Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 9, S. 75-83, 8, 2025

Araştırma Makalesi

IJANSER
https://as-proceeding.com/index.php/ijanser
ISSN:2980-0811

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 9, pp. 75-83, 8, 2025

Research Article

Derin Öğrenme Uygulamaları ile Kötü Amaçlı Yazılım Tespitinde Model Performans Karşılaştırılması

Abdulhak ALASULU[1*], Bashar ALHAJAHMAD[2]

[1]*Computer Engineering / Graduate School of Natural and Applied Sciences, Siirt University, Turkey*
[2]*Computer Engineering / Faculty of Engineering and Architecture, Siirt University, Turkey*

[*]*(aalasulu73@gmail.com)*
(bashar.aptech@gmail.com)

(Received: 25 July 2025, Accepted: 31 July 2025)

(3rd International Conference on Modern and Advanced Research ICMAR 2025, July 25-26, 2025)

ATIF/REFERENCE: Alasulu, A. & Alhajahmad, B. (2025). Derin Öğrenme Uygulamaları ile Kötü Amaçlı Yazılım Tespitinde Model Performans Karşılaştırılması, *International Journal of Advanced Natural Sciences and Engineering Researches*, 9(8), 75-83.

Özet- Bu çalışmada, günümüzün hızla artan dijitalleşme ve mobil cihaz kullanımıyla paralel olarak yükselen kötü amaçlı yazılım tehditlerine karşı derin öğrenme tabanlı yaklaşımların etkinliği incelenmiştir. Kötü amaçlı yazılım tespiti amacıyla, yaygın makine öğrenimi algoritması olan Uzun Kısa Vadeli Bellek (LSTM) ile derin öğrenme mimarilerinden Tekrarlayan Sinir Ağı (RNN) ve Derin Sinir Ağı (DNN) modelleri test edilmiştir.
Deneylerde, gerçek dünya kötü amaçlı yazılım senaryolarını yansıtan ve %50 kötü amaçlı, %50 iyi huylu bellek dökümlerinden oluşan CIC-MalMem-2022 veri seti kullanılmıştır. Modellerin performansı, doğruluk (accuracy) ve F1 skoru metrikleri üzerinden değerlendirilmiştir.
Elde edilen bulgular, derin öğrenme modelleri olan Derin Sinir Ağı (DNN) ve Tekrarlayan Sinir Ağı (RNN) nın, Uzun Kısa Vadeli Bellek (LSTM)'e kıyasla kötü amaçlı yazılım tespitinde önemli ölçüde daha başarılı olduğunu göstermiştir. Her iki derin öğrenme modeli de %100'e yakın doğruluk oranları ve 0.9998'lik F1 skorları ile üstün performans sergilemiştir. Uzun Kısa Vadeli Bellek (LSTM)de yüksek performans gösterse de, derin öğrenme modellerinin çok daha az yanlış sınıflandırma yaparak daha güvenilir sonuçlar ürettiği belirlenmiştir.
Bu çalışma, derin öğrenme tabanlı yaklaşımların kötü amaçlı yazılım tespitinde son derece etkili olduğunu ve siber güvenlik sistemlerinin güçlendirilmesinde kritik bir rol oynayabileceğini ortaya koymaktadır. Bu bulgular ışığında, gelecekte daha güvenli bilişim sistemleri oluşturmak için derin öğrenme uygulamalarının siber güvenlik alanına entegrasyonu büyük önem taşımaktadır.

Anahtar Kelimeler –Derin Öğrenme, Kötü Amaçlı Yazılım, Makine Öğrenmesi, Siber Güvenlik, Yapay Sinirağları.

I. GİRİŞ

Son zamanlarda insanların internet ortamındaki etkileşimleri dikkate alındığında gözle görülebilir büyük bir artış yaşanmıştır. Bilgi tabanlı sistemlerin artması, dijital ortamların yoğun ilgi görmesi, sosyal medya üzerinden işlemlerin gerçekleşmesi, vb., durumlardan dolayı internet dünyası kötü amaçlı yazılımcıların hedefi haline gelmiştir. Kötü amaçlı yazılımcılar tarafından geliştirilen yazılımlar, çeşitli saldırı türleri ile

sistemleri ele geçirmek, verileri yedeklemek veya değiştirmek, varlığını gizleyerek sürdürmek, kullanıcı hesaplarını ele geçirmek, ağ ortamına bağlı sistemlere zarar vermek, vb. amaçlar doğrultusunda tasarlanmıştır [1].

