



**T.C. İSTANBUL T CARET
 NİVERSİTESİ**

LİSANS ST  E TİM ENSTİT S 

**YAPAY ZEKA DESTEKLİ BLOKZİNCİR TABANLI
FİNANSAL BİLGİ PAYLAŞIM PLATFORMU**

Bulut KARADA 

**Danışman
Prof. Dr. Abd l Halim Z İM**

**İkinci Tez Danışmanı
Do . Dr. Akhan AKBULUT**

**DOKTORA TEZİ
BİLGİSAYAR M HENDİSLİ İ ANABİLİM DALI
İSTANBUL – 2024**

KABUL VE ONAY SAYFASI

Bulut KARADAĞ tarafından hazırlanan " **Yapay Zeka Destekli Blokzincir Tabanlı Finansal Bilgi Paylaşım Platformu**" adlı tez çalışması 03/09/2024 tarihinde aşağıdaki jüri üyeleri önünde başarı ile savunularak, İstanbul Ticaret Üniversitesi Fen Bilimleri Enstitüsü **Bilgisayar Mühendisliği Anabilim Dalı**'nda **DOKTORA TEZİ** olarak kabul edilmiştir.

Danışman **Prof. Dr. Abdül Halim ZAİM**

İstanbul Teknik Üniversitesi

Jüri Üyesi **Doç. Dr. Mustafa Cem KASAPBAŞI**

İstanbul Ticaret Üniversitesi

Jüri Üyesi **Dr. Öğretim Üyesi Metin TURAN**

İstanbul Ticaret Üniversitesi

Jüri Üyesi **Prof. Dr. Fatih ALAGÖZ**

Boğaziçi Üniversitesi

Jüri Üyesi **Doç. Dr. Muhammed Ali AYDIN**

İstanbul Üniversitesi - Cerrahpaşa

Onay Tarihi :

Prof. Dr. Hanifi PARLAR
Enstitü Müdürü

AKADEMİK VE ETİK KURALLARA UYGUNLUK BEYANI

İstanbul Ticaret Üniversitesi, Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

Bulut KARADAĞ

İÇİNDEKİLER

	Sayfa
İÇİNDEKİLER.....	i
ÖZET	iii
ABSTRACT	iv
TEŞEKKÜR.....	v
ŞEKİLLER.....	vi
ÇİZELGELER.....	vii
SİMGELER VE KISALTMALAR.....	viii
1. GİRİŞ	1
1.1 Problemin Tanımı.....	4
1.2 Araştırma Önemi.....	4
1.3 Araştırma Amacı	5
1.4 Araştırma Kapsamı	5
1.5 Hedef Kitle.....	6
1.6 Araştırma Metodolojisi.....	6
1.7 Tez Organizasyonu.....	7
2. LİTERATÜR ÖZETİ.....	9
2.1 Blokzincir Tabanlı Finans Çalışmaları.....	9
2.2 Yapay Zeka Destekli Kredi Skor ve Otomatik Karar Çalışmaları.....	11
3. YAPAY ZEKA İLE KREDİ SKORU TESPİTİ.....	20
3.1 İş Akışı.....	21
3.2 Veri Seti	22
3.3 Makine Öğrenmesi Algoritmaları	23
3.4 Uygulama	27
3.4.1 Veri ön işleme.....	27
3.4.2 Veri dönüşümü	32
3.4.3 Makine öğrenmesi ile modelleme.....	33
3.5 Kredi Skor Tespiti Önyüz Çalışmaları.....	35
3.6 Sonuçlar ve Tartışma	36
4. SOSYAL MEDYA ANALİZİ.....	38
4.1 Mimari Topolojisi	39
4.2 Veri Kaynağı.....	39
4.3 Üretken Yapay Zeka.....	40
4.4 Uygulama	41
4.5 Sosyal Medya Analiz Sonuçları ve Kazanımlar.....	42

5. BLOKZİNCİR TABANLI FİNANSAL VERİ PAYLAŞIMI.....	43
5.1 Mimari Topolojisi	45
5.2 Blokzincir Teknolojisi	46
5.2.1 Blokzincir kalite özellikleri	46
5.2.2 Blokzincir türleri.....	50
5.2.2.1 Halka açık Blokzincir ağı.....	50
5.2.2.2 Özel Blokzincir ağı.....	50
5.2.2.3 Hibrit Blokzincir ağı	51
5.2.2.4 Konsorsiyum Blokzincir ağı	52
5.2.3 Akıllı kontrat.....	53
5.2.4 Uzlaşma.....	53
5.3 Uygulama	54
5.4 Sonuçlar ve Tartışma	57
6. SİSTEMİN GENEL İŞLEYİŞİ.....	58
6.1 Mimari Topolojisi	58
6.2 İş Akışı.....	59
6.3 Uygulama	60
7. SONUÇ VE ÖNERİLER.....	63
7.1 Öneriler.....	64
KAYNAKLAR	66
ÖZGEÇMİŞ.....	72

ÖZET

Doktora Tezi

YAPAY ZEKA DESTEKLİ BLOKZİNCİR TABANLI FİNANSAL BİLGİ PAYLAŞIM PLATFORMU

Bulut KARADAĞ

İstanbul Ticaret Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. Abdül Halim ZAİM

Eş Danışman: Doç. Dr. Akhan AKBULUT
2024, 73 sayfa

Günümüzde bankacılığın en temel konularından biri kredilerin tahsilatıdır. Kredilerin gecikmeye düşmeden tahsil olmasını sağlamanın ilk şartı ise, doğru müşteriye doğru kredinin kullandırılmasıdır. Kredi değerlendirme aşamasında, tahsis mercilerine ışık tutacak en kritik veriler istihbarat sorgulamaları ile elde edilmektedir. İstihbarat verileri genellikle merkezi kurumlardan elde edilmektedir. Bu verilerin merkezi tek bir kuruluştaki bulunması manipülasyona açıktır. Bu problemi odağına alan bu çalışma üç aşamadan oluşmaktadır. Birincisi, sentetik olarak üretilen finansal ve demografik veriler makine öğrenmesi yoluyla analiz edilerek kredi tahsisi gerçekleşip gerçekleşmeyeceğinin belirlenmesidir. Bu kısımda makine öğrenmesi algoritmalarından LightGBM ile %93.82 başarı oranı sağlandı. İkinci olarak müşterinin sosyal medya verileri üretken yapay zeka yardımıyla analiz edilerek kredi tahsisi için kullanıldı. Üçüncü olarak ise kredi tahsisi gerçekleştirildikten sonra krediye ait limit, risk ve teminat gibi finansal istihbarat verileri detaylı olarak blokzincir ağında saklanmasıdır. Blokzincir ağındaki bu veriler ileride bankanın finansal risk değerlendirmesinde kullanılabilir. İstihbarat birimleri bu bilgilere ulaşmak istediğinde ise hızlı bir şekilde kimlik numarası üzerinden sorgulayarak kredi tahsis verilerine ulaşır. Blokzincir altyapısı üzerine tasarlanmış bir çalışma olması sebebiyle güvenilir, şifrelenmiş bir veriye çok hızlı bir şekilde ulaşılmış olur. Finans kuruluşları veya bankalar arasında istihbarat verilerinin paylaşılması için kurulacak altyapıda, her bir düğümü bir finans kuruluşu olarak düşünebiliriz. Bir finans kuruluşu tarafından yapılacak finans veri girişi, ağa bağlı tüm bankalarda eş zamanlı olarak görülebilecektir. Daha sonra ağa dahil edilecek yeni bir finans kuruluşu olursa, o kuruluş da geriye dönük tüm blokları kendi tarafına çekebilecektir.

Anahtar Kelimeler: Bankacılık, blokzincir, dağıtık finansal veri paylaşımı, kredi skor tahmini, makine öğrenmesi

ABSTRACT

Ph.D. Thesis

ARTIFICIAL INTELLIGENCE SUPPORTED BLOCKCHAIN BASED FINANCIAL INFORMATION SHARING PLATFORM

Bulut KARADAĞ

**İstanbul Ticaret University
Institute of Graduate School
Department of Computer Engineering**

Supervisor: Prof. Dr. Abdül Halim ZAİM

**Co-Supervisor: Assoc. Prof. Dr. Akhan AKBULUT
2024, 73 pages**

One of the core challenges faced by banks today is ensuring timely loan repayment. A crucial factor in achieving this is matching borrowers with suitable loans. During credit assessment, intelligence gathered through inquiries plays a vital role in informing loan allocation decisions. The fact that these data are stored in a single centralized entity makes them susceptible to manipulation. Focusing on this issue, the current study is structured in three stages. First, synthetic financial and demographic data are analyzed using machine learning to determine whether credit allocation should occur. In this phase, a success rate of 93.82% was achieved using the LightGBM algorithm. Second, the customer's social media data were analyzed with the aid of generative artificial intelligence for use in credit allocation decisions. Third, after the credit allocation, financial intelligence data such as limits, risks, and collateral related to the credit are stored in detail on a blockchain network. These data stored on the blockchain network can be utilized for the bank's future financial risk assessments. When intelligence units wish to access this information, they can quickly retrieve credit allocation data by querying through the identity number. Since this study is designed on a blockchain infrastructure, it allows for fast access to reliable and encrypted data. In the infrastructure that will be established for the sharing of intelligence data among financial institutions or banks, each node can be considered as a financial institution. Any financial data entry made by one financial institution will be simultaneously visible to all banks connected to the network. Additionally, if a new financial institution is added to the network later, it will be able to retrieve all previous blocks onto its own system.

