Teknoloji Fakültesi

BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Mobil Platformlarda Phishing Saldırılarına Karşı Federe Öğrenme Tabanlı Etkin Savunma Sistemi Geliştirilmesi

Bitirme Projesi 1. Ara Raporu

Bilgisayar Mühendisliği Bölümü

DANIŞMAN

Doç. Dr. Kazım YILDIZ

İSTANBUL, 2025

MARMARA ÜNİVERSİTESİ
TEKNOLOJİ FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Marmara Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Öğrencileri Eren DOĞAN tarafından "**Mobil Platformlarda Phishing Saldırılarına Karşı Federe Öğrenme Tabanlı Etkin Savunma Sistemi Geliştirilmesi**" başlıklı proje çalışması, xxx tarihinde savunulmuş ve jüri üyeleri tarafından başarılı bulunmuştur.

Jüri Üyeleri

Dr. Öğr. Üyesi xxx xxx Marmara Üniversitesi	**(Danışman)**	(İMZA)……..
Prof. Dr. Xxx xxx Marmara Üniversitesi	(Üye)	(İMZA)……..
Prof. Dr. Xxx xxx Marmara Üniversitesi	(Üye)	(İMZA)……..

ÖNSÖZ

Proje çalışmam süresince karşılaştığım her türlü sorunda sabırla yardım ve bilgilerini esirgemeyen, tüm desteğini yanımda hissettiğim değerli hocam Arş. Gör. Büşra Büyüktanır'a en içten teşekkürlerimi sunarım. Kendisinin yol gösterici tavsiyeleri ve sürekli motivasyon sağlaması, bu çalışmanın başarıya ulaşmasında büyük bir paya sahiptir.

Ayrıca, proje fikrinin oluşmasında sağladığı katkılardan dolayı Doç. Dr. Kazım Yıldız'a da teşekkür ederim. Onun yönlendirici önerileri sayesinde çalışmanın kapsamı ve uygulama süreci daha planlı bir şekilde yürütülmüştür.

İçindekiler Tablosu

ÖZET

Mobil Platformlarda Phishing Saldırılarına Karşı Federe Öğrenme Tabanlı Etkin Savunma Sistemi Geliştirilmesi

Bu araştırma önerisinin temel amacı, internet kullanıcılarını artan phishing (oltalama) saldırılarına karşı korumak için mobil platformlarda gerçek zamanlı çalışan, kullanıcı gizliliğini ihlal etmeden sürekli güncellenebilen ve kullanıcının tarama alışkanlıklarına adapte olarak değişen saldırı tekniklerine uyum sağlayabilen, federe öğrenme tabanlı derin öğrenme modeli geliştirmektir. Geliştirilecek model, sayfa kaynak kodlarından öznitelik çıkararak phishing web sitelerini etkin ve gerçek zamanlı bir şekilde tespit edecek şekilde tasarlanacaktır. Böylece, kullanıcıların kişisel bilgilerinin korunması, internet güvenliğinin artırılması ve phishing saldırıları sonucu oluşabilecek mali ve operasyonel zararların minimize edilmesi sağlanacaktır.

Derin öğrenme modeli olarak Convolutional Neural Network (CNN) tabanlı bir yapı geliştirilecektir. Model için en iyi parametreleri bulmak amacıyla hiperparametre optimizasyonu aşaması için Bayesian optimizasyon yöntemi kullanılacaktır. Model eğitimi ve optimizasyonu, Microsoft Azure bulut platformu üzerinde gerçekleştirilecektir. Eğitimde Alexa, PhishTank, OpenPhish, PhishRepo, PhishStats ve Kaggle gibi kaynaklardan elde edilen veri setleri kullanılacaktır. Bu çeşitlilik, modelin farklı phishing saldırı türlerine karşı dayanıklılığını artırmayı hedeflemektedir.

Federe öğrenme yaklaşımı, kullanıcının verilerini cihazda tutarak gizliliği korurken modelin güncellenmesini sağlar. Bu amaçla, TensorFlow'un Federated Learning kütüphanesi kullanılacaktır. Eğitim süreçlerinde TensorFlow ve Jupyter Notebook gibi araçlar tercih edilecektir. Araştırmanın nihai hedefi, geliştirilen modelin yalnızca teorik düzeyde kalmayıp, toplumun genel faydasına sunulabilecek, gerçek zamanlı çalışan ve kullanıcı gizliliğini ön planda tutan bir güvenlik çözümü olarak hayata geçirilmesidir. Bu doğrultuda, yapay zeka modeli, doğrudan kullanıcı cihazında lokal olarak çalışacak ve Kotlin ve Jetpack Compose kullanılarak geliştirilecek anti-phishing özellikli bir web tarayıcısına entegre edilecektir. Tarayıcı, kullanıcıların güvenli bir internet deneyimi yaşamalarını sağlayarak, ziyaret edilen sitelerin güvenilirliğini anlık olarak

değerlendirecek ve olası tehditleri tespit ettiğinde kullanıcıyı uyaracaktır. Yapay zeka modelinin cihaz üzerinde lokal olarak çalışması ve tarayıcıya entegre edilmesi sayesinde, bu güvenlik çözümünün yıllar sonra bile kullanıcılar tarafından güvenli bir şekilde kullanılabilir olması hedeflenmektedir.

Modelin etkinliğini değerlendirmek için bir kullanıcı çalışması planlanmaktadır. Katılımcılara çeşitli phishing ve yasal web siteleri gösterilerek bu sitelerin phishing olup olmadığını değerlendirmeleri istenecektir. Elde edilen sonuçlar, modelin insan performansına kıyasla ne kadar etkili olduğunu belirlemek için kullanılacaktır.

Proje, dokuz aylık bir süre boyunca iş paketleri ve zaman çizelgesine uygun olarak yürütülecektir. İş paketleri; veri toplama ve ön işleme, CNN geliştirilmesi, hiperparametre optimizasyonu, model eğitimi, federe öğrenme altyapısının kurulması, web tarayıcısının geliştirilmesi, performans testleri, güvenlik önlemleri, kullanıcı çalışması ve pilot uygulama aşamalarını içermektedir. Her iş paketi için olası riskler analiz edilmiş ve alternatif planlar oluşturulmuştur.

Projenin yaygın etkileri, kullanıcıların phishing saldırılarına karşı korunmasına ve siber güvenlik bilincinin artırılmasına katkı sağlamayı içermektedir. Araştırma sonuçları, bilimsel makaleler ve konferans bildirileri ile paylaşılacaktır. Geliştirilen sistemin, kullanıcı güvenliğini artırması ve siber saldırılardan kaynaklanan kayıpları azaltması beklenmektedir.

Mart, 2025 **Eren DOĞAN**

ABSTRACT

Developing a Federated Learning Based Effective Defense System Against Phishing Attacks on Mobile Platforms

The main purpose of this research proposal is to develop a federated learning-based deep learning model that works in real time on mobile platforms to protect internet users against increasing phishing attacks, can be continuously updated without violating user privacy, and can adapt to changing attack techniques by adapting to user browsing habits. The model to be developed will be designed to detect phishing websites effectively and in real time by extracting features from page source codes. Thus, personal information of users will be protected, internet security will be increased, and financial and operational losses that may occur as a result of phishing attacks will be minimized. A Convolutional Neural Network (CNN)-based structure will be developed as a deep learning model. Bayesian optimization method will be used for the hyperparameter optimization phase in order to find the best parameters for the model. Model training and optimization will be performed on the Microsoft Azure cloud platform. Data sets obtained from sources such as Alexa, PhishTank, OpenPhish, PhishRepo, PhishStats and Kaggle will be used in the training. This diversity aims to increase the model's resistance to different types of phishing attacks.