Panda Security'in raporuna göre bilgisayar korsanları tarafından günlük 230 bin kötü amaçlı yazılım tasarlanmaktadır. Bu sayı her geçen gün daha da artmaktadır. Kaspersky firmasının raporuna göre kimlik avı dolandırıcılığında istenmeyen e-postalar ilk sırada yer almaktadır ve bu eğilimin yakın zamanda değişmesi mümkün değildir. Kötü amaçlı yazılımlar, kurbanları taklit etmek, dolandırıcılık yapmak veya ek kaynaklara erişim sağlamak için kullanılabilen kişisel verileri çalmak için kullanılabilir . IBM X-Force Threat Intelligence Index 2024'e göre, 2023'te çalıntı kimlikler kullanılarak yapılan siber saldırılarda bir önceki yıla göre %71'lik bir artış yaşandı.Ayrıca siber saldırılara maruz kalan küçük şirketler ekonomik anlamda olumsuz etkilenmişlerdir. Accenture firması, web tabanlı saldırılara karşı her yıl yaklaşık 2.4 milyon dolar harcama yaparak tedbirler almıştır. Siber saldırılar sadece bir grup tarafından gerçekleşebilen eylemler olarak tanımlanmayıp, devletler tarafından da desteklenmektedir. Bu durumun en iyi örneği, İsrail tarafından desteklenmiş ve İran Nükleer Santrali'nin çalışmasını engellemek amacıyla geliştirilmiş Stuxnet adı verilen kötü amaçlı yazılımdır [2].

Kötü amaçlı yazılımların etkisini azaltabilmek ve tespitini kolaylaştırmak için bu yazılımların türlerine göre gerçek zamanlı stratejiler geliştirilmektedir. Günümüzde yapay zeka tabanlı yaklaşımlar gerçek zamanlı sistemlere entegre edilerek işlemlerin daha kolaylaştırılması hedeflenmiştir. İnsanüstü karar verebilen bu sistemler gerçek zamanlı işlemlerde daha da etkili rol üstlenebilmektedir [3].

Literatür de bazı çalışmalar bu durumu destekler niteliktedir; M. Akhtar ve ark. [4] botnet saldırılarını tespit etmek için hibrit bir derin öğrenme yaklaşımı önerdi. Önerdikleri yaklaşım evrişimsel sinir ağı (ESA) ile uzun-kısa süreli bellek (LSTM) modelinden oluşmuştu. Kötü amaçlı yazılım sınıflarının tespitinde önerdikleri yaklaşım ile %99 genel doğruluk başarısı elde etmişlerdi [4]. Ding Yuxin ve Zhu Siyi çalışmasında derin inanç ağlarını (DBN) kullanarak kötü amaçlı yazılımların tespitini başarılı bir şekilde gerçekleştirdi. Sınıflandırma sürecinde önerdikleri DBN modelini destek vektör makineleri (SVM), karar ağaçları ve k-en yakın komşu (kNN) yöntemleri ile kıyaslamışlar ve en iyi performansı DBN modeli ile yaklaşık %98 oranında genel doğruluk başarısı sağlamışlardı [5]. Vinayakumar Ravi ve ark. çalışmasında kötü amaçlı yazılımların tespiti için çok görünümlü dikkat tabanlı derin öğrenme modelini önerdiler. Önerilen yaklaşımı hem windows tabanlı verilerde hem de android tabanlı verilerde uyguladılar ve sırasıyla %98 ile %97 genel doğruluk başarısı elde ettiler [6]. J. Pavithra ve S.Samy çalışmasında web işlemlerinin yer aldığı veri kümesini kullanarak web saldırılarının tespitini gerçekleşmişlerdir. Önerdikleri yaklaşımda makine öğrenme yöntemlerini (SVM, rastgele orman) ve Bayes modelini kullandılar. Deneysel analizlerde en başarılı sonucu rastgele orman yöntemi vermiştir ve bu yöntem ile yaklaşık %99 genel doğruluk başarısı sağlamışlardır [7]. Bu çalışmada kötü amaçlı yazılımları tespit etmek için derin öğrenme uygulmaları kullanılarak en iyi performansı veren modeli tespit etmektir.

II. MATYERYAL VE YÖNTEMLER

Bu bölümde, veri seti, model ve eğitim ayrıntıları ele alınmaktadır.

A. Veri Seti

CIC-MalMem-2022 [8] Gizlenmiş kötü amaçlı yazılım, tespit edilmekten ve yok edilmekten kaçınmak için gizlenen kötü amaçlı yazılımdır. Gizlenmiş kötü amaçlı yazılım veri kümesi, bellek aracılığıyla gizlenmiş kötü amaçlı yazılım tespit yöntemlerini test etmek için tasarlanmıştır. Veri kümesi, gerçek dünyada yaygın olan kötü amaçlı yazılımları kullanarak gerçek dünya durumuna mümkün olduğunca yakın bir şekilde temsil etmek için oluşturulmuştur. Casus Yazılım, Fidye Yazılımı ve Truva Atı kötü amaçlı yazılımlarından oluşan bu veri kümesi, gizlenmiş kötü amaçlı yazılım tespit sistemlerini test etmek için kullanılabilecek dengeli bir veri kümesi sağlıyor.