Keywords: Banking, blockchain, credit score prediction, distributed financial data sharing, machine learning.

TEŞEKKÜR

Bu araştırma yolunda bana ilham veren ve karşılaştığım engelleri aşmama bilgi ve tecrübesiyle yardımcı olan değerli hocam Prof. Dr. Abdül Halim ZÂİM'e derin minnettarlığımı sunmak isterim. Değerli Eş Danışmanım Doç. Dr. Akhan AKBULUT'a, araştırmanın yürütülmesinde sağladığı yardımlar için teşekkür ederim.

Araştırmanın yürütülmesi sırasında maddi ve manevi yönden desteklerini esirgemeyen tüm İstanbul Ticaret Üniversitesi personeline teşekkür ederim.

Ailem, özellikle de eşim, bu tezin tamamlanması için bana her zaman inandı ve beni motive etti. Destekleri için onlara minnettarım.

Bulut KARADAĞ
İSTANBUL, 2024

ŞEKİLLER

	Sayfa
Şekil 1.1. Tez metodolojisinin akış şeması.....	7
Şekil 3.1. Yapay Zeka çalışmasının adımları.....	22
Şekil 3.2. Eksik alanların grafiksel gösterimi.....	28
Şekil 3.3. Eksik ve tekil alanların kaldırılması sonrası veri setinin gösterimi	29
Şekil 3.4. Risk Grubu alanının toplam sayılarının grafiksel gösterimi	29
Şekil 3.5. Maaş alanının grafiksel gösterimi	30
Şekil 3.6. Maaş alanının Risk Grubu ile ilişkili grafiksel gösterimi	30
Şekil 3.7. Maaş alanının Risk Grubu ile ilişkisinin tek grafiksel gösterimi	31
Şekil 3.8. Maaş alanının normal ve logaritmik gösterimi	32
Şekil 3.9. Yas alanının normal ve logaritmik gösterimi	33
Şekil 3.10. LightGBM modeline ait karışıklık matrisi	35
Şekil 3.11. Kredi skor çalışmasının önyüzü	36
Şekil 4.1. Sosyal medya analiz modülünün mimari topolojisi	39
Şekil 4.2. Mesleklerin risk grubuna göre oranları.....	40
Şekil 4.3. Sosyal medya analiz uygulamasının çıktısı	41
Şekil 5.1. Blokzincir tabanlı finansal veri girişi ve sorgulama topolojisi	45
Şekil 5.2. Blokzincir ağında akıllı sözleşmenin yaygınlaştırılması	46
Şekil 5.3. Blokzincir türleri	52
Şekil 5.4. Blokzincir tabanlı finansal veri paylaşım önyüzü.....	56
Şekil 6.1. Sistemin genel topolojisi.....	59
Şekil 6.2. Sistemin genel iş akışı.....	60
Şekil 6.3. Kredi skor tahmini	61
Şekil 6.4. Sosyal medya analizi	62
Şekil 6.5. Blokzincir tabanlı finansal veri paylaşımı.....	62

ÇİZELGELER

	Sayfa
Çizelge 3.1. Veri seti alanları	23
Çizelge 3.2. Veri setinde boş satırları bulunan alanlar	27
Çizelge 3.3. Veri setinden çıkarılan aykırı değerler.....	31
Çizelge 3.4. Model sonuçları	33
Çizelge 3.5. Skor sonuçları	36
Çizelge 4.1. Üretken yapay zeka modeline verilen metin	41
Çizelge 5.1. Blokzincir platformlarının kalite özellikleri.....	46
Çizelge 5.2. Ethereum, Corda ve Hyperledger Fabric değerlendirme sonuçları	47
Çizelge 5.3. Nem, Stellar ve Ripple değerlendirme sonuçları.....	48
Çizelge 5.4. Cardano, Eos ve Tron değerlendirme sonuçları	49
Çizelge 7.1. Geleneksel ve blokzincir tabanlı finansal veri paylaşım sistemlerinin değerlendirilmesi	64

SİMGELER VE KISALTMALAR

ABI	Application Binary Interface
ACSS	An Automatic Credit Scoring Strategy
ANN	Artificial Neural Networks
AUC	Area under the ROC Curve
BACS	Blockchain and AutoML-based technology for efficient credit scoring classification
BP	Back Propagation
CDS	Credit Default Swaps
CNN	Convolutional Neural Network
DBN	Deep Belief Networks
DL	Deep Learning
ELM	Extreme Learning Machine
ETH	Ether
EVM	Ethereum Virtual Machine
FICO	Fair, Isaac and Company
KKB	Kredi Kayıt Bürosu
KNN	K-Nearest Neighbours
LSTM	Long short-term memory
MEA	Mean Absolute Error
MLP	Multi-layer Perceptron
NFT	Non-fungible token
NLP	Natural Language Processing
OICSM	Online Integrated Credit Scoring Model
PNN	Probabilistic Neural Network
PoS	Proof of State
PoW	Proof of Work
P2P	Peer to Peer
RBFN	Radial Basis Neural Network
RBM	Restricted Boltzmann Machines
RMSE	Root-Mean-Squared Error
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
SELM	Spiking Extreme Learning Machine
SELM(u)	U2 ile Spike Generating Function
SELM ϵ	exp(u) ile Spike Generating Function
SVM	Support Vector Machine
SMS	Kısa mesaj servisi
TCKN	Türkiye Cumhuriyeti Kimlik Numarası
V-QELM	Voting-based Q-generalized ELM
WEKA	Waikato Environment for Knowledge Analysis

1. GİRİŞ

Finansal kuruluşların temel faaliyetlerinden biri olan kredi tahsilatı, modern bankacılık sisteminin omurgasını oluşturmaktadır. Doğru müşteriye doğru kredinin tahsis edilmesi, gecikme ve ödenmeme risklerini minimize ederek bu sürecin etkinliğini belirleyen anahtar faktördür. Bu bağlamda, kredi değerlendirme aşamasında, tahsis mercilerinin karar verme süreçlerine rehberlik edecek kritik verilerin elde edilmesi hayati önem taşımaktadır. İstihbarat sorgulamaları, bu verilerin kaynağını oluşturmakta ve genellikle merkezi kurumlardan sağlanmaktadır. Fakat merkezi kredi kayıt kurumlarının sunduğu veriler her zaman ihtiyaca tam olarak cevap verememekte ve işlem den belli bir süre sonra bankalar ile paylaşılmaktadır. Ek olarak, tüm verilerin tek bir kuruluştaki tutuluyor olması, veri manipölasyonuna ve suistimale açık bir ortam yaratmaktadır.

Müşteriler, finansal ihtiyaçlarını karşılamak için bankalara kredi başvurusunda bulunurlar. Bu süreçte bankalar, kredi riskini değerlendirmek için borçlulara ait çeşitli bilgilere ihtiyaç duyarlar. Bu bilgiler, borçluluk geçmişi, gelir durumu, kredi kartı kullanımı gibi unsurları kapsar. Bu bilgiler, asimetrik bilgi sorunu olarak bilinen bir durumu ortaya çıkarır. Borçlular, kendi mali durumları hakkında bankalara nazaran daha fazla bilgiye sahipken, bankalar bu bilgilere erişimde zorlanırlar.

Bu asimetrik bilgi sorununu çözmek için KKB ve Risk Merkezleri gibi kuruluşlar kurulmuştur. Bu kuruluşlar, bankalardan topladıkları kredi ve borçluluk bilgilerini birleştirerek, her bir borçlu için kapsamlı bir risk profili oluştururlar. Bu profiller, bankaların kredi riskini daha etkin bir şekilde değerlendirmelerine ve kredi verme kararlarını daha bilinçli bir şekilde vermelerine olanak tanır. KKB'den elde edilen veriler sadece risk değerlendirmesi için değil, aynı zamanda her bir borçluya özgü bir kredi skoru oluşturmak için de kullanılır. Findeks olarak da adlandırılan bu skor, 0 ile 1900 arasında değişen bir puanlama sistemidir (Kredi Kayıt Bürosu 2024). Yüksek Findeks skoru, borçlunun kredi riskini düşük tuttuğunu ve krediyi geri ödeme olasılığının yüksek olduğunu gösterir. Bu sayede

bankalar, kredi faiz oranlarını ve kredi limitlerini belirlerken Findeks skorunu önemli bir parametre olarak kullanırlar.

Türkiye'de Findeks skoru tek bir kuruluş tarafından üretilirken, Amerika Birleşik Devletleri'nde FICO, Vantage, CE, Educational credit score gibi farklı türde kredi skorları üretilmektedir. Bu durum, kredi piyasasının işleyişinde ve kredi riskini değerlendirme yöntemlerinde farklılıklara yol açmaktadır.