The federated learning approach allows the model to be updated while preserving privacy by keeping the user's data on the device. For this purpose, TensorFlow's Federated Learning library will be used. Tools such as TensorFlow and Jupyter Notebook will be preferred in the training processes. The ultimate goal of the research is to implement the developed model not only at a theoretical level, but also as a security solution that can be offered to the general benefit of society, works in real time and prioritizes user privacy. In this direction, the artificial intelligence model will be integrated into an anti-phishing web browser that will run locally on the user's device and will be developed using Kotlin and Jetpack Compose. The browser will instantly evaluate the reliability of the visited sites, ensuring that users have a safe internet experience, and will warn the user when it detects potential threats. Thanks to the artificial intelligence model running locally on the device and being integrated into the browser, it is aimed for this security solution to be used securely by users even years later.

A user study is planned to evaluate the effectiveness of the model. Participants will be shown various phishing and legitimate websites and asked to evaluate whether these sites are phishing or not. The results obtained will be used to determine how effective the model is compared to human performance.

The project will be carried out in accordance with the work packages and timeline for a period of nine months. The work packages include data collection and preprocessing, CNN development, hyperparameter optimization, model training, establishment of federated learning infrastructure, development of web browser, performance tests, security measures, user study and pilot application stages. Possible risks were analyzed for each work package and alternative plans were created.

The widespread impacts of the project include contributing to the protection of users against phishing attacks and increasing cybersecurity awareness. The research results will be shared through scientific articles and conference proceedings. The developed system is expected to increase user security and reduce losses caused by cyber attacks.

March, 2025 **Eren DOĞAN**

KISALTMALAR

- **CNN (Convolutional Neural Network)**: Evrişimli Sinir Ağı
- **CSS (Cascading Style Sheets)**: Basamaklı Stil Şablonları
- **GDPR (General Data Protection Regulation)**: Genel Veri Koruma Tüzüğü (AB)
- **HTML (HyperText Markup Language)**: Köprü Metni Biçimlendirme Dili
- **KVKK (Kişisel Verilerin Korunması Kanunu)**: Türkiye'deki Veri Koruma Yasası
- **SMS (Short Message Service)**: Kısa Mesaj Servisi
- **SSL (Secure Sockets Layer)**: Güvenli Yuva Katmanı (Genelde TLS/SSL şeklinde geçer)
- **TLS (Transport Layer Security)**: Taşıma Katmanı Güvenliği
- **TÜİK (Türkiye İstatistik Kurumu)**
- **URL (Uniform Resource Locator)**: Tekdüzen Kaynak Bulucu

ŞEKİL LİSTESİ

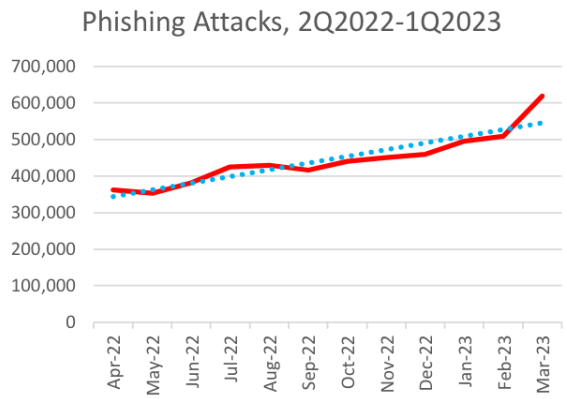
1. GİRİŞ

Bilişim teknolojilerinin modern yaşamda entegrasyonu giderek daha karmaşık ve derin bir yapı kazanmaktadır. Türkiye İstatistik Kurumu'nun (TÜİK) 2023 tarihli Hanehalkı Bilişim Teknolojileri Kullanım Araştırması verilerine göre, Türkiye'deki hanelerin %95,5'i internet erişimine sahiptir; bu oran, 2018 yılında %83,8 olarak ölçülmüştür. Aynı çalışma, bireylerin internet kullanım oranının son beş yıl içinde %14,2 oranında artarak %87,1'e ulaştığını ortaya koymaktadır. Bu bulgu, artan internet erişimiyle birlikte daha çok bireyin dijital dünyaya entegre olduğunu ve internet kullanımının yaygınlaştığını göstermektedir (TÜİK, 2023).

Artan internet kullanımı, kişisel bilgi güvenliği ihlalleri, zararlı yazılım saldırıları ve benzeri çeşitli siber tehditlerin de paralel olarak artmasına yol açmıştır (Ömer ve ark., 2023). Bu tehditler, hem bireysel kullanıcılar hem de ulus-devletlerin güvenlik yapıları için önemli riskler teşkil etmektedir (Bıçakcı, Ergun ve Çelikpala, 2016). Küresel düzeyde, kişisel bilgi güvenliğine yönelik saldırıların sayısı her geçen gün artmakta ve bu durum çok daha tehlikeli bir boyut kazanmaktadır.

Kişisel bilgi güvenliğinin ihlalinde en çok kullanılan yöntemlerden biri phishing (oltalama) saldırılarıdır. Phishing saldırılarının yaygınlığı, dünya genelinde 45 ülkede Anti-Phishing Working Group (APWG, 2016) tarafından yapılan bir çalışmada ayrıntılı olarak incelenmiştir. Bu çalışmanın bulgularına göre, phishing saldırılarının en yaygın olduğu ülkeler sıralamasında ilk sırada %47,09 ile Çin yer almaktadır. Bu ülkeyi %42,88 ile Türkiye ve %38,98 ile Tayvan takip etmektedir.

Phishing saldırıların zaman içinde nasıl arttığını ortaya koyan güncel verilere göre (APWG, 2023), bu saldırılar 2022'nin ikinci çeyreğinden 2023'ün ilk çeyreğine kadar sürekli bir artış göstermiştir. Şekil 1'de sunulan verilere göre, 2022'nin Nisan ayında yaklaşık 400.000 olan saldırı sayısı, 2023'te 600.000'i aşmış ve yeni bir zirveye ulaşmıştır. Bu trend, phishing saldırılarının artan bir yaygınlık ve tehlike arz ettiğini göstermektedir.



Şekil 1.1 Oltalama Saldırıları İstatistiği (APWG, 2023)

Phishing saldırıları, kişisel ve gizli verilere yetkisiz erişim elde etmek ve bu verileri kötü niyetli amaçlarla kullanmak için bilişim teknolojilerinin sunduğu fırsatları değerlendiren bir dolandırıcılık yöntemidir (Ünver ve Mirzaoğlu, 2011). Bu saldırılar, genellikle sosyal mühendislik teknikleri kullanılarak bireylerin güvenini kötüye kullanmaya dayanmaktadır (Hekim, 2015). Phishing girişimleri genellikle e-posta, SMS ya da sosyal medya kanalları üzerinden zararlı bağlantılar gönderilmesi yoluyla yapılmaktadır (Buber, Diri ve Şahingöz, 2017; Jakobsson ve Myers, 2007). Birleşik Krallık Kamu Sektörü Bilgileri tarafından 2023 yılında yapılan Cyber Security Breaches Survey, işletmelerin ve STK'ların en yaygın maruz kaldığı siber tehdit türünün %79 oranıyla phishing olduğunu ortaya koymuştur. Bu çalışma, önceki yıllarda elde edilen bulgularla da uyumludur (Birleşik Krallık Kamu Sektörü Bilgileri, 2021, 2022, 2023). Bu bulgular, kişisel bilgi güvenliğinin ihlalinde phishing saldırılarının merkezi bir rol oynadığını göstermektedir.

Bu siber tehditler, belirli birey ya da gruplardan öte tüm toplumu etkileyen büyük bir sorun haline gelmiştir. Finansal kayıplar, kişisel verilerin çalınması, iş süreçlerinin kesintiye uğramı ve geniş çaplı güvenlik ihlalleri gibi çok sayıda olumsuz sonuca neden olabilmektedir (Wassermann, Meyer, Goutal ve Riquet, 2023). Finansal kayıplar, phishing saldırıları sonucunda kullanıcıların banka hesap bilgilerinin, kredi kartı numaralarının ve diğer finansal bilgilerin ele geçirilmesi sonucunda oluşmaktadır (Sharif ve Mohammed, 2022).