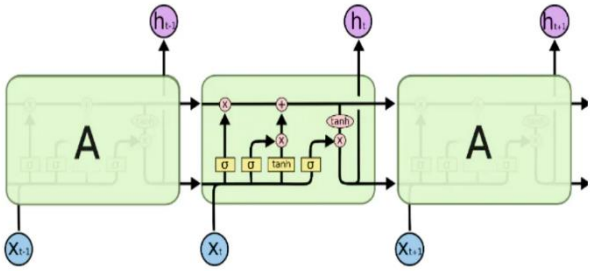
Veri seti %50 kötü amaçlı bellek dökümleri ve %50 iyi huylu bellek dökümlerinden oluşacak şekilde dengelenmiştir. Veri seti 29.298 iyi huylu ve 29.298 kötü huylu olmak üzere toplam 58.596 kayıt içerir. Veri setinde toplam 57 öznitelik bulunuyor. Sayısal veri içermeyen veriler sayısal verilere dönüştürüldü. Bu çalışmada veri seti %70 eğitim %30 test olmak üzere ikiye ayrılmıştır.

B. Modeller

Bu bölüm, deneysel çalışmalarda kullanılan Uzun Kısa Süreli Bellek (LSTM), Tekrarlayan Sinir Ağları (RNN) ve Derin Sinir Ağları (DNN) transfer öğrenme modellerini açıklanmaktadır.

a. Uzun Kısa Süreli Bellek (LSTM)

LSTM ağları, tekrarlayan sinir ağları (RNN) gibi peş peşe bağlanmış sinir hücrelerinden meydana gelmektedir. Hochreiter ve Schmidhuber (1997) tarafından RNN' lerin bir türevi olarak tasarlanmıştır [9]. RNN' lerde gözlenen kaybolan gradyan sorununa çözüm getirebilmesi sebebiyle zaman serilerinde etkili bir şekilde kullanılabilmektedir. Bunu sinir hücresi içerisinde kapı olarak isimlendirilen yapılar vasıtasıyla gerçekleştirmektedir. Bu kapılar ve görevleri şekil 1 de görüldüğü gibi şu şekilde özetlenebilir: İstenilen bilgilerin güncellenmesinden sorumlu Giriş Kapısı, sinir hücresi tarafından kullanılması istenilmeyen bilgilerin yok edildi Unutma Kapısı ve sinir hücresine ait çıkış bilgisinin tutulduğu Çıkış Kapısıdır.



Şekil 1. Bir LSTM Hücresi

Şekil 2' de LSTM ağındaki bir sinir hücresine ait yapı gösterilmekle birlikte bahsedilen kapı yapılarına ait matematiksel eşitlikler aşağıda paylaşılmaktadır [10]:

Giriş Kapısı: $\boldsymbol{it} = \boldsymbol{\sigma}\,(\boldsymbol{wi} \times [\boldsymbol{ht}\!-\!\mathbf{1}, \boldsymbol{xt}] + \boldsymbol{bi})$ (1)

$\bar{\boldsymbol{ct}} = \boldsymbol{tanh}\,(\boldsymbol{wc} \times [\boldsymbol{ht}\!-\!\mathbf{1}, \boldsymbol{xt}] + \boldsymbol{bc}$ (2)

Unutma Kapısı: $\boldsymbol{ft} = \boldsymbol{\sigma}\,(\boldsymbol{wf} \times [\boldsymbol{ht}\!-\!\mathbf{1}, \boldsymbol{xt}] + \boldsymbol{bf})$ (3)

Çıkış Kapısı: $\boldsymbol{ct} = \boldsymbol{ft} * \boldsymbol{ct}\!-\!\mathbf{1} + \boldsymbol{it} * \bar{\boldsymbol{ct}}$ (4)