Blokszincir, merkezi olmayan bir ağda birçok cihaz veya “düğüm” tarafından paylaşılan dağıtılmış bir defter olduğundan, paylaşılan bir kayıt tutma türüdür. Bu düğümlerin her biri, tüm kayıt blokszincirinin bir kopyasını tutar (Lemieux 2016). Ademi merkeziyetçilik, şeffaflık, sağlamlık, denetlenebilirlik ve güvenlik gibi ayırt edici özellikleri nedeniyle, blokszincirinin birçok geleneksel iş yapış şeklini etkilediği görülmektedir (Hoonsopon 2019). Blokszincir yeni bir konu olmamasına rağmen Bitcoin ile bilinirliği artmıştır. Bitcoin'in teknik dokümanı, Satoshi Nakamoto adlı bir kişi veya grup tarafından 2008 yılında yazılan “Bitcoin: A Peer-to-Peer Electronic Cash System” başlıklı makaledir (Nakamoto 2008). Makale, bir finansal aracı kullanmadan blokszincir teknolojisi kullanılarak dijital varlıkların transferini öne sürmektedir. Elektronik para transferlerinin en ciddi sorunlarından biri olan çifte harcama sorununu bu transferleri gerçekleştirirken nasıl çözdüğü de yazıda detaylı olarak anlatılıyor. Bu transfer sisteminin altyapısı, dağıtılmış bir defter yapısı sağlayan blokszincir teknolojisi üzerine kuruludur. Her ağ katılımcısı, blokszincirinde bir düğüm olarak kabul edilir ve düğümler tarafından yapılan işlemler, bloklar halinde birbirine bağlanır. Satoshi Nakamoto, Bitcoin teknik incelemesini 2008'de ilk kez yayınladığından beri, üç nesil blokszinciri kabul edildi. Blokszincir 1.0 kripto para işlemleri için, Blokszincir 2.0 finansal uygulamalar için ve Blokszincir 3.0 finans dışındaki hükümet, sağlık ve bilim gibi alanlardaki uygulamalar için kullanılıyor (Perera vd., 2020).

Blokszincir benzeri bir sistem ilk olarak kriptograf David Chaum tarafından 1982 tarihli “Karşılıklı Şüpheli Gruplar Tarafından Kurulan, Korunan ve Güvenilen Bilgisayar Sistemleri” tezinde önerildi (Sherman vd., 2019). Blokszincir, Bitcoin'den önce önerilmiş olsada merkezi olmayan blokszincir kavramı ilk olarak

Bitcoin ile kavramsallaştırıldı. 2013 yılında Ethereum'un piyasaya sürülmesiyle Vitalik Buterin, yalnızca bir kripto para birimi değil, aynı zamanda geliştiricilerin kendi akıllı sözleşmelerini programlayabilecekleri açık kaynaklı, merkezi olmayan bir blokzincir ağının altyapısını da önerdi. Ethereum blokzincir ağı artık programlanabilir ve uygulanabilen bir akıllı sözleşme yapısına sahiptir.

Bu araştırma üç fazdan oluşmaktadır. İlk aşama olan kredi değerlendirmesinde, makine öğrenmesi algoritmaları kullanılarak banka müşterilerinin finansal ve demografik verileri, kredi riskini tahmin etmek için kapsamlı bir şekilde analiz edilir. Bunun sonucunda, müşteriye ait kredi riskleri daha net bir biçimde tespit edilir, müşteriye bir skor verilir, kredi tahsis süreci doğru ve hızlı bir şekilde gerçekleşir. İkinci fazda müşterinin sosyal medya üzerinden analizi yapılır. Bu yapılırken, müşterinin beyan ettiği meslek ve işyeri bilgileri referans alınarak analiz gerçekleştirilir. Üçüncü fazda ise kredi onaylandıktan sonra kredi tahsisi yapılan müşteriye ait limit, risk ve teminat gibi istihbarata konu olan bilgileri detaylı olarak blokzinciri altyapısı yoluyla eş zamanlı olarak paydaş olan tüm kuruluşlarla paylaşılır. İstihbarat birimleri bu bilgilere ulaşmak istediğinde ise hızlı bir şekilde kimlik numarası üzerinden sorgulayarak finans verilerine ulaşır. Blokzincir altyapısı üzerine tasarlanmış bir çalışma olması sebebiyle güvenilir, şifrelenmiş bir veriye çok hızlı bir şekilde ulaşılmış olur.

Finans kuruluşları arasında istihbarat verilerinin paylaşılması için kurulacak altyapıda, her bir düğüm bir finans kuruluşu olarak düşünülmektedir. Bir finans kuruluşu tarafından yapılacak finans veri girişi, ağa bağlı tüm diğer kuruluşlarda eş zamanlı olarak görülebilmektedir. Daha sonra ağa dahil edilecek yeni bir finans kuruluşu olursa, o kuruluş da geriye dönük tüm blokları kendi tarafına çekebilmektedir. İlerleyen zamanda ağdan bir finans kuruluşu ayrılrsa bile mevcut işleyişte bir sorun olmayacaktır.

1.1 Problemin Tanımı

Günümüzde bankalar için en önemli konuların başında doğru kredi tahsisi ve sonrasındaki risk durumunun tespiti gelir. Günümüzde kredi tahsis çalışmaları bankalar tarafından çoğunlukla karar ağaçları ile belli parametre setleri ile yapılmaktadır. Tahsis sonrası müşteri risk değerlendirme süreci ise KKB ve Risk Merkezi gibi kurumların bankalardan toplayıp tekrar bankalara verdiği müşteri finansal verileri ile gerçekleşmektedir. Bankacılıkta kredi tahsisi yapılırken genellikle kural tabanlı basit yöntemler uygulanır. Son yıllarda makine öğrenmesi algoritmaları da denenerek yeni modeller ortaya koyulmaktadır. Ancak bu çalışmalar için kullanılan veri setleri Alman Kredi Risk (Hofmann 1994), American Express (Express 2022) ve Kredi sınıflandırma (Kaggle 2023) veri setlerinden oluşmaktadır. Türkiye'deki banka müşterilerinin davranışlarına özgü sentetik olarak hazırlanan veri seti üzerinden böyle bir çalışma yapılmamıştır. Ayrıca kredi tahsis sonrası KKB ve Risk Merkezi gibi merkezi sistemler tarafından sağlanan veriler, şeffaflık ve değiştirilemez ilkelerine tam olarak uymamaktadır. Bu araştırma, yukarıda belirtilen kaygıların çözümü için yapay zeka destekli ve blokzincir tabanlı bir finansal bilgi paylaşım platformu sunmaktadır.

1.2 Araştırma Önemi

Bu çalışmada, finansal kuruluşlar için kredi riskini daha etkin bir şekilde yönetmenin ve tahsilat süreçlerini optimize etmenin önemi vurgulanmaktadır. Geleneksel kredi değerlendirme sistemlerinde, merkezi kurumlardan gelen verilerin yetersizliği, veri manipülasyonu riski ve asimetrik bilgi sorunu gibi zorluklar mevcuttur. Blokzincir teknolojisi, merkezi olmayan, şeffaf ve güvenli bir ortam sunarak bu sorunların üstesinden gelmeyi amaçlamaktadır. Önerilen sistemde, kredi riskini daha doğru bir şekilde tahmin etmek için makine öğrenmesi algoritmaları denenerek daha hızlı ve daha adil kredi tahsisi olanağı sunulur. Nihayetinde doğru kişiye doğru kredi sağlanmış olur. Bu durumda bankalar doğru müşteriye kredi imkanı sunarak kendilerini de güvence altına almış olurlar. Ayrıca, kredi tahsisi sonrası müşteri kredisine ait (limit, risk, teminat) verilerin blokzincir tabanlı bir sistem üzerinden tüm paydaşlarla gerçek

zamanlı olarak paylaşılması sağlanarak, veri güvenliği ve şeffaflık artırılır ve bu sayede daha güvenilir ve etkin bir finansal sistemin yolu açılır.

1.3 Araştırma Amacı

Bu araştırmanın amacı, makine öğrenmesi algoritmalarını kullanarak kredi riskini daha doğru bir şekilde tahmin etmek ve kredi tahsisi süreçlerini otomatikleştirmektir. Bunun yanında, finansal kuruluşlar arasındaki kredi tahsisi sonrası risk değerlendirme amacıyla ihtiyaç duyulan istihbarat verilerinin paylaşımını iyileştirmek için blokzincir teknolojisinin nasıl kullanılabileceğini araştırmaktadır. Geleneksel kredi değerlendirme ve istihbarat paylaşım sistemlerinin sunduğu zorlukları ele alarak, blokzincirin merkezi olmayan yapısının, şeffaflığının ve güvenilirliğinin bu süreçleri daha verimli ve güvenli hale getirme potansiyeli gösterilmektedir. Son olarak bu sistem, tüm paydaşlarla gerçek zamanlı olarak kredi bilgileri paylaşımı sağlayarak istihbarat verisi güvenliğini ve erişilebilirliğini artırmayı amaçlamaktadır.

Bu tezin hipotezleri ise şu şekilde sunulmaktadır;

H0: Makine öğrenmesi algoritmalarının kullanılması, geleneksel kredi risk değerlendirme yöntemlerine kıyasla, kredi riskinin daha doğru bir şekilde tahmin edilmesini sağlar.

