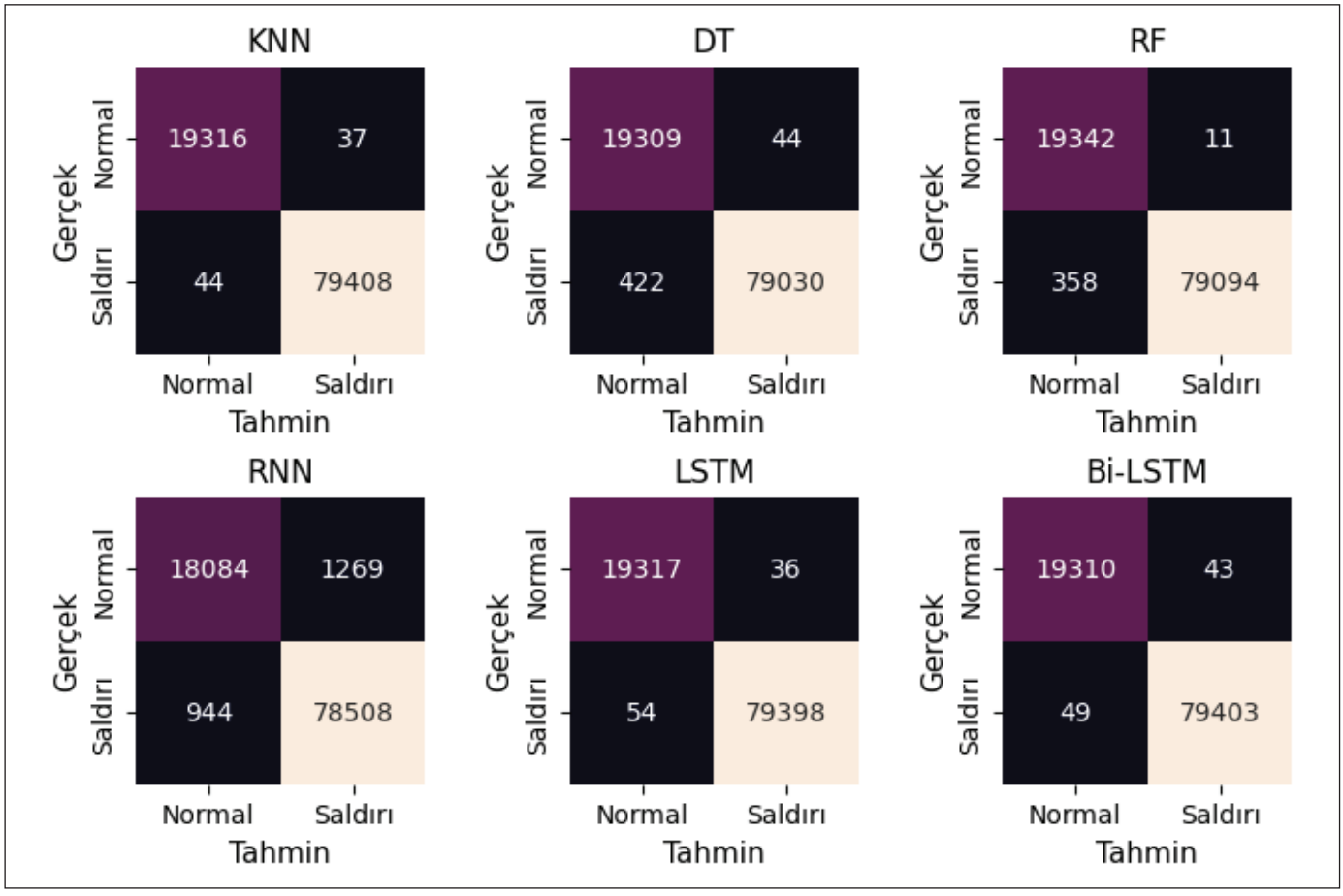
Çizelge 1. Tasarlanan RNN model mimarileri.