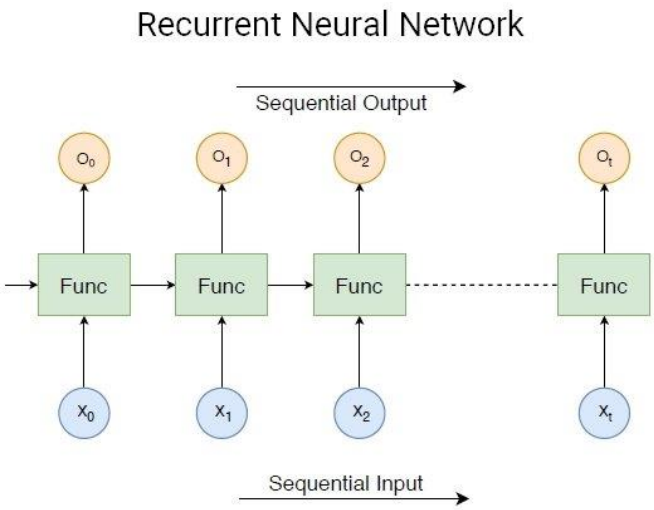
$\boldsymbol{ot} = \boldsymbol{\sigma}\,(\boldsymbol{wo} \times [\boldsymbol{ht}\!-\!\mathbf{1}, \boldsymbol{xt}] + \boldsymbol{bo})$ (5)

$\boldsymbol{ht} = \boldsymbol{ot} * \boldsymbol{tanh}(\boldsymbol{ct})$ (6)

b. Tekrarlayan Sinir Ağları (RNN)

Tekrarlayan sinir ağı (RNN), sıralı veri beslemesini kullanan başka bir yapay sinir ağı sınıfıdır. RNN'ler, sıralı giriş verilerinin zaman serisi sorununu ele almak için geliştirilmiştir.

Şekil 2 de gösterildiği gibi bir RNN'nin girdisi, geçerli girdi ve önceki örneklerden oluşur. Bu nedenle, düğümler arasındaki bağlantılar zamansal bir dizi boyunca yönlendirilmiş bir grafik oluşturur. Ayrıca, bir RNN'deki her nöron, önceki örneklerden gelen hesaplamanın bilgilerini tutan bir dahili belleğe sahiptir.



Şekil 2 RNN algortimasının çalışma prensibi

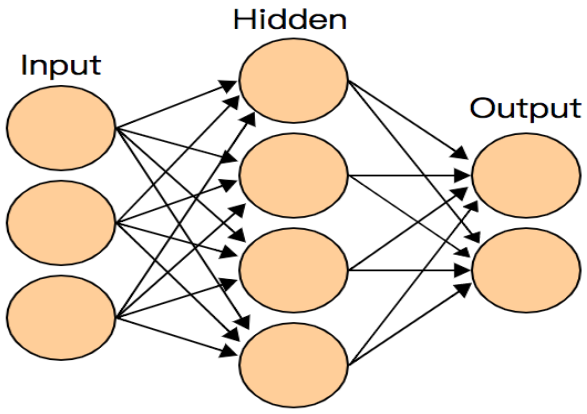
c. Derin Sinir Ağları (DNN)

Derin Sinir Ağları (DNN), karmaşık örüntüleri öğrenerek tahminler yapan veya kararlar veren bir modelleridir. Bunu şöyle başarır:

Veri Alma ve Katmanlı İşleme: DNN'ler, resimdeki pikseller veya metindeki kelimeler gibi ham veriyi alır. Bu veri, bir dizi gizli katmandan geçer. Her katman, bir önceki katmandan aldığı bilgiyi işleyerek daha karmaşık ve soyut özellikler çıkarır (örneğin, bir resimde önce kenarları, sonra şekilleri, sonra nesneleri tanıma).

Ağırlıklar ve Aktivasyonlar: Her katmandaki "nöronlar", gelen bilgilere ağırlıklar uygular. Bu ağırlıklar, bilginin önemini belirler ve bir aktivasyon fonksiyonu ile dönüştürülerek ağın karmaşık ilişkileri öğrenmesini sağlar. Bu ağırlıklar, modelin aslında "öğrendiği" parametrelerdir.

Kayıp ve Geri Yayılım ile Öğrenme: Model bir tahmin yaptığında, bu tahmini doğru sonuçla karşılaştırır ve ne kadar hata yaptığını bir kayıp fonksiyonuyla hesaplar. Bu hata, ağın içinde geri yayılır (backpropagation). Geri yayılım sırasında, hatayı azaltmak için nöronların ağırlıkları ve sapmaları otomatik olarak ayarlanır. Şekil 3 te gösterildiği gibi girdileri gizli bir katmana aldıktan sonra işleme sokup çıktı olarak vermektedir.



Şekil 3 DNN algortimasının çalışma mantığı

C. *Eğitim Detayları*

Model, epoksi 10 olacak şekilde eğitilmiştir. Modelin genelleme performansını ve aşırı öğrenme durumunu izlemek amacıyla, her dönem sonunda doğrulama veri kümesi üzerinde çeşitli metrikler hesaplanmıştır. Bu metrikler arasında doğrulama kaybı doğrulama doğruluğu verileri bulunmaktadır.