Phishing saldırıları sırasında kullanılan aldatıcı teknikler sayesinde, genellikle e-posta, mesajlaşma uygulamaları ve sosyal medya aracılığıyla kullanıcılar meşru olmayan web sitelerine yönlendirilmekte ve bu süreçte kişisel ve hassas bilgiler ele geçirilmektedir. Bu sürecin sonuçları finansal kayıplar, kişisel bilgilerin çalınması, iş süreçlerinin aksaması ve

güvenlik ihlalleri şeklinde tezahür etmektedir (Goel, Williams ve Dincelli, 2017).

Kişisel verilerin çalınması, bu siber tehditlerin önemli bir sonucu olarak ortaya çıkmıştır. Kullanıcıların isim, adres, telefon numarası ve sağlık kayıtları gibi hassas bilgilerinin phishing saldırıları yoluyla ele geçirilmesi, bireylerin mahremiyetini ihlal etmekte ve kimlik hırsızlığı gibi daha büyük sorunlara yol açabilmektedir (Kamiya ve diğerleri, 2019).

Kurumsal seviyede ise, phishing saldırılarının iş süreçlerinin aksamasına neden olduğu gözlemlenmiştir. Büyük şirketlerde çalışan bireylerin phishing saldırılarına maruz kalması, şirket içi ağların güvenliğini tehdit etmekte, veri ihlallerine yol açabilmekte ve iş sürekliliğini tehlikeye sokabilmektedir. Bu durum, şirketlerin itibarına ve finansal sağlığına zarar verebilmektedir (Kamiya ve diğerleri, 2019).

Mobil internetin ve akıllı telefon kullanımının yaygınlaşması, mobil cihazları oltalama saldırıları için popüler bir hedef haline getirmiştir. Mobil kullanıcılar, masaüstü kullanıcılara kıyasla daha az dikkatli olabilmekte ve bu nedenle mobil cihazlardaki oltalama saldırıları daha yüksek bir başarı oranına sahip olabilmektedir (Wu, Du ve Wu, 2015). Mobil cihazların ekran boyutlarının sınırlı olması, saldırganların URL'leri gizlemesine ve sahte siteleri maskelemesine olanak tanımakta, bu da kullanıcıların bu tür saldırıları fark etmelerini zorlaştırmaktadır (Canova ve diğerleri, 2015). Ayrıca, mobil tarayıcılar genellikle azalan güvenlik özelliklerine sahiptir, bu da güvenlik açıklarını daha belirgin hale getirmektedir (Varshney ve diğerleri, 2016).

Araştırmacılar, kullanıcıları phishing saldırılarına karşı korumak için çeşitli stratejiler geliştirmiştir. Bu stratejiler genellikle eğitsel ve teknik çözümler olarak ikiye ayrılmaktadır. Eğitsel teknikler, internet kullanıcılarını eğitim programları, atölye çalışmaları ve farkındalık kampanyaları ile phishing e-postalarını ve web sitelerini tanımaları konusunda bilgilendirmeyi amaçlamaktadır (Aleroud ve Zhou, 2017). Teknik çözümler ise genellikle otomatik sınıflandırma ve kara listeleme gibi tekniklerle phishing tespitini hızlandırmayı ve insan müdahalesini en aza indirmeyi hedeflemektedir (Jain ve Gupta, 2022).

Web sitesi görünümü üzerinden saldırının tespiti, birçok uzman için bile zor bir işlem olabilmektedir çünkü saldırganlar, farklı teknikler kullanarak bilgili kullanıcıları dahi aldatabilmektedir. Bu nedenle, oltalama saldırılarının tespit edilmesi için yazılım desteği almak büyük önem taşımaktadır (Buber, Diri ve Şahingöz, 2017). Phishing tehditlerini

azaltmak amacıyla, araştırmacılar URL'lerin, ana bilgisayar bilgilerinin ve web sitesi içeriklerinin el ile oluşturulmuş özelliklerini kullanan çeşitli liste tabanlı ve makine öğrenimi tabanlı yöntemlerin doğruluğunu artırmak için çalışmaktadır (Dou ve arkadaşları, 2017). Bununla birlikte, phishing tespiti, phishing'in sürekli evrimi nedeniyle kesin bir çözüm sunmayan bir silahlanma yarışı olarak kalmaktadır. Bu durum, phishing özelliklerini otomatik olarak çıkarabilen ve phishing web sitelerini doğru bir şekilde tespit edebilen tekniklerin geliştirilmesini sürekli olarak zorunlu kılmaktadır. Yapay zekâ ve makine öğrenimi gibi gelişmiş teknolojiler, saldırıların otomatik olarak tespit edilmesi ve engellenmesi için gittikçe daha fazla önem kazanmaktadır. Bu teknolojiler, phishing saldırılarının karmaşıklığını ve sürekli değişen doğasını anlamak ve etkili savunma mekanizmaları geliştirmek için kritik öneme sahiptir (Brad, 2021).

Mobil uygulama mağazalarında sunulan benzer güvenlik uygulamalarının büyük bir çoğunluğunun statik kara liste yöntemlerine dayandığı gözlemlenmiştir. Bu yöntem, önceden tanımlanmış zararlı URL'leri içeren sabit listeler kullanarak, kullanıcıların ziyaret etmeye çalıştığı web sitelerini bu listelerle karşılaştırmaktadır. Ancak bu listeler, sıklıkla güncellenmedikleri için yeni ortaya çıkan tehditlere karşı pasif kalmaktadır. Örneğin, bir kara liste oluşturulduktan kısa bir süre sonra, siber saldırganlar tarafından yeni alan adları devreye sokulabilmekte veya mevcut zararlı siteler, algılanmamak için adreslerini değiştirebilmektedir. Bu tür değişiklikler, statik listelerin hızla eskimesine ve dolayısıyla yeni tehditleri tespit etme kapasitesinin azalmasına neden olmaktadır. Ayrıca, statik listeler yalnızca bilinen tehditleri barındırdığı için, henüz tanımlanmamış veya daha önce rapor edilmemiş zararlı siteler bu kontrollerden sıyrılabilmektedir. Bu durum, kullanıcıları evrim geçiren siber tehditlere karşı savunmasız bırakmaktadır.

1.1. Proje Çalışmasının Amacı ve Önemi

Bu araştırma önerisinin temel amacı, internet kullanıcılarını artan phishing (oltalama) saldırılarına karşı korumak için mobil platformlarda gerçek zamanlı çalışan, kullanıcı gizliliğini ihlal etmeden sürekli güncellenebilen ve kullanıcının tarama alışkanlıklarına adapte olarak değişen saldırı tekniklerine uyum sağlayabilen, federe öğrenme tabanlı derin öğrenme modeli geliştirmektir. Geliştirilecek model, sayfa kaynak kodlarından öznitelik çıkararak phishing web sitelerini etkin ve gerçek zamanlı bir şekilde tespit edecek şekilde tasarlanacaktır. Böylece, kullanıcıların kişisel bilgilerinin korunması, internet güvenliğinin

artırılması ve phishing saldırıları sonucu oluşabilecek mali ve operasyonel zararların minimize edilmesi sağlanacaktır.

Araştırmanın nihai hedefi, geliştirilen modelin teorik düzeyde kalmayıp, toplumun genel faydasına sunulabilecek, gerçek zamanlı çalışan ve kullanıcı gizliliğini ön planda tutan bir güvenlik aracı olarak sunulmasıdır. Bu amaçla, modelin entegrasyonuyla özel bir anti-phishing özellikli web tarayıcısı geliştirilecektir. Bu tarayıcı, kullanıcıların güvenli ve kesintisiz bir internet deneyimi yaşamalarını sağlayacak, ziyaret edilen sitelerin güvenilirliğini anlık olarak değerlendirecek ve olası tehditleri tespit ettiğinde kullanıcıya uyarılar gönderecektir.

