Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi
Afyon Kocatepe University – Journal of Science and Engineering
https://dergipark.org.tr/tr/pub/akufemubid

Araştırma Makalesi / Research Article
e-ISSN: 2149-3367
DOI: https://doi.org/10.35414/akufemubid.1566769
AKÜ FEMÜBİD **26** (2026) 015103 (89-101)
AKU J. Sci. Eng. **26** (2026) 015103 (89-101)

Yalova İli Bilişim Suç Türlerinin Yapay Zekâ Yöntemleriyle Tahmini

Makale Bilgisi / Article Info
Alındı/Received: 15.10.2024
Kabul/Accepted: 27.09.2025
Yayımlandı/Published: 03.02.2026

The Prediction of Cybercrime Types in Yalova Province Using Artificial Intelligence Methods

Güneş HARMAN[1*], Sezer YURDUSEVEN[2]

[1] *Yalova Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Yalova, Türkiye*
[2] Yalova Üniversitesi, Lisansüstü Eğitim Enstitüsü/Adli Bilişim Anabilim Dalı/Adli Bilişi, Yalova, Türkiye

Öz

Çağımız dünyasında gelişen teknolojilerle birlikte meydana gelen bilişim suçu türlerinin de hızla arttığı gözlenmektedir. İşlenmekte olan suç çeşitliliğinin artması gerek hukuki altyapıyı gerekse kolluk birimlerinin işlerini zorlaştırmaktadır. Elde edilen dijital veri miktarının artması, analiz ve raporlama gibi süreçlerin uzamasına sebep olmaktadır. Bu çalışma, 2020-2023 yılları arasında Yalova Cumhuriyet Başsavcılığı'na intikal etmiş bilişim suçlarının makine öğrenmesi algoritmaları kullanılarak analiz edilmiştir. Elde edilen sonuçlar doğrultusunda, meydana gelmesi olası/muhtemel bilişim suçu türleri ile bu olaylara müdahil olan kişi veya kişilerin profilleri tespiti yapılmaya çalışılmıştır. Çalışma kapsamında toplam 2500 veriden oluşan bir veri seti kullanılmış ve 11 adet öznitelik belirlenmiştir. Bu veri seti üzerinde Rastgele Orman (RO), Destek Vektör Makineleri (DVM), Çok Katmanlı Yapay Sinir Ağları (Çok Katmanlı YSA), K-En Yakın Komşular (IBK) gibi çeşitli makine öğrenmesi algoritmaları kullanılarak performansları karşılaştırılmıştır.

Algoritmalarının performans değerlendirmeleri sonucunda, bilişim suçu olaylarına karışan kişilerin müşteki (mağdur) mi yoksa şüpheli mi olduğuna karar verebilmek açısından en yüksek doğruluk oranı %83 ile DVM-Sigmoid algoritmasından elde edilmiştir. Öte yandan, bilişim suçu türlerine ait makine öğrenmesi algoritmalarının performans değerlendirmeleri karşılaştırıldığında, en yüksek başarı %85 doğruluk oranı ile RO algoritmasından alınmıştır.

***Anahtar Kelimeler**: Yapay zekâ, Makine öğrenmesi, Bilişim suçları, Yalova ili.*

Abstract

In today's world, it is observed that the types of cybercrimes that occur with the developing technologies are also increasing rapidly. The increase in the variety of crimes committed complicates both the legal infrastructure and the work of law enforcement units. The increase in the amount of digital data obtained causes the processes such as analysis and reporting to be prolonged. In the study, an attempt was made to interpret the cybercrimes that were transferred to the Yalova Chief Public Prosecutor's Office between 2020-2023 using machine learning algorithms. At the same time, in line with the results, an attempt was made to determine the types of cybercrimes that may occur and the profiles of the people involved in the incidents. Within the scope of the study, 11 features were determined from a total of 2500 pieces of data. The performances of machine learning algorithms such as Random Forest (RF), Support Vector Machines (SVM), Multilayer Artificial Neural Networks (Multilayer ANN), K-Nearest Neighbors (NNN) were compared on the data set used. When the performance evaluations of the machine learning algorithms used are compared, the best result in deciding whether the people involved in cybercrime incidents are complainants (victims) or suspects was obtained from the SVM-Sigmoid algorithm with an accuracy rate of 0.83. When the performance evaluations of the machine learning algorithms belonging to cybercrime types are compared, the best result was obtained from the RO algorithm with an accuracy rate of 0.85.

***Keywords:** Artificial Intelligence, Machine learning, Cybercrime, Yalova province.*

1. Giriş

Bilişim ve iletişim alanlarındaki teknolojilerin hızla gelişmesiyle birlikte, bilişim kavramının önemi artmaya başlamıştır. Bu teknolojik ilerlemeler, bilişim teriminin günlük hayatta ve akademik çalışmalarda daha sık kullanılmasına neden olmuştur. Kendi iktisadi ve kültürel yeniliklerini oluşturmaya başlayan bilişim, tüm dünyanın haberleşmesinde kullanılan teknolojik sistemler sayesinde düzenli bir şekilde bilginin işlenmesini ve yönetilmesini öngören bir bilim dalı haline gelmiştir. Bilişim, teknoloji ile bilgi arasında bir köprü işlevi görmeye başlamış; veri toplama, saklama, işleme, iletişim kurma ve analiz etme süreçlerini kolaylaştırarak bireyler ve işletmelerin verimliliğini artırmıştır. Bu süreçte teknolojiye ulaşmanın en kolay yolu olan internet kullanımda ciddi oranda artmış, buna paralel olarak dijital ortamlarda işlenen suçlarda da önemli ölçüde artışlar meydana gelmiştir. Teknolojik gelişmelerin en büyük avantajı insan hayatını kolaylaştırması iken, en büyük dezavantajı ise kötü niyetli kişilerin çok sayıda kaynağa erişim sağlaması ve sistem açıklarını istismar edebilmesidir (Yılmaz vd. 2016). Bu nedenle, teknolojik

*Sorumlu Yazar/Corresponding Author: Güneş HARMAN e-posta/e-mail: gunes.guclu@yalova.edu.tr

cihazların ve ağ teknolojilerinin yaygınlaşmasıyla birlikte, bu sistemleri kötü amaçları için kullanan kişiler ortaya çıkmıştır. Bu durum, "bilişim suçları" olarak adlandırılan yeni bir suç türünün doğmasına yol açmıştır (Ermeydan 2018). Bilişim suçlarının tüm yönleriyle incelenmesi ve şüphelilerin tespit edilmesinde adli bilişim en büyük destekleyici unsur olmuştur.

Adli bilişim, bilgisayar sistemlerinde saklanan ya da bu sistemler aracılığıyla iletilen bilgilerin toplanması, incelenmesi ve raporlanmasını sağlayan bir bilim dalı olmanın yanında, belirli bilgi, ilke ve disiplinlere de sahiptir (Bilgi Teknolojileri ve İletişim Kurumu, 2019). Bilgi Teknolojileri ve İletişim Kurumu 2019). Diğer bir ifadeyle adli bilişim; meydana gelen bir bilişim suçu kapsamında bilişim ağlarında bulunan dijital delillere ulaşıldığı andan, bu delillerin adli mercilere ulaştırılmasına kadar geçen sürecin bütünüdür.

Adli bilişim kavramının kanuni uygulamalarından ziyade teknolojik boyutu ön plandadır. Bilişim sistemlerinden elde edilen delillerin ayrıştırılarak hukuki bir delil olma özelliğine dönüştürme süreci oldukça zordur. Aynı zamanda teknolojik altyapıya sahip teknik bilgi gerektiren bir iştir. Bu nedenle, adli bilişim sürecinde ve uygulamalarında çeşitli sorunlarla ve zorluklarla karşılaşılmaktadır. Bu zorluklar arasında, bilişim suçlarında işlenen veri miktarının fazlalığı, teknolojik altyapıların yetersizliği ve bilgi teknolojilerindeki hızlı değişimlere karşı hukuki düzenlemelerin gecikmesi sayılabilir.

Adli bilişim suçlarıyla ilgili karşılaşılan sorunlara ilişkin adli bilişim uzmanlarıyla yapılan görüşmeler sonucunda, bilişim suçlarında delil niteliği taşıyan verilerin çoğunlukla gelişmiş şifreleme algoritmalarıyla korunduğu, bu algoritmaların sürekli olarak kendini güncellediği ve teknolojik sistemlerin belirli bir ekosistem içinde entegre çalıştığı anlaşılmıştır. Bu durum, delil niteliği taşıyan verilerin farklı cihazlar aracılığıyla kısmen ya da tamamen silinmesine ve bu verilerin elde edilmesini zorlaştıran ciddi sorunların ortaya çıkmasına neden olmaktadır. Adli bilişim uzmanları, meydana gelen bilişim suçlarında kullanılan şifreleme algoritmalarının çoğunlukla genç ve eğitim düzeyi ilköğretim seviyesinin üzerinde olan bireyler tarafından tercih edildiğini tespit etmişlerdir. Bu nedenle, siber suçlara yönelik farkındalık ve bilinçlendirme seminerlerinin planlanmasında önceliği evli ve genç nüfusa verdiklerini ifade etmişlerdir. Bilişim Suçlarının işlenme oranlarındaki artışın sebepleri arasında işsizlik, gelir seviyesi, yaş, cinsiyet ve eğitim durumu gibi etkenler büyük bir etkiye sahiptir. Adli bilişim alanında gerçekleştirilen çalışmalarda, yapay zekânın bir alt dalı olan makine öğrenmesi ve derin öğrenme yöntemlerinden faydalanılmıştır. Son beş yıl içinde bu alanda yapılmış olan çalışmalar literatür araştırmasında sunulmuştur.