Eğitim, epoksi ile belirlenen tüm dönemler tamamlanana kadar devam etmiş ve sonrasında F1 skorları raporlanmıştır.

III. DENEY VE BULGULAR

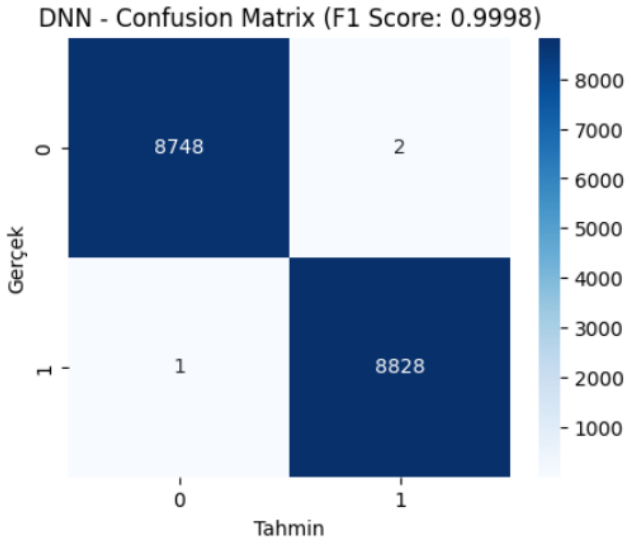
Bu bölümde, kötü amaçlı yazılım tespiti için transfer öğrenme modellerinin performansları CIC-MalMem-2022 veri seti üzerinde incelenmiştir .Bu bağlamda, farklı kapasitelere sahip modeller aynı eğitim parametreleri kullanılarak karşılaştırılmıştır. Çalışmanın amacı, CIC-MalMem-2022 veri seti için en yüksek performansı veren modeli tespit etmektir.Yapılan deney sonucunda Tablo 1 de gösterildiği gibi LSTM F1 skoru 0.9991, RNN F1 skoru 0.9998 ve DNN için F1 skoru 0.9998 olarak tespit edilmiştir Her modelin doğruluk değerleri ve F1 skorları aşağıdaki Tablo 1'de gösterilmiştir.

Tablo 1. Transfer öğrenme modellerinin doğruluk ve F1 skor verileri açısından karşılaştırılması.

Models	Accuracy	F1 Skor
LSTM	0.9991	0.9991
RNN	1.000	0.9998
DNN	1.000	0.9998

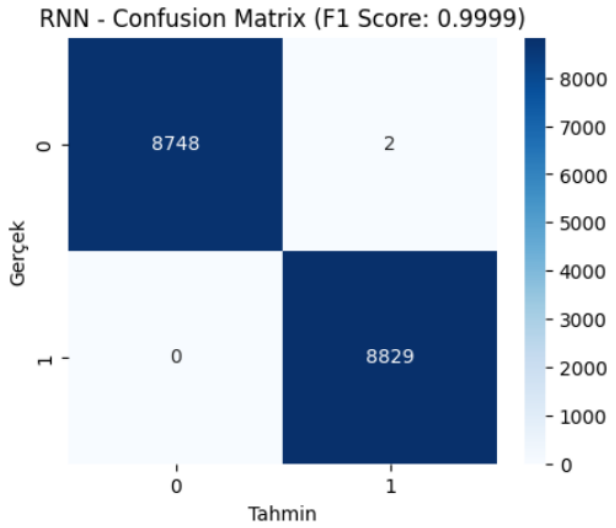
Tablo 1 incelendiğinde tüm modellerin başarı gösterdiği görülmüştür. LSTM ,DNN ve RNN'ye göre daha az başarılıdır. Bu sonuçlar ışığında DNN , RNN ve LSTM için aşağıdaki grafikler oluşturulmuştur.

Şekil 4 te DNN algoritması ile eğitim sonucunda confusion matrixinde 17576 veriyi doğru sınıflandırmış sadece 3 adet veriyi yanlış sınıflandırıldığı tespit edilmiştir.



