

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

GELİŞMİŞ SOSYAL MÜHENDİSLİK SALDIRILARI
ANALİZİ VE TESPİTİ

ŞEYDANUR AHI
YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
SİBER GÜVENLİK PROGRAMI

GEBZE
2020

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

GELİŞMİŞ SOSYAL MÜHENDİSLİK
SALDIRILARI ANALİZİ VE TESPİTİ

ŞEYDANUR AHI
YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
SİBER GÜVENLİK PROGRAMI

DANIŞMANI
PROF. DR. İBRAHİM SOĞUKPINAR

GEBZE

2020

T.R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

**ANALYSIS AND DETECTION OF
ADVANCED SOCIAL ENGINEERING
ATTACKS**

ŞEYDANUR AHI
**A THESIS SUBMITTED FOR THE DEGREE OF
MASTER OF SCIENCE**
DEPARTMENT OF COMPUTER ENGINEERING
CYBER SECURITY PROGRAM

THESIS SUPERVISOR
PROF. DR. İBRAHİM SOĞUKPINAR

GEBZE
2020



YÜKSEK LİSANS JÜRİ ONAY FORMU

GTÜ Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 10/07/2020 tarih ve 2020/33 sayılı kararıyla oluşturulan jüri tarafından 13/07/2020 tarihinde tez savunma sınavı yapılan Şeydanur AHİ'nin tez çalışması Bilgisayar Mühendisliği Anabilim Dalı Siber Güvenlik Programında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

JÜRİ

ÜYE

(TEZ DANIŞMANI) : Prof. Dr. İbrahim Soğukpınar

ÜYE

: Doç. Dr. Mehmet Göktürk

ÜYE

: Doç. Dr. Ali Gökhan Yavuz

ONAY

Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun
...../...../..... tarih ve/..... sayılı kararı.

İMZA/MÜHÜR

Prof. Dr. Ümit DEMİR

Gebze Teknik Üniversitesi

Fen Bilimleri Enstitüsü Müdürü

ÖZET

Sosyal mühendislik teknolojiyi kullanarak ya da teknolojiyi kullanmadan insanlardan bilgi edinme (aldatma) sanatıdır. Günümüzde karşı karşıya olduğumuz saldırıların çok büyük bir kısmı insan kaynaklıdır ve aynı şekilde sistemleri değil, onları kullanan insanları hedef almaktadır. Güvenlik zincirindeki en zayıf halka olan insan, farklı zamanlarda farklı davranışlar sergilemesinden dolayı güvenlik sürecinde çeşitli zafiyetler gösterebilmektedir. Kimlik avı teknik olarak tüketicilerin finansal veya kişisel bilgilerini ele geçirmek için oluşturulmuş bir tür sosyal mühendislik saldırısıdır. Kimlik avı, bugün e-ticaret dünyasının karşılaştığı en büyük zorluklardan biridir. Kimlik avı saldırıları yüzünden birçok şirket ve birey milyarlarca dolar kaybetmektedir. Kimlik avı saldırılarının bu küresel etkisi artmaya devam edecektir ve bu nedenle tehditleri azaltmak için daha etkili kimlik avı algılama tekniklerinin geliştirilmesi gerekmektedir. Bu çalışmada, Sosyal Mühendislik teknik ve taktiklerinin çeşitli biçimlerini ve sosyal mühendislerin insan güvenlik açıklarından nasıl yararlandıklarını, bu tür saldırılara karşı alınacak önlemlerle ilgili bazı yolları tartışmakta ve bu tür saldırıları önlemek için bilinçli olmanın önemi vurgulanmış ayrıca kimlik avı e-posta saldırılarına karşı derin öğrenme modelleri kullanılarak oluşturulan bir tespit yöntemi önerilmiştir. Önerilen yöntemde gelen e-posta iletilerinin başlık ve gövde bölümlerinden elde edilen özellikler kullanılarak çeşitli derin öğrenme modelleri eğitilmiştir. Yapılan testler sonucunda kimlik avı saldırılarına karşı önerilen bu tespit yöntemi %96,84'lük bir başarı oranı elde edilmiştir.

Anahtar sözcükler: Sosyal Mühendislik Saldırısı, Oltalama E-posta Tespiti, Derin öğrenme, Çok Katmanlı Algılayıcı, LSTM, Word2Vec.

SUMMARY

Social engineering is the art of getting information (deception) from people with using technology or without using technology. The vast majority of the attacks facing today are human origin, and likewise, these attacks target computer users. Human being who is the weakest link in the security chain shows various weaknesses in the security process, due to human being's variable behavior in different times. Phishing that is a kind of social engineering attack is technically created to capture consumers' financial or personal information. Phishing is the one of the biggest challenges for the e commerce world. Many companies and individuals lose billions of dollars because of phishing attacks. This global impact of phishing attacks will continue to increase, therefore, more effective phishing detection techniques need to be developed to reduce threats. In this regard, various forms of Social Engineering techniques and tactics and how social engineers benefited from human security deficits, to discuss information about the measures to be taken against such attacks, and to identify such attacks, a detection method was proposed. Various deep learning models were trained using the features obtained from the head and body parts of incoming e-mails in the proposed method. As a result of the tests, a 96.84% success rate was achieved with this detection method proposed against phishing attacks.

Keywords: Social Engineering Attack, Phishing E-mail Detection, Deep learning, Multi Layer Perceptron, LSTM, Word2Vec.

TEŞEKKÜR

Başta, yüksek lisans eğitimimde ve akademik hayatımda desteğini ve yardımlarını hiçbir zaman esirgemeyip bilgisi ile bu çalışmanın oluşmasının yolunu açan danışmanım Prof. Dr. İbrahim Soğukpınar'a

Bütün çalışmam boyunca yanımda olan, bilgi ve tecrübelerini benimle paylaşan sayın Atahan DUMAN'a,

ve göstermiş olduğu desteklerinden dolayı sevgili aileme en içten teşekkürlerimi sunarım.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	v
SUMMARY	vi
TEŞEKKÜR	vii
İÇİNDEKİLER	viii
SİMGELER VE KISALTMALAR DİZİNİ	x
ŞEKİLLER DİZİNİ	xii
TABLolar DİZİNİ	xiii
1. GİRİŞ	1
2. KURAMSAL TEMELLER VE İLGİLİ ÇALIŞMALAR	4
2.1. Sosyal Mühendislik Nedir ?	4
2.2. Sosyal Mühendislik Saldırılarının Döngüsü	7
2.3. Sosyal Mühendisin Nitelikleri	8
2.4. Sosyal Mühendislik Yaklaşımları	9
2.4.1. Fiziksel Yaklaşımlar	10
2.4.2. Sosyal Yaklaşımlar	10
2.4.3. Tersine Sosyal Mühendislik	10
2.4.4. Teknik Yaklaşımlar	11
2.4.5. Sosyo-Teknik Yaklaşımlar	11
2.5. Bilgisayar destekli iş birliği	12
2.5.1. Ofis iletişimi	12
2.5.2. Dış İletişim	13
2.6. Sosyal Mühendislik Saldırısında Kullanılan Yöntemler	13
2.6.1. Tersine Sosyal Mühendislik (RSE)	14
2.6.2. Sirtına Binme (Piggyback Rides)	16
2.6.3. Teknik Konuşma	16
2.6.4. Balina Avı(Whaling)	17
2.6.5. Çöp Karıştırma	17
2.6.6. Telefon ile Yapılan Kimlik Avı (Vishing)	18
2.6.7. Sosyal Medya Ağı	19

2.6.8. Nöro-dilbilimsel Programlama (NLP)	20
2.6.9. Karşıt Cinsiyet Çekiciliği	21
2.6.10. Dost Toplantıları ve Konuşmalar	21
2.7. Bazı Popüler Sosyal Mühendislik Teknikleri	21
2.7.1. Yalan	22
2.7.2. Kısmi Gerçeği söylemek	22
2.7.3. Sebep Vermek	23
2.7.4. Kaçınma ve Derivasyon	24
2.7.5. Karşılıklı İlişki	24
2.7.6. Mizah Kullanımı	24
2.7.7. Evet-Evet Tekniği	24
2.7.8. Top Atma Tekniği	25
2.8. İnsan Zayıflıklarının Önemli Hedefleri	25
2.8.1. Korku	25
2.8.2. Yardım İsteği	26
2.8.3. Dikkatsizlik	26
2.8.4. Konfor Bölgesi	27
2.9. Gelişmiş Sosyal Mühendislik Saldırıları	28
2.9.1. Kimlik Avı (Phishing)	29
2.9.2. Kimlik Avı E-postası	35
2.9.3. Kimlik Avı Önlemleri	40
2.10. Makine Öğrenmesi	42
2.10.1. Makine Öğrenmesi Algoritmalarının Türleri	43
2.10.2. Tanımlar	44
2.10.3. Makine Öğrenmesi Modelleri	49
2.11. İlgili Çalışmalar	58
2.11.1. Kimlik Avı E-postasının Rastgele Orman Makine Öğrenme Tekniği Kullanılarak Sınıflandırılması	58
2.11.2. PySpark kullanarak e-posta kimlik avının analizi ve tespiti	59
2.11.3. Kimlik Avı Tespiti için Makine Öğrenmesi Tekniklerinin Karşılaştırılması	59
2.11.4. Çoğunluk Oylamasını Kullanarak Spam E-postalarını Tespit Etme Yaklaşımı	60

2.11.5.Kimlik Avı E-postalarının Tespiti ve Sınıflandırılması için Makine Öğrenmesi Tekniklerinin Kullanılması	61
2.11.6.Makine Öğrenmesine Dayalı Kimlik Avı Web Sitesi Algılamasında En Etkili Özellik Kategorisini Belirleme	62
2.11.7.Metin Sınıflandırması için Unsur Çıkarma veya Unsur Seçimi: Kimlik Avı E-posta Tespiti Üzerine Bir Vaka Çalışması	64
2.11.8.Kimlik Avı E-posta Filtreleme Teknikleri Araştırması	65
2.12. E-posta Yapısı	67
2.12.1. E-posta İstemcileri(posta kullanıcı aracı)	68
2.12.2. E-posta Başlıkları	69
2.12.3. E-posta Dosyası Biçimleri	70
3. ÖNERİLEN YÖNTEM	72
3.1. Analiz	74
3.1.1. Başlık Analizi	74
3.1.2. Gövde Analizi	78
4. DENEY SONUÇLARI VE DEĞERLENDİRME	85
4.1. E-posta Veri Kümesi	85
4.2. Değerlendirme ve Deney Sonuçları	86
5. SONUÇLAR VE GELECEK ÇALIŞMALAR	91
KAYNAKLAR	92
ÖZGEÇMİŞ	99

SİMGELER ve KISALTMALAR DİZİNİ

<u>Simgeler ve</u>	<u>Açıklamalar</u>
<u>Kısaltmalar</u>	
APWG	: Kimlik Avı Koruması Çalışma Grubu
CSS	: Basamaklanmış Stil Katmanları
FC	: Tamamen bağlı
FP	: Yanlış Pozitif
GTÜ	: Gebze Teknik Üniversitesi
HTML	: Hiper metin işaret dili
LSTM	: Uzun Kısa Süreli Bellek
LR	: Lojistik regresyon
MDL	: Minimum açıklama uzunluğu
MLP	: Çok katmanlı algılayıcı
RNN	: Tekrarlayan Sinir Ağı
RSE	: Tersine sosyal mühendislik
SMS	: Kısa Mesaj Servisi
SVM	: Destek Vektör Makinesi
TF-IDF	: Terim frekansı-ters metin frekansı
TP	: Gerçek Pozitif
TN	: Gerçek Negatif
URL	: Düzgün Kaynak Bulucu
YSA	: Yapay Sinir Ağı

ŞEKİLLER DİZİNİ

<u>Şekil No:</u>	<u>Sayfa</u>
2.1: Sosyal mühendislik saldırı yolu.	5
2.2: Sosyal mühendislik ataklarının izlediği döngü.	8
2.3: Tersine sosyal mühendislik adımları.	15
2.4: Kredi kartı ortalama.	30
2.5: Kredi kartı ortalama yorumları.	31
2.6: 2019 için APWG kimlik avı etkinliği eğilimi.	32
2.7: Kimlik avı saldırılarına göre en çok hedeflenen sektörler.	33
2.8: Kimlik avı e-posta örneği.	35
2.9: HTTPS’de barındırılan kimlik avı saldırıları.	37
2.10: Çapraz doğrulama.	47
2.11: Eğitim,validasyon ve test setlerini kullanarak eğitim süreci.	48
2.12: Yapay sinir ağı modeli.	52
2.13: Tekrarlayan sinir ağı.	53
3.1: Bu çalışmada geliştirilen karma modelin mimarisi	73
3.2: Ölçeklendirilmiş veri kümesi.	78
3.3: Vektör haline getirilmiş kelime dizileri.	80
3.4: Derin öğrenme modelinin görselleştirilmesi.	82
3.5: LSTM mimarisi.	83
4.1: Karışıklık matrisi.	88
4.2: ROC-AUC eğrisi	89

TABLÖLAR DİZİNİ

<u>Tablo No:</u>	<u>Sayfa</u>
2.1: E-posta örneđi.	68
3.1: Başlık kısmından çıkarılan özellik kümesi.	74
3.2: Başlık analizi sonrası oluşan özellik matrisi gösterimi.	77
3.3: Derin öğrenme modeli özeti.	84
4.1: Veri kümesinin özeti.	85
4.2: Sınıflandırma karmaşıklık matrisi.	86
4.3: Sınıflandırma raporu.	88
4.4: Makine öğrenmesi algoritmaları ile sınıflandırma yapan diğer araştırmalar ile karşılaştırma.	90

1. GİRİŞ

Bilgi güvenliği alanı hızla büyüyen bir disiplindir. Hassas bilgileri korumak için güvenlik önlemlerinin etkinliği artıyor olsa da insanlar manipülasyona açık olduğu için insan unsuru zincirin en zayıf halkası olmaya devam etmektedir. Bilgilerin korunması kuruluşlar ve hükümetler için hayati öneme sahiptir ve bilgiye yasadışı erişime karşı önlemlerin geliştirilmesi, gittikçe daha fazla ilgi gören bir alandır. Teknoloji tek başına bilgi hırsızlığına karşı yeterli bir koruma değildir; personel genellikle bir bilgi güvenliği sistemindeki zayıf halkadır. Personel üyeleri, daha sonra yetkisiz kişilerin korunan sistemlere erişmesine izin veren hassas bilgileri açığa çıkarabilir. İnsanları hassas bilgileri ifşa etmenin “sanatı” sosyal mühendislik olarak bilinir ve bunu yapma süreci sosyal mühendislik saldırısı olarak bilinir. Sosyal mühendisliğin çeşitli tanımları ve aynı zamanda bir dizi farklı sosyal mühendislik saldırısı modeli vardır Bir sosyal mühendislik saldırısı, hassas bilgileri ortaya çıkarmak için çeşitli manipülasyon tekniklerini kullanarak bu zayıflığı hedefler.

Dünya üzerindeki internet aktivitesinin ve aktif kullanıcı sayısının artması, insanların birbirleriyle kolay bir şekilde bağlantı kurmalarına, birçok veriye erişebilmelerine, hizmet ve ürünlerini bu şekilde temin etmelerine neden olmuştur. Büyük bir para trafiğinin bu sanal topluluklarda dönmesi dolandırıcıların ve hırsızların da iştahını kabartmaktadır. Eski zamanlarda daha çok fiziksel aktivitelere dayanan hırsızlık ve dolandırıcılık yöntemlerinin sanal ortama aktarılması ile hırsızların işlerinin kolaylaşması ve kısa süre zarfında çok fazla olay gerçekleştirebilme imkanlarının olması bu alanı daha da güçlendirilmeye muhtaç bırakmaktadır. İnsanlara ait bu değerli bilgileri çalmanın yollarından bir tanesi de kimlik avı saldırılarıdır. Kimlik avı, bugün yapılan farklı (ve kazançlı) dolandırıcılık türlerinden biridir. Ceza hukukunda sahtekârlık, yalnızca kişisel kazanımların veya bireyin imajının sarsılmasına yönelik kasıtlı bir aldatma olarak tanımlanır. Genel anlamda sahtekârlık, finansal veya kişisel kazanımlar amacıyla insanları kendi kişisel bilgilerini ifşa etmelerini sağlamak ve onları kandırmak için bir dizi eylem olarak tanımlanabilir. Kimlik avı, yasal bir kuruluşun kopya web sitesini oluşturarak kullanıcılardan (genellikle hırsızlık amacıyla) hassas veya gizli bilgilerini elektronik olarak elde etmeye çalışan bir eylemdir. Kimlik avı genellikle bir elektronik aygıt (tablet ve bilgisayar gibi) ve bir bilgisayar ağı yardımıyla yapılır; son kullanıcıların (güvenlik zincirindeki en zayıf

unsur olduđu düşünölen) çeřitli tespit sistemlerinde mevcut olan zayıflıkları hedef alınır [1, 2]. Kimlik avı saldırıanları, hedeflenen kullanıcının hesabına yetkisiz erişim elde etmek amacıyla sistemde kullandığı kişisel bilgilerini açığa çıkarmaya ikna etmek için kullanıcılara yönelik titiz bir şekilde oluşturulmuş iletileri (sosyal mühendislik iletileri olarak bilinir) ileterek kötölüklerini sürdürürler. Örneğın, bir kullanıcıya gönderilen sahte bir e-posta bir kötü amaçlı yazılım (tarayıcıda man-in-the-browser (MITB) olarak adlandırılır) içerebilir; bu kötü amaçlı yazılım, web tarayıcı ActiveX bileşenleri, eklentileri veya e-posta ekleri şeklinde olabilir; Bu kullanıcı farkında olmadan bu eki bilgisayarına indirirse, kötü amaçlı yazılım kullanıcının bilgisayarına kendini yükler ve kullanıcı (yani banka hesabının meşru sahibi) bir çevrimiçi işlem yapmaya çalıştığında bu bilgileri dolandırıcılara iletir [1]. Bu saldırıları durdurmak ve savunma mekanizmalarının kullanıcıların tekrar dolandırmalarını önlemek için daha güvenli hale getirilmeleri gerekir ki bu da mevcut savunma sistemlerinin (tasarımlarının ve teknolojisinin) büyük ölçüde geliştirilmesi gerektiğı anlamına gelmektedir [3]. Behdad ve diğeri [3] savunma sisteminin iyileştirilmesinin sahtekârları durdurmak için yeterli olmadığını ve bazılarının hâlâ nüfuz edebileceğine dikkat çekti; sistem aynı zamanda hileli dolandırıcılık faaliyetlerini tanımlayabilmeli ve gerçekleşmesini engelleyebilmelidir. Bugün çeřitli e-posta filtreleri tarafından kullanılan birçok geleneksel yaklaşım doğası gereğı statiktir; yeni ve ortaya çıkan kimlik avı modellerini kaldırabilecek kadar sağlam değildir; yalnızca mevcut kimlik avı kalıplarını işleme yetenekleri vardır, böylece e-posta kullanıcılarını yeni kimlik avı saldırılarına karşı savunmasız bırakmaktadırlar. Bu bir kısır döngü olarak devam etmektedir, çünkü dolandırıcıların faaliyetlerini sürekli hale getirme istekleri; tespit edilmeyi önlemek amacıyla çalışma biçimlerini mümkün olduğunca sık değiştirmektedirler. Bu, birçok araştırmacıyı hem bilinen hem de ortaya çıkan sahtekarlığı ele alabilecek diğeri etkili teknikler aramaya motive etti ve bu da makine öğrenme algoritmalarının keşfedilmesine yol açmıştır. Bu makine öğrenme algoritmaları, sınıflandırma amacıyla kullanılan bir veri kümesinden yeni veya mevcut kalıpları (özellikleri) keşfetmek için veri madenciliğı yöntemini kullanan yapay zekanın bir alt dalıdır. Derin öğrenme ise makine öğrenmesi algoritmalarında kullanılan yöntemlerin katmanlar haline getirilerek her katmanın birbiri arasında bilgi paylaşımını yapmasını sağlayan çalışma alanıdır. [4]

Bu çalışmada, sosyal mühendislik teknik ve taktiklerinin çeşitli biçimlerini ve sosyal mühendislerin insan güvenlik açıklarından nasıl yararlandıklarını, bu tür saldırılara karşı alınacak önlemlerle ilgili bazı yolları tartışmakta ve bu tür saldırıları önlemek için bilinçli olmanın önemi vurgulanmış ayrıca kimlik avı e-postalarından kaynaklanan sosyal mühendislik saldırılarından etkilenen kullanıcı sayısını azaltmak ve kimlik avı e-postalarını sınıflandırmak amacıyla karma derin öğrenme modelleriyle oluşturulmuş bir sınıflandırıcı yöntem önerilmiştir. Önerilen bu sınıflandırıcıda karma derin öğrenme modellerinin eğitilmesi için özellik çıkarılabilecek iki temel kısımdan oluşturulmuştur: E-postadaki, url sayısı veya e-postadaki ek sayısı gibi e-postanın özelliklerini barındıran başlık kısmı ve e-postanın alıcısı tarafından okunması amaçlanan e-posta metnini içeren e-posta gövde kısmı.

Bu yöntemde, 4512 e-postadan oluşan bir veri kümesinden 11 önemli kimlik avı özelliği (literatürden tanımlanmıştır) çıkarılmıştır [5] ve bu özellik çıkarımından sonra her e-posta için bu özelliklerin bir vektör temsili oluşturulmuştur; bu daha sonra modeli eğitmek için kullanılmıştır. İkinci kısımda bulunan e-posta metni ise çeşitli işlemlerden geçirilerek vektör haline getirilmiş ve farklı bir modeli eğitmek için kullanılmıştır. Daha sonra bu 2 modelden çıkarılan veriler farklı bir özellik matrisinde birleştirilerek son karar verici modele iletilmiş ve sınıflandırma işlemi tamamlanmıştır.

Tezin geri kalan bölümleri şu şekilde düzenlenmiştir: 2. Bölüm, sosyal mühendisliğe giriş, sosyal mühendislik ve kimlik avı saldırılarından ve bu alandaki ilgili çalışmaları özetlemektedir. Sonraki 3. Bölüm gelişmiş sosyal mühendislik saldırısı olan kimlik avı e-posta türlerini ve e-postaların derin öğrenme tekniklerini kullanarak kimlik avı veya meşru olarak sınıflandırılmasını, önerilen H-OLTA yöntemi hakkında detaylı bilgi verirken, ardından gelen Bölüm 4'te değerlendirme ve deney sonuçları vardır. Son olarak, Bölüm 5 sonuç ve önerilerdir.

2. KURAMSAL TEMELLER VE İLGİLİ ÇALIŞMALAR

Günümüzde sosyal mühendislik saldırıları büyük tehlikeler oluşturmaktadır. Sosyal mühendislik saldırılarına karşı koymak ve zarar risklerini en düşük boyuta indirmek için bilinçli olmak gerekir. Bu bölümde sosyal mühendislik saldırılarıyla mücadele etmek ve onları anlamak için temel anlamda gerekli kuramsal bilgiler verilecektir.

2.1. Sosyal Mühendislik Nedir?

Bilgi çağında yaşadıkça savaşlar bile maddi olmayan soyut ve siber alanlarda yapıldı. Bu saldırı türü sadece eğitilmiş birlikler, yaşam riski veya mühimmat gerektirmediğinden düşmanınızı kolayca gizlemeniz de mümkün olduğu için bu günlerde çok popüler oldu. Tüm bu nedenler siber savaşları bugün birçok ülke için çok çekici kılmaktadır. Bahsedilen birçok siber savunma konferansında olduğu gibi; saldırılar hem teknik hem de teknik olmayan alanlarda olabilir. Sosyal Mühendislik bu sınıflandırmanın “teknik olmayan” kısmına yerleştirilmiş ve savaş literatüründe “silah türü” olarak tanımlanmıştır [6].

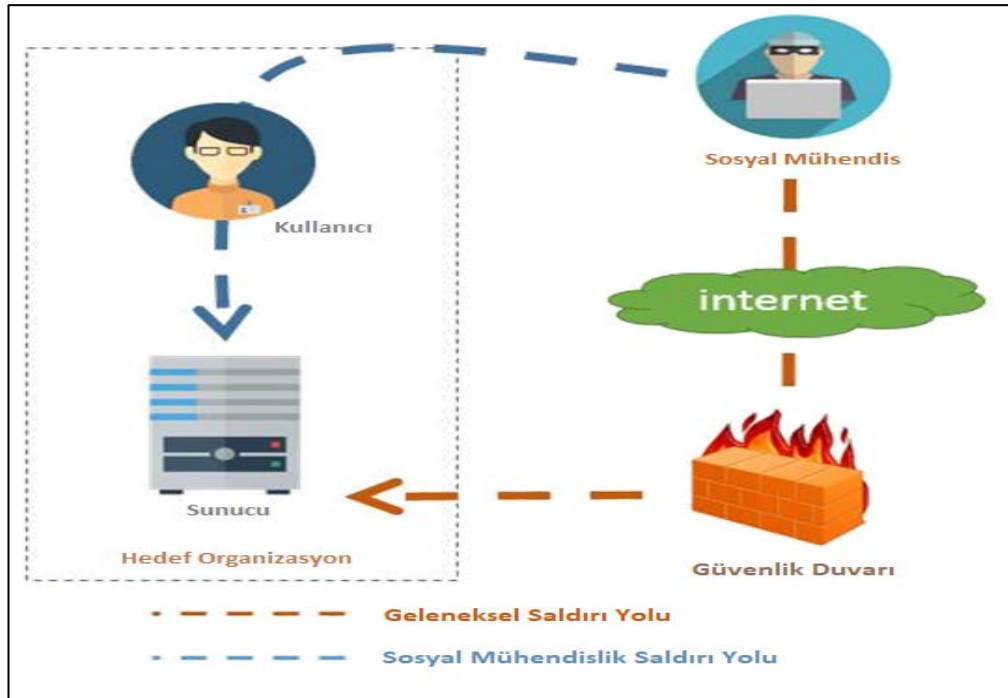
Geleneksel saldırı ve hedef sisteme kötü amaçlı kod gönderme gibi saldırılar çoğunlukla teknik tabanlıdır. Bu saldırı türleri, yazılım ve donanım sistemlerinin güvenlik açıklarından yararlanmaya çalışan hedefleme sisteminin kendisine veya uygulamalarına yöneliktir. Bununla birlikte, teknik ve teknolojik karşı önlemler giderek daha fazla kuruluş tarafından kademeli olarak kabul edildiğinden, bu teknik tabanlı saldırıların etkinliği azalmıştır. Bu durum teknik bilgisayar korsanlarını alternatif yolu seçmeye teşvik etti; sosyal mühendislik adı verilen ve hem insanların hem de teknolojinin güvenlik açıklarını hedefleyen teknik olmayan bir yöntemdir [7].

Sosyal mühendislik, insanların zafiyetlerinden faydalanarak çeşitli ikna ve kandırma yöntemleriyle istenilen bilgileri elde etmeye çalışmaktır. İnsanların karar verme süreçlerini değiştirmeye yönelik teknikler içerir. Saldırıları karşı zafiyet içeren davranışlar şunlardır:

- İnsanların birbirlerine güvenme eğiliminde olma istekleri

- Panik halinde insanların kritik bilgileri verme eğiliminde olma
- Diğer insanlara yardımcı olma istekleri
- Kolay yoldan para kazanma isteği

Sosyal mühendislik, kullanıcıların bilgi sistemlerinden ödün vermelerini sağlama sanatıdır [8]. Sistemlere yapılan teknik saldırılar yerine, sosyal mühendisler insanları bilgiye erişerek, gizli bilgileri ifşa etmeye, hatta etki ve ikna yoluyla kötü niyetli saldırılarını gerçekleştirmeye yönlendirir. Teknik koruma önlemleri genellikle bu tür saldırılara karşı etkisizdir.[9] Buna ek olarak, insanlar genellikle bu tür saldırıları tespit etmede iyi olduklarına inanırlar. Bununla birlikte araştırmalar insanların yalanları ve aldatmayı tespit etme konusunda zayıf performans gösterdiğini göstermektedir. Sosyal mühendislik saldırıları genellikle yabancılar(kuruluş için çalışmayan kişiler) tarafından yapılır; ancak bu içeriden gelenler için bir risk olmadığı anlamına gelmez. Saldırganlar, çoğunlukla ihtiyaç duydukları bilgileri vermek için “sistem kullanıcılarını” kötüye kullanarak hedef sistemlerinden yararlanmak için psikolojik hileler kullanırlar. Bu saldırı türleri çoğunlukla, teknik olarak güvenlik altına alınmış sistemleri kırmak için potansiyel risk almadan ve zaman kaybetmeden istedikleri bilgiye erişirler.



Şekil 2.1: Sosyal mühendislik saldırı yolu.

Şekil 2.1'de gösterildiği gibi, sistemleri internet bağlantısı üzerinden kesmek ve hedef kuruluş tarafından alınan güvenlik duvarlarını veya diğer güvenlik önlemlerini geçmek yerine; sosyal mühendisler sistem kullanıcılarını bilgilerini ele geçirmeye çalışırken daha kısa ve daha az riskli yolu seçerler. Ünlü sosyal mühendis Kevin Mitnick'in yukarıda belirttiği gibi, bu ikinci yol daha az zaman alacak ve daha az riskli olacaktır.

Birçoğumuz bilgisayar ağı ihlallerinin sadece teknik yollardan kaynaklandığını düşünüyoruz. Bilgisayar sistemlerindeki ve ağlarındaki teknik kusurlar veya güvenlik açıkları çoğu zaman davetsiz misafirlerin veya dışarıdaki kişilerin sömürülmesine veya istismar edilmesine neden olabilir. Çoğunlukla sistemdeki her güncelleme yeni tür kusurları veya güvenlik açıklarını doğurduğundan, teknik önlemler hiçbir zaman tam güvenli sistem ihtiyaçlarını karşılamak için yeterli değildir. Buna ek olarak, sosyal mühendislik türleri saldırıları da sistemlerin kullanılmasında önemli bir rol oynamaktadır. Bu saldırı türleri çoğunlukla bir saldırganın hedef sistemlerde herhangi bir veya daha az koda veya teknik kusura ihtiyaç duymadan teknik yazılım ve donanım güvenliği engelleri ve önlemleri arasında gezinmesine yardımcı olmayı amaçlar. Saldırgan, kurbanlarının kullanıcı adlarını ve şifrelerini öğrenmeyi veya sosyal mühendislik taktiği ve tekniklerini kullanarak herhangi bir alarm vermeden hedef sistemlerinin kontrolünü ele geçirmek için kimlik bilgilerini almayı amaçlamaktadır. Saldırılarından önce sosyal mühendisler kurbanlarının kırılganlıkları veya zayıflıkları hakkında derin araştırmalar yaparlar. Doğru zaman ve anı bekledikten sonra saldırgan seçtiği yaklaşımı uygun teknikle uygular.

Sosyal mühendislik saldırısı türleri, bu teknikleri basitçe doğrudan mağdurlara veya mağdura yakın olan kişilere yöneliktir. Sosyal Mühendisler çoğunlukla aşağıdaki parametreleri seçiyor:

- Aile
- Arkadaş
- Sosyal çevre
- Patron

Sosyal mühendisler, doğru saldırı doğru insanlara uygulandığında başarı düzeyi olasılıklarının saldırılarının daha yüksek olacağını bilmektedir. Bu saldırıları neden

gerçekleştirdiklerine gelince arkasında birçok sebep ve motivasyon vardır ve aşağıda gösterildiği gibi sınıflandırılabilir.

- Kişisel menuniyet
- Para kazanmak
- Sivil zarar
- Veri çalmak
- Siber suçlar
- Siber savaş
- Araştırma

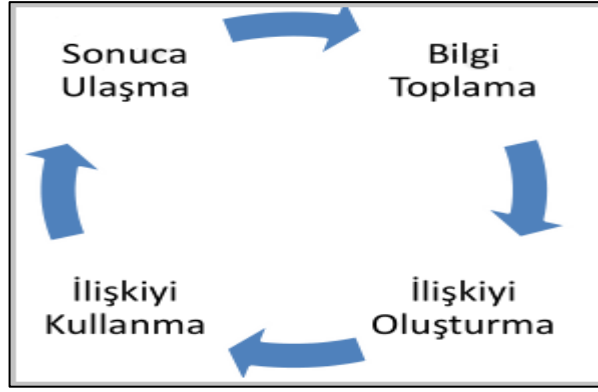
Bir sosyal mühendis için hangi tekniğin hangi insanlara uygulanacağına karar vermek çok kritik bir sorudur. Mağdurları hakkında her yönüyle ayrıntılı soruşturma yapmak zorundadır. İhtiyaç duyulan bilgiyi aldıktan sonra hedefe göre soruşturmaları harekete geçirmenin zamanı geldi. Saldırıların geri bildirimlerine göre saldırganın daha iyi sonuçlar elde etmek için yolunu güncelleyeceği açıktır.