Bilişim Suçları ve psikolojik etkileri açısından Türkiye'de telefon dolandırıcılığının analizine ilişkin yapılan çalışmada; dijital teknolojilerin hızla gelişmesiyle birlikte kişilere ait bilgiler elektronik ortamlarda bulunabilir ve erişilebilir duruma geldiği belirtilmiştir. Bu durumun dolandırıcılık olaylarının meydana gelmesinin en büyük sebeplerinden bir tanesi olduğu öne sürülmüştür. Bilişim dünyasında dolandırıcıların kendilerini mağdurları ikna etmek için kamu ve özel sektöre ait çeşitli meslek gruplarından biri gibi tanıttıkları, İnsanların zaafları ve korkularından faydalanarak maddi ve manevi zarar verdiklerini gözlemlemişlerdir. Teknolojik yeniliklerin artmasıyla birlikte meydana gelen Bilişim Suçları çeşitliğinin de arttığı ancak hukuki altyapının aynı hızda değişkenlik gösterememesinden dolayı dolandırıcıların hukuki boşlukları kullandıklarını tespit etmişlerdir. Çalışmada, veri madenciliği kullanılarak ülkemizde meydana gelen telefon dolandırıcılığı olayları incelenmiş ve meydana gelen olaylar arasında ilişkiler ve bağlantılar ortaya konmuştur. Kullanılan verilerde 63 polis, 33 savcı, 19 pazarlamacı, 2 asker, 12 MIT personeli, 20 banka çalışanı ve 3 sigortacı telefon dolandırıcılığı ile ilişkilendirilmiş, bunların cinsiyet dağılımları ise 123 erkek ve 29 kadın olarak incelenmiştir. Dolandırıcılar tarafından hedef alınan meslek gruplarına göre kişilerin dağılımları ise şu şekildedir: 40 özel sektör çalışanı, 34 kamu çalışanı, 29 işsiz, 38 emekli, 5 çocuk ve 6 öğrenci olarak belirlenmiştir.

Son olarak, bu dolandırıcılık faaliyetlerinde dolandırıcıların elde ettikleri başarı oranları 94 kez başarılı ve 58 kez başarısız olarak değerlendirilmiştir (Tekkanat vd. 2018). Bu çalışmanın bulguları, dolandırıcılık vakalarının önemli bir bölümünde mağdurların kişisel zaafları, maddi menfaat beklentileri ve bilinçsiz karar mekanizmaları doğrultusunda kendi istekleriyle kişisel bilgilerini dolandırıcılara sağladıklarını ortaya koymaktadır. Ayrıca, hukuki düzenlemelerdeki yetersizlikler ve hızla gelişen teknolojilerin bireysel yaşamda sağladığı kolaylıkların, özel bilgilerin dijital ortamda korunmasını zorlaştırdığı tespit edilmiştir. Bu durum, dijital güvenlik alanında ciddi bir açık yaratarak, kişisel verilerin kötü niyetli aktörler tarafından ele geçirilme riskini artırmaktadır. Araştırma kapsamında elde edilen veriler, yaşlı bireylerin genç nüfusa kıyasla dolandırıcılık mağduru olma ihtimallerinin daha yüksek olduğunu göstermektedir. Bu durum, dijital okuryazarlık düzeyinin düşük olması, teknolojik gelişmelere adaptasyon sürecinin daha uzun sürmesi ve güvenlik önlemlerine ilişkin farkındalığın yeterli düzeyde

olmaması gibi faktörlerle ilişkilendirilmektedir. Bunun yanı sıra, teknolojik yeniliklerin güvenlik açıklarını da beraberinde getirdiği ve bireylerin bu riskler karşısında bilinçsiz veya tedbirsiz davranmalarının, dolandırıcılık eylemlerini kolaylaştırdığı görülmektedir. Özellikle, siber güvenlik konularında yeterli bilgiye sahip olmayan ya da gerekli önlemleri almayan bireylerin, dolandırıcılar tarafından manipüle edilme olasılığının daha yüksek olduğu anlaşılmıştır. Bu bağlamda, bireysel farkındalığın artırılmasına yönelik eğitim politikalarının geliştirilmesi ve dijital güvenlik önlemlerinin güçlendirilmesi, dolandırıcılık vakalarının azaltılması açısından kritik bir gereklilik olarak değerlendirilmektedir.

Nüfus yoğunluğunun çok fazla olduğu yerlerde meydana gelen; olayların tespiti, nesne tespiti, olaylardaki şiddetin derecesinin tespiti ve kalabalık alanların analizi konularında çalışma yapılmıştır. Analizin yapılmasında Rastgele Orman (RO), Destek Vektör Makineleri (DVM), Çok Katmanlı Yapay Sinir Ağları (Çok Katmanlı YSA), K-En Yakın Komşular (IBK) algoritmalarından faydalanılmıştır. Çalışmalar kapsamında kullanılan analiz ve video tespitleri uygulamalarının birbirlerine engel olduğu, hava şartları gibi etkenlerin de hesaba katılması durumunda yapılan çalışmaların büyük çoğunluğunda görüntü verilerinin analiz edilmesinde yetersiz kalındığı gözlemlenmiştir. Video analizinde kullanılmakta olan uygulamaların gerçek dünya şartlarının da ele alınmasıyla birlikte çözüm ürettiği, meydana gelen sebeplerin yanında aksi durumların hepsinin oluşması durumunda da tespiti yapılacak problemlerin çözümünü yapabilecek bir analiz uygulamasının olmadığı belirtilmiştir (Sreenu ve Durai 2019).

Dijital görüntü manipülasyonunun tespiti tekniklerine ilişkin yapılan çalışmada; Teknolojilerin gelişmesiyle birlikte çok sayıda dijital fotoğrafların oluşturduğu verilerin meydana geldiği belirtilmiştir. Bununla birlikte mevcutta bulunan fotoğraf düzenleyici programlar sayesinde görüntüler üzerinde değişiklik yapmak çok basit hale geldiği gözlemlenmiştir. Tahrip edilmiş görüntüler suçlular tarafından insanları yanıltmak ya da meydana gelen suçlarla ilgili soruşturmaların akıbetini değiştirmek amacıyla kullanıldığı tespit edilmiştir. Görüntüler üzerinde çok kez oynandıktan sonra dijital veri üzerinde bıraktığı izlerin yok edilmesi, orijinal görüntü verilerinin ortaya konulma süreci çok zorlaştırdığını gözlemlemişlerdir. Bu çalışma ile görüntüler üzerinde yapılan manipülasyonların tespitini yapmak amacıyla derin öğrenme yöntemleri kullanılmış, görüntüler üzerinde yapılabilecek çeşitli görüntü sahteciliğini ve çeşitli görüntü manipülasyonu tespitinin yapılması ele alınmıştır. Çalışma kapsamında ImageNet ile farklı derin öğrenme algoritmaları kullanılmıştır. Görüntüler üzerinde yapılan manipülasyon ve sahteciliğin tespit edilmesinde kullanılabilecek yazılım programlarının oluşturulmasının zor ve zaman aldığı belirtilmiştir. Görüntüler üzerindeki manipülasyon yöntemlerinin sürekli değişip geliştiği, bu değişikliğe ve gelişmeye paralel olarak meydana gelen sahteciliklerin ortaya konulmasını sağlayacak sistem ve yazılımların geliştirilmesi gerektiği ortaya konulmuştur (Thakur ve Rohilla 2020).

Video ve görüntü analizi ile ilgili yapılan çalışmada; Gelişmiş teknolojiler tüm insanlığa fayda sağlamış, insanlığın büyük buluşlar yapmasına olanak sağladığı gözlemlenmiştir. Bunlara örnek olarak görüntü ve video çekebilen üst düzey mobil cihazlar , dijital kameraların bulunduğu ancak bu teknolojilerin kullanımı da beraberinde büyük riskler getirdiği tespit edilmiştir. Mobil cihazlar ve dijital kameralarda depolanan görüntü ve videolar genellikle insanlar tarafından sosyal medya platformlarında paylaşılmakta bu durum her türlü manipülasyona açık olduğundan sosyal medya kullanıcılarını zor durumda bırakılabileceği değerlendirilmiştir. Bu çalışmayla sosyal medya platformlarında yer alan görüntü ve videoların hukuki süreçte delil olarak kullanılması muhtemel 3GP, MOV ve MP4 formatına sahip görsellerin, orijinal görsellik yapıları bozularak görüntü ve videolar üzerinde oynandığını ortaya koyabilen bir yöntem önerilmiştir. Görüntü ve videolar üzerinde yapılan düzenlemeler ile arkada bıraktığı izler makine öğrenmesi yöntemleri ile analiz edilerek hangi görüntü ve videolar üzerinde düzenlemenin yapıldığı ortaya çıkarılmaya çalışılmış, Çalışmada 4979 adet orijinali ile oynanmış video görüntüsü kullanılmıştır. Sonuçların değerlendirilmesinde RO modelinin en iyi başarıyı gösterdiği tespit edilmiştir. Hukuki süreç içerisinde delil niteliği taşıyan görüntü ve videoların üzerinde oynanması nedeniyle masum kişiler suçlu durumuna sokulabilir iken bir suçluyu da tüm suç unsurlarından arındırabileceği gözlemlenmiştir. Bu nedenle tüm bu hukuki süreçlerde doğru karar verebilmek adına adli tıp biliminin geliştirilmesi gerektiği önerilmiştir (Orozco vd. 2020)**.**

Makine öğrenmesi algoritmaları kullanılarak siber suçları analiz etmek ve sınıflandırılması için yapılan bu çalışmada; Hindistan'da faaliyet gösteren çeşitli resmi kaynaklardan toplanmış gerçek siber suç vakalarına ait kayıtlar kullanılmıştır. Veri seti hem yapılandırılmış (suç tipi, yer, tarih vb.) hem de yapılandırılmamış (suç açıklamaları, yazılı raporlar) veri türlerini içermektedir. Veri seti dolandırıcılık, kimlik hırsızlığı, yasa dışı erişim vb. suç türlerini temsil eden örneklerden oluşmaktadır. Kullanılan veri setinde veri temizleme, öznitelik seçimi ve

normalizasyon gibi çeşitli ön işleme adımları uygulanmıştır. Naive Bayes, Karar Ağaçları, DVM, RO, IBK ve K-Means gibi hem denetimli hem de denetimsiz öğrenme algoritmaları kullanılarak suç türleri sınıflandırılmıştır. Modeller çapraz doğrulama yöntemiyle test edilmiş ve özellikle DVM algoritmalarının %99'a varan doğruluk oranlarına ulaşmıştır. Çalışmadan elde edilen sonuçlar, geliştirilen sistemin siber suç analizinde güvenilir ve etkili bir araç olabileceğini göstermekte; bu doğrultuda, kolluk kuvvetlerine ve güvenlik uzmanlarına karar destek sağlayacak bir altyapı sunduğu belirtilmektedir (Chandra vd. 2020).