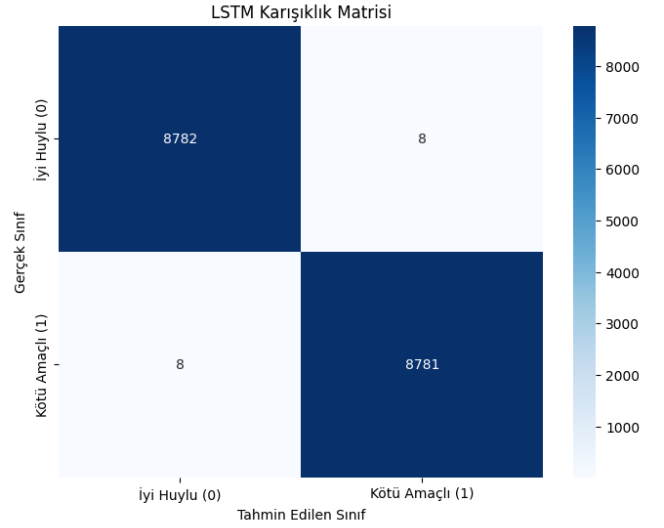
Şekil 4 DNN için Confusion Matrixi

Şekil 5 te RNN algoritması ile eğitim sonucunda confusion matrixinde 17577 veriyi doğru sınıflandırmış sadece 2 adet veriyi yanlış sınıflandırıldığı tespit edilmiştir.



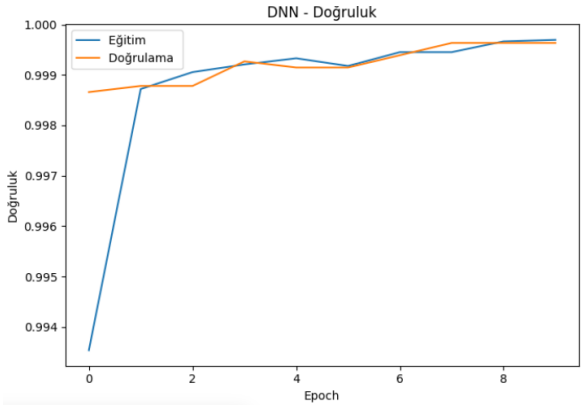
Şekil 5 RNN için Confusion Matrixi

Şekil 6 da LSTM algoritması ile eğitim sonucunda confusion matrixinde 17563 veriyi doğru sınıflandırmış sadece 16 adet veriyi yanlış sınıflandırıldığı tespit edilmiştir.



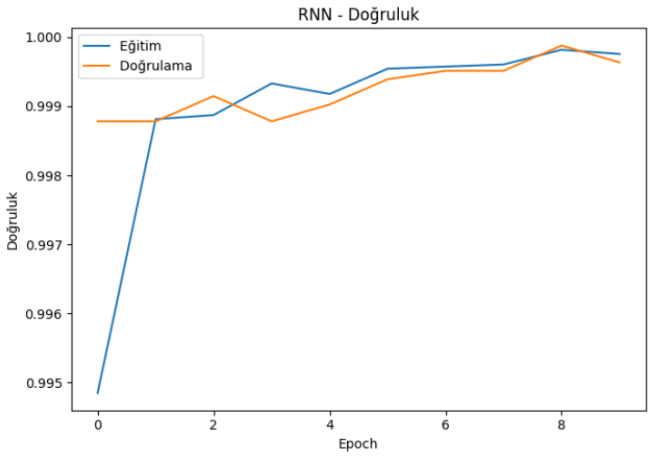
Şekil 6 LSTM Confusion Matrixi

Şekil 7 de DNN algoritması için doğruluk grafiği yer almaktadır. Grafiğe göre doğruluk değeri kabul edilebilir ve yüksek bir değerdedir.



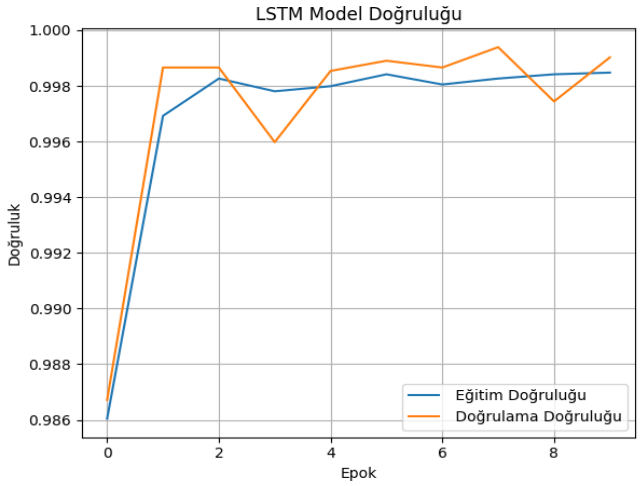
Şekil 7 DNN doğruluk Grafiği

Şekil 8 de RNN algoritması için doğruluk grafiği yer almaktadır. Grafiğe incelendiğinde doğruluk değeri yeterince iyi bir değerdedir.