Model	Mimari
Basit RNN	Model: "sequential" _____ Layer (type) Output Shape Param # ===== simple_rnn (SimpleRNN) (None, 100) 10200 dense (Dense) (None, 1) 101 ===== Total params: 10301 (40.24 KB) Trainable params: 10301 (40.24 KB) Non-trainable params: 0 (0.00 Byte)
LSTM	Model: "sequential_2" _____ Layer (type) Output Shape Param # ===== lstm_1 (LSTM) (None, 50) 10400 dense_2 (Dense) (None, 1) 51 ===== Total params: 10451 (40.82 KB) Trainable params: 10451 (40.82 KB) Non-trainable params: 0 (0.00 Byte) _____
Bi-LSTM	Model: "sequential_1" _____ Layer (type) Output Shape Param # ===== bidirectional (Bidirection al) (None, 100) 20800 dense_1 (Dense) (None, 1) 101 ===== Total params: 20901 (81.64 KB) Trainable params: 20901 (81.64 KB) Non-trainable params: 0 (0.00 Byte) _____

maktadır. Karmaşıklık matrisleri analiz edildiğinde; saldırı örneklerinden 79408'i KNN tarafından doğru şekilde sınıflandırılmıştır. Yüksek duyarlılıkla saldırıların tespitini gerçekleştiren diğer bir yöntem Bi-LSTM olmuştur ve saldırıya ilişkin örneklerin 79403'ünü doğru tespit etmiştir. RF, normal trafiğe ilişkin örneklerin sadece 11'ini yanlış şekilde saldırı olarak nitelendirmiştir. Bununla birlikte saldırıya ilişkin 358 örneği de normal olarak tahmin etmiştir.

Model performanslarını karşılaştırmak için doğruluk, kesinlik, duyarlılık, F1 ölçütü ve yanlış tespit oranı hesaplanmıştır. Her bir modele ilişkin performans sonuçları Çizelge 2'de sunulmuştur. Sonuçlar incelendiğinde, bütün modellerin %97.76 ve üzeri sınıflandırma doğruluğu sunduğu görülmektedir. KNN ve Bi-LSTM, %99.94 ile en yüksek tespit oranına (duyarlılığa) sahiptir.

Çalışma sonuçları, literatürde aynı veri seti üzerinde yapılan benzer amaçlı çalışmalar ile karşılaştırılmıştır (Çizelge 3). Çakır ve Angin (2021)'nin çalışmalarında en yüksek doğrulukla saldırı tespiti LSTM yöntemi ile elde edilmiştir. Bununla birlikte LSTM kullanılarak elde edilen sınıflandır-



Şekil 3. Karmaşıklık matrisleri.

Çizelge 2. Model performans sonuçları.

Model	Doğruluk	Kesinlik	Duyarlılık	F1-ölçütü	Yanlış Tespit Oranı
Basit RNN	0.9776	0.9841	0.9881	0.9861	0.0656
LSTM	0.9991	0.9995	0.9993	0.9994	0.019
Bi-LSTM	0.9991	0.9995	0.9994	0.9994	0.0022
KNN	0.9992	0.9995	0.9994	0.9995	0.0019
DT	0.9953	0.9994	0.9947	0.9971	0.0023
RF	0.9963	0.9999	0.9955	0.9977	0.0006

Çizelge 3. Aynı veri seti üzerinde literatür karşılaştırması.

Referans	Yöntem	Doğruluk	Kesinlik	Duyarlılık	F1-ölçütü	Yanlış Tespit Oranı
Bu çalışma	**Basit RNN**	0.9776	0.9841	0.9881	0.9861	0.0656
	LSTM	0.9991	0.9995	0.9993	0.9994	0.019
	Bi-LSTM	0.9991	0.9995	0.9994	0.9994	0.0022
	KNN	0.9992	0.9995	0.9994	0.9995	0.0019
	RF	0.9963	0.9999	0.9955	0.9977	0.0006
Çakır ve Angin 2021	**FCN**	0.935	0.927	0.992	0.958	-
	LSTM	0.942	0.939	0.989	0.963	
	TCN	0.941	0.939	0.988	0.963	
Gao vd. 2021	**FNN**	-	~0.9911	~0.8763	~0.9253	-
	LSTM		0.95±0.01	0.84±0.02	0.88±0.01	
Iwendi vd. 2020	**AdaBoost+RF**	0.9908	0.9910	0.9910	0.9910	0.009
	Bagging+RF	0.9940	0.9940	0.9940	0.9990	0.0057
Liu ve Zhang 2020	**CNN**	0.9802	0.9998	0.9981	0.9989	0.02

ma sonuçları incelendiğinde bu çalışmada, (Çakır ve Angin 2021) ve (Gao vd. 2021) çalışmalarından daha yüksek başarım sağlanmıştır. Yanlış tespit oranı açısından en başarılı model, 0.0006 ile bu çalışmada sunulan RF modeli olmuştur. Iwendi ve diğerleri (2020) tarafından önerilen Bagging+RF yaklaşımı da düşük yanlış tespit oranına (0.0057) sahiptir. Çizelge 3'te verilen aynı veri seti kullanılarak yapılan çalışmalar incelendiğinde, en yüksek duyarlılığa sahip modeller ise bu çalışmada sunulan Bi-LSTM (0.9994) ve RF (0.9955) modelleridir.