Siber şuçların tespit edilmesi ve sınıflandırmak için yapılan bir diğer çalışmada; sosyal medya platformu Twitter üzerinden toplanan veriler aracılığıyla Hindistan'daki özellikle Jamtara gibi bilinen suç bölgelerinde siber suç faaliyetlerinin mekânsal analizini gerçekleştirmeyi amaçlamaktadır. Çalışmada en yüksek doğruluk oranı %96,82 ile sigmoid çekirdek fonksiyonundan elde edilmiştir (Mahor vd. 2021).

Makine öğrenmesi algoritmaları ile terör olaylarının tahmin edilmesi ile ilgili yapılan çalışmada; Günümüzün en önemli tehditlerinden bir tanesinin terörizm olduğu, ayrıca terörizmin tüm dünyada iş gücü ve güvenlik harcamalarını artırarak ülkelerin eğitim ve sağlık gibi alanlara aktarması gereken fonları terörizm faaliyetlerinden dolayı güvenliğe aktarmak zorunda kaldıkları belirtilmiştir. Tüm bunların yanı sıra terörizmin çok sayıda can kaybına sebep olduğundan dolayı insanoğlunun sosyal yaşamlarını olumsuz yönde etkilediği gözlemlenmiştir. Teknolojilerin gelişmesiyle birlikte bilgiye ulaşılması kolaylaşmış, bu sayede ülkeler için öngörülebilirlik derecelerinin arttığı, stratejik karar mercilerinin, güvenlik ve istihbarat gibi alanlarda karar almalarının kolaylaştığı değerlendirilmiştir. Bu çalışmada meydana gelen terör olaylarının hangi terörist grubu veya grupları tarafından yapıldığı ile ilgili tahminde bulunması ve incelemelerin detaylandırılması için makine öğrenmesi yöntemleri ile yeni modeller oluşturulmuştur. Sonuçların değerlendirilmesinde, YSA, RO, Lojistik Regresyon ve Karar Ağacı algoritmalarının uygun olduğu görülmüştür (Görmez 2021).

Adli Bilişim İnceleme süreçlerinde yapay zekâ kullanımı ile ilgili yapılan çalışmada; teknolojilerin hızla gelişmesine bağlı olarak bilgiyi depolayan ve kullanan bilişim cihazları ve uygulamaların kullanım oranlarının arttığı gözlemlenmiştir. Dijital cihazlarda depolanan verilerin artması suç oranlarının artmasına sebep olduğu tespit edilmiştir. Dijital dünyanın gelişmesiyle bilişim suçlarındaki veri hacminin büyümesiyle birlikte doğru analizini zorlaştırdığı anlaşılmıştır. İşlenen verilerin artması yargılama süreçlerini de olumsuz etkilediği belirtilmiştir. Çalışmada toplam 4000 veri seti kullanılmış olup, bunların 2000'i tabanca, 2000'i ise bıçak görüntülerinden oluşmaktadır. Elde edilen sonuçlara göre, geliştirilen model %97,8 doğruluk oranı ile başarılı bir performans sergilemiştir. Bu sonuç, derin öğrenme alanında gerçekleştirilen diğer sınıflandırma çalışmaları ile karşılaştırılarak değerlendirilmiştir (Dilber, 2022).

Ses incelemeleri için yapay zekâ ile cinsiyet modelinin geliştirilmesi ile ilgili yapılan çalışmada; analiz aşamalarında doğruluğun arttırılması ve insan müdahalesinin minimum indirgenmesi adına görsel verileri yüksek doğruluk oranı ile sınıflandırmayı sağlayan VGG16 ağ tabanlı bir model önerilmiştir. Suç kapsamında elde edildiği varsayılan veriler, TensorFlow ve Keras derin öğrenme kütüphaneleri desteği ile FloydHub geliştirme ortamında önerilen model ile teste tabi tutulmuş ve söz konusu veriler üzerinde sınıflandırma işlemleri gerçekleştirilmiştir. Sonuç olarak suçların ortaya çıkarılması aşamalarında yapılan çalışma bir etken olabilir ancak suç delillerinin ortaya konulmasında tek başına yeterli olmadığı anlaşılmıştır. Teknolojik ses cihazlarının gelişmesiyle birlikte insan hayatını her alanda kolaylaştıran ses verilerine rastlamakta olduğu, ses verileri ile işlenen suçların tespitinin yapılabilmesi için ses tanıma sistemlerinin geliştirilmesi gerektiği belirtilmiştir. Bu sistemler sayesinde cinsiyet, yaş ve insanların duygu durumlarına dair birçok özelliğin tespitinin sağlanmakta olduğu gözlemlenmiştir. Tespiti yapılan bu tür özellikler sayesinde konuşmacıların veya ses kayıtlarındaki kişilerin tüm özellikleri veri haline dönüştürülerek meydana gelen bilişim suçlarında delil niteliği taşıyabileceği aktarılmıştır. Bu çalışmada yapay zekada kullanılan sınıflandırma yöntemleri karşılaştırılmış toplanan veriler ile konuşmacının sesinden cinsiyet tespitinde kullanılabilecek bir model geliştirilmiştir (Demirel 2022).

Ses verileri; düşük maliyetli olması, veri girişinin kolaylığı ve uzaktan kolayca erişilebilmesi gibi konularda avantaj sağladığından dolayı bu konudaki çalışmaların artmasına sebep olduğu anlaşılmıştır. Modelde yaş ortalaması 28,6 olan kadın ve 37,4 olan erkeklerden Youtube üzerinden toplanan 170 ses verisi ile veri seti oluşturulmuş, bu veri seti kullanılarak cinsiyet sınıflandırmasında kullanılabilecek bir model geliştirilmiştir. Bilişim suçlarında yapay zekâ yöntemleri kullanılarak elde edilen ses delilleri ile kişilerin yaşı ve cinsiyeti gibi özelliklerinin tespit edilmesi çok önemli olduğu görülmüştür. Değişik yaş ve cinsiyet grupları arasında uygulanan modelin doğruluk oranlarında değişiklikler yaşandığı gözlemlenmiştir. Modelde makine öğrenmesi

modellerinden DVM kullanılmış, çalışmada doğruluk oranı %99,59 olarak hesaplandığı tespit edilmiştir (Demirel 2022).

Kablosuz ağlar üzerinden gerçekleştirilen siber tehditlerin analizi ile ilgili yapılan çalışmada; Teknolojilerin gelişmesiyle birlikte kullanmakta olduğumuz cihazlar siber suç oranlarının artmasıyla tehdit altında olduğu belirtilmiştir. Birbirleri ile etkileşim halinde bulunan cihazlardan herhangi birinin tehdit unsurlarına maruz kalması diğer etkileşim halindeki cihazları da tehlikeye düşürebildiği aktarılmıştır. Çalışmada kablosuz ağlar üzerinde, sistemleri tehdit edebilecek muhtemel tüm saldırılar gerçekleştirilmiş, sınıflandırma için Navie Bayes (NB), DVM, gradyan arttırma ve IBK algoritmaları kullanılmıştır. Sınıflandırma algoritmaları sonucunda veri setinin elde ettiği skorlar toplanmış, en iyi skoru 0.91'le gradyan arttırma algoritması verdiği gözlemlenmiştir. Uygulanan tüm farklı saldırılar ağlarda farklı sonuçlara yol açmakta olduğu, bununda elde edilen verilerin birbirinden farklı olmasını sağladığı gözlemlenmiştir. Uygulanan bazı saldırıların özelliklerinin belirlenemediği gözlemlenmiştir. Bu nedenle meydana gelebilecek saldırıların tüm özelliklerini belirleyebilecek bir yazılıma ihtiyaç olduğu ya da halihazırda performansı daha yüksek yazılımların kullanılması gerektiği tespit edilmiştir. Gerçekleştirilen siber saldırıların çeşitliğinin ve sayısının arttırılarak çalışmanın genişletilmesi, farklı makine öğrenmesi algoritmaları ile performansların artabileceği önerilmiştir. Hızla gelişmekte olan teknolojiyle birlikte siber suç çeşitliliği ve oranlarında artış meydana gelmekte olduğundan bu hususta bilimsel araştırmaların arttırılarak siber güvenlik konusunda bilinçli insan sayısının arttırılması sağlanması gerektiği aktarılmıştır (Kaçan 2023).

