



T.C.
KONYA TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**BANKACILIKTA UZAKTAN MÜŞTERİ
EDİNİMİNDE MÜŞTERİ KİMLİĞİNİN
DOĞRULANMASI VE CANLILIK TESPİTİNİN
MAKİNE ÖĞRENMESİ İLE
GERÇEKLEŞTİRİLMESİ**

Şeyma Nur KARAKAYA

YÜKSEK LİSANS

Bilgisayar Mühendisliği Anabilim Dalı

Ocak - 2023
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Şeyma Nur KARAKAYA tarafından hazırlanan “Bankacılıkta Uzaktan Müşteri Ediniminde Müşteri Kimliğinin Doğrulanması ve Canlılık Tespitinin Makine Öğrenmesi ile Gerçekleştirilmesi” adlı tez çalışması 20/01/2023 tarihinde aşağıdaki jüri tarafından oy birliği / ~~oy çokluğu~~ ile Konya Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

Başkan

Prof. Dr. Halife KODAZ

Danışman

Dr. Öğr. Üyesi Hazim İŞCAN

Üye

Dr. Öğr. Üyesi Ahmet ÖZKIŞ

İmza

.....

.....

.....

Yukarıdaki sonucu onaylarım.

Prof. Dr. Saadettin Erhan KESEN
Enstitü Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Şeyma Nur KARAKAYA

Tarih: 20.01.2023

ÖZET

BANKACILIKTA UZAKTAN MÜŞTERİ EDİNİMİNDE MÜŞTERİ KİMLİĞİNİN DOĞRULANMASI VE CANLILIK TESPİTİNİN MAKİNE ÖĞRENMESİ İLE GERÇEKLEŞTİRİLMESİ

Şeyma Nur KARAKAYA

**Konya Teknik Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı**

Danışman: Dr. Öğr. Üyesi Hazim İŞCAN

2023, 52 Sayfa

Jüri

**Prof. Dr. Halife KODAZ
Dr. Öğr. Üyesi Hazim İŞCAN
Dr. Öğr. Üyesi Ahmet ÖZKİŞ**

Uzaktan Müşteri Edinimi süreci, kişinin fiziksel olarak yapması gereken işlemlerinin (kişinin kendisinin şubeye gitmesi, ıslak imza, aslen kimliğinin gösterilmesi vb.) olmadığı, sürecin tamamen yönetmeliğe uyacak şekilde dijital ortamlar üzerinden gerçekleştirildiği bir süreçtir. Bu işleymde müşteri olmak isteyen kullanıcılar; kimlik kartı ve yüz tanıma teknolojileri ile zaman ve mekân fark etmeksizin banka müşterisi olabilmektedirler. Kolay ve güvenilir bir şekilde bir bankanın müşterisi olmayı sağlayan bu sistem, şubede bulunma ve evraklara bağımlı işlem mecburiyetini ortadan kaldırmaktadır. Bu çalışmada, Bankalardaki Uzaktan Müşteri Edinimi sürecinin, makine öğrenmesi kullanılarak yürütülmesi amaçlanmıştır. Böylece uzaktan kimlik tespiti için, tüm adımlar sonrasında yapılan müşteri ile müşteri temsilcisinin görüntülü görüşüp kişinin gerçekliğinin teyit edilmesine ihtiyaç duyulmayacaktır. Süreç otomatik olarak ilerletilip, müşterinin güvenilirliğine doğruluk analizi yapılarak karar verilecektir. Müşteri olma süreci için; Python dilinde PyQt kullanılarak bir uygulama kodlanmıştır. Bu uygulama üzerindeki kimlik tespit aşamalarında makine öğrenmesi algoritmalarından faydalanılmıştır. Aynı zamanda OCR, NLP ve OpenCV gibi teknolojiler kullanılarak doğruluk analizi yapılmıştır. Bu çalışmanın özellikle bankacılık ve finans alanında uzaktan müşteri edinimi süreçlerine sağlayacağı katkı değerlendirilmektedir.

Anahtar Kelimeler: Makine Öğrenmesi, NLP, OCR, OpenCV, Uzaktan Müşteri Edinimi

ABSTRACT

REALIZING WITH MACHINE LEARNING CUSTOMER IDENTITY VERIFICATION AND VITALITY DETECTION IN THE BANKING REMOTE CUSTOMER ACQUISITION

Şeyma Nur KARAKAYA

Konya Technical University
Institute of Graduate Studies
Department of Computer Engineering

Advisor: Assist. Prof. Dr. Hazim İŞCAN

2023, 52 Pages

Jury

Prof. Dr. Halife KODAZ
Assist. Prof. Dr. Hazim İŞCAN
Assist. Prof. Dr. Ahmet ÖZKİŞ

The Remote Customer Acquisition process is a process in which the person does not have any physical transactions (going to the branch, wet signature, showing her original identity, etc.). Users who want to be customers in this process; With ID card and face recognition technologies, they can become bank customers regardless of time and place. This system, which enables to be a customer of a bank in an easy and reliable way, eliminates the necessity of being in a branch and dependent on documents. In this study, it is aimed to carry out the Remote Customer Acquisition process in banks using machine learning. Thus, for remote identification, it will not be necessary to make a video call between the customer and the customer representative after all the steps and confirm the authenticity of the person. The process will be automatically advanced and the reliability of the customer will be decided by conducting an accuracy analysis. For the process of becoming a customer; An application is coded in Python using PyQt. Machine learning algorithms were used in the identification stages on this application. At the same time, accuracy analysis was performed using technologies such as OCR, NLP and OpenCV. The contribution of this study to remote customer acquisition processes, especially in the field of banking and finance, is evaluated.

Keywords: Digital Onboarding, Machine Learning, NLP, OCR, OpenCV

ÖNSÖZ

Yüksek Lisans eğitimime kendisinin destekleri ile başladığım ancak bugün aramızda olmayan saygı değer hocam Doç. Dr. Barış KOÇER'i saygı ve özlemle anıyorum.

Gösterdiği yollarla çalışmama farklı bir boyut kazandıran ve her türlü desteğini benden esirgemeyen kıymetli hocam Dr. Öğr. Üyesi Hazim İŞCAN' a minnettarlığımı ve teşekkürlerimi sunarım.

Değerli vaktini bana ayırarak her konuda desteğini benden esirgemeyen kıymetli hocam Doç. Dr. Abdullah Engin ÖZÇELİK'e teşekkür ederim.

Ayrıca gösterdikleri sabırlardan ve teşviklerinden dolayı, her zaman yanımda olan aileme sonsuz teşekkür ederim.

Şeyma Nur KARAKAYA
Tarih: 20.01.2023

İÇİNDEKİLER

ÖZET	iv
ABSTRACT.....	v
ÖNSÖZ	vi
İÇİNDEKİLER	vii
SİMGELER VE KISALTMALAR	viii
1. GİRİŞ.....	1
1.1. Tezin Amacı.....	2
1.2. Tezin Önemi	3
1.3. Tezin Organizasyonu	3
2. KAYNAK ARAŞTIRMASI	5
3. MATERYAL VE YÖNTEM.....	12
3.1. Makine Öğrenmesi Yöntemleri	12
3.1.1. Destek vektör makinesi.....	13
3.1.2. Lineer regresyon	14
3.1.3. Karar ağaçları.....	15
3.1.4. Yapay sinir ağları.....	16
3.1.5. Rastgele orman	18
3.2. Kimlik Tespit Yöntemleri.....	20
3.2.1. OCR (Optik karakter tanıma)	20
3.2.2. NLP (Doğal dil işleme).....	21
3.2.3. Biyometrik doğrulama	21
3.2.4. Güvenlik Sorusu ve Sesli Yanıt	22
3.3. Kütüphaneler ve Algoritmalar	23
3.3.1. TesseractOCR	23
3.3.2. FuzzyWuzzy	23
3.3.3. SpeechRecognition	24
3.3.4. OpenCV	25
3.3.5. FaceRecognition	25
3.3.6. CNN (Convolutional neural network)	26
4. UYGULAMA	27
4.1. Veri Kümesi.....	37
4.2. Veri Ön İşleme.....	38
5. SONUÇLAR VE ÖNERİLER.....	40
KAYNAKLAR	41

SİMGELER VE KISALTMALAR

Kısaltmalar

BDDK	: Bankacılık Düzenleme ve Denetleme Kurumu
TBMM	: Türkiye Büyük Millet Meclisi
NFC	: Near Field Communication (Yakın Alan İletişimi)
OCR	: Optical Character Recognition (Optik Karakter Tanıma)
NLP	: Natural Language Processing (Doğal Dil İşleme)
ML	: Machine Learning (Makine Öğrenmesi)
DVM	: Destek Vektör Makinesi
LR	: Logistik Regresyon
YSA	: Yapay Sinir Ağları
CNN	: Konvolüsyonel Sinir Ağları (Convolution Neural Network)



1. GİRİŞ

Covid19 dönemiyle birlikte ihtiyaç haline gelen dönüşümlerin başında bankaların müşteri edinimi süreçlerinin dijital ortama taşınması gelmektedir. Müşterilerin fiziksel olarak şubede bulunmasına ihtiyaç duyulmadan bankanın müşterisi olabilmesi ve bütün işlemlerinin elektronik platformda güvenilir bir şekilde gerçekleştirmesi için öncelikle finans şirketleri ve akabinde birçok kuruluştta bu teknolojiye geçiş sağlanmaya çalışılmaktadır. Bu konuda geçtiğimiz yıl adım atılarak, sonucunda finansal kuruluşların müşteri edinimi süreçlerinde “Kimlik Tespiti” ve “Sözleşme” işlemlerini dijital platformlar üzerinden gerçekleştirebilmesine ilişkin yasa teklifi TBMM’de kabul görmüştür. Bu kanunun kabul görmesi üzerine bankacılık ve finans sektöründe fiziksel olarak müşteri olmanın yanı sıra artık tamamıyla dijital ortam üzerinden de müşteri olunabilecektir.

BDDK tarafından 01.04.2021 tarihinde yayımlanan; Bankalarca Kullanılacak Uzaktan Kimlik Tespiti Yöntemlerine ve Elektronik Ortamda Sözleşme İlişkisinin Kurulmasına İlişkin Yönetmelik kapsamında bankalarca yeni müşteri kazanımında ve müşteri kimliğinin tespit edilmesinde kullanılacak yöntemlerle ilişki ayrıntılar duyurulmuştur (Ekonomi Haberleri, 2021).

Uzaktan Müşteri Edinimi işlemi, gerekli fiziksel işlemlerin (aslen şubede bulunma, sözleşmeye imza atma, fiziksel olarak kimlik gösterme vb.) olmadığı, işlemlerin tamamıyla yönetmeliğe uygun olacak biçimde çeşitli elektronik ortamlardan gerçekleştirildiği bir süreçtir (Resmî Gazete Sayı 31441, 2021). Her bankanın kendi uygulaması ile işlemler başlamakta, NFC (Yakın Alan İletişimi), OCR (Optik Karakter Tanıma) ve yüz doğrulama sistemleri kullanılmakta, sonrasında ise müşteri temsilcisi ile yapılan kısa bir görüntülü video görüşme sonrası süreç tamamlanmaktadır. Bu süreçte kullanıcılar; kimlik kartları ve yüz tanıma doğrulamalarıyla her zaman ve her yerde banka müşterisi olabilmektedirler. Kolay ve güvenli bir şekilde banka müşterisi olunabilen bu sistem, şube ve evrağa bağlı işlem zorunluluklarını da ortadan kaldırmaktadır.

Çizelge 3.1, “Bankalarca Kullanılacak Uzaktan Kimlik Tespiti Yöntemlerine ve Elektronik Ortamda Sözleşme İlişkisinin Kurulmasına İlişkin Yönetmelik” mevzuatına yönelik veriler, müşteri sayılarına ilişkin istatistiki bilgileri içermektedir. Mayıs 2021 döneminden itibaren aylık dönemlerde toplanan rapor, Kasım 2022 dönemi itibarıyla 16 banka verisinden oluşmaktadır (TBB, 2022).

Çizelge 3.1. Uzaktan ve Şubeden Müşteri Edinim İstatistikleri
(https://www.tbb.org.tr/Content/Upload/istatistikraporlar/ekler/3968/Uzaktan_ve_Subeden_Musteri_Edinimi_Istatistikleri-Kasim_2022.pdf)

	Uzaktan - Başvuru Sayısı	Uzaktan - Müşteri Temsilcisi ile Sonuçlandırılan Müşteri Sayısı	Toplu Edinim - Sonuçlandırılan Müşteri Sayısı*	Online Başvuru - Kurye ile Sonuçlandırılan Müşteri Sayısı*	Şubeden - Sonuçlandırılan Müşteri Sayısı*
Mayıs 2021	278.053	69.760	102.251	102.543	418.838
Haziran 2021	228.303	63.829	96.423	123.985	725.093
Temmuz 2021	234.948	61.092	95.373	111.949	550.785
Ağustos 2021	273.959	73.712	84.462	115.342	686.747
Eylül 2021	370.838	100.476	167.601	119.043	1.008.034
Ekim 2021	467.139	117.233	113.627	125.184	692.434
Kasım 2021	686.183	137.787	93.820	148.882	727.535
Aralık 2021	537.821	147.615	171.517	161.152	787.115
Ocak 2022	406.629	113.692	86.777	118.843	609.463
Şubat 2022	454.657	136.212	101.174	127.637	666.513
Mart 2022	733.258	244.015	115.054	133.284	787.038
Nisan 2022	1.409.652	331.771	118.889	160.849	735.912
Mayıs 2022	906.998	313.497	87.033	154.779	666.876
Haziran 2022	1.132.700	365.542	80.708	137.174	870.766
Temmuz 2022	742.869	247.234	63.993	113.100	839.418
Ağustos 2022	1.041.907	327.647	96.815	193.271	1.728.280
Eylül 2022	1.043.518	364.703	220.593	143.904	1.514.245
Ekim 2022	1.041.496	355.178	187.813	151.422	986.143
Kasım 2022	1.036.792	381.051	142.991	139.101	759.836

* Gerçek müşteri edinimleridir.

Kasım 2022 tarihinden geçmişe dönük kanun teklifinin yürürlüğe girdiği Mayıs 2021 boyunca uzaktan müşteri olma yolundaki başvuru sayısı yaklaşık olarak 1 milyon 37 bin olmuştur. Müşteri temsilcisi ile görüntülü görüşülerek onay alınan ve neticelendirilen uzaktan müşteri olan kişi sayısı 381 bindir. Kasım 2022 dönemi itibarıyla toplu müşteri olma, online müracaat ve şubelerden olunulan müşteri sayısı yaklaşık 1 milyon 42 bin olmuştur.

1.1. Tezin Amacı

Bu tez kapsamında, Bankalardaki Uzaktan Müşteri Edinimi'ndeki tüm sürecin (kimlik tespiti ve doğruluk analizinin) makine öğrenmesi yöntemleri kullanılarak yürütülmesi amaçlanmıştır. Böylece uzaktan kimlik tespiti için, tüm adımlar sonrasında yapılan müşteri ile müşteri temsilcisinin görüntülü görüşüp kişinin gerçekliğini teyit etme adımına ihtiyaç duyulmayacaktır. Süreç otomatik olarak ilerletilip, müşterinin güvenilirliğine doğruluk analizi yapılarak karar verilecektir. Sürecin tamamıyla otomatik yürütülmesiyle yüz tanıma ve canlılık tespiti için kullanılan teknolojiler ve kütüphanelerde makine öğrenmesi algoritmalarından yararlanılacaktır. Öncelikle mevcutta kullanılan NFC teknolojisinin yerine daha kolay ve güvenilir olan OCR Teknolojisi kullanılarak müşteri kimlik bilgileri alınacak, sonrasında müşterinin bir

önceki adımda doldurmuş olduğu bilgiler ile NLP tekniği kullanılarak karşılaştırma yapılacak ve eşleşme belirli bir oranı sağladığında, kişinin sesli yanıt verebileceği bir güvenlik sorusu sorulacak ve doğru cevap verildiği takdirde son yöntem olarak pasif canlılık analizi yani müşterinin kimlik kartından okunmuş olan biyometrik fotoğrafı ile kişinin anlık canlı bir görüntüsü alınıp karşılaştırılarak doğruluk tahminleme yapılacaktır. Her aşamada elde edilen doğruluk oranı bir sonraki adım için ön şart olarak kullanılacaktır.