Bu araştırma kapsamında geliştirilecek yapay zekâ modeli, kullanıcı verilerini merkezi bir sunucuya aktarmadan doğrudan cihaz üzerinde çalışarak phishing tespiti yapabilen, gizlilik odaklı ve gerçek zamanlı bir güvenlik çözümü olarak tasarlanmaktadır. Bu model, mobil cihazlarda kullanıcıların tarama alışkanlıklarına adapte olacak şekilde çalışarak sürekli erişilebilir ve kullanımı kolay bir güvenlik çözümü sunacaktır. Kullanıcı verilerini yerel olarak işleyerek gizliliği garanti altına alırken, phishing saldırılarını tespit etmede yüksek doğruluk oranları sağlamayı hedeflemektedir.

Sonuç olarak, bu proje yalnızca teorik bir model sunmakla sınırlı kalmayıp, kullanıcılara doğrudan ulaşabilen bir güvenlik aracı ortaya koymayı hedeflemektedir. Modelin dinamik ve sürekli gelişen yapısı, phishing saldırılarına karşı proaktif bir savunma mekanizması oluşturacaktır. Böylelikle mevcut anti-phishing çözümlerine kıyasla hem etkinlik hem de gizlilik açısından daha üst düzeyde bir koruma sağlayacak olan bu entegre güvenlik sistemi, bireysel kullanıcıların güvenliğini artırmanın ötesinde, internet güvenliği alanında toplumsal fayda sağlayacak ve bu bağlamda yeni nesil bir güvenlik yaklaşımı olarak kendini gösterecektir.

1.2 Araştırma Soruları ve Hipotezler

Bu çalışmada, söz konusu siber tehditlere yönelik teknik çözümler geliştirmek amacıyla aşağıdaki araştırma soruları ve hipotezler belirlenmiştir:

Araştırma Sorusu 1: Mobil platformlar için yapay zekâ tabanlı phishing sayfa kaynak kodu tespit modelinin geliştirilmesi sırasında karşılaşılan zorluklar nelerdir ve bu zorlukların üstesinden gelmek için hangi stratejiler uygulanabilir?

Hipotez 1: Mobil platformlarda sınırlı işlem gücü ve enerji kaynakları gibi donanımsal kısıtlamalar, yapay zekâ tabanlı phishing tespit modelinin geliştirilmesinde temel zorluklardır. Bu zorlukları aşmak için modelin optimize edilmiş ve hafif bir yapıda tasarlanması ile federe öğrenme yöntemlerinin kullanılması etkili bir strateji olacaktır.

Araştırma Sorusu 2: Mobil platformlarda phishing sitelerini gerçek zamanlı olarak analiz eden yapay zekâ destekli bir modelin phishing sitelerini tespit etme başarı oranı nedir ve bu başarıyı etkileyen faktörler nelerdir?

Hipotez 2: Yapay zekâ destekli modelin phishing sitelerini tespit etme başarı oranı %85'in üzerinde olacaktır. Bu başarıyı etkileyen faktörler arasında modelin eğitim verisinin kalitesi, güncelliği ve modelin kullanıcı alışkanlıklarına adapte olma yeteneği yer almaktadır.

Araştırma Sorusu 3: Kullanıcıların kendilerine gösterilen phishing sitelerini tespit etme sürecindeki karar verme performansları ile yapay zekâ destekli modelin analiz sonuçları arasındaki doğruluk ve performans farklılıkları nelerdir?

Hipotez 3: Yapay zekâ destekli modelin phishing sitelerini tespit etme doğruluğu, kullanıcıların bireysel karar verme performansından daha yüksek olacaktır. Kullanıcılar, sosyal mühendislik teknikleri nedeniyle yanıltılabilecek iken, model nesnel analiz yeteneği sayesinde daha yüksek doğruluk oranına ulaşabilecektir.

Araştırma Sorusu 4: Federe öğrenme ile kullanıcının alışkanlıklarına adapte olan modelin performansı nasıl değişecektir?

Hipotez 4: Federe öğrenme yöntemiyle kullanıcının alışkanlıklarına adapte olan modelin performansı, standart modele kıyasla anlamlı bir artış gösterecektir. Bu artış, modelin kullanıcı spesifik verilerle güncellenmesi ve daha hassas tespit yapabilmesi sayesinde elde edilecektir.

1.3 Araştırmanın Özgün Değeri

Phishing tespiti konusunda literatürde pek çok çalışma bulunmasına rağmen, mobil cihazlarda sayfa kaynak kodlarını federe öğrenme yöntemiyle analiz ederek kullanıcı gizliliğini ön planda tutan kapsamlı bir araştırmaya henüz rastlanmamıştır. Mevcut federe öğrenme temelli çalışmaların büyük çoğunluğu, e-posta, SMS veya mesaj tespiti gibi dar kapsamlı alanlara odaklanmakta; ancak bunların önemli bir kısmı teorik düzeyde kalıp

gündelik hayatta uygulanabilir bir çözüm sunamamaktadır. Bu çalışma, söz konusu eksikliği gidermek ve federe öğrenme tabanlı sayfa kaynak kodu analiziyle yenilikçi bir yapay zekâ çözümü geliştirmeyi amaçlamaktadır.

Literatür taraması sırasında, “federated learning in phishing detection”, “phishing website source code analysis” ve “anti-phishing solutions on mobile devices” gibi anahtar kelimeler kullanılarak IEEE Xplore, SpringerLink, ScienceDirect, ACM Digital Library ve Google Scholar gibi önde gelen akademik veri tabanlarında kapsamlı araştırmalar yapılmıştır. Elde edilen sonuçlar, phishing konusuna dair çok sayıda çalışmayı ortaya koysa da, özellikle sayfa kaynak kodlarını analiz etme sürecinde federe öğrenmeyi kullanan bir yaklaşıma rastlanmadığını göstermiştir. Bu durum, araştırmamızın hem teknik yenilik hem de kullanıcı gizliliği odaklı özgün bir katkı sunma potansiyelini açıkça ortaya koymaktadır.

2. MATERYAL VE YÖNTEM

Bu çalışmada, söz konusu probleme yönelik teknik bir çözüm önerisi olarak, derin öğrenme tabanlı bir yapay zekâ modeli geliştirilmesi hedeflenmektedir. Bu yapay zekâ modeli, kullanıcıların ziyaret ettiği web sitelerinin sayfa kaynak kodlarını analiz ederek bu sitelerin güvenilirliğini değerlendirecektir. Modelin, yenilikçi bir özellik olarak, federe öğrenme yöntemini kullanması planlanmaktadır. Federe öğrenme yaklaşımı sayesinde, yapay zekâ modeli cihazda çalışacak ve kullanıcı verilerini merkezi bir sunucuya aktarmadan öğrenme sürecini sürdürecektir. Bu yöntem, kullanıcıların veri gizliliğini korurken, modelin kullanıcı alışkanlıklarına ve değişen phishing saldırı tekniklerine adapte olmasına olanak tanıyacaktır.

Yapay zekâ modeli, başlangıçta merkezi bir sunucuda eğitilecek ve phishing ile yasal siteler arasındaki farkları ayırt edebilecek temel yeteneklere sahip olacaktır. Eğitim sürecinde model, sayfa kaynak kodlarında bulunan HTML, JavaScript ve CSS gibi unsurları inceleyerek phishing sitelerini yasal sitelerden ayırt edebilecek öznitelikleri öğrenecektir. Bu aşamanın amacı, modelin web sitelerinin yapısal ve işlevsel özelliklerinden elde edilen verileri anlamlandırarak doğru tespit yeteneğini geliştirmektir. Bu temel yetenekleri kazandıktan sonra model, kullanıcı cihazlarına dağıtılacak ve yerel ortamda çalışarak gerçek zamanlı phishing tespiti yapabilecektir.