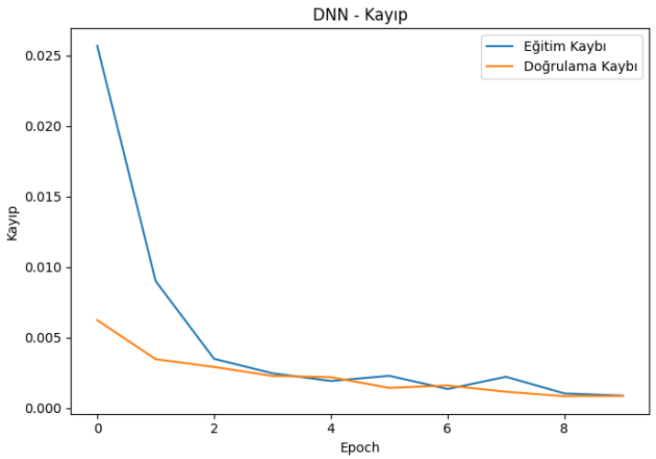
Şekil 8 RNN doğruluk Grafiği

Şekil 9 de LSTM algoritması için doğruluk grafiği yer almaktadır. Grafik incelendiğinde doğruluk değeri RNN ve DNN ye göre kısmen daha düşük bir değerde olmasına karşın yine iyi bir seviyededir.



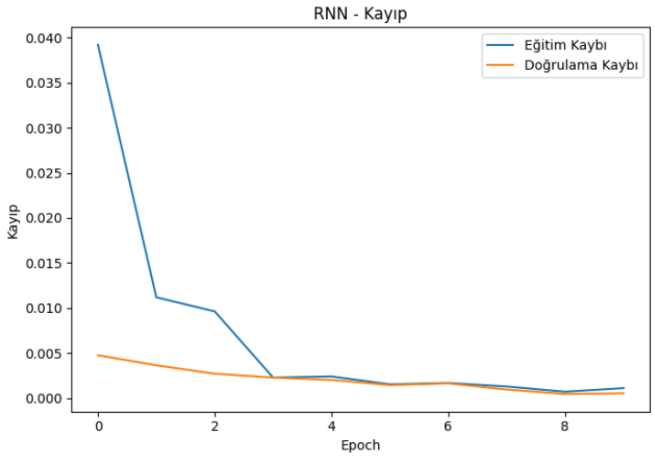
Şekil 9 LSTM doğruluk grafiği

Şekil 10 da DNN algoritması için kayıp grafiği gösterilmiştir. Grafik incelendiğinde makul birkayıp olduğu gözlemlenmiştir.



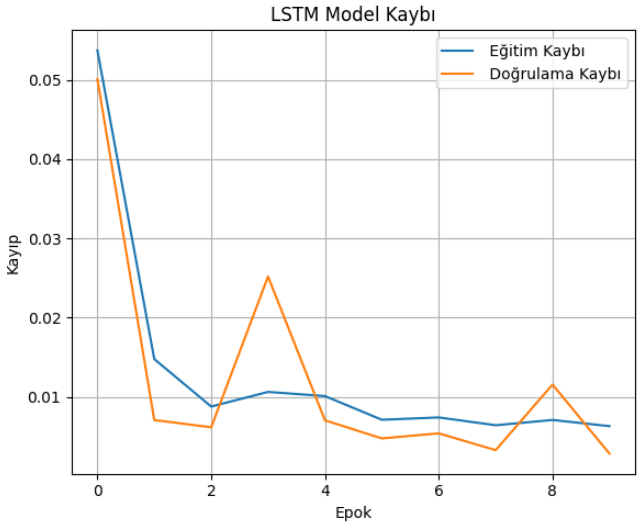
Şekil 10 DNN kayıp grafiği

Şekil 11 de RNN algoritması için kayıp garfiği gösterilmiştir. Graifğe bakıldığında önemli derecede bir kayıp olmadığı saptanmıştır.



Şekil 11 RNN kayıp grafiği

Şekil 12 de LSTM algoritması için kayıp grafiği gösterilmiştir. Dnn ve RNN ye göre kısmen daha fazla kayıp yaşanmış ama yine de kabul edilebilecek bir kayıp söz konusudur.



Şekil 12 LSTM kayıp grafiği

IV. SONUÇ

Bu çalışmada, günümüzün hızla artan dijitalleşme ve mobil cihaz kullanımına paralel olarak yükselen kötü amaçlı yazılım tehditlerine karşı derin öğrenme tabanlı yaklaşımların etkinliği incelenmiştir. Kötü amaçlı yazılım tespiti amacıyla Uzun Kısa Süreli Bellek (LSTM) , Tekrarlayan Sinir Ağı (RNN) ve Derin Sinir Ağı (DNN) olmak üzere üç farklı modelin performansları, CIC-MalMem-2022 veri seti üzerinde karşılaştırılmıştır. Bu veri seti, dengeli bir dağılımla (%50 kötü amaçlı, %50 iyi huylu) gerçek dünya senaryolarını temsil eden bellek dökümlerinden oluşmaktadır.