Belirtilen bu çalışmaların dışında bilgi güvenliği kapsamında siber saldırıların tespit edilmesinde de makine öğrenmesi yöntemleri kullanılmıştır. (Harman ve Cengiz 2024, Ay ve Yolaçan ,2022). Yapılmış olan bu çalışmada, 2020-2023 yılları arasında Yalova İli Cumhuriyet Başsavcılığı'na intikal etmiş bilişim suçlarının analizi yapılmıştır. Çalışmada kullanılan veri setinin oluşturulması sırasında bilişim suçlarıyla ilgili, Yalova İlinde meydana gelen bilişim suçu türleri, bilişim suçlarına müdahil olan kişilerin genel özellikleri, bilişim suçları tespitinin nasıl yapıldığı, meydana gelen suçların aydınlatılma aşamaları, bilişim suçlarının ortaya çıkartılmasında ne tür bilgilere ihtiyaç duyulduğu, bilişim suçları analizinde kullanılan yöntemler hakkında bilgi alınmıştır. Alınan bilgiler doğrultusunda olay tarihi, olay ilçesi, suç türü, müşteki(mağdur), şüpheli, cinsiyet, yaş, medeni hal, meslek, uyruk, öğrenim durumu özelliklerini de kapsayacak şekilde düzenlenen toplam 2500 adet veri, özel hayatın gizliliğini ihlal etmeyecek şekilde hazırlanarak Yalova Cumhuriyet Başsavcılığı tarafından çalışmada kullanılmak üzere verilmiştir.

Bu çalışmada RO, DVM, Çok Katmanlı YSA ve IBK makine öğrenmesi algoritmaları kullanılarak bilişim suçları analiz edilmiş ve yorumlanmaya çalışılmıştır. Elde edilen sonuçlar doğrultusunda, meydana gelmesi muhtemel bilişim suçu türleri ile bu olaylara müdahil olan kişi profillerinin tespiti yapılmaya çalışılmıştır.

Çalışmanın en önemli kazanımlarından ve katkılarından biri, gerçek veriler kullanarak Bilişim Suç türlerine göre alınması gereken önleyici tedbirlerin belirlenmesine yönelik yol gösterici nitelikte bulgular sunmasıdır.

Şekil 1. Uygulanan sistem

2. Materyal ve Metot

Bu çalışmada uygulanan sistem Şekil 1' gösterilmiştir. Çalışmada kullanılan veri seti Yalova ili sınırları içerisinde 2020-2023 yılları arasında meydana gelen bilişim suçlarını kapsamaktadır. Veri setinin oluşturulması safhasında Yalova Cumhuriyet Başsavcılığı siber suçlarla mücadele savcısı ile görüşülmüş meydana gelen bilişim suçları ile ilgili,

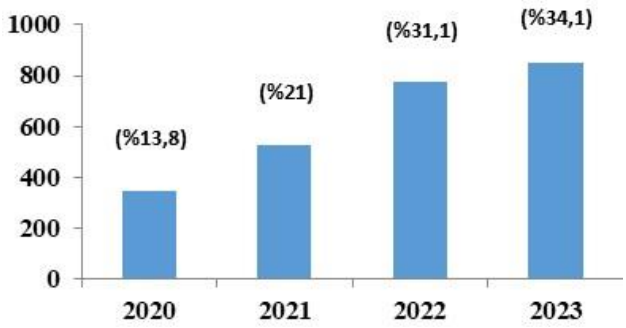
- Yalova ilinde meydana gelen bilişim suçu türleri,
- Bilişim suçlarına müdahil olan kişilerin genel özellikleri,
- Bilişim suçları tespitinin nasıl yapıldığı,
- Meydana gelen suçların aydınlatılma aşamaları,
- Bilişim suçlarının ortaya çıkartılmasında ne tür bilgilere ihtiyaç duyulduğu,
- Bilişim suçları analizinde kullanılan yöntemler hakkında bilgi alınmıştır.

Alınan bilgiler doğrultusunda çalışmada kullanılmak üzere belirlenen özellikler Tablo 1'de verilmiştir.

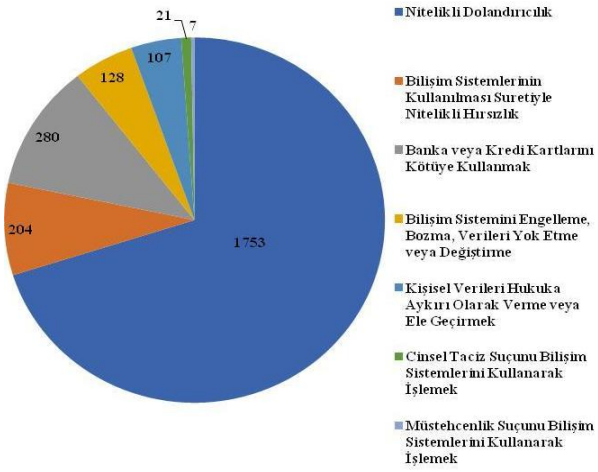
Çizelge 1. Veri setine ait özellikler.

1	Olay Tarihi
2	Olay İlçesi
3	Suç Türü
4	Müşteki (Mağdur)
5	Şüpheli
6	Cinsiyet
7	Yaş
8	Medeni Hal
9	Meslek
10	Uyruk
11	Öğrenim Durumu

Bilişim suçu olaylarının veri seti içerisindeki yıllara göre dağılımı incelendiğinde; 2020'de 346 (%13.8), 2021'de 524 (%21), 2022'de 778 (%31.1) ve 2023'te 852 (%34.1) olayın meydana geldiği Şekil 2'de gösterilmiştir.



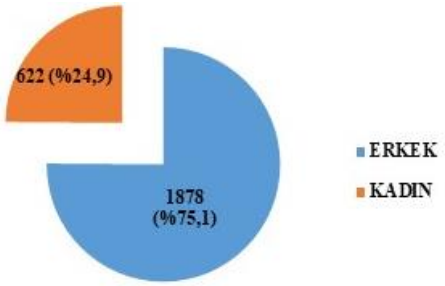
Şekil 2. Bilişim suçu olaylarının yıllara göre dağılımı



Şekil 3. Bilişim suçu türlerinin dağılımı

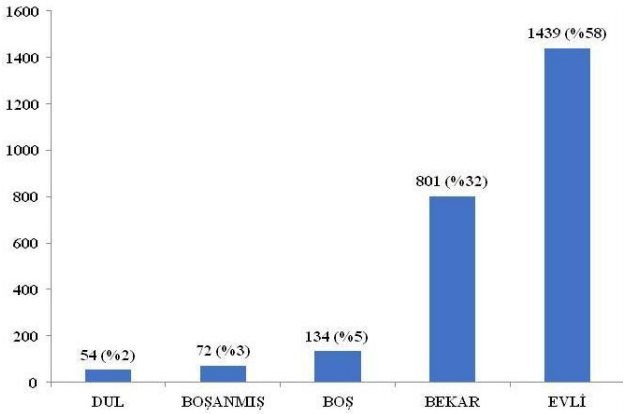
Bilişim suçu türlerinin dağılımına bakıldığında; nitelikli dolandırıcılık suçu oranı %70.12, bilişim sistemlerinin kullanılması suretiyle nitelikli hırsızlık suçu oranı %8.16, banka veya kredi kartlarını kötüye kullanmak suçu oranı %11.2, bilişim sistemini engelleme, bozma, verileri yok etme veya değiştirme suçu oranı %5.12, kişisel verileri hukuka aykırı olarak verme veya ele geçirmek suçu oranı %4.28, cinsel taciz suçunu bilişim sistemlerini kullanarak işlemek suçu oranı %0.84 ve müstehcenlik suçunu bilişim sistemlerini kullanarak işlemek suçu oranı %0.28 olduğu görülmüştür. Bilişim suçu türlerinin dağılımı Şekil 3'de verilmiştir.

Bilişim suçuna karışan kişilerin cinsiyet dağılımı Şekil 4'de verilmiştir.

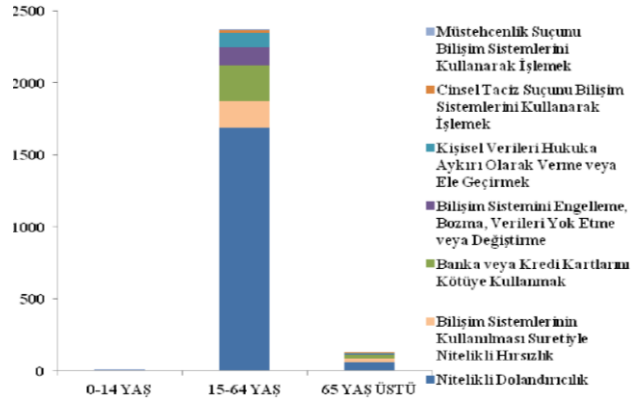


Şekil 4. Bişilim suçlarına karışmış kişilerin cinsiyet dağılımı ve oranı

Veri setinin medeni durumlarına göre dağılımları incelendiğinde; 1439 evli, 801 bekar, 134 medeni durumu belli olmayan (boş), 72 boşanmış ve 54 dul kişilerden oluştuğu tespit edilmiştir. Evlilerin veri seti içerisindeki oranı %58, bekarların %32 medeni durumu belli olmayanların (boş) %5, boşanmışların %3 ve dulların %2 olduğu anlaşılmış medeni durum dağılımları Şekil 5'de verilmiştir. Bilişim suçlarına karışan kişilerin suç türüne göre yaş dağılımı Şekil 6'da verilmiştir.