Federe öğrenme yöntemi, modelin yalnızca başlangıç eğitimiyle sınırlı kalmadan, her bir kullanıcı cihazında öğrenim sürecini sürdürebilmesini sağlamaktadır. Model, cihaz üzerinde çalışırken kullanıcının tarama alışkanlıklarını ve karşılaşılan sayfa kaynak kodlarını analiz ederek kendini güncelleyebilecektir. Kullanıcı hangi tür sitelere sıklıkla giriş yapıyorsa, model bu tür sitelerin karakteristik özelliklerini daha iyi öğrenerek phishing sitelerini daha hassas bir şekilde tespit edebilecektir. Cihazlar üzerindeki yerel güncellemeler, belirli periyotlarla şifrelenmiş bir biçimde merkezi sunucuya iletilecek ve burada tüm kullanıcıların katkıları birleştirilerek genel bir model oluşturulacaktır. Merkezi sunucuda bir araya getirilen bu güncellemeler, daha güçlü bir genel modele dönüştürülerek her bir kullanıcı cihazına tekrar gönderilecektir. Bu döngü, yapay zekâ modelinin kullanıcı gizliliğini ihlal etmeden sürekli olarak güncellenebilmesini ve phishing tespitinde güncel tehditlere karşı daha yüksek doğruluk oranlarına ulaşabilmesini sağlamaktadır.

Bu yapay zekâ modeli, kullanıcı deneyimini iyileştirmek üzere geliştirilecek özel bir web

tarayıcısına entegre edilecektir. Anti-phishing özelliğine sahip bu web tarayıcı, kullanıcıların güvenli bir tarama deneyimi yaşamasını sağlayarak ziyaret edilen sitelerin güvenliğini anlık olarak değerlendirecek ve olası tehditleri tespit ettiğinde kullanıcıya uyarılar gönderecektir. Böylelikle, kullanıcı gizliliğine saygılı, sürekli güncellenen ve kişiselleştirilmiş güvenlik özellikleri sunan bir araç oluşturulacaktır. Federe öğrenme tabanlı bu yaklaşım, kullanıcıların verilerinin cihazdan dışarıya çıkmasını gerektirmeden bireysel güvenlik çözümleri sunarak veri gizliliğini en önemli öncül olarak benimsemektedir.

Bu bölümde, araştırmanın amacına ve hedeflerine ulaşmak için uygulanacak yöntem ve teknikler detaylı bir şekilde açıklanacaktır. Araştırma, derin öğrenme ve federe öğrenme yöntemlerini kullanarak mobil platformlarda gerçek zamanlı phishing tespiti yapabilen bir yapay zekâ modeli geliştirmeyi hedeflemektedir. Ayrıca, modelin performansını insanlarla kıyaslamak amacıyla bir kullanıcı çalışması gerçekleştirilecektir.

2.1. Araştırmanın Tasarımı

Araştırma, deneysel ve uygulamalı bir tasarıma sahiptir ve aşağıdaki aşamalardan oluşmaktadır:

1. **Veri Toplama ve Ön İşleme**
2. **Derin Öğrenme Modelinin Geliştirilmesi**
3. **Hiperparametre Optimizasyonu**
4. **Model Eğitimi ve Değerlendirmesi**
5. **Federe Öğrenme Uygulaması**
6. **Anti-Phishing Web Tarayıcısının Geliştirilmesi**
7. **Modelin ve Sistemin Performansının Ölçülmesi ve Optimizasyonu**
8. **Kullanıcı Çalışması ile İnsan Performansının Ölçülmesi**

2.2. Veri Toplama ve Ön İşleme

2.2.1. Veri Kaynakları

Araştırmada kullanılacak veri seti, maksimum gerçeklik ve gerçek dünya senaryolarını

yansıtmak amacıyla farklı kaynaklardan elde edilen phishing ve yasal web sitelerinin sayfa kaynak kodlarından oluşacaktır.

- **Yasal Web Siteleri (Legal):**
    - **Eğitim Verisi:** Alexa Top Sites
    - **Test Verisi:** Webpage Rank ve Google Arama sonuçları
- **Phishing Web Siteleri:**
    - **Eğitim Verisi:** PhishTank, OpenPhish, PhishRepo
    - **Test Verisi:** PhishTank, PhishStats, Kaggle

Ayrıca, van Dooremaal ve arkadaşlarının hazırladığı "Phishing Website Dataset" de kullanılacaktır.

2.2.2. Veri Toplama Süreci

- **Web Crawler Geliştirilmesi:** Python dilinde bir web crawler geliştirilecektir. Bu araç, belirlenen URL'lerin sayfa kaynak kodlarını otomatik olarak indirecektir.
- **Veri Doğrulama ve Temizleme:** İndirilen sayfa kaynak kodları geçerlilik ve bütünlük açısından kontrol edilecek; bozuk, eksik veya yinelenen veriler temizlenecektir.

2.2.3. Veri Ön İşleme

- **Ön İşleme Teknikleri:** Sayfa kaynak kodları üzerinde tokenizasyon, stop-word temizleme ve karakter kodlaması gibi ön işleme adımları uygulanacaktır.
- **Öznitelik Çıkarma:** Derin öğrenme modeline uygun girdi verisi oluşturmak için öznitelik çıkarma işlemleri yapılacaktır. Bu aşamada, n-gram ve Word Embedding yöntemleri kullanılacaktır.

2.3. Derin Öğrenme Modelinin Geliştirilmesi

2.3.1. Model Seçimi

- **Convolutional Neural Network (CNN):** Model olarak sadece Convolutional Neural Network (CNN) kullanılacaktır. CNN'ler, metin verilerinin

sınıflandırılmasında başarılı sonuçlar vermektedir.

- **Modelin Hafifletilmesi:** Mobil cihazlarda çalışabilecek hafif ve optimize edilmiş bir CNN modeli tasarlanacaktır. Model mimarisi ve katman sayısı mobil cihazların işlem gücüne uygun şekilde düzenlenecektir.

2.3.2. Bağımlı ve Bağımsız Değişkenler

- **Bağımsız Değişkenler:** Web sitelerinin sayfa kaynak kodlarından çıkarılan öznitelikler (örneğin, HTML etiketleri, JavaScript fonksiyonları, URL yapısı).
- **Bağımlı Değişken:** Web sitesinin phishing veya yasal olarak sınıflandırılması (ikili sınıflandırma).

2.4. Hiperparametre Optimizasyonu

- **Bayesian Optimizasyon:** Modelin hiperparametre optimizasyonu için Bayesian optimizasyon yöntemi kullanılacaktır. Bu yöntem, arama alanını verimli bir şekilde keşfederek en uygun hiperparametreleri belirlemektedir.
- **Optimizasyon Parametreleri:**
    - Öğrenme oranı
    - Katman sayısı
    - Filtre sayısı
    - Kernel boyutu
    - Dropout oranı
- **Optimizasyon Ortamı:** Hiperparametre optimizasyonu ve model eğitimi işlemleri yüksek performans ve yüksek bellek gerektirdiği için bu işlemler **Microsoft Azure** bulut platformu üzerinde gerçekleştirilecektir.

2.5. Model Eğitimi ve Değerlendirmesi

2.5.1. Model Eğitimi

- **Eğitim Ortamı:** Model eğitimi ve diğer yapay zekâ işlemleri için **TensorFlow** kütüphanesi ve **Jupyter Notebook** kullanılacaktır.

- **Veri Setinin Bölünmesi:** Eğitim ve test verileri, belirtilen kaynaklardan elde edilen veri setlerine göre ayrılacaktır.
- **Eğitim Parametreleri:**
    - Hiperparametre optimizasyonu sonucunda belirlenen en uygun parametreler kullanılacaktır.
    - Optimizasyon algoritması olarak Adam kullanılacaktır.

2.5.2. Model Değerlendirme

- **Performans Metrikleri:**
    - Doğruluk oranı
    - Precision
    - Recall
    - F1 skoru
    - ROC-AUC eğrisi
- **Hedef Başarı Oranı:** Modelin **%85'in üzerinde doğruluk oranına** ulaşması hedeflenmektedir.
- **Yanlış Pozitif Oranı:** Yanlış pozitif oranının **%15'in altında** tutulması amaçlanmaktadır.

2.6. Federe Öğrenme Uygulaması

2.6.1. Federe Öğrenme Mimarisi

- **TensorFlow Federated Learning Kütüphanesi:** Federe öğrenme için **TensorFlow Federated** kütüphanesi kullanılacaktır.
- **Merkezi Sunucu ve İstemci Yapısı:** Merkezi sunucu, modelin global güncellemelerini tutarken, istemci cihazlar yerel verileri kullanarak modeli güncelleyecektir.