Çalışma kapsamında KDD'99 veri setinin farklı versiyonları üzerine yapılan çalışmalar da incelenmiştir. Saldırı tespit için LSTM yaklaşımının benimsendiği NSL-KDD veri seti üzerinde yapılan bir çalışmada (Kasongo 2023) %88.13 doğruluk, KDD'99 veri seti üzerinde yapılan bir diğer çalışmada (Laghrissi vd. 2021) %99.49 doğruluk elde edilmiştir. Bi-LSTM kullanan bir saldırı tespit çalışmasında (Pooja ve Shrinivasacharya 2021) ise KDD'99 veri seti için %99 doğruluk sağlanmıştır. Alenazi ve Mishra tarafından önerilen XGBoost modeli ile KDD'99 veri setinde %99.98 ve NSL-KDD veri setinde %99.97 ise doğruluk sağlanmıştır (Alenazi ve Mishra 2024). En yüksek doğrulukla saldırı tespitinin XGBoost modeli tarafından sunulduğu görülmüştür.

4. Sonuç ve Öneriler

Bu çalışmada önerilen modellerin başarımını incelemek için saldırı tespiti alanındaki çalışmalarda yoğunlukla kullanılmış temel kıyaslama (benchmark) veri setlerinden biri olan %10 KDD'99 veri seti kullanılmıştır. Çalışmamızda modellerin uygulandığı bu veri seti sınırlı sayıda saldırı kategorisi içermektedir. Önerilen modeller çok basamaklı saldırıları içeren farklı bir veri seti üzerine uygulandığında modellere ilişkin başarımın daha net bir şekilde gözlemleneceği Çakır ve Angin (2021)'nin çalışmasında sonuçlar kısmında verilmiştir. Bu araştırmada incelenen modeller ele alınarak başarım kıyaslaması yapıldığında KNN, daha yüksek bir test doğruluğu (%99.92), duyarlılık (%99.94) ve F1-ölçütü (%99.95) sunmuştur.

Önerilen modeller, siber güvenliğin sağlanması noktasında siber saldırıların tespit edilmesinde yardımcı olabilir. Bununla birlikte siber güvenlik, teknolojinin gelişmesiyle birlikte sürekli olarak değişen bir alan olduğundan, siber tehditlere karşı savunma stratejilerini sürekli olarak güncellemek ve iyileştirmek önemlidir.

KDD'99 veri setinin, makine öğrenmesi yöntemleri için sınırlamalarının üstesinden gelmek amacıyla Tavallaee ve diğerleri (2009) tarafından NSL-KDD veri seti tanıtılmıştır. Bu nedenle gelecek çalışmada, önerilen yöntemler daha kullanışlı ve kapsamlı farklı veri setleri üzerinde test edilebilir. Gelecek çalışmada, %10 KDD'99 veri seti üzerinde topluluk öğrenme yöntemlerinin başarımı detaylı olarak değerlendirilebilir. Ayrıca bu çalışma, saldırı tespiti için çoklu sınıflandırma yapılacak şekilde genişletilebilir.

Yazar katkıları

Nesibe Yalçın: Çalışmayı planlamış ve tasarlamış, veri setini düzenlemiş, deneyleri gerçekleştirmiş, sonuçları analiz etmiş ve makaleyi yazmıştır, **Semih Çakır:** Sonuçları analiz etmiş ve makaleyi yazmıştır. **Sibel Ünaldı:** Sonuçları değerlendirmiş ve makaleyi yazmıştır.

5. Kaynaklar

Alenazi M., Mishra, S. 2024. Cyberatttack detection and classification in IIoT systems using XGBoost and Gaussian Naïve Bayes: A comparative study. Eng. Technol. Appl. Sci. Res., 14, 4, 15074-15082. Doi: 10.48084/etasr.7664

Bouktif, S., Fiaz, A., Ouni, A., Serhani, MA. 2020. Multi-sequence LSTM-RNN deep learning and metaheuristics for electric load forecasting. Energies, 13, 391. Doi: 10.3390/en13020391

Cakir, S., Toklu, S., Yalcin, N. 2020. RPL attack detection and prevention in the Internet of Things networks using a GRU based deep learning. IEEE Access, 8, 183678-183689. Doi: 10.1109/ACCESS.2020.3029191

Çakır, B., Angin, P. 2021. Zamansal evrişimli ağlarla siber saldırı tespiti: karşılaştırmalı bir analiz. Avrupa Bilim ve Teknoloji Dergisi, 22, 204-211. Doi: 10.31590/ejosat.848784)

Dolgun, MÖ., Özdemir, T., Oğuz, D. 2009. Veri madenciliğinde yapısal olmayan verinin analizi: metin ve web madenciliği. İstatistikçiler Dergisi. (2), s.48-58.