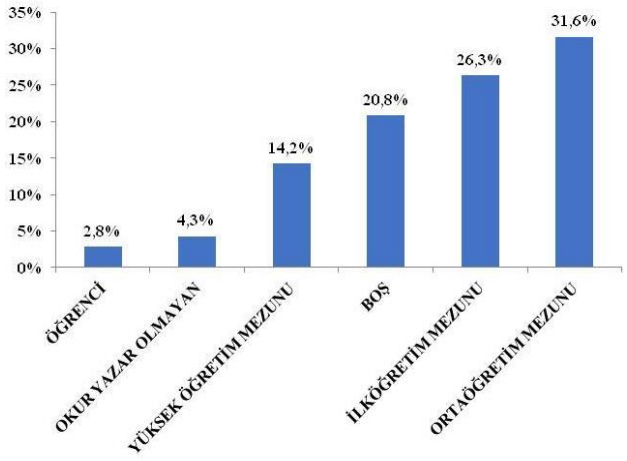


Şekil 5. Medeni durum dağılımı



Şekil 6. Bilişim suçu türlerinin yaş dağılımı

Bilişim suçlarına karışan kişilerin suç türlerine göre eğitim durumları Şekil 7'de verilmiştir.



Şekil 7. Bilişim suçu türlerine ait eğitim durumları dağılımı

2.1 Model Eğitimi

Makine öğrenmesi yöntemlerinden biri olan Rastgele Orman (RO) algoritması, birbirinden bağımsız olarak çalışan birçok karar ağacının bir araya gelerek en yüksek puanı alan değeri seçmesi ile çalışır. Kullanım kolaylığı ve esnekliği nedeniyle yaygın olarak benimsenmiştir. Algoritma, birden fazla karar ağacı oluşturarak sınıflandırma işlemi sırasında sınıflandırma doğruluğunu artırmayı amaçlar. Bu yöntem, veri kümesi üzerinde çeşitli modeller oluşturarak, kümenin yeniden ve daha ayrıntılı incelenmesini sağlar. RO, sınıflandırma ve regresyon problemleri kullanılan güçlü bir makine öğrenmesi algoritmasıdır. Birden fazla modelin birleşimini kullanarak en iyi bir performans elde etmeye çalışır. Algoritmada eğitim süreci tamamlandıktan sonra, her ağaç bağımsız olarak tahminler yapar. Sınıflandırma problemlerinde, her ağacın tahmin ettiği sınıf bir oy olarak değerlendirilir ve en çok oyu alan sınıf nihai tahmin sonucu olarak seçilir. Regresyon problemlerinde ise, her ağacın tahmin ettiği değerlerin ortalaması alınarak nihai tahmin sonucu elde edilir.

Destek Vektör Makineleri (DVM), sınıflandırma ve regresyon (sayısal tahmin) problemlerinde yaygın olarak kullanılan güçlü bir makine öğrenmesi algoritmasıdır. Vektörler yardımı ile verileri birbirinden ayırt etmemizi sağlar. İstatistiksel öğrenim teorisine dayanır. Temel amacı, veri noktalarını sınıflandırmak veya bir doğru veya eğri ile regresyon yapmaktır. Bir düzlemdeki noktaları ayırmak için bir doğru çizer ve bu doğrunun iki sınıfın noktalarına olan uzaklığını maksimum yapmayı hedefler. Bu yöntem, karmaşık ancak küçük ve orta ölçekli veri setleri için uygundur. Özellikle lineer olmayan ve ayrılamaz veri kümelerini sınıflandırmak için etkilidir. Bu algoritmanın amacı sınıfları birbirinden ayıran bir karar sınırı oluşturmaktır.

Çok katmanlı yapay sinir ağları (Çok Katmanlı-YSA), ileri beslemeli bir mimariye sahip olup, giriş katmanı, bir veya birden fazla gizli katman ve çıkış katmanından oluşur. Bu ağlar, geri yayılım algoritması kullanılarak eğitilir ve ağırlıklar hata fonksiyonunu minimize edecek şekilde güncellenir. Çok katmanlı sinir ağları, sınıflandırma, regresyon ve örüntü tanıma gibi birçok görevde etkin bir şekilde kullanılabilir. Özellikle derin öğrenme uygulamalarında temel yapı taşlarından biri olan bu algoritmalar, karmaşık ilişkileri öğrenme ve temsil etme kapasitesiyle öne çıkmaktadır.

IBK algoritması anlaşılması en kolay sınıflandırma algoritmalarından biridir. Geniş bir kullanım alanına sahiptir. Bu sınıflandırma algoritmasında, örnek verilerin veri özniteliklerinin sayısı olduğu n boyutlu bir uzayda çizilmektedir. n boyutlu uzaydaki her nokta, sınıf değeri ile etiketlenmektedir. Sınıflandırılmamış bir verinin sınıflandırılmasını keşfetmek için, n boyutlu uzayda çizilir ve en yakın k veri noktasının sınıf etiketleri not edilir. Genellikle k bir tek sayıdır. En yakın k veri noktası arasında maksimum sayıda oluşan sınıf, yeni veri noktasının sınıfı olarak alınır. Yani, karar k komşu noktanın oylanmasıyla verilmektedir.

Yapılmış olan çalışmada Makine öğrenmesi algoritmalarının parametre ayarlarında çeşitli değişiklikler ve denemeler yaparak elde edilen sonuçlar analiz edilmiştir. Çalışmanın bulguları, algoritmaların varsayılan (default) parametre değerleriyle en iyi performansın sağlandığını göstermektedir.

Çalışmada, sınıflandırma modellerinin başarısını ölçmek için kullanılan performans metriklerinden aşağıda bahsedilmiştir. Tablo 2'de gösterilen karışıklık matrisi kullanılarak çalışmada uygulanan modellerde yapılan hatalar ve doğruluk oranları hakkında bilgi vermektedir.

Veri seti (x ve y) eğitim ve test olarak ayrılmıştır. Train_test_split fonksiyonu kullanılarak x ve y veri kümesi rastgele seçilen veriler ile %80 eğitim ve %20 test verisi olarak bölünmüştür. Model oluşturma işlemi RO, DVM, çok katmanlı YSA ve IBK makine öğrenmesi algoritmaları için uygulanmış ve eğitim işlemi gerçekleştirilmiştir. Eğitim veri seti (x_train ve y_train) üzerinden eğitilmiştir. Eğitilen model kullanılarak test veri kümesi üzerinden tahminler yapılmıştır. Tahmin sonuçları ile kullanılan modellerin doğruluk değerleri hesaplanmış ve modeller arası karşılaştırma yapılabilmesi için gerekli performans metrikleri elde edilmiştir.

2.2 Performans metrikleri

Çizelge 2. Karışıklık matrisi

		TAHMİN EDİLEN DEĞERLER	
		Pozitif	Negatif
GERÇEK DEĞERLER	Pozitif	Doğru Pozitif (TP)	Yanlış Negatif (FN)
	Negatif	Yanlış Pozitif (FP)	Doğru Negatif (TN)

- Kesinlik; pozitif olarak tahmin edilen örneklerin toplam veri sayısına bölünmesidir.

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

- Duyarlılık; pozitif olarak tahmin edilmesi gereken verilerin oransal olarak kaçının doğru tahmin edildiğini göstermektedir.

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

- Doğruluk uygulanan modellerin tüm veri üzerinde ne kadar doğru tahmini yapıldığını göstermektedir.

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

- F1 skoru; kesinlik ve duyarlılık değerlerinin harmonik ortalamasının alınmasıdır. F1 skoru en yüksek 1, en düşük olarak da 0 değerini alır.

$$F_1 = 2 * \frac{Kesinlik*Duyarlılık}{Kesinlik+Duyarlılık}$$

3. Bulgular

2020-2023 yılları arası Yalova ilinde meydana gelmiş bilişim suçlarıyla ilgili olaylara karışan kişilerin; müşteki (mağdur) veya şüpheli olup olmadıkları durumu ile bilişim suçu türlerinden olan; banka veya kredi kartlarını kötüye kullanmak, bilişim sistemlerinin kullanılması suretiyle nitelikli hırsızlık, bilişim sistemini engelleme, bozma, verileri yok etme veya değiştirme, kişisel verileri hukuka aykırı olarak verme veya ele geçirmek ve nitelikli dolandırıcılık suçlarından hangisine müdahil oldukları durumu makine öğrenmesi algoritmalarından RO, DVM, Çok Katmanlı YSA ve IBK algoritmalarıyla analiz edilerek performansları test edilmiştir.

Çalışmada, sınıflandırma modellerinin başarısını ölçmek için; Kesinlik, duyarlılık, doğruluk ve F1 skor kullanılan performans metrikleri kullanılmıştır.

Veri seti içerisinde bulunan bilişim suçu türlerinden cinsel taciz suçunu bilişim sistemlerini kullanarak işlemek ve müstehcenlik suçunu bilişim sistemlerini kullanarak işlemek suçlarının veri sayısının az olmasından dolayı performans değerlerini etkilediğinden veri setinden çıkarılmıştır.

Yalova ilinde meydana gelmiş bilişim suçlarına karışan kişilerin, müşteki (mağdur) veya şüpheli olup olmadıkları durumu ile ilgili makine öğrenmesi algoritmalarıyla yapılan analizlerde; Müşteki (mağdur) özniteliğine ait DVM sınıflandırma algoritmasından Radyal Temel Fonksiyonunun doğruluk değeri 0.79, *Polinom'un* doğruluk değeri 0.69, Doğrusal'ın doğruluk değeri 0.80 ve Sigmoid 'in doğruluk değeri 0.83 olup en yüksek doğruluk değeri Sigmoid' de elde edilmiştir. Çok Katmanlı YSA algoritmasının doğruluk değeri 0.79, RO algoritmasının doğruluk değeri 0.83, ve son olarak IBK algoritmasının doğruluk değeri 0.79'dur. Sınıflandırma algoritmalarının performans değerleri karşılaştırması Çizelge 1'de verilmiştir. Çizelge 1 incelendiğinde en yüksek doğruluk değerleri DVM-Sigmoid ve RO algoritmalarında alınırken, en düşük doğruluk değeri DVM-Polinom algoritmasından alınmıştır.