2.6.2. Federe Öğrenme Süreci

- **Yerel Model Güncellemeleri:** Her kullanıcı cihazı, kendi tarama verileri

üzerinden yerel olarak modeli güncelleyecektir.

- **Model Ağırlıklarının Birleştirilmesi:** Merkezi sunucu, kullanıcı cihazlarından gelen model güncellemelerini birleştirerek global modeli güncelleyecektir.
- **Gizlilik ve Güvenlik:** Kullanıcı verileri cihaz dışına çıkmadan, sadece model güncellemeleri şifreli bir şekilde sunucuya iletilecektir.

2.7. Anti-Phishing Web Tarayıcısının Geliştirilmesi

2.7.1. Tarayıcı Geliştirme

- **Platform Seçimi:** Tarayıcı uygulaması, **Android** platformu için geliştirilecektir.
- **Yazılım Teknolojileri: Kotlin** programlama dili ve **Jetpack Compose** kullanılacaktır.

2.7.2. Model Entegrasyonu

- **Yerel Model Entegrasyonu:** Geliştirilen CNN tabanlı yapay zekâ modeli, tarayıcı uygulamasına entegre edilecektir.
- **Gerçek Zamanlı Analiz:** Tarayıcı, ziyaret edilen web sitelerinin sayfa kaynak kodlarını anlık olarak analiz edecek ve phishing tespiti yapacaktır.

2.7.3. Kullanıcı Arayüzü ve Deneyimi

- **Kullanıcı Dostu Arayüz:** Basit ve anlaşılır bir arayüz tasarlanacaktır.
- **Uyarı ve Bildirimler:** Phishing tespiti durumunda kullanıcılara net ve yönlendirici uyarılar sunulacaktır.

2.8. Modelin ve Sistemin Performansının Ölçülmesi ve Optimizasyonu

2.8.1. Performans Testleri

- **Hız ve Verimlilik Ölçümleri:** Sayfa kaynak kodlarının analiz süresi ve modelin işlemci kullanımı ölçülecektir. Hedef, analiz süresinin **5 saniyeden kısa** olmasıdır.
- **Cihaz Kaynak Kullanımı:** Modelin cihaz üzerinde maksimum **200 MB** depolama alanı kullanması ve enerji tüketiminin optimize edilmesi sağlanacaktır.

2.8.2. İstatistiksel Analiz

- **Veri Analizi:** Toplanan veriler istatistiksel yöntemlerle analiz edilecektir (örneğin, t-test, ANOVA).
- **Hipotez Testleri:** Araştırma sorularına yönelik hipotezler test edilecektir.

2.8.3. Optimizasyon

- **Model İyileştirmeleri:** Performans verilerine göre modelin hiperparametreleri ve mimarisi optimize edilecektir.
- **Sistem Güncellemeleri:** Kullanıcı geri bildirimleri doğrultusunda tarayıcı ve model güncellenecektir.

2.9. Kullanıcı Çalışması ile İnsan Performansının Ölçülmesi

2.9.1. Çalışmanın Amacı

Modelin performansını insanlarla kıyaslamak amacıyla bir kullanıcı çalışması gerçekleştirilecektir. Bu çalışma ile kullanıcıların phishing ve yasal web sitelerini tespit etme becerileri ölçülerek modelin performansı ile karşılaştırılacaktır.

2.9.2. Katılımcılar

- **Katılımcı Sayısı:** En az **50 gönüllü** katılımcı.
- **Demografik Çeşitlilik:** Farklı yaş, cinsiyet ve eğitim seviyelerinden katılımcılar seçilecektir.

2.9.3. Prosedür

- **Web Sitelerinin Seçilmesi:** Katılımcılara gösterilecek web siteleri, test veri setinden rastgele seçilen eşit sayıda phishing ve yasal web sitesinden oluşacaktır.
- **Görev Tanımı:** Katılımcılardan her bir web sitesini inceleyerek phishing veya yasal olduğuna karar vermeleri istenecektir.
- **Veri Toplama:** Katılımcıların her bir web sitesi için verdikleri kararlar ve karar verme süreleri kaydedilecektir.

2.9.4. Etik Onay ve Bilgilendirilmiş Onam

- **Etik İlkeler:** Çalışma, etik kurallara uygun olarak yürütülecek ve katılımcıların gizliliği korunacaktır.
- **Onam Formu:** Katılımcılardan bilgilendirilmiş onam alınacaktır.

2.9.5. Veri Analizi

- **Performans Metrikleri:**
    - Doğruluk oranı
    - Doğru pozitif ve doğru negatif oranları
    - Karar verme süreleri
- **Karşılaştırma:** Katılımcıların performansı, modelin performansı ile karşılaştırılacaktır.
- **İstatistiksel Testler:** İstatistiksel testler kullanılarak (örneğin, bağımsız örneklem t-testi), performans farklarının anlamlılığı değerlendirilecektir.

2.10. Güvenlik ve Gizlilik Önlemleri

- **Şifreli İletişim Protokolleri:** Model güncellemeleri ve veri transferleri için TLS/SSL protokolleri kullanılacaktır.
- **Yasal Uyumluluk:** Sistem, **GDPR** ve **KVKK** gibi veri koruma yasalarına tam uyumlu olacaktır.
- **Güvenlik Testleri:** Penetrasyon testleri ve güvenlik denetimleri yapılarak olası açıklar kapatılacaktır.

2.11. Ön Çalışmalar ve Fizibilite

Araştırma ekibi, daha önce phishing tespiti ve derin öğrenme üzerine çalışmalar yapmış olup, bu alanda deneyime sahiptir. Ön fizibilite çalışmaları kapsamında, küçük ölçekli bir veri seti üzerinde CNN modeli denenmiş ve olumlu sonuçlar elde edilmiştir. Ayrıca, TensorFlow ve TensorFlow Federated kütüphaneleri kullanılarak temel federe öğrenme uygulamaları gerçekleştirilmiştir.

EKLER

EK-1: İş Zaman Çizelgesi

EK-2: Risk Yönetimi Tablosu

İŞ-ZAMAN ÇİZELGESİ

İP No	İş Paketlerinin Adı ve Hedefleri	Gerçekleştiren	Zaman Aralığı	Başarı Ölçütü ve Projenin Başarısına Katkısı	Tamamlanma Oranı
1	**Veri Toplama ve Ön İşleme** - Farklı kaynaklardan phishing ve yasal web sitelerinin sayfa kaynak kodlarının toplanması - Verilerin ön işlenmesi ve öznitelik çıkarımı	Araştırma Yürütücüsü (Eren DOĞAN)	Ay 1-2	- En az **30.000** phishing ve **30.000** yasal web sitesi verisinin toplanması - Model eğitimi için hazır veri setinin oluşturulması - **Projenin temelini oluşturacak veri setinin hazırlanması (Projenin başarısına katkısı: %20)**	%100
2	**CNN Modelinin Geliştirilmesi** - Mobil cihazlarda çalışabilecek hafif ve optimize edilmiş bir CNN modelinin tasarlanması	Araştırma Yürütücüsü (Eren DOĞAN)	Ay 2-3	- Model mimarisinin belirlenmesi ve kodlanması - Modelin mobil uyumluluğunun sağlanması - **Phishing tespiti için temel yapay zekâ modelinin oluşturulması (Projenin başarısına katkısı: %15)**	%100
3	**Hiperparametre Optimizasyonu** - Bayesian optimizasyon yöntemi ile modelin hiperparametrelerinin en iyi değerlerinin bulunması	Araştırma Yürütücüsü (Eren DOĞAN)	Ay 3-4	- En uygun hiperparametrelerin belirlenmesi - Modelin doğruluk oranının **%85'in üzerine** çıkması - **Model performansının artırılması ve en iyi sonuçların elde edilmesi (Projenin başarısına katkısı: %10)**	%100
4	**Model Eğitimi ve Değerlendirmesi** - Modelin belirlenen veri seti üzerinde eğitilmesi ve test edilmesi	Araştırma Yürütücüsü (Eren DOĞAN)	Ay 4-5	- Modelin eğitim sürecinin tamamlanması - Test verisi üzerinde **%85'in üzerinde doğruluk** elde edilmesi - Yanlış pozitif oranının **%15'in altında** olması - **Yüksek performanslı bir phishing**	%100