Gao, J., Gan, L., Buschendorf, F. et al. 2021. Omni SCADA intrusion detection using deep learning algorithms. IEEE Internet of Things Journal, 8(2), 951-961. Doi: 10.1109/JIOT.2020.3009180

Graves, A., Schmidhuber, J. 2005. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Netw., 18(5-6), 602-610. Doi: 10.1109/ijcnn.2005.1556215

Hochreiter, S., Schmidhuber, J. 1997. Long Short-Term Memory. Neural Comput. 9(8), 1735-1780. Doi: 10.1162/neco.1997.9.8.1735

Iwendi, C., Khan, S., Anajemba, JH., Mittal, M., Alenezi, M., Alazab, M. 2020. The use of ensemble models for multiple class and binary class classification for improving intrusion detection systems. Sensors, 20, 2559. Doi: 10.3390/s20092559

İlgün, EG., Samet, R. 2024. Veri setine uygulanan ön işlemler ile makine öğrenimi yöntemi kullanılarak geliştirilen saldırı tespit modellerinin performanslarının arttırılması. Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 39(2), 679-692.

Kasongo, SM. 2023. A deep learning technique for intrusion detection system using a Recurrent Neural Networks based framework. Comput. Commun., 199, 113-125. Doi: 10.1016/j.comcom.2022.12.010

Liu, G., Zhang, J. 2020. CNID: research of network intrusion detection based on convolutional neural network. Discrete Dynamics in Nature and Society, 2020(1), 4705982.

Laghrissi, F., Douzi, S., Douzi, K., Hssina, B. 2021. Intrusion detection systems using long short-term memory (LSTM). J. Big Data, 8, 65(1-16). Doi: 10.1186/s40537-021-00448-4

Mikolov, T., Karafiát, M., Burget, L., Černocký, J., Khudanpur, S. 2010. Recurrent neural network based language model. 11th Annual Conference of the International Speech Communication Association - Interspeech 2010, Japan, 1045-1048. Doi: 10.21437/Interspeech.2010-343

Özger, F. 2023. Makine öğrenmesi algoritmalarının hibrit yaklaşımı ile ağ anomalisi tespiti. Yüksek lisans tezi, Sakarya Uygulamalı Bilimleri Üniversitesi, s.18-19.

Pooja, TS., Shrinivasacharya, P. 2021. Evaluating neural networks using Bi-Directional LSTM for network IDS (intrusion detection systems) in cyber security. Global Transitions Proceedings, 2(2), 448-454. Doi: 10.1016/j.gltp.2021.08.017

Schuster, M., Paliwal, KK. 1997. Bidirectional recurrent neural networks. IEEE Trans. Signal Process., 45(11), 2673-2681. Doi: 10.1109/78.650093

Şeker, A., Diri, B., Balık, HH. 2017. Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. GJES, 3(3), 47–64.

Tavallaee, M., Bagheri, E., Lu, W., Ghorbani, AA. 2009. A detailed analysis of the KDD CUP 99 data set. 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, Ottawa, ON, Canada, 1-6. Doi: 10.1109/CISDA.2009.5356528

Thant, YM., Su Thwin, MM., Htwe, CS. 2023. IoT network intrusion detection using long short-term memory recurrent neural network. 2023 IEEE Conference on Computer Applications (ICCA), Yangon, Myanmar, 334-339. Doi: 10.1109/ICCA51723.2023.10182005

UCI KDD Archive. 1999. KDD Cup 1999 Data. Available on April 2024, https://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

Ünaldı, S., Yalçın, N. 2022. Hava kirliliğinin makine öğrenmesi tabanlı tahmini: Başakşehir örneği. Mühendislik Bilimleri ve Araştırmaları Dergisi, 4(1), 35-44. Doi:10.46387/bjesr.1055946

World Economic Forum (WEF). 2024. Global Risks Report 2024. https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf

Yalçın, N., Çakır, S., Ünaldı, S. 2024. Attack detection using artificial intelligence methods for SCADA security. IEEE Internet Things J. Doi: 10.1109/JIOT.2024.3447876