1.2. Tezin Önemi

Uzaktan müşteri ediniminde kimlik tespit yöntemlerinin makine öğrenimi ile yapılması, bu yöntemlerin etkinliğini ve verimliliğini artırma potansiyeline sahiptir. Örneğin, makine öğrenmesi algoritmaları, müşterinin gerçek kimliğini doğrulamada kullanılan belge taramaları ve fotoğrafları gibi görsel verileri işleyebilir ve bu verileri doğrulama sürecine yardımcı olur. Ayrıca, makine öğrenmesi algoritmaları, müşteri görüşmeleri veya anketler gibi sözel verileri de işleyebilir ve müşterinin gerçek kimliğine ilişkin daha fazla bilgi toplayabilir. Makine öğrenmesi algoritmalarının kullanımının yanı sıra diğer kimlik tespit yöntemlerine de ihtiyaç duyulmuştur. Kimlik tespit yöntemlerinin seçiminde, çalışmanın amacı ve hedeflerine uygun olarak en uygun yöntemlerin seçilmesine dikkat edilmiştir.

Bu çalışmada, tamamen dijital bir platformda banka müşterisi olunacağı için bu süreci en verimli ve güvenilir bir şekilde gerçekleştiren bir yöntem ortaya konulması hedeflenmektedir. Müşteri gizliliği ve güvenliğine ilişkin endişeler de dikkate alınmıştır. Bu hedef doğrultusunda, Python programlama dilinde; hem makine öğrenmesi algoritmalarından hem de kimlik tespit aşamalarında kullanılan teknolojiler için güvenilir Python kütüphanelerinden faydalanılmıştır.

1.3. Tezin Organizasyonu

Tez, beş başlıktan meydana gelmektedir. İlk başlıkta, teze giriş kısmı yer almaktadır. Bankalarda kullanılan Uzaktan Müşteri Ediniminin bu anlamda tez kapsamında önerilen kimlik tespit yöntemlerinin içeriği hakkında bilgi verilmiştir. Tezin amacı ve önemine ilişkin yer alan bilgiler de bu bölüm içerisinde verilmiştir.

İkinci bölümde kaynak araştırması yapılmış ve benzer çalışmalar incelenmiştir. Kaynak araştırmasında, dijital bankalar, bankalardaki uzaktan müşteri edinimindeki

kimlik tespit yöntemlerinde kullanılan teknolojilere ilişkin ve makine öğrenmesi algoritmalarının uygulandığı çalışmalar incelenmiştir

Üçüncü bölümde; materyal ve yöntem başlığı adı altında, kimlik tespiti ve canlılık analizi hakkındaki bilgilerin tümü, kullanılan algoritmalar ve kütüphaneler sunulmuştur.

Dördüncü bölümde, incelenen ve üzerinde geliştirmeler yapılan yöneme ait uygulamaya ait aşamalardan bahsedilmiş ve uygulamada kullanılan algoritmaların analizleri sunulmuştur.

Son olarak beşinci bölümde ise tez kapsamında yapılan çalışmalar, algoritmalar ve analizlerden kazanılmış neticeler geniş bir çerçevede değerlendirilmiştir.



2. KAYNAK ARAŞTIRMASI

Finans sektöründe elektronik platformlara tamamen geçiş hızlı bir şekilde sürdürülmektedir. Daha iyi müşteri tecrübesi, genç neslin toplum üzerindeki sözünün geçirgenliğinin artışı, akıllı cihazların kullanımının sürekli artması, online işlem yapma ahlakı ile yapay zekâ ve büyük veri teknolojileri dijitalleşmeyi sürekli gündemde tutmaktadır. Tamamıyla dijital olan ve fiziksel düzeneklere ihtiyaç duyulmayan neobanklar da mobil bankacılıkta hızla yerini almakta ve fiziksel finans kuruluşlarına potansiyel rakipler olarak görülmektedirler. Finans kuruluşlarında uzaktan kimlik tespiti ve fiziki evrakların yerini dijital platformda mevzuat oluşturulmasına yönelik düzenlemeler gündem edilmiştir. Bu çalışmada mobil bankacılık macerası, sektörün evrimi ve mobil bankacılık kavramları incelenmiş ve fiziksel bankalarla ilişkisi açısından challenger bankalar derinlemesine incelenmiştir. Ardından neobank örnekleri incelenmiş ve geleceğin finans sektöründe neobankların yeri tartışılmıştır (Balkan, 2021).

Turan'ın yaptığı çalışmada, ilk olarak dijital bankacılık ve şubesiz bankacılığın faydalarından bahsedilmiş ve şubesiz bankacılıkta mobil şube üstünden videolu canlı görüşme ile hesap açma gibi finans işlemlerinin yapılması, internet şubeden müracaat edilerek hesap açma gibi müşteri kazanım yöntemleri anlatılmaktadır. Sürekli gelişen teknoloji, bankacılık sektörünü de etkileyerek yenilikçi kılmakta ve bankacılığa yeni bakış açıları kazandırmaktadır. Şubesiz veya minimum sayıda şube ile finansal hizmetlerin sağlanması şubesiz (dijital) bankacılık olarak bilinir. Bankacılığın sonsuza kadar süreceği, ancak artık banka şubelerine ihtiyaç olmayacağı anlayışıyla müşteri kazanımı için çeşitli yaklaşımlar geliştirilmiştir. Kurye ile sözleşme imzalama ve XTM cihazı ile hesap açma konuları incelenmiştir. Kuveyt Türk Katılım Bankası'nın bu metodolojileri kullanan örnek kanalları incelenmiştir. Tez kapsamında, mobil bankacılık sayesinde bankanın işlevsel verimliliğinin arttığı gözlemlenmiştir. Bu sistem ile yapılan işlem sürelerinin ciddi bir azalış sağlaması sebebiyle ihtiyaç duyulan eleman sayısının aşağı yönde, bankaya ait müşterilerin memnuniyetinin ise yukarı yönde ivme kazanacağı düşünülmektedir (Turan, 2022).

Goode'nin çalışmasında, bankalar, yeni müşterilerini daha iyi belirlemek, mevcut müşterilerin kimliklerini güvenli bir şekilde doğrulamak, yüksek değerli işlemleri korumak ve dolandırıcılıkla mücadele etmek için biyometrik teknolojinin giderek daha fazla benimsenmesi gerektiğinden bahsedilmektedir. Aslında, bankalar geleneksel fiziksel şubelerden en son dijital platformlara kadar her yerde biyometrik doğrulama

kullandığından, bu teknolojinin, tüm kanallarda bankacılık müşterilerinin kimliğini doğrulamak ve güvenliğini sağlamak için tek güvenilir araç olduğunu kanıtladığı anlatılmaktadır (Goode, 2018).

Dhawle ve arkadaşları, derin öğrenmenin bir parçası olan Python programlama dili ile OpenCV kullanarak insan yüzünü algılamak ve tanımak için ideal bir yol sunmaktadır. OpenCV yardımıyla, yüz kolayca algılanabilmekte ve benzer görünümdeki veri kümesinin eşleşmesi ile kişi analizi yapılmaktadır. Yüz tanıma teknolojisi, bir kişinin kimliğini belirleme tekniklerinden biridir. Bir insanın kimliği, kişinin kendisine ait bireysel yüzü kullanılarak tanımlanır. Çalışmada aynı zamanda gerçek zamanlı yüz tespiti yapılmasına yardımcı olacak bir sistem önerilmektedir. Geliştirilmiş olan bu sistem güçlü Python algoritması kullanmaktadır. Tez kapmasında geliştirilen uygulama çeşitli makinelerde, akıllı telefonlardaki platformlarda ve birçok yazılım uygulamalarında kullanılabilir. Bu yöntem askeriye, güvenlik, okullar, kolejler ve üniversiteler, havayolları, bankacılık, çevrimiçi web uygulamaları, oyun vb. kullanım alanlarında yüzün algılanması ve tanınması açısından birçok alanda kolaylık ve verimlilik sağlamaktadır (Dhawle ve ark., 2020).

Bibi ve arkadaşları, kimlik tespit yöntemlerinden biri olan el yazısı ile imza doğrulama teknolojisini incelemiştir. Biyometrik doğrulama sistemleri, çok sayıda uygulama alanında kullanılmaktadır. Genellikle bilimsel araştırmalara konu olan bu sistem kimlik tespiti gibi analizlerde önemli bir rol izlemektedir. Biyometrik doğrulama sistemlerinin birden fazla çeşidi olmakla birlikte bu çeşitler arasında bulunan özellikle el yazısı ile imza doğrulama son yıllarda büyük ilgi görmüştür. El yazısı imza doğrulaması, önceden kayıt altına alınmış verilerden orijinal imzayı ayıran davranışsal bir biyometrik doğrulama yöntemidir. Bankacılık, finansal, ticari işlemler, çek işleme, erişim kontrolü ve e-ticaret gibi birçok uygulama alanında bu konu incelenip sınıflandırma taksonomisi modeline göre çevrimdışı ve çevrimiçi imza doğrulama sistemlerinin teknikleri araştırılmıştır. İmza doğrulama sisteminin mevcut veri kümeleriyle birlikte ayrıntılı ve kapsamlı bir şekilde okuyucuya mevcut eğilimleri ve gelecekte oluşacak zorluklarla birlikte çeşitli yönleri sunulmuştur (Bibi ve ark., 2020).

Biggio ve arkadaşları, bir yüz tanıma sisteminde sahte bir yüzün bilinmeyen bir güvenlik açığından faydalanarak yüz şablonlarını tehlikeye atan bir saldırıya nasıl karşı koyulacağını örneklerle anlatmışlardır. Makine öğrenmesi alanında önerilen yeni bir özellik kapsamında biyometrik doğrulama sistemlerinin güvenlik açıkları ile ilgili çalışmalar işlenmektedir. Normal sistem işleyişini tehlikeye atmak için verileri manipüle

eden akıllı ve uyarlanabilir saldırganların bulmaya çalışmaktadır. Bu çalışma, aynı zamanda biyometrik tanıma sistemlerinin daha önce bilinen ve yeni güvenlik açıklarına karşı yapılan saldırılara karşı alınan önlemlerle birlikte bir savunma mekanizması geliştirmektedir. Bu geliştirilen savunma mekanizmaları ile bir sınıflandırma yapıp bu sınıflandırmanın detayları açıklanmıştır (Biggio ve ark., 2015).

Balasubramanian ve arkadaşları, NFC ve OCR teknolojisi kullanılarak COVID-19 semptomları olan hastaların hızlı bir şekilde teşhis edilmesini sağlayan ve durumu uzaktan izleyen akıllı bulut tabanlı bir yapılandırma aracını entegre eden, web tabanlı ilk temassız tele-sağlık klinik hizmetinden bahsetmişlerdir. COVID-19 olmuş hastanın sağlık çalışanlarına yayılmadan verimli bir şekilde uzaktan izlenmesi gerekmektedir. Uzaktan izleme ile karantina koğuşundaki sağlık çalışanlarının, hastanelerde bulunan sağlık profesyonelleri ile iletişim kurabileceği, etkin ve hızlı bir iletişim sistemi sağlanması için hastane ile karantina servisleri arasında temassız akıllı bulut tabanlı bir bağlantı kurulmuştur. Yapılan çalışma üç aşamadan oluşur: öncelikle NFC teknolojisi kullanılarak hastadan temassız sağlık parametresi çıkarılır, devamında OCR teknolojisi ile tıbbi raporun dijital metne dönüştürülmesi işlemi gerçekleştirilir ve NLP teknolojisi kullanılarak ilgili tıbbi parametrelerin değerlerinin çıkarılması sağlanır. Son olarak bu tıbbi parametreler görselleştirilir. Sistemin NFC okuyucusundan bulutta dağıtılan bir yapılandırma algoritması kullanılarak yapılan analize kadar olan doğruluğu %94 'ün üzerindedir. Önerilen web tabanlı sistemin çeşitli yetenekleri benzer sistemlerle karşılaştırılmış ve gerçek ya da sahte klinik kurulumları test edilmiştir ve doktorlar sistemin güvenilir ve kullanıcı dostu olduğu kanısına varmışlardır (Balasubramanian ve ark., 2022)

Koukoumidis ve arkadaşları, çağrı merkezlerinde yaygın sorunlardan biri olan yüksek çalışan devir hızını incelemişlerdir. Etkileşimi yumuşatmak ve müşteri deneyimini geliştirmek için tanıtılan Yapay Zekâ teknikleri, örneğin müşterinin karmaşık menü seçeneklerinde gezinmesini gerektirerek ters etkilere sahip olabilir. Bu açıklama, temsilci eğitimine ve müşteri çağrılarında uygulanan Yapay Zekâ tabanlı teknikleri açıklar. Teknikler, iletişim merkezlerindeki ciroyu azaltabilir ve müşteri hizmetleri temsilcileriyle etkileşime giren son kullanıcıların deneyimini iyileştirebilir. Tekniklere göre, insan müşteri temsilcilerini eğitmek için uygun Yapay Zekâ teknikleri uygulanır ve sırayla Yapay Zekâ tekniklerini eğitmek için insan geri bildirim kullanılır. Etkileşim deneyimini geliştirmek için müşterilerin iletişim tarzlarını yansıtmak için insan-Yapay Zekâ güçlendirmesi kullanılabilir. Teknikler, güvenliği artırmak için de kullanılabilir.

Örneğin, dolandırıcılık aramalarını otomatik olarak algılayarak kullanıcıları uyararak. Teknikler, ölçeklenebilir, bağımsız, yapay veya insan-Yapay Zekâ ile güçlendirilmiş müşteri hizmetleri araçlarının oluşturulmasını sağlar (Koukoumidis ve ark., 2021).

Bektaş ve arkadaşları, karakterlerin belirleyici ifadelerinden yararlanılarak herhangi bir fotoğraf uzantılı bir belgenin, sınıflandırma metodları yardımıyla karakterlerini en iyi doğruluk oranı ile tespit etmeyi hedeflemiştir. OCR teknolojisi, resim üzerindeki string ifadeleri belirleyip ayrıştırarak, harflerden kelimeleri, kelimelerden de cümleleri oluşturmaya çalışır ve bu sayede metnin ayıklanmasına olanak tanır. Sayısal ifadeler kullanılarak resim üzerindeki ifadelerin çözümlenmesi işlemi kâğıt üzerindeki metinlerin belgelenmesi için önem arz etmektedir. Karakterlerin hızlı ve doğru bir şekilde tahmin edilebilmesi için makine öğrenmesi algoritmalarından faydalanılmıştır ve bu alan geliştirilmeye açık bir araştırma konusudur. Bu nedenle farklı sınıflandırma yöntemlerinin performansları değerlendirilmiş ve bu yöntemlerin zaman aralığı kıyaslanıp en iyi yol belirlenmeye çalışılmıştır (Bektaş ve ark., 2016).

Liados ve arkadaşları, kimlik kartlarının ve pasaportların otomatik olarak okunması için bir uygulama olan ICAR sistemini açıklamaktadır. Sistem belgenin görüntüsünü düz yataklı bir tarayıcı ile alır ve renk bilgilerini kullanarak önceden tanımlanmış bir dizi model arasından belgenin türünü tanır. Metinsel alanlar, bağlantılı bir bileşen analizi ile görüntüde bulunur ve yapısal düzenlemeleri açısından tanımlanır. Bir dizi tamamlayıcı istatistiksel ve yapısal OCR tekniği, her metin görüntü bölgesini okumak için bir oylama stratejisiyle birleştirilir. Bilinmeyen giriş belgeleri için, makine tarafından okunabilen ICAO 9303 biçimiyle uyumlu satırlar bulunur ve tanınır. Sistem başlangıçta İspanyolca belgeler için tasarlanmış olsa da denetimli bir öğrenme prosedürü ile yeni biçimlerin entegrasyonuna izin verir (Liados ve ark., 2001).

Tskhadadze'nin çalışmasında, modern koşullarda bankacılık sektörünün gelişimi, uzaktan bankacılık müşteri hizmetleri sistemlerinin geliştirilmesi ve bankacılık hizmet düzeyini artırmak gibi önemli bir alan ele alınmaktadır. Bugün bir banka müşterisi, hesapları hakkında uzaktan bilgi alabilir ve hesaplarını yönetir, çeşitli banka işlemlerini gerçekleştirir. Bir banka için, herhangi bir uzaktan bankacılık hizmet sistemi, performans göstergelerinin büyümesi, maliyetlerin düşürülmesi, ürün kalitesinde azalma olmadan ve en önemlisi bankalar için önemli olan müşteri tabanını artırmak için rekabetçi pazardır (Tskhadadze, 2018).

Bankacılık dünyasının kendisinin günlük yaşamda önemli bir katkısı vardır. Ancak kullanılan prosedürlerin uygulanmasına bağlı olarak ortaya çıkan bazı eksikliklere halen sıklıkla rastlanmaktadır. Bu durum, halen manuel olarak yapılan müşteri öz-kayıt verisi sayısından görülebilir. Vatandaşlık Kartlarında bulunan Optik Karakter Tanıma (OCR) teknolojisi, doğruluk, hızlı ve en iyi okuma sonuçlarını almak için kullanılabilir. Bu çalışmanın amacı, Kimlik Kartı (KTP) okumada Optik Karakter Tanıma (OCR) teknolojisini kullanmaktır. OCR okumaları, renkli ve gri tonlamalı görüntülerdeki orijinal, orta ve küçük görüntülerin boyutunu karşılaştırmak için kullanılabilir. Böylece Optik Karakter Tanıma (OCR) Özellikli Kimlik Kartlarında kişisel verilerin işlenmesinde en iyi sonuçlar elde edilebilir. Dwi ve arkadaşlarının çalışmasında, gri tonlamalı Kimlik Kartı (KTP) verileri daha doğru okunmaktadır. Renkli Kimlik Kartı (KTP) doğruluğu için %86,32 ve gri tonlamalı Kimlik Kartı (KTP) doğruluğu için %88,58 olmuştur (Dwi ve ark., 2019).

Pires'in çalışmasında, banka kullanıcı profilinin çok değiştiği ve günümüzde mesai saatleri içinde eski müşteri hizmetleri modelinin düşünülmesinin mümkün bile olmadığından bahsedilmektedir. Müşteri, bırakın bankaya gitmeyi, üçüncü taraf yetkilendirmesine güvenmeden hesabına erişmek ve günün 24 saati işlem yapmak istemektedir. Dijital bir bankaya kaydolmanın hızlı ve kolay olması, ancak güvensiz olmaması amaçlanmıştır. Aksine, teknolojik gelişmeler sayesinde belge doğrulama ve profil analizi sürecinin her iki taraf için de kusursuz, şeffaf ve güvenilir bir şekilde yürütülmesi gerekmektedir. Digital Onboarding, tüm tüketici kayıt döngüsünü %100 elektronik hale getirir. Bir hesap açmak için daha fazla evrak, şirket tanıma ve makbuz gerekmez. Doğrudan insan etkileşimi, yalnızca ara sıra, kaydın daha ileri aşamalarında ve yeni bir müşterinin onaylanmasıyla ilgili bekleyen sorunlar veya şüpheler olduğunda ortaya çıkar. Müşterileri çekmek ve elde tutmak için verimli iletişim ile güvenli ortamlar yaratma zorluğu Fintech'lere kalmıştır. İdari Modernizasyon Ajansı (AMA), fiziksel ve elektronik güvenliği garanti etmeyi, dolandırıcılık riskini azaltmayı ve kişisel verilerin gizliliğini ve mahremiyet haklarını korumayı amaçlayan kimlik doğrulama çözümlerinden sorumludur. Bir kişinin gerçekte kim olduğunu iddia edebileceği bir sistem oluşturmak için, bu devlet kurumunun hizmetlerini, yani Vatandaş Kartı (CC) veya Dijital Cep Anahtarı (CMD) aracılığıyla kullanmaktadır. Bu proje ile, AMA'nın ve ePortugal veya Portal da Empresa gibi diğer devlet kurumlarının uygulama programlama arayüzünün (API) kullanımı yoluyla yeni müşterilerin kayıt sürecini otomatikleştiren bir web platformu geliştirme amaçlanmaktadır. Bu API'ler, tüketicilerin Vergi Kimlik

Numarası (TIN) veya cep telefonu numarası gibi hassas verileri girerek veya CC'lerini okuyarak, özel işlerini veya kişinin kendisini tanımlayan diğer hassas verileri hızlı ve güvenli bir şekilde aramasına olanak tanır. Bu şekilde müşteri, daha sonra kaydedilecek ve müşterinin profilini oluşturmak için kullanılacak olan verilerinin kullanılmasına izin verdiğini bilir. Hesap kaydına ek olarak, bu API'ler, kullanıcının arka ofiste kimlik doğrulaması yapması veya belgelerin içeriğini doğrulaması ve doğrulaması için alternatif bir yol olarak da kullanılabilir (Pires, 2022).

Rusli ve arkadaşlarının çalışmasında, Endonezya Kimlik kartının hesap sahtekarlığını önlemek için OCR ve NLP teknikleri kullanılarak kimlik kartı doğrulaması yapılmıştır. Bilgi Teknolojilerinin gelişmesi, bilgi alışverişi araçlarını giderek daha fazla değiştirmektedir. Ve bu durum basılı belgelerin sayısallaştırılması ihtiyacına yol açar. Günümüzdeki çağda, sık sık meydana gelen birçok dolandırıcılık vardır. Hesap dolandırıcılığını önlemek için NLP ve OCR kullanılarak kimlik kartı çıkarma doğrulaması yapılmaktadır. Optik Karakter Tanıma (OCR), görüntüden metin oluşturmak için kullanılır. Endonezya kimlik kartı veya Kartu Tanda Penduduk (KTP) da OCR ile ayıklanır ve metne dönüştürülür. Bu, servis operatörünün veri girişi yapmasını kolaylaştırmak için kullanılmaktadır. Doğruluğu artırmak için NLP (Doğal Dil İşleme) yöntemi kullanarak metin düzeltilmesi yapılmıştır. Deneyde 50 tane Endonezya kimlik kartı görüntüsü ile test yapılmıştır. 0.78 F skor elde edilmiş ve Kimlik kartı başına ayıklamak için gerekli süre 4510 milisaniyedir. (Rusli ve ark., 2020).

Yenen ve Şahin'in çalışmalarında, Covid19 'un getirdiği sosyal uzaklaşma zamanlarında Telekom endüstrisi için bir zorunluluk hale gelen uzaktan müşteri kazanımı ve dijital işe alım süreçlerinin işleyişinde OCR (Optik Karakter Tanıma) ile kimlik taraması, derin öğrenme yapay zekâ algoritmaları ile doğrulama, kapsamlı yazılım projelerinde uzaktan satıcıların ve diğer paydaşların yönetimi anlatılmaktadır. Pandemi sürecinde klasik satış kanalları beklenen performansla çalışmadığı için dijital kanallar daha da önem kazanmıştır. Bu kapsamda Telekom sektöründe gerçekleştirilen fiziki evrak kullanılmadan gerçekleştirilen dijital satış projesi ele alınmaktadır (Yenen ve Şahin, 2021).

Aslandođdu ve arkadaşlarının çalışmasında, makine öğrenmesi modelinden faydalanılarak doğruluk tahminleme ve doğruluk analizi yapan yüz tanıma sistemi uygulaması geliştirmesi yapılmıştır. Yüz tanıma teknolojisi, bir insanın yüzünü kullanarak kimliğini tanımlamanın veya doğrulamanın bir yolu olarak kullanılan bir biyometrik doğrulama yöntemidir. Çalışmada birçok Yapay Zekâ tekniklerinden

yararlanılmıştır. Veri kümesi eğitilmiş, test edilmiş ve bu süreç en iyi başarı sağlanıncaya kadar sürdürülmüştür. Bu çalışma herhangi bir yazılım geliştirmesi yapmadan kişinin kendi derin öğrenme modellerini geliştirmesine yarayan Öğretilmiş Makine teknolojisini kullanarak, Tensorflow Javascript teknolojileri ile bulut tabanlı bir yüz tanıma sistemi geliştirilmiştir. Çalışma kapsamında uygulanan algoritmaların başarı oranı %99,8 olarak görülmüştür. Bu makalenin yapay zekâ modellerinin kullanıcılar tarafından geliştirilmesi açısından yapay zekâ bilgi havuzuna sağlayacağı katkı incelenmiştir (Aslandođdu ve ark., 2022).

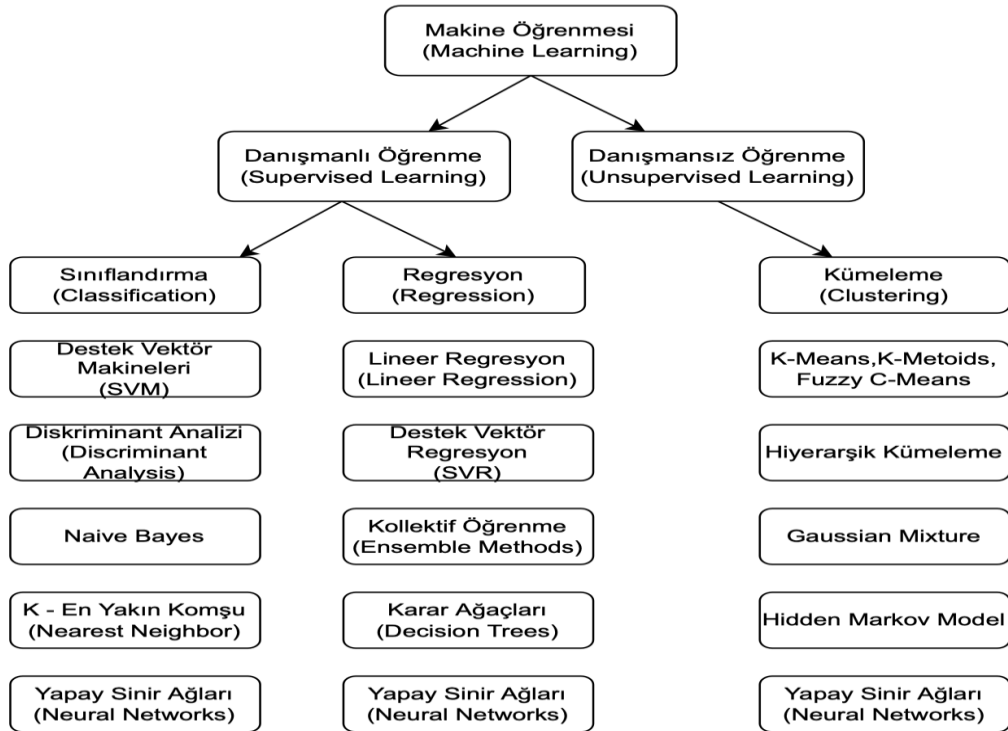


3. MATERYAL VE YÖNTEM

3.1. Makine Öğrenmesi Yöntemleri

Makine öğrenmesi kavramı, Arthur Samuel tarafından 1959'da "bilgisayarın özel olarak programlanmadan ayarlamalar yapma becerisi" olarak ifade edilmiştir (Samuel, 1959). Makine Öğrenmesi, yapay zekanın bir alt kategorisidir ve bilgisayarların örüntü tanıma geliştirme sürecini veya sürekli olarak verilerden öğrenmeye, öğrendiği bu verilere dayalı tahminler yapma ve bu kazandığı becerisini diğer alanlarda kullanmayı ifade eder.

Makine öğrenmesi, bilgisayar programlarının veri setlerinden öğrendiği bilgileri kullanarak, veri setlerinin özelliklerine göre tahmin yapmayı hedefleyen bir yöntemdir. Bu yöntemde, veri setleri etiketlenmiş olarak kullanılır ve veri setlerinden öğrendiği bilgileri kullanarak, veri setlerinin özelliklerine göre tahmin yapılır. Makine öğrenmesi, öğrendiği verilerle bir sorunu çözmeyi veya bir görevi yerine getirmeyi hedefler. Bu tahminler, genellikle veri setinin özelliklerine göre bir modele uygun hale getirilir ve bu modele göre, veri seti üzerinde bir tahmin yapılır. Şekil 3.1, makine öğrenimi yöntemlerinin genel bir özetini vermektedir.



Şekil 3.1. Makine Öğrenmesi Yöntemleri (MathWorks, 2022)

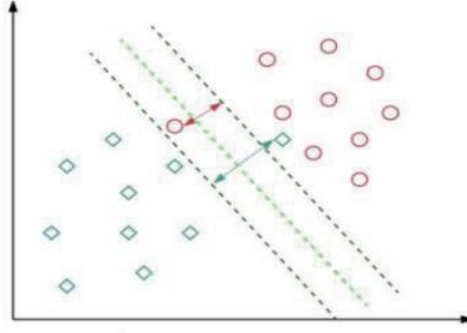
Şekil 3.1’de gösterilen yöntemlerden en yaygın kullanılanları destek vektör makinesi yöntemi, naive bayes, lineer ve lojistik regresyon, yapay sinir ağı yöntemi, karar ağaçları ve rastgele orman algoritmalarıdır. Bu yöntemler, veri setlerini incelerken, veri noktaları arasındaki ilişkiyi modellemeyi, veri noktalarını sınıflara ayırmayı ve veri noktalarının özelliklerine göre tahmin yapmayı hedefler. Veri setinin özelliklerine göre makine öğrenmesi algoritmaları denetimli öğrenme ve denetimsiz öğrenme olmak üzere iki farklı kategoriye ayrılmaktadır (VanderPlas, 2017).

Denetimli Öğrenme: Denetimli öğrenme yöntemleri, veri setlerinin büyüklüğü ve özelliklerine göre seçilerek kullanılır. Denetimli öğrenmede, sınıflandırma yapılacak özneliği bilinen veri seti özelinde, belirlenen algoritma eğitilir. Sonrasında eğitimi tamamlanmış olan algoritma ile tahmini yapılmaya çalışılan veri setinin sonuçlarını belirlemeye çalışır. Veri setlerinin özellikleri arasında anlamlı bir ilişki olduğu durumlarda denetimli öğrenme yöntemleri tercih edilebilir. Denetimli öğrenmede sınıflandırma sayısı ve veri setindeki özneliklerin hangi sınıfa ait olduğu bilgisinin, öncesinde tahmin ediliyor olması gerekmektedir.

Denetimsiz Öğrenme: Denetimsiz öğrenme yöntemlerinde, veri setindeki hangi özneliğin hangi sınıflandırmaya dahil olacağı bilgisi verilmemekte, verinin kendi başına öğrenmesi ve kendisine uygun, en makul sınıflandırmaya yerleşmesi beklenmektedir. Denetimsiz öğrenme, herhangi bir sınıflandırma yapılmadan veri setinin tipine göre modelleme yapmaktadır. Veri setlerinin özellikleri arasında anlamlı bir ilişki olmadığı durumlarda denetimsiz öğrenme yöntemleri tercih edilebilir.

3.1.1. Destek vektör makinesi

Destek Vektör Makinesi (DVM), denetimli makine öğrenmesi algoritmalarındandır. Düzlem üzerine konumlandırılmış noktaları ayırt etmek için düz bir çizgi çizer ve çizginin, sınıflandırılmış iki verinin noktaları için de maksimum mesafede olmasını hedefler. Eğer tahmin edilmesi gereken veri, bu düzlem üzerindeki ikili sınıflandırmadan biri ise bu yöntem kullanılabilir. Çözümlemesi karışık, aynı zamanda küçük ve orta ölçüdeki veri kümeleri için uygun bir makine öğrenmesi yöntemidir. Destek Vektör Makinesi yönteminin kullanım açısından yapısı itibariyle basit olması ve kullanıldığı uygulamalarda yüksek performans elde etmesi sebebiyle kullanımı oldukça yaygındır. Örnek bir DVM Şekil 3.2’de gösterilmiştir.

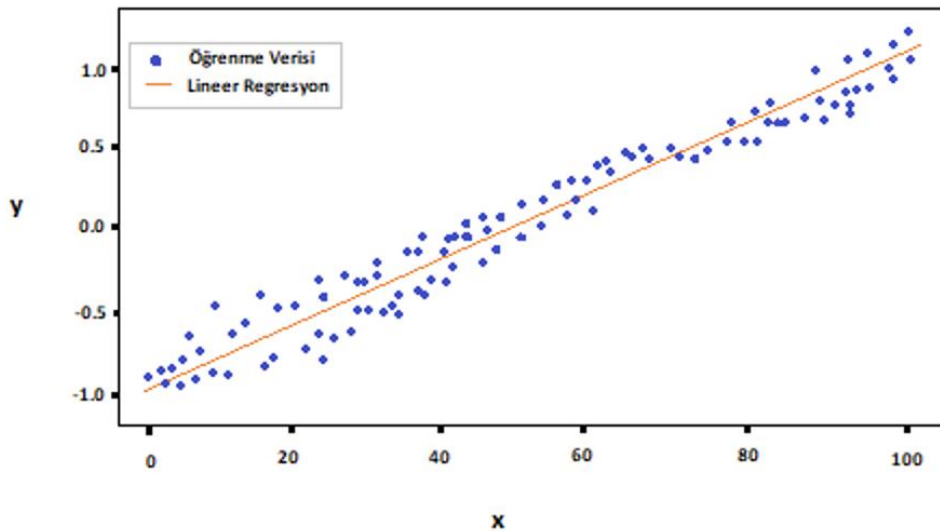


Şekil 3.2. Örnek DVM modeli

3.1.2. Lineer regresyon

Lineer regresyon, bir veri setinde bulunan iki değişken arasındaki ilişkiyi modelleyen bir makine öğrenimi yöntemidir. Bu yöntem, veri setlerinde bulunan bir bağımlı değişken (y değişkeni) ve bir bağımsız değişken (x değişkeni) arasındaki ilişkiyi modelleyerek, y değişkeninin değerini x değişkene göre tahmin etmeyi hedefler.

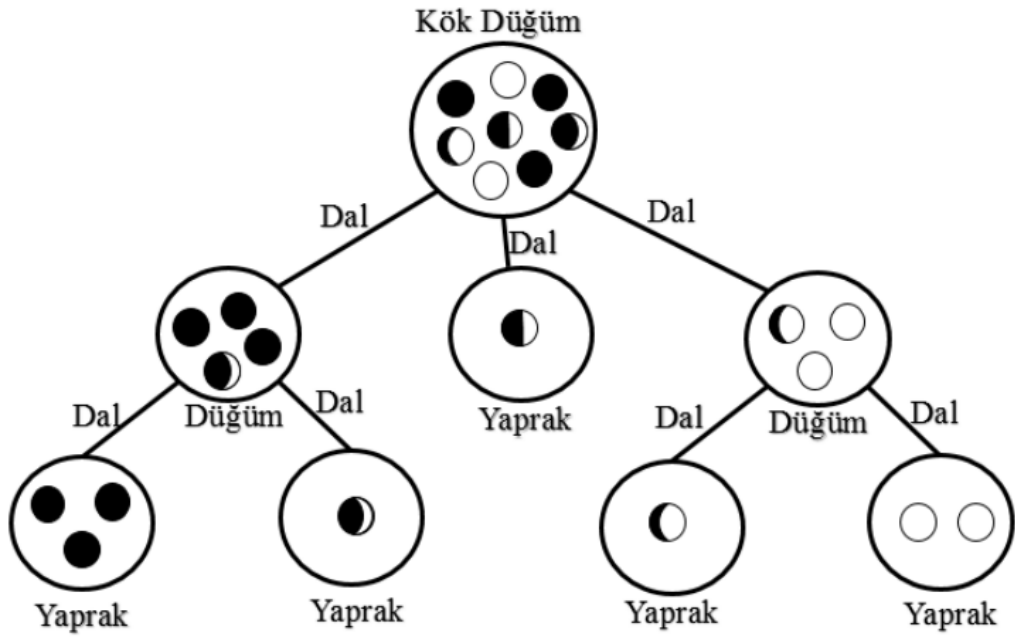
Lineer regresyon yöntemi, veri setlerinde bulunan x ve y değişkenleri arasındaki ilişkiyi en iyi şekilde modelleyen bir doğru çizerek, veri setlerinin özellikleri arasındaki ilişkiyi anlatır. Veri kümelerinin en iyi ihtimale sahip olan kısmını kapsayacak şekilde bir doğru denklemi oluşturulur ve oluşturulan bu denklem ile gelecekteki veriler tahmin edilir. Öznitelik sayısı birden fazla olursa Multilineer Regresyon şeklinde adlandırılır (Kardaş ve Güveni, 2019). Bağımlı değişkenin kategori edilmiş olduğu senaryolar için oldukça kullanışlı bir yöntemdir. Örnek bir regresyon modeli Şekil 3.3'te gösterilmiştir.



Şekil 3.3. Örnek lineer regresyon modeli

3.1.3.Karar ağaçları

Karar ağaçları, denetimli makine öğrenimi yöntemlerinden biridir. Gruplama işleminin, ağaç modeli üzerinde bir sistemde sınıflandırarak, belirlenen öznelik üstünde bazı karar verme aşamaları kullanılarak elde edildiği bir sınıflandırma algoritmasıdır. Bu yüzden karmaşık veri modellerinde kullanımı tavsiye edilmektedir. Karar ağaçlarının en baştaki düğümlerine root adı verilir. Her bir inceleme kökte yer alan koşula göre “Evet” veya “Hayır” olarak bölümlendirilir. Yapraklar ise sonucu vermektedir. Şekil 3.4’te örnek karar ağacı modeli gösterilmektedir.



Şekil 3.4. Örnek karar ağacı yapısı

Karar ağaçları yöntemi, düğüm, dal ve yapraklardan meydana gelmektedir. Ağaçtaki eğitim için kullanılacak olan ve verileri içeren en tepedeki düğüm, kök olarak adlandırılmaktadır (Krzysztof Cios ve ark., 2007). Bu yüzden bir karar ağacı kökü yukarıda olacak şekilde ters olarak çizilir. Şekil 3.4’deki resimde, ağacın dallara kenarlara bölünmesine bağlı olarak bir koşulu temsil eder. Artık ayrılmayan dalın sonu yaprak olarak ifade edilir. Başka bir deyişle kök ile diğer düğüm ortasında bulunan yapıya dal, en son bölümde meydana gelen yapı da yaprak olarak isimlendirilmiştir. Karar ağaçlarının oluşumunda dallanmanın belirlenmesi, hangi koşullar baz alınarak

oluşturulacağı ya da hangi yöntemlere göre ağaç yapısının belirleneceği konusunda farklı öneriler vardır. Bunlardan bazıları Bilgi Kazancı, Ki-Kare olasılık tablo istatistiği, Twoing kuralı, Gini İndeksi, gibi yaklaşımlardır (Kavzoğlu ve Çölkesen, 2010).

Karar ağaçları yönteminde oluşturulan dalların amacı eğitimde kullanılacak olan veri setini en uygun şekilde benzer özellikleri baz alarak gösteren alt veri setlerine bölmektir (Aggarwal, 2015). Bu bölme işlemi kesme kuralı koşulu elde edilene kadar sürdürülmektedir. Fakat fazla öznitelik ortaya çıktığında, çok fazla bölünmeye sebep olunabilir ve oldukça büyük bir ağaç modelinin oluşumuna sebep olur. Böyle ağaçlar karmaşaya ve hatalı sonuçlar elde edilmesine yol açabilir. Ne zaman bölünmenin durdurulması gerektiğine karar verilmesinin yollarından biri, tüm yapraklarda kullanılacak minimum düzeyde eğitim girdisini ayarlamaya çalışmaktır. Diğer bir yöntem ise modelin “maksimum derinliğini” hesaplamaktır. Maksimum derinlik kavramı, bir kökten bir yaprağa giden en uzun mesafenin uzunluğunu ifade eder.

Karar ağaçları yönteminde, modelin başarı oranı budama işlemi ile artırılabilir. Budama, modele katkısının hiç olmadığı ya da az olduğunu düşündüğümüz dalların kaldırılmasına fayda sağlar. Bu şekilde, modelin karmaşıklığı azaltılmış ve böylece tahminlemesi artırılmış olur (Kavzoğlu ve Çölkesen, 2010). Budamaya, kök ya da yapraklardan olacak şekilde istenilen düğümden başlanabilir. En kolay budama yöntemlerinden biri; yapraklardan başlayarak ve o yapraktaki en bilindik sınıfa ait her düğümü çıkararak döngüye girmesidir. Alt modelin büyüklüğüne göre düğümlerin kaldırılıp kaldırılamayacağını anlamak adına bir öğrenme parametresinin (alfa) kullanıldığı maliyet karmaşıklığı budaması gibi daha karmaşık budama yöntemleri kullanılabilir. Bu durum, zayıf noktanın model üzerinden çıkarılması olarak da bilinir.

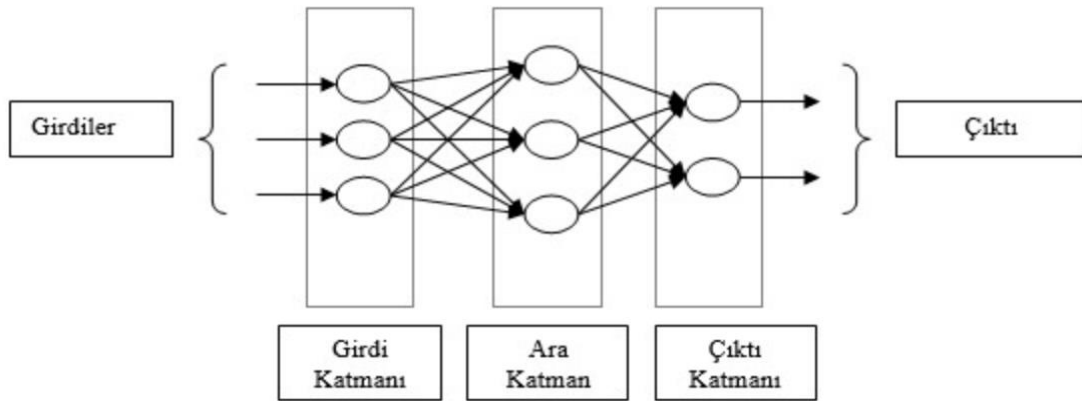
3.1.4. Yapay sinir ağları

Yapay sinir ağları yöntemi, insanın beyninin işleviyle alakalı prensiplerinin matematiksel hesaplar yapan cihazlar üzerinde benzer şekilde kodlama düşüncesi ile meydana gelmiş ve beynin nöronlarının sayısal ifadelerle modellenmesi ile oluşturulmuş bir yöntemdir (Efe ve Kaynak, 2000). Yapay Sinir Ağı algoritmaları insan beyninden esinlenmiştir. Yapay nöronlar birbirine bağlıdır ve birbirleriyle iletişim kurar. Her bağlantı, önceki öğrenme olayları tarafından ağırlıklandırılır ve her yeni veri girişi ile daha fazla öğrenme gerçekleşir. Yapay sinir ağları yöntemi ile insan beyninin idrak edebildiği birçok fonksiyon ve insanların duygusal yönden açık (acıkmak, yorulmak,

mutsuz olmak, ağlamak vb.) sayılan özellikleri yok sayılarak kolay bir biçimde ortaya koyulmaktadır (Aytekin, 2017).

Yapay sinir ağı modeli, bu modeli meydana getiren yapay sinir hücrelerinin birbirleriyle birçok layer içerisinde paralel iletişimler kurarak bir araya gelmeleriyle oluşturulmaktadır (Aydın, 2005). Yapay sinir ağı algoritmalarında, süreç insan beyninin işleyişindeki gibi belirli bir süre bilgiyi öğrenmek ve elde etmek ile sonrasında bu öğrendiği bilgiler için sinaptik ağırlık şeklinde ifade edilen ilişkiler kullanılmaktadır (Haykin, 1994; Tiryaki ve ark., 2015).

Verileri analiz etmek ve öğrenmek için farklı katmanlardan oluşurlar. Her gizli katman, resimdeki desenleri algılamaya çalışır. Bir desen algılandığında, bir sonraki gizli katman etkinleştirilir ve bu böyle devam eder. Sinir Ağları, ağın verileri her işlediğinde farklı nöronlar arasındaki bağlantıları öğrenir ve ağırlıkları ilişkilendirir. Şekil 3.5'teki görselde yapay sinir ağı örneği gösterilmiştir.



Şekil 3.5. Yapay sinir ağı örneği

Yapay sinir ağı yöntemi uygulamalarından çoğunlukla sınıflandırma, veri ilişkilendirme, tahmin, sınıflandırma, veri filtreleme ve veri yorumlama gibi işlemlerde yararlanılmaktadır. Bu yöntemde Şekil 3.5'teki örnek modelde de görüldüğü üzere giriş değerleri ve bu giriş değerlerine göre tahmin edilerek oluşturulan sonuç yani bir çıkış değeri oluşmaktadır. Veri modelini sınıflandırma yapmak amacıyla kullanılan yapay sinir ağları, giriş değerlerini bölümlene yapmak için kullanılır. Mesela bir makine üzerinde fark edilen yanlış işleyişin sınıflandırılması örnek olarak gösterilebilir. Yapay sinir ağları veri ilişkilendirme için modellendiğinde, eğitimle öğrenmiş olduğu bilgiler ışığında hatalı ve eksik kısımları düzeltmektedir. Örneğin eksik bir görüntünün bütün haline getirilmesi veri ilişkilendirilmesi ile sağlanır. Yapay sinir ağları, veri yorumlamak için

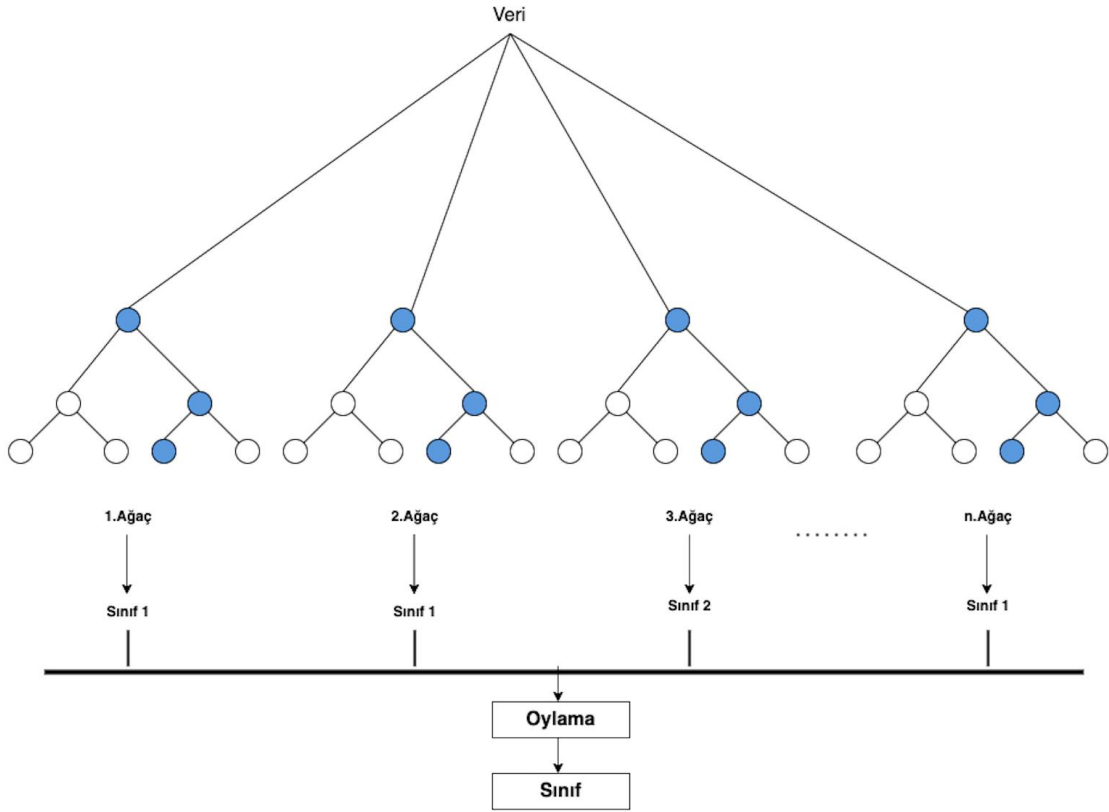
modellendiğinde ise giriş verilerini analiz etmektedir. Yeni olayların yorumlanması, bir durum karşısında elde edilen tüm örnek veriler ile verinin eğitilmesi ile ortaya çıkan bilgiler kullanılarak sağlanır. Birçok veri kümesi içerisinde en elverişli verileri belirli duruma getirme işlevini veri filtrelemede eğitilen ağlar sağlamaktadır.

3.1.5. Rastgele orman

Rastgele Orman Algoritması, son derece popüler olan ve makine öğrenimindeki sınıflandırma ve regresyon problemlerinde kullanılan denetimli bir makine öğrenmesi algoritmasıdır. Temeli bir ormanın çok sayıda ağaçtan oluştuğunu ve ne kadar çok ağaç olursa ormanında o kadar sağlam olacağı teorisine dayanmaktadır. Bu teoriye göre rastgele orman algoritmasındaki ağaç sayısı arttıkça, doğruluğu ve problem çözme yeteneği de artmaktadır. Rastgele orman yöntemi, istenilen veri setinin farklı birçok alt segmentinde karar ağacı algoritması içeren ve bu veri setinin yaklaşık olarak doğruluk analizini geliştirmek için ortalaması alınan bir sınıflandırıcı görevini üstlenir. Karmaşık bir sorunu çözmek ve modelin performansını iyileştirmek için birden çok sınıflandırıcıyı birleştirme süreci olan topluluk öğrenme kavramına dayanır. Aynı zamanda bu algoritma Bagging yönteminin aşamalarına dayanmaktadır. Breiman aracılığıyla 1996'da geliştirilmiş olan Bagging yöntemi veri eğitim kümesi kullanarak farklı alt veri setleri oluşturur ve bu alt veri setlerinde elde edilen ağaçlar ile ayırma işlemi yapılmaktadır olarak açıklanır (Breiman, 1996).

Rastgele orman yöntemi, veri setlerinin özelliklerini kullanarak birçok farklı karar ağacı oluşturur. Oluşturulacak karar ağacı sayısı kullanıcıdan alınır. Ağaç sayısının fazla olması sınıflandırma işleminin başarısının da artmasını sağlar. Makine öğrenmesinde sınıflandırma işlemi, öğeleri önceden kategorize edilmiş bir eğitim veri kümesine göre kategorilere ayırma sürecidir. Bu yüzden rastgele orman yönteminde ağaç sayısı arttıkça algoritmanın çalışma süresi de aynı oranda artış göstermektedir. Bunun yanında rastgele orman yönteminde, ağaç sayısındaki artış ne kadar olursa olsun ezberleme olarak adlandırılan ve literatürdeki karşılığı overfitting olan koşul ile karşılaşılacaktır. Overfitting (ezberleme), veri kümesinin aşırı oranda uyum göstermesi ve dolayısıyla bu veri kümesinde yer almayan yeni verilere uyum sağlayamaması problemidir. Rastgele orman algoritması, kategorik değerler ile çalışmaktadır. Kategorik değer, veri tipleri sayısal ifadelerle sahip olmayan verilerdir. Ancak makine öğrenmesinin temeli matematiksel ifadelerle dayandığı için kategorik verilerin sayısal olarak ifade edilmesi

gerekmektedir. İki tür kategorik değer bulunmaktadır. Bunlar ordinal ve nominal verilerdir. Nominal veri tipi için Evli/Bekar, Kadın/Erkek örnekleri verilebilir. Nominal değişken içinde yer alan değerlerin birbirinden üstün olma gibi bir durumu söz konusu olmadığı için aralarında herhangi bir sıralama olamaz. Ordinal veri tipinde ise bir üstünlükten bahsedilebilir. Ordinal veri tipine ‘az gelişmiş’, ‘gelişmekte olan’, ‘gelişmiş’ gibi ülke gelişmişlik düzeyini belirten nitelikler örnek gösterilebilir. Rastgele orman algoritmasında kullanıcıdan alınan karar ağaçları, veri setlerinin özelliklerine göre eğitilir ve bu eğitim sırasında, her bir karar ağacının veri setlerinin özelliklerini en iyi şekilde modelleyen bir çıktı üretmesi hedeflenir. Daha sonra, bu karar ağaçlarının çıktıları bir araya getirilerek, veri setlerinin özelliklerini en iyi şekilde modelleyen bir tahmin üretilir. Şekil 3.6’da örnek bir rastgele orman yöntemi verilmiştir.



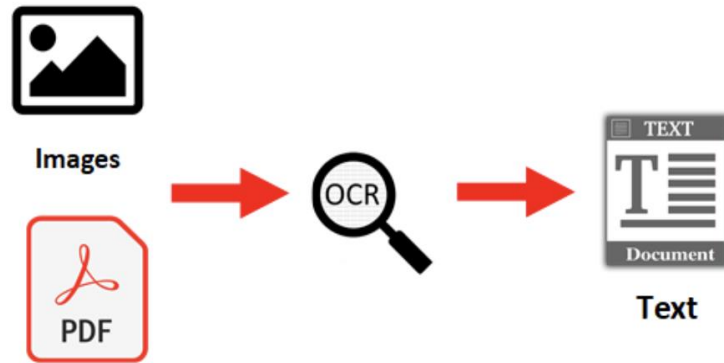
Şekil 3.6. Rastgele orman yöntemi (Soylu, 2018)

3.2. Kimlik Tespit Yöntemleri

3.2.1. OCR (Optik karakter tanıma)

OCR teknolojisi, basılı veya yazılı belgelerdeki karakterlerin tanınması olarak ifade etmektedir. Örneğin; dergi, gazete ya da el yazısı içeren belgelerden metinleri ayırt edip, işleyerek istenilen ortamda dosya formatında aktarılan, başka bir söylemle dijital forma dönüştüren bir yazılım teknolojisidir (Eser, 2021). Başka bir tanımla, makine baskısı veya el yazısı metin görüntüsünü düzenlenebilir bir dijital bilgisayar formatına dönüştürme işlemidir. OCR teknolojisi, belge taramaları ve fotoğraflar gibi görsel verileri işleyerek, bu verilerdeki metinleri tanımlayabilmektedir. OCR teknolojisiyle işlenen verinin doğruluğunu artırmak için, önceden belirlenmiş doğrulama kuralları kullanılabilir. Bu sayede, OCR teknolojisinin doğruluğu ve verimliliği artırılabilir. OCR teknolojisi, örüntü tanıma ve yapay zekâ alanında zorlu bir araştırma alanı olarak kabul edilir (Nanonets, 2019).

OCR teknolojisi, günümüzde birçok alanda kullanılmaktadır. Bu alanlardan biri de kimlik tespit yöntemleridir. Örneğin, müşterinin kimliğini doğrulamada kullanılmaktadır. Müşterinin bir pasaport veya kimlik belgesi göndermesi durumunda, bu belgenin taranmış bir kopyası OCR teknolojisiyle işlenerek, belgedeki bilgilerin doğruluğu kontrol edilebilir. OCR teknolojisinin kullanımının birçok avantajı vardır. OCR teknolojisi hızlı ve etkili bir şekilde veri işler ve bu sayede, uzaktan müşteri edinim süreci hızlandırılır (Ormanlı, 2020). Ayrıca, OCR teknolojisi, müşterinin fiziksel olarak bir belge göndermesine gerek kalmadan, belge taramasının yapılmasını mümkün kılar. Bu sayede, müşteri edinim süreci için zaman ve para tasarrufu sağlanmış olur. Şekil 3.7'de OCR teknolojisi modellenmiştir.



Şekil 3.7. OCR modellemesi

3.2.2. NLP (Doğal dil işleme)

NLP teknolojisi, sözel verileri işleyen ve bu verileri anlamlı bir ifadeye dönüştüren yöntemdir. Bu teknoloji, dillerin kurallı yapısını küçük parçalara ayırıp, işleyerek, anlaşılmasını veya tekrardan türetilmesi amacını gütmektedir. Bu çözümlenmeye örnek olarak konuşma sentezi, bilgi sağlama, yazılı belgelerin otomatik çevrilmesi, komut anlama, otomatik konuşma ve konuşma üretme, otomatik metin özetleme, doğruluk analizi gibi farklı kullanım alanlarında birçok örnek verilebilir.

NLP teknolojisi girdi ve çıktı verilerine göre iki kategoride işlenmektedir. Bu kategoriler yazılı metinlerin ve ses kayıtlarının üzerinde yapılan işlemlerdir. Aynı zamanda doğal dil işlemenin çoğunlukla iki kaynağı bulunmaktadır. Bu iki kaynak, sesli konuşma çerçevesinde ona yine sesli cevap vermek veya yazılı bir ifadenin analiz edilmesi olarak ilişkilendirilebilir. Ses özelinde NLP çalışmaları; ses üzerinden daha çok sesli ifadelerin yazılı belgeler haline dönüştürülmesi ve daha sonra yazılı hale dönüştürülmüş bu metinlerin işlenmeleriyle oluşmaktadır (Seker, 2015).

3.2.3. Biyometrik doğrulama

Yüz tanıma teknolojileri, bir insanın yüzünü kullanarak kimliğini ayırt etmenin yollarından biridir. Yüz tanıma teknolojileri kişileri, resimlerden, video kayıtlarından veya canlı olarak tanımayı sağlamaktadır. Yüz tanıma, aynı zamanda biyometrik doğrulamanın da yöntemlerinden biridir. Diğer biyometrik doğrulamalara; sesli tanıma, parmak izi tanıma sistemi ve retina tanıma ya da iris tanıma örnek verilebilir. Bu teknoloji, çoğunlukla güvenlik, emniyet ve bankacılık uygulamalarında kullanılsa da teknolojiyle ilgili diğer alanlarda da kullanımı her geçen gün artmaktadır.

Yüz tanıma yazılımı, yüz verileri için şifreli bir dijital model oluşturmak üzere gözler arasındaki mesafe, çene ve burun arasındaki mesafe vb. dahil olmak üzere yüzün geometrisini analiz eder. Kimlik doğrulaması yapılırken, yüz tanıma aracı yüzünüzü gerçek zamanlı olarak tarar ve modeli sistemde depolanan modelle karşılaştırır.

Sunulan biyometrik özelliğin gerçek bir insandan mı yoksa dijital veya üretilmiş bir gösterim mi olduğunu belirlemek için kullanılan iki tür canlılık tespiti vardır. Bunlar; aktif canlılık tespiti ve pasif canlılık tespitidir. Aktif canlılık tespiti, kullanıcının gözlerini kırpmasını veya başını çevirmesini gerekli kılar. Pasif canlılık tespitinde ise kullanıcıya ait yakalanan bir biyometrik numuneyi analiz etmek için algoritmalar kullanır. Yüz tanıma, kişinin benzersiz yüz verisi ile yakalanan görüntüsünün algoritmik hesaplamalar

sonucunda karşılaştırılarak benzerlik oranı elde edilmesi ile gerçekleşir. Şekil 3.8’de yüz tanıma ait görsel bulunmaktadır.



Şekil 3.8. Yüz tanıma modeli

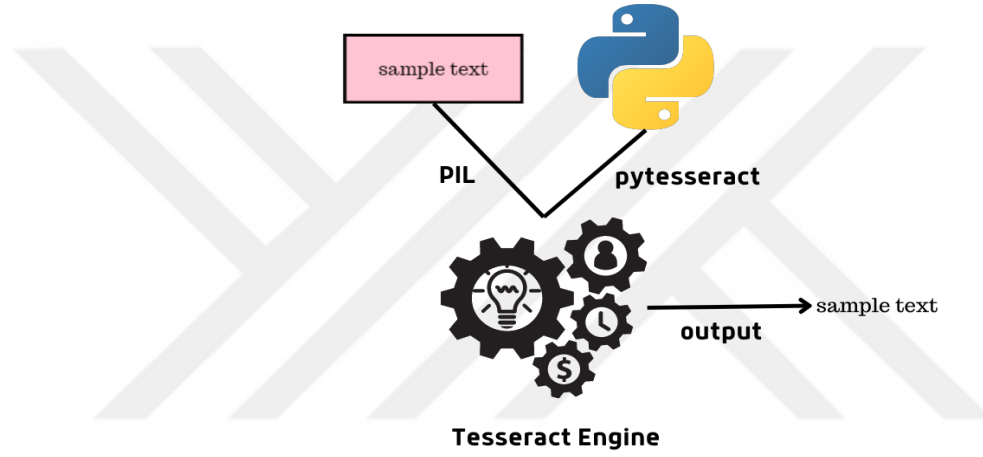
3.2.4. Güvenlik sorusu ve sesli yanıt

Güvenlik sorusu ve sesli yanıt sistemi, kişinin gerçekliğinden emin olmak için kullanılan bir yöntemdir. Güvenlik sorusu olarak, müşteri kimliğini ayırt etmeyi sağlayan ve resmi kimlik belgesi yerine geçen belgeler üzerinde yer alan kişiye ait bilgiler, dijital hizmetlerin kullanımı boyunca hiçbir adımda kimlik doğrulama gayesiyle kullanılmamalıdır. Kurum ya da kuruluşların kimlik doğrulama yaparken müşterinin bildiği kendisine ait bilgileri güvenlik sorusu olarak kullanmak istemesinde, bu güvenlik sorusunun resmi kimlik kartı ya da onun yerine geçen dokümanlar üzerinde bulunan bilgilerden biriyle ilişkili olmaması ve cevabın sadece müşterinin kendisi tarafından biliniyor olması gerekmektedir. Sesli yanıt sistemi ile de gerekli belgeleri onaylama, kişinin kendi sesiyle onayının alınmasını sağlamış olur.

3.3. Kütüphaneler ve Algoritmalar

3.3.1. TesseractOCR

TesseractOCR, açık kaynak kodlu optik karakter tanıma(OCR) platformudur. OCR, metin katmanı olmadan görüntülerden ve belgelerden metin çıkarır ve belgeyi yeni aranabilir bir metin dosyasına, PDF'e veya diğer birçok formata dönüştürebilir (Smith, 2007). Görüntülerden basılı metni çıkarmak için bir API kullanır ve çok çeşitli dilleri desteklemektedir. Şekil 3.9'da kütüphanenin işleyişine ilişkin görsel bulunmaktadır.



Şekil 3.9. TesseractOCR kütüphanesi işleyişi

3.3.2. FuzzyWuzzy

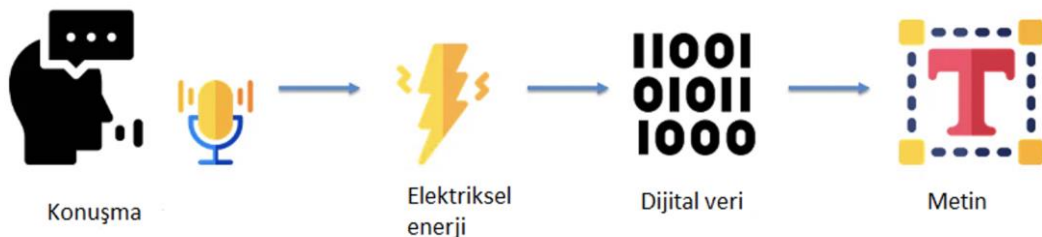
String ifadelerin birbirleriyle olan benzerliklerini anlamak için farklı çözüm yöntemleri mevcuttur. Bu benzerlikleri yakalamak ve arama motorlarında alakalı sonuçlar döndürmek için basit bir yapıdan daha çok yapay zekâ tabanlı bir yapıya ihtiyaç duyulmaktadır. FuzzyWuzzy, Python dilinde bu benzerlikleri yakalamak için kullanılan kütüphanedir. Inputlar arasındaki farkları hesaplamak için Levenshtein Distance algoritması kullanılır. Levenshtein Distance algoritması, özellikle arama motorlarında (search engine) kullanılmaya uygun bir model sağlamaktadır. Kullanıcıların tarayıcı üzerinden arama yaptığı sözcükleri tam olarak ifade edemedikleri, kelimeye benzer ya da kelimeye ait sadece birkaç harfi hatırladıkları koşullarda, öneri olarak sunulan sözcüklerin sunulmasını sağlayan bir metin işleme algoritmasıdır. Levenshtein Distance algoritması, girilen bir string değer başka bir string değere dönüşmesi için

belirli adımlar uygular ve matris hesaplamalar gerçekleştirir (Senyurt, 2012). Amacı, iki stringi eşleştirmektir. Sonuç olarak tek bir matematiksel ifade dönmektedir ve iki kelimedenden birinin diğerine dönüştürülebilmesi için gerekli olan işlem sayısını ya da maliyetini vermektedir. Bu değer küçük olması, istenilen bir durumdur. Bir sözcüğün, bir öneri veri seti içindekiler ile karşılaştırılması sonucu ortaya çıkan sayısal ifadelerden en küçüğü veya küçükleri, amaca ulaşılması ve doğru tavsiyelerde bulunulması bakımından oldukça önem arz etmektedir. Vladimir Levenshtein tarafından geliştirilen bu algoritma, aşağıda gösterilmekte olan özyinelemeli (recursive) bir eşitliği kullanmaktadır. Eşleştirme yapılması için FuzzyWuzzy kütüphanesinde yer alan Levenshtein Distance algoritmasının gerçekleştirdiği kontroller şunlardır;

- **Insert:** Başlangıç string'ine harf eklemesi yapar.
- **Substitution:** Kelimeler eşleştirilirken yanlış yazılan harfin bitiş string'indeki harf ile değiştirilmesini gerçekleştirir.
- **Delete:** Kelimeler eşleştirilirken başlangıç string'inde fazladan harf var ise harfin silme işlemini gerçekleştirir.

3.3.3. SpeechRecognition

SpeechRecognition, bir Python kütüphanesidir. Dilbilimsel ve akustik modelleme gerçekleştiren algoritmalarla çalışır. Akustik modelleme, konuşmanın daha önemli kısımlarını kelime ve cümle olarak elde etmek için konuşmadaki fononları/fonetiği tanımak için kullanılır. SpeechRecognition kütüphanesi, ses sinyallerini seslere ayırır ve bu sese uyan en olası sözcüğü bulmak için algoritmaları kullanarak sesleri analiz eder. Tüm bu işlemler, doğal dil işleme ve yapay sinir ağları kullanılarak yapılır (Kamath ve ark., 2019). Özetle, ses dosyasından veya mikrofondan alınan ses yazıya dönüştürülür. Sesin metne dönüştürülmesi Şekil 3.10' daki görselde modellenmiştir.



Şekil 3.10. Sesin metne dönüştürülme adımları

Konuşma tanıma, konuşan kişinin ürettiği ses enerjisinin mikrofon yardımıyla elektrik enerjisine dönüştürülmesi ile başlar. Daha sonra bu elektrik enerjisini analogdan dijitale ve son olarak da metne dönüştürür.

3.3.4. OpenCV

OpenCV kütüphanesi, intel firması aracılığıyla 1999 senesinde öne sürülmüş açık kaynak kodlu bir görüntü işleme kütüphanesidir. Birçok görüntü işleme alanında, özellik çıkarma algoritmalarında, nesnelerin tanınmasında, videoların analizlerinde ve makine öğrenmesi gibi alanlarda OpenCV kütüphanesi kullanılmaktadır (Pisarevsky, 2007). OpenCV kütüphanesi içerisinde görüntü işlemeye ve makine öğrenmesine ilişkin 2500'den fazla algoritma ve yöntem yer almaktadır. Bu yöntemler ile yüz tanıma, nesnelere belirleme, nesnelerin sınıflandırılması, plaka tanıma işlemleri, resimlerin karşılaştırılması, OCR ile metin tanımlama gibi işlemleri yapabilme kabiliyeti yüksektir (Pişkin, 2017).

3.3.5. FaceRecognition

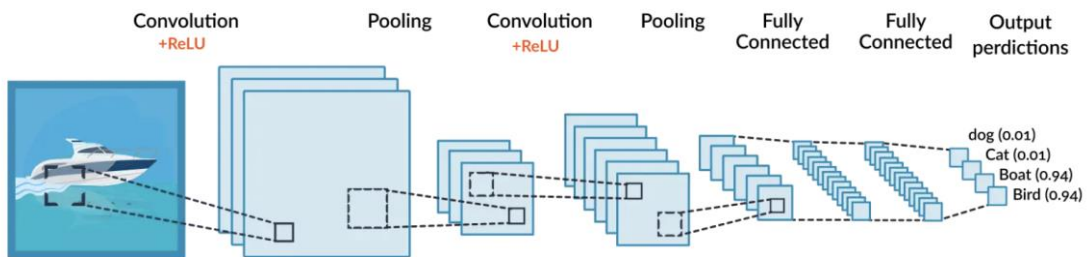
Yüz tanıma teknolojisi, yüzleri fotoğraf ve videolardan tanıma işlemi yöntemidir. Yüz tanıma sistemleri fotoğraflarda, videoda veya gerçek zamanlı olarak kişileri tanımlamak için kullanılır (Çınar, 2019). FaceRecognition kütüphanesi, bir kişinin yüzünü kullanarak kimliğini belirleme veya doğrulamada kullanılan kütüphanelerden biridir. Bu kütüphanede, bulunan yüzlerin her birinin tam koordinatlarını belirlenir ve bu koordinatlardaki resim ayrı bir görüntüye çevrilir. Bu sayede bir fotoğraf üzerindeki kişilere ait yüzler tespit edilebilir ve işlenebilir. Kütüphane, fotoğraflar içerisinde bulunan yüzlerin konumlarını belirler ve belirlenmiş bu yüz fotoğrafları bilgisayarın anlayabilmesi için matrisler halinde ifade edilir. Yüz tanıma algoritmalarında verinin doğrulama oranlarının artırılması parçaya dayalı metodlar ile sağlanmaktadır. Bu algoritmalarda, kişinin yüzü belirli bölgelere bölünerek, bölgesel olarak tanıma sağlanmakta ve yüzün farklı bölümlerinden edinilen neticeler sonucunda bir araya getirilmektedir. Bölgeye yönelik parçalı yapılan tanıma yöntemlerinde, bütün yüzü kullanılarak bir modele dönüştürülmektedir (Kakadiaris ve ark., 2004). Model üzerindeki belirlenmiş bölgelerden o yüze ait kısımlar elde edilmekte ve her parça ayrı bir dalgacık dönüşümüyle analiz yapılarak, sonuçlar ağırlıklı toplamla birleştirilmektedir.

3.3.6. CNN (Convolutional neural network)

Konvolüsyonel Sinir Ağları (CNN) kütüphanesi, çok katmanlı algılayıcıların (Multi Layer Perceptron-MLP) türlerinden biridir. Hücrelerin görme merkezindeki; tüm görüntüyü içine alacak biçimde alt bölgelere bölünmüş, basit hücreler, kenar benzeri özelliklere, karışık hücreler ise daha büyük alıcılar ile tüm görüntüye konsantre olduğu düşünülmektedir (Fukushima, 1980). CNN algoritması, görüntüyü çeşitli katmanlarla işler. Bu katmanlar;

- **Convolutional Layer:** Özellikleri saptamak için kullanılır.
- **Non-Linearity Layer:** Sisteme doğrusal olmayanların (non-linearity) tanıtılması.
- **Pooling (Downsampling) Layer:** Ağırlık sayısını azaltır ve uygunluğu kontrol eder.
- **Flattening Layer:** Klasik Sinir Ağı için verileri hazırlar.
- **Fully-Connected Layer:** Sınıflamada kullanılan Standart Sinir Ağı katmanlarıdır.

CNN algoritması, sınıflandırma sorununun çözümü için temel olarak Sinir Ağı algoritması kullanmaktadır. Fakat bu bilgileri ayarlamak ve bazı özellikleri ayırt etmek için diğer katmanlar kullanılmaktadır (Hubel ve Wiesel, 1968). Şekil 3.11’de CNN ağına ait mimari bulunmaktadır.

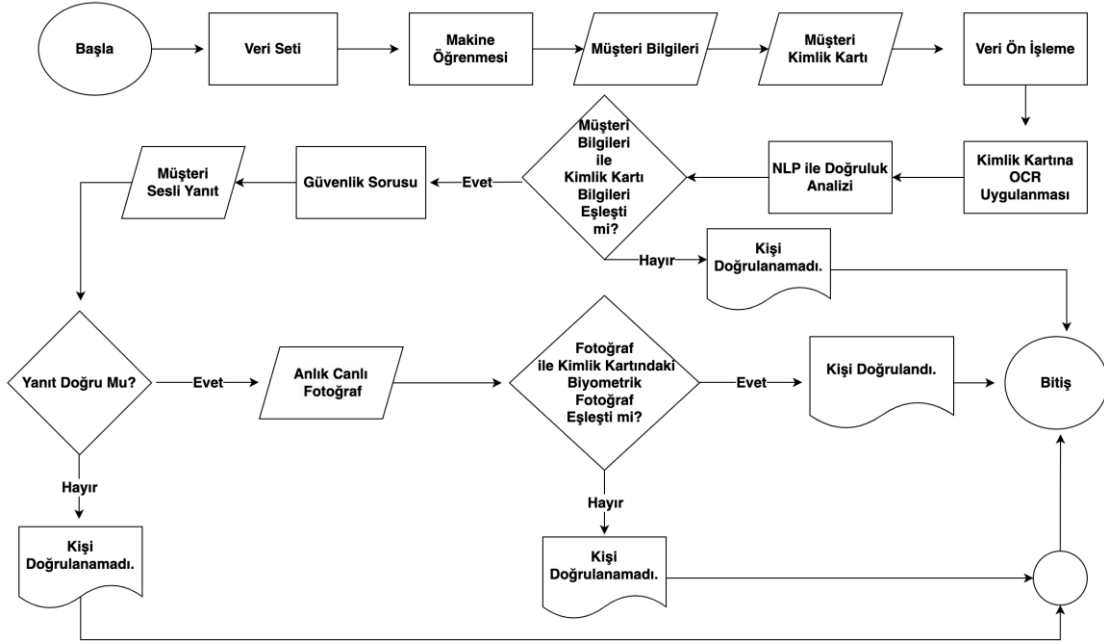


Şekil 3.11. CNN ağı mimarisi

CNN çoğunlukla görüntü işleme projelerinde kullanılır ve giriş değeri olarak görüntüleri alan bir makine öğrenme algoritmasıdır (Kılınç ve Başepmez, 2018). Farklı tekniklerle görüntülerdeki nitelikleri yakalayan ve sınıflandıran CNN kütüphanesi, farklı katmanlardan oluşmaktadır. Convolutional Layer, Pooling ve Fully Connected katmanlarından geçen görüntü, çeşitli işlemlere maruz bırakılarak derin öğrenme modeline girecek şekle getirilmektedir.

4. UYGULAMA

Çalışma kapsamında, müşteri canlılık tespiti için PyQt eklentisi kullanılarak Python dilinde bir arayüz geliştirilmiş ve bu arayüz ile müşteri, gerçekleştireceği işlemler için adım adım yönlendirilmiştir. Bu çalışmaya ait uygulamanın, algoritma şeması Şekil 4.1’de gösterilmiştir.



Şekil 4.1. Akış Şeması

Öncelikle müşterinin kendisine ait kişisel bilgileri doldurulur. Sonrasında kimlik kartının ön ve arka yüzlerinin fotoğrafı çekilir. Ve kimlik kartı bilgileri OCR teknolojisi kullanılarak ayrıştırılır. İlk adımdaki müşterinin doldurmuş olduğu bilgiler ile OCR teknolojisi kullanılarak elde edilmiş kimlik kartı üzerindeki bilgiler NLP kütüphaneleri kullanılarak karşılaştırılır ve bir doğruluk oranı hesaplanır. Bu doğruluk oranı %100 ve üzerinde değilse canlılık tespiti adımına hiç geçilmez. %100’ün üzerinde ise; müşteri güvenlik sorusu ile karşılaşır ve bu soruya sesli yanıt verir. Soruya doğru yanıt verilmiş ise OpenCV kütüphanesi kullanılarak kamera açılır ve kişinin anlık görüntüsü yakalanır. Bu anlık canlı görüntüsü ile kimlik kartı üzerindeki biyometrik fotoğrafı FaceRecognition ve CNN yüz tanıma algoritmaları kullanılarak yüzdelik olarak benzerlik oranı hesaplanır ve analizi yapılır. Bu analiz sonucunda kişi gerçektir ya da değildir sonucuna ulaşılır.

Uygulama Anaconda IDE'si üzerindeki Spyder geliştirme ortamında Python dili ile birlikte, arayüzü PyQt kullanılarak geliştirilmiştir. Geliştirilen uygulama, bu çalışma özelinde Bankacılıkta Uzaktan Müşteri Edinimi için müşterinin gerçekleştirilmesi gereken adımların sunulduğu arka planda ise tez kapsamında belirtilen tüm doğrulama ve analiz işlemlerinin yapıldığı bir çıktıdır. Şekil 4.2'den itibaren Şekil 4.10'a kadar uzaktan müşteri ediniminde kimlik tespiti için gerçekleştirilecek adımların uygulama üzerindeki her bir aşamaya ait ekran görüntülerinin görselleri yer almaktadır. Her bir sayfada müşterinin uygulaması gereken işlemler açıklayıcı bir metin ile belirtilmiştir. Ayrıca her bir sayfada gerekli işlemler yapıldıktan sonra, bir sonraki adıma geçilmesi için buton yer almaktadır.

The image shows a mobile application interface for a customer registration process. The screen is titled 'STEP 1' and contains a form with the instruction 'Please fill in your identity information'. The form has the following fields: Given Name, Surname, TR Identity No, Date of Birth, Gender (M/F), Mother's Name, Father's Name, and Nationality (TR). Below the form are two buttons: 'CLEAN' (orange) and 'NEXT STEP' (purple). The background is a light gray with a large, faint 'X' watermark.

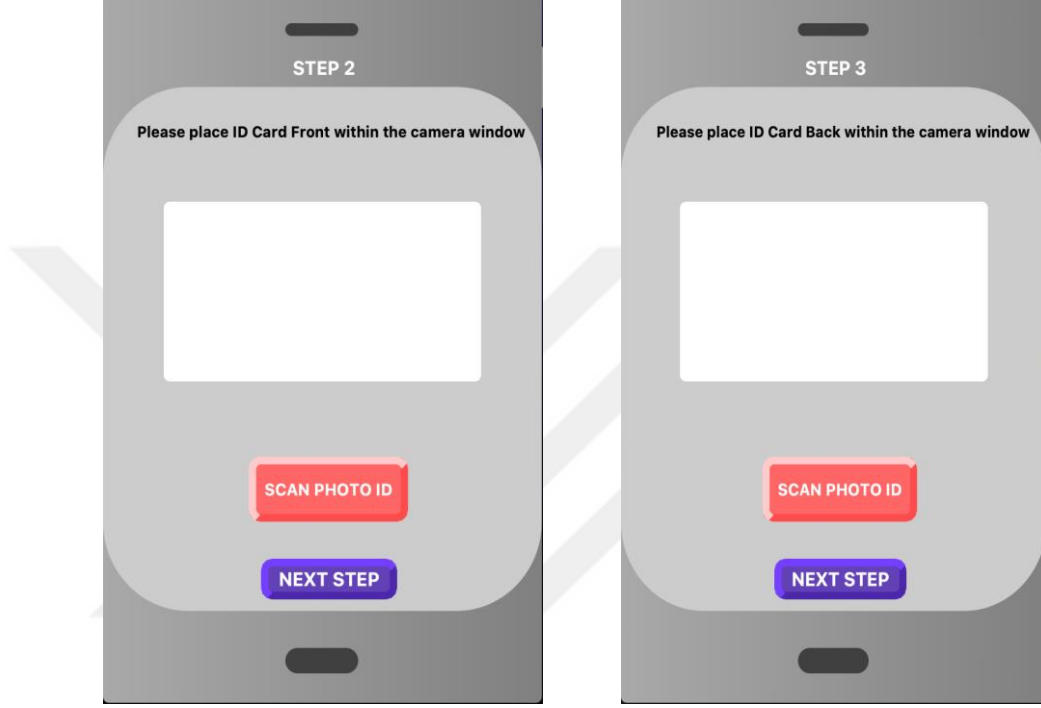
Şekil 4.2. 1.Adım

Şekil 4.2' de yer alan görsel, müşterinin uzaktan müşteri edinimi yaklaşımı ile dijital bir platform üzerinden müşteri olmak istemesi talebine istinaden karşılaştığı ilk adımdır. Bu adımda; kişinin kendisine ait bilgileri belirtilen alanlara uygun bir biçimde doldurması beklenmektedir. Doldurulan kişi bilgileri, .csv uzantılı bir belgeye kaydedilir. Bu belgeye yazılırken Python'ın DictWriter kütüphanesinden faydalanılmıştır. Belge formatı Şekil 4.3' teki gibidir. Dosyada yer alan 'Anne Önceki Soyadı' bilgisi kişinin kimlik kartı bilgilerinden alınacağı için ilk adımda boş görünmektedir.

personinfo.csv

Id	Name	Surname	Bdate	Gender	Nationality	Mname	Fname	OldMotherSurname	CardOrPerson
60013015040	şeyma nur	alkan	1996/06/08	f	tr	salıha	hüseyin		P

Şekil 4.3. Kişi bilgilerinin kaydedildiği .csv dosyası



Şekil 4.4. 2. Ve 3. Adım

Şekil 4.4.'te çalışmanın 2. aşaması olan müşterinin T.C. Kimlik Kartı bilgilerinin alındığı, uygulamaya ait ekran resimleri gösterilmektedir. Müşteri, kendisine ait kimlik kartının ön ve arka yüzünü ekranda belirtilen çerçeveye yerleştirerek taramayı başlatır. Taranan kimlik belgesindeki bilgiler bir sonraki adımda müşteriye gösterilecektir. Şekil 4.5'te ise bir önceki adımda müşteriden alınan kimlik kartı ön ve arka yüz fotoğraflarına OCR (Optik Kart Tanıma) teknolojisi uygulanarak, kart üzerindeki bilgiler sayfa üzerinde yer alan bilgilerin ayrılmış alanlarına doldurulur. Python üzerinde OCR (Optik Karakter Tanıma) tekniği için PyTesseract kütüphanesi kullanılmıştır. Kurulumu için komut satırından;

- *pip install pytesseract*

komutu kullanılır. Aşağıda, indirilen kütüphanenin tanımlanması ve fotoğrafın ön işleme yapılmış hali 'image_to_data' fonksiyonu ile metne dönüştürülme işleminin yapıldığına dair kısa bir örnek kod bloğu gösterilmiştir.

```
import pytesseract
img_data = pytesseract.image_to_data(
    out_rgb,
    lang="tur",
    config=config,
    output_type=pytesseract.Output.DATAFRAME,
)
```

Şekil 4.5. 4. Adım

Aynı zamanda kimlik kartı resimlerinden alınan bu bilgiler, ilk adımdaki müşterinin kendisinden alınan bilgiler ile karşılaştırılmak üzere .csv uzantılı belgeye kaydedilir. Belge örneği Şekil 4.6'da gösterilmektedir.

personinfo.csv

Id	Name	Surname	Bdate	Gender	Nationality	Mname	Fname	OldMotherSurname	CardOrPerson
60013015040	şeyma nur	alkan	1996/06/08	f	tr	salıha	hüseyin		P
60013015040	ŞEYMA NUR	ALKAN	1996/06/08	F	TR	SALIHA	HÜSEYİN	Işık	C

Şekil 4.6. Kimlik kartından okunan bilgilerin de eklendiği .csv dosyası

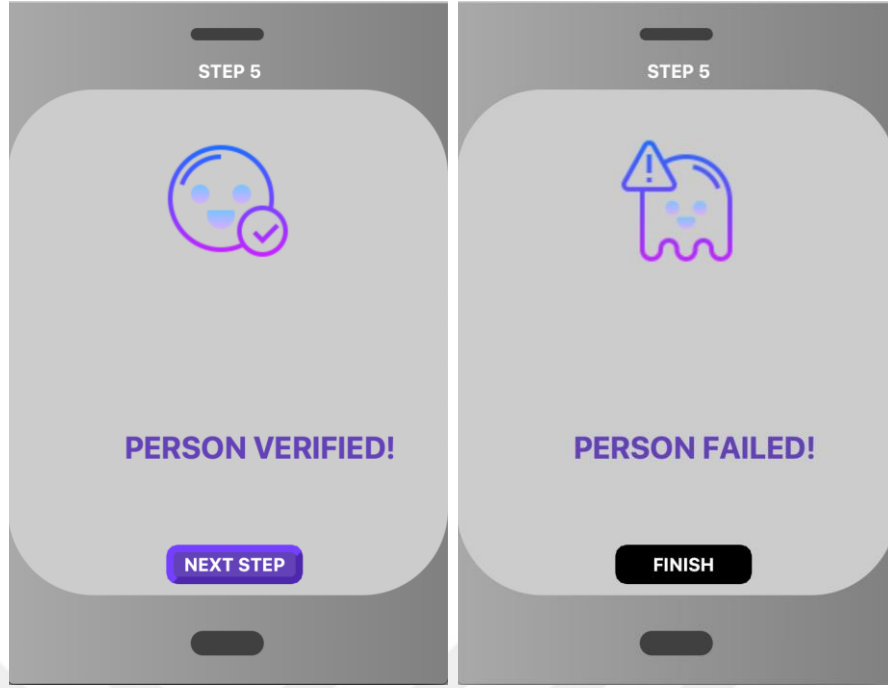
Dosyanın 2.satırında yer alan OCR (Optik Karakter Okuma) teknolojisi kullanılarak kimlik kartı üzerinden okunan bilgilerde ayrıca ‘Anne Eski Soyadı’ bilgisi de doldurulmuştur. Bu bilgi, kimlik kartının arkasında bulunan çip üzerinde yer alan bilgilerden biridir ve bu çip okunarak elde edilecektir. ‘Anne Eski Soyadı’ kişisel bilgisine, uygulamanın sonraki adımlarında güvenlik sorusu olarak kullanılmak için ihtiyaç duyulmuştur. Uygulamanın 4. adımından 5. adımına geçerken, Şekil 4.6’daki belge örneğinde yer alan 2 satır, NLP (Doğal Dil İşleme) tekniği olan FuzzyWuzzy kütüphanesi kullanılarak karşılaştırılmıştır. FuzzyWuzzy kütüphanesinin kullanılması için öncelikle kütüphanelerin indirilmesi gerekmektedir. Bahsedildiği üzere bu kütüphanenin arka tarafında Levenshtein metriği kullanıldığı için ikisinin birlikte kurulması gerekmektedir. Komut satırından sırasıyla;

- *pip install fuzzywuzzy*
- *pip install python-Levenshtein*

komutları yazıldığı takdirde gerekli kütüphaneler ve fonksiyonlar otomatik olarak indirilecektir. Sonrasında, kütüphane aşağıdaki kod parçasığında gösterildiği gibi tanımlanarak çeşitli fonksiyonları kullanılabilir.

```
from fuzzywuzzy import fuzz
import csv
file = open('personinfo.csv')
csvreader = csv.reader(file)
rows = []
for row in csvreader:
    rows.append(row)
verify = fuzz.WRatio(rows[1],rows[2])
```

Kod örneğinde, öncelikle .csv uzantılı dosyada yer alan satırlar okunmuş, sonrasında dosyada yer alan 2 satır ‘WRatio’ fonksiyonuna parametre olarak verilerek benzerlik oranı elde edilmiştir. WRatio fonksiyonu, iki cümle arasındaki büyük-küçük harf farklılıklarına takılmadan bir benzerlik oranı hesaplamaktadır. Elde edilen benzerlik oranı %100 olduğunda Şekil 4.7’de yer alan ilk görseldeki gibi kişinin ilk doğruluk aşaması tamamlanmış sayılır. Eğer benzerlik oranı, %100 altında ise Şekil 4.7’de yer alan 2. görseldeki gibi süreç sonlanmaktadır. Müşterinin kendisi tarafından doldurması istenen kişisel bilgilerinin, kimlik kartı üzerindeki bilgileri ile eşleşmediği bu aşamada hata kabul edilmemektedir.



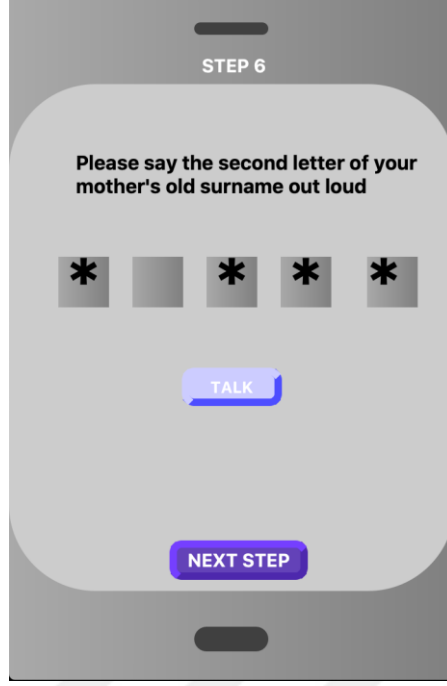
Şekil 4.7. 5. Adım

Müşterinin ilk adımda doldurmuş olduğu kişisel bilgileri ile kimlik kartından okunan bilgiler eşleştğinde Şekil 4.8’de ekran görüntüsü yer alır ve sürecin 6. Adımı olan güvenlik sorusu adımına geçilir. Bu adımda güvenlik sorusu olarak ‘anne önceki soyadının ikinci harfi’ sorulmuştur ve bu sorunun cevabının sesli olarak söylenmesi istenmiştir. Yanıtın sesli olarak alınması ve konuşmanın string bir ifadeye dönüştürülmesinde NLP (Doğal Dil İşleme) tekniği kütüphanelerinden olan SpeechRecognition kütüphanesi kullanılmıştır. Bu kütüphane bir ses dosyasını veya mikrofondan aldığı sesi yazıya dönüştürür. Kurulumu komut satırından;

- *pip install SpeechRecognition*

komutu kullanılarak yapılmaktadır. Aşağıdaki kod parçacığında ise, SpeechRecognition kütüphanesi tanımlandıktan sonra mikrofondan alınan sesin Türkçe dil desteği ile yazıya dönüştürülmesi aktarılmaktadır.

```
import speech_recognition as sr
r = sr.Recognizer()
with sr.Microphone() as source:
    r.adjust_for_ambient_noise(source)
    data = r.record(source, duration=5)
    r.recognize_google(data, language="tr")
text =
```

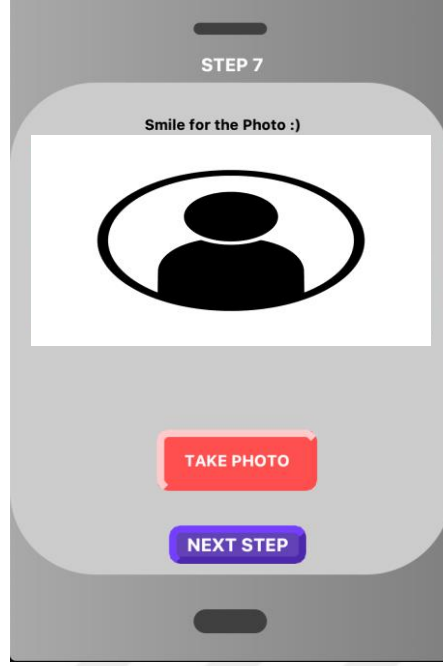
Şekil 4.8. 6. Adım

Bu adımda, müşterinin 3 kere yanlış yanıt vermesi halinde süreç sonlanmakta ve süreci baştan başlatmasına engel olunmaktadır. Yanıtın doğru olması durumunda ise son aşama olan pasif canlılık testi adımına geçilmektedir. Şekil 4.9'da ekran görüntüsünde görüldüğü gibi belirtilen çerçevede kamera açılmaktadır. Kişi yüzünü net gösterecek biçimde çerçeveye yerleştirir ve kişinin anlık bir görüntüsü yakalanır. Kameranın açılması ve anlık fotoğrafın çekilmesinde OpenCV kütüphanelerinden yararlanılmıştır. Kurulum için;

- *pip install opencv-python*

komutu kullanılmıştır. Kod parçacığı; kütüphane tanımı yapılması, kameranın açılması ve kişinin anlık yakalanan fotoğrafının .jpg uzantısı ile kaydedilmesini ifade etmektedir.

```
import cv2
video_capture = cv2.VideoCapture(0)
ret, frame = video_capture.read()
cv2.imshow('Video', frame)
if cv2.waitKey(1):
    check, frame = video_capture.read()
    cv2.imwrite(filename='x.jpg', img=frame)
```



Şekil 4.9. 7. Adım

Kaydedilen müşteriye ait anlık görüntü ile karşılaştırılmak üzere FaceRecognition kütüphanesinden faydalanılarak kimlik kartının ön yüzünde yer alan biyometrik resim, kimlik kartı fotoğrafı üzerinden ayırt edilir. FaceRecognition kütüphanesinin kurulumu için;

- *pip3 install face_recognition*

komutu kullanılır. Kod bloğunda, FaceRecognition kütüphanesi tanımlanması ve kimlik kartı fotoğrafı üzerinde yer alan biyometrik resmin, 'load_image_file' ve 'face_locations' fonksiyonları kullanılarak ayırt edilmesi ifade edilmektedir.

```
import face_recognition
image = face_recognition.
    load_image_file("kimlikkart.jpg")
face_locations = face_recognition.face_locations(image)
```

Kişinin biyometrik fotoğrafı ile kişinin anlık fotoğrafının karşılaştırılması için CNN (Konvolüsyonel Nöral Ağ) algoritması kullanılmıştır. Biyometrik fotoğrafın konvolüsyonel nöral ağ ile özeti (embedding) çıkarılarak, anlık yakalanan fotoğrafın özeti ile karşılaştırılacak ve benzerlikleri bir eşik değerden daha düşük ise eşleştirilmiş olacaktır. Konvolüsyonel nöral ağın çıktısı, çıktı boyutunda bir vektör olarak ifade edilir. Dolayısıyla iki ayrı fotoğraf, aslında iki vektör olarak düşünülür.

Vektörler yönü ve uzunluğu olan niceliklerdir. İki vektörün benzerlikleri aralarındaki açı yani kosinüs benzerliği (Denklem 4.1) veya aralarındaki uzaklık yani Öklid uzaklığı (Denklem 4.2) kullanılarak bulunmaktadır.

$$\text{similarity}(A, B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n A_i^2} \times \sqrt{\sum_{i=1}^n B_i^2}} \quad (4.1)$$

İki vektör arasındaki kosinüs benzerliği iki vektörün çarpımının, iki vektörün boylarının çarpımına oranı olarak ifade edilmektedir.

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (4.2)$$

X ve y olarak iki nokta arasındaki Öklid Uzaklığı ise noktaların karşılıklı gelen koordinatları arasındaki değerlerin farkların kareleri toplamının karekökü olarak tanımlanmaktadır.

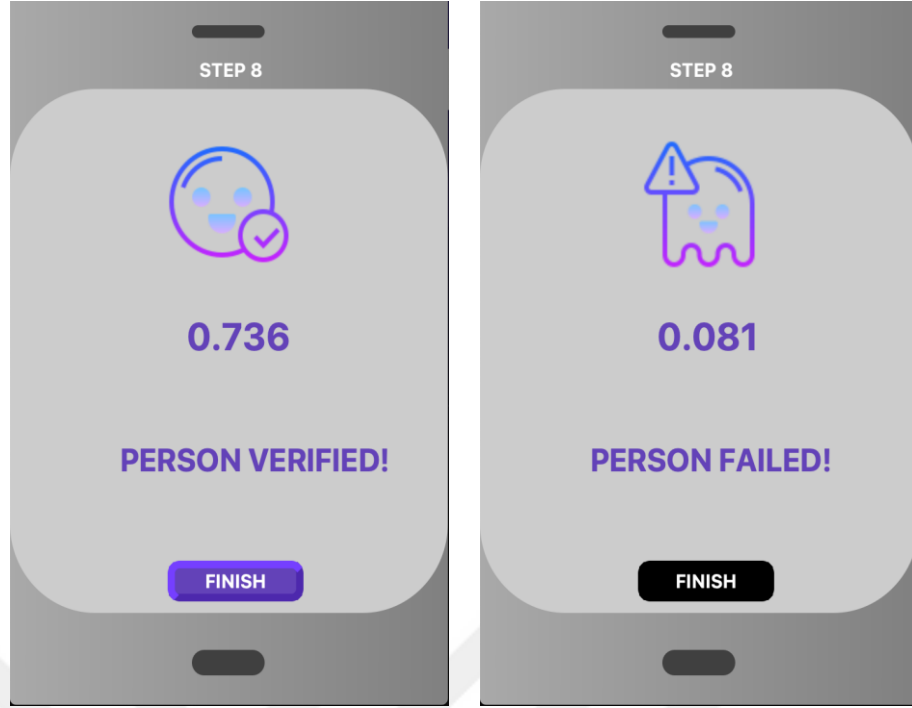
Bu çalışmada, çoklu testler sonucunda kosinüs benzerliğinden daha doğru sonuçlar elde edildiği için, benzerlik oranı hesaplanmasında kosinüs benzerliği kullanılmıştır. Python'ın numpy kütüphanesinden faydalanılarak aşağıdaki kod parçacığında görüldüğü gibi kosinüs benzerliğinin hesaplanması için fonksiyon yazılmıştır.

```
def findCosineDistance(source_representation, test_representation):
    a = np.matmul(np.transpose(source_representation), test_representation)
    b = np.sum(np.multiply(source_representation, source_representation))
    c = np.sum(np.multiply(test_representation, test_representation))
    return 1 - (a / (np.sqrt(b) * np.sqrt(c)))
```

Karşılaştırılacak resimler bu fonksiyona parametre olarak gönderilir ve benzerlik oranı hesaplanır.

```
cosine_distance = findCosineSimilarity(img1, img2)
```

Benzerlik oranı 0.40'dan küçükse kişinin fotoğrafları eşleşmiş olur ve yüzde kaç oranında benzerlik yakalandığı hesaplanır. Şekil 4.10'da yer alan ilk görseldeki gibi müşteriye doğrulandığına dair bilgi verilir. Kişi eşleşmediyse Şekil 4.10'da 2. görseldeki gibi uyarı verilir ve süreç sonlanır.



Şekil 4.10. 8. Adım

4.1. Veri Kümesi

Veri setinde kişiye ait kimlik kartımızda da yer alan temel bilgiler yer almaktadır. Bu veriler; T.C. kimlik numarası, ad, soyad, doğum tarihi, cinsiyet, uyruk, anne adı ve baba adı bilgileridir.