				tespit modelinin elde edilmesi (Projenin başarısına katkısı: %15)	
5	**Federe Öğrenme Altyapısının Kurulması** - TensorFlow Federated kullanarak federe öğrenme mimarisinin oluşturulması - Modelin cihazlarda yerel olarak güncellenebilmesi	Araştırma Yürütücüsü (Eren DOĞAN)	Ay 5-6	- Federe öğrenme altyapısının başarılı bir şekilde kurulması - Kullanıcı verilerinin **%100 gizlilikle** cihaz dışına çıkmadan modelin güncellenebilmesi - **Kullanıcı gizliliğinin korunması ve modelin adaptif hale getirilmesi (Projenin başarısına katkısı: %10)**	%100
6	**Anti-Phishing Web Tarayıcısının Geliştirilmesi** - Android platformunda Kotlin ve Jetpack Compose kullanarak web tarayıcısının geliştirilmesi - Modelin tarayıcıya entegre edilmesi	Araştırma Yürütücüsü (Eren DOĞAN)	Ay 4-6	- Tarayıcının ilk versiyonunun geliştirilmesi - Modelin tarayıcıya başarılı bir şekilde entegre edilmesi - **Kullanıcıların güvenli ve kesintisiz bir internet deneyimi yaşamalarını sağlayacak aracın oluşturulması (Projenin başarısına katkısı: %10)**	%0
7	**Performans Testleri ve Optimizasyon** - Modelin ve sistemin hız ve verimlilik açısından test edilmesi - Gerekli optimizasyonların yapılması	Araştırma Yürütücüsü (Eren DOĞAN)	Ay 6-7	- Analiz süresinin **%90 olasılıkla 5 saniyeden kısa** olması - Modelin cihaz kaynaklarını etkin kullanması (maksimum **200 MB** depolama alanı) - **Sistemin kullanıcılar için verimli ve hızlı çalışmasının sağlanması (Projenin başarısına katkısı: %5)**	%0
8	**Güvenlik ve Gizlilik Önlemleri** - Şifreli iletişim protokollerinin uygulanması - GDPR ve KVKK uyumluluğunun sağlanması - Güvenlik testlerinin yapılması	Araştırma Yürütücüsü (Eren DOĞAN)	Ay 6-8	- Şifreli protokollerin başarılı bir şekilde uygulanması - Yasal düzenlemelere **%100 uyumun** sağlanması - Güvenlik açıklarının tespit edilip kapatılması - **Kullanıcı verilerinin güvenliğinin ve gizliliğinin sağlanması (Projenin başarısına katkısı: %5)**	%0

9	**Kullanıcı Çalışması ve İnsan Performansının Ölçülmesi** - Katılımcılara phishing ve yasal web sitelerinin gösterilmesi - İnsan performansının ölçülmesi	Araştırma Yürütücüsü (Eren DOĞAN)	Ay 7-8	- En az **50** katılımcı ile çalışmanın gerçekleştirilmesi - Katılımcıların kararlarının ve sürelerinin kaydedilmesi - Modelin insan performansına göre **en az %10 daha yüksek doğruluk** göstermesi - **Modelin performansının insan performansı ile karşılaştırılarak üstünlüğünün gösterilmesi (Projenin başarısına katkısı: %10)**	%0
10	**Pilot Uygulama ve Kullanıcı Testleri** **-** Geliştirilen sistemin en az 100 kullanıcı ile test edilmesi - Geri bildirimlerin toplanması	Araştırma Yürütücüsü (Eren DOĞAN)	Ay 7-9	- Pilot uygulamanın başarılı bir şekilde tamamlanması - Kullanıcılardan **%80'in üzerinde memnuniyet skoru** alınması - **Sistemin son kullanıcıya hazır hale getirilmesi ve kullanıcı deneyiminin iyileştirilmesi (Projenin başarısına katkısı: %5)**	%0

RİSK YÖNETİMİ TABLOSU

İP No	En Önemli Riskler	Risk Yönetimi (B Planı)
1	**Veri Toplama ve Ön İşleme** - Farklı kaynaklardan yeterli miktarda ve kalitede veri toplanamaması - Veri setlerinin güncelliğinin sağlanamaması	- **Alternatif Veri Kaynakları:** Veri toplama sürecinde zorluk yaşanması durumunda, ek veri kaynakları araştırılacak ve kullanılacaktır (örneğin, farklı phishing veri tabanları, web arşivleri). - **Sentetik Veri Oluşturma:** Gerekirse, gerçek veriye benzer sentetik veri oluşturularak veri seti genişletilecektir. - **Veri Temizleme ve Doğrulama Süreçlerinin İyileştirilmesi:** Veri kalitesini artırmak için ek kontroller ve otomasyon sağlanacaktır.
2	**CNN Modelinin Geliştirilmesi** - Modelin mobil cihazlarda çalışacak kadar hafif olmaması - Modelin istenen doğruluk seviyesine ulaşamaması	- **Model Mimarisi Optimizasyonu:** Modelin hafifletilmesi için farklı mimari yaklaşımlar denenecektir (örneğin, MobileNet, SqueezeNet). - **Alternatif Modellerin Kullanılması:** CNN dışında hafif ve etkili diğer modeller (örneğin, LightGBM, XGBoost) değerlendirilecektir.
3	**Hiperparametre Optimizasyonu** **- Hiperparametre optimizasyonunun beklenenden uzun sürmesi veya uygun hiperparametrelerin bulunamaması**	- **Paralel İşleme ve Kaynak Artırımı:** Microsoft Azure üzerinde daha fazla kaynak kullanarak optimizasyon süreci hızlandırılacaktır. - **Grid Search veya Random Search Kullanımı:** Bayesian optimizasyon sonuç vermezse, alternatif optimizasyon yöntemleri kullanılacaktır.
4	**Model Eğitimi ve Değerlendirmesi** - Modelin eğitim sürecinde aşırı uyum (overfitting) veya yetersiz öğrenme (underfitting) sorunlarının ortaya çıkması	- **Düzenlileştirme Teknikleri:** Dropout, L1/L2 regularizasyonu gibi yöntemlerle overfitting engellenecektir. - **Veri Artırma (Data Augmentation):** Eğitim verisi çeşitlendirilecektir. - **Model Kompleksitesinin Ayarlanması:** Modelin katman sayısı ve büyüklüğü yeniden düzenlenecektir.
5	**Federe Öğrenme Altyapısının Kurulması** - Federe öğrenme sürecinde iletişim gecikmeleri veya model güncellemelerinin senkronize edilememesi	- **Asenkron Federe Öğrenme:** Model güncellemelerinin asenkron şekilde işlenebileceği bir mimari uygulanacaktır. - **İletişim Protokollerinin Optimizasyonu:** Veri transferi için daha verimli protokoller kullanılacaktır. - **Yerel Model Eğitiminin Güçlendirilmesi:** Merkezi sunucuya bağımlılığı azaltmak için cihazlarda daha fazla işlem yapılması sağlanacaktır.
6	**Anti-Phishing Web Tarayıcısının Geliştirilmesi** - Tarayıcının geliştirme sürecinde teknik zorluklar veya gecikmeler	- **Mevcut Tarayıcı Tabanlarının Kullanılması:** Sıfırdan geliştirmek yerine mevcut açık kaynak tarayıcı projeleri üzerine inşa edilecektir.