Deneyler sonucunda elde edilen bulgular, derin öğrenme modelleri olan DNN ve RNN'in, LSTM ye kıyasla kötü amaçlı yazılım tespitinde önemli ölçüde daha başarılı olduğunu göstermiştir. Özellikle DNN ve RNN modelleri, neredeyse %100'e yakın doğruluk oranları ve 0.9998 gibi yüksek F1 skorları ile üstün bir performans sergilemiştir. LTSM da yüksek bir doğruluk (0.9991) ve F1 skoru (0.9991) elde etse de, karmaşıklık matrisleri incelendiğinde, DNN'in yalnızca 3, RNN'in ise sadece 2 yanlış sınıflandırma

yaparken, LSTM'nin 16 yanlış sınıflandırma yaptığı görülmüştür. Bu durum, derin öğrenme modellerinin çok daha az hata oranıyla daha güvenilir sonuçlar ürettiğini teyit etmektedir.

Bu bulgular, kötü amaçlı yazılımların giderek karmaşık hale geldiği günümüzde, geleneksel makine öğrenmesi yöntemlerine kıyasla derin öğrenme uygulamalarının siber güvenlik sistemlerinde daha etkili ve faydalı olabileceğini açıkça ortaya koymaktadır. Elde edilen yüksek performans değerleri, derin öğrenme tabanlı yaklaşımların, gelecekteki kötü amaçlı yazılım tespit ve önleme stratejileri için güçlü bir temel oluşturduğunu göstermektedir. Bu tür modellerin entegrasyonuyla, daha güvenli bilişim sistemleri kurulması ve dijital ortamdaki kullanıcıların kötü amaçlı yazılımlardan korunması mümkün olacaktır.

Gelecekteki çalışmalar, farklı derin öğrenme mimarileri ve daha geniş, çeşitlendirilmiş veri setleri üzerinde ek deneyler yaparak bu modellerin genellenebilirliğini ve sağlamlığını daha da artırmaya odaklanabilir. Ayrıca, gerçek zamanlı kötü amaçlı yazılım tespiti için bu modellerin performans optimizasyonları da önemli bir araştırma alanı olacaktır.

KAYNAKLAR

[1] M.D. Yılmaz "Malware classification with using deep learning", Comput. Informatics, vol. 2, no. 2, pp. 21–40, 2022.

[2] U.H. Tayyab, F.B. Khan, M.H. Durad, A. Khan, and Y.S. Lee "A Survey of the Recent Trends in Deep Learning Based Malware Detection", J. Cybersecurity Priv., vol. 2, no. 4, pp. 800–829, 2022.

[3] M. Toğaçar "Siber Saldırılara Karşı Kullanılan Makine Öğrenme Yöntemlerinin Web Uygulamalarında Güvenlik Etkinliğinin Ölçümü", Gazi Üniversitesi Fen Bilim. Derg. Part C Tasarım ve Teknol., vol. 9, no. 4, pp. 608–620, 2021.

[4] M.S. Akhtar and T. Feng "Detection of Malware by Deep Learning as CNN-LSTM Machine Learning Techniques in Real Time", Symmetry (Basel)., vol. 14, no. 11, pp. 2308, 2022.

[5] D. Yuxin and Z. Siyi "Malware detection based on deep learning algorithm", Neural Comput. Appl., vol. 31, no. 2, pp. 461–472, 2019.

[6] V. Ravi, M. Alazab, S. Selvaganapathy, and R. Chaganti "A Multi-View attention-based deep learning framework for malware detection in smart healthcare systems", Comput. Commun., vol. 195, pp. 73–81, 2022.

[7] J. Pavithra and S. Selvakumara Samy "A Comparative Study on Detection of Malware and Benign on the Internet Using Machine Learning Classifiers", Math. Probl. Eng., vol. 2022, pp. 1–8, 2022.

[8] Tristan Carrier, Princy Victor, Ali Tekeoglu, Arash Habibi Lashkari," Detecting Obfuscated Malware using Memory Feature Engineering", The 8th International Conference on Information Systems Security and Privacy (ICISSP), 2022

[9]Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[10]Wen, Y., Lin, P., & Nie, X. (2020, March). Research of stock price prediction based on PCA-LSTM model. In IOP Conference Series: Materials Science and Engineering (Vol. 790, No. 1, p. 012109). IOP Publishing.