KVKK (Kişisel Verileri Koruma Kanunu)'ya ait kişisel verilerin işlenmesine ilişkin hüküm ihlal edilmemiş ve kullanılan veri seti rastgele belirlenmiş kişi bilgilerinin kombinasyonları şeklinde üretilmiştir. Gerçek kişi veya kurumlarla herhangi bir ilişkisi yoktur. Çizelge 4.1'de bu çalışmada kullanılan veri setine ait örnek kayıtlar gösterilmiştir.

Çizelge 4.1. Müşteri Bilgileri Veri Seti

Id	Name	Surname	Bdate	Bplace	Gender	Nationality	Mname	Fname
10612935850	Yusuf	Alkan	21.03.1999	TURHAL	M	TR	ZEKİYE	ERDAL
12272522586	Yiğit	Koç	01.11.1963	KATRANCI	M	TR	ZEYNEP	HÜSEYİN
11743831110	Mehmet Ali	Işık	17.07.1991	SÖKE	M	TR	SÜHEYLA	MUSTAFA
10022546684	Ali	Kaya	31.05.1979	İSTANBUL	M	TR	AYŞE	GÖKHAN
12010816212	Havva	Tekin	18.01.2000	FATİH	F	TR	SEYHAN	İHSAN İLKER
12496276210	Zeynep	Aslan	24.08.2001	BOLU	F	TR	SANIYE	KAMİL SAMİ

Şekil 4.3'de yer alan görsel, üzerinde işlemler yapılan diğer bir veri olan örnek kimlik kartıdır. OCR ile müşteri bilgilerinin eşleştirilmesi ve canlılık tespitinde kimlik kartı bilgileri kullanılmıştır.



Şekil 4.3. Müşteri Kimlik Verisi

4.2. Veri Ön İşleme

Veriler düzensiz halde bulunabilir, bozuk kayıtlar içerebilir ya da uygulayacağımız analize uygun halde bulunmayabilirler. Daha iyi sonuçlar elde edebilmek üzere çeşitli filtrelerden geçirmek gerekmektedir. Bu ve benzeri birçok duruma müdahaleyi içeren yöntemler bütününe veri ön işleme denir. Veri kalitesi, veri yönetimindeki en önemli sorunlardan biridir. Ön işleme yapılmamış veriler genellikle yanlış veri analitiği sonuçlarına yol açmaktadır.

Bu aşamada sınıflandırma için kullanılacak veri setindeki verileri elde etme ve ön işleme için Python kütüphanelerinden faydalanılmıştır. Aynı zamanda bir makine öğrenimi modeli oluşturmak için ilk adımdır. Şekil 4.4'te kimlik kartı üzerinde OCR işlemi yapılmadan önce görüntünün önce gri, sonrasında binary görüntüsüne dönüştürülmesi gösterilmiştir.



Şekil 4.4. Kimlik Kartı Veri Ön İşlemleri

Bu çalışmada NLP tekniği müşterinin kendi doldurduğu bilgiler ile kimlik kartından OCR ile okunmuş verilerin karşılaştırılıp analiz edilmesinde kullanılmıştır. Benzerlikleri yakalamak için ise FuzzyWuzzy Doğal Dil İşleme kütüphanesinden faydalanılmıştır. FuzzyWuzzy, diziler ve desenler arasındaki farkları hesaplamak için sadece Levenshtein Distance metriğini kullanan ve SeatGeek tarafından geliştirilen açık kaynaklı bir Python kütüphanesidir. Fuzzy (Bulanık) string eşlemesi, kullanıcılar kelimeleri yanlış yazdığı veya yalnızca kısmi kelimeler girdiğinde bile eşleşmeleri bulabilen bir arama türüdür. Bu yüzden bu kütüphane tercih edilmiş ve çeşitli testler sonucunda doğru sonuçlar verdiği teyit edilmiştir. Sesli yanıt adımında Speech Recognition kütüphanesinden faydalanılmış ve kullanıcının verdiği yanıt texte dönüştürülerek doğru cevap ile karşılaştırma yapılmıştır. Karşılaştırma sonucuna göre diğer bir doğruluk adımı olan yüz tanımaya geçilmiştir. Kişi belirtilen çerçeve alanına yüzünü yerleştirmeye yönlendirilmiş ve kişinin anlık bir görüntüsü yakalanmaya çalışılmıştır. Kamera açılması ve kişinin anlık fotoğrafının yakalanmasında OpenCv

kütüphanelerinden yararlanılmıştır. Yakalanan anlık görüntü ile müşterinin kimlik kartı görüntüsünden, FaceRecognition kütüphanesi ile kişinin biyometrik fotoğrafı ayırt edilmiştir. Ve bu görüntü ile anlık yakalanan görüntü CNN algoritması kullanılarak karşılaştırma yapılmış ve bir doğruluk oranı hesaplanmıştır. Bu doğruluk oranına göre kişinin gerçek kişi olup olmadığına karar verilmiştir.



5. SONUÇ

Uzaktan Müşteri Edinimi süreci, fiziki prosedürlerin (aslen şubede bulunma, sözleşmeye imza atma, fiziken kimlik gösterme vb.) olmadığı, sürecin tamamen mevzuata uygun bir şekilde dijital platformlardan gerçekleştirildiği bir süreçtir. Bu çalışmada, Bankalardaki Uzaktan Müşteri Edinimi sürecinin uygulama üzerinde tamamen otomatik olarak yürütülmesi amaçlanmıştır. Böylece uzaktan kimlik tespiti için, tüm adımlar sonrasında yapılan müşteri ile müşteri temsilcisinin görüntülü görüşüp kişinin gerçekliğinin teyit edilmesine ihtiyaç duyulmayacaktır. Müşteri olma süreci için; Python dilinde PyQt kullanılarak bir uygulama kodlanmıştır. Bu uygulama üzerindeki kimlik tespit aşamalarında Makine öğrenmesi algoritmalarından faydalanılmıştır. Aynı zamanda OCR, NLP, OpenCV gibi teknolojiler kullanılarak doğruluk analizi yapılmıştır. Uygulama çıktılarına göre; kişinin doğruluk modellemesi en iyi oranlarda tahmin edilmektedir. Yapılan analiz sonuçlarına göre; OCR ve NLP algoritmaları kullanılarak bilgilerin eşleşmesi adımında; doğru verileri 90% ve üzeri oranında doğru, yanlış verileri %0,09 oranında yanlış bularak doğru bir şekilde tahminlemiştir. Son adım olan yüz tanıma kısmında ise; CNN algoritmasında kosinüs benzerliği kullanarak doğru verileri %73,6 ve üzeri oranda doğru, yanlış verileri ise %0,081 oranında yanlış bularak doğru bir şekilde tahminlemiştir. Yüz tanıma kısmı aşmasına geçilmeden hemen öncesinde, kimlik bilgilerinin sahibi olan müşteri dışında başka bir kişinin herhangi bir ortamdan ya da belge üzerinden erişemeyeceği; yani doğru cevabı kişinin sadece kendisinin bilebileceği bir güvenlik sorusu sorularak, bu soruya sesli yanıt verilmesi istenmiştir. Bu sayede müşteri doğruluğu bir kez daha teyit edilmiş ve doğruluk tahmini güçlendirilmiştir.

Bu çalışma ile bankacılık ve finans sektöründe dijital platformlar üzerinden bankalara müşteri sağlama sürecindeki işlem sürelerinde, fark edilir düzeyde bir azalış beklenmektedir. Hem zaman tasarrufu hem de mevcut sistemdeki fazladan efor gerektiren görüntülü görüşme adımına gerek duyulmayacağından dolayı, müşterilerden daha hızlı, daha verimli ve daha güvenilir işlem sağlandığına dair pozitif yönde geri bildirimlerin alınacağı öngörülmektedir.

KAYNAKLAR

- Aggarwal, C. C. (2015), Data mining: the textbook, Springer, p.
- Aslandođdu Ö.F. & Aydın H. & Çetinkaya A. (2022), “Bulut üzerinde Tensorflow Javascript ile geliştirilen Öğretilebilir Makine (ÖM) Modelinin Doğruluk Tahmini ve Analizi: Yüz Tanıma Sistemi Uygulaması” , Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi,(2022), 6(1): 66-77
- Aydın Ö. (2005). “Yapay Sinir Ağlarını Kullanarak Bir Ses Tanıma Sistemi Geliştirilmesi.” Yüksek Lisans Tezi, Trakya Üniversitesi, Edirne, 74 s.
- Aytekin A. (2017). Basit Düşün, Akış Diyagramları ile Programlama, Detay Yayıncılık, Ankara, 326 s.
- Balasubramanian, V., Vivekanandhan, S., & Mahadevan, V. (2022). Pandemic tele-smart: a contactless tele-health system for efficient monitoring of remotely located COVID-19 quarantine wards in India using near-field communication and natural language processing system. *Medical & biological engineering & computing*, 60(1), 61-79.
- Balkan H. (2021) , “Dijital (Şubesiz) Bankaların Regülasyonu”, Bankacılar Dergisi, Sayı 116, 2021
- Bektaş B. & Babur S. & Turhal U. & Köse E. (2016) “Makine Öğrenmesi Yardımıyla Optik Karakter Tanıma Sistemi” , Conference: 5. Uluslararası Matbaa Teknolojileri Sempozyumu ,(Nov,2016)
- Bibi, K., Naz, S., & Rehman, A. (2020). Biometric signature authentication using machine learning techniques: Current trends, challenges and opportunities. *Multimedia Tools and Applications*, 79(1), 289-340.
- Biggio, B., Russu, P., Didaci, L., & Roli, F. (2015). Adversarial biometric recognition: A review on biometric system security from the adversarial machine-learning perspective. *IEEE Signal Processing Magazine*, 32(5), 31-41.
- Breiman, L. (1996), Bagging predictors, *Machine learning*, 24 (2), 123-140.
- Çınar U.K. (Eylül 1, 2019), Veri Bilimi Okulu, (<https://www.veribilimiokulu.com/cv2-ile-yuz-tanima-ve-belirleme/>) [Ziyaret Tarihi: 25.12.2022]
- Deniz V. (Eylül 21, 2020) (<https://www.procompliance.net/bankalarca-kullanilacak-uzaktan-kimlik-tespiti-yontemlerine-iliskin-teblig-taslagi/>) [Ziyaret Tarihi: 20.12.2022]
- Dhawle, T., Ukey, U., & Choudante, R. (2020). Face detection and recognition using OpenCV and Python. *Int. Res. J. Eng. Technol*, 7(10).
- D. H. Hubel and T. N. Wiesel, (1968) ,“Receptive fields and functional architecture of monkey striate cortex,” *J. Physiol.*, vol. 195, no. 1, pp. 215–243, (Mar. 1968).

- Dwi H, Purnomo H, Purwanto H. (2019), “Utilization of Optical Character Recognition Technology in Reading Identity Cards”, *Int. Journal of Information Technology and Business*, Vol. 2, No. 1(2019) 38-46
- Efe, & Kaynak, O. (2000). *Yapay Sinir Ağları ve Uygulamaları*: Boğaziçi Üniversitesi.
- Ekonomi Haberleri. (Kasım 10, 2021), (<https://www.alomaliye.com/2021/11/10/ekim-2021-uzaktan-ve-diger-kanallardan-musteri-edinimi/>)
[Ziyaret Tarihi: 22.12.2022]
- Eser, A. (2021), “Yapay Sinir Ağları ile Bankacılık Dökümanlarının Sınıflandırılması “(Doctoral dissertation, Necmettin Erbakan University (Turkey))
- Goode A. (2018), “Biometrics for banking: best practices and barriers to adoption”, *Biometric Technology Today* Vol.(2018), No. 10
- Haykin S (1994). *Neural Networks: A Comprehensive Foundation*, Prentice Hall, NJ.
- J. Lladós & F. Lumbreras & V. Chapaprieta & J. Queralt.(2001) “ICAR: Identity Card Automatic Reader” , *Proceedings of Sixth International Conference on Document Analysis and Recognition* , (13 Sept. 2001)
- Kakadiaris, I. A., G. Passalis, G. Toderici, M. N. Murtuza, Y. Lu, N. Karampatziakis, T. Theoharis, (2004), “Threedimensional face recognition in the presence of facial expressions: an annotated deformable model approach,” *IEEE Trans. PAMI*, vol. 29, no. 4, pp. 640-649,
- Kamath, U., Liu, J., & Whitaker, J. (2019). “Deep learning for NLP and speech recognition” (Vol. 84). Cham, Switzerland: Springer.
- Kardaş K. & Güveni A. , (2019) “Analysis of the Effects of Quizzes, Homeworks and Projects on Final Exam with Different Machine Learning Techniques” *EMO Bilimsel Dergi - Haziran 2020 - Cilt: 10 Sayı:1*
- Kavzoğlu, T. ve Çölkesen, İ., (2010), “Karar ağaçları ile uydu görüntülerinin sınıflandırılması: Kocaeli örneği” , *Harita Teknolojileri Elektronik Dergisi*, 2 (1), 36-45.
- Kılınç, D., & Başeğmez, N. (2018). *Uygulamalarla Veri Bilimi*. İstanbul: Abaküs.
- K.N. Fukushima, (1980) , “A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position.,” *Biol. Cybern.*, vol. 36, no. 4, pp. 193–202.
- Koukoumidis, Emmanouil & Hasan, Shiblee & Abramson, Dustin & Johnson Jr., Joseph Edwin. (2021) "Improving Remote Customer Interaction Experiences Using Machine Learning", *Technical Disclosure Commons*, (July 21, 2021)
- Nanonets, (<https://nanonets.com/blog/ocr-with-tesseract/>)

[Ziyaret Tarihi: 15.12.2022]

Nelli V. Tskhadadze, Financial University of the Government of the Russian Federation, Moscow, Russia, “Use of Remote Banking Technology” Advances in Social Science, Education and Humanities Research, volume 273 International Conference on Communicative Strategies of Information Society, (CSIS 2018)

Ormanlı C. (Ocak 30, 2020) , Medium ,(https://medium.com/kronnika/python-ile-ocr-optical-character-recognition-uygulamas%C4%B1-963d031fb036)
[Ziyaret Tarihi: 05.12.2022]

Pires M. R. (2022), “Digital Onboarding in the Customer Registration Process for Financial Services and Products”, Repositório Aberto da Universidade do Porto FEUP- Faculdade de Engenharia FEUP - Dissertação (2022-07-21)

Pisarevsky, V., (2007), “Introduction to OpenCV”, Intel Corporation, Software and Solutions Group, 2-67.

Pişkin, M. (2016), “Opencv ile görüntü işleme”.

Resmi Gazete Sayı 31441, (Nisan 1, 2021, Perşembe)
(https://www.resmigazete.gov.tr/eskiler/2021/04/20210401-7.htm)
[Ziyaret Tarihi: 25.12.2022]

Rohit Thakur , (Nov 23,2020)
(https://towardsdatascience.com/step-by-step-face-recognition-code-implementation-from-scratch-in-python-cc95fa041120) adresinden alındı.

Rusli F. M., Adhiguna K. A., Irawan H., “Indonesian ID Card Extractor Using Optical Character Recognition and Natural Language Post-Processing”, Cornell University arXiv:2101.05214v1 [cs.CV] (15 Dec 2020)

Samuel, A. L., (1959), Some studies in machine learning using the game of checkers, IBM Journal of research and development, 3 (3), 210-229.

Seker, S.E., (2015), Doğal Dil İşleme(Natural Language Processing) , YBS Ansiklopedi, Cilt 2, Sayı 4, (Aralık 2015)

Senyurt,B.S., (2012), “https://www.buraksenyurt.com/post/Levenshtein-Distance-Algoritmasi”

Smith, R. (2007, September). “An overview of the Tesseract OCR engine. In Ninth international conference on document analysis and recognition “ , (ICDAR 2007) (Vol. 2, pp. 629-633). IEEE.

TBB, Uzaktan ve Şubeden Müşteri Edinimi İstatistikleri, (Kasım,2022)
(https://www.tbb.org.tr/Content/Upload/istatistikraporlar/ekler/3968/Uzaktan_ve_Subeden_Musteri_Edinimi_Istatistikleri-Kasim_2022.pdf)

VanderPlas, J., 2017, Python data science handbook, oreilly, p. 331-333.

Wikipedia , 2021 , OCR , https://tr.wikipedia.org/wiki/Optik_karakter_tan%C4%B1ma :
[Ziyaret Tarihi: 25.11.2022].

Yenen, S. Ü., & Şahin, G. (2021, November). Digital Customer Acquisition and Onboarding During the Pandemic. In 2021 15th Turkish National Software Engineering Symposium (UYMS) (pp. 1-4). IEEE.