	yaşanması - Tarayıcının kullanıcı dostu olmaması veya performans sorunları	- **Kullanıcı Arayüzü Tasarımında Uzman Desteği:** UI/UX konusunda uzmanlardan destek alınacaktır. - **Performans Optimizasyonu:** Kod ve kaynak yönetimi iyileştirilerek performans sorunları giderilecektir.
7	**Performans Testleri ve Optimizasyon** - Modelin hedeflenen hız ve verimlilik seviyesine ulaşamaması - Cihaz kaynaklarının beklenenden fazla kullanılması	- **Modelin Daha Fazla Optimizasyonu:** Modelin boyutu ve işlem yükü daha da azaltılacaktır. - **Hafif Modellerin Entegrasyonu:** Gerekirse daha hafif modeller kullanılacaktır. - **Fonksiyonelliklerin Gözden Geçirilmesi:** Gereksiz veya kaynak tüketen özellikler kaldırılacaktır.
8	**Güvenlik ve Gizlilik Önlemleri** - Şifreli iletişim protokollerinde güvenlik açıklarının tespit edilmesi - Yasal uyumluluk süreçlerinde beklenmedik engellerin ortaya çıkması	- **Güvenlik Denetimlerinin Artırılması:** Uzmanlardan güvenlik testleri için destek alınacaktır. - **Yedek Protokollerin Kullanılması:** Alternatif ve daha güvenli iletişim protokolleri uygulanacaktır. - **Yasal Danışmanlık Alınması:** Veri koruma yasalarına tam uyum için hukuk danışmanlarından destek alınacaktır.

KAYNAKLAR

Aleroud, A., & Zhou, L. (2017). Phishing environments, techniques, and countermeasures: A survey. *Computers & Security*, 68, 160-196. https://doi.org/10.1016/j.cose.2017.04.006 Erişim Tarihi: Mart 14, 2025.

APWG. (2016). Phishing activity trends report, 4th quarter 2016. https://docs.apwg.org/reports/apwg_trends_report_q4_2016.pdf Erişim Tarihi: Mart 14, 2025.

APWG. (2023). Phishing activity trends report, 1th quarter 2016. https://docs.apwg.org/reports/apwg_trends_report_q1_2023.pdf Erişim Tarihi: Mart 14, 2025.

Bıçakcı, S., Ergun, F. D., & Çelikpala, M. (2016). Türkiye'de Siber Güvenlik. In *Türkiye'de Siber Güvenlik ve Nükleer Enerji* (pp. 28-74). İstanbul: Edam. https://www.researchgate.net/publication/289326451_Turkiye'de_Siber_Guvenlik Erişim Tarihi: Mart 14, 2025.

Birleşik Krallık Kamu Sektörü Bilgileri. (2021). Siber Güvenlik İhlalleri Anketi. https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2021 Erişim Tarihi: Mart 14, 2025.

Birleşik Krallık Kamu Sektörü Bilgileri. (2022). Siber Güvenlik İhlalleri Anketi. https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2022 Erişim Tarihi: Mart 14, 2025.

Birleşik Krallık Kamu Sektörü Bilgileri. (2023). Siber Güvenlik İhlalleri Anketi. https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2023 Erişim Tarihi: Mart 14, 2025.

Brad. (2021). AI-Based Cybersecurity Systems' Benefits. *International Journal of Smart Sensors and Ad Hoc Networks (IJSSAN)*, 3(3), 67-69. ISSN No. 2248-9738. Erişim Tarihi: Mart 14, 2025.

Buber, E., Diri, B., & Sahingoz, O. K. (2017). NLP teknikleri kullanarak URL'den phishing saldırılarını tespit etme. In *2017 Uluslararası Bilgisayar Bilimleri ve Mühendisliği Konferansı (UBMK)* (pp. 337-342). IEEE. https://ieeexplore.ieee.org/abstract/document/8093406 Erişim Tarihi: Mart 14, 2025.

Canova, G., Volkamer, M., Bergmann, C., Borza, R., Berens, B., Stockhardt, S., & Tenberg, R. (2015). Learn to Spot Phishing URLs with the Android NoPhish App. *IFIP Advances in Information and Communication Technology*, 453, 87-100. https://doi.org/10.1007/978-3-319-18500-2_8 Erişim Tarihi: Mart 14, 2025.

Dou, Z., Khalil, I., Khreishah, A., Al-Fuqaha, A., & Guizani, M. (2017). Systematization of knowledge (sok): A systematic review of software-based web phishing detection. IEEE Communications Surveys & Tutorials, 19(4), 2797–2819. Erişim Tarihi: Mart 14, 2025.

Goel, S., Williams, K., & Dincelli, E. (2017). Got Phished? Internet Security and Human Vulnerability. *J. Assoc. Inf. Syst.*, 18, 2. https://doi.org/10.17705/1jais.00447. Erişim Tarihi: Mart 14, 2025.

Hekim, H. (2015). Oltalama (Phishing) Saldırıları. In *Siber Suçlar, Tehditler, Farkındalık ve Mücadele* (pp. 57-83). Ankara: Global Politika ve Strateji. http://www.academia.edu/35136881/Oltalama_Phishing_Saldirilari. Erişim Tarihi: Mart 14, 2025.

Jain, A. K., & Gupta, B. B. (2022). A survey of phishing attack techniques, defence mechanisms and open research challenges. *Enterprise Information Systems*, 16(4), 527-565. https://doi.org/10.1080/17517575.2021.1896786. Erişim Tarihi: Mart 14, 2025.

Jakobsson, M., & Myers, S. (2007). Delayed password disclosure. *ACM SIGACT News*, 38(3), 56-75. https://dl.acm.org/doi/abs/10.1145/1324215.1324228. Erişim Tarihi: Mart 14, 2025.

Kamiya, S., Kang, J., Kim, J., Milidonis, A., & Stulz, R. (2019). Risk Management, Firm Reputation, and the Impact of Successful Cyberattacks on Target Firms. *Information Systems & Economics eJournal*. https://doi.org/10.2139/ssrn.3135514. Erişim Tarihi:

Mart 14, 2025.

Ömer, A., Serkant, S., Aktuğ, M., Ozkan-Okay, M., Yilmaz, A. A., & Akin, E. (2023). A Comprehensive Review of Cyber Security Vulnerabilities, Threats, Attacks, and Solutions. *Electronics, 12*(6), 1333-1333. https://doi.org/10.3390/electronics12061333. Erişim Tarihi: Mart 14, 2025.

Sharif, M., & Mohammed, M. (2022). A literature review of financial losses statistics for cyber security and future trend. *World Journal of Advanced Research and Reviews*. https://doi.org/10.30574/wjarr.2022.15.1.0573. Erişim Tarihi: Mart 14, 2025.

Türkiye İstatistik Kurumu. (2023). Hanehalkı Bilgi Teknolojileri (BT) Kullanım Araştırması. https://data.tuik.gov.tr/Bulten/Index?p=Hanehalki-Bilisim-Teknolojileri-(BT)-Kullanim-Arastirmasi-2023-49407 Erişim Tarihi: Mart 14, 2025.

Ünver, M., & Mirzaoğlu, A. G. (2011). Yemleme ("Phishing"). Bilgi Teknolojileri ve İletişim Kurumu. https://www.academia.edu/24841831/Yemleme_Phishing Erişim Tarihi: Mart 14, 2025.

Varshney, G., Misra, M., & Atrey, P. K. (2016). A survey and classification of web phishing detection schemes. *Security and Communication Networks*, 9, 6266–6284. https://doi.org/10.1002/sec.1674 Erişim Tarihi: Mart 14, 2025.

Wassermann, S., Meyer, M., Goutal, S., & Riquet, D. (2023). Targeted Attacks: Redefining Spear Phishing and Business Email Compromise. *ArXiv*, abs/2309.14166. https://doi.org/10.48550/arXiv.2309.14166. Erişim Tarihi: Mart 14, 2025.

Wu, L., Du, X., & Wu, J. (2015). Effective Defense Schemes for Phishing Attacks on Mobile Computing Platforms. *IEEE Transactions on Vehicular Technology,* 65, 1-1. https://doi.org/10.1109/TVT.2015.2472993 Erişim Tarihi: Mart 14, 2025.