Yalova ilinde meydana gelmiş bilişim suçlarına karışan kişilerin, bilişim suçu türlerinden;

- ✔ Banka veya kredi kartlarını kötüye kullanmak,
- ✔ Bilişim sistemlerinin kullanılması suretiyle nitelikli hırsızlık,
- ✔ Bilişim sistemini engelleme, bozma, verileri yok etme veya değiştirme,
- ✔ Kişisel verileri hukuka aykırı olarak verme veya ele geçirmek,
- ✔ Nitelikli dolandırıcılık suçlarından hangisine müdahil oldukları durumuyla ilgili makine öğrenmesi algoritmalarıyla yapılan analizlerde;

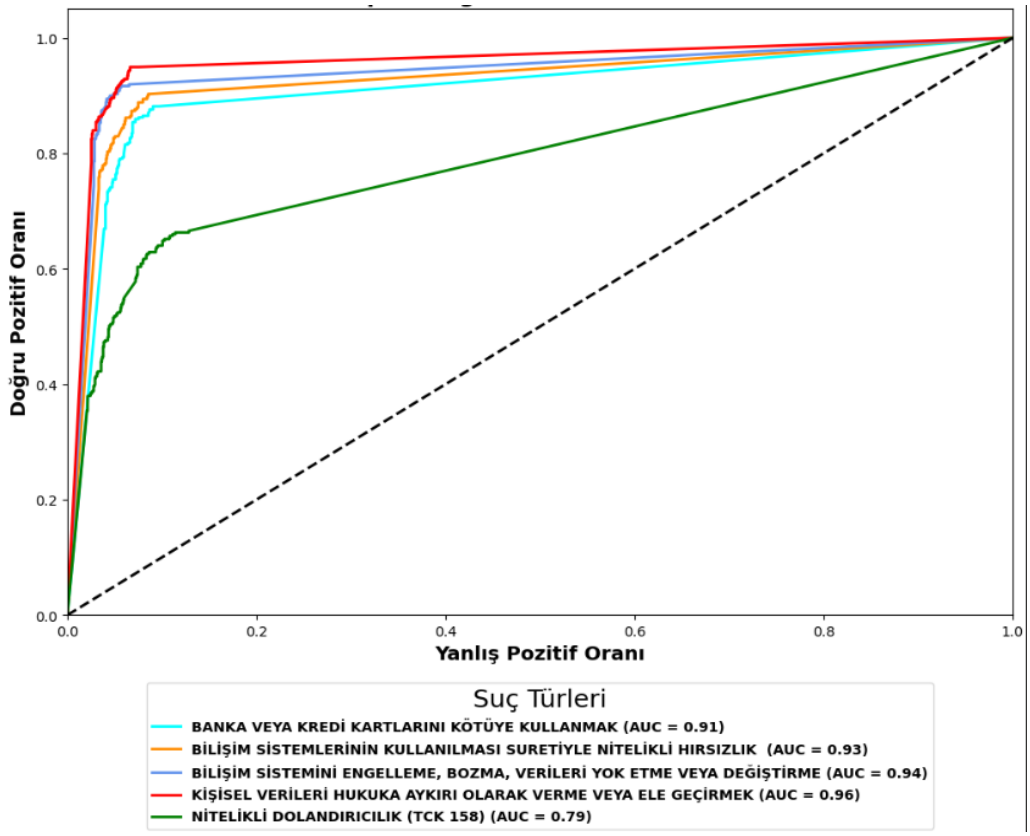
Bilişim suçu türlerine ait; DVM algoritmasının doğruluk değeri 0.82, Çok katmanlı YSA algoritmasının doğruluk değeri 0.82, RO algoritmasının doğruluk değeri 0.85 ve IBK algoritmasının doğruluk değeri 0.81'dir. En yüksek doğruluk değerleri RO algoritmasından alınırken, en düşük doğruluk değeri IBK algoritmasından alınmıştır. Makine öğrenmesi algoritmalarına ait diğer performans değerleri Çizelge 2 'de verilmiştir.

Çizelge 1. Kişilerin şüpheli veya müşteki olmasına göre sınıflandırma algoritmalarının performans değerleri karşılaştırması.

	Doğruluk	**Kesinlik**	**Duyarlılık**	**F1 Skor**
DVM Radyal Temel Fonksiyonu (RBF)	0.79	0.77	0.88	0.82
DVM- Polinom (Poly)	0.69	0.65	0.96	0.77
DVM- Doğrusal (Linear)	0.80	0.79	0.88	0.83
DVM- Sigmoid	0.83	0.82	0.88	0.85
Çok Katmanlı YSA	0.79	0.80	0.84	0.82
RO	0.83	0.86	0.84	0.85
IBK	0.79	0.81	0.82	0.82

Çizelge 2. Bilişim suç türlerine göre DVM, çok katmanlı YSA, IBK ve RO algoritması performans değerleri

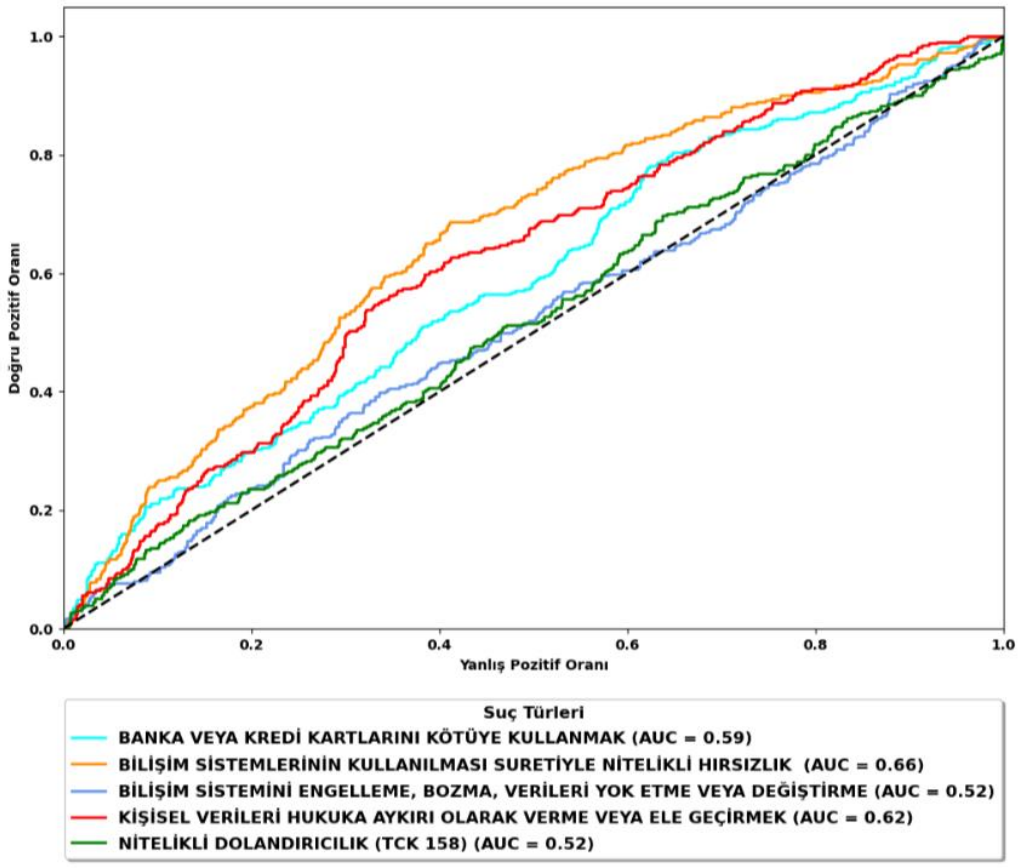
	Suç Türü	Kesinlik	Duyarlılık	F1 Skoru
DVM	Banka veya Kredi Kartlarını Kötüye Kullanmak	0.80	0.85	0.83
	Bilişim Sistemlerinin Kullanılması Suretiyle Nitelikli Hırsızlık	0.80	0.87	0.83
	Bilişim Sistemini Engelleme, Bozma, Verileri Yok Etme veya Değiştirme	0.88	0.89	0.88
	Kişisel Verileri Hukuka Aykırı Olarak Verme veya Ele Geçirmek	0.82	0.89	0.85
	Nitelikli Dolandırıcılık	0.77	0.58	0.66
Çok Katmanlı YSA	Banka veya Kredi Kartlarını Kötüye Kullanmak	0.82	0.79	0.81
	Bilişim Sistemlerinin Kullanılması Suretiyle Nitelikli Hırsızlık	0.84	0.86	0.85
	Bilişim Sistemini Engelleme, Bozma, Verileri Yok Etme veya Değiştirme	0.86	0.90	0.88
	Kişisel Verileri Hukuka Aykırı Olarak Verme veya Ele Geçirmek	0.85	0.94	0.89
	Nitelikli Dolandırıcılık	0.71	0.60	0.65
IBK	Banka veya Kredi Kartlarını Kötüye Kullanmak	0.77	0.87	0.82
	Bilişim Sistemlerinin Kullanılması Suretiyle Nitelikli Hırsızlık	0.78	0.87	0.82
	Bilişim Sistemini Engelleme, Bozma, Verileri Yok Etme veya Değiştirme	0.88	0.90	0.89
	Kişisel Verileri Hukuka Aykırı Olarak Verme veya Ele Geçirmek	0.85	0.87	0.86
	Nitelikli Dolandırıcılık	0.75	0.54	0.62
RO	Banka veya Kredi Kartlarını Kötüye Kullanmak	0.82	0.85	0.84
	Bilişim Sistemlerinin Kullanılması Suretiyle Nitelikli Hırsızlık	0.86	0.82	0.84
	Bilişim Sistemini Engelleme, Bozma, Verileri Yok Etme veya Değiştirme	0.90	0.90	0.90
	Kişisel Verileri Hukuka Aykırı Olarak Verme veya Ele Geçirmek	0.87	0.91	0.89
	Nitelikli Dolandırıcılık	0.80	0.78	0.79