H1: Blokzincir teknolojisinin merkezi olmayan yapısı, şeffaflığı ve güvenilirliği, finansal kuruluşlar arasındaki kredi tahsisi sonrası risk değerlendirmesi için ihtiyaç duyulan istihbarat verilerinin paylaşımını iyileştirir ve bu verilerin güvenliğini artırır.

1.4 Araştırma Kapsamı

Araştırmanın kapsamı, finansal kuruluşlar arasında kredi tahsilatı ve risk değerlendirmesi süreçlerini iyileştirmek için makine öğrenmesi algoritmaları kullanılması ve istihbarat sürecinde blokzincir teknolojisinin potansiyel

kullanımını ele almaktır. Özellikle, makine öğrenmesi algoritmalarının, kredi başvurularını değerlendirmek ve kredi skorlarını belirlemek için finansal ve sosyal medya verilerini analiz etmek amacıyla kullanıldığı vurgulanmaktadır. Ayrıca, araştırma kredi riskinin daha doğru bir şekilde tahmin edilmesi ve müşteri kredi bilgileri gibi istihbarat verilerinin paylaşımı için blokzincir tabanlı bir sistem önermektedir. Bu sistem, blokzincir aracılığıyla istihbarat verilerini paylaşmanın güvenliği, şeffaflığı ve verimliliği gibi avantajları da ele almaktadır.

1.5 Hedef Kitle

Bu araştırmanın hedef kitlesi öncelikli olarak finansal kuruluşlardır. Özellikle, konvansiyonel bankalar, dijital bankalar, faktöring şirketleri ve diğer kredi veren kurumlar, araştırmanın bulgularından doğrudan yararlanabilir. Araştırma, kredi değerlendirme süreçlerini iyileştirmek ve istihbarat paylaşımını kolaylaştırmak için blokzinciri teknolojisinin nasıl kullanılabileceğine dair bilgiler sunmaktadır. Bu nedenle, finansal kuruluşlardaki risk yöneticileri, kredi analistleri, veri bilimcileri ve blokzincir teknolojisi uzmanları araştırmanın hedef kitlesi içerisinde yer almaktadır. Bunun yanı sıra, düzenleyici kurumlar ve politika yapıcılar da araştırmanın bulgularını, finansal sistemin güvenliğini ve istikrarını artırmak için kullanabilirler.

1.6 Araştırma Metodolojisi

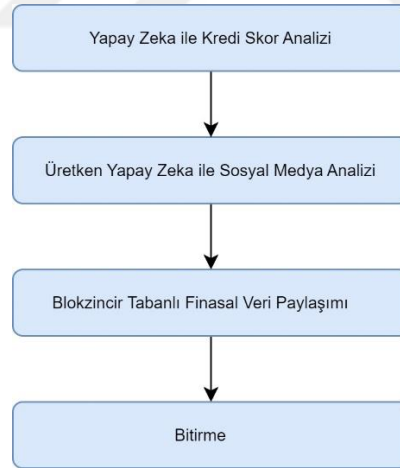
Bu araştırma temel olarak dört aşamada ele alınmaktadır:

- Sentetik veri seti kullanılarak ve makine öğrenmesi algoritmaları denenerek kredi skor çalışması yapılmaktadır. Bu kısımda on makine öğrenmesi algoritması farklı parametreler ile denenerek sonuçlar karşılaştırılmaktadır.
- Kredi skor tespiti yapılan bir müşterinin sosyal medya analizi ikinci aşamada ele alınmaktadır. Sosyal medya uygulaması LinkedIn üzerinden otomatik olarak çekilen veriler üretken yapay zeka modeli Gemini A.I. ile yorumlanarak karar vericiye destek olması açısından bir çıktı oluşturmaktadır.
- Kredi tahsisi gerçekleştirildikten sonra müşterinin kredi bilgilerinin bir blokzincir ağına kaydedilmesi ve istenildiği takdirde okunması araştırılmaktadır. Kredi

sağlanan müşterinin limit, risk ve teminat gibi finansal bilgileri güvenli, merkeziyetsiz ve değiştirilemez bir şekilde Ethereum ağında saklanır. Blokzincir ağında olan başka bir finansal kuruluş tarafından istenildiği zaman ulaşılır ve bilgiler elde edilir.

- Son aşamada ise yukarıdaki tüm araştırma sonuçlarının bir araya getirilmesi ve sistemin genel işleyişinin ortaya koyulmasıyla birlikte tez yazımı gerçekleştirilmektedir. Şekil 1.1, tez metodolojisinin akış şemasını göstermektedir.

Şekilde görüldüğü üzere öncelikle yapay zeka ile kredi skor analizi yapılır. Bunu yaparken makine öğrenmesine ait 10 farklı algoritma denenmektedir. Sonrasında üretken yapay zekadan faydalananarak müşterinin sosyal medya analizi gerçekleştirilir. Bu kısımda Gemini üretken yapay zeka uygulamasından faydalanılmaktadır. Bir sonraki aşamada ise kredi tahsisi yapılan müşteriye ait finans bilgilerinin blokzincir tabanlı olarak paylaşılması gerçekleştirilmektedir. Son olarak tüm bu üç kısım bir bütün halinde bitirme tezi başlığı altında sunulmaktadır.



Şekil 1.1 Tez metodolojisinin akış şeması

1.7 Tez Organizasyonu

Bu tez beş bölümden oluşmaktadır. Bölüm 2, yapay zeka destekli kredi risk ve skor çalışmaları ile blokzincir tabanlı finansal veri paylaşımları konusundaki literatür çalışmalarına yer vermektedir. Bölüm 3, sentetik veri seti kullanılarak

makine öğrenmesi algoritmaları ile kredi skor tespitini sunmaktadır. Kredi skoruna göre bir kişiye kredi verilip verilmeyeceği tespit edilebilmektedir. Bölüm 4, bir banka müşterisinin sosyal medyadan edinilen bilgiler ile kredi uygunluğunu tespit eden sosyal medya analizine yer vermektedir. Bölüm 5, kredi tahsisi gerçekleştikten sonra müşteri finansal verilerinin blokzincir tabanlı bir sistem üzerinde işlenmesini sunmaktadır. Bölüm 6, 3. 4. ve 5. bölümlerde yer alan kısımları bir araya getirerek sistemin tüm işleyişini mimari topoloji ve iş akışı çerçevesinde ele alarak sunmaktadır.



2. LİTERATÜR ÖZETİ

Bu bölümde hem blokzincir tabanlı finans çalışmaları hem de blokzincirden bağımsız olarak genel kredi süreçlerinden başlayarak yapay zeka destekli kredi skor ve otomatik karar yöntemleri üzerinde durulmaktadır.

2.1 Blokzincir Tabanlı Finans Çalışmaları

Dumitru ve arkadaşlarının yaptığı çalışmada, kredi notu hesaplanırken verilerle ilgili sorunlardan bahsedilmektedir. Temel sorunlardan biri, kişiler ve üçüncü kişiler arasındaki güven eksikliği nedeniyle sınırlı veri paylaşımıdır. Bu durum, kredi puanlamasında yetersiz veriye ve doğru hesaplanamayan sonuçlara neden olur. Bu sorunu çözmek için çalışmalarında “Güvenilir Veri Pazarı”nı tanıtmaktadırlar. Blokzincir ile entegre olabilecek bu sistem kredi puanlamasına katkı sağlamaktadır (Roman ve Stefano 2016).

Karayılan çalışmasında, blokzinciri altyapısını incelemiş ve Müşterini Tanı çözümleri için blokzinciri altyapısı kullanılması ve blokzincirinin finansal uygulamalarda kullanımı üzerinde çalışılmaktadır. Araştırma kapsamında bilgi paylaşım ağının oluşturulması için iki model ortaya konulmakta ve modeller karşılaştırılmaktadır. Birinci modelde akıllı kontratlar geliştirilerek veri paylaşımı ve depolaması blokzinciri üzerinde sağlanmıştır. İkinci modelde veriler harici bir veritabanında tutularak blokzincirde anahtarlanmıştır (Karayılan 2019).

Finansal alanda yapılan başka bir çalışmada blokzinciri, gerçek zaman muhasebesi ve kredi risk modellemesi üzerindedir. Byström çalışmasında, blokzincirinin kredi risk modelleme şeklini nasıl etkileyeceğini ve blokzincirinde gerçek zamanlı muhasebe ile güven ve zamanın nasıl iyileştirilebileceği konusunda çalışmıştır. Çalışmada blokzincirdeki kayıtların asla değiştirilemiyor olması özelliği vurgulanmıştır. Blokzinciri kullanarak bir firmanın muhasebe kayıtlarını depolamak için güvenilir ve sürekli güncellenen bir yapı olan blokzincir uygulanmıştır. Mali veriler düzenli aralıklarla hazırlanıp firmanın defterine eklenir. Bir denetçi tabloların doğruluğu hakkında bir görüş bildirir.