Sonuç olarak bu durum günümüzde birçok kuruluşun ve bireyin karşılaştığı en büyük siber güvenlik tehditlerinden biri olarak kabul edilmektedir [12].

Günümüzde güvenlik şirketleri, müşterilerinin sisteminin siber güvenlik seviyesini teknik güvenlik açıkları açısından test etmektedir. Ne yazık ki bir süre sonra kullanıcılar kuruluşun BT politikaları ve bu tür farkındalık testleri hakkında bilgi sahibi olduklarında bu tür e-postalar gönderildikten sonra güvenlik politikalarına göre davranıyormuş gibi davranırlar. Dahası sistem kullanıcıları herhangi bir anket veya röportaj uyguladıktan sonra hangi seçeneği işaretlemeleri veya seçmeleri gerektiğini zaten biliyorlar ve sonuç olarak bu tür testler kuruluşların farkındalık düzeylerinin gerçek resmini yansıtmaz.

2.2. Sosyal Mühendislik Saldırılarının Döngüsü

Sosyal mühendislerin saldırılarını yaparken dört aşaması vardır.



Şekil 2.2: Sosyal mühendislik ataklarının izlediği döngü.

Şekil 2.2’de belirtildiği gibi; ilk adım bilgi toplama ve ön-anlatımdır. Bu aşamada sosyal mühendisler hedefleri hakkında yeterli bilgi toplamak için ellerinden geleni yaparlar. Kurbanlarının kişiliğini, alışkanlıklarını, beğenilerini, hoşlanmadığı şeyleri vb. keşfetmek için birçok araç, program, sosyal medya ve web sitesi kullanıyorlar [45]. Bir sonraki adım bir şekilde hedefle ilişki geliştirmektir. Bu teknikler kurban, duruma ve çevreye göre değişir. Bu ilişki güvene, saygıya, korkuya, karşılıklı iyilikler ve yardıma dayanabilir [10]. Mağdurla iletişim kurduktan sonra bir sonraki adım sömürüdür. Sosyal mühendis, doğru zaman olduğuna inandıktan sonra inşa edilen ilişkiden faydalanmaya çalışır. Saldırgan, kurbanı hakkında çok fazla bilgiye sahip olduğundan bu aşamada tekniğini ve taktiklerini daha ayrıntılı olarak hazırlayacaktır. Döngüdeki son adım yürütmedir. Bu aşamada saldırı taktiklerini ve tekniklerini eyleme geçirir. Daha sonra ihtiyaç duyacağı daha fazla bilgi almak için algılanıncaya kadar istediği şeyi almayı ve sistemden çıkmayı veya sistemde kalmayı seçebilir [11].

2.3. Sosyal Mühendisin Nitelikleri

Mağduru başarılı bir şekilde manipüle etmek ve saldırıların hedefine ulaşmak için sosyal mühendis kötü niyetlerini gizleme yeteneğine sahip olmalıdır. Ayrıca oyunu mağdurda akıllıca oynamak için hedeflerin zayıf yönlerini de bilmek zorundadır. Sonunda hangi tekniğin hedef üzerinde gerçekleştirileceğine karar verebilecek. Aynı zamanda ilişki geliştirme aşamasında mağdurun zarar görmesinden çekinmeyecek kadar acımasız olmalıdır. Dostlukları sadece hedeflerini ellerine geçene kadardır. Bu noktadan sonra başka hiçbir şeyi umursamıyor.

Bir sosyal mühendis çevresi ile ilişkilerinin samimiyet seviyesini ayarlamayı tercih eder. Bir kapıyı asla o kadar sert kapatmazlar çünkü bir gün o kapıya geri dönmeleri gerekebileceğinin farkındalar. Bu şekilde davrandığında gerektiğinde herhangi bir bilgi veya yardım isteme hakkına sahip olacaktır. Hediye almayı severler, insanların kalbini nasıl çalacağını bilirler. İnsanların hissedebileceği şüpheli hisleri öldürmek için gerçekten çok kibarlar. Güvenlerini kazandıklarında saldırırlar ve bu onların en acımasız tarafıdır. Büyük sırların küçük detaylar altında saklandığını biliyorlar bu yüzden bu konuya büyük önem veriyorlar. Çalışırken çok ciddi davranır ancak kendini ve başkalarını neşelendirmek için şakalar yapar. İhtiyaç duydukları zaman onları istismar etmek için başkalarının büyük çabasını kullanmaktan asla kaçınmazlar. Birisi kesin olarak onlar için “iletişim uzmanı” diyebilir [13].

2.4. Sosyal Mühendislik Yaklaşımları

Sosyal mühendislik saldırıları çok yönlüdür ve eylemin farklı aşamalarında kullanılan fiziksel, sosyal ve teknik yönleri içerir.

Bir sistemi ihlal etmenin en kolay yolu şu anda sorumlu olandan izin istemektir. Ne kadar şifreleme ve yazılım güvenliği önlemleri uygulanırsa uygulansın, bir ağ teknik veya teknik olmayan saldırılara karşı hiçbir zaman tamamen güvenli değildir. Tamamen güvenli sistemler söz konusu olduğunda en zayıf halka olan insan bileşeni asla göz ardı edilemez. Çalışanlar bilmeden veya amaçlarıyla isteyen sistemlerine erişim vermek isterse kaç tane sanal özel ağın (VPN), güvenlik duvarının, antivirüsün, kötü amaçlı yazılımdan koruma yazılımının veya şifreleme cihazının etkin olduğu önemli değildir [14].

Sosyal mühendis, bu en zayıf halkalardan doğru zamanda bilgi almak için aldatma, ikna, taktik ve teknikleri kullanan kişidir. Doğru kişiye doğru yaklaşımla doğru teknik uygulanırsa, bu organizasyona karşı başarılı bir saldırı olması anlamına gelir. Sosyal mühendis olmak, insan psikolojisini anlamak ve manipüle etmek ve birisini hassas, kritik bilgiler vermek veya saldırıya yetkisiz erişim vermek için yöntemsel bir şekilde kullanmakla ilgilidir. Başka bir deyişle, bu iyi bir yalancı olmakla ilgili değildir; insanları çok fazla zaman kaybetmeden ve yüksek risk almadan güvenli sistemlere daha kolay girebilme avantajı için etkilemenin yollarını keşfeden

bir mühendis olmakla ilgilidir. Bu alt bölüm saldırganların kullandığı farklı yaklaşımları açıklamayı amaçlamaktadır.

2.4.1. Fiziksel Yaklaşımlar

Adından da anlaşılacağı gibi, fiziksel yaklaşımlar saldırgan gelecekteki bir kurban hakkında bilgi toplamak için bir tür fiziksel eylem gerçekleştirir. Bu, kişisel bilgilerden (sosyal güvenlik numarası, doğum tarihi gibi) bilgisayar sistemi için geçerli kimlik bilgilerine kadar değişebilir. Sık kullanılan bir yöntem, çöp karıştırma[15], yani bir kuruluşun ya da kişinin çöpünü aramaktır. Çöp kutusu, çalışanlar, kılavuzlar, notlar ve hatta kullanıcı kimlik bilgileri gibi hassas bilgilerin çıktıları hakkında kişisel veriler bulabilen saldırganlar için değerli bir bilgi kaynağı olabilir. Saldırgan hedeflenen bir kuruluşun ofislerine erişebiliyorsa örneğin, açık planlı çalışma alanlarında, post-it notlarına yazılan şifreler gibi bilgileri bulabilirler. Daha az karmaşık fiziksel saldırılar, bilgi elde etmek için hırsızlık veya gasp gerektirir.[43]

2.4.2. Sosyal Yaklaşımlar

Sosyal mühendisliğin en başarılı saldırıları sosyal yaklaşımlardır. Saldırganlar bu vesile ile Cialdini'nin ilkeleri gibi sosyopsikolojik tekniklere güvenmektedir. Kurbanlarını manipüle ederek ikna ederler. İkna yöntemleri (iddia edilen) otoritenin kullanımını içerir. Cialdini [15] tarafından açıkça ele alınmayan yaygın bir sosyal vektör, hedef odaklı kimlik avı ve yemleme saldırılarında kullanılan meraktır. Bu tür saldırıların başarı şansını artırmak için, genellikle gelecekteki kurbanları ile bir ilişki geliştirmeye çalışırlar. Granger'e [15] göre en yaygın sosyal saldırı türü telefonla yapılmaktadır.

2.4.3. Tersine Sosyal Mühendislik

Potansiyel bir mağdurla doğrudan temasa geçmek yerine saldırgan onları güvenilir bir varlık olduğuna inandırmaya çalışabilir. Amaç potansiyel mağdurların ona yaklaşmasını sağlamak, örneğin yardım istemek. Bu dolaylı yaklaşım “tersine

sosyal mühendislik” olarak bilinir [13,15] ve üç ana bölümden oluşur: sabotaj, pazarlama ve yardım [16]. Bunun ilk adımı şirketin bilgisayar sistemini sabote etmektir. Bu herhangi birinin kullandığı sisteme basit erişimler elde eder, böylece sistemi bozar. Durumu fark eden kurban sistemin düzeltilmesi için yardım aramaya başlar. Saldırganlar daha sonra sorunu çözebileceklerini bildirir. Kurban sorunun çözülmesi için saldırgana istediği her bilgiyi vermeye başlar ve böylece kritik bilgilere erişim sağlanmış olur.

2.4.4. Teknik Yaklaşımlar

Teknik saldırılar çoğunlukla internet üzerinden yapılır. Granger [15], internetin sosyal mühendisler için şifreleri toplamada özellikle ilginç olduğunu, kullanıcıların farklı hesaplar için genellikle aynı (basit) şifreleri kullandığını belirtiyor. Çoğu insan, saldırganlara (veya arayacak olanlara) bol miktarda kişisel bilgi verdiklerinin farkında değildir. Saldırganlar genellikle gelecekteki kurbanlar hakkında kişisel bilgi toplamak için arama motorlarını kullanır. Farklı Web kaynaklarından bilgi toplayabilen araçlar da vardır. Bu tür en popüler araçlardan biri Maltego’dur [17]. Sosyal ağ siteleri de değerli bilgi kaynakları haline geliyor. Böylece daha rahat bir şekilde araçlar ile tarama yaparak kişisel bilgilere kolaylıkla erişilir.

2.4.5. Sosyo-Teknik Yaklaşımlar

Başarılı sosyal mühendislik saldırıları genellikle yukarıda tartışılan farklı yaklaşımların birkaçını veya tümünü birleştirir. Bununla birlikte, sosyo-teknik yaklaşımlar sosyal mühendislerin en güçlü silahlarını yaratmıştır. Buna bir örnek, yemleme(baiting) saldırısıdır: Saldırganlar, kötü amaçlı yazılım bulaşmış depolama ortamlarını gelecekteki kurbanlar tarafından bulunabilecek bir yerde bırakır. Örneğin bir Truva atı içeren bir USB sürücü olabilir [18]. Saldırganlar ayrıca caziplik katarak insanların merakından yararlanırlar. Depolama ortamlarına ilişkin etiketler, örneğin “gizli” veya “çalışanların işten çıkarılması 2019” adında etiket verirler. Teknik ve sosyal yaklaşımların bir diğer yaygın kombinasyonu kimlik avıdır. E-dolandırıcılık genellikle e-posta veya anlık mesajlaşma yoluyla yapılır ve spam'e benzer şekilde büyük bir kullanıcı grubuna yöneliktir. Sosyal mühendislik, genellikle bireylere veya

küçük insan gruplarına yöneliktir. Dolandırıcılar çok sayıda kullanıcıya mesaj göndererek kimlik avı saldırılarını karlı hale getirecek kadar insanı kandırabileceklerini umarlar. Herley ve Florencio [19], klasik kimlik avının kazançlı olmadığını, kimlik avı saldırılarının neden daha karmaşık “hedef odaklı kimlik avı” saldırılarına doğru ilerlediğini açıklayabilir. Hedef odaklı kimlik avı saldırıları, ilk veri madenciliğinden sonra yüksek oranda hedeflenmiş mesajlardır. Jagatic ve diğ.[20] sosyal ağ sitelerini öğrencilerle ilgili veriler çıkarmak ve arkadaşlarından biri tarafından gönderilmiş gibi görünen bir mesaj göndermek için kullandı. Böyle bir “sosyal veri” kullanarak kimlik avının başarı oranını yüzde 16'dan yüzde 72'ye çıkardılar. Bu nedenle, hedef odaklı kimlik avı, teknolojik yaklaşımların ve sosyal mühendisliğin bir kombinasyonu olarak kabul edilir.

2.5. Bilgisayar Destekli İş Birliği

İşletmeler ve çalışanlar günlük işleri kolaylaştırmak, otomatikleştirmek ve geliştirmek için çok çeşitli teknolojiler kullanır. Ayrıca iş birliğine dayalı iş yapılarının ortaya çıktığını da görüyoruz: Dosya paylaşımı veya işbirlikçi çalışma alanları, iç veya dış iletişim, bloglar, wiki'ler vb. için bilgisayar destekli iş birliği araçları, şirket içindeki müşterilerle bağlantı kurmaya, tüm müşterileri ve ortakları için sürekli bir iletişim kanalı oluşturulmuştur.

Bu bilgisayar destekli iş birliği araçları tarafından oluşturulan çok çeşitli iletişim kanalları göz önüne alındığında, sosyal mühendislik saldırıları büyük bir saldırı potansiyeline sahiptir. Ancak, iş bağlamında ofis iletişimi ile dış iletişim arasında ayrım yapıyoruz. Bu, bir kurbanın sosyal mühendislik saldırısını tespit etme yeteneği hakkında tahminlerde bulunmamızı sağlar.[44]

2.5.1. Ofis İletişimi

Modern iletişim araçları, personel arasındaki iletişim akışlarını büyük ölçüde değiştirdi ve yüksek hızlı bilgi alışverişini mümkün kıldı. Veri aktarımının güvenliğini koruyan gelişmiş teknolojiler vardır. Bununla birlikte, bu önlemlerin çoğu teknik saldırıları kapsamakla birlikte, sosyal mühendislik saldırıları dikkate alınmamıştır.

Kurumsal ortamlarda yüz yüze iletişimin yerini genellikle e-postalar veya anlık iletiler alır ve sosyal mühendisler için yeni bir saldırı yüzeyi oluşturur.

Açıkçası dahili hesaplardan veya dahili adresleri olan e-postalardan gelen sosyal mühendislik saldırılarının potansiyel bir mağdurun kimlik avı e-postasını ayırt edememe olasılığı daha yüksektir. Örneğin Parsons ve ark [21], 117 katılımcının kimlik avı e-postaları ile iyi huylu e-postaları ayırt etme yetenekleri üzerinde test edildiği bir rol oynama senaryosu denemesi gerçekleştirmiştir. Sonuçları daha yüksek bir farkındalık düzeyine sahip kişilerin daha fazla kimlik avı e-postalarını belirleyebildiğini gösterdi. Sosyal mühendislik saldırıları yoluyla kazanılan değerli kişisel bilgilerin, bir banka hesabının kullanılması veya itibar kaybı gibi dolaylı sonuçları gibi doğrudan sonuçları olabilir [22]; ayrıca daha fazla sosyal mühendislik saldırısının etkinliğini artırmak için de kullanılabilir. Genel olarak bir saldırı başarılı olduktan sonra çok yönlü sosyal mühendislik saldırılarıyla karşı karşıyadır, düşman bilgiyi daha başarılı sosyal mühendislik saldırıları gerçekleştirmek için kullanabilir.

2.5.2. Dış İletişim

Ofis içi iletişimde olduğu gibi e-posta hizmetlerinin, bulutun, blogların, vb. dış iletişim için de iç iletişimdeki ile aynı zorlukları yaratma eğilimi vardır. Ancak kurumsal sınır giderek bulanıklaştıkça, hangi bilgilerin yayınlanabileceğine veya harici bir iletişim ortağına iletebileceğine karar vermek zordur. Örneğin pazarlama blogları reklam amaçlı yararlıdır, ancak istenmeyen bilgi sızıntısı riski de taşır. Başka bir örnek potansiyel bir düşmanın birkaç yıl içinde kaç kişinin çalıştığını bulabileceği ve ilgili şirketin ekonomik durumunu çıkarabileceği, örneğin personel hakkında LinkedIn'deki bilgilerin yayınlanmasından çoğu bilgiye erişir [24]. Dış iletişimin en güçlü potansiyel riski çok çeşitli olası iletişim kanallarında yatmaktadır.

Elbette bu cihazların çoğuna güvenlik sistemleri kurulur; ancak bu sistemler sosyal mühendislik saldırılarına karşı koruma sağlamaz.

2.6. Sosyal Mühendislik Saldırısında Kullanılan Yöntemler

Sosyal mühendislik saldırılarında bilgi edinmek bazen kolay olabiliyor. En başta güven kazanmaya hal hatır sorarak başlar ve sonrasında onu överek kendi tuzağına

iyice çeker (herkes başkalarının kendisini övmesini sever biraz poh pohlanmak hoşuna gider insanların). Peki aynı durum sizin başınıza gelseydi ne yapardınız? Şirket içinde işe yeni başlayan personeller yeterli eğitimlerden geçirilmezse bir saldırgan sizin yerinize bu eğitimi verebilir ama ufak bir farkla artık sisteminiz de rahatça dolanabilecek bir şifreyi ele geçirmiştir. Bu saldırılarda sıkça kullanılan yöntem aşağıda verilmiştir.

- Kurumun herhangi bir bölümünde ki çalışan gibi davranmak
- Üst düzey yetkili gibi davranmak
- Yardıma ihtiyacı olan bir personel gibi davranmak
- Kendini acındırmak
- Omuz sörfü
- Personele içinde zararlı yazılım bulunan bir usb bellek hediye etmek
- Kurbanı zararlı olan bedava bir yazılımı indirmeye ikna etmek
- Güven kazanmak için şirket içi terimler kullanmak
- Onu önceden tanıyormuş gibi yapmak
- Karşı cinsi etkilemeye çalışmak
- E-Posta ekinde keylogger ya da trojen gibi zararlı yazılımlar göndermek
- Kullanıcının yeniden parola bilgilerini girmesini sağlayan sahte pencereler açmak
- Kendini Emniyet, İstihbarat mensubu gibi tanıtmak
- Teknik destek sağlamak için aradığını ve belli başlı adımları izlemesi gerektiğine ikna etmek
- Kurumun içerisine fiziksel olarak sızmak.

Sosyal mühendislerin kullandığı en sık kullanılan yöntemler yukarıdaki gibidir. Saldırganlar kendilerine bunlar gibi yüzlerce yöntem belirtebilir. Kurum içi personellerinin bu gibi süreçlerde özellikle dikkat etmesi gerekmektedir.

2.6.1. Tersine Sosyal Mühendislik (RSE)

Tersine sosyal mühendislik (RSE), aşağıdaki Şekil 2.3'te gösterildiği gibi üç basamağa sahiptir.



Şekil-2.3: Tersine sosyal mühendislik adımları.

İlk adımda, bir sosyal mühendis hedef sisteme teknik bir saldırı yaparak veya virüs ya da kötü amaçlı yazılımın bulaştığını bildiren bir e-posta göndererek hedef ağı sabote etmenin bir yolunu bulur. Uygun bir süre bekledikten sonra saldırgan hizmetlerini bir güvenlik danışmanı olarak tanıtır veya yetkisiz sistem ihlallerine karşı savaşmak için kurulmuş bir şirkettenmiş gibi davranır. Hedef organizasyon reklamı gördükten sonra sosyal mühendisle iletişime geçer ve sistemlerindeki hatayı düzeltmek için sosyal mühendisle görüşür. Bu aslında saldırganın istediği gibi ağlarında çalışmasına izin vermek anlamına gelir. Sosyal mühendis, içeri girdikten sonra sistemde neyin yanlış olduğunu zaten bildiği için problemi çözme izlenimi veriyor. Ancak saldırgan gizli verileri çalmak, keylogger eklemek veya sisteme ulaşmak için kimlik bilgileri almak gibi gerçekten kötü bir şey yapacaktır.

Kısaca RSE'de; sosyal mühendis hedeflenen ağda bir sorun yaratır ve kendini duruma karşı bir pozisyona sokar. Ardından sorunu çözdükten sonra kuruluşun güvenilir üyesi olur. Böylece saldırganın hedef sistemin kontrolünü ele geçirmek için yeni bir taktik veya teknik üretmesi gerekmez. Saldırgan sadece kendi başına bir sorun yaratır ve hedef sistemden yararlanmak için önceden oluşturduğu sorunu çözer. Sonunda şirket için “iyi, yararlı” bir adam gibi görünür. Doğru zaman geldiğinde, saldırgan hedeflerine ulaşmak için tüm haklarını kullanmaktan çekinmeyecektir.

2.6.2. Sırtına Binme (Piggyback Rides)

Sırtına Binme (Piggyback Rides), tekniği ile bir sosyal mühendis yetkili bir çalışan gibi davranır ve o binaya erişimi olan birini takip ederek güvenli bir binaya veya tesise girer. Bu saldırının klasik bir örneği; yağmurlu bir günde güvenli bir tesisin ön kapısında ağır bir kutu taşıyan bir sosyal mühendis ortaya çıkıyor. Yetkili bir çalışan yürürken sosyal mühendis, “Benim için kapıyı açar mısınız? Bu kutu yüzünden rozetime ulaşamıyorum.” İnsanlar genellikle başkalarına yardım etmek istedikleri için, yetkili çalışan güvenli kapıyı açar ve saldırgan kendi elleriyle erişim izni verir ve sosyal mühendisi ağır kutu dışında hiç çaba harcamadan kendini tesiste bulur. Bunun bir başka yaygın örneği, genellikle kuruluşun dışındaki çalışanların sigara içme alanlarında ortaya çıkmasıdır. Sosyal mühendis, diğer çalışanlarla sigara dışında durur; ve çalışanlar sigara içmeyi bitirdiğinde kart okuyucular gibi fiziksel güvenlik kontrollerini atlayarak binanın hemen arkasına geçerler. Burada kilit nokta; tüm bunlar sisteme karşı gerçekleşirken sistem uyarısı yoktur, anormal davranış için alarm veya işaret yoktur.

2.6.3. Teknik Konuşma

Bu saldırı türünde sosyal mühendis, kullanıcıların güvenli parolaların ne kadar kullanıldığını bulmaya çalışan bir yardım masası operatörü olarak davranır. Çalışanları ararlar ve şifrelerini “güçlü” bir şifre ile değiştirmeye ikna ederler. Buradaki saldırgan, çoğu kullanıcının muhtemelen bilmediği veya herhangi bir fikri olmayan çok teknik terimler kullanır. Parolalarını güncelledikten sonra yeni parolanın yeterince güçlü olup olmadığını “kontrol etmesini” isterler. Parolayı veritabanına kaydederken kontrol etmek için biraz bekledikten sonra kullanıcı hiçbir şeyden şüphelenmez. Bu saldırı türünde çalışanlar BT personelinin şirketin güvenlik saldırılarına karşı gerçekten iyi bir iş yaptığını düşünmesini sağlar.

2.6.4. Balina Avı(Whaling)

Bu saldırı türünde sosyal mühendislerin hedefleri yöneticiler, yöneticiler gibi yüksek profilli hedeflerdir. Bu kişiler hakkındaki kişisel bilgilere zaten birçok sosyal medya sitesinde veya şirketin resmi web sitesinde erişilebilir ve bulunabilir. Örneğin bir şirket yöneticilerinin kurumsal web sitelerinde hakkındaki çoğu bilgiler mevcuttur. Bu bilgiler bir sosyal mühendis tarafından şirket yetkilisine kullanılabilir, çünkü tüm bilgiler zaten şirketin kendisi tarafından doğrulanır. Eğer biyoloji müdürü X Üniversitesi'nden mezun oldu ve basketbol oynamaktan hoşlanıyorsa (bazı yöneticiler hobilerini bile biyografilerine koyuyorlarsa), bir sosyal mühendis bu bilgiyi kötüye kullanabilir. Üniversite mezunları bölümünden mezunlar için özel bir mezunlar basketbol turnuvasına gelmesini teklif eden şirket yetkilisine bir e-posta gönderebilir. Yönetici muhtemelen davetiyenin gerçek ve geçerli olduğunu düşünecektir çünkü otantik görünecektir. E-posta, kişiden turnuvada iyi bir yer ayırmak için kredi kartı bilgilerini girmek üzere bir web sitesine erişmesini istemek için devam edebilir. Kurumsal memurlar ve diğer yüksek profilli hedefler hakkında çok fazla bilgi olması nedeniyle, balina avcılığı giderek daha popüler hale geliyor. Bu bilgi, sosyal mühendislerin bunları doğrulanmış bilgilerle ikna edici bir şekilde hedeflemelerini kolaylaştırır. Dikkatsizce verilen bilgileri kullanarak bile ilgi alanlarını veya neyi sevip neleri sevmediklerini bilen belirli hedeflere saldırabilirler.

2.6.5. Çöp Karıştırma

Çöp karıştırma, saldırı esnasında işe yarar bilgiler bulmak için hedefteki kurum veya kişinin çöpünü karıştırmadır. Genellikle kurumun kurum sınırları dışında kullandığı çöpler tercih edilir. Bu saldırılarda saldırganın girdiği risk hemen hemen yok denecek kadar azdır (bir tek kirlenme riski vardır) çünkü kurum dışına attığımız çöpler halka açık alanlarda bulunuyorsa çöp dalışı tamamen yasaldir, kişiler ve kurumlar çöplerinden kendileri sorumludur. Çöp bidonları kurum sınırları içerisinde bulunmalı ve dışarıdan erişime mutlak kapalı olmalıdır. Aksi takdirde çöp kutumuzdan faydalanmak üzere hali hazırda bekleyen saldırganlar ve rakip firmanın elemanları çöpünüzden menfaat sağlayabilir. Çöplerimizi atarken de kurum içinde belli başlı kurallar olması gerekir. Mesela kağıtlar mutlaka okunamayacak derecede ayrıştırılmalı

ya da kurum için önemli bilgileri içeren bilgiler taşınabilir veya sabit sürücüler verileri tamamıyla okunamaz hala gelmeden çöpe atmamalıyız. Bizim bilinçsizce attığımız çöpler düşmanlarımızın hazinesi olabilir. Peki bu saldırganlar bizim çöplerimizde neler arıyor birkaç örnek vermek gerekirse;

- Çalışan bilgileri
- Şifreler
- İmla hatasından dolayı atılan raporlar
- Güncelliğini yitirmiş telefon rehberi
- Müşteri listeleri
- Faturalar
- Girilecek ihale ile ilgili bilgiler
- Şirket içindeki donanımların kullanım kılavuzu
- Projelerimiz, programlarımız ve cihazlarımız hakkındaki dokümantasyonlar

2.6.6. Telefon ile Yapılan Kimlik Avı (Vishing)

Vishing, kimlik avı saldırısının eşdeğerini gerçekleştirmek için telefonu kullanan bir saldırı türüdür. Bu saldırıda saldırgan mağdurun yanıtını beklemek yerine, mağdurdan aldığı tepkilere göre kendini ayarlayabilir. Sosyal mühendis, konuşmayı ve kurbanı ikna etme fırsatını yönetme konusunda bir girişime sahiptir. Mağdurun telefondaki ses tonu bile sosyal mühendise hedef şirketten hedefine ulaşmak için bir sonraki adımları hakkında bir ipucu verecektir. Yaygın bir örnek ve oldukça etkili olan, bunu önceden kaydedilmiş bir mesaj yardımıyla gerçekleştiriyor. Telefon cevaplandığında kaydedilen mesaj aramanın kişinin bankasından olduğunu ve kredi kartlarının güvenliğinin ihlal edildiğini söyleyebilir. “Kurbanlardan” sorunu çözmek için bir numarayı aramaları istenir. Kullanıcı numarayı aradıktan sonra kurbanın kredi kartı numarasını, pin'ini, adresini ve sosyal mühendisin istediği veya ihtiyaç duyduğu her şeyi girmesini isteyen başka bir otomatik mesaj duyar. Bir sosyal mühendis diyalogu kendi başına yönetmeye karar verirse, ofis arka plan seslerini kullanmak işini kolaylaştıracaktır. Böylece kurbanlar şirket ya da bankadaki meşru insanlarla konuştukları konusunda daha ikna olacaklar.

Telefon ile yapılan kimlik avı saldırısının bir diğer popüler varyasyonu, kişiyi doğrudan aramak yerine kurbanı kısa mesaj göndermektir. Bu başka bir kimlik avı saldırısı türü olarak düşünülebilir. Günümüzde çoğumuz bu tür mesajlara aşinayız. Türkiye'de Emniyet Müdürlüğü son derece güvenilir organizasyonlardan biri olduğu için sosyal mühendisler bu departmanın adını para da dahil olmak üzere ihtiyaç duydukları bilgileri almak için kullanıyorlar. Bu bölümün personelinin kimliğine bürünüyorlar ve bazı kritik bilgileri veya doğrudan elde etmek istediklerini almaya çalışıyorlar.

Telefon ile yapılan kimlik avı onları desteklemek için diğer saldırı türleriyle kullanılabilir. Böylece mühendisin kurbanı karşılaştığı durum hakkında bir şeyler yapması gerektiğine ikna olur. Mağdura herhangi bir zaman sınırı verilirse başarı olasılığı önemli ölçüde artar.

2.6.7. Sosyal Medya Ağı

Facebook veya twitter gibi sosyal ağ siteleri bir sosyal mühendis bilgi havuzudur. Bir sosyal mühendis, yalnızca bazı sayfalara tıklayarak bu tür sitelerden hedefleri hakkında çok fazla ayrıntı keşfedebilir. İnsanlar nerede çalıştıkları, ne yapmayı sevdikleri, hangi müzik gruplarını sevdikleri, şu anda nasıl hissettikleri, sevmedikleri veya korktukları ve daha fazlası hakkında bilgi paylaşırlar. Bir sosyal mühendis, yayınlanan bu bilgileri kendi sayfalarından kurbanlarına karşı kötüye kullanabilir. İnsanların özel hayatlarını sosyal sayfalarından açıkça paylaşmaya istekli olması sosyal mühendislerin ellerini bu kadar güçlü kılar. İnsanlar şu anda nerede olduklarını paylaşıyor ve her saniye konumlarını güncelliyorlar. Dolayısıyla bir sosyal mühendis kurbanının yakınlarda olduğunu bilirse, taktiklerini doğrudan onlara karşı harekete geçirebilir. Mağdurun çalışma ortamından sosyal mühendis istediği şeyi elde etme şansını denerken daha rahat ve özgür olacaktır.

Bu yaklaşım çeşitli şekillerde kullanılabilir:

- Mağdurun sayfasında listelenen bir arkadaşınızı taklit eden e-posta gönderme,
- Popüler hangout'ları keşfetmek için bir kişinin resimlerini görüntülemek ve daha sonra yakınlarda veya sadece aynı alanlarda göstermek,

- Kişinin yaşını, okulunu, önceki şirketlerini, doğum yerini ve bunların tümünü hedeflemek için kullanılabileceğini keşfetmek,
- Güven oluşturmak için kurbanı bir kişiyle çevrimiçi ilişki kurmak üzere arkadaş olarak eklemek.

Sosyal mühendis daha sonra kurbanla ilgili edindiği bilgileri başka bir saldırı başlatmak için kullanmakta tereddüt etmeyecektir.

2.6.8. Nöro-dilbilimsel Programlama (NLP)

Hadnagy tarafından 2010'da belirtildiği gibi iyi bir sosyal mühendis insan yazılımını nasıl manipüle edeceğinizi veya bilgilerini ele geçirmeyi çok iyi kavrar. Nöro-dilbilimsel programlama (NLP), birçok sosyal mühendis tarafından kurbanlarını manipüle etmek için kullanılan güçlü psikolojik yaklaşımlardan biridir. Bu teknik doğru uygulandığında hedefi manipüle etmeyi kolaylaştırır. NLP, bir kişinin nörolojik süreçleri, dili ve öğrenilmiş davranış yanıtlarıyla ilgilenir. NLP başlangıçta terapötik hedeflerde kullanılmak üzere tasarlanırken, sosyal mühendisler tarafından kurbanlarını yapmak için insan aklını manipüle etmek ve bilgilerini ele geçirmek için kullanıyor [26]. Bu teknikte sosyal mühendis, manipüle etmeye çalıştığı kişiye bilinçaltı mesajlar vermek için beden dilini ve dikkatli bir sözcük seçimi kullanmanın yollarını arayacaktır. Aday, beden dilini kurbanın beden dili ile eşleştirerek başlar. Ayrıca nefes oranı, ses seviyesi, aksan ve kelime hazinesiyle mağdurla eşleşir. Bunu yaparak bilinçaltı bir seviyede bir ilişki kurmaya çalışır. Daha sonra beden dilini değiştirerek, omuzlarına veya kollarına hafifçe dokunarak olumlu düşünceleri, duyguları ve imgeleri ifade eden kelimeleri kullanarak başka bilinçaltı mesajları da verebilir. Tüm bu dokunsal, görsel ve sözlü eylemler (NLP terimlerinde demirleme ve yeniden çerçeveleme denir), kişiyi olumlu duygulara sahip olmak ve sosyal mühendis ile uyum hissi kazanmak için etkileyen bilinçaltı mesajları verir. Bu noktadan sonra bir şirketin sırları veya kritik bilgileri hakkında bilgi toplamak gibi sonrasında neyin peşinde olduğu ile iletişim kurabilir.

2.6.9. Karşıt Cinsiyet Çekiciliği

İnsani çekiciliği kullanmak, sosyal mühendise ilgi duyan birini bulmak ve kurbanı duygularının karşılıklı olduğu izlenimini vermekle ilgilidir. Bu kişiyi kurum hakkında kritik bilgi toplamadan, bir binanın çöp toplama anahtarlarına kadar saldırganlara karşı savunmasız bırakıyor, mağdur dikkat etmiyor. Bir sosyal mühendis kurbanının karşı cinsiyetten etkilenme konusunda kırılgan olduğunu öğrenirse bu tekniği istediğini elde etmek için hedefe karşı kullanır.

2.6.10. Dost Toplantıları ve Konuşmalar

Bir sosyal mühendis hedeflenen şirket veya kuruluş hakkında daha fazla bilgi edinmek isterse, barlara veya eğlence merkezlerine gitmeyi seven hedef kişiyle dışarıda kalabilecekleri bir anı arayabilir. Sosyal mühendis, işten sonra hangilerinin bara gideceğini görmek için işten eve gelen insanları takip edebilir ya da bar veya kulüplerin isimlerini açığa çıkarabilecek resim veya başka bir bilgi olup olmadığını görmek için insanları sosyal ağ sitelerinde arayabilir. Bu bilgilerle donanmış olan sosyal mühendisi, bar / kahve dükkanındaki kurbanla bir konuşma yapabilir ve bazı kritik bilgileri açığa çıkarmaya çalışabilir. Bu görevin yerine getirilmesi için bir sosyal mühendisin birçok yolu vardır. Sosyal mühendis, hedef kişinin hangi bar / kahvede sık sık ziyaret ettiğini öğrendiğinde, barmenle konuşmaya başlamak için erken gelebilir. Barmenlere her zaman alkolsüz içkiler hazırladığından emin olmak için bir miktar nakit para karşılığında rüşvet verebilir. Bu şekilde sosyal mühendis ayık kalır ve hedef kişi sarhoş olurken hedefine odaklanabilir. O gecenin ilerleyen saatlerinde sosyal mühendis hedef kişiyle bir görüşme yapabilir ve hedef kişisini sarhoş etmeye çalışabilir. Hedef sarhoş olduktan sonra sosyal mühendis iş konusunu gündeme getirebilir ve bir binaya nasıl girileceği, şifreler, ticari sırlar ve daha fazlası gibi kişinin asla paylaşmayacağı bilgileri almaya devam edebilir. Böyle bir ortam olmanın bir başka olumlu etkisi, sosyal mühendislerin ofis ortamına göre daha rahat olmalarıdır. Mağdurun ise şirket kurallarının ve katı olacağı kendi çalışma alanı olmayacaktır.

Ortak sosyal mühendislik saldırı tipleri hakkında bazı ipuçları verdikten sonra günümüzde popüler sosyal mühendislik tekniklerine girmenin zamanı gelmiştir. Saldırı türleri kurbanların dünyasına girerken aşağıdaki teknikler kullanılıyor.

2.7. Bazı Popüler Sosyal Mühendislik Teknikleri

Bu bölümde bir sosyal mühendis için insanları manipüle etme ve başkalarının istedikleri gibi davranmalarını sağlama konusunda bir beyin fırtınası çalışmasını içermektedir. İşte bu bölümde en yaygın kullanılan tekniklerden bazıları verilecektir.

2.7.1. Yalan

Bütün insanların dürüst olmadıklarının farkında olmalıyız. İnsanların çoğu başkalarını kendi iç niteliklerine, kendi niteliklerine göre değerlendirir. Bir insan yalan söyleme alışkanlığına sahip değilse, muhtemelen başkalarının da yalan söylemesini beklemeyecektir. Bu bir insanın normal bir doğal davranışıdır.

Önceki bölümlerde belirtildiği gibi, sosyal mühendisler kurbanları hakkında bilgi edinir ve kurbanlarına karşı elde ettikleri bilgileri kullanırlar. Davranışlarının ahlaki sınırlarda olup olmadıklarını kendileri sorgulamazlar. Sadece hedeflerine odaklanmışlar. Bu alışkanlık, onları çok profesyonelce taklit etmeye yönlendirir. Bir süre sonra başkalarını manipüle etmek için garip bir şey olarak görülmemesi gereken yalancılık tekniğini kullanabilirler.

2.7.2. Kısmi Gerçeği Söylemek

Başka bir teknik ise gerçeği söylemek, ancak bazı kısımlarını atlamaktır. Bu yalan söylemeye alışık olmayan ve daha önce söyledikleri tüm yalanları takip edemeyenler için iyi bir tekniktir. Sadece belirli gerçekleri belirtirler ve tüm durumlar ve olasılıklar hakkında başka bir şey söylemezler. Burada saldırganın ana fikri önce güven kazanıyor. Onayladıktan sonra mağdurla aynı şekilde düşündüklerini, eylemlerini gerçek kötü niyetlerine göre yaptıklarını belirlediler. Fakat o aşamaya kadar kurbanlarının gerçeği söylediklerini düşünmelerini sağlamalıdır.

2.7.3. Sebep Vermek

Eğer saldırgan bir rasyonalizasyon verirse ve “çünkü” kelimesini kullanırsa, sebep saçma olsa bile mağdurun kendisinden istediği gibi davranması daha olasıdır. Burada “çünkü” sihirli kelimesinin anlamını genişletmek daha iyi olacaktır. Akılcılaştırma ile ilgili bir çalışma yapılmış. Deney, fotokopi makinesini kullanmak için bir kütüphanede gerçekleştirilir. Makalelerin kopyalarını almak için sıraya giren birçok insan var ve bir kişi gelir ve sıraya girmelerini ister. Birinci grupta, sırayı bozmak isteyen kişi diğerlerine sorar:

- “Affedersiniz, beş sayfam var. Fotokopi makinesini kullanabilir miyim çünkü acelem var” der.

Bu teklif sonrasında katılımcıların %94’ü sayfalarını onlardan önce kopyalamasına izin veriyor. Farklı bir grupta aynı kişi soruyor:

- “Affedersiniz. Beş sayfam var. Fotokopi makinesini kullanabilir miyim?”

Sebepsiz böyle bir tekliften sonra; sadece %60’ı sıraya girmesine izin veriyor. Yeni adamın kopyalarını almasına izin vermeyi kabul etmiyorlar ve sadece bahsettiği aynı sebepten beklediklerinden şikayet ediyorlardı. Son grupta, aynı kişi tekrar sorar:

- “Affedersiniz, beş sayfam var. Fotokopi makinesini kullanabilir miyim, çünkü kopya çıkarmam gerekiyor.”

Böyle garip bir sebep duyduktan sonra %93 kişi onun sıraya girmesine izin veriyor. Bu açıkça, “çünkü” kullanımının, insanların çizgiyi kesmek için geçerli bir nedeni olduğunu ve nedenini kendisi bile işleyemediklerini düşünmeleri için yeterli olduğunu göstermektedir [27].

2.7.4. Kaçınma ve Derivasyon

Kaçınma ve sapma, sosyal mühendisler tarafından kullanılan iki yaygın tekniktir. Bu teknikte; saldırgan, konuyu değiştirerek veya sorunun dikkatini yönlendirmek için bir cevap olarak belirsiz veya ilgisiz bir şey söyleyerek cevap vermeye çalışan sorulardan kaçınır. Ayrıca, kurbanı kandırmak veya hedefte suçluluk oluşturmaya çalışmak için masumiyet, öfke veya kafa karışıklığı gibi davranabilirler.

2.7.5. Karşılıklı İlişki

Bu numara, kurbanı cevap vermek zorunda kalacağı bir hediye vermeyi ya da bir nevi iyilik yapmayı içerir. Burada sosyal mühendis, kendisi ve kurban arasında duygusal bir bağ kurmaya çalışıyor. Bu tekniğin, mağdurun mühendisin kendisine rüşvet verdiğini hissetmesi durumunda uygularken bazı önemli noktaları vardır. Eğer mağdur mühendisin kötü bir niyeti olduğunu hissediyorsa, o zaman mühendise normal olarak uygun olmayacaktır. Bir süre bekler ancak çok fazla değil, çünkü potansiyel mağdurun borçlu hissini bırakması durumunda karşılık alma duygusunun geçerliliği sona erebilir.

2.7.6. Mizah Kullanımı

Sosyal mühendislerin çoğu kurbanlarını güldürmek ve ilişki kurma aşamasında rahat ve mutlu hissetmelerini sağlamak için önceden hazırlanmış bazı “anlık” şakalara sahiptir. Çoğu insan iyi bir mizah anlayışı olan kişilerden zevk aldığından bu tür bir teknik kullanıldığında başarılı saldırı olasılığı artar. Bu tekniğin bir başka avantajı, eğer bir şeyler ters giderse mühendisin sorunlu durumu komik bir duruma dönüştürme şansı vardır. Saldırı mizah dolu olduğu için daha görünmez.

2.7.7. Evet-Evet Tekniği

Bir sosyal mühendisin hedeflerini olumlu bir zihin durumuna getirmesi önemlidir. Bunu yapmayı başarır ise olası yüksek başarı anlamına gelir. Bu teknikte kurbanı

muhtemelen “evet” ile cevaplanan çok fazla soru soruldu. Mesela “bu hayatta başarılı olmak istiyorsun, değil mi?”, “Emekli olduğunda daha iyi koşullarda yaşamak istiyorsun, değil mi?” Mağdura bir tür soru soruluyor ve “evet” ile cevap vermesini bekliyor. Saldırgan, birçok sorudan sonra “şirketimizle çalışmak ister misin?” Veya “bize güveniyorsanız, yarın sisteminizi yarın test etmemize izin verebilir misiniz?” Gibi ana soruyu sorar ve normal olarak kurbanından bekler tekrar “evet” demesini.

2.7.8. Top Atma Tekniği

Bu saldırı türünde mağdura birçok harika fırsat verilir ve saldırgan tarafından gelecekte neler yapabileceğini hayal etmeye zorlanır. Mağdurla rüyalar arasında duygusal bağlantı sağlandığından emin oluncaya kadar saldırgan bekler. Saldırgan bir şekilde şartlar nedeniyle konuştukları her şeyi yapamadıklarını söyleyerek ortaya çıkıyor. Saldırgan, kurbanı eşyalarla ilişkilendirdikten sonra o kadar ileri gittikten sonra geri dönüşün zor olacağını biliyor. Böylece mağdur saldırganın yeni teklifini kabul eder.

2.8. İnsan Zayıflıklarının Önemli Hedefleri

Sosyal mühendisler, aşağıda anlatıldığı gibi, SANS Enstitüsüne (siber güvenlik eğitimi ve sertifika satış) göre insanda yaşayan dört psikolojik sorundan birini veya birkaçını kullanır [28].

2.8.1. Korku

Korku, en çok sömürülen insan özelliklerinden biridir. Bu genellikle mağdurda hemen bir hareket hissi yaratan birçok saldırganın kullandığı diğer tekniklere göre e-posta ve sesli kimlik avından farklı olabilir. Bu özellik mağdurun kullanıcı adının ve şifresini ele geçirdiği, bilgilerin çalındığı veya banka hesabında şüpheli faaliyetler veya benzeri bir şey olduğu iddiasıyla kullanılabilir.

Mağdurun işini kaybetme veya adli yaptırımlarla tehdit etme gibi senaryolar olabilir. İnsandaki bu özellik aslında bir sosyal mühendis için önemli. İnsanların çoğu

hâkim veya mahkeme hakkında bir şeyler duymaktan bile korktukları için kendilerini derhal hareket edeceklerini düşünüyorlar. Tehdit edildikten sonra sakin olmaları zor olacak. Ayrıca tehdit mağdurun aile üyelerini içeriyorsa saldırganın amacına ulaşmak için yüksek bir başarı olasılığı anlamına gelir.

2.8.2. Yardım İsteği

Bir sonraki en çok sömürülen özellik “yardım etme arzusu” veya “başkalarına karşı nazik olmak”. Bu saldırı türü çoğunlukla iki yolla uygulanır: sırtlama veya kimliğe bürünme. Sırtına binme metodu yukarıda tartışılmıştır. Saldırgan temizleyici olarak giyinebilir. Kullanması ve taşınması gereken çok şey olduğu için hedef departmana girerken ve onun için kapıyı açık tutmalarını isteyerek izin verilenleri durdurur. Çoğu durumda yetkili eşyalar sadece temizlik işini rozeti veya giriş kartı için kontrol etmeden gitmesine izin verir. Bunların hepsi insan doğasında sahip olduğu özelliklerden yardım etme arzusundan kaynaklanıyor, tıpkı birileri bizim için aynı şeyi yapsaydı bizim de sevdiğimiz gibi. Ayrıca kullanıcılar departman dışına çıktıklarında binadaki temizlikçilere izin vermek gibi bazı senaryolar vardır. Bu tür davranışların temel nedeni temizleyicilerin işlerini daha iyi yapmalarına yer açmaktır, ancak kuruma zarar vermek isteyen kötü niyetli herhangi bir kişi bu özelliği kötüye kullanabilir. Temizlik aletleri ve bakım elemanı olmak zorunda değildir. Sosyal mühendis bile bir teslimatçı olarak giyebilir ve büyük, ağır bir kutu taşıyabilir. İsim bırakma tekniğini bile kullanabilir ve teslimatın patrandaki veya şirketteki bazı üst sıradaki kişiler için olduğunu söyleyebilir. İnsanlar patronlarının adını duyduklarında kanatlarını saldırgana doğru indiriyor ve saldırganların içeri girmesine izin veriyorlar. Yardım etme isteğinden yararlanmanın bir başka yolu da içeriden kimliğe bürünmek. Bu yukarıda da tartışılmıştır. Kimliğe bürünme genellikle keşif yöntemleri veya çöp tenekesi dalışı gibi diğer tekniklerden sonra kullanılır. Saldırı için gerekli bilgileri aldıktan sonra sosyal mühendis hedeflerine ulaşmak için adımlarını attı.

2.8.3. Dikkatsizlik

İnsanların dikkatsiz özellikleri çoğu insanın örgütlerinde güvenlik politikası kurallarına kayıtsız hissettiği gibi sömürülebilir. Çöp karıştırma çöp ve birliğin

araştırılmasını veya imha edilmiş belgelerin parçalanmadan atılması gibi yararlı bilgilerin sağlanması için atıkları yeniden bir araya getirir; bu özelliğe karşı yapılan saldırılara iyi bir örnek olabilir [14]. Bu özelliğin sömürülmesi genellikle daha karmaşık bir saldırının sadece bir adımıdır. Bu özellik genellikle saldırıyı daha başarılı bir başarı oranı haline getirmek için gerekli bilgileri toplamak amacıyla keşif aşamasının bir parçası olarak kötüye kullanılır. Şifre hırsızlığı bu istismarın bir başka örneği olarak görülebilir. Çoğu şirket çalışanlarını güvenlik politikalarına göre düzenli olarak şifrelerini değiştirmeye zorlar. Bu politikayı uygularken en azından bazı harfleri, özel karakterleri veya sayıları içeren yeni şifrenin gücüyle ilgili katı kurallara sahiptirler. Bu durum çalışanların her değişiklikten sonra güncellenmiş şifrelerini hatırlama mücadelesi veriyor. Bir süre sonra kendi şifrelerini takip edemezler. Çalışanların çoğu ne yazık ki güncellenmiş şifrelerini bir kağıda yazarak bu soruna bir çözüm buluyor. Kağıdı gizlemeyi ya da kağıdın altında kolayca tahmin edilebileceği bir yere koymayı unuturlar. Çoğu çalışanın bu davranışı parola hırsızlığını kolaylaştırır, çünkü saldırgan sadece ofis bilgisayarının çevresini örneğin sandalye, masa çekmecesini, klavyenin altında vb. kontrol etmesi yeterli olmaktadır. Sistem kullanıcılarının yaptığı başka bir teknik ise bir program kurmaktır. İnternette indirilen güncellenmiş kullanıcı adlarını ve şifreleri, muhtemelen P2P programlarını depolar. Bu şekilde bir saldırganın kullanıcı adlarını ve şifrelerini ele geçirmesi için mükemmel bir şans veriyorlar.

2.8.4. Konfor Bölgesi

Sosyal mühendislerin sömürdüğü bir sonraki özellik kurbanlarının konfor bölgesidir. İnsanlar işyerlerinde kendilerini evlerinde güvende hissediyorlar. Bu duygu bu güvenlik hissinden yararlanabilecek olası tehditlere, dolandırıcılığa karşı daha az algılayıcı olmalarını sağlar. Bu özelliğin yaygın olarak kötüye kullanılması içeriden bir kişiyi taklit etmektir. En yaygın kimliğe bürünme BT personelidir. BT personelini seçmenin nedeni insanların genellikle BT sistemlerinin nasıl korunduğuna dair çok fazla bilgiye sahip olmamaları veya bu alanda bir şekilde bilgisiz olmalarıdır [29]. Çalışanlar BT personeli gibi gömlek giyen bir kişi gördüğünde, insanlar büyük olasılıkla BT departmanında çalıştığını varsayacaktır. Sonuç olarak kendilerini korumaları normal şekilde düşecektir. Bu şekilde mağdurun güvenini kazandıktan

sonra mağdurun bilgisayarında bazı kritik yazılım güncellemelerini kolayca yapmasını isteyebilir. Mağdurun bilgisi dışında, saldırgan kendisine uzaktan erişim hakkı verebilir.

Başka bir senaryo, saldırganın çalışanlardan tamamen normal gibi görünen yüksek düzeyli güvenli şifre kullanıp kullanmadığını kontrol etmelerini istemesi gibi olabilir. Böyle bir saldırıdan sonra bile bazı durumlar olmuştur; çalışanlar saldırgana güvenliği için iyi ve yardımsever olduğu için teşekkür eder.

Omuz sörfü, birçok sosyal mühendis tarafından başkalarının konfor bölgesini kötüye kullanabilecekleri bir başka yöntemdir. Aslında bu teknik hem konfor bölgesinin sömürülmesini hem de dikkatsizliğini içerir. İnsanlar çoğunlukla sırtlarını kontrol etmezler. Birisinin görevde yoğunlaştıklarında klavyede kullanıcı adlarını ve şifrelerini yazarken birisinin arkasında durup onları izleyeceğini düşünmüyorlar. İşyerindeki konfor bölgeleri nedeniyle kendilerini güvende hissetmek de burada sömürülebilen bazı nedenler olarak görülebilir.

Sosyal mühendislerin yapabileceği başka bir saldırı türü, bir dizüstü bilgisayar, harici sabit sürücü, USB, rozet, cüzdan, çanta, akıllı telefon ve giriş kartları gibi işle ilgili diğer makineleri çalmaktır. Sosyal mühendislerin çoğu bu tür şeyleri çalmak için çoğunlukla cuma gününü seçiyor çünkü bu sosyal mühendisin keşfedilmesi için biraz daha zaman tanıyacaktır. Bu tür davranışların ana nedeni mağdurun bildirisinin pazartesi sabahı ve hafta sonu yaklaşımı olacaktır.

Fiziksel güvenlik bir başka konfor bölgesi tehdididir. Örneğin, sigara içilen alanlarda, insanlar genellikle kartlarını tekrar tekrar basmak zorunda kalmadan dışarı çıkabilmeleri için kapıyı açık bırakırlar. Çalışanlar sadece birkaç dakika dışarı çıkıyor ama bu kesinlikle bir pencere veya kapıya sızmak için sosyal mühendislere fırsat tanıyor.

2.9. Gelişmiş Sosyal Mühendislik Saldırıları

Sosyal mühendislik saldırılarında saldırgan her zaman telefonla ya da kurumun içine sızarak tek başına bir şeyler yapmaz. Sosyal mühendisimizin ihtiyacına göre kullanacağı bir takım araçlar mevcuttur. Örnek olarak bir kuruma fiziksel olarak sızmış saldırgan yönetici konumdaki personelin bilgisayar ile klavye arasına fark edilemeyecek kadar ufak bir usb keylogger yerleştirebilir. Tabi bunun bir takım riskleri

vardır, kaydedilen şifreleri görmek için tekrar kuruma sızıp usb keyloggerı alması gerekir. Ya da başka bir örnek vermek gerekirse saldırgan kurumda çalışan üst düzey personellerden birine ortam dinleyicisi görevi gören bir mouse ya da kamera görevi gören şık ve pahalı bir kalem hediye edebilir. Bunun gibi akla gelmeyecek kadar yaratıcı birçok araç piyasada mevcut ve düşük bir ücret karşılığında herkesin sahip olabileceği şeyler mevcuttur. Özellikle büyük ölçekli şirketler ve kurumların bu gibi araçlara karşı ciddi güvenlik önlemleri alması gerekmektedir.

Bir sistemi ihlal etmenin en kolay yolu, sorumlu olan birinden izin istemek. Ne kadar şifreleme ve yazılım güvenliği önlemleri uygulanmış olursa olsun, bir ağ hiçbir zaman teknik veya teknik olmayan saldırılara karşı tamamen güvenli değildir. En zayıf halka olan insan bileşeni, tamamen güvenli sistemler söz konusu olduğunda asla göz ardı edilemez. Çalışanların bilmeden veya amacına göre istediği kişilere sistemlerine erişim sağlamaya istekli olmaları durumunda, kaç tane sanal özel ağ (VPN), güvenlik duvarı, antivirüs, zararlı yazılım önleme yazılımı veya şifreleme cihazlarının çalıştığı önemli değildir. Sosyal mühendis, şu anda bu en zayıf bağlantılardan bilgi almak için aldatma, ikna etme veya taktikleri ve teknikleri kullanan kişidir. “Her dakika doğan bir emici var” gerçeği sosyal mühendisler için birçok güvenli sisteme sızma fırsatı veriyor. Doğru yaklaşıma sahip doğru teknik doğru kişiye uygulanırsa bu organizasyona karşı yüksek olasılıkla başarılı bir saldırı anlamına gelir. Sosyal mühendis olmak insan psikolojisini anlamak ve manipüle etmek ve birisine hassas, kritik bilgiler vermek veya saldırgana yetkisiz erişim vermek için birini kullanmanın metodik bir yoludur. Başka bir deyişle, bu iyi bir yalancı olmakla ilgili değildir. Bu insanları çok fazla zaman kaybetmeden ve yüksek risk almadan güvenli bir sisteme daha kolay girebilmeleri için avantajlarından etkilemenin yollarını keşfeden bir mühendis olmakla ilgilidir.

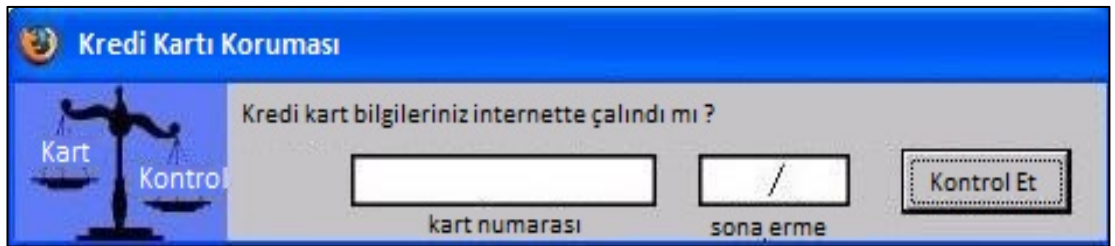
2.9.1. Kimlik Avı (Phishing)

Kimlik avı, bir kullanıcıdan gelen kişisel ve hassas bilgileri çalmak için sahte bir eylemdir. Siber saldırı olarak kimlik avı, e-posta, anlık mesajlaşma uygulamaları, SMS ve Sosyal Ağ uygulamaları gibi internet ve mobil teknolojiler aracılığıyla gerçekleştirilir. Bu teknikte sosyal mühendis, PayPal Inc., Ebay Inc., şirket BT personeli veya bankacılık sitesi gibi resmi bir siteden gelmiş gibi görünen bir kişiye

bir web sitesini ziyaret etmesini ve kullanıcı adı, şifre gibi bu türden hassas bilgileri girmesini isteyen e-posta gönderir. Web sitesi resmi web sitesi gibi görünüyor, ancak bunun yerine saldırgan tarafından aynı görünümlü bir site oluşturulur. Saldırganın PayPal Inc. çalışanını taklit ettiği gerçek bir kimlik avı e-postasından bir örnek: “Tüm hileli işlemlerin yüzde 98’inin, çalınan kredi kartlarını kullanan üyelerin var olmayan ürünleri satın almak veya satmaktan kaynaklandığını fark edilmektedir. Bu nedenle hesabımızı koruma ve web sitemizdeki dolandırıcılık örneğini azaltma taahhüdümüzün bir parçası olarak üyelerimizin fatura kayıtlarına bir banka / çek kartı eklemelerini istiyoruz. Banka / çek kartınız yalnızca sizi tanımlamak için kullanılacaktır. Lütfen çevrimiçi deneyiminizden 5-10 dakika sürebilir ve kayıtlarınızı yenileyebilirsiniz, PayPal hizmetiyle ilgili ileride herhangi bir sorun yaşamayacaksınız. Ancak kayıtlarınızı onaylamamanız hesabınızın askıya alınmasına neden olur.” [14]

Bu e-posta, e-posta alıcısının kredi kartı bilgilerini girmesi için sahte ancak aynı görünümlü bir web sitesine bağlantı sağlamak gibi devam eder. PayPal Inc. çoğunlukla, dünyanın her yerinde bir kredi kartı havuzu ve herhangi bir ödeme yaparken en çok kullanılan site olduğu için tercih edilir.

Bu kimlik avı saldırıları, yalnızca kullanıcıyı başka bir web sitesini yönlendirmekle kalmaz, aynı zamanda bazı kötü amaçlı kodlar veya programlar göndermeyi de içerir. Bu tür altında görülen en yaygın saldırı türü “.exe” dosyaları göndermek ve kullanıcıların bilgisayarlarında çalışmasını sağlamaktır.



Şekil 2.4: Kredi kartı oltalama.

Şekil 2.4’de, saldırganın kredi kartı bilgilerinin çalınıp çalınmadığına dair insanlara yardım eden iyi niyetli bir adam olduğu açıkça görülmektedir. Burada saldırgan kurbanların dikkatini çekmek istediğinden “çalındı” kelimesini yazdı. “Kredi kart bilgileriniz çalındı mı?” gibi bir soru sorarak insanları meraklandırmaya çalışıyor.

Buna ek olarak saldırgan tüm insanların anlayabileceği şekilde pencereyi çok basit bir şekilde tasarlar. Ayrıca adaleti ve yasayı temsil etmek için kullanılan “ölçek” i kullanır. Tüm bu semboller, kelimeler dikkatsiz insanları oltaya düşürmek için özenle seçilmiş ve bu insanların kredi kartı numaralarını ve son kullanma tarih bilgilerini ele geçirmiş olur. Bu bilgilerle, saldırganın amazon.com'dan kolayca birçok şey satın alabileceği bir kredi kartı numaraları havuzu olacaktır.



Şekil 2.5: Kredi kartı oltalama yorumları.

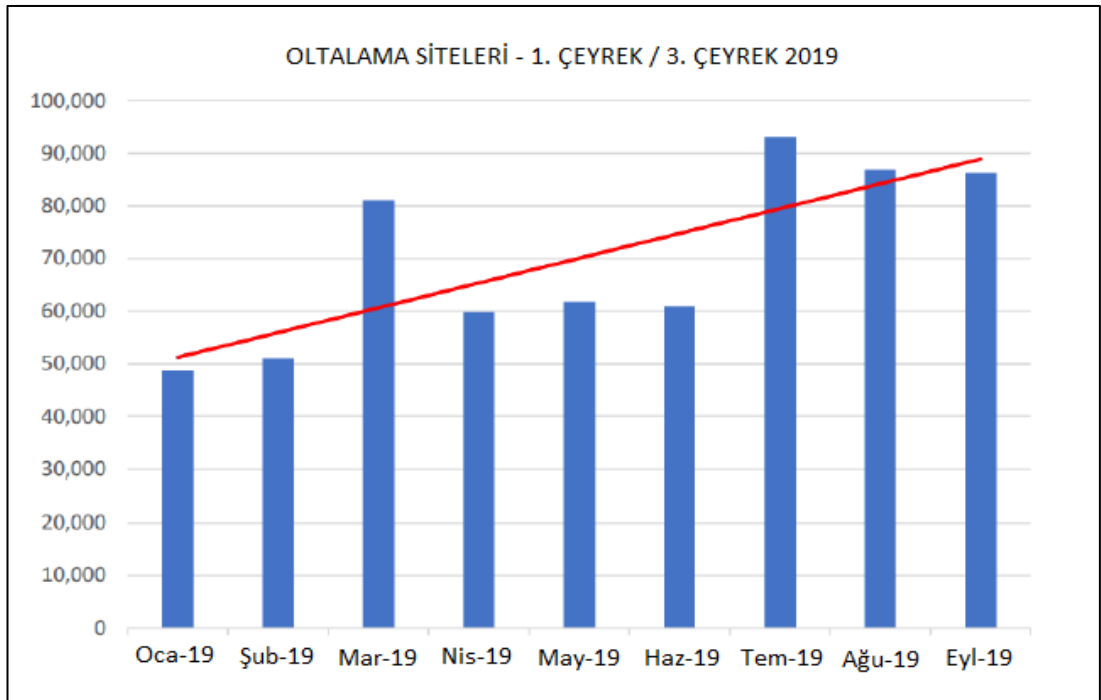
Bu saldırı çok basit görünebilir, ancak birçok kullanıcı bu ekranı posta kutularında gördükten sonra kredi kartı bilgilerini veriyor. Şekil 2.5'te kullanıcılardan kredi kartlarını verdikten sonra bazı yorumlar var. Şekil 2.5'te 135 yorumun sadece 3'ü görülebilir [14]. Yorumlardan da görülebileceği gibi, sosyal güvenlik numaraları gibi kontrol etmek için daha fazla bilgi vermek isteyen bazı insanlar var. Bu insanlar "kontrol et" butonu hakkında hiçbir şey bilmiyorlardı aslında bu buton kredi kartı bilgilerini saldırganın veritabanına kaydediyor.

Kimlik avı, bir saldırı olarak başarılı olmak için hem sosyal mühendislik hem de teknik alt bölüm kullanır. Sosyal mühendislik, şüphelenmeyen kullanıcıyı, aldıkları mesajın (e-posta, SMS vb.) veya ziyaret ettikleri web sitesinin güvenilir ve meşru bir kuruluş / kaynaktan olduğuna inanmaya kandırmaya çalışır. Diğer yandan, teknik alt bölüm teknikleri, kişisel bilgileri doğrudan çalmak ve kullanıcının bir fidye yazılımı girişimi ile erişmesini engellemek için doğrudan kurbanların bilgisayarlarına kötü amaçlı yazılım yüklemeyi içerir.

Daha önce de belirtildiği gibi bir kimlik avı saldırısının amacı, şüpheli olmayan kullanıcının finansal kazanç için bankacılık uygulamalarından alınan şifreler ve kimlik

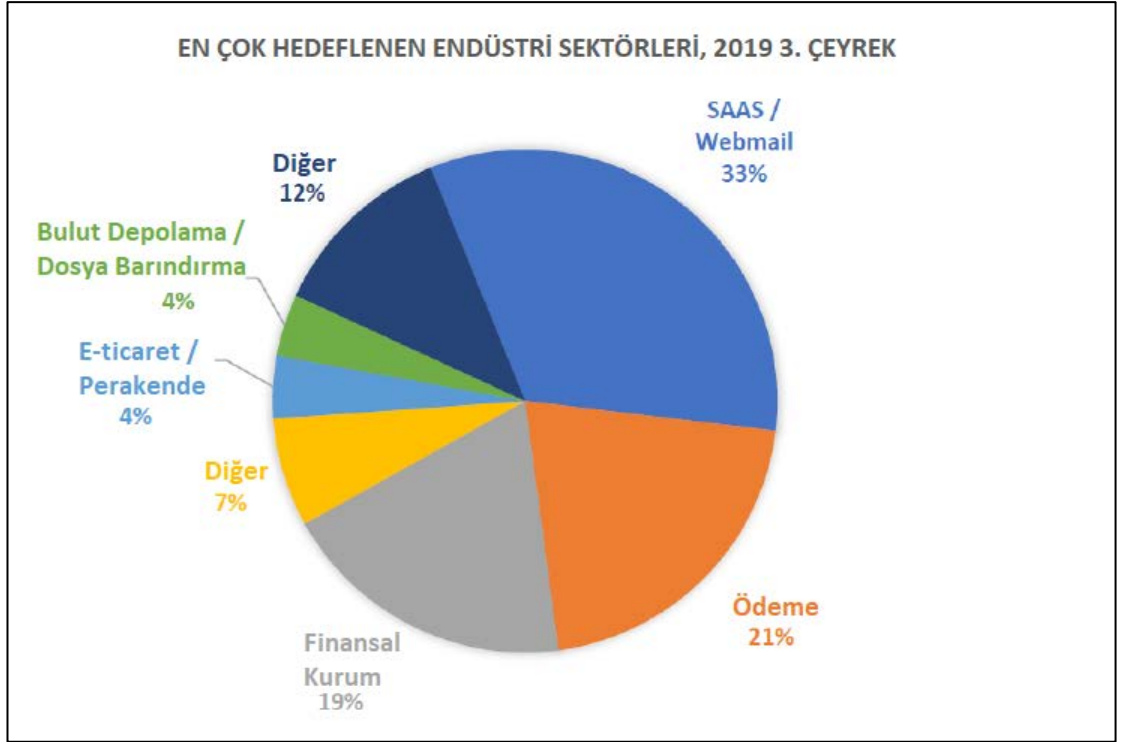
bilgileri veya saldırganın para istediği bir tür fidye yazılımı için sosyal uygulamalardan kimlik bilgileri gibi kişisel bilgileri çalmaktır. Başarılı bir kimlik avı saldırısının getirebileceği tehlikeler, yasadışı faaliyetleri yürütmek için kurbanın kimliğini taklit etmek ve belirli hassas bilgileri çalmayı veya belirli bir sisteme girmeyi başaran suçlular olarak şöhret ve haysiyet kazanmaktır.

Kimlik Avı Koruması Çalışma Grubu'na (APWG) [31] ve Phishing Etkinliği Eğilimleri ile ilgili en son raporuna göre, APWG tarafından 2019'un üçüncü çeyreğinde tespit edilen kimlik avı sitesi sayısı 266.387 idi. Bu, ikinci çeyrekte görülen 182.465'ten yüzde 46 artış gösterdi ve 2018'in 4. çeyreğinde görülen 138.328'in neredeyse iki katı oldu. APWG'ye göre alınan benzersiz kimlik avı raporlarının toplam sayısı, 2005 yılında 173.063 ve 2018'de 1.040.654 idi. Her biri çok sayıda gerçek kimlik avı saldırısı ile sonuçlanır.



Şekil 2.6: 2019 için APWG kimlik avı etkinliği eğilimleri.

2019'un üçüncü çeyreğinde en çok hedeflenen sanayi sektörlerinin aşağıdaki Şekil 2.7'de ki gibi olduğu rapor edilmektedir.



Şekil 2.7: Kimlik avı saldırılarına göre en çok hedeflenen sektörler.

2.9.1.1. Kimlik Avı Saldırı Teknikleri

Farklı kimlik avı saldırı türleri şunlardır [3].

- Hedef Odaklı Kimlik Avı

Geleneksel kimlik avı saldırısı, maskelenmiş mesajı onlara göndererek mümkün olduğunca çok hedefe saldırmak mantığına sahiptir. Hedef odaklı kimlik avı ise, daha meşru görünümlü bir mesaj oluşturmak için bir veya birkaç hedef hakkında mümkün olduğunca fazla bilgi edinmeye çalışır. Üst ve üst yönetim pozisyonlarını hedefleyen bir hedef odaklı kimlik avı alt kategorisidir.

- Sesli Kimlik Avı

Saldırgan telefondaki kurbanı arar ve yasal bir kuruluştan taklit eder. Daha sonra bu bilgileri çalmak ve daha sonra kendi yararına kullanmak amacıyla kullanıcının doğru olduğunu “doğrulamak” için kişisel bilgiler talep ederler.

- SMS Kimlik Avı

Bu kimlik avı türü, Kısa Mesaj Servisi (SMS) kullanılarak yürütülür. E-posta kimlik avı ilkelerini izleyen saldırgan, güvenilir bir kuruluştan görünüşte meşru görünen bir SMS gönderir ve kurbandan belirli bilgiler vermesini ister. Son on yılda akıllı telefonların eklenmesiyle, bu kimlik avı SMS mesajları kötü amaçlı web sitelerine yönlendiren URL'ler de içerebilir.

- E-Posta/Spam

En yaygın kullanılan yöntemlerden biri bir bankacılık uygulamasına kimlik bilgileri gibi kişisel bilgi isteği ile birçok kullanıcıya kimlik avı e-postası göndermektir. Klonlama kimlik avı, daha önce güvenilir bir kaynaktan gönderilen bir e-postanın klonlandığı ve içine kötü amaçlı URL'lerin veya eklerin eklendiği ve yeniden gönderildiği e-posta kimlik avının bir alt kategorisidir.

2.9.1.2. Kimlik Avı Saldırısının Yaşam Döngüsü

Kimlik avı saldırısının yaşam döngüsü aşağıdaki adımlardan oluşur:

Planlama ve Kurulum: Bu adım kuruluşun özelliklerini ve saldırganın başarılı bir saldırı gerçekleştirmesine yardımcı olabilecek her şeyi belirleyerek saldırganın hedef veya hedefler hakkında bilgi toplamasını içerir.

İnşa etmek: Mağdurları ziyaret ettikleri web sitesinin meşru olduğuna inandırmak için saldırgan istismar edilecek meşru web sitesinin bir kopyası olan bir web sitesi oluşturur. Tek fark hassas bilgileri girerken işlevselliktir, örneğin mağdur sisteme giriş yapmak için kimlik bilgilerini girmeye çalıştığında meşru web sitesindeymiş gibi zanneder fakat hesap bilgilerini kimlik avı web sitesi üzerinden saldırıya gönderilir. Bu saldırı kimlik avı web sitesinin mümkün olduğunca çalmak için konum, gizli sorular vb. gibi ek bilgiler istemesini de içerebilir.

Yayma: Saldırgan daha sonra kimlik avı dağıtma yöntemini (e-posta, SMS, anlık mesajlaşma vb.) seçer ve kötü amaçlı bağlantılar içeren kimlik avı iletilerini toplu olarak göndermeye başlar.

Kurulum: Sahte bağlantı kimlik avı web sitesine yönlendirir. Bağlantıyı tıkladığınızda ister akıllı telefon ister kişisel bilgisayar olsun mağdurun sistemine kötü amaçlı yazılım yükleme denemesi de olabilir.

Veri Toplama: Saldırgan kötü niyetli web sitesinden ve yüklü herhangi bir yazılımdan giren kurbanlardan hassas bilgiler alır ve bulabildiği kadar hassas bilgiyi otomatik olarak göndermeye çalışır.

Ayrılma: Son adım saldırganın tüm kanıtları kaldırmasıdır. Bu kimlik avı web sitesinin sunucusunu kapatmayı, sahte bağlantı göndermek için kullanılan herhangi bir mesajlaşma / e-posta hesabını silmeyi ve hatta saldırının düzeyine bağlı olarak daha aşırı önlemlere ulaşmayı içerir.

2.9.2. Kimlik Avı E-postası

Kimlik avı e-postası (adından da anlaşılacağı gibi) gizli bir kaynaktan geliyormuş gibi görünecek şekilde gizlenmiş bir e-postadır. Örneğin, bir bankadan veya çevrimiçi bir mağazadan e-posta olarak gizlenebilir. Bu tür e-postalar alıcıdan hassas bilgiler elde etmek amacıyla kötü niyetli kullanıcılar tarafından oluşturulur. Bunun çalışma şekli kimlik avı e-postasına en basit haliyle potansiyel kurbandan banka hesabı kimlik bilgileri, kredi kartı gibi bir tür kişisel / hassas bilgi girmesini isteyecek kötü amaçlı bir URL ekleyerek web sitesine yönlendirebilir. Kimlik avı e-postasının kolay ve temsili bir örneği Şekil 2.8’ de gösterilmektedir.

Sayın web posta kullanıcısı,

Veritabanımızı ve e-posta hesap merkezimizi güncelliyoruz. Kullanılmayan tüm E-posta hesaplarını siliyor ve yeni hesaplar için daha fazla alan oluşturmuyoruz. Bu süre zarfında hizmet kesintisi yaşamamanızı sağlamak için aşağıdaki ayrıntıları sağlamanız gerekir:

ASAGIDAKI WEBMAIL HESABINIZI ONAYLAYIN

1. E-posta: @gtu.edu.tr

2. Kullanıcı Adı:

2. Parola:

3. Parolayı onaylayın:

Yeni ve geliştirilmiş web posta işlevselliğimiz ve geliştirmelerimizle size daha iyi hizmet verebilmemiz için lütfen bu iletiyi yanıtlayın.

Bunun size mal olabileceği rahatsızlıktan dolayı özür dileriz.

Saygılarımızla,

Webmail Yöneticisi ©2020

Şekil 2.8: Kimlik avı e-posta örneği.

Kimlik avı e-postası olup olmadığını tespit etmesine yardımcı olmak için kullanılabilecek birkaç özellik vardır. Bu özelliklerden bazıları şunlardır:

- Genel Karşılama

Kimlik avı e-posta örneğinin görüntüsünde görülebileceği gibi genel bir karşılama vardır, yani alıcının belirli bir adı yoktur. Genellikle değişken adın eklenmediği gibi bir "düz" selamlama vardır ve başka bir olasılık bir isim olarak alıcının adresinin yerel kısmının (e-posta işaret adının önündeki kısmı) olmasıdır.

- Adresi talep edilen gönderenden farklı yanıtlama

Güvenilen ve yasal bir kaynağın kimliğine bürünmek amacıyla saldırgan gönderenin adını taklit eder ve güvenilir kuruluşun adına değiştirir ve gerçek gönderen adresini gizler [33].

- Sahte Bağlantı

E-postada sağlanan bağlantılarda güvenli ve iyi bilinen bir adrese yol açmış gibi görünseler bile gösterilen metin bağlantı tıklandığında ziyaret edilecek gerçek adresten farklı olabilir. Bu e-postaya HTML öğeleri ekleyerek halledilir.

- Alan Adları

Genel olarak bir e-posta gerçekten banka gibi güvenilir bir kaynaktan geliyorsa e-postadaki URL'lerde herhangi bir alan adı varsa bunlar bankanın alan adından olmalıdır. Aksi takdirde e-posta şüpheli olur.

- URL Sayısı

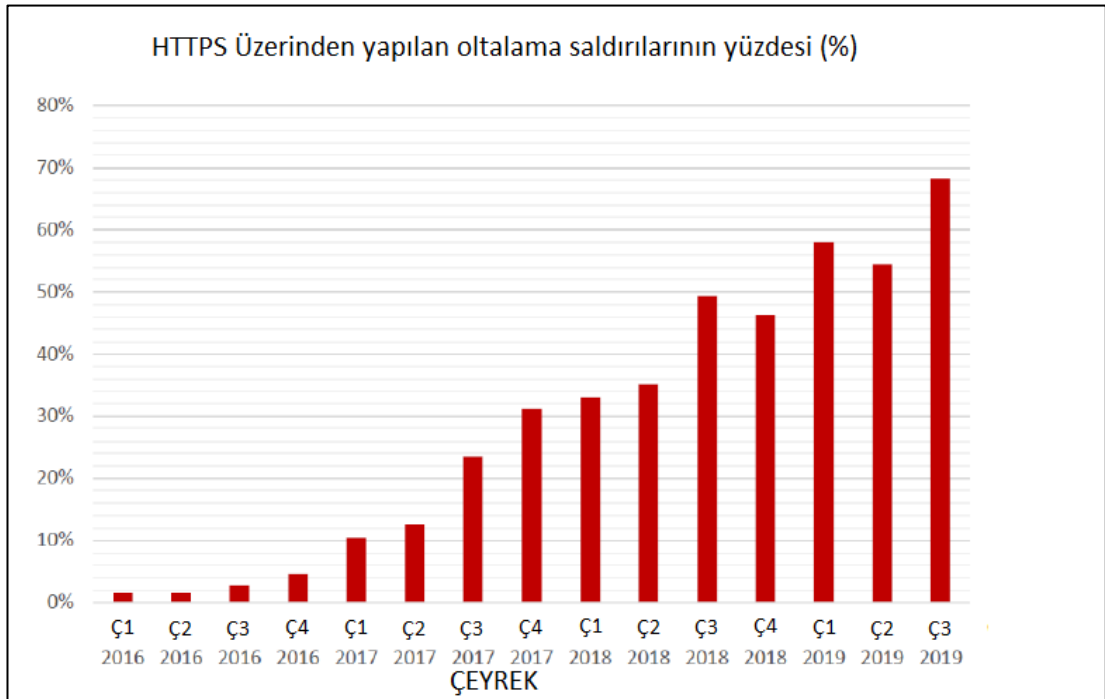
Güvenilen kaynaklar genellikle e-postalarına bağlantılar eklemekten kaçınır, çünkü önceki madde işaretlerinde gösterildiği gibi bağlantılar adresleri ziyaret etmek için çok güvenilir bir yol değildir. Bir web sitesini ziyaret etme ihtiyacı varsa, bağlantıyı sağlamak yerine kullanıcıdan web sitesini kendi başına ziyaret etmesini istemek tercih edilir.

- URL Yeniden Yönlendirme Hizmetlerini Kullanma

Bağlantıya tıklayarak ziyaret edilecek gerçek web sitesini gizlemek için saldırgan TinyURL [34] gibi hizmetleri kullanarak daha küçük bir URL'ye sıkıştırabilir. Bu şekilde ziyaret edilen web sitesinin alan adı yalnızca ziyaret ettikten sonra görülür ve bu da kullanıcının meşruiyet açısından kontrol etmemesine yol açabilir.

- SSL Sertifikaları Güvenli Değil

APWG'nin kimlik avı eğilimleri hakkındaki en son raporuna [31] ve katılımcı PhishLabs'a göre, 2019'un üçüncü çeyreğinde tüm kimlik avı sitelerinin yüzde 68'i SSL sertifikaları kullanıyordu. Bu bir önceki çeyrekte yüzde 54'ten yükseldi. Bu kullanıcıların bir web sitesinin güvenli olduğunu göstermek için yalnızca SSL sertifikalarına güvenemeyeceğinin açık bir göstergesidir.



Şekil 2.9: HTTPS'de barındırılan kimlik avı saldırıları.

- @ Sembolünü Kullanma

Bir web sitesini ziyaret ettiğinizde ve URL'de @ simgesi varsa ziyaret edilen gerçek web sitesi @ simgesinden sonraki sitedir. Bu şekilde saldırganın @ simgesinden önce güvenilir bir web sitesi olabilir ve ardından kimlik avı web sitesini ekleyebilir.

- Onaltılık Karakter Kodlarını Kullanma

Bir URL yazarken en yaygın kullanılan normal karakterler olan ASCII karakterleri kullanılabilir veya onaltılık gösterimleri kullanılabilir. Her karakter için eşdeğer bir onaltılık sayı vardır. Bu durum saldırgan tarafından kullanılabilir ve şüpheli olmayan kullanıcı tarafından kolayca okunamayan karakterler içerebilir.

- Bilgi İsteme

Meşru bir e-posta hiçbir zaman kişisel bilgi istememelidir. E-postada böyle bir deneme varsa bu bir kimlik avı girişimi olduğunun açık bir göstergesidir.

- Aciliyet Duygusu

Genellikle kötü niyetli kullanıcıların popüler web siteleri tarafından ("kara listeye alınan" olarak da bilinir) kaldırılmadan veya kötü amaçlı olarak işaretlenmeden önce mümkün olduğunca fazla bilgi edinmeleri istenir. Bu nedenle, kurbanı ikna etmek için psikolojik bir faktör eklemenin yanı sıra kimlik avı e-postalarının metninde aciliyet hissine neden olur. Örneğin, e-posta bilgileri sağlamak için sınırlı bir zaman aralığı olduğunu veya mağdurun hesabının feshi gibi sonuçların olabileceğini söyleyebilir.

Yukarıda belirtilenler, bir kullanıcı her e-postadan geçerse bir e-postayı kimlik avı veya yasal olarak tespit etmek için kullanılabilecek bir kimlik avı e-postasının birkaç özelliğidir. Kimlik avı e-postasının içeriği veya saldırganın kurbanı meşru bir kaynak olduğuna ikna etme girişimi söz konusu olduğunda farklı kimlik avı e-postaları türleri vardır [35].

- Hükümet Manevrası

Bu kimlik avı e-postasında saldırgan, kurbanı korkutmak ve herhangi bir bilgi vermeye ikna etmek amacıyla federal bir organı taklit etmeye çalışmaktadır.

Arkadaş Taktiği: Saldırgan, bu durumda kurbanın bir arkadaşını taklit eder. Saldırganın büyük olasılıkla kurbanın kimliğine bürünmek için herhangi bir ilgili bilgisi yoktur (hedef odaklı kimlik avı saldırısı olmadığı sürece), genel bir arkadaş kullanmaya çalışır ve kurbanın böyle bir e-posta gönderecek bir arkadaşına sahip olma şansına güvenir.

- Faturalama Sorunu

Bu kimlik avı e-postası yakın tarihli bir çevrimiçi satın alma işleminin farklı nedenlerle (kredi kartının kullanım süresi dolmuş, yanlış fatura adresi vb.) tamamlanamayacağını ve sorunun çözülmesi için sağlanan web sitesini ziyaret etmenin gerekli olduğunu belirtir. Sağlanan bağlantı bir kimlik avı web sitesine yönlendirir.

- Sona Erme Tarihi

Bu tür kimlik avı e-postasıyla, mağdurun belirli bir hizmete veya şirkete ait hesabının süresinin dolmak üzere olduğuna dair bir bildirim vardır. Ancak son kullanma tarihi e-postada sağlanan kimlik avı web sitesini ziyaret edip kimlik bilgilerini sağlayarak genişletilebilir.

- Virüs veya Tehlikeli Hesap ile Korkutmak

Bu durumda saldırgan bilgisayarının veya telefonunun tehlikeye atıldığını veya hesabının ihlal edildiğini belirterek kurbanı korkutmaya çalışır. Sorunun çözülebilmesi için kimlik avı web sitesini ziyaret etmesi veya kötü amaçlı bir ek indirmesi istenir.

- Yarışma Kazanma

Bu kurbanı bir ödül kazandığını düşünmeye ikna etme girişimidir. Ödülü almak için bir web sitesini ziyaret etmeli ve daha sonra onlara karşı kullanılan konum gibi belirli bilgileri girmelidirler.

- Banka Dostu

Bankalar belirli miktarda para çekildiğinde hesap bildirimleri sunar. Bu kimlik avı e-postası bir bankayı taklit eder ve para çekme sınırının aşıldığını ve doğrulama amacıyla belirli bilgilerin girilmesi gerektiğini belirtirler.

Mağdur: Bu durumda saldırgan kurban olduğu söylenen şirkete belirli bir miktar para iade eden bir müşteriyi taklit eder. Cevap yoksa yetkilileri dahil etme tehdidi söz konusudur.

2.9.3. Kimlik Avı Önlemleri

Aşağıdaki bölümde, her türlü karşı önlemin açıklaması sunulmaktadır [36].

- Kullanıcı Eğitimi

Hem siber güvenlik uzmanları hem de siber suçlular birbirlerinden bir adım önde olma çabası içindeler. Bu her iki tarafın da onu kullanmaya ve güvence altına almaya ya da sömürmeye hazır görünen her yeni teknoloji ve güvenlik açığı ile sürekli olarak eğitilmesine neden olur. Siber güvenlik uzmanları siber güvenlik sektöründe meydana gelen herhangi bir değişikliğin çabucak farkında olsalar da normal kullanıcı temel hedefidir. Ve normal kullanıcı sonunda bir kimlik avı saldırısının kurbanı olacak kullanıcıdır. Teknik bir kusur nedeniyle meydana gelen güvenlik açıkları ve istismarlar olabilse de insan siber güvenlik zincirinin en zayıf halkasıdır.

Kullanıcıların kimlik avı girişimlerini tanıyabilmesi için ilgili eğitimin olması gerekir. Halihazırda yerleşik eğitim programları ve hizmetleri bulunmaktadır [37,38]. Kullanıcı eğitimi bir karşı önlem olarak çok etkilidir çünkü problemin tabanına (saldırıların başladığı insan faktörü) güçlendirme sağlanmaktadır.

- Makine Öğrenmesi Teknikleri

Makine öğrenimi, yeterli sayıda örnek girerek bir sınıflandırıcıyı eğitme işlemidir. Örnekler istenen sınıflandırma ile ilgilidir. Örneğin, bir e-postanın kimlik avı olup olmadığını belirleyebilecek bir kimlik avı e-posta sınıflandırıcısı tasarlanacaksa, örnekler kimlik avı veya yasal olarak el ile sınıflandırılmış e-postalar olacaktır. Sonuç, henüz bir sınıflandırma yapılmayan yeni örnekleri tespit etmenin (sınıflandırmanın) mümkün olduğu eğitilmiş bir sınıflandırıcıdır. Bu tezde önerilen araç, derin öğrenme yöntemlerine dayanmaktadır. İlgili çalışma, yaklaşım ve uygulama bölümlerinde daha fazla ayrıntı verilecektir.

- Arama Motoru Tabanlı Teknikler

Bir web sitesini kimlik avı olarak belirlemenin yolu, belirli anahtar kelimelerin ve bilgi parçalarının (resimler) web sitesinden çıkarılması ve arama motorlarına sorgulanmasıdır. Kuruluşun alan adı, ilk belirli sayıda arama sonucu içinde döndürülürse, web sitesi yasal olarak kabul edilebilir. Ancak belirlenecek web sitesi

bir kimlik avı web sitesiye, alan adı bu ilk sonuçlarda görünmez. Bu yaklaşımın dezavantajı, yeni oluşturulan meşru web sitelerinin arama sonuçlarında yüksek olmama şansının yüksek olması nedeniyle yüksek yanlış pozitif / alarm oranının olmasıdır.

- Liste Tabanlı Kimlik Avı Önleme Çözümleri

Bu tür bir koruma yaklaşımı, kimlik avı web sitesini ziyaret etmekten korumak için kara ve beyaz listeler kullanmıştır. Kara liste, kimlik avı web siteleri olarak tanımlanmış ve onaylanmış web sitelerinin listesidir. Aksine bir beyaz liste ise yasal web siteleri olarak tanımlanan web sitelerinin bir listesidir.

Liste tabanlı yaklaşımın avantajı, kara liste yaklaşımı kullanılırken gerçek pozitif (doğru alarm) yüzde 100 olmasıdır, çünkü kimlik avı olarak belirlenen her web sitesi kara listede olacaktır, yani kimlik avı olarak onaylanmıştır. Bir beyaz liste kullanırken, gerçek negatif oran yüzde 100'dür, çünkü meşru olarak belirlenecek her web sitesi aslında beyaz listeye dahil edilecektir, bu yüzden meşru olarak onaylanmıştır. Öte yandan, bu yöntemlerin dezavantajları vardır. İlk olarak, listelerin bakımı kendi başına bir hizmettir, yani işlem gücü gerektirir. İkincisi, yukarıda belirtilen alarmlar yüzde 100 doğru olsa da kara listeler için yanlış negatif oran ve beyaz listeler için yanlış pozitif oran önemli olacaktır. Bunun nedeni kara listeler için kara listeye henüz eklenmemiş bir kimlik avı web sitesi varsa, yanlış bir şekilde meşru bir kategori olarak sınıflandırılacaktır. Beyaz liste ile, beyaz listede olmayan bir web sitesi gerçekte bir kimlik avı veya yasal bir site olup olmadığına bakılmaksızın kimlik avı olarak kategorize edilir. Uygulamalar ve normal kullanım için kullanma olanağı sağlayan listeye dayalı hizmetler vardır (Google Güvenli Tarama [39], PhishTank [40])

- Görsel Benzerlik Temelli Yaklaşımlar

Bu teknik, bir web sayfasını doğrulanmış kimliği doğrulanmış web sayfalarıyla karşılaştırmak için görsel özellikler kullanır. Görsel özellikler arasında metin içeriği, metin biçimlendirme, CSS varlığı ve kullanımı, web sitesinin ekran görüntüleri ve içerdiği resimler yer alır. Bu yaklaşımın temeli iki sayfanın benzer görünümlere sahip olamayacağı varsayımdır, böylece bir web sitesi bu görsel özelliklere göre karşılaştırıldığında ve başka bir doğrulanmış web sayfasına benzer olduğunda bunun bir girişim olduğu düşünülebilir.

2.10. Makine Öğrenmesi

Kimlik avı e-postalarını algılamaya yönelik önerilen çözüm, sınıflandırma işlemine yardımcı olmak için Makine Öğrenimi teknolojilerini kullanır. Bu bölümde makine öğrenimi temellerinin, makine öğrenme tekniklerinin türleri ve literatürde bulunan en çok kullanılan makine öğrenimi sınıflandırıcılarının bir açıklaması olacaktır. Makine öğrenmesi yapay zeka çalışmalarının bir alt kümesi olarak kabul edilir. Makine öğrenimi, belirli komutlar olmadan belirli görevleri tamamlamak için değil, verilen örneklerden gelen kalıpları algılayarak bilgisayarlar tarafından kullanılmak üzere tasarlanmış algoritmalarından oluşur.

Makine öğreniminin son yıllarda başarılı olmasının nedeni şudur. Makine öğrenimi algoritmaları genellikle çok fazla işlem gücü gerektirir, çünkü sadece belirli komutları izlemeye ve yürütmeye değil, çok sayıda örnek üzerinde yineleme yapmaya ve belirli bir deseni algılayabilmek için belirli parametreleri ve ağırlıkları uyarlamaya dayanırlar. Nihai sonuç, ağırlıkları minimum bir hata ile "genelleme" ve yeni örnekleri (henüz sınıflandırılmamış) tanımlayabilmeleri için ayarlanmış bir model / sınıflandırıcı olmalıdır. Bilgisayarların işlem gücü önemli ölçüde artmış ve tüm süreci daha uygulanabilir hale getirmiştir [41].

Bir programlama ve problem çözme teknolojisi olarak makine öğrenimi aşağıdaki nedenlerden dolayı faydalıdır:

- Programlama

Geleneksel bir algoritma ve program tasarlarken, geliştiricinin meydana gelebilecek her olası durumu dikkate alması ve bunları çözmek için bir parça kod sağlaması gerekir. Bununla birlikte, bir makine öğrenimi algoritması kullanıldığında tek temel gereksinim, istenen örneklerle veri kümesini toplamak veya bulmak, oluşturmak ve bunları eğitmek için sınıflandırıcıya karşı kullanmaktır. Tek ek gereklilik, makine öğrenme algoritmasının sahip olabileceği ek hiperparametrelerin ayarlanmasıdır.

- Nihai Ürünlerin Kişiselleştirilmesi ve Ölçeklendirilmesi

Makine öğrenme algoritmaları belirli örneklerle eğitildiği için belirli bir tür örnek için tasarlanmış olan gelişmiş bir tasarım sadece makine öğrenme algoritmasını

sıfırdan kullanarak başka bir örnek seti için kolayca ayarlanabilir. Örneğin, bir program dilbilgisi veya metin tabanlı işlevsellik için yapılmışsa, o dilde yeterli veri örneği alınarak anında başka bir dile taşınabilir.

- **Daha Önce Çözölemeyen Sorunların Çözümü**

Makine öğreniminin eklenmesi ile yüz tanıma ve ses tanıma gibi sorunlara yeni bir yön verilmiştir.

2.10.1. Makine Öğrenmesi Algoritmalarının Türleri

Günümüzde üretim sanayi ve elektronik sektörünün hızla gelişmesi nedeniyle eski yıllara göre sorunlar ve hesaplama problemleri katlanarak büyümüştür. Bu büyümeler aynı zamanda ortaya çok fazla verinin çıkmasına neden olmuştur. Süregelen bu süreçte çok miktarda verinin işlenmesi hatta eski veri yığınlarının içerisinde gelecek bir durum veya olay tahmini yapılması mümkün hale gelmiştir. Makine öğrenmesi bu büyük veri yığınlarının içerisinde belirli problemleri çözmek veya belirli durumlar için tahminlerde bulunabilmek için ortaya çıkmıştır. Derin öğrenme ise bir makine öğrenmesi yöntemidir, belirleyici farkı ise katmanlar halinde çalışması ve bu katmanların birbirleri arasında bilgi aktarımı olmasıdır. Yani herhangi bir katman kendinden önceki katmanlardan bilgiler alarak kendi bilgisini revize edebilir. Verilen bir veri kümesiyle çıktıları tahmin edebilmek için farklı öğrenme yöntemleri kullanılabilir. Farklı öğrenme türleri takip eden satırlarda verilmiştir.

2.10.1.1. Denetimli Öğrenme

Denetimli bir makine öğrenme algoritması, önceden sınıflandırılmış örneklerden istenen modeli eğitir ve öğrenir. Algoritma, bir hatayı hesaplamak ve parametreleri uyarlamak için örneği giriş olarak ve her örneğin önceden belirlenmiş çıktısını kılavuz olarak kullanır. Böyle bir sınıflandırıcının performansını değerlendirmenin yaygın bir yolu (daha sonraki bir bölümde performans değerlendirme yöntemleri hakkında daha fazla bilgi verilecektir), sınıflandırılmış veri setini iki bölüme ayırmaktır: sınıflandırıcıyı eğitmek için bir kısım (eğitim seti) ve bir performansını değerlendirme

kısmı (test testi). İlk olarak sınıflandırıcı eğitim seti kullanılarak eğitilir ve daha sonra sınıflandırıcı önceden belirlenmiş sınıflandırmalarıyla ilgili bilgileri kullanmadan test seti örneklerini sınıflandırmaya çalışır. Bir sonraki adım, performans değerlendirme amacıyla sınıflandırıcının tahminlerini gerçek belirlenmiş sınıflandırmalarla karşılaştırmaktır.

Denetimli öğrenme algoritmaları sınıflandırma ve regresyon algoritmalarına ayrılır. Sınıflandırma algoritmaları söz konusu olduğunda, bunların amacı her örnek için önceden belirlenmiş bir kümeden bir değer seçmektir. Örneğin, bir e-postayı kimlik avı veya meşru olarak sınıflandırma durumunda, sınıflandırma algoritması yasal bir e-posta tahmini için sıfır ve kimlik avı e-posta tahmini için bir değer verir. Regresyon algoritmaları, çıktılar bir aralık içinde sayısal bir değere sahip olduğunda kullanılır.

2.10.1.2. Denetimsiz Öğrenme

Denetimsiz algoritmalar, sınıflandırılmamış bir dizi girdi örneği alır ve bunları bir araya getirerek veya veri noktalarının komşularına kümeleyerek bir model bulmaya çalışır. Bu algoritmaların amacı, sağlanan verilere ilişkin bilgi sağlamak için örneğin herhangi bir kümede gruplandırılmayan aykırı değerleri vurgulayarak veya belirli eğilimleri bağlı olarak göstererek, işlemin başında bilinmeyen bilgi veya modelleri bulmaktır.

2.10.2. Tanımlar

Bir tasarım veya bir modelin algılanması için kullanılacak bir makine öğrenme modeli seçerken, en iyi sınıflandırıcıyı tasarlamak ve üretmek için seçilecek ve ince ayar yapacak çok sayıda parametre ve hiperparametre vardır. Ayrıca, “en iyi” sınıflandırıcı terimi birçok farklı yoruma ve hesaplama yoluna sahiptir. Bu bölümde makine öğrenimi performans değerlendirmesinin temelleri ve bazı gerekli tanımlar sunulmaktadır.

Bu tez yoluyla çözülmeye çalışılan durum, e-postaların kimlik avı veya yasal olarak sınıflandırılmasıdır. Bu bir sınıflandırma problemidir ve denetimli öğrenmedir.

Bu yeni sınıflandırılmamış e-postaları sınıflandırabilmek için bir modeli eğitmek için kullanılacak önceden sınıflandırılmış bir veri kümesi olacağı anlamına gelir.

Denetimli makine öğreniminde çıktıları tahmin etmek ve bir modelin eğitilmesi için girişler sağlanır. Farklı modelleri ve makine öğrenmesinin nasıl çalıştığını daha ayrıntılı bir şekilde açıklayabilmek için ilgili birkaç tanım sunulmaktadır.

- Etiket, tahmin edilecek y değişkenidir. E-postalar söz konusu olduğunda, etiket “is_Phishing” olarak adlandırılabilen bir değişkendir ve e-posta kimlik avı yapıyorsa değeri 1 ve yasal ise 0’dır.
- Örnek, verilerin belirli bir örneğidir. Tezin amaçları için örnek, veri kümesindeki bir e-postadır. Örnekler, daha önce bir kaynak tarafından sınıflandırılmış olup olmadıklarına bağlı olarak etiketlenebilir veya etiketsiz olabilir (çoğu zaman elle sınıflandırılır).
- Özellikleri, makine öğrenme modelinin toplam özellik sayısı olan x_i , $i = 1, 2, \dots, n$ giriş değişkenleridir. Özellikler, sınıflandırılacak her örnekten alınan e-postaların özellikleridir. Her özellik, belirli bir bilgi parçasını kendi yolunda sağlar ve toplamda e-postanın tüm resmini sağlar, bu da algılanacak örüntüyü vurgulayacaktır. E-postalar söz konusu olduğunda özellikler e-postanın özelliklerini (e-postadaki URL sayısı, eklerin sayısı, e-postanın HTML içeriğine sahip olup olmadığı vb.) ve metinsel özellikleri (belirli kelimelerin sıklığı, cümlelerin benzerliği) içerebilir. Alınan özelliklerin ayrıntılı bir sunumu analiz bölümünde yer almaktadır. Bir model, makine öğrenme sürecinden çıkan sınıflandırıcıdır ve her örneği y tahminiyle eşleştirir.

Daha iyi anlamak için bu noktada Doğrusal regresyon modeli örnek olarak kullanılacaktır. Bu modelde, n özellik x_i , bir özellik alma algoritması ile her örnekten alınır, her örnek için sağlanan etiketler / çıkışlar vardır ve temel model şu formüle sahiptir.

$$y' = b + w_1x_1 + w_2x_2 + \dots + w_nx_n \quad (2.1)$$

Burada b yarıllıktır ve w_i , $i = 1, 2, \dots, n$ her bir ilgili özellik için ağırlıktır. Denetimli öğrenmede algoritma, her örnek geçtikten sonra önyargı ve ağırlıkların değerlerini

değiştirerek bir model oluşturur, böylece belirli bir kayıp fonksiyonu kullanılarak bir hata en aza indirilir.

Yaygın olarak kullanılan bir kayıp fonksiyonu, her tahminin kare hatalarının toplamının ortalaması olan Ortalama Kare Hatası'dır (MSE).

$$MSE = \frac{1}{n} \sum (y - \hat{y})^2 \quad (2.2)$$

Aşırı öğrenme (overfitting), eğitim verilerini kullanarak hatanın son derece düşük olması için modeli eğitme işlemidir, ancak sonuçta ortaya çıkan model genellemede kötüdür, yani eğitim ve test verilerindeki hata çok düşük olmasına rağmen model iyi değildir. Eğitimin doğru yolu, eğitim verisine iyi uyum sağlamak (kayıp fonksiyonundan nispeten küçük bir değere sahip olmak), ancak aynı zamanda basit bir şekilde öğrenmektir, böylece elde edilen model “genelleme” yapabilir, yani algılanması istenen deseni algılar.

- Test ve Eğitim Setleri

Modeli eğitmek ve performansını doğru bir şekilde değerlendirmek için önceden sınıflandırılmış mevcut veri setinin iki sete ayrılması gerekir: bir test seti ve bir eğitim seti. Bu iki kümenin birbirinden bağımsız olması ve orijinal veri kümesinden rastgele seçilmesi gerekir. Bu, modelin herhangi bir önyargı olmadan eğitilmesini ve performans değerlendirmesinin uygun şekilde yürütülmesini sağlayacaktır.

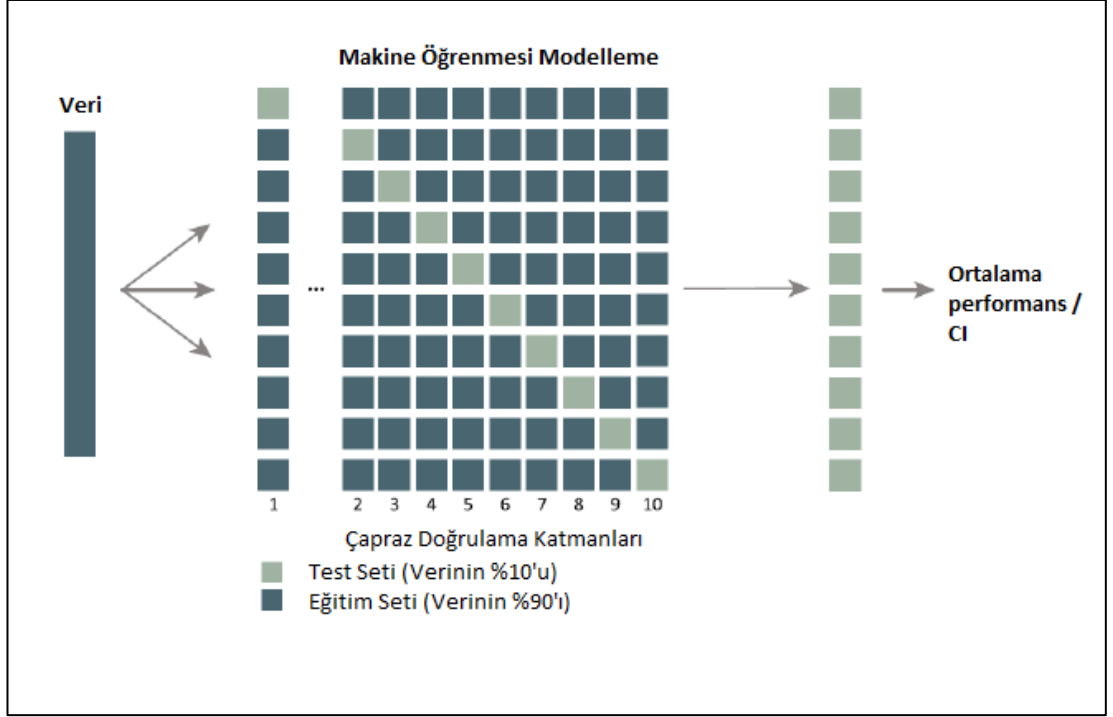
- Çapraz Doğrulama

Test ve eğitim setlerinin rastgele örnek seçimi olması nedeniyle ortaya çıkabilecek herhangi bir rastgele yanlılığı gidermek amacıyla bir eğitim ve performans değerlendirme tekniğidir. İşlem aşağıdaki Şekil 2.10’da gösterildiği gibidir [42]:

- Veri kümesi n eşit, bağımsız ve rastgele alt kümeye bölünmüştür.
- Alt kümelerden birini test seti olarak seçin ve diğer alt kümeleri tek bir eğitim setinde birleştirilir.
- Sınıflandırıcıyı eğitin, tahminler yapın ve performansı değerlendirin.
- Her bir sonraki altküme için bir eğitim seti olarak ayarlayıp diğer altkümeleri bir eğitim setiyle birleştirerek işlemi tekrarlayın.

- Sonunda, veri kümesinin her alt kümesi için performans sonuçları vardır. Toplam performans sonuçları, her alt küme performansının ortalama değerlerinden oluşur.

Bu test alt kümelerinin herhangi birinde herhangi bir yanlışlık varsa, diğer alt kümelerin performans sonuçlarıyla ortalama değerleri hesapladıktan sonra en aza indirgenmesini sağlayacaktır.



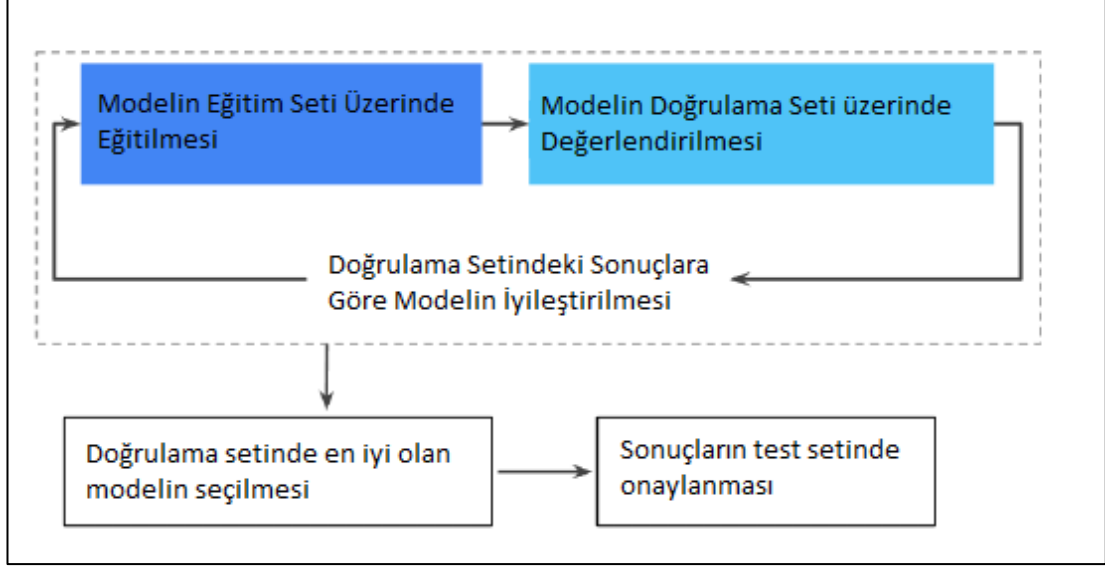
Şekil 2.10: Çapraz doğrulama.

- Doğrulama Seti

Doğrulama seti, makine öğrenme modelinin hiperparametreleri optimize edilirken kullanılır. Orijinal veri kümesinden alınan ek bir kümedir, yani kullanılırsa orijinal veri kümesi 3 alt kümeye bölünecektir: test, eğitim ve doğrulama. Ek doğrulama seti kullanılırken süreç Şekil 2.11’de ki gibidir:

Eğitim veri seti modeli eğitmek için kullanılır. Doğrulama veri seti modelin performansını değerlendirmek ve bu bilgiyi eğitim sürecinin ve modelin kendisinin hiperparametrelerini ayarlamak için kullanılır. Tüm ayarlar tamamlandıktan sonra test veri seti, makine öğrenme modelinin performansını son bir kez değerlendirmek için kullanılır. Ek veri setinin amacı, doğrulama setinde istisnai olarak iyi performans

göstermesi için her parametreyi çok fazla ayarlarken ortaya çıkabilecek aşırı öğrenme olasılığını azaltmaktır, ancak model test seti kullanılarak değerlendirildiğinde, gerçekten aşırı öğrenme meydana gelir[41].



Şekil 2.11: Eğitim, validasyon ve test setlerini kullanarak eğitim süreci.

- Özellik Mühendisliği

Özellikleri alma ve bir veri kaynağından hangi özellikleri alacağımızı seçme işlemidir (ham veri e-postalarından). Veri kaynağından alınmak üzere seçilen özelliğin türüne bağlı olarak bazı işlemlerden sonra yapılması gerekir. Sayısal değerler olduğu gibi aktarılabilir. Sınıflandırıcının bunları doğru bir şekilde yorumlayabilmesi için dizelerin kodlanması gerekir. Dizeleri kodlamak için sık kullanılan bir algoritmaya One-hot encoding denir. Bu algoritmayı kullanarak, özelliğin alabileceği her farklı dize değeri, bir boole vektöründe dizin olarak temsil edilir. Vektör, her dizinde sıfır değerine ve yalnızca belirli örnekte bulunan dize değerini temsil eden dizinde bir değere sahiptir. Bir özellik seçmeye gelince sınıflandırıcıya deseni daha net bir şekilde vurgulayacak bir dizi özellik sağlamak için izlenmesi gereken bir dizi iyi yöntem vardır. İyi bir özellik şu özelliklere sahip olmalıdır:

Çoğu örneğin bir değeri olmalıdır. Örnekten bir özelliğin alınamadığı durumlar vardır; bunlardan biri veri kümesinin biraz bozuk olması nedeniyle çeşitli nedenlerle ortaya çıkabilir. Örneğin, bir e-posta veri kümesinde, alıcıların adreslerinin kaldırıldığı bir e-posta olabilir. Bu durumda, bir özellik alınamaz. Eksik değerlerin doldurulabileceği

yöntemler vardır. Sayısal tipte bir özellik söz konusu olduğunda, eksik değer diğer örneklerin değerlerinin ortalaması, minimum, maksimum vb. ile değiştirilebilir. Ancak, çoğu örnekten alınamayan ve sonuçta ortaya çıkan özellikler yapay değerlerle değiştirildiğinde çok az bilgi sağlanır ve örüntü tespiti için yararlı değildir. Anlamlı bir değeri olmalı. Belirli bir modeli tespit etmek için özellikler seçerken, gerçekten ilgili bilgileri sağlayan özelliklerin seçilmesi gereken kritik bir düşünme süreci olmalıdır. Örneğin, bir e-postayı kimlik avı veya yasal olarak sınıflandırmak için özellikler seçerken, e-postada kaç URL'nin bulunduğu veya URL'lerden kaçının alan adı veya düz IP olduğu gibi bir özellik sınıflandırıcı için değerli bilgiler sağladığından, bir IP URL'si genellikle ilgili web sayfasının aceleyle oluşturulduğu ve bunun bir kimlik avı web sitesine yönlendirilme olasılığı olduğu anlamına gelir. Önceki yöntemin zihniyetini takiben, benzersiz tanımlayıcılar (ID'ler) hiçbir zaman özellik olarak kullanılmamalıdır, çünkü belirli örnekler için veritabanında arama yaparken yararlı olmalarına rağmen, her örnek için farklıdır ve örüntü tanıma açısından yardımcı olur. Özelliklerin belirli bir aralıkta farklı bir anlamı olan ve başka bir belirli aralıkta başka bir anlamı olan değerleri olmamalıdır. Örneğin, bir evin piyasada bulunduğu yılları temsil eden bir özellik olması gerekir. Bir ev hiç pazar olmamışsa, bu gerçeği göstermek için örneğin -1 değeri vermek iyi bir uygulama değildir. Evin piyasada olup olmadığını temsil edecek bir boolean değere sahip ek bir özellik ayarlamak iyi bir uygulamadır. Özelliklerin temsil edilen değerleri zaman içinde değişmemelidir. Örneğin, bir özellik bir zaman değerini açıklıyorsa ve kaynak modelde belirli değişiklikler varsa, sürenin çıktığı birimi değiştirmeyin. Eğer kaynak model bu zaman değerini saniye cinsinden vermek için kullanılırsa, örneğin dakika olarak değişmemelidir, çünkü bu zamanın temsil edilme şeklini değiştirir. Özellikler aşırı uç değerlere sahip olmamalıdır. Çünkü aşırı değerlere sahip örnekler, kayıp fonksiyonunun aşırı bir çıkış değerine sahip olmasına neden olacaktır. Bu değer, parametrelerde aşırı bir ayarlama yapılmasına ve muhtemelen neredeyse eğitilmiş modeli bozmasına neden olacaktır.

2.10.3. Makine Öğrenmesi Modelleri

Bu bölümde, bazı avantaj ve dezavantajları olan temel makine öğrenme modellerinin bir açıklaması olacaktır.

2.10.3.1. Lojistik Regresyon

Regresyon, üretilen modeli sınırlı bir değer kümesinden değil, sürekli bir değer uzayından değerleri tahmin etmeyi amaçlayan makine öğrenme algoritması türüdür. Bununla birlikte regresyon modelleri sürekli çıkış değeri kullanılarak ve çeşitli eşikler ayarlanarak sınıflandırma için de kullanılabilir. Eşik değerlere ve sürekli çıkışa bağlı olarak ayarlanan sınıflandırma değerlerinden bir değer nihai çıktı olarak ayarlanır. Örneğin, kimlik avı e-postaları söz konusu olduğunda bir eşik belirlenir ve sürekli çıktı eşik değerden yüksekse e-posta kimlik avı olarak sınıflandırılır ve aksi halde meşru olur.

Daha önce açıklandığı gibi doğrusal regresyon doğrusal bir fonksiyonu mümkün olduğu kadar deseni temsil edecek ve hataları en aza indirecek şekilde uyarlamaya çalışır. Ancak doğrusal regresyon sadece sayısal değerlerle sınırlıdır. Lojistik regresyon kategorik verileri kabul edebilen ve logit fonksiyonunu kullanarak meydana gelme olasılığını tahmin edebilen doğrusal olmayan bir regresyon türüdür. Lojistik regresyon, değişkenler / özellikler arasındaki doğrusal ilişkiyi dikkate almaz. Yanlılık ve özellik ağırlıkları hatayı en aza indirecek şekilde ayarlanır.

Lojistik regresyonun avantajları [47]:

- Küçük veri kümeleriyle iyi performans
- Çıktısı olasılık olarak yorumlanabilir
- Lojistik regresyonun dezavantajları şunlardır:
- Veri varsayımlarına uyulması gerekir.
- Sadece doğrusal çözümler sağlayabilir.

2.10.3.2. K-En Yakın Komşu

K-en yakın komşu parametrik olmayan, örnek tabanlı bir öğrenme yöntemidir [48]. Örnekler, her biri ilgili etiket değeriyle (örneğin yasal e-posta için 0 ve kimlik avı e-postası için 1) birleştirilmiş bir özellik alanında puan olarak depolanır. Yeni bir örnek sınıflandırılacaksa, özellik değerleri onu özellik alanına yerleştirmek için kullanılır. Daha sonra en yakın komşuları dikkate alınır ve sınıflandırma değeri çoğunluk olan etiket değerine eşittir. Örneğin, $k = 5$ olduğunu varsayalım. En yakın 5 komşusunun 3

meşru e-posta ve 2 kimlik avı e-postasından oluştuğu özellik alanına yeni bir e-posta eklenir. Yeni e-posta yasal bir e-posta olarak sınıflandırılacaktır [49].

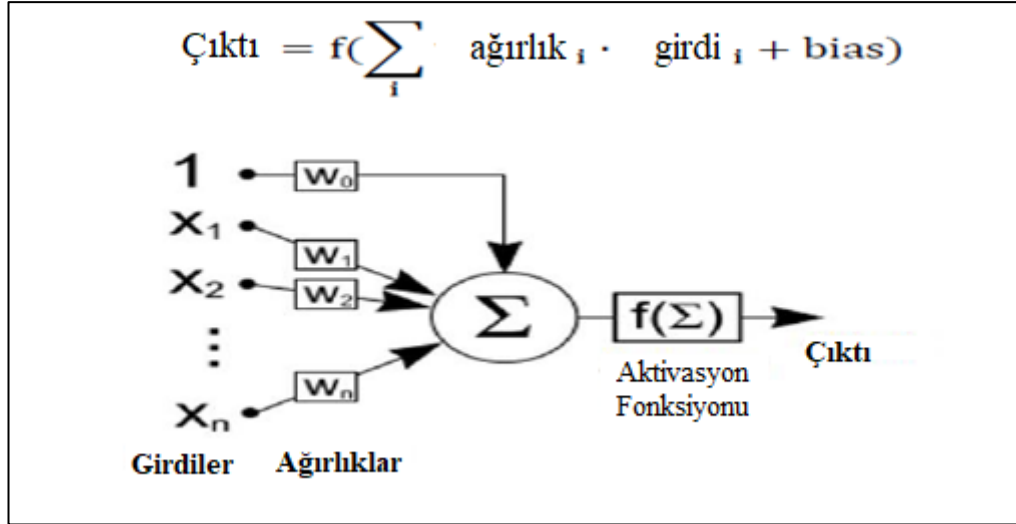
K seçimi en yakın komşu algoritmanın performansını etkiler. K değeri çok küçükse ve gürültü varsa belirli bölgelerde gürültü çoğunluğu olabilir ve yanlış tahminlere neden olabilir. Öte yandan k değeri çok büyükse diğer etiket değerlerinden gelen komşular olabilir ve bu da yine yanlış sınıflandırmalarla sonuçlanır.

En yakın k komşu algoritmasının avantajları:

- Anlaşılması ve uygulanması kolay bir algoritmadır.
- Bir numunenin birçok sınıf etiketine sahip olduğu uygulamalarda iyi performans gösterir.
- Bununla birlikte bu algoritmanın birkaç dezavantajı vardır:
- Eğitim süresi kısa olsa da bunun bir dezavantajı vardır. Yeni örneklerin sınıflandırılması özellikle de k değeri yüksekse genellikle hesaplama açısından pahalıdır, çünkü yeni örnek ile her komşu arasındaki mesafenin hesaplanması gerekmektedir.
- Eğitim sırasındaki her örnek saklandığı için büyük depolama gereksinimlerine sahiptir.
- k için en uygun değeri seçmenin belirli bir yolu yoktur. Tek yol deneme yanılma yöntemidir.

2.10.3.3. Yapay Sinir Ağları

Yapay bir sinir ağı, her bağlantının bir ağırlığa ve her düğümün bir sapmaya sahip olduğu bir dizi bağlı giriş / çıkış birimidir. Yapay Sinir Ağları, girişlerin / özelliklerin YSA'ya beslendiği bir giriş katmanından sınıflandırma işlemine yardımcı olan bir veya daha fazla gizli katmandan ve istenen sınıflandırma değerini çıkaran bir çıkış katmanından oluşur [50].



Şekil 2.12: Yapay sinir ağı modeli.

F fonksiyonuna aktivasyon fonksiyonu denir. Uygulamaya bağlı olarak farklı etkinleştirme işlevleri vardır. Aktivasyon fonksiyonları çıkış aralığı 0 ile 1 arasındadır. En basit aktivasyon fonksiyonu adım fonksiyonudur [51]. Daha tutarlı ve doğru sonuçlar elde etmek için daha karmaşık işlevler kullanılır. Yaygın olarak kullanılan aktivasyon fonksiyonu sigmoid fonksiyonudur [52].

YSA yetiştirilmenin popüler bir yoluna Geri Yayılım denir. Eğitim süreci aşağıdaki gibidir:

- Ağırlıkları ve önyargıları başlatılır.
- Eğitim örneğini YSA'ya beslenir.
- İstenen çıktı ile çıktı ve hatasını hesaplanır.
- Her nöron için çıktının ne olması gerektiğini ve aralarındaki farkı hesaplanır.
- Hata, ağırlığı yansıtacak şekilde ağırlık ve sapmalar güncellenerek geriye doğru yayılır.

Belirli bir hata eşiğine veya belirtilen sayıda ağı ulaşılana kadar işlem tekrarlanır.

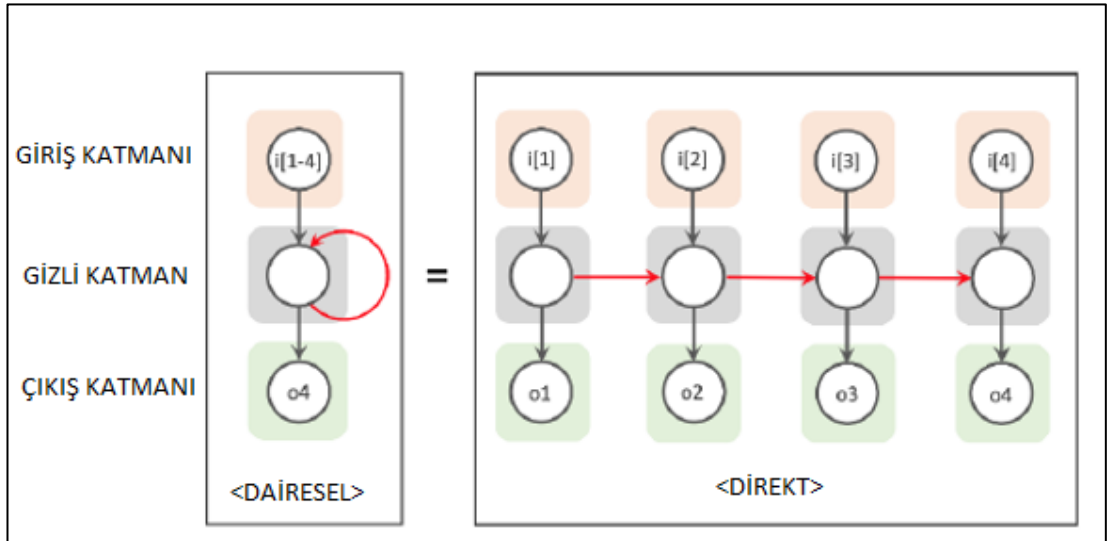
Yapay sinir ağlarının avantajları şunları içerir:

- Gürültülü verileri tolere edebilir.
- Özellikler ve etiket arasında net bir ilişki olmadığında kullanılabilir.
- İşlemi hızlandırmak için paralelleştirme teknikleri kullanılabilir.

- El yazısı metin tanıma, görüntü ve ses tanıma gibi zorlu görevlerde başarılı.
- YSA'ların dezavantajları şunları içerir:
- Uzun öğrenme süresi
- Yorumlanması zor (ağırlıkların ve yanlılıkların belirli bir şeyi temsil etmediği karmaşık ağ)

2.10.3.4. Tekrarlayan Sinir Ağı

Tekrarlayan Sinir Ağları geleneksel yapay sinir ağlarının hafıza yeteneklerini geliştiren bir tür ANN (yapay sinir ağı)'dır [53].



Şekil 2.13: Tekrarlayan sinir ağı.

RNN' lerin ek özelliği, gizli katmandaki geri besleme döngüsüdür. Bu sıralı veriler hakkında ek bilgi sağlar. Örneğin, sinir ağındaki girdi bir kelime cümlesiyse ve her bir girdi bir kelimeyse, gizli katmandaki geri besleme döngüsü ile RNN düğümlerinin ağırlıklarını eğitmek ve güncellemek için tüm cümleyi dikkate alabilir.

2.10.3.5. Karar Ağaçları

Karar ağacı, ağaç yapısı algoritmalarını kullanarak verileri ayrık olanlara ayırır [48]. Bu makine algoritmalarının ardındaki fikir, verilerin / özelliklerin arkasındaki

yapıyı vurgulamaktır. Eğitim seti özyinelemeli olarak orijinalden daha saf olan altkümelere bölünür, yani altkümeler mümkün olduğunca eşit bir şekilde bölünür. Üretilen karar ağacının her bir düğümü bir özelliği temsil eder. Her dal, bu özelliğin belirli bir değer aralığını temsil eder. Veri kümesinin bölünmesi için optimal bir durum (neredeyse hiç gerçekleşmeyen) etiketle 1-1 ilişkisi olan bir özellik olacaktır, yani bu özellik ile her örneği sınıflandırmak mümkün olacaktır. Bununla birlikte bu durum hiçbir zaman belirgin değildir ve aynı zamanda makine öğrenmesine pratik bir ihtiyaç olmadığı anlamına da gelir, çünkü sadece if-else örneğidir. Yaprak düğümleri bir sınıflandırma sonucunu temsil eder. Bir sınıflandırma algoritması olarak karar ağacı, ilgili özellik değerlerine bağlı olarak her örneği ağaçtan alarak ve örneği yaprak düğümü değerine göre sınıflandırarak çalışır.

Verileri bölme söz konusu olduğunda, kök düğümü ve iç düğümlerde, verileri bölmek için kullanılmak üzere seçilen özellik eğitim verilerini en iyi bölen özelliktir. Bilgi kazancı, Kazanç oranı, Gini endeksi, Ki kare, C-SEP, G-istatistikleri ve Minimum açıklama uzunluğu (MDL) gibi hangi özelliğin bunu yaptığını bulmak için farklı metrikler vardır.

- Ön Budama

Karar ağacı, belirli bir eşik değerine ulaşıldığında gelişimini erken durdurarak budanır.

- Budama Sonrası

Karar ağacı, ağaç tamamlandıktan sonra doğrulukla yardımcı olmayan bazı alt ağaçları kaldırarak veya ağacı hafif doğrulukta kazanmak için çok fazla bölerek budanır.

Karar Ağaçlarının avantajları şunlardır:

- Karar ağaçları sınıflandırmada basit ve hızlıdır.
- Yüksek boyutlu verileri işleyebilirler.
- İnşa edilmiş karar ağacı kolayca anlaşılabilir.

- Farklı değerlerin özelliklerini işleyebilirler, çünkü ayırma her özelliğin sahip olduğu değere değil, her özelliğin eğitim kümesini ne kadar iyi böldüğüne bağlı olarak gerçekleşir.
- Karar Ağaçlarının dezavantajları şunlardır:
- Yüksek eğitim süresi, çünkü ağacın her bir düğüm dallaması için hangi özelliğin seçileceğine karar vermek için her örneğin kullanılması gerekir.
- Büyük veritabanlarıyla uğraşırken kullanılabilir bellek eksikliği.

2.10.3.6. Destekli Vektör Makineleri

SVM' ler istatistiksel öğrenme teorisi ve yapısal risk minimizasyonu prensibine dayanır ve karar sınırlarının yerini belirlemeyi amaçlar. Temel fikir bir destek vektör sınıflandırıcısının mümkün olan maksimum mesafe ile farklı sınıfları ayıran düz bir çizgi bulmaya çalışmasıdır. Doğrusal olarak ayrılabilir verilerin en basit durumunda, verileri düzgün bir şekilde ayırmak ve son sınıflandırıcıyı belirlemek için sadece bir düz çizgi gerekir. Verilerin tek bir düz çizgiyle ayrılamaması durumunda SVM, yanlış sınıflandırma hatalarının sayısını en aza indirirken kenar boşluğunu en üst düzeye çıkaran hiper düzlemi bulur.

SVM'lerin avantajları:

- Diğer sınıflandırıcılara kıyasla sağlam ve doğru yöntem
- Teorik temel, küçük eğitim veri seti gerektirir, yüksek boyutlarda iyi çalışır.

SVM'lerin dezavantajları şunları içerir:

- Yüksek hesaplama gücü ve işlem süresi gerektirir.
- Uzun eğitim süresi.
- Bellek gereksinimleri, eğitim örneği sayısının karesine analogdur.
- Sonuçların yorumlanması zordur.

2.10.3.7. Naif Bayes

Naive bayes istatistiksel sınıflandırıcıdır. Bir örneğin belirli bir sınıfa ait olma olasılığını hesaplayarak sınıflandırır. Naive Bayes sınıflandırıcısı tüm değişkenlerin sınıflandırmaya katkıda bulunduğunu ve karşılıklı olarak ilişkili olduğunu varsayar. Bayes teoremine dayanır; burada H, belirli X verilerinin C sınıfına ait olduğu hipoteziyse,

$$P(H | X) = \frac{P(X | H)P(H)}{P(X)} \quad (2.3)$$

$P(H | X)$, X doğruysa H hipotezinin olasılığıdır, $P(H)$ H'den X'ten bağımsız H olasılığıdır ve $P(X | H)$ H doğruysa X olasılığıdır.

Algoritma her sınıfın olasılığını hesaplar. Hipotez ve örneği en yüksek olasılığa sahip sınıfla sınıflandırır.

Naive Bayes sınıflandırıcısının avantajları şunlardır [48]:

- Düşük eğitim ve sınıflandırma süresi, kolay inşa edilebilir.
- Her örnek üzerinden geçmesi gereken herhangi bir yinelemeli parametre içermez, bu nedenle büyük veri kümelerine karşı iyi çalışır.

Naive Bayes sınıflandırıcısının dezavantajları şunlardır:

- Genellikle diğer sınıflandırıcılardan daha az doğrudur.

2.10.3.8. Topluluk Sınıflandırması

Topluluk Sınıflandırması, birden fazla sınıflandırıcı kullanır, çıktılarını alır ve nihai tahmini oluşturmak için bunları farklı şekillerde birleştirir.

2.10.3.9. Rastgele Orman

Rastgele orman, her ağacın bağımsız olarak örneklenen rastgele bir vektörün değerlerine bağlı olduğu birçok karar ağacı sınıflandırıcısını birleştiren bir sınıflandırıcıdır. Eğitim sırasında, daha sonra sınıf tahmini için kullanılan bir dizi karar ağacı oluşturulur. Bu karar ağaçlarının oylanan sınıfları dikkate alınarak elde edilir ve en yüksek oyu alan sınıf çıktı olarak kabul edilir [54].

2.10.3.10. AdaBoost

Yükseltme, her yeni ağacın orijinal veri setinin değiştirilmiş bir versiyonu üzerinde eğitildiği karar ağaçlarını kullanan bir tür topluluk eğitimidir [55]. AdaBoost ilk karar ağacını her gözlem için bir ağırlık değeri ile eğitir. İlk ağaçta her örnek aynı ağırlık değerine sahiptir. Birinci ağacın değerlendirilmesinden sonra doğru bir şekilde sınıflandırılması zor olan örneklerin ağırlık değerleri artırılmış ve kolayca sınıflandırılan örnek ağırlıkları azaltılmıştır. İşlem tekrarlanır ve sınıflandırma çıktısının önceki oluşturulan alt ağaçların ağırlıklı toplamı olduğu bir topluluk ağacı modeli oluşturulur.

2.10.3.11. Gradyan Arttırılmış Ağaçlar

Gradyan yükseltme makineleri AdaBoost ile aynı şekilde çalışır, tek fark yeni oluşturulan ağaçların adaptasyonunun nasıl yapıldığıdır. GBM, farklı örnekler için ağırlık değerlerini uyarlamak yerine, kayıp fonksiyonundaki azalan gradyanları kullanır. Kayıp fonksiyonu, daha az kontrol sağlayan genel kontrol edilemeyen bir kayıp fonksiyonu yerine kullanıcı tarafından belirlenen bir maliyet fonksiyonu olabilir.

2.10.3.12. Çoğunluk Oylama

Çoğunlukla oy kullanıldığında, sınıflandırıcılar bağımsız olarak eğitilir ve daha sonra sınıflandırma uzmanlarının çoğunun örnek olduğunu tahmin ettiği sınıfın nihai çıktı olacağını tahmin eden bir oylama süreci yürütülür [57].

2.10.3.13. İstif

Yığınlama, ilk adımın çeşitli farklı basit sınıflandırıcıları eğitmeyi ve daha sonra çıktılarını son tahmini çıkaracak bir son sınıflandırıcıyı eğitmek için özellikler olarak kullanmayı içeren bir meta sınıflandırıcıdır [58].

2.11. İlgili Çalışmalar

Bu bölümde kimlik avı e-postası alanında ilgili çalışmalar sunulacaktır.

2.11.1. Kimlik Avı E-postasının Rastgele Orman Makine Öğrenme Tekniği Kullanılarak Sınıflandırılması

Bu makale [54], e-postaları kimlik avı veya yasal olarak sınıflandırmada rastgele orman tekniği makine öğrenme algoritmasının kullanımını araştırmaktadır. Önerilen sınıflandırma yönteminde, özellikler doğrudan her e-postadan çıkarılır. Ağ üzerinden gönderilen sorgu yok veya büyük miktarda veri depolanıyor, böylece gerekli bellek ve işlem süresi azalıyor. Alınan özellikler e-postanın özelliklerini temel alır ve e-postanın metnini dikkate almaz. 15 özellik belirlenmiştir:

- IP adresi içeren URL'ler
- “href” özelliği ile bağlantı metni arasındaki farklar: HTML kodu yazarken, bağlantı metninin ziyaret edilecek gerçek URL'den farklı olması mümkündür.
- Bağlantı metninde “link”, “click” ve “here” kelimelerinin bulunması
- Alan adındaki nokta sayısı: Sayı üçten fazla olmamalıdır. Eğer öyleyse 1 ikili değeri kaydedilir.
- HTML e-postası: MIME standardı, e-postada bulunan içerik türünü tanımlar. “Text / html” ise, 1 ikili değeri kaydedilir.
- HTML e-postası: MIME standardı, e-postada bulunan içerik türünü tanımlar. “Text / html” ise, 1 ikili değeri kaydedilir.
- Javascript'in varlığı
- Bağlantı sayısı

- Etki alanına bağlı sayısı: URL'lerde bulunan farklı etki alanı adlarının sayısı sayılır.
- Kelime listesi özellikleri: Bu sınıflandırıcının yalnızca metin özelliklerini açıklar.

E-postanın tüm metnini dikkate almaz, bunun yerine e-posta alıcısının olası bir ikna girişimini gösteren belirli kelimelerin varlığını sayar. Farklı alt kategorilere ayrılırlar:

- Güncelle, onayla
- Kullanıcı, müşteri, müşteri
- Askıya al, kısıtla, tut
- Doğrulama, hesap, bildirim
- Giriş, kullanıcı adı, şifre, tıklama, günlük
- Sosyal güvenlik numarası, menkul kıymet gelir.

Rastgele orman algoritmasının karar ağaçları oluşturulmadan önce 15 özelliğin tümünden bilgi kazancı (IG) hesaplanır ve eğitim ve sınıflandırma süreci için en iyi IG'ye sahip sekiz özellik seçilmiş ve %98 başarı elde edilmiştir.

2.11.2. PySpark kullanarak e-posta kimlik avının analizi ve tespiti

Bu makalede [58] Spark aracı ve daha özel olarak Spark'ın python tabanlı bir uygulaması olan PySpark, kimlik avı e-postalarını sınıflandırmak için makine öğrenimi sınıflandırıcılarıyla birlikte kullanılmaktadır.

2.11.3. Kimlik Avı Tespiti için Makine Öğrenmesi Tekniklerinin Karşılaştırılması

Bu çalışmada [59] kimlik avı e-postalarını tespit etmek için çeşitli farklı makine öğrenme algoritmaları karşılaştırılmıştır. Toplam 2889 e-posta kullanıcı %59.5 'i meşru ve geri kalanı kimlik avı e-postalarıydı. Sınıflandırma işlemi için 43 özellik kullanılmıştır.

Kimlik Avı Koruması Çalışma Grubu'na (APWG) göre, kimlik avına karşı üç kategori savunma mekanizması vardır: dedektif, önleyici ve düzeltici.

Kimlik avı araç çubuklarından, kimlik avı saldırılarına karşı ilgili çalışma olarak bahsedilir, ancak sonuçlar tatmin edici değildir (% 50 - % 75 algılama oranları).

Test edilen makine öğrenme algoritmaları şunlardır:

- Lojistik Regreyon
- Sınıflandırma ve Regresyon Ağaçları(karar ağaçlarının kategorisi)
- Rastgele Orman
- Nöral Ağlar
- Vektör Makineleri Desteklemek
- Bayes Katkılı Regresyon Ağaçları

Bu makalede kullanılan özellikler, kimlik avı ve yasal e-postalarda en sık kullanılan terimlerin sıklığını temsil eder. Özellik olarak kelime seçme yöntemi metin madenciliğinde yaygın olarak uygulanır ve “kelime çantası” yöntemi olarak adlandırılır. Bu makaledeki özellikler kelimelerdir. Her e-postanın özelliklerinin alınması için geçtiği işlem şudur:

- E-posta ekleri kaldırılır.
- Başlık bilgileri, HTML etiketleri ve öğeleri ayıklanır
- Durdurma sözcükleri (ben, sen, orada, içeri, açık, vb.) metinden kaldırılır.
- Metine budama işleme uygulanır (örneklerin / soneklerin kaldırılması, örn. “koşmak” kelimesi “koş” olur).
- En sık kullanılan terimler TF-IDF (terim frekansı - ters belge sıklığı) ile tespit edilir ve özellikler bunlardan oluşturulur. TF-IDF, bir örnekte sıkça görünen ve tüm verilerde sık olmayan kelimelere daha yüksek ağırlık vererek en yüksek lojistik regresyon ile %95.11 başarı elde etmiştir.

2.11.4. Çoğunluk Oylamasını Kullanarak Spam E-postalarını Tespit Etme Yaklaşımı

Bu makale [60] spam e-postalarını tespit etmek için bir topluluk modeli (çoğunluk oylaması) önermektedir. Makalede kimlik avı e-postalarını değil, spam e-postaları

sınıflandırmaya çalışsa da, çoğunluk oylama yönteminin yararlı olduğunu ve bu çalışmanın amaçları açısından yararlı kabul edilmiştir.

Makalede kullanılan sınıflandırma algoritmaları, Saf Bayes, Destek Vektör Makineleri, Rasgele Orman, Karar Gütük ve k-En Yakın Komşulardır. Bir önceki bölümde belirtildiği gibi çoğunluk oylaması önce birkaç basit sınıflandırıcıyı eğiterek ve daha sonra nihai sınıflandırmayı yapmak için çıktılarını oy olarak kullanarak çalışır. Kullanılan örnek veri seti UCI'den [61] alınmıştır ve modelleri oluşturmak ve eğitmek için kullanılan araç RapidMiner'dir [62]. Önerilen model aşağıdaki şekilde gösterilmiştir. Beş iç sınıflandırıcı vardır ve beş kişiden dördü e-postayı spam olarak tahmin ederse, son çıktıda e-posta spam olarak etiketlenir. Başlangıçta birkaç farklı sınıflandırıcı eğitilmiş ve test edilmiştir.

En iyi performans gösteren 5 sınıflandırıcı seçildi (en yüksek doğruluk): Karar Gütük, k-En Yakın Komşular, Rastgele Orman, Destek Vektör Makinesi ve Naif Bayes. Son olarak oylama tekniğine sahip topluluk modeli oluşturuldu, eğitildi ve test edildi. Performans sonuçları, çoğunluk oylamasının bir performans iyileştirmesi olduğunu kanıtladı.

2.11.5. Kimlik Avı E-postalarının Tespiti ve Sınıflandırılması için Makine Öğrenmesi Tekniklerinin Kullanılması

Bu makalede [63], önışleme aşamaları kelime gömme veya vektörleştirme olmak üzere sinir ağı tabanlı bir sınıflandırıcı önerilmiştir. Model altı bölümden oluşur ve beş özellik kullanır. Eğitim süreci için on kat çapraz doğrulama yapıldı. Kullanılan e-posta veri setleri, yasal e-postalar için Spam Assassin'den [64] ve Jose Navario kimlik avı e-posta kuruluşundan [65] alınmıştır.

E-postalardan alınan beş özellik şunlardır:

- URL sayacı
- E-postanın HTML içeriği olup olmadığı
- E-postanın içinde JavaScript kodu olup olmadığı
- E-posta bölümlerinin sayısı(ekler, HTML metni, düz metin vb.)
- Vektör Ortalaması

Vektör ortalaması için: Önce e-postalar, e-postaları temizleyen ve vektörleştirmeye hazır hale getiren E-posta Temizleyici'den geçer. Vektörleştirme veya Kelime Gömme, Doğal Dil İşleme'deki (NLP) sözcük dağarcığındaki sözcüklerin veya ifadelerin gerçek sayıların vektörleriyle eşleştirildiği bir dizi dil modelleme ve özellik öğrenme tekniğidir. Tüm kelimeler vektörleştirildikten sonra, bir mesaj için vektörler toplanır ve ortalama hesaplanır.

E-postalardan özellikler alındıktan sonra elde edilen veri kümesi, 6 sütun (5 özellik ve 1 etiket) ve 14370 satır / e-posta (6656 iyi huylu ve 7714 kimlik avı e-postası) içeren bir csv dosyasıdır. Sinir Ağı'nın uygulanmasında Matlab [66] kullanılmıştır. Verilerin %70' i eğitim, %15'i doğrulama ve %15'i test için kullanılmıştır.

2.11.6. Makine Öğrenmesine Dayalı Kimlik Avı Web Sitesi Algılamasında En Etkili Özellik Kategorisini Belirleme

Bu çalışmada [67], çeşitli farklı özellikler web sitelerinden alınmıştır ve doğruluk açısından test edilmiştir (bu özellik grubu en iyi doğrulukla sonuçlanır). Özellikler makine öğrenimi tabanlı kimlik avı web sitesi algılamasıyla ilgili çalışmalardan seçilmektedir. Bunlar 2006 ve 2016 yılları arasında yayınlanan 18 araştırma makalesine dayanmaktadır. Hangi kategorinin en iyi sonuçları verdiğini belirlemek için özellikler beş farklı kategoriye ayrılmıştır. Bu özellikler şunlardır:

- İçerik URL'sinin Gizlenmesi ile İlgili Özellikler
 - HTML kodundaki harici ve kaynak URL'lerin yüzdesini bakılır.
 - Favicon'un farklı bir alan adından yüklenip yüklenmediğini kontrol edilir.
 - HTML, eylemleri farklı bir etki alanı adına yol açar.
 - Boş değerler, otomatik doğrudan değer, geçerli bir web sayfasının URL'si veya bazı anormal değerler içeren köprülerin yüzdesine bakılır.
 - Durum çubuğunda sahte bir URL görüntülemek için onMouseOver ile JavaScript kodu olup olmadığı kontrol edilir.
 - Herhangi bir HTML formunda anormal eylem olup olmadığını kontrol edilir.

- Alan Adı Gizleme ile İlgili Özellikler

- URL'nin ana bilgisayar adı bölümündeki “-“ karakter sayısı
- URL bir IP adresidir.
- TLD'nin (Üst düzey alan adı) veya ccTLD'nin (ülke kodu TLD) alt alan adının bir parçası olarak kullanılıp kullanılmadığını kontrol edilir.
- TLD veya ccTLD'nin URL'de yol olarak kullanılıp kullanılmadığını kontrol edilir.
- Gizlenmiş HTTPS
- HTML kaynak kodundaki en sık kullanılan alan adının, web sayfasının URL alan adıyla eşleşip eşleşmediğini kontrol edin.

- HTML İçeriği ile İlgili Özellikler

- HTTPS protokolü olmayan URL'ler içeren HTML form işlemleri.
- Form işlemi görelili bir URL içeriyor.
- Form işlemi “#”, “about: blank”, boş dize veya “javascript: true” içerir.
- Sağ tıklama yeteneklerini devre dışı bırakan JavaScript kodu.
- Açılır pencereleri başlatan JavaScript kodu.
- HTML “mailto” işlevini içeren HTML kodu.
- Iframe veya çerçeve kullanımı.
- Başlık etiketinin boş olup olmadığını kontrol edilir.
- Form kapsamının metin değil yalnızca resim içerip içermediğini kontrol edilir.

- Sembol İstismarıyla İlgili Özellikler

- URL'deki nokta sayısı
- URL'deki “-” sayısı
- URL'de “@” sembolünün bulunup bulunmadığını kontrol etme.
- “~” sembolünün URL'de olup olmadığını kontrol etme.
- URL'de “_”, “%”, “&”, “#” ve sayısal karakter sayısı (her biri için farklı özellik).
- URL'de “/” olup olmadığını kontrol edilir.
- URL'deki hassas kelimelerin sayısı (yani “güvenli”, “hesap”, “onayla” vb.).

Sonuçlar, URL gizlemenin en iyi performans gösteren sınıflandırıcıyı oluşturduğunu gösterdi.

2.11.7. Metin Sınıflandırması için Unsur Çıkarma veya Unsur Seçimi: Kimlik Avı E-posta Tespiti Üzerine Bir Vaka Çalışması

Bu makalede [68], sınıflandırma performansı açısından özellik boyutsallık azaltma yönteminin en iyi sonuçları verdiği bir test yapılmıştır. Bu deneylerin tamamlandığı sınıflandırma kimlik avı e-posta sınıflandırmasıydı. Kullanılan sınıflandırıcı, nihai bir tahmin oluşturmak için rastgele oluşturulmuş eğitim setleri tarafından eğitilmiş sınıflandırıcıların bir kombinasyonu olan bir torbalama sınıflandırıcısıdır. Temel sınıflandırıcı olarak J48 karar ağacı algoritması kullanılmıştır.

Boyut azaltma için iki temel kategori vardır: özellik çıkarma ve özellik seçimi.

- **Özellik Çıkarma**

Orijinal özellik alanı daha kompakt yeni bir alana dönüştürülür.

- **Temel Bileşenler Analizi(PCA)**

Amaç orijinal değişkenlerin birleşimi olan ve ilişkisiz yeni değişkenler türetmektir. İlk olarak veri kümesinin ortalaması hesaplanır ve daha sonra orijinal niteliklerin kovaryans matrisi hesaplanır. Sonra özvektörler çıkarılır. Bunlar, orijinal nitelik uzayından niteliklerin ilişkisiz olduğu yeni bir alana doğrusal bir dönüşümdür. En iyi n özvektörler yeni özellikler olarak seçilir ve geri kalanı atılır.

- **Gizli Semantik Analizi(LSA)**

Benzer anlamları olan birden fazla kelime ve birden fazla anlamı olan sözcüklerin üstesinden gelmek için kullanışlıdır.

- **Özellik Seçimi**

Orijinal özelliklerin bir alt kümesi seçilir ve Makine Öğrenme modelini eğitmek için yalnızca bu alt küme kullanılır.

- Ki-Kare

Bu terim ve sınıf arasındaki bağımlılığa bağlı olarak her özellik için bir puan üretir.

- Bilgi Kazancı

Rastgele bir değişken veya başka bir rastgele değişkeni gözlemlemekten sinyal hakkında elde edilen bilgi miktarıdır. Sonuç her özellik için bir puandır ve en yüksek puanlara sahip özellikler seçilir.

E-postaların sınıflandırılmasına yönelik atılan adımlar şunlardır:

- Veri kümesi, kimlik avı veya yasal olarak etiketlenmiş genel e-posta şirketlerinden e-posta toplayarak hazırlanır.
- Jetonlama, e-postanın sözcüklerini bir diziye ayırmak için yapılır.
- Sınıflandırıcının oluşturulmasında önemli bir önemi olmayan kelimeleri kaldırmak için durdurma sözcüğü kaldırma işlemi uygulanır.
- Esnek sonlamayı kelimelerden çıkarmak için köklenme uygulanır.
- Terimler-Belge-Frekansı (TDF) matrisi, satırların farklı örnekleri (e-postalar) temsil ettiği ve sütunların farklı terimleri (kelimeleri) ve karşılık gelen örnekte göründükleri frekansları temsil ettiği şekilde oluşturulur.
- Yukarıda belirtilen olası tekniklerden neredeyse bir vektör olan her bir satıra boyutsallık azalması uygulanır.
- E-postalarda sınıflandırma yapılır.

Modelde tüm e-postanın kendisini dikkate alır, bu nedenle özellik vektörü başlık, URL ve gövde tabanlı özellikler gibi farklı özellikler içerir. Performans sonuçları yaklaşık 1500 özellikte maksimum performans olduğunu göstermektedir.

2.11.8. Kimlik Avı E-posta Filtreleme Teknikleri Araştırması

Bu çalışma [69], kimlik avı e-postalarına karşı filtreleme tekniklerindeki ilgili çalışmaları araştırdı. Farklı teknikler arasında, makine öğrenimi temelli birkaç filtreden bahsedilmiştir. Bergholz ve diğ. [70], dinamik bir Markov zincirinden ve sınıf-konu modellerinden elde edilen özelliklere sahip bir sınıflandırıcı önermiştir. E-postalardan çıkarılan özellikler aşağıdaki kategorilerden alınmıştır:

- Dört yapısal (toplam metin parçası sayısı)
- Sekiz bağlantı özelliği (toplam bağlantı sayısı)
- Dört element özelliği (HTML veya JavaScript kodu)
- İki spam filtresi özelliği (e-postayı yerleşik spam filtresinden geçirme ve çıktılarını özellik olarak alma)
- Dokuz kelime listesi özelliği (kelimeler: hesap, güncelleme, onaylama, doğrulama, güvenli, bildirim, günlük, tıklama, elverişsiz)

İkinci adım, dinamik bir Markov zincirinin, belirli bir sınıfa ait bir mesajı yakalama olasılığına bağlı olan özelliklerdir. Üçüncü olarak, 50 gizli konu modeli özelliği alınır (e-postalarda kelime kümeleri birlikte görünür).

Ma, Ofoghi ve diğ. [71] bilgi kazanımı algoritmasını kullanarak birçok özelliği sıraladıktan sonra sadece yedi özellik kullanan bir model önermişlerdir. Bu özellikler en güçlü özellikleri temsil eder. Modeli oluşturmak için izlenen adımlar şunlardır:

- Yedi özelliği içeren bir özellik üreticisi
- Uyarlanabilir beş makine öğrenme algoritması ile bir makine öğrenme seçim yöntemi
- Bilgi kazanımı tümevarım ile yaratılır
- Daha küçük bir özellik vektör alanı seçen bir özellik değerlendirme yöntemi
- Özellik kümelerini optimize etmek için rafine bir özellik matrisi
- Sınıflandırıcı olarak karar ağaçları algoritması kullanıldı ve ulaşılan doğruluk %99,8'e eşitti.

Castillo ve diğ. [72], FRALEC adlı bir hibrit sistem önermişlerdir. Hibrit bir sistem, sonuçları geliştirmek için farklı algoritmaları bir araya getirmeye dayanan çok katmanlı sistemlerdir. Sistem aşağıdakilerden oluşur:

- 1.E-postaların metin içeriği için bir Naif Bayes sınıflandırıcısı
- 2.E-postaların dil bilgisel olmayan özelliklerinin üç kategoriye ayrılması için kural tabanlı bir sınıflandırıcı: hileli, yasal ve şüpheli
- 3.URL web sitelerinden gelen yanıtları sınıflandıran öykünücü tabanlı bir sınıflandırıcı.

Daha önce yasal olarak sınıflandırılmış bir e-postanın gövdesi formlar içeriyorsa, e-posta kimlik avı olarak atanır. Yasal sınıflandırıcı e-postanın gövdesi formlar, bağlantılar veya resimler içermiyorsa, e-posta yasal olarak atanır. Model meşru ve geri kalan kimlik avı olan 1038 e-postaya karşı test edildi ve %96'lık bir hassasiyet elde edilmiştir. İslam ve ark. [73] çok katmanlı bir sınıflandırma yöntemi önermişlerdir. E-posta özellikleri, uyarlanabilir çok katmanlı bir sınıflandırıcı tarafından sıralı bir şekilde çıkarılır, sınıflandırılır ve çıktılar bir karar sınıflandırıcısına gönderilir. Mesaj sınıflandırıcıların herhangi bir katmanı tarafından yanlış sınıflandırılırsa, üçüncü katman işlemdeki son kararı verecektir. Üç aşamada kullanılan modeller SVM, AdaBoost ve Naive Bayes idi. Model %97' lik bir doğruluk elde etti.

Makine öğrenimi sınıflandırıcılarının arkasındaki teori hakkında kapsamlı bir anlayışa sahip olan ve birkaç ilgili çalışma sunduktan sonra önerilen hibrid sınıflandırıcıyı oluşturmanın bir sonraki adımı kısa ama açık bir problem ifadesi ve çözülmeye çalışıldığı yaklaşımdır.

2.12. E-posta Yapısı

Bu projenin ana odak noktası e-postaları kimlik avı veya iyi huylu olarak sınıflandırmaktır. Bir e-postanın sınıflandırılmasına yönelik ilk adım, e-postaları analiz edecek ve bir makine öğrenimi sınıflandırıcısının minimum hata oranına sahip e-postaları tanımlaması için ihtiyaç duyacağı tüm yararlı bilgileri alacak olan uygun bir ön işleme sürecidir. E-postayı düzgün bir şekilde ön işleme için, tezin bu bölümünde bir e-postanın temel ve gerekli özellikleri sunulacaktır.

Tablo 2.1, e-postanın düz metin olarak basit bir şeklini göstermektedir.

Tablo 2.1: E-posta örneği.

Kimden: Mendez, Nicole
Gönderilen: 30 Mart 2001 Cuma 10:51
Nereye: Rance, Susan
Konu: RE: HATIRLATMA: Lütfen zaman çizelgelerinizi bugün doldurun
Harika. Teşekkürler. Bunu nasıl yapacağınız aşağıda açıklanmıştır:
Araçlar
Seçenekler
Posta Biçimi
Bu mesaj biçiminde gönder: HTML
(Kırtasiye ve Yazı Tipleri Altında)
Varsayılan olarak bu kırtasiye malzemesini kullanın (kırtasiye malzemenizi seçmek için açılır
menüyü kullanın)
Nicole Mendez

İnternet İleti Biçimi (IMF), bilgisayar kullanıcıları arasında “elektronik posta” iletileri (e-postalar) olarak gönderilen ve RFC 5322 (Yorum İsteği) standardı [74] tarafından tanımlanan metin iletileri için bir sözdizimidir. ASCII dışı veriler ve multimedya içerik ekleri RFC 2045-2049'da tanımlanmıştır. Toplamda Çok Amaçlı İnternet Posta Uzantılarını (MIME) oluştururlar.

Çok parçalı e-postalar, bir veya daha fazla veri kümesinin tek bir gövdede birleştirildiği birden çok parçalı mesajlardır. Bu durumda, varlığın başlığında "çok parçalı" bir İçerik Türü dosyalaması görünmelidir. Gövde her biri bir öncekine ve bir kapanış sınırına sahip olacak şekilde bir veya daha fazla "gövde parçası" içerecektir. Çok parçalı e-postaların bölümleri, yaprak düğümleri olabilen veya diğer çoklu bölümler için kök olabilecek diğer bölümlere ayrılan bir kök bölümün bulunduğu bir ağaç yapısını takip eder.

2.12.1. E-posta İstemcileri (posta kullanıcı aracı)

Bir e-posta istemcisi, e-posta okuyucu veya mail kullanıcı aracı (MUA), bir kullanıcının e-postasına erişmek ve bunları yönetmek için kullanılan bir bilgisayar programıdır. Birincil rolü e-posta istemcileri olarak çalışmak olan web uygulamaları ve belirli bilgisayar donanımı veya yazılım parçaları vardır.

Bir e-postada, orijinal e-postayı hazırlamak ve göndermek için kullanılan programı belirten başlık “X-Mailer” başlığıdır.

2.12.2. E-posta Başlıkları

Daha önce de belirtildiği gibi, e-posta başlıkları belirli bir biçimde yazılır ve e-postanın belirli bir özelliğini belirler. Bunlar e-postanın değişkenleri olarak yorumlanabilir. Çok sayıda farklı e-posta başlığı vardır ve bir kullanıcı kendi özel ihtiyaçları için özel başlıklar oluşturabilir.

“İçerik Tipi” başlığı mesaj içeriğinin medya tipini belirtir. Üstbilgi değeri, bir tür ve "tür / alt tür" biçimindeki bir alt türden oluşur (ör. "Metin / düz" veya "metin / html").

Bu mekanizma aşağıdaki türleri destekler:

Çoklu Bölüm: Bu, e-postanın başka bölümleri olduğunda kök olarak kullanılan türdür. Bu bölüm, e-postanın kök bölümü olarak birden fazla bölüm içerdiği her durumda dahil edilmelidir. Ağaç yapısını karmaşıktırmak için bir e-postada birden çok parçalı parçalar kullanılabilir.

Metin: Bu tür, her türlü metin bilgisini içerir. Diğer alt türler, uygulamanın metnin görünümünü artırabileceği zenginleştirilmiş metin için kullanılır.

Mesaj: Bu tür kapsüllenmiş bir iletidir. “Mesaj” türünün bir parçası RFC 822 uyumlu bir mesajdır ve kendi farklı tür başlık alanını içerebilir.

Resim: Bu adın da belirttiği gibi, bu tür resim verilerini içerir. Bilgileri görüntülemek için bir görüntüleme cihazına (monitör, yazıcı vb.) İhtiyaç vardır. Alt türler görüntü dosyası formatlarını içerir:

Ses: Bu tür bir ses dosyası içerir. Birincil alt tip “ses / temel” dir ve daha karmaşık alt tipler “ses / mpeg” gibi dosya biçimini içerir. Ses, içeriği üretmek için bir ses aygıtı gerektirir.

Video: Bu tür bir video dosyası içerir. Birincil alt tür “video / mpeg” dir. Video, içerikleri görüntülemek için bir görüntüleme cihazı gerektirir.

Uygulama: Bu diğer türlere eklenemeyen verileri içerir. Tipik bir veri türü yorumlanmamış ikili veriler veya posta tabanlı bir uygulama tarafından işlenecek diğer verilerdir. Diğer kullanımlar arasında e-tablolar ve posta tabanlı zamanlama sistemleri için veriler bulunur. “İçerik-Eğilimi” başlığı bir posta iletisinin sunulduğunu açıklar. Bir MIME parçası şunları içerebilir:

Satır içi içerik düzenleme: mesaj görüntülendiğinde mesaj kısmı otomatik olarak görüntülenmelidir.

Ek içeriği düzenleme: parça otomatik olarak görüntülenmez ancak açmak için işlem gerektirir (açmak / göstermek veya indirmek için tıklanması gereken geleneksel ek).

“İçerik Aktarımı Kodlaması” üstbilgisi, bir e-postanın (ve bu konuyla ilgili herhangi bir dosyanın) ikili verilerini temsil etme yöntemini tanımlar. E-posta için kullanılan bir tür ikili-metin kodlaması varsa bu başlık dönüşüm algoritmasının ne kullanıldığını tanımlar, böylece posta istemcisi kodu çözebilir ve e-posta bilgilerini sunabilir. Herhangi bir dönüşüm yoksa bu başlık içeriğin biçimi için açıklayıcı bir etiket sağlar. İçerik Aktarımı Kodlaması üstbilgisinin en yaygın kullanılan değerleri ve buna bağlı olarak en yaygın kullanılan e-posta kodlamaları şunlardır:

Geleneksel e-posta kodlama biçimi olan 7 bit. Biçim ASCII numarasının her satırı için 998 karaktere kadar 1 ile 127 arasında değişebilir, CRLF karakterleri yalnızca birlikte ve yalnızca satırın sonunda görünür.

Base64: Baytları 7 bit kurallarını karşılayan bir forma kodlamak için kullanılır. Bu kodlama, US-ASCII karakterlerinden (metin olmayan 8 bit ve ikili veriler) oluşmayan baytlar için daha kullanışlıdır.

8bit: Bu, 8BITMIME SMTP uzantısını destekleyen SMTP sunucularıyla kullanım içindir, 7bit ile aynı türdedir. 8 bit'in varlığının nedeni ondan önce orijinal SMTP'nin yalnızca metin verilerinin taşınmasını desteklemesidir. Bu nedenle, herhangi bir ikili verinin bir metne kodlanması gerekiyordu. 8 bit kodlama bu sınırlamayı hafifletmek için geliştirilmiştir.

Binary: Bu herhangi bir bayt dizisidir.

2.12.3. E-posta Dosyası Biçimleri

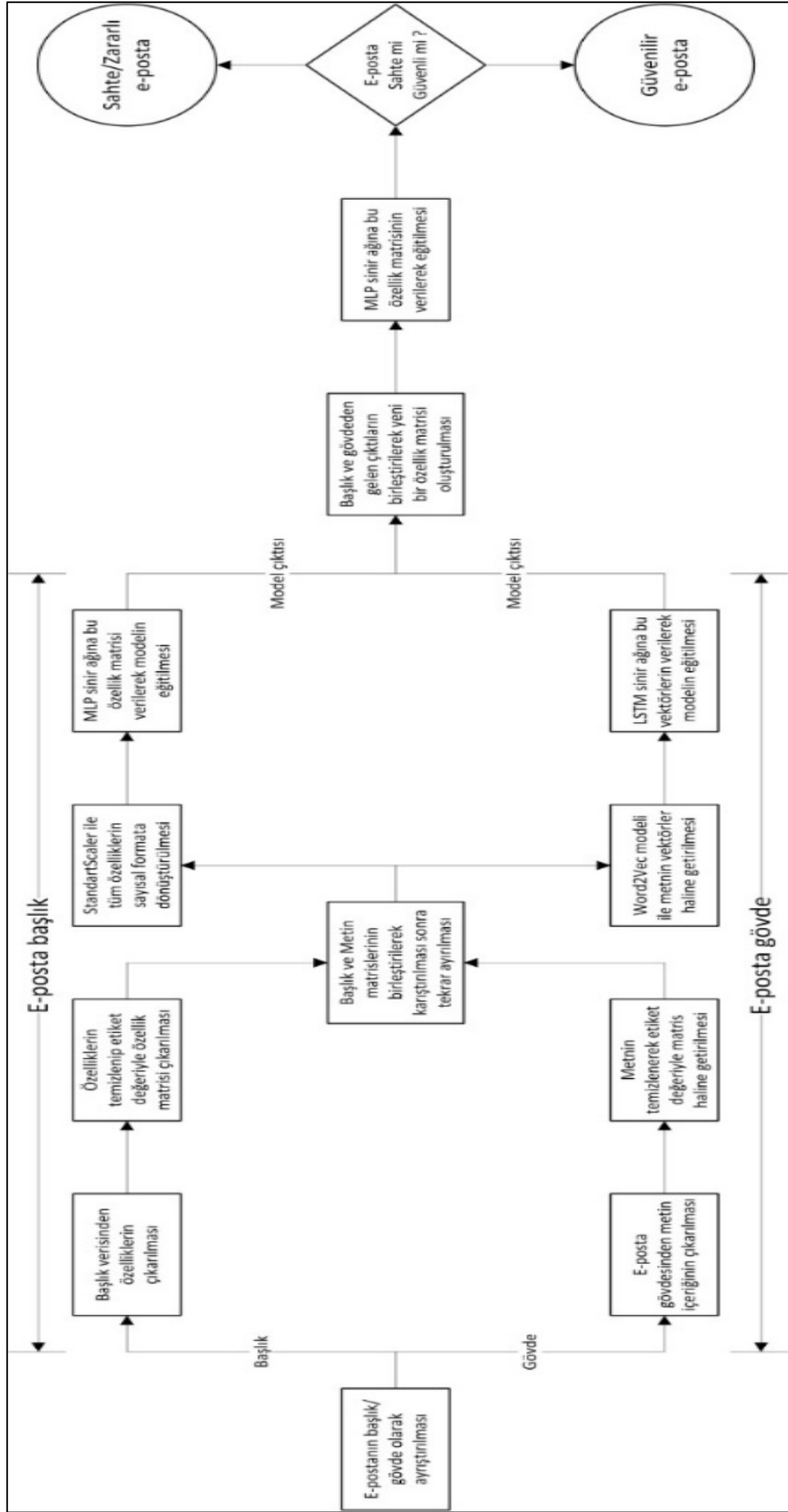
E-posta depolamak söz konusu olduğunda birçok farklı dosya formatı vardır. En popüler dosya biçimleri şunlardır [76]:

- EML (E-Posta İletisi Dosyası): Outlook ve diğer ilgili uygulamalar kullanılarak kaydedilen e-postalar. Microsoft Outlook, EML dosyalarını açmak için varsayılan yazılımdır.
- VCF (Sanal Kart Formatı), kişi bilgilerinin saklanması için dijital bir dosya formatıdır.

- TNEF (Aktarım Nötr Kapsülleme Biçimi), Messaging Application Programming Interface (MAPI) tabanlı e-posta eklerini kapsüllemek için Microsoft tabanlı bir biçimdir.
- PST (outlook Kişisel Depolama Dosyası), çeşitli kullanıcı bilgilerini depolayan bir dosyadır.
- OST (Çevrimdışı Depolama Dosyası), Microsoft Outlook kullanılarak Exchange Server'a kaydedildiğinde kullanıcının yerel makinedeki posta kutusu verilerinden oluşur.
- EMLX, Apple tarafından geliştirilen ve geliştirilen bir dosya biçimidir. Apple posta uygulaması, e-postaları dışa aktarmak için EMLX dosya biçimini kullanır.
- MBOX (Posta kutusu), e-postaların toplanması için bir dosya biçimidir. E-postalar ekleriyle birlikte kabın içinde saklanır. MBOX dosyası, e-postaları olan, birbiri ardına yazılmış ve birkaç yeni satırla ayrılmış bir metin dosyasıdır. Bu dosyada, “Kimden” başlık alanında yeni bir e-posta başlatılır ve bu, uygulamaların ve programların eriştiklerinde ve MBOX dosyasına eriştiklerinde aradığı satırdır. Bu proje, mbox biçimindeki e-posta veri kümelerinde çalışır.
- Maildir (mdir), her iletinin benzersiz bir adla ayrı bir dosyada tutulduğu ve her posta klasörünün bir dosya sistemi dizini olduğu e-postaları saklamak için kullanılan bir biçimdir. Maildir dizininde genellikle üç alt dizin bulunur: tmp, new ve cur. Maildir dosya formatı yapısı aşağıdaki şekilde gösterilmiştir. Kimlik avı e-postalarına karşı makine öğrenimi literatüründe bulunan birkaç veri kümesi bu dosya biçimindeydi ancak tutarlılık için mbox'a dönüştürüldü.

3. GELİŞTİRİLEN YÖNTEM(H-OLTA)

Kimlik avı e-posta saldırılarını tespit edebilmek için geliştirilen H-OLTA(Hibrit ve OLTA lama kelimelerinin birleşiminden türetilmiştir)yöntemde alınan e-postalar, incelenmesi için 2 parçaya bölünmüştür. Bu parçalar gelen e-postanın başlık ve gövde kısımlarıdır. E-postalar başlık kısmında gönderici, alıcı, içerik tipi, konu, ek dosyalar gibi özellikleri barındırır. Gövde kısmında ise genel mesaj içeriği bulunmaktadır. Fakat gönderici bu mesaj içeriğine resimler, dosyalar ve bağlantı adresleri ekleyebilir. Mesaj içeriğinde olan bu özelliklerde başlık analizi olarak adlandırılacak ilk analiz modeline gönderilmiştir. Bu sayede metin işleme yalnızca bir modelde gerçekleştirilirken, eposta içerisinden çıkarılan diğer tüm özellikler başlık analizi içerisinde toplanmıştır. Geliştirilen yöntem e-postanın başlık ve gövde bölümlerini ayrı olarak incelemiş, daha sonra her iki kısım için oluşturulan özellik matrislerine göre farklı derin öğrenme modellerine göndererek bu modeller eğitilmiştir. Başlık değerlerine göre oluşturulan özellik matrisi Çok katmanlı algılayıcı (Multilayer perceptron-MLP) modeline girdi olarak verilmiştir. Gövdeden çıkarılmış metin değerleri ise temizlenip word2vec modeli ile vektörler haline getirilerek LSTM sinir ağına girdi olarak verilmiştir. Bu iki modelden alınan çıktılar işlenerek yeni bir özellik matrisine dönüştürülerek, karar verecek olan yine tamamen bağlı (fully connected-FC) çok katmanlı algılayıcı bir modele verilmiştir. Bu çok katmanlı model ile diğer iki modelden gelen işlenmiş sonuçlar üzerinden daha doğru ve keskin tahminler yapılmaya çalışılmıştır.Yalnızca metin veya yalnızca başlık verilerinin kullanılmamasının sebebi 2 özellik kümesinin de birbirini destekleyip tek bir modelin veriyi ezberlemesinin önüne geçmektir. Şekil-3.1’de sistemin genel yapısı gösterilmiştir.



Şekil-3.1: Bu çalışmada geliştirilen karma modelin mimarisi.

3.1. Analiz

Bu kısımda geliştirilen H-OLTA yöntemi için başlık ve gövde analizi detaylı bir şekilde anlatılacaktır.

3.1.1. Başlık Analizi

E-postalar başlık kısımlarında gönderici ve niyeti hakkında belirgin izler taşıyabilirler. Oluşturulan yöntem e-postanın kimlik avı niyetiyle gönderilip gönderilmediğini kontrol etmek amacıyla geliştirilmiştir. Bu sebeple eposta başlığında çıkarılacak doğru ve yol gösterici öznitelikler e-postanın yüksek doğrulukla sınıflandırılması için çok önemlidir. Yöntem dahilinde daha önceki çalışmalardan da edinilen istatistikler doğrultusunda, kimlik avı e-postalarının en çok oranda barındırdığı ve tespit için en belirgin özellikler baz alınarak bir özellik seti oluşturulmuştur. Bu özellik seti ve özellikleri Tablo 3.1’de ve devamında verilmiştir.

Tablo 3.1: Başlık kısmından çıkarılan özellik kümesi.

Özellik No	Özellik Adı
1	URL içerisinde @ işareti bulunması
2	E-postada bulunan eklerin sayısı
3	CSS kontrolü
4	Harici kaynaklar
5	Flash içeriği kontrolü
6	HTML içeriği kontrolü
7	HTML Form kontrolü
8	HTML iFrame kontrolü
9	URL de IP kontrolü
10	Javascript kontrolü
11	URL sayısı

- URL içerisinde “@” işareti bulunması

E-posta metni içerisinde veya başlık içerisinde bulunan tüm URL’lerin “@” işareti barındırıp barındırmadığını kontrol eden özellik. Eğer kontrol edilen URL’ler içerisinde “@” işareti barındırıyorsa doğru (true) olarak, bulundurmuyorsa yanlış (false) olarak işaretlenmiştir.

- E-postada bulunan eklerin sayısı

E-postalar ile ek dosya gönderimi yapılabilmektedir. Bu dosyalar zararlı yazılım içerebilirler. Bu özelliğe gönderilen e-postanın kaç adet ek dosya içerdiği eklenmiştir.

- CSS Kontrolü

E-postalar yalnızca düz metinler olarak gönderilmemektedir. Zararlı e-postalar da kullanıcıyı aldatmak amacı ile genellikle görsel e-postalar hazırlanmaktadır. Bu hazır e-posta kalıpları içerisinde CSS stil dosyaları barındırmaktadır. Bu özelliğe e-posta içerisinde bulunan stil dosyalarının sayısı eklenmiştir.

- Harici Kaynaklar

E-posta metni içerisinde bulunan CSS stil ve Javascript etiketleri harici bir kaynağı referans gösterebilir. CSS stil dosyaları için kullanılan “style” etiketi “href” özelliği ile, Javascript dosyaları için kullanılan “script” etiketi ise “src” özelliği ile harici kaynakları referans gösterir. Bu özelliğe e-posta metni içerisinde yukarıdaki şekilde belirtilen yöntem ile referans gösterilen harici kaynakların sayısı eklenmiştir.

- Flash İçeriği Kontrolü

E-postalar içerisinde flash içerikler bulunabilmektedir. Geçmiş zamanlarda birçok kez flash dosyalar ile dağıtılan zararlı yazılımlar ve flash dosyaların barındırdığı zafiyetler üzerine çalışmalar yapılmıştır. Bu özellik e-posta içerisinde flash içeriği bulundurup bulundurmadığını kontrol etmektedir, eğer bulunduruyorsa doğru (true) olarak, bulundurmuyorsa yanlış (false) olarak işaretlenmiştir.

- HTML İçeriği Kontrolü

E-postalar yalnızca düz metin olarak gönderilmemektedir. Kimlik avı dolandırıcıları kullanıcıları aldatmak için görsel e-posta içerikleri

hazırlayabilmektedir. Bu hazırlanan içerikler ile bazı kurum ve kuruluşların taklit edilmiş içerikleri ile kullanıcılar aldatılabilir. Bu özellik e-posta içerisinde HTML içeriği bulundurup bulundurmadığını kontrol etmektedir, eğer bulunduruyorsa doğru (true) olarak, bulundurmuyorsa yanlış (false) olarak işaretlenmiştir.

- HTML Form Kontrolü

E-postalar içerisinde bulunan HTML içerikleri formlar barındırabilmektedir. Bu formlar ile kimlik avı dolandırıcıları kullanıcılardan aksiyonlar alarak bilgilerini ele geçirmeye çalışırlar. Bu özellik e-posta içerisindeki HTML içeriğinin form etiketi bulundurup bulundurmadığını kontrol etmektedir, eğer bulunduruyorsa doğru (true) olarak, bulundurmuyorsa yanlış (false) olarak işaretlenmiştir.

- HTML iFrame Kontrolü

iFrame, HTML içeriğinde belirli bir çerçeve içerisine farklı bir internet sayfasını çağırıp, görüntülenmesine yardımcı olan bir HTML etiketidir. Bu özellik e-posta içerisinde bulunan HTML içeriğinin iFrame etiketini barındırıp barındırmadığının kontrol edilmesi ile doğru (true) ya da yanlış (false) olarak işaretlenmiştir.

- URL de IP Kontrolü

E-posta içerisinde bulunan URL'lerin etki alan adı yerine IP adresi barındırabilmektedir. Kimlik avı dolandırıcıları etki alan adı yerine IP adresi ekleyebilmektedir. Bu özellik e-postaların etki alan adı yerine IP adresi barındırıp barındırmadığının kontrol edilmesi ile doğru (true) ya da yanlış (false) olarak işaretlenmiştir.

- Javascript Kontrolü

E-postalar kullanıcıyı aldatmak amaçlı çeşitli Javascript betikleri kullanabilirler. Bu betikler ile kullanıcıların birtakım aksiyonlar almalarını sağlayabilirler. Bu yüzden e-postaların içerisindeki Javascript betiklerinin sayısı önemlidir. Bu özellikte e-posta içerisindeki script etiketlerinin sayısı eklenmiştir.

- URL Sayısı

Bu özellik e-posta içerisinde bulunan toplam aktif URL sayısını kontrol etmektedir.

Tespit edilen bu özellikler veri kümesi içerisindeki tüm e-postalar için çıkartılmış ve özellik matrisi oluşturulmuş olup, Tablo 3.2 ‘de görselleştirilmiştir. Bu özellikler yalnızca e-postaların başlıklarında bulunan özellikler değildir, gövdesinde bulunan özellikler de buraya eklenmiştir. Çünkü e-posta gövdesinde bulunan metinler gövde analizi başlığı altında anlatılacağı şekilde kelime vektörleri haline getirilerek LSTM sinir ağı ile eğitilmiştir. Başlık analizinde çıkarılan özellikler, bu haliyle derin öğrenme yöntemine gönderilmeden önce bir dizi ön işleme tabi tutulmuştur. Bu işlemde standart ölçekleyici kullanılmıştır. Standart ölçekleyici standart skor hesaplama yöntemini kullanmaktadır. Bu ölçekleyici özellik matrisindeki değerlerin standart bir değere (puan) dönüştürülmesini sağlamaktadır.

Tablo 3.2: Başlık analizi sonrası oluşan özellik matrisi gösterimi.

	@inURLs	Attachments	Css	ExternalResources	Flashcontent	HTMLcontent	HtmlForm	HtmlIframe	IPsinURLs	Javascript	URLs
4147	False	0	0	0	False	True	False	False	True	0	47
4186	False	0	0	0	False	True	False	False	True	0	8
2958	False	0	0	0	False	True	False	False	False	0	2
2738	False	0	0	0	False	True	False	False	False	0	2
4210	False	0	0	0	False	True	False	False	False	1	7
...	...	...	...	...	...	...	...	...	...	...	...
2081	False	0	0	0	False	False	False	False	False	0	0
239	False	0	0	0	False	False	False	False	False	0	0
3826	False	0	0	0	False	True	False	False	False	0	10
3873	False	0	2	2	False	True	False	False	True	12	43
2778	False	0	0	0	False	True	False	False	False	0	2

4512 rows x 11 columns

Dönüşüm sonrasında ortaya çıkacak bu standart puan, ölçeklenecek değerden diğer tüm değerlerin ortalamasının çıkarılıp, tüm değerlerin standart sapmasına bölünmesiyle ifade (3.1) ile hesaplanmaktadır:[79]

$$z = \frac{x - \mu}{\sigma} \quad (3.1)$$

Burada;

μ değerlerin ortalamasını,

x standart puana dönüştürülecek değeri,

σ ise değerlerin standart sapmasını temsil etmektedir. Mutlak z değeri ise standart sapma birimlerinde dönüştürülecek değer ile ortalama değer arasındaki mesafeyi temsil etmektedir. z değeri, dönüştürülecek değer ortalamanın altında ise negatif, üstünde ise pozitif olacaktır. [79,80]

İfade (1)'de tanımlanan standart ölçekleyici sayesinde Çizelge-2' de gösterilen e-posta içerisinden çıkarılan 11 adet özelliğin standart puan yöntemine göre dönüştürülmesi sonucunda yeni bir özellik matrisi elde edilmiştir. Çizelge-2'de görünen doğru (true) ve yanlış (false) olarak işaretlenen verilere ise önce, doğru = 1, yanlış = 0 tam sayı dönüşümü yapılmıştır. Standart ölçekleyici dönüşümü sonrasında oluşan özellik matrisinden bir kesit Şekil-3.2'te gösterilmiştir.

```
[[-0.02105847, -0.059655 , -0.10824129, ..., 2.69232778,
-0.0998656 , 3.12055807],
[-0.02105847, -0.059655 , -0.10824129, ..., 2.69232778,
-0.0998656 , 0.15387414],
[-0.02105847, -0.059655 , -0.10824129, ..., -0.3714258 ,
-0.0998656 , -0.30253878],
...,
[-0.02105847, -0.059655 , -0.10824129, ..., -0.3714258 ,
-0.0998656 , 0.30601178],
[-0.02105847, -0.059655 , 4.58776554, ..., 2.69232778,
17.92387874, 2.8162828 ],
[-0.02105847, -0.059655 , -0.10824129, ..., -0.3714258 ,
-0.0998656 , -0.30253878]])
```

Şekil-3.2: Ölçeklendirilmiş veri kümesi.

3.1.2. Gövde Analizi

Başlık analizi bölümünde e-posta içerisindeki sınıflandırmayı en çok etkileyen özellik kümesi oluşturulmuştur. Gövde analizi içerisinde bulunan özellikler de başlık analizinde oluşturulan özellik matrisine eklenebilmesi için başlık analizi bölümüne gönderilmiştir. Gövde analizinde ise gelen e-postanın metin içeriği kısmı üzerinden bir özellik matrisi çıkarılmıştır. Öncelikle e-postanın metin değerinin içerisinde eğer

varsa bütün html etiketleri temizlenmiştir. Html etiketleri oluşturulacak kelime vektörlerini olumsuz yönde etkilediği ve çok sık tekrarlar gerçekleştirdikleri için temizlenmeleri gerekmektedir. Bu html etiket temizliğini Python dili içerisinde bulunan “BeautifulSoup” kütüphanesi ile yapılmıştır.[81] Metin içeriği html etiketlerinden arındırıldıktan sonra içerisinde URL bulundurup bulundurmadığı kontrolü yapılmıştır. Bulunan URL’ler başlık analizindeki özelliklere eklendikten sonra metin içerisinden çıkarılmıştır. Metin içerisindeki gereksiz boşluklar ve tüm sayılar python dili ve regex yapısı ile temizlenmiştir. Bu temizleme işlemlerinin sonucunda metin içerisinde hiçbir sayı, özel karakter, tanımlanamayan ifade, URL ve html etiketi bırakılmamıştır. Saf metin olarak kalan değerler küçük harflere dönüştürüldükten sonra, özellik matrisi oluşturulmak üzere word2vec olarak adlandırılan modele verilmiştir. Word2vec verilen metin içerisindeki karakter veya kelimeleri vektör uzayında ifade etmek için vektörlere dönüştüren, denetimsiz tahmin etmeye dayalı bir modeldir. Google tarafından 2013 yılında geliştirilmiştir. 2 tür yöntem kullanmaktadır. Sürekli kelime torbası (Continuous bag-of-words) ve skip-gram mimarilerini verimli bir şekilde uygulanmasını sağlar [82]. Yüksek boyutlu bir vektörün, düşük boyutlu bir alana dönüştürülmesi işlemine yerleştirme (embedding) işlemi denilmektedir. Bu Word2vec modeli ile yapılan işleme ise kelime yerleştirme (word embedding) denilmektedir. Metin kalıplarının vektör haline getirilmesi için keras kütüphanesinde [83] bulunan Tokenizer sınıfı ile metin değerleri önce kelime kelime ayrılarak belirteç (token) haline getirilmiştir. Tokenize işleminden sonra tam sayı dizilerine dönüştürülmüştür. Bu da keras’ta bulunan text_to_sequence yöntemi ile gerçekleştirilmiştir. Metinlerden alınan kelime sayısı 300 ile sınırlandırılmıştır. Bu sayede tüm e-postalardan eşit uzunlukta kelime vektörleri oluşturulmuştur. Şekil-3.3’te vektör haline getirilmiş kelime dizilerinden bir kesit gösterilmiştir.

```
array([[ 1, 290, 129, ..., 522, 291, 2891],
       [ 0,  0,  0, ..., 488, 5444, 776],
       [ 0,  0,  0, ..., 311, 5446, 3545],
       ...,
       [ 0,  0,  0, ...,  7, 1835, 1261],
       [2350,  8, 2356, ..., 1255, 797, 272],
       [ 0,  0,  0, ..., 10, 311, 1301]], dtype=int32)
```

Şekil-3.3: Vektör haline getirilmiş kelime dizileri.

Son olarak ise Google tarafından 3 milyon kelime ile eğitilmiş Word2Vec modelinden [82] alınan vektörler ile, Şekil-3'te gösterilen dizileri sabit boyutlu yoğun vektörlere dönüştürme işlemi için yerleştirme katmanı (embedding layer) oluşturulmuştur.

Bu çalışma ile e-postaların kimlik avı veya güvenli olarak sınıflandırılmasıdır. Bu bir sınıflandırma problemidir ve denetimli öğrenme kullanılır. Bu, yeni sınıflandırılmamış e-postaları sınıflandırabilmek için bir modeli eğitmek için kullanılacak önceden sınıflandırılmış bir veri kümesi olacağı anlamına gelir. Denetimli öğrenmede çıktıları tahmin etmek için bir modelin eğitilmesi için girişler sağlanır. Bir model, derin öğrenme sürecinden çıkan sınıflandırıcıdır ve her örneği y tahminiyle eşleştirir. Temel olarak derin öğrenme modelinin çalışma prensibi doğrusal regresyon örneği üzerinden verilmiştir. Bu modelde, n özellik sayısı, x veri setinde ki bir örneği ifade eden değişken, tüm örnekler için bu hesaplama yapılmaktadır. Her örnek için sağlanan etiketler/çıkışlar vardır ve temel model şu formüle sahiptir:

$$y' = b + w_1x_1 + w_2x_2 + \dots + w_nx_n \quad (3.2)$$

(3.2)'de b önyargı (bias)' dır ve w her örnek için bir ağırlık değeridir. Denetimli öğrenmede algoritma, her örnek geçtikten sonra önyargı ve ağırlıkların değerlerini güncelleyerek bir model oluşturur, böylece belirli bir kayıp fonksiyonu kullanılarak tahmin hatası en aza indirilir. Deneylerde kullanılan kayıp fonksiyonu, her tahmin ve

o tahminin gerçek değeri arasındaki ortalama kare farkını hesaplayan ortalama kare hatası'dır (3.3).

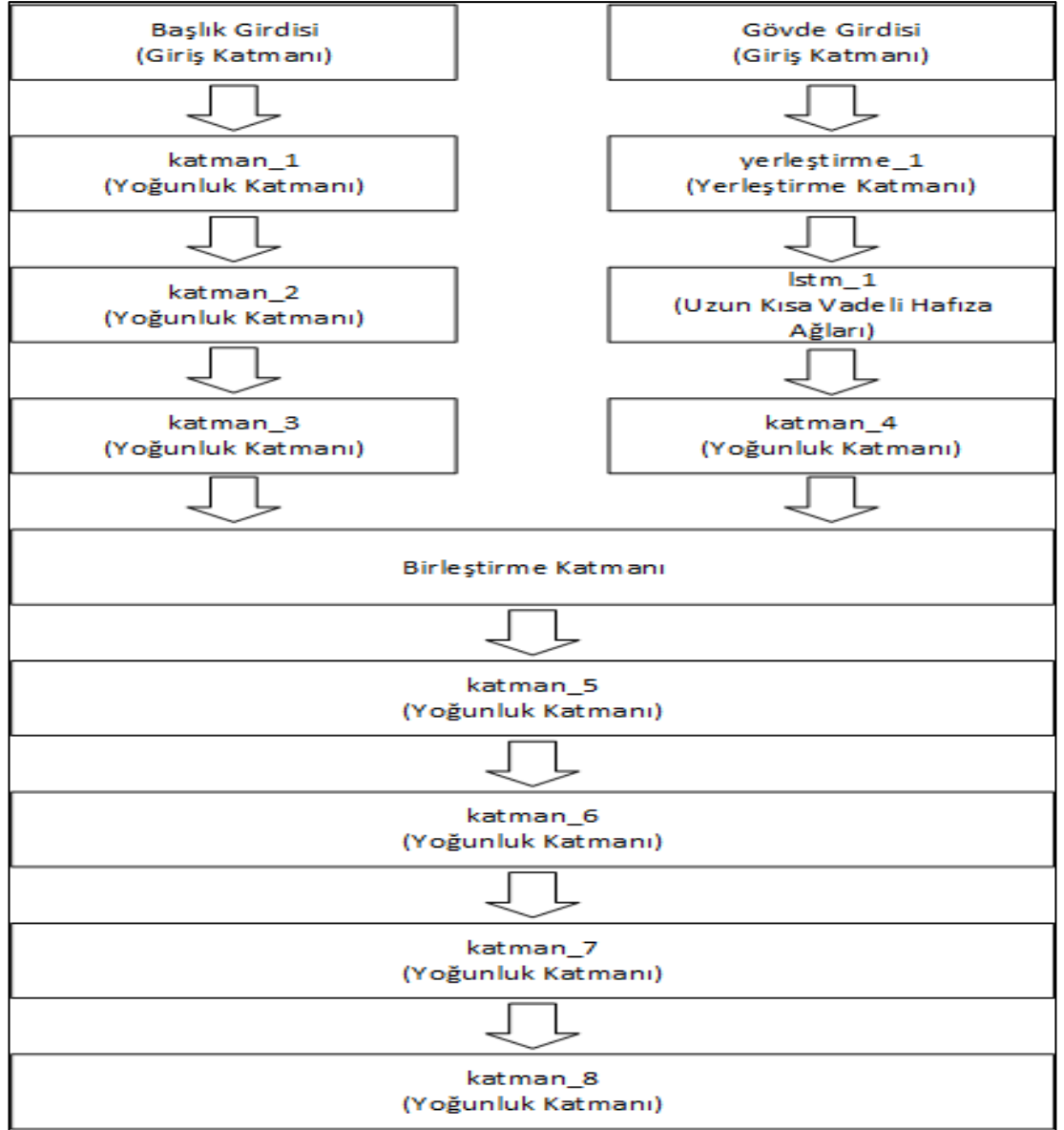
$$MSE = \frac{1}{n} \sum (y - \hat{y})^2 \quad (3.3)$$

Aktivasyon fonksiyonları ise her nöronda hesaplanan değerin çıktısını istenilen aralığa sıkıştırma veya istenilen şekilde almak için kullanılan fonksiyonlardır. Deneylerde kullanılan aktivasyon fonksiyonları ise ReLu (Rectified linear unit) ve Sigmoid fonksiyonlarıdır. ReLu şu şekilde ifade edilmektedir:

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (3.4)$$

(3.4) ile gelen negatif değerleri 0 değerine sabitlerken diğer değerleri aynı bırakmaktadır. Bu aktivasyon fonksiyonunu kullanma nedenimiz gövde analizinden yalnızca pozitif değerlerin gelmesi, modelin dengesiz karar vermesine neden olmaktadır. Bu yüzden başlık analizindeki negatif değerleri (3.4) fonksiyonu ile dengeleyip Sigmoid fonksiyonu (3.5) ile 0 ve 1 aralığına dönüştürülmüştür. Bu işleme normalizasyon işlemi denilmektedir. (3.5) fonksiyonu ile verilen değerler 0 ve 1 arasına normalize edilmiştir. Modelimiz tahmin sonucunda kimlik avı veya güvenli olarak işaretlemek istediği için bir sınıflandırma yapması gerekmektedir. Sonuçlarımız 0 veya 1 olduğu için sigmoid aktivasyon fonksiyonu bu işlemi gerçekleştirmiştir. Matematiksel tanımını (3.5)'deki şekildedir:

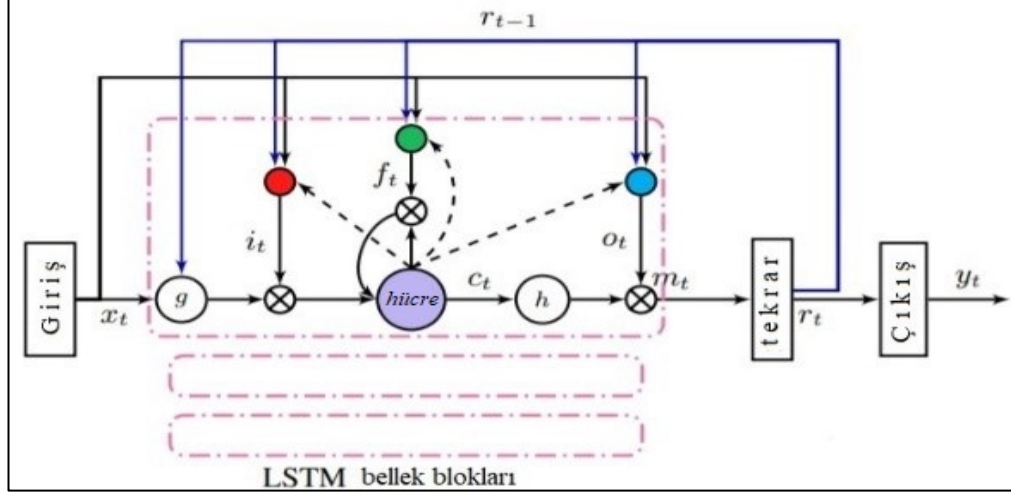
$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (3.5)$$



Şekil-3.4: Derin öğrenme modelinin görselleştirilmesi.

Son olarak model eğitiminde her adım da güncellenen değerler ile birlikte tahmin edilen değerın gerçek değere daha fazla yaklaşabilmesi için bir takım optimizasyon işlemlerinden geçmesi gerekmektedir. Bu çalışmada Adam optimizier [84] kullanılmıştır. Adam optimizier ile değişkenler optimize edilmiştir. Gövde analizinde oluşturulan yerleştirme katmanından sonra kullanılacak olan Uzun Kısa Vadeli Bellek (Long Short Term Memory - LSTM) ağları 1997 yılında tanıtılmıştır. LSTM mimarisi giriş, unutma ve çıkış olmak üzere 3 kapı, blok girişi, Sabit Hata Döngüsü, çıkış aktivasyon fonksiyonu ve gözetleme (peephole) bağlantılarına sahiptir [85]. Bloğun çıktısı tekrar tekrar bloğun girişine ve tüm kapılarına bağlanır. Gözetleme bağlantıları

ve unutma kapısı ilk geliştirilen mimaride bulunmamaktadır. LSTM'in kendi durumunu sıfırlamak için unutma kapısı, kesin zamanlamaları öğrenmeyi kolaylaştırmak için ise gözetleme bağlantıları eklenmiştir [85].



Şekil-3.5: LSTM mimarisi.

Başlık analizinde oluşturulan özellik matrisi ilk oluşturulan modele verilmeden önce eğitim ve test verisi olarak ayrılmıştır. Toplam 4512 veri içerisinde %80 (3609) eğitim, %20 (903) test verisi olacak şekilde bölündükten sonra, modele girdi olarak verilmiştir. Başlık analizi modelinde ilk katmanda 6 adet nöron bulunmaktadır, ardından gelen gizli yoğunluk katmanında da 6 adet nöron bulunmaktadır ve bu katmanda (4) fonksiyonu kullanılmıştır çünkü standart ölçekleme işleminden sonra oluşan özellik matrisi negatif değerler içerebilmektedir bu negatif değerleri optimize edebilmek için bu aktivasyon fonksiyonu kullanılmıştır. Daha sonra son çıkış katmanında 1 adet nöron bulunup aktivasyon olarak (5) fonksiyonu kullanılmıştır. Bu sayede çıktı 0 ve 1 arasında ölçeklendirilmiştir. Gövde analizi modelinde vektör dizilerine dönüştürülen ve ardından özellik matrisi haline getirilen veriler öncelikle yerleştirilme katmanına gönderilmiştir. Bu katmanda 300 adet nöron bulunmaktadır. Fakat bu katmandaki değişkenler eğitilemez değişkenlerdir ve sadece yerleştirme işleminde kullanılmıştır. Ardından LSTM katmanı gelmektedir ve bu katmanda 128 adet nöron bulunmaktadır ve onu, (5) fonksiyonu kullanan bir çıkış katmanı takip etmektedir. Bu iki analiz modelinden çıkan değerler birleştirilerek 2 adet 6 nöronlu gizli yoğunluk katmanına ardından (5) fonksiyonu kullanan bir çıkış katmanına sahip

modele bağlanmıştır. Bu modelde iki özellik grubu için tek derin öğrenme yöntemi kullanmak yerine, birden fazla model kullanılmasının nedeni sayısal ve niteliksel değerlerin, metin değerleriyle birlikte eğitilerek düşük ve optimal olmayan sonuçlara neden olacağı düşünülmektedir. Bunun yerine sayısal değerler için en iyi sonuçları alabileceğimiz modeller, metin değerleri için ise kelime vektörleri ve LSTM modellerinin kullanımı alınan sonuçları iyileştirilmiştir. İkinci bir neden ise yalnızca başlık analizi veya yalnızca gövde analizi ile yapılacak hatalar veya sınıflandıramama problemini birbirlerini destekleyerek çözmeleri beklenmiştir. Geliştirilen bu derin öğrenme modelinin özeti Tablo 3.3'te gösterilmiştir.

Tablo 3.3: Derin öğrenme modeli özeti.

Katman (tür)	Çıktı	Değişken
Başlık Analizi		
Yoğunluk katmanı	6	78
Yoğunluk katmanı	6	42
Yapay sinir ağı	1	7
Gövde Analizi		
Yerleştirme katmanı	300	900000000
LSTM	128	219648
Yapay sinir ağı	1	129
Karar Katmanı		
Yoğunluk katmanı	128	384
Yoğunluk katmanı	64	8256
Yoğunluk katmanı	32	2080
Yapay sinir ağı	1	33
Eğitilen Toplam Değişken Sayısı: 230,657		

4. DENEY SONUÇLARI ve DEĞERLENDİRME

4.1. E-posta Veri Kümesi

Bir derin öğrenme sınıflandırıcısı oluşturmanın ilk adımı uygun bir veri kümesine sahip olmaktır. Bu çalışmada yapılan araştırmalardan sonra aşağıdaki veri kümeleri bulunmuştur:

Jose Nazario kimlik avı e-postası veri kümesi [65]: Bu veri kümesi mbox biçimindeki dosyalarda kimlik avı e-postalarını içermektedir. E-postalar, veri kümesinin yaratıcısı tarafından sınıflandırılmıştır ve kendisine gönderilen e-postalardır. Literatürde birçok çalışmada kimlik avı posta kümesi olarak kullanılan bir veri kümesidir.

Enron Email Veri Kümesi [77]: Enron, dünyanın en büyük entegre doğal gaz ve elektrik şirketlerinden biriydi [78]. 2001 yılında iflas etmeden önce Enron, 2000 yılında yaklaşık 101 milyar dolar gelir ile 29.000 çalışan istihdam etti. 2001 yılı sonunda Enron'un raporlanan finansal durumunun, o zamandan beri Enron skandalı olarak bilinen sistematik ve planlanmış muhasebe sahtekarlığı tarafından sürdürüldüğü ortaya çıktı. Skandal nedeniyle, şirketin tüm e-postaları söz konusu skandalı araştırması sırasında Federal Enerji Düzenleme Komisyonu tarafından halka duyuruldu. Bu veri seti, çoğunlukla Enron'un üst düzey yönetimi olan ve klasörlerde düzenlenmiş yaklaşık 150 kullanıcının verilerini içerir. Şirket toplamda yaklaşık 500.000 e-posta içeriyor. Bu e-postalar, literatür tarafından makine öğrenimi algoritması eğitimi ve testi amacıyla meşru bir e-posta veri kümesi olarak kullanılmıştır. Deneylerde kullanılan veri kümesi bu 2 kaynaktan derlenmiştir ve 2256 güvenli ve 2256 kimlik avı e-postası olarak ayrılmıştır. Toplamda 4512 e-postadan oluşmaktadır. Tablo 4.1'te Veri kümesi görselleştirilmiştir.

Tablo 4.1: Veri kümesinin özeti.

Veri kümesi	Güvenli	Kimlik Avı	Toplam
Eğitim	1,801	1,808	3,609
Test	455	448	903
Toplam	2,256	2,256	4,512

4.2. Değerlendirme ve Deney Sonuçları

Test kümesini daha önce elde ettiğimiz modele girecek ve bir dizi gösterge elde edeceğiz. Bu göstergeler, modelin performansını özel olarak değerlendirmek için kullanılır. Bu göstergeleri şöyle açıklayabiliriz [86]:

True Positive (TP): Doğru algılanan kimlik avı e-postaların sayısı.

False Negative (FN): Kimlik avı olan fakat güvenli olarak algılanan e-postaların sayısı.

False Positive (FP): Güvenli olan fakat kimlik avı olarak algılanan e-postaların sayısı.

True Negative (TN): Doğru algılanan güvenli e-postaların sayısı.

Tablo 4.2’de sınıflandırma karışıklık matrisini gösterilmiştir.

Tablo 4.2: Sınıflandırma karmaşıklık matrisi.

Gerçek \ Tahmin	0 (Güvenli)	1 (Kimlik Avı)
	0 (Güvenli)	1 (Kimlik Avı)
0 (Güvenli)	TN	FP
1 (Kimlik Avı)	FN	TP

Tüm modeli test etmek için aşağıdaki değerlendirme ölçütlerini göz önünde bulunduruyoruz:

- Doğruluk (Accuracy)

Doğruluk, modelin veri kümesindeki gerçek sınıflandırmalarla karşılaştırıldığında elde ettiği doğru tahminlerin oranıdır. Şu şekilde hesaplanır:

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \quad (3.6)$$

- Kesinlik (Precision)

Kesinlik, doğru tahmin edilen pozitif örneklerin toplam pozitif tahminlere oranıdır:

$$Precision = \frac{TP}{TP+FP} \quad (3.7)$$

- Hatırlama (Recall)

Hatırlama veya algılama olasılığı (Recall), doğru tahmin edilen kimlik avı e-postalarının toplam kimlik avı e-postaları miktarına oranıdır:

$$Recall = \frac{TP}{TP+FN} \quad (3.8)$$

- F-Skor

F-skoru, hassasiyet ve hatırlamanın harmonik ortalamasıdır. F-skoru 0 ile 1 arasında bir değer aralığına sahiptir, burada 1 mükemmel hassasiyete eşittir ve her ikisi de 1'e eşittir. F1 skoru, hassasiyet ve hatırlama arasında bir ilişki metriği sağlamak için yararlıdır, çünkü bu iki değer bir şekilde ters benzetmeye sahiptir. aşağıda tarif edilen ROC eğrisi ile temsil edilmektedir:

$$F - measure = 2 * \frac{precision * recall}{precision + recall} \quad (3.9)$$

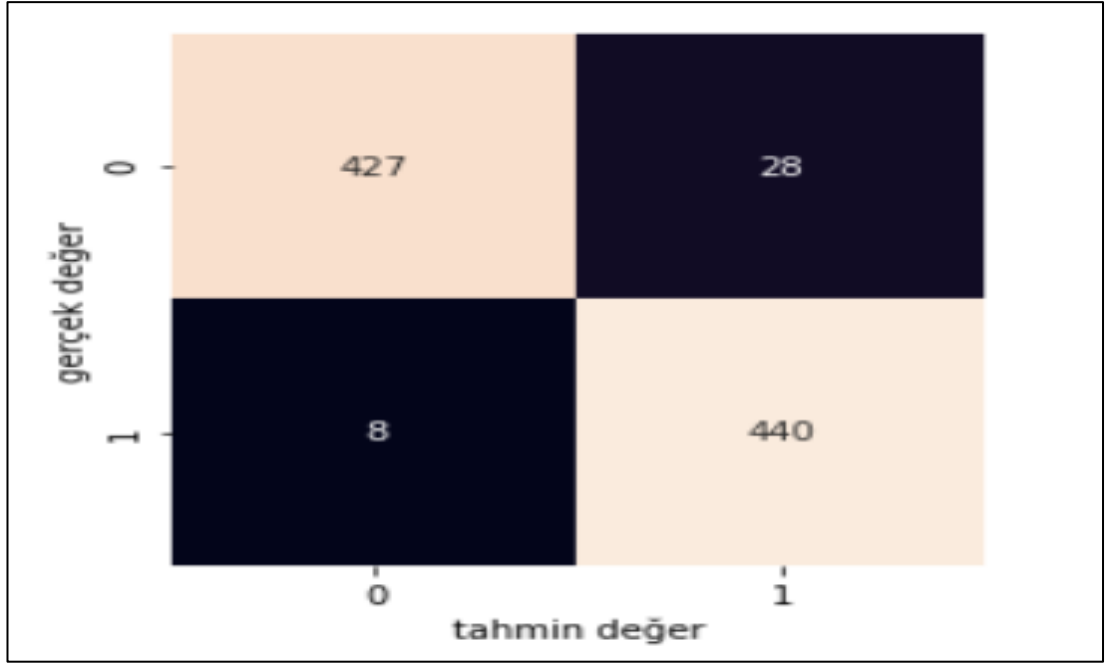
- Yanlış Alarm Olasılığı

Yanlış alarm olasılığı (FPR), pozitif bir tahminin yanlış olma olasılığıdır:

$$FPR = \frac{FP}{FP + TN} \quad (3.10)$$

- Karışıklık Matrisi

Bir karışıklık matrisi, tüm olası tahmin ve gerçek sonuç kombinasyonlarının listelendiği 2x2'lik bir matristir. Karışıklık matrisinin her bir hücresinin değerleri, bir test setinde yürütülen ilgili tahminlerin sayısına eşittir. Test işlemlerinden sonra modelimizin karmaşıklık matrisi Şekil-4.1'de gösterilmiştir.



Şekil-4.1: Karışıklık matrisi.

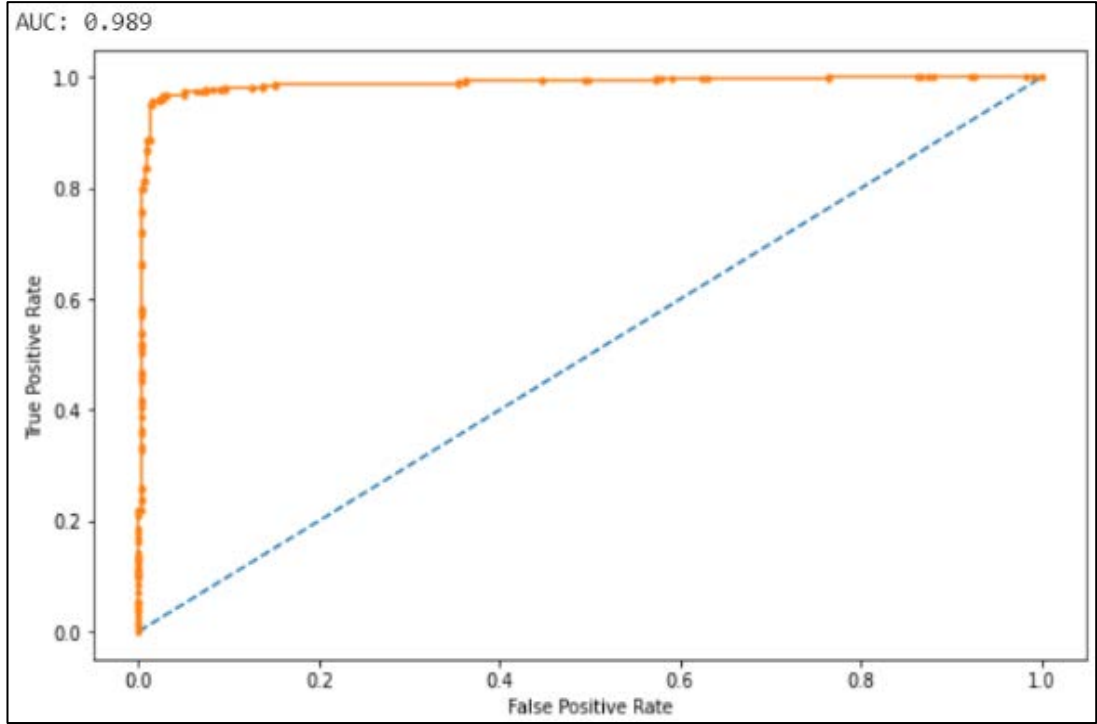
Yukarıda tanımlanan değerlendirme ölçütleri ise Tablo 4.3'te gösterilmiştir. Bu ölçütlerin hesaplanmasında ise oluşturulan modelin sınıflandırılma raporu özelliğinden faydalanılmıştır. Bu özellik sklearn [87] kütüphanesinde bulunmaktadır.

Tablo 4.3: Sınıflandırma raporu.

#	Precision	Recall	F-score	Accuracy
0	0.97	0.96	0.96	0.96
1	0.96	0.97	0.96	0.96

- ROC-AUC Eğrisi

ROC eğrisi, ikili sınıflandırma sistemlerinde ayırım eşik değerinin farklılık gösterdiği durumlarda, hassasiyetin kesinliğe olan oranıyla ortaya çıkmaktadır. ROC daha basit anlamda doğru pozitiflerin, yanlış pozitiflere olan kesri olarak da ifade edilebilir. Sınıflandırıcıların performans ölçümü için ROC analizi yapılarak doğruluk, hassasiyet ve kesinlik değerleri ile ROC puanı hesaplanabilir.[88]Yapılan deneyler sonrasında edinilen ölçüt değerleriyle birlikte oluşturulan ROC-AUC eğrisi Şekil-4.2'de gösterilmiştir.



Şekil-4.2: ROC-AUC eğrisi.

Tablo 4.4'te gösterilen yöntemler bu çalışmada kullanılan veri kümesinin birebir aynısını kullanmamışlardır. Fakat bu çalışmada kullanılan veri kümelerinin birçok varyasyonu bu yöntemlerde kullanılmaktadır. Veri kümesinin dengeli olması için 500.000 adet güvenli e-posta içerisinden sadece 2256 tanesi kullanılmıştır. Çünkü 2256 adet kimlik avı e-postası mevcuttur. Diğer yöntemlerde daha fazla veya az sayıda örnek kullanılmıştır. Ayrıca Tablo 4.4'te farklı kamu kurum ve kuruluşlarından özel izinler ile alınan çeşitli kimlik avı e-postası veri kümeleri de kullanılmıştır.

Tablo 4.4: Makine öğrenmesi algoritmaları ile sınıflandırma yapan diğer araştırmalar ile karşılaştırma.

Referans	Yöntem	Doğruluk
[29]	SVM ,NaiveBayes,AdB	86
[10]	Fisher Ayrımsalık Analizi	93
[26]	Word embedding + CNN	94,2
[34]	Doc2vec+SVM	88,4
[28]	RNN	91
[35]	Multi-filter approach	92,72
[6]	SVM+NNet	97,99
[36]	NaiveBayes-DecisionTree	96,77
H-OLTA	Hibrit MLP - LSTM	96,84

Bu karşılaştırmalara ek olarak derin öğrenme yöntemlerinde yalnızca bu değerler üzerinden değerlendirme yapmak yeterli olmayacaktır. Aynı zaman da hesaplanan kayıp (loss) değeri de önemli bir değerlendirme ölçütüdür. Düşük kayıp oranları modelin ne kadar hassas tahminlerde bulunabildiğini göstermektedir. Bu sebeple modelin geliştirilirken hedeflenen başarı kriterlerinden bir tanesi de düşük kayıp değeridir. Yöntemimizde %3'lük bir kayıp değeri oluşmaktadır ve bu kayıp değeri diğer birçok yönteme göre düşük bir seviyededir.

5. SONUÇLAR ve GELECEK ÇALIŞMALAR

Bu çalışmada, sosyal mühendislik saldırılarına, mevcut algılama tekniklerine genel bir bakış sağladık. Ayrıca bu çalışmada e-postaları kimlik avı ve düşük hata oranıyla güvenli olarak sınıflandırmak amacıyla derin öğrenme yöntemlerini birleştirerek hibrit bir yöntem olan ‘H-OLTA’ (Hibrit ve OLTA kelimelerinin birleşiminden türetilmiştir.) yöntemi önerilmiştir. Kimlik avı e-posta tespit sistemlerinde doğruluk yüksek bir paya sahiptir. Geliştirilen yöntemlerin bir kısmı yalnızca metin bölümüyle ilgilenerek uzun süren metin işleme yöntemleri önermektedir. Bu yöntemlerin uzun sürme nedeni büyük kelime havuzlarının verileri dönüştürme ve öğrenme süreleri oldukça uzun sürmesidir. Yalnızca özellik çıkarımı yöntemi metin değerlendirme yöntemlerine göre çok daha hızlı öğrenilebilse de tek başlarına tespit etmede hata oranları çok yüksek olabilmektedir. Bu yüzden geliştirilen H-OLTA yöntemi ile hem bu hata oranlarını düşürmek hem de kabul edilebilir hızlarda modeller eğiterek tahmin üretilmeye çalışılmıştır. Gövde analizi kısmında nispeten uzun süren tespit işlemini başlık analizinde kısa sürede eğitilen özellik çıkarımı modeli ile destekleyerek daha doğru ve tutarlı tespitlerde bulunulmuştur. Aynı zamanda H-OLTA yönteminde dengeli bir veri kümesi kullanılmıştır. Dengeli veri kümesi ile modelin eğitilmesi, modelin tek bir sınıfa eğilimli olmasını engellemek için önemlidir. H-OLTA yöntemi umut verici bir sonuç elde etmiştir. Özellik seçimi için yeni özellikler eklenebilir, fakat gereksiz özellik eklenmesi başarı oranının düşmesi gibi sorunları da beraberinde getireceğinden özellik çıkarımı yapılırken tespit için en fazla etkiye sahip özellikler alınmalıdır. Deneysel sonuçlar daha fazla eğitim verisinin dahil edilmesinin kimlik avı tespit doğruluğunu iyileştireceğini gösterdiğinden gelecekteki çalışmalarda kimlik avı tespit sisteminin performansını daha da artırmak için eğitim verilerine daha fazla veri kümesi eklenebilir. Kimlik avı e-postalarını gerçek zamanlı olarak tespit etmek için kullanılan paralel stratejiler nedeniyle model yapım süresi önemli ölçüde azalsa da hesaplama süreleri tam anlamıyla minimize edilemediği için daha verimli paralel algoritmalar yazılabilir. Modern bir toplumda bilginin korunması son derece önemlidir ve bilginin etrafındaki güvenlik seviyesi sürekli olarak iyileştirilse de zayıf olan tek nokta manipülasyon tekniklerine duyarlı olan insan olmaya devam edecektir. Bu yüzden insanların bilinçlendirilmesi yapılması gereken en önemli çalışmadır.

KAYNAKLAR

- [1] Khonji M., Iraqi Y., Jones A., (2013), "Phishing detection: a literature survey", IEEE Communications Surveys Tutorials, 15(4), 2091-2121.
- [2] Sheng S., Holbrook M., Kumaraguru P., Cranor L. F., Downs J., (2010), "Who falls for phish? A demographic analysis of phishing susceptibility and effectiveness of interventions", in Proceedings of the 28th Annual SIGCHI Conference on Human Factors in Computing Systems ,373-382, Atlanta, Georgia,10-15 April.
- [3] Behdad M., Barone L., Bennamoun M., French T., (2012), "Nature inspired techniques in the context of fraud detection", IEEE Transactions on Systems, Man, and Cybernetics Part C: Applications and Reviews, 42(6), 1273-1290.
- [4] Akinyelu A. A., Adewumi A. O., (2014), "Classification of phishing email using random forest machine learning technique", Journal of Applied Mathematics, 1(6),425731.
- [5] Mohammad R. M., Thabtah F., McCluskey L., (2014), "Intelligent rule-based phishing websites classification", IET Information Security, 8(3), 153-160.
- [6] Thornburgh T., (2004), "Social engineering: the dark art" ,Proceedings of the 1st Annual Conference on Information Security Curriculum Development, 133–135, Kennesaw, GA, USA, 4-8 October.
- [7] Irani D., Balduzzi M., Balzarotti D., Kirde E., Pu C., (2011), "Reverse social engineering attacks in online social networks", In International conference on detection of intrusions and malware, and vulnerability assessment 55-74, Springer, Berlin, Germany,2011.
- [8] Web 1, (2015), <https://www.social-engineer.org/> , (Date of Access: 17/10/2019).
- [9] Ivaturi K., Janczewski L., (2011), "A taxonomy for social engineering attacks", In International Conference on Information Resources Management, Centre for Information Technology, Organizations, and People,1-12,New Zeland,24 August.
- [10] Kvedar D., Nettis M., Fulton S. P., (2010), "The Use of Formal Social Engineering Techniques To Identify Weaknesses During a Computer Vulnerability Competition". Journal of Computing Sciences in Colleges, 26(2),80-87.
- [11] Long J., (2011), "No tech hacking: A guide to social engineering, dumpster diving, and shoulder surfing", 1st Edition , Syngress.
- [12] Janczewski L.J., Fu L., (2010), "Social Engineering-Based Attacks: Model and New Zealand Perspective", in proceedings of the international multiconference

on computer science and information technology ,847–853, Wisla,Poland,18-20 October.

- [13] Mitnick K. D., Simon W. L., (2003), “The art of deception: Controlling the human element of security”, 1st Edition, John Wiley Sons.
- [14] Abraham S., Chengalur-Smith I., (2010), “An overview of social engineering malware: Trends, tactics, and implications”, *Technology in Society*, 32(3), 183-196.
- [15] Granger S., (2001), “Social engineering Fundamentals, part I: hacker tactics”, 1st Edition, Security Focus.
- [16] Nelson R., (2001)., “Methods of hacking: Social engineering”, 10 ,the Institute for Systems Research, University of Maryland, ABD.
- [17] Web 2, (2020), “<http://www.paterva.com/maltego/>”, (Date of Access: 06/05/2020).
- [18] Web 3, (2019), <http://www.darkreading.com/security/perimeter/show> ,(Date of Access: 21/07/2020).
- [19] Herley C., Florêncio D., (2008), “A profitless endeavor: phishing as tragedy of the commons”, In Proceedings of the 2008 New Security Paradigms Workshop ,59-70.
- [20] Jagatic T. N., Johnson N. A., Jakobsson M., Menczer F., (2007), “Social phishing”, *Communications of the ACM*, 50(10), 94-100.
- [21] Parsons K., McCormac A., Pattinson M., Butavicius M., Jerram C., (2013), “Phishing for the truth: A scenario-based experiment of users’ behavioural response to emails” In IFIP International Information Security Conference 366-378, Springer, Berlin, Germany, 8-10 July.
- [22] Tam L., Glassman M., Vandenwauver M., (2013), “The psychology of password management: a tradeoff between security and convenience”, *Behaviour Information Technology*, 29(3), 233-244.
- [23] Bader G., Anjomshoaa A., Tjoa A., (2010), “Privacy aspects of mashup architecture. In: Social Computing (SocialCom)”, IEEE Second International Conference on, 1141-1146, Minneapolis, MN, USA, 20-22 August.
- [24] Miller K., Voas J., Hurlburt G., (2012), “Byod: security and privacy considerations”. *IT Professional*, 2012;14(5), 53-55
- [25] Thompson H., (2012), “The human element of information security”, *IEEE Security Privacy*, 11(1):32-35.
- [26] Hadnagy C., (2010), “Social engineering: The art of human hacking”, 1st Edition, John Wiley Sons.

- [27] Ghafir I., Prenosil V., Alhejailan A., Hammoudeh M., (2016), “Social engineering attack strategies and defence approaches”, In 2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud),145-149, Vienna, Austria,22-24 August.
- [28] Knapp K. J. (Ed.), (2009), Cyber Security and Global Information Assurance: Threat Analysis and Response Solutions: Threat Analysis and Response Solutions,1st Edition, IGI Global.
- [29] Greitzer F. L., Strozer J. R., Cohen S., Moore A. P., Mundie D., Cowley J., (2014), “Analysis of Unintentional Insider Threats Deriving from Social Engineering Exploits”, IEEE Security and Privacy Workshops, 236–250.
- [30] Web 4, (2018), <https://financialcryptography.com/mt/archives/000609.html>, (Date of Access:25/03/2020).
- [31] Web 5, (2020), <http://www.antiphishing.org/trendsreports/> ,(Date of Access:05/05/2020).
- [32] Web 6, (2019), <https://www.phishing.org/phishing-techniques>, (Date of Access:13/05/2020).
- [33] Drake C. E., Oliver J. J., Koontz E. J., (2004), “Anatomy of a Phishing Email”, In CEAS, 216-154.
- [34] Web 7, (2019), <https://tinyurl.com/> , (Date of Access:24/07/2019).
- [35] Jaquith A., (2007), “Security metrics: replacing fear, uncertainty, and doubt,1st Edition, Pearson Education.
- [36] Purkait S., (2012), “Phishing counter measures and their effectiveness–literature review”, Information Management Computer Security,382-420.
- [37] Web 8, (2018), <https://www.vadesecure.com/en/phishing-awareness-training-8-things employees-understand/>, (Date of Access:12/05/2019).
- [38] Web 9, (2017), <https://symtrex.com/security-solutions/user-education/> (Date of Access:27/06/2019).
- [39] Gerbet T., Kumar A., Lauradoux C., (2016), “A privacy analysis of Google and Yandex safe browsing”, In 2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, 347-358, Toulouse,France,28 June-1 July.
- [40] Bell S., Komisarczuk P., (2020), “An Analysis of Phishing Blacklists: Google Safe Browsing, OpenPhish, and PhishTank”, In Proceedings of the Australasian Computer Science Week Multiconference ,1-11.

- [41] Rasthofer S., Arzt S., Bodden E., (2014), "A machine-learning approach for classifying and categorizing android sources and sinks", In NDSS,14, 1125.
- [42] Schaffer C., (1993), "Selecting a classification method by cross-validation", Machine Learning, 13(1), 135-143.
- [43] Krombholz K., Hobel H., Huber M., Weippl E., (2015), "Advanced social engineering attacks", Journal of Information Security and applications, 22, 113-122.
- [44] Mouton F., Leenen L., Venter H. S., (2016), "Social engineering attack examples, templates and scenarios", Computers Security, 59, 186-209.
- [45] Edwards M., Larson R., Green B., Rashid A., Baron A., (2017), "Panning for gold: automatically analysing online social engineering attack surfaces", Computers Security, 69, 18-34.
- [46] Halgas L., Agrafiotis I., Nurse J. R. (2019), "Catching the Phish: Detecting Phishing Attacks using Recurrent Neural Networks (RNNs)". arXiv preprint arXiv:1908.03640.
- [47] Juarez-Orozco L. E., Martinez-Manzanera O., Nesterov S. V., Kajander S., Knuuti J., (2018), "The machine learning horizon in cardiac hybrid imaging", European Journal of Hybrid Imaging, 2(1), 15.
- [48] Bhavsar H., Ganatra A., (2012), "A comparative study of training algorithms for supervised machine learning", International Journal of Soft Computing and Engineering (IJSCE), 2(4), 2231-2307.
- [49] Web 10, (2018), <https://www.coursehero.com/file/32960124/A-Complete-Tutorial-on-Tree-Based-Modeling-from-Scratch-in-R-Pythopdf/> ,(Date of Access:02/24/2019).
- [50] Sibi P., Jones S. A., Siddarth P., (2013), "Analysis of different activation functions using back propagation neural networks", Journal of Theoretical and Applied Information Technology, 47(3), 1264-1268.
- [51] Ninness C., Rumph R., McCuller G., Harrison C., Ford A. M., Ninness S. K. (2005), "A functional analytic approach to computer-interactive mathematics", Journal of Applied Behavior Analysis, 38(1), 1-22.
- [52] Leibovich-Raveh T., Lewis D. J., Al-Rubaiey Kadhim S., Ansari D., (2018), "A new method for calculating individual subitizing ranges",Journal of Numerical Cognition, 4(2):429-447.
- [53] Mandic D., Chambers J., (2001), "Recurrent neural networks for prediction: learning algorithms, architectures and stability",1st Edition, Wiley.
- [54] Jagatic T. N., Johnson N. A., Jakobsson M., Menczer F., (2007), "Social phishing", Communications of the ACM, 50(10), 94-100.

- [55] Natekin A., Knoll A., (2013), "Gradient boosting machines, a tutorial", *Frontiers in neurorobotics*, 7, 21.
- [56] Tama B. A., Rhee K. H., (2017), "An extensive empirical evaluation of classifier ensembles for intrusion detection task" *Computer Systems Science and Engineering*, 32(2), 149-158.
- [57] Divina F., Gilson A., Gómez-Vela F., García Torres M., Torres J. F., (2018), "Stacking ensemble learning for short-term electricity consumption forecasting", *Energies*, 11(4), 949.
- [58] Sawant A. A., Game P. S. (2018, August), "Approaches for Anomaly Detection in Network: A Survey", In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) ,1-6, Pune, India, 16-18 August.
- [59] Abu-Nimeh S., Nappa D., Wang X., Nair S., (2007, October), "A comparison of machine learning techniques for phishing detection", In *Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit* ,60-69.
- [60] Hussain R., Qamar U., (2014), "An Approach to Detect Spam Emails by Using Majority Voting", In *International Conference on Data Mining, Internet Computing and Big Data (BigData2014)* ,76-83, Kuala Lumpur, Malaysia,17-19 November.
- [61] Web 11, (2019), <https://archive.ics.uci.edu/ml/datasets/Spambase> , (Date of Access:30/09/2019).
- [62] Hofmann M., Klinkenberg R. (Eds.), (2016), "RapidMiner: Data mining use cases and business analytics applications", 1st Edition,CRC Press.
- [63] Moradpoor N., Clavie B., Buchanan B., (2017), "Employing machine learning techniques for detection and classification of phishing emails", In *2017 Computing Conference* ,149-156, London, England,18-20 July.
- [64] Web 12, (2018), <https://spamassassin.apache.org/old/publiccorpus/>, (Date of Access:23/05/2019).
- [65] Web 13, (2018), <https://monkey.org/~jose/phishing/>, (Date of Access:22/05/2019).
- [66] Web 14, (2019), <https://www.mathworks.com/>, (Date of Access:19/03/2019).
- [67] Tan C. L., Chiew K. L., Musa N., Ibrahim D. H. A., (2018), "Identifying the Most Effective Feature Category in Machine Learning-based Phishing Website Detection", *International Journal of Engineering Technology*, 7(4.31), 1-6.

- [68] Zareapoor M., Seeja K. R., (2015), "Feature extraction or feature selection for text classification: A case study on phishing email detection", International Journal of Information Engineering and Electronic Business, 7(2), 60.
- [69] Almomani A., Gupta B. B., Atawneh S., Meulenberg A., Almomani E., (2013), "A survey of phishing email filtering techniques", IEEE communications surveys tutorials, 15(4), 2070-2090.
- [70] Bergholz A., Chang J. H., Paass G., Reichartz F., Strobel S. (2008), "Improved Phishing Detection using Model-Based Features", In CEAS.
- [71] Ma L., Ofoghi B., Watters P., Brown S., (2009), "Detecting phishing emails using hybrid features", In 2009 Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing ,493-497.
- [72] Del Castillo M. D., Iglesias Á., Serrano J. I. (2007, February), "An integrated approach to filtering phishing e-mails". In International Conference on Computer Aided Systems Theory,4739,321-328. Springer, Berlin, Germany,12-16 February.
- [73] Islam R., Abawajy J. (2013), "A multi-tier phishing detection and filtering approach", Journal of Network and Computer Applications, 36(1), 324-335.
- [74] Web 15, (2008), <https://tools.ietf.org/html/rfc5322> ,(Date of Access:16/08/2019).
- [75] Wood D., (1999), Programming Internet Email: Mastering Internet Messaging Systems ",1st Edition, O'Reilly Media, Inc."
- [76] Capiel G. J., (2002), "Method and system for remotely sensing the file formats processed by an E-mail client", U.S. Patent ,No. 6,449,634.
- [77] Shetty J., Adibi J., (2004), "The Enron email dataset database schema and brief statistical report", Information sciences institute technical report, University of Southern California, 4(1), 120-128
- [78] Steffensmeier D. J., Schwartz J., Roche M., (2013), "Gender and twenty-first-century corporate crime: Female involvement and the gender gap in Enron-era corporate frauds", American Sociological Review, 78(3), 448-476.
- [79] Zill D., Wright W. S., Cullen M. R., (2011), "Advanced engineering mathematics",1st Edition, Jones Bartlett Learning.
- [80] Milton J. S., Arnold J. C., (1994), "Schaum's Outline of Introduction to Probability Statistics: Principles Applications for Engineering the Computing Sciences",1st Edition, McGraw-Hill Higher Education.
- [81] Nair V. G., (2014), "Getting Started with Beautiful Soup",1st Edition,Packt Publishing Ltd.

- [82] Mikolov T., Grave E., Bojanowski P., Puhřsch C., Joulin A., (2017), “Advances in pre-training distributed word representations”, arXiv preprint arXiv:1712.09405.
- [83] Patel A. A., (2019), “Hands-On Unsupervised Learning Using Python: How to Build Applied Machine Learning Solutions from Unlabeled Data”,1st Edition, O'Reilly Media.
- [84] Kingma D. P., Ba J., (2014), “Adam: A method for stochastic optimization”, ArXivpreprint arXiv:1412.6980.
- [85] Şeker A, Diri B , Balık H., (2017), “Derin Öğrenme Yöntemleri ve Uygulamaları Hakkında Bir İnceleme”, Gazi Mühendislik Bilimleri Dergisi (GMBD) , 3 (3) , 47-64.
- [86] Vinayakumar R., Barathi Ganesh, H. B., Anand Kumar M., Soman K. P., (2018), “DeepAnti-PhishNet: applying deep neural networks for phishing email detection”, CEN-AISecurity@ IWSPA, 40-50.
- [87] Kramer Oliver., (2016), "Scikit-learn." Machine learning for evolution strategies, Springer, Cham, 45-53.
- [88] Fawcett T., (2006), "An Introduction to ROC Analysis", Pattern Recognition Letters, 861-874.

ÖZGEÇMİŞ

Şeydanur AHİ 1994 yılında Erzincan’da doğdu. 2012 yılında başladığı Cumhuriyet Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümünü 2016 yılında başarıyla tamamlayarak 2017 yılında yüksek lisans eğitimine Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı Siber Güvenlik Programında başladı. 2019 yılından bu yana Gebze Teknik Üniversitesi’nde araştırma görevlisi olarak çalışmaktadır.