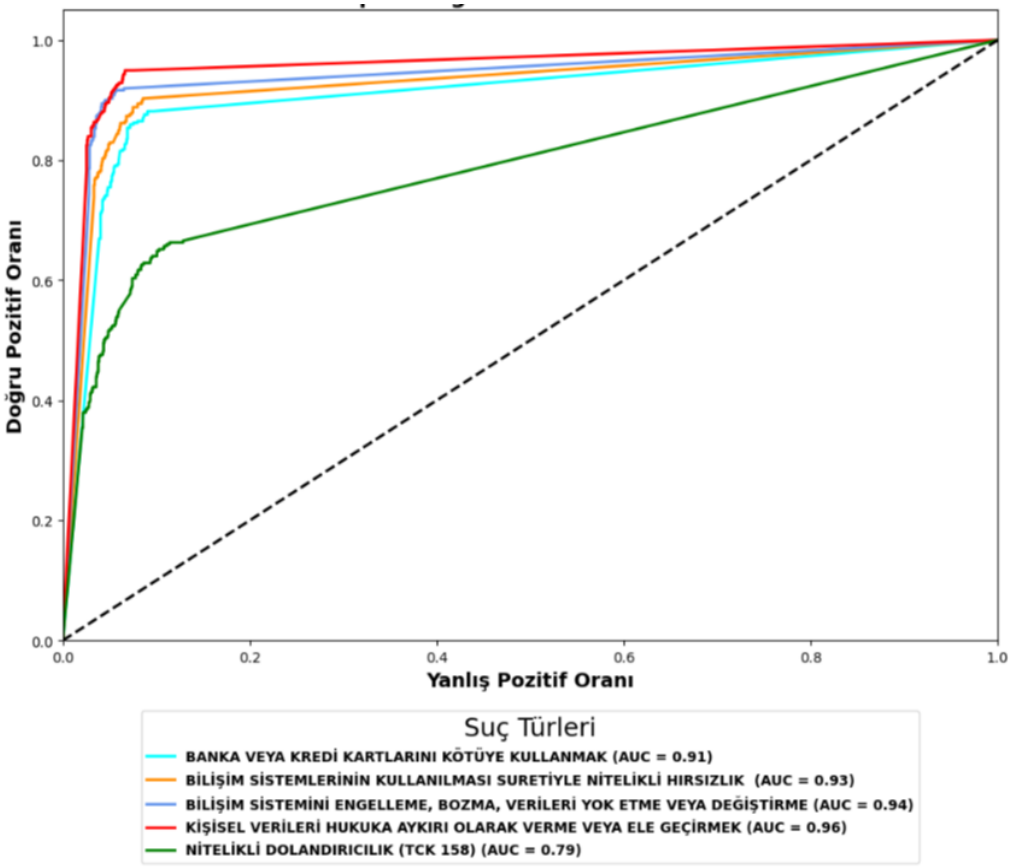
Şekil 8. Çok katmanlı YSA algoritması çoklu sınıf ROC eğrisi

Çok Katmanlı YSA algoritmasına ait çok sınıflı ROC eğrilerinde en yüksek performans (AUC) değeri 0.96 ile kişisel verileri hukuka aykırı olarak verme veya ele geçirmek suçundan alınırken, en düşük 0.79 ile nitelikli dolandırıcılık suçundan alınmıştır. Çok Katmanlı YSA algoritmasına ait çok sınıflı ROC eğrileri Şekil 8'de verilmiştir.

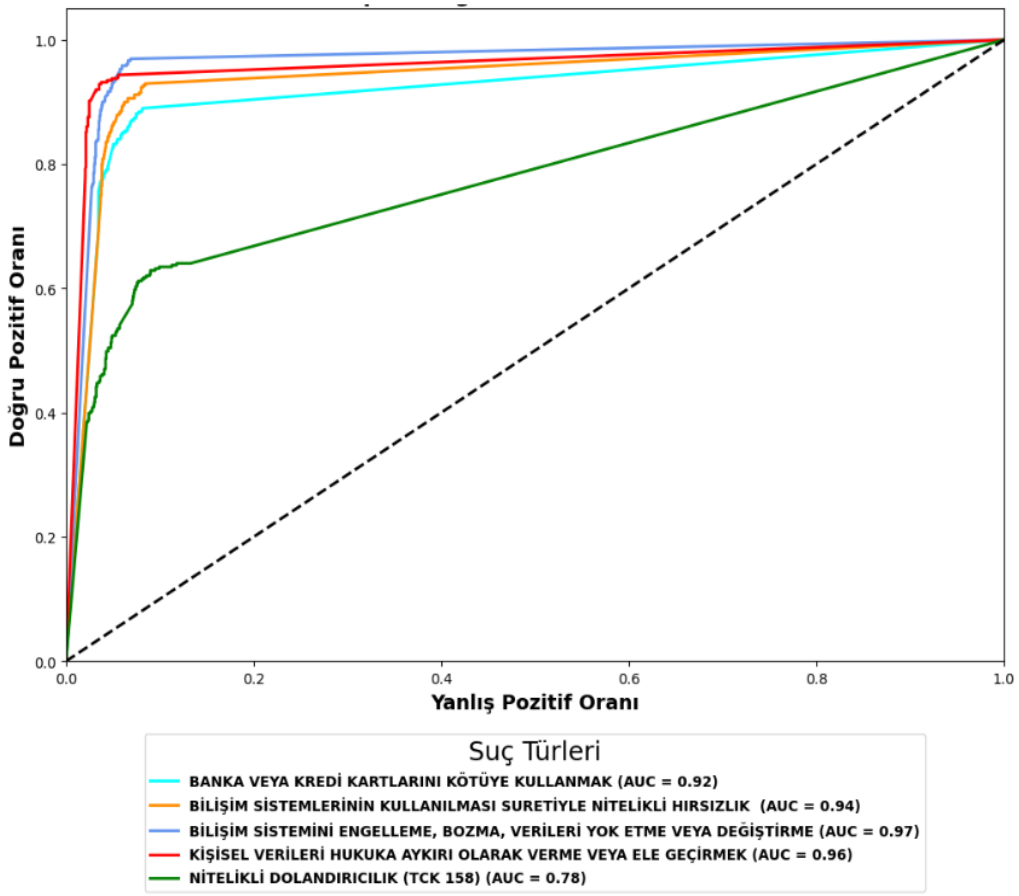
DVM algoritmasına ait çok sınıflı ROC eğrilerinde en yüksek performans (AUC) değeri 0.66 ile bilişim sistemlerinin kullanılması suretiyle nitelikli hırsızlık suçundan alınırken, en düşük 0.52 ile bilişim sistemini engelleme, bozma, verileri yok etme veya değiştirme ve nitelikli dolandırıcılık suçlarından alınmıştır. DVM algoritmasına ait çok sınıflı ROC eğrileri Şekil 9'da verilmiştir.



Şekil 9. DVM algoritması çoklu sınıf ROC eğrisi



Şekil 10. KNN algoritması çoklu sınıf ROC eğrisi



Şekil 11.RO algoritması çoklu sınıf ROC eğrisi

IBK algoritmasına ait çok sınıflı ROC eğrilerinde en yüksek performans (AUC) değeri 0.96 ile kişisel verileri hukuka aykırı olarak verme veya ele geçirmek suçundan alınırken, en düşük 0.79 ile nitelikli dolandırıcılık suçundan alınmıştır. IBK algoritmasına ait çok sınıflı ROC eğrileri Şekil 10.'de verilmiştir RO algoritmasına ait çok sınıflı ROC eğrilerinde en yüksek performans (AUC) değeri 0.97 ile bilişim sistemini engelleme, bozma, verileri yok etme veya değiştirme suçlarından alınırken, en düşük 0.78 ile nitelikli dolandırıcılık suçundan alınmıştır. RO algoritmasına ait çok sınıflı ROC eğrileri Şekil 11.'de verilmiştir

4. Sonuçlar ve Tartışma

Yapılmış olan çalışmada kullanılan makine öğrenmesi modellerin 2020-2023 yılları arasında Yalova ilinde meydana gelmiş bilişim suçları üzerindeki performansları karşılaştırmıştır. Bu analiz ile hangi modellerin suç tahmininde daha etkili olduğu ve modeller arasındaki farklılıkların neler olduğu tespit edilmiştir. Bu süreçte meydana gelmiş bilişim suçları ile mücadelede hangi makine öğrenmesi algoritmalarının daha etkili olduğunu belirlenmiştir.

Yalova'da meydana gelmiş bilişim suçu olaylarına karışan kişilerin müşteki (mağdur) mi şüpheli mi olduğuna dair yapılan analiz ile ilgili olarak Müşteki (mağdur) özniteliğine ait;

- ✔ En yüksek doğruluk değerleri 0.83 ile DVM-Sigmoid ve RO algoritmalarında alınırken, en düşük doğruluk değeri 0.69 ile DVM-Polinom algoritmasından alınmıştır.
- ✔ Kesinlik değerlerine bakıldığında en yüksek 0.86 ile RO algoritmasında alınırken, en düşük 0.65 ile DVM-Polinom algoritmasından alınmıştır.
- ✔ Duyarlılık değerlerine bakıldığında en yüksek 0.96 ile DVM-Polinom algoritmasından alınırken, en düşük 0.82 ile IBK algoritmasından alınmıştır.
- ✔ F1 skoru değerlerine bakıldığında en yüksek 0.85 ile DVM-Sigmoid ve RO algoritmalarından alınırken, en düşük 0.77 ile DVM-Polinom algoritmasından alınmıştır.