Yatırımcılar ve kredi risk yöneticileri bu bilgileri kullanırken denetçinin doğru bilgi verdiği ve firmanın mali verilerini kayıt defterine doğru kaydettiğine güvenmek zorundadır. Bu süreçte güven kavramı mali tabloların hazırlanmasından denetçinin onayına kadar son derece önemlidir. Bu güvenin sağlanması için Byström blokzincirinin bir çözüm olabilmesi için çalışma yapmıştır. Çalışmasında firmanın gönüllü olarak mali verilerini blokzincire yazması verilere değiştirilmezlik ve zaman damgası vuracaktır. Bu sayede oluşturulan mali veriler defterinin tamamı görünür ve bu verilerin tutarlılığını kanıtlayacaktır (Byström 2019).

S. Chakraborty ve arkadaşları tarafından kredi risk yönetimi için bir blokzinciri tabanlı kredi analizi altyapısı önerilmektedir. Blokzincir teknolojisi kullanılarak, kredi verenlerin ve borçluların güvenli ve şeffaf bir şekilde işlem yapması sağlanarak finansal sistemlerin verimliliği amaçlanmaktadır. Çalışmada, kredi verenlerin ve borçluların kredi skorlarını belirlemek için makine öğrenmesi algoritmaları kullanılmıştır. Önerilen altyapı, geleneksel yöntemlere göre daha doğru ve hızlı sonuçlar verdiği gözlemlenmiştir (Chakraborty 2019).

Shivani Bharatbhai Patel ve arkadaşları tarafından KiRTi olarak isimlendirilen, finansal kurumlar için blokzincir tabanlı bir kredi öneri sistemi sunulmaktadır. Çalışmanın amacı, kredi skoru üretimi için üçüncü taraf kredi derecelendirme kuruluşlarına ihtiyaç duymadan, potansiyel borçlular ve potansiyel kreditorler arasında akıllı kredi işlemlerini kolaylaştırmaktır. KiRTi sistemi, potansiyel borçluların zaman serisi sıralı verilerini açık blokzincirinde saklamakta ve LSTM modeli ile kredi skoru üretmektedir. Kredi skorunun güncellenmesi için kenar ağırlıkları boolean göstergelere göre ayarlanmaktadır. Akıllı sözleşmeler ise potansiyel borçlular ve potansiyel kreditorler arasında otomatik geri ödeme sağlamaktadır. Sistem, Alman kredi veri setinde %97,5 doğruluk ve 0.98304 F-ölçütü elde etmiştir. Sistem ayrıca hesaplama maliyetinin 20.96 ms ve iletişim maliyetinin 121 bayt olduğunu kanıtlamıştır (Patel vd., 2021).

Bu çalışmada, blokzincir ve otomatik makine öğrenimi temelli bir kredi skora sınıflandırma yöntemi olan BACS incelenmiştir. Kredi skora sınıflandırması, modern

ekonomide önemli bir rol oynamaktadır ancak kredi skorlama modeli oluşturma süreci zahmetli ve zaman alıcıdır. Bu nedenle, bu süreci otomatikleştirmek için blokzincir ve otomatik makine öğrenimi teknolojileri kullanılmıştır. Blokzincir, kredi verilerinin güvenli, izlenebilir ve değiştirilemez bir şekilde saklanmasını sağlarken, otomatik makine öğrenimi veri örnekleme, özellik çıkarımı, özellik seçimi ve hiperparametre optimizasyonu gibi veri madenciliği adımlarını etkin bir şekilde bütünleştirmektedir. BACS yöntemi, Alman Kredi Verisi adlı bir veri seti üzerinde test edilmiş ve %75,7 doğruluk oranı elde etmiştir. Ayrıca, BACS yöntemi diğer sınıflandırma modellerinden daha yüksek performans göstermiştir. Bu sonuçlar, BACS yönteminin kredi skorlama sınıflandırması için verimli ve doğru bir yöntem olduğunu göstermektedir (Yang, Qiao ve Qi vd., 2022).

2.2 Yapay Zeka Destekli Kredi Skor ve Otomatik Karar Çalışmaları

Finansal kuruluşlar, kredi başvurusunda bulunan kişi ve kredi geçmişi hakkında toplam hesap sayısı ve türleri, hesaplarının yaşı, işlem sıklığı, borçlarını zamanında ödeyip ödemediği gibi bilgileri toplar ve kredi raporları oluşturur. Bu ayrıntılara dayanarak, yeni müşteriler için karar verilmektedir. Kredi ile ilgili risk değerlendirilirken, farklı yönlerin dikkate alınması gerekmektedir (Tsai ve Wu. 2008).

Erkut Eliuz, Tüketici Kredilerinde Risk Yönetimi ve Bir Skorkart Modeli Önerisi adlı yüksek lisans tezinde, banka müşterilerinin temerrüt olasılıklarını, geçmişte tüketici kredisi talep etmiş başvuru formlarındaki bilgilerden faydalanarak lojistik regresyona dayalı basit bir skorkart modeli ile geliştirdi. Model meslek, müşteri süresi, yaş, ikamet durumu, medeni durum, cinsiyet, çocuk sayısı, otomatik talimat geçen gün sayısı, eğitim durumu, gelir-kredi tutarı oranı, istenilen taksit sayısı ve temerrüt durumu parametrelerinden oluşmaktadır. Veri seti toplamı 22182 olup, 745'i temerrüde düşmüş ve 21437'si temerrüde düşmeyen müşteriden elde edildi. Veri seti rastgele %70 ve %30 olmak üzere iki kısma ayrılmıştır. Modelin kalan %30'u, %70'i "Stepwise logistic regression"- "forward LR" yöntemi kullanılarak oluşturulduktan sonra test edildi. Modelin

veri seti ve kontrol grubu ROC deęerleri sırasıyla 0.725 ve 0.740 çıktı. Her iki grup için Gini katsayıları hesaplandığında, sonuçlar $|2*0.725|-1=0.45$ ve $|2*0.740|-1=0.48$ 'dir. Veri seti ile kontrol grubu veri setinin Gini katsayılarının birbirine yakın olması bu modelin tüm örneklere uygulanabilir olduğunu göstermektedir (Eliuz 2009).

Son yıllarda geleneksel hizmetler, borç verenlerin ve borç alanların finansal kurumları dahil etmeden aralarında etkileşimde bulunduğu kitlesel fonlama platformları olarak adlandırılan çevrimiçi yerler tarafından kesintiye uğradı. Bununla birlikte, kredi verenlerin deneyim eksikliği ve borçlunun kredi geçmişi hakkında eksik veya belirsiz bilgiler, doğru bir kredi riski puanlaması gerektiren kitle fonlama platformlarında riskleri artırabilir. Bu platformlar, sırasıyla lojistik regresyon kullanarak uygunluklarını deęerlendirmek ve Random Forest sınıflandırma yoluyla borçlu durumunu tahmin etmek için Riza Emekter ve arkadaşları tarafından önerilen yöntemleri kullanabilmektedir (Emekter vd., 2015).

Yoichi Hayashi'nin üzerinde çalıştığı kural bazlı risk deęerlendirme modelinde, kredi riskini Pareto optimal perspektifinden deęerlendirmek için Re-RX algoritmasının üç versiyonu ile bir dizi kural oluşturuldu. Ancak, kural oluşturma sürecindeki farklı sorunlar nedeniyle bu yaklaşımın büyük miktarda veriye uygulama zorlukları gösterilmektedir (Hayashi, Application of a rule extraction algorithm family based on the Re-RX algorithm to financial credit risk assessment from a Pareto optimal perspective 2016).

Cuicui Luo ve arkadaşları, CDS veri setlerine uygulanan kredi skortlama modellerinin performanslarını araştırmaktadırlar. Çalışmada, derin öğrenme algoritması DBN ile RBM kullanılmıştır. Bu algoritma, veri setlerindeki öznitelikleri otomatik olarak öğrenmek ve sınıflandırmak için kullanılmıştır. Çalışma, DBN'nin performansını lojistik regresyon, MLP ve SVM gibi popüler kredi skortlama modelleri ile karşılaştırmıştır. DBN ile RBM, CDS veri setlerindeki 11 özniteliği otomatik olarak öğrenmek ve 2 sınıfa ayırmak için kullanılmıştır. Veri setleri, standartlaştırma ve normalleştirme gibi ön işlemlerden geçirilmiştir.

Modellerin performansları, 10 katlı çapraz doğrulama ile ölçülmüştür. Sonuçlar, DBN'nin doğruluk ve AUC açısından diğer yöntemlerden daha iyi performans gösterdiğini sunmuştur. Özellikle DBN'nin doğruluğu %87,5 ve AUC'si %0,91 olarak bulunmuştur (Luo, Wu ve Wu. 2017).

Jay Lohokare ve arkadaşları tarafından finansal verilere ek olarak sosyal medya verileri de dahil edilerek yapılan çalışmada, banka müşterilerine ait kredi skoru hesaplama yoluna gidildi. Standart finansal verilere ek olarak müşteriye ait Twitter duygu durumu, takipçi sayısı ve arkadaş sayısı da dahil edildi. Finansal veriler akıllı telefonlarına her işlem sonrası gelen SMS üzerinden elde edildi. SMS verisi, işlemin tutarı, kart tipi, toplam cari tutar, banka ve işlem tarihi gibi verilerden oluşmaktadır. Bu finansal veriler ve sosyal medyadan elde edilen veriler kullanılarak yapay sinir ağları ile bir skor elde edilmektedir (J. Lohokare 2017).

Tüketici kredi skorlaması, hızla büyüyen tüketici finansmanı sektöründe kredi riski yönetiminin önemli bir parçasıdır ve çeşitli veri madenciliği teknikleri bunun üzerinde önerilmiş ve kullanılmıştır. Son zamanlarda, derin öğrenme teknikleri önemli bir popülerlik kazanmış ve görüntü tanıma, bilgisayar görüşü gibi birçok alanda mükemmel performans göstermiştir. Bu çalışmada, derin öğrenmenin tüketici kredi skorlaması üzerinde kullanılması amaçlanmıştır. Evrişimli sinir ağı ile Relief adlı öznitelik seçimi algoritmasını birleştiren hibrit bir model önerilmektedir. Deneyler, Çinli bir tüketici finansmanı şirketinden gerçek bir veri kümesi üzerinde gerçekleştirilmiş ve sonuçlar, önerilen modelin lojistik regresyon ve Random Forest gibi diğer karşılaştırma modellerine göre üstün performans elde ettiğini göstermiştir (Zhu vd., 2018).

Melih Emre Turan tarafından hazırlanan Doğrusal ve lojistik regresyon yöntemleri ile kredi skorlama sistemleri ve bir uygulama adlı yüksek lisans tezinde, mevcut skorlama sistemleri incelenip günümüzde en sık kullanılan doğrusal ve lojistik regresyon yöntemleri ile iki farklı sonuç elde edilmiştir. Çalışmada, meslek, riskli durum kaydı, borç gelir oranı, ilk kredi tahsisi üzerinden geçen süre, son dönemde yapılan başvuru sayısı gecikmede bulunan kredi adedi

ve kredinin teminata oranı parametreleri kullanıldı. Nihayi olarak bir skor kart elde edildi. Veri seti toplamda 5960 kayıttan oluşmaktadır. Önerilen iki farklı doğrusal ve lojistik modeller için elde edilen ROC değerleri sırası ile %90,14 ve %90,18' dir. ROC değerinin bu kadar yüksek olması, modellerin iyi ve kötü müşteriyi ayırt etmede son derece etkili olduğunu göstermektedir. Sonuçların birbirine yakın olması sebebiyle İki yöntemin de kullanılabileceği belirtilmektedir. Ancak kredilerin değerlendirilmesinin yanı sıra kredi kötülüğü olasılığının tahmin edilmesi de isteniyorsa lojistik regresyonun daha uygun bir yöntem olduğu görüşü sunulmaktadır (Turan 2019).

Bu çalışmada sadece müşterilere ait finansal verilere değil aynı zamanda sosyal medya verilerini de hesaplama dahil ederek model geliştirildi. Makine öğrenmesi algoritmalarından NaiveBayes, RandomTree, RandomForest ve MultilayerPerceptron açık kaynak kodlu WEKA platformu üzerinde çalıştırıldı. Algoritmalar içinde en yüksek doğruluk oranını %84.63 ile NaiveBayes verirken en düşük oran RandomTree algoritmasından %75.23 oranı ile elde edildi. Çalışmayı farklı kılan kısım sosyal medya verilerinin de hesaplama dahil edilmesidir. Hesaplama dahil edilen finans kuruluşu müşterisinin LinkedIn, Facebook, Twitter ve Quora gibi sitelerdeki bilgileri, yorumu ve fotoğrafları kullanılarak güvenilirlik yüzdesi elde edilmektedir. Bu yüzde, müşteriye ait geçmiş kredi bilgileri ile değerlendirilmektedir (Kulkarni ve Dhage. 2019).

Kredi skortlama modelleri, geleneksel finans kurumlarında yıllardır yaygın olarak kullanılmaktadır. P2P kredi vermenin kredi işlemi çok sık olur ve yeni veriler veri dağılımının değişmesine neden olur. Veri güncellemesini dikkate almayan bir kredi skortlama modeli, sonraki kredi değerlendirmesinde ciddi bir sapmaya veya hatta başarısızlığa neden olur. Bu çalışmada Zaimei Zhang ve arkadaşları tarafından, P2P kredi verme için çevrimiçi bütünleşik kredi skortlama modeli (OICSM) önerilmiştir. OICSM, kredi skortlama modelinin iki tür özneliği daha etkili bir şekilde ele alabilmesi ve çevrimiçi olarak güncellenebilmesi için gradyan arttırma karar ağacını ve sinir ağını bütünleştirir. Gerçek ve temsili kredi veri kümelerine dayalı olarak çevrimdışı ve çevrimiçi deneyler yapılmıştır ve önerilen modelin etkinliğini ve üstünlüğünü doğrulamak için kullanılmıştır. Deneysel

sonular, OICSM'nin iki znitelik zerinde derin ğrenmedeki avantajı nedeniyle performansını nemli lde arttırabildiğini ve evrimii dinamik gncelleme yeteneđi sayesinde model bozulmasını dzelterbildiğini gstermiřtir (Zhang, Niu ve Liu. 2020).

Kredi risk skoru zerine yapılan bir bařka alıřmada, makine ğrenme algortimalarından K-Nearest Neighbor, Decision Tree, Random Forest, Naive Bayes ve Logistic Regresyon karřılařtırıldı. Ticari bir bankanın kredi bilgileri kullanarak oluřturulan modellerde en yksek dođruluk % 96.53 ile Random Forest algoritması olurken en dřk dođruluk %79.99 ile Naive Bayes algoritmasıdır (Wang vd., 2020).

Paul Diaconescu ve Victor-Emil Neagoe tarafından nerilen derin ğrenme modelinde, kredi skoru iin en uygun hiperparametreler bulunmaya alıřıldı. Derin ğrenme ařamasında Scikit-optimize, bayesian optimizasyon ve gradient descent optimizasyon algortimaları kullanıldı. alıřmada, Alman kredi veri setine ait 20 parametrelilik 1000 kayıt %70 eđitim %30 test olacak řekilde sınıflandırıldı. alıřma sonucunda, finansal kuruluřların kredi deđerlendirmesi yaparken elde edilen kt kredi durumunu %3 oranında azalttıđı grld. Ortaya ıkan model, finansal kuruluřların kendi modelleri ile entegre edilebilmektedir (Diaconescu ve Neagoe. 2020).

Xolani Dastile ve arkadaşlarının kredi skorlaması zerine yaptıđı bu alıřmada, istatistiksel ve makine ğrenmesi modellerini sistematik bir literatr taraması ile incelemektedirler. alıřmanın amacı, literatrdeki sınırlılıkları belirlemek, bir makine ğrenmesi erevesi nermek ve ortaya ıkan ynleri iřaret etmektir. alıřma, kredi skorlamasında kullanılan modelleri  ana gruba ayırmaktadır: istatistiksel modeller, makine ğrenmesi modelleri ve derin ğrenme modelleri. alıřmada ayrıca modellerin avantajları, dezavantajları, performansları ve uygulamaları karřılařtırılmaktadır. alıřmada, kredi skorlaması iin 76 istatistiksel model, 63 makine ğrenmesi modeli ve 17 derin ğrenme modeli sistematik olarak incelenmiřtir. Sonular, makine ğrenmesi ve derin ğrenme modellerinin istatistiksel modellere gre daha iyi performans gsterdiđini ortaya

koymuřtur. zellikle XGBoost, LightGBM ve CatBoost gibi kmeleme yntemleri ile ANN, CNN ve LSTM gibi derin ğrenme yntemleri kredi skorlamasında en iyi sonuları vermiřtir (Dastile, Celik ve Potsane. 2020).

Makine ğrenmesi yoluyla yapılan bir bařka kredi skor alıřmasında, Decision Tree, KNN ve SVM modelleri uygunlandı. Yař, meslek, medeni durum, eğitim seviyesi gibi kiřisel zelliklerin yanında yıllık gelir, bor durumu, ev kredisi mi, ihtiya kredisi mi gibi maddi zelliklerle birlikte toplam 16 zellik ile eğitim gerekleřti. Eğitilen sistemin en yksek doğruluk oranına SVM modeli %90.4 ile sahipken en dřk orana %79.4 ile KNN modeli sahiptir. Sistemin test doğruluğunda ise en yksek değere %83.2 ile Decision Tree ulařmaktadır (Singh ve Goel 2020).

Bjrn Rafn Gunnarsson ve arkadařlarının yaptıėı alıřmada, kredi skorlaması iin derin ğrenme algoritmalarının uygunluėu arařtırılmaktadır. Bu amala, iki derin ğrenme mimarisi (ok katmanlı algılayıcı aėı ve derin inan aėı) ile iki geleneksel yntem ve iki kmeleme yntemi karřılařtırılmıřtır. Aynı zamanda, eřitli kredi skorlama veri setleri ve performans ltleri kullanarak modeller deėerlendirilmiřtir. Bunun yanında, kredi skorlaması baėlamında Bayes istatistiksel test prosedrleri tanıtılmıř ve geleneksel olarak kredi skorlamasında en iyi uygulama olarak kabul edilen parametrik olmayan test prosedrleri ile karřılařtırılmıřtır. alıřmada, kredi skorlaması iin altı farklı sınıflandırma algoritması karřılařtırılmıřtır. Bu algoritmaların performansları, 12 farklı veri seti zerinde 10 katlı apraz doğrulama ile llmřtr. Performans ltleri olarak ACC, AUC, F1 Score ve Gini katsayısı kullanılmıřtır. Sonular, Bayes istatistiksel test prosedrleri ile analiz edilmiřtir. Sonulara gre, kmeleme yntemi olan XGBoost, tm veri setleri ve performans ltleri iin en iyi performans gsteren yntemdir. Derin sinir aėları ise daha yzeysel karřılıklarından iyi performans gstermemiř ve daha fazla hesaplama maliyetine neden olmuřtur (Gunnarsson vd., 2021).

alıřma, memetik evrimsel algoritma ve nral mimari arama kullanarak otomatik bir ACSS nermektedir. Kredi skorlama, mikro ve kk iřletmelerin yanı sıra

bireyler için artan sayıda kredi işlemiyle giderek daha kritik bir rol oynamaktadır. Birçok araştırma, kredi skorlama modellerini uzman yöntemlerine dayandırmaktadır. Ancak, farklı kredi verileri modeller için farklı gereksinimlere sahiptir. Kredi skorlama araştırması için mevcut zorluklara yanıt olarak, Fan Yang ve arkadaşları tarafından ACSS adında bir otomatik kredi skorlama stratejisi önerilmiştir. Çalışmada, veri içe aktarma, veri madenciliği seçimi, özellik mühendisliği, parametre optimizasyonu ve diğer bölümleri içeren bir kredi değerlendirme platformu tasarlanmıştır. Kamuya açık ve kendilerine ait kredi veri setlerini kullanarak deneyler yapıldı ve sonuçları en son kredi değerlendirme yöntemleriyle karşılaştırıldı. Yapılan geniş çaplı deneyler, ACSS yönteminin Almanya kredi veri seti, Tayvan kredi veri seti ve kendilerine ait veri seti için kredi skorlamada göreceli olarak dikkat çekici performans iyileştirmeleri sağladığı görülmüştür. Önerilen yöntem tarafından elde edilen en iyi sonuçlar sırasıyla MAE için 0.98, RMSE için 0.99, doğruluk için 0.895 ve kesinlik için 0.901'dir. Deneysel bulgular, önerilen ACSS'nin otomasyon ve doğruluk arasında bir denge sağladığını ve tüketim endüstrisine uygulanarak kredi değerlendirmelerinin güvenilirliğini artırabileceğini göstermektedir (Yang, Qiao ve Huang vd., 2021).

Durgesh Kumar Singh ve Noopur Goel'in önerdiği modelde, veri seti üzerinde en iyi özellik seçimi uygulanarak müşterilerin kredi puanlarının tahmin edilmesi için en uygun olan random forest algoritmaları ile lojistik regresyon arasında bir karşılaştırma yapıldı. Önerilen modelin amacı, bankanın iflasını, takipteki varlıklarını ve banka kayıplarını azaltmaktır. Veri seti 45211 kayıttan oluşmaktadır ve yaş, meslek, evlilik durumu, yıllık hesap tutarı, borç gibi 17 parametreden oluşmaktadır. Uygulanan random forest ve lojistik regresyon algoritmalarına ait doğruluk oranı sırasıyla 0.988 ve 0.825 olarak hesaplandı (Goel 2021).

Ashwani Kumar ve arkadaşlarının derin öğrenme ve K-Means algoritmalarını kullanarak önerdikleri kredi skor tahmini modeli, geçerli ve gerçek dünya senaryosuna uygun farklı kredi skor veri setleri ile de doğrulandı. Önerdikleri modeli geliştirirken müşterinin kredi tarihçesi, yıllık geliri, talep edilen kredi türü

gibi verilerin yanında kredi veren alıřanın tecrbesi de veri setine dahil edildi. nerilen model sonu olarak sadece temerrt olasılıđını deđil aynı zamanda bor veren kuruma sađlayacađı karı da tahmin etmektedir. nerilen derin đrenme modelinin dođruluđu %87 olarak ifade edilmektedir (Kumar, Shanthi ve Bhattacharya. 2021).

Yoichi Hayashi tarafından hazırlanan bu alıřmada kredi skorlaması iin derin đrenme tekniklerinin ortaya ıkan eđilimleri ve potansiyeli incelenmektedir. Kredi skorlaması, bankacılık sektrnde risk ynetiminin temel bir unsuru olarak yıllardır kullanılmaktadır. Kredi skorlaması problemlerinde, bađımlı deđiřken ikili olup bařarısız kredilere '0', bařarılı kredilere '1' atanır. Bu nedenle, borlunun temerrt olasılıđını tahmin etmek iin lojistik regresyon ve sinir ađları gibi teknikler kullanılabilir. Son yıllarda, derin đrenme algoritmalarının daha yksek dođruluk sađlaması nedeniyle makine đrenmesi algoritmalarının kısmen DL algoritmalarıyla deđiřtirildiđi grlmektedir. Bu sistematik inceleme alıřması, geliřmiř derin đrenme tekniklerinin ortaya ıkan eđilimleri ve potansiyeli hakkında derin bir bakıř aısı sađlamayı amalamaktadır. Bu alıřma ayrıca, DBN daha sıđ ađlara gre daha yksek dođruluk sađlama nedenlerini aıklamakta, DL tabanlı sınıflandırıcıların potansiyel sınıflandırma yeteneklerini tartıřmakta ve DL ile aıklanabilir kredi skorlaması arasında bir kpr kurmaktadır. DBN'lerin teorik zellikleri de sunulmakta ve daha sıđ ađlara gre daha yksek dođruluk sađlama nedenleri anlatılmaktadır. Bu alıřmada, 2019 ve 2022 yılları arasında yayınlanan alıřmalar analiz edilerek, kredi skorlama topluluđuında yaygın olarak kullanılan Avustralya, Almanya (kategorik), Almanya (sayısal), Japonya ve Tayvan veri kmeleri iin dođruluk ve alıcı iřletim karakteristik eđrisi altındaki alan aısından deđerlendirilen ve karřılařtırılan en son DL teknikleri gzden geirilmiş ve karřılařtırılmıştır. Bu derleme makalesi ayrıca, tablo verilerinin iki boyutlu CNN uygulamak iin nasıl grntlere dnřtrldđn ve yerel ve kresel kural ıkarma ve kural tabanlı yntemler kullanarak "kara kutu" modellerin kredi skorlamasında nasıl uygulandıđını aıklamaktadır. Son olarak, kredi skorlama veri kmeleri iin DL tabanlı sınıflandırıcıların tasarımına iliřkin yeni bir bakıř aısı sunulmakta ve gelecekteki

arařtırma ynleri zerine bir tartıřma yapılmaktadır (Hayashi, Emerging Trends in Deep Learning for Credit Scoring: A Review 2022).

Venkatareshbabu Kuppli ve arkadaşlarının yapmıř olduėu bu alıřmada, kredi skor sınıflandırması iin SELM olarak adlandırılan yeni bir makine ėrenmesi algoritması sunuldu. nerilen model, biyolojik gereki sivri ulardan esinlenerek tasarlandı. Ayrıca nerilen SELM algoritmasının performansı BP, PNN, V-QELM, RBFN, SELM(u), SELM€ ve ELM gibi eřitli diėer standart algoritmalar ile Avustralya veri seti, Alman veri seti (kategorisel), Alman veri seti (rakamsal), Japon kredi ekran veri seti ve İflas veri seti kullanılarak karřılařtırıldı. nerilen SELM'in veri setlerinde sınıflandırma doėruluėu, AUC ve hesaplama sresi aısından BP, PNN, V-QELM, RBFN, SELM(u), SELM€ ve ELM'den daha iyi performans gsterdiėi gzlemlendi (Kuppli, Tripathi ve Edla. 2022).

3. YAPAY ZEKA İLE KREDİ SKORU TESPİTİ

Finansal piyasaların dinamik yapısı, müşterilerin kredi riskini doğru bir şekilde değerlendirmeyi her zamankinden daha kritik hale getirmektedir. Bu bağlamda, makine öğrenimi teknikleri, kredi riskini ölçmek ve optimal karar vermeyi teşvik etmek için güçlü bir araç olarak öne çıkmaktadır (Sabato 2010). Bu teknikler, finansal kuruluşların kredi değerlendirme süreçlerini optimize etmelerine ve genel risk yönetimi uygulamalarını geliştirmelerine imkan tanımaktadır. Geleneksel kredi puanlama modelleri, temerrüt riskini tahmin etmek için doğrusal ve yorumlanabilir bir yaklaşım sunan lojistik regresyon modellerine dayanmaktadır (Sabato, 2010). Fakat, daha karmaşık makine öğrenimi algoritmalarının ortaya çıkmasıyla birlikte finans sektörü, geleneksel yöntemlerden daha yüksek performans gösteren karmaşık ve doğrusal olmayan modellere geçiş eğilimindedir.

Makine öğrenimi tabanlı kredi puanlama modellerinin en önemli avantajlarından biri, tüketici kredi raporları gibi, bireyin mali geçmişi ve davranışları hakkında kapsamlı bilgiler içeren büyük veri kümelerini işleme yeteneğidir. Bu modeller, geleneksel istatistiksel yöntemlerle kolayca tespit edilemeyen karmaşık kalıpları ve ilişkileri verilerden ortaya çıkarabilir (Stefania ve Vamossy 2019).

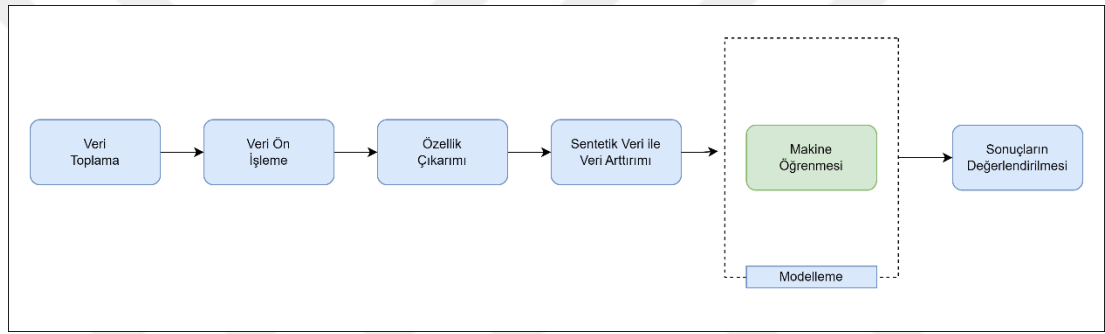
Sürekli değişen bankacılık dünyasında, kredi riskinin etkin bir şekilde değerlendirilmesi, finansal istikrarın sağlanmasında ve yatırımların getirisinin maksimize edilmesinde kritik bir unsur haline gelmiştir. Bu bağlamda, makine öğrenmesi algoritmaları, kredi skortlama uygulamalarını köklü bir şekilde dönüştürme potansiyeline sahip güçlü bir araç olarak ortaya çıkmıştır. Bankacılıkta geleneksel kredi skortlama modelleri büyük ölçüde regresyon analizi, karar ağaçları ve istatistiksel yöntemlere dayanmaktadır. Ancak, bu modeller genellikle modern tüketici davranışlarının ve finansal verilerin karmaşıklıklarını ve nüanslarını yakalamakta zorlanmaktadır. Öte yandan, makine öğrenmesi algoritmaları, finansal kurumların elindeki geniş veri yığınlarını kullanarak, geleneksel yaklaşımlar tarafından gözden kaçırılacak desenler ve içgörülerden faydalanabilir. Bu çalışmada, makine öğrenmesi yoluyla

banka müşterileri için kredi skor çalışması yapılmaktadır. Bu skor müşterilere kredi verilir verilmemesine karar veren makine öğrenmesi ile geliştirilmiş bir yöntemdir. Bunun için öncelikle 16151 satır sentetik veri edinildi. Bu veriler uzman görüş yoluyla elde edilen verilerdir. Bu verilerin bazı kolonları modele girdi olabilecek düzeyde olmadığı için veri ön işleme işlemi gerçekleşti. Ön işlemeye tabi olan veriler noksan, sıfır değeri taşıyan ve normalizasyonun dışında olan değerledi. Bir sonraki aşamada mevcut parametrelere ek olarak One Hot Encoding yönetimi uygulanarak yeni özellikler eklendi. Veriler SMOTE (Kevin W. Bowyer, 2011) isimli sentetik veri kütüphanesi kullanılarak 28,095 satıra ulaştı. Modelleme aşamasında hem makine öğrenmesine ait on farklı algoritma denendi. Bu algoritmaların içinde en yüksek başarı oranına ulaşan LightGBM algoritması görselleştirilmek üzere .pkl uzantılı model haline getirildi. Bu model web üzerinden gösterilecek şekilde Python dilinde Flask kütüphanesi kullanarak hazırlandı.

3.1 İş Akışı

Bu çalışmanın ilk adımı veri toplanmasıdır. Sürecin ilk ve en önemli adımıdır. Bu aşamada, 16151 satır veri uzman görüşü ile sentetik yolla toplandı. Bu veriler ikinci aşamada analiz için uygun hale getirilmeleri gerekmektedir. Bu işlem, veri ön işleme olarak adlandırılır ve veri temizleme, düzeltme ve standardizasyonu içerir. Veri temizlemede, eksik, tutarsız veya gürültülü veriler tespit edildi ve düzeltildi. Değer taşımayan alanlar kaldırıldı. Veri standardizasyonunda ise, farklı formatlarda ve birimlerde olan veriler tek bir formata ve birime dönüştürüldü. Örneğin maaş bilgileri için nokta ve virgül kullanılan sayısal değerler tek bir formata çevrildi. Veriler ön işlendikten sonra, analiz için önemli olan anlamlı özellikler ve temsil edici nitelikler elde edildi. Bu işlem, özellik çıkarımı olarak adlandırılır ve istatistiksel yöntemler, veri madenciliği teknikleri ve alan bilgisine dayalı yaklaşımlar kullanılarak gerçekleştirildi. Doğru özelliklerin seçilmesi, modelin performansını ve doğruluğunu önemli ölçüde arttırmaktadır. Mevcut verilerin yetersiz veya dengesiz olduğu durumlarda, modelin genelleme yeteneğini geliştirmek için sentetik veriler oluşturulabilir. Bu işlem, veri arttırma olarak adlandırılır. Sentetik veri oluşturmada, SMOTE (Kevin

W. Bowyer, 2011) isimli sentetik veri kütüphanesi kullanılarak 28,095 satır elde edildi. Uygun veri ön işleme, özellik çıkarımı ve sentetik veri oluşturma adımlarından sonra makine öğrenmesi algoritmaları ile farklı modeller oluşturuldu. Bu aşamada, modelin hiperparametreleri optimize edildi ve modelin öğrenmesi sağlandı. Algoritma seçimi ve hiperparametre optimizasyonu, modelin performansını önemli ölçüde etkilemektedir. Eğitilmiş modeller, test verileri üzerinde değerlendirildi ve performansı çeşitli metrikler kullanılarak analiz edildi. Bu metrikler, Accuracy, Precision, Recall ve F1 skoru değerleri içermektedir. Elde edilen sonuçlar yorumlandı ve birbirleriyle karşılaştırıldı. Şekil 3.1, tüm bu sürecin iş akış adımlarını göstermektedir.



Şekil 3.1 Yapay Zeka çalışmasının adımları

3.2 Veri Seti

Bu çalışma için yapılan araştırmalar sonrasında farklı veri setlerine ulaşıldı. Alman Kredi Risk (Hofmann, 1994), American Express (Express, 2022) ve Kredi sınıflandırma (Kaggle, 2023) veri setleri incelendi. Bu veri setleri ile daha önce çalışmalar yapılmış ve yapay zeka modelleri ortaya konulmuştu. Bu veri setlerinin haricinde sentetik yolla oluşturulan veri seti de incelemeye dahil edildi. Tüm makine öğrenmesi çalışmaları bu sentetik veri seti üzerinde yapılmaktadır.

Bir finans kuruluşundan elde edilen veriler anonim hali ile 16151 satır kayıttan oluşmaktadır (Karadag, <https://github.com/> 2024). Bu veri seti içerisinde maaş, yaş ve ikametgah yılı gibi sayısal verilerin yanında medeni durum, eğitim durumu, cinsiyet, ünvanı ve mesleği gibi kategorisel veriler de bulunmaktadır. Bunların haricinde müşterinin geçmiş dönemlere ait

“Son_3_Ay_Icinde_Otomatik_Ret_Basvurusu_VarMi” gibi finansal alanlar da bulunmaktadır. Çizelge 3.1’de veri setinin ham halinde bulunan alanlar görülmektedir. Veriler Python programlama dile ile işlendi. Bunun için visual studio code geliştirme ortamı üzerinde Python 3.9.13 sürümü ile yazılan kodlar çalıştırıldı.

Çizelge 3.1 Veri seti alanları

Maas_Belgesiz	Dogdugu_Ulke
Maas_Belgeli	Yas
Diger_Belgesiz	Askerlik_Durumu
Diger_Belgeli	Talebinin_Altinda_Teklif_VarMi
Es_Belgeli	Son_3_Ay_Icinde_Otomatik_Ret_Basvurusu_VarMi
Es_Belgesiz	Son_3_Ayda_Olumsuz_Istihbarattan_Dolayi_Ret_Edilmis_Teklifi_VarMi
Medeni_Durum	Son_1_Yil_Icinde_GM_Tarafından_Ret_Edilmis_Tahsis_VarMi
Cinsiyet	KKB_Sorgusunda_Kanuni_Takibi_Sorunu_Var_Mi
Egitim_Durumu	Yazilmis_Cek_Kontrolu_UygunMu
Unvani	Müsteri_Terkin_Olmus_Avukat_Icra_Takibine_Dusmus_Mu
Meslegi	Ikametgah_Yili
Sektor	Oturdugu_Ev_Durumu
	Risk_Grubu

3.3 Makine Öğrenmesi Algoritmaları

Kredi skortlama tespitinde on farklı