Yalova ilinde meydana gelmiş bilişim suçlarına karışan kişilerin, bilişim suçu türlerinden banka veya kredi kartlarını kötüye kullanmak, bilişim sistemlerinin kullanılması suretiyle nitelikli hırsızlık, bilişim sistemini engelleme, bozma, verileri yok etme veya değiştirme, kişisel verileri hukuka aykırı olarak verme veya ele geçirmek ve nitelikli dolandırıcılık suçlarından hangisine müdahil oldukları durumuyla ilgili olarak;

- ✔ En yüksek doğruluk değeri 0.85 ile RO algoritmasından alınırken, en düşük doğruluk değeri 0.81 ile IBK algoritmasından alınmıştır.
- ✔ En yüksek Kesinlik değeri 0.90 ile bilişim sistemini engelleme, bozma, verileri yok etme veya değiştirme suçuna ait RO algoritmasından alınırken, en düşük 0.71 ile nitelikli dolandırıcılık suçuna ait Çok Katmanlı YSA algoritmasından alınmıştır.
- ✔ En yüksek Duyarlılık değeri 0.94 ile kişisel verileri hukuka aykırı olarak verme veya ele geçirmek suçuna ait çok katmanlı YSA algoritmasından alınırken, en düşük 0.54 ile nitelikli dolandırıcılık suçuna ait IBK algoritmasından alınmıştır.
- ✔ En yüksek F1 skoru değeri 0.90 ile bilişim sistemini engelleme, bozma, verileri yok etme veya değiştirme suçuna ait RO algoritmasından alınırken, en düşük 0.62 ile nitelikli dolandırıcılık suçuna ait IBK algoritmasından alınmıştır.
- ✔ En yüksek Roc Eğrisi performans (AUC) değeri 0.97 ile bilişim sistemini engelleme, bozma, verileri yok etme veya değiştirme suçuna ait RO algoritmasından alınırken, en düşük 0.52 ile nitelikli dolandırıcılık suçuna ait DVM algoritmasından alınmıştır.

Bu çalışmada bilişim suçu olaylarına karışan kişilerin müşteki (mağdur) mi şüpheli mi durumuyla ilgili yapılan performans değerlendirmesi sonucu müşteki (mağdur) özniteliğine ait en yüksek doğruluk oranı 0.83 ile DVM-Sigmoid algoritmasından alınırken, bilişim suçu türlerine ait performans değerlendirmesi sonucu en yüksek doğruluk değeri 0.85 ile RO algoritmasından alınmıştır. RO algoritması içerisinde performans değerlendirmesi yapılan bilişim suçu türleriyle ilgili diğer performans değerlerine bakıldığında, en yüksek skor 0.90 ile bilişim sistemini engelleme, bozma, verileri yok etme veya değiştirme suçundan alınmıştır. Veri setinde *Nitelikli dolandırıcılık* suçunun veri içerisindeki oranı fazla olmasına rağmen elde edilen performans değerlerinin düşük olması veri setinde yer alan diğer değişkenlerin yüksek çeşitliliğe ve yapısal karmaşıklığa sahip olmasından kaynaklandığı değerlendirilmektedir. Nitelikli dolandırıcılığa özgü belirleyici özelliklerin, diğer suç türleriyle benzerlik göstermesi veya yeterince ayırt edici olmaması, bu suçun doğru şekilde tanımlanmasını ve analiz edilmesini zorlaştırmakta; dolayısıyla elde edilen sonuçların beklenen düzeyde olmamasına neden olmaktadır

Bu kapsamlı çalışma sayesinde, Yalova'da meydana gelen bilişim suçu olaylarının her yıl hızla artış gösterdiği gözlenmiştir. Olaylara müdahil olan kişilerin genel olarak evli, erkek, ortaöğretim mezunu ve genç nüfusa sahip oldukları, meydana gelen bilişim suçu olaylarının ilçelere göre dağılımına bakıldığında olayların genelde Yalova ili Merkez ilçesinde meydana geldiği görülmüştür. Olaylara müdahil olan kişi profillerinin etkin bir şekilde tahminin yapılmasına olanak tanınmıştır.

Bu tahminler gelecekte meydana gelebilecek suçların çözümü ve önlenmesi aşamalarında belirlenecek stratejilerin geliştirilmesine katkı sağlayacaktır. Kullanılmış olan makine öğrenmesi modellerinin başarılı bir performans göstermeleri yeni modellerin geliştirilmesi sürecinde önemli katkılar sunacaktır.

Yapılan çalışmada oluşturulan veri sayısı ve tespiti yapılan özellik sayısının arttırılması uygulanan makine öğrenmesi modellerinin performans değerlerini arttıracağı anlaşılmıştır. Ayrıca kullanılan makine öğrenmesi model sayılarının arttırılması da önerilmektedir.

Sonuç olarak, hızla gelişen teknolojilerle birlikte bilişim suçlarının çeşitliliği de artmaktadır. Bu durum elde edilen adli veri delillerini de arttırmakta adli mercilerin iş yükünü arttırmaktadır. Bu nedenle, gelişen teknolojiye paralel olarak gerekli hukuki altyapı oluşturulmalıdır. Bilişim suçlarını iyi analiz edebilen bilinçli ve eğitimli bireylerin sayısı artırılmalı, bu doğrultuda yapılacak eğitim ve bilimsel çalışmalar desteklenmelidir.

Etik Standartlar Bildirgesi
Yazarlar tüm etik standartlara uyduklarını beyan ederler.

Yazarlık Katkı Beyanı
Yazar 1: Kaynaklar, Araştırma, Deney, Yazma – orijinal taslak Görselleştirme, Yazma – orijinal taslak
Yazar 2: Kaynaklar, Araştırma, Deneyleme, Biçimsel analiz, Doğrulama, Metodoloji, Görselleştirme, Yazma – orijinal taslak,

Çıkar Çatışması Beyanı
Yazarların bu makalenin içeriğiyle ilgili olarak beyan edecekleri hiçbir çıkar çatışması yoktur.

Verilerin Kullanılabilirliği
Bu çalışma sırasında oluşturulan veya analiz edilen tüm veriler, yayınlanan bu makaleye dahil edilmiştir.

5. Kaynaklar

Ay, A. K., Yolaçan, E. (2022). Yeniden Örnekleme Metotlarının Kredi Kartı Sahtecilik Tespiti için Topluluk Öğrenmesine Kapsamlı Analizi. Afyon Kocatepe Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi, 22(5), 1005-1015. https://doi.org/10.35414/akufemubid.1066453

Bilgi Teknolojileri ve İletişim Kurumu (11 Haziran 2019). Bilişim Hukuku ve Bilişim Suçu. https://internet.btk.gov.tr/bilisim-hukuku-ve-bilisim-sucu

Chandra, R., Gadekallu, T. R., Abidi, M. H., & Al-Ahmari, A. (2020). Computational system to classify cyber crime offenses using machine learning. Sustainability, 12(10), 1–18. https://doi.org/10.3390/su12104087

Demirel, F.B. (2022). Adli Bilişim Ses İncelemeleri İçin Yapay Zeka Tabanlı Cinsiyet Tespiti Modeli Geliştirilmesi. [Yayımlanmış yüksek lisans tezi]. Fırat Üniversitesi.

Ermeydan, D. (2018). Türk ceza kanunu'nda bilişim suçları. (Master's thesis, Çağ Üniversitesi Sosyal Bilimler Enstitüsü).

Görmez, B. (2021). Adli Bişimde Makine Öğrenmesi:Makine Öğrenmesi Algoritmaları İle Terör Olaylarının Tahmin Edilmesi Çalışması. [Yayımlanmış yüksek lisans tezi]. Ankara Üniversitesi.

Harman, G., Cengiz, E. (2024). Deep Learning Based Network Intrusion Detection. Journal of Engineering Sciences and Design, 12(3), 517-530. https://doi.org/10.21923/jesd.1417622

Kaçan, İ. (2023). Kablosuz ağlar üzerinden gerçekleştirilen siber tehditlerin makine öğrenmesi yöntemleri ile ağ adli bilişim analizinin gerçekleştirilmesi. [Yayımlanmış yüksek lisans tezi]. Fırat Üniversitesi.

Mahor, V., Rawat, R., Telang, S., Garg, B., Palimkar,P. (2021). Machine learning-based detection of cyber crime Hub analysis using Twitter data. IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), Kuala Lumpur, Malaysia, 2021, pp. 1-5. https://doi.org/10.1109/GUCON50781.2021.9573736

Orozco, A.L.S., Huaman, C.Q., Alvarez, D.P., and Villalba, L.J.G. (2020). A machine learning forensics technique to detect post-processing in digital videos. Future Generation Computer Systems, 111, 199–212.

Sreenu, G. and Durai, M.A.S. (2019). Intelligent video surveillance: a review through deep learning techniques for crowd analysis. J Big Data

Tekkanat, E., Topaloğlu, M., & Yılmaz, O (2018) Bilişim Suçları ve Psikolojik Etkileri Açısından Türkiye'de Telefon Dolandırıcılığının Etkin Analizi. Ege Eğitim Teknolojileri Dergisi,2(2), 44-54

Thakur, R. and Rohilla, R. (2020). Recent advances in digital image manipulation detection techniques: A brief review. Forensic Science International, 312, 110311.

Yılmaz, E. N., Gönen, S., & Ulus, H. İ. (2016). Bilişim alanında işlenen suçlar üzerine bir inceleme. Bilişim Teknolojileri Dergisi, 9(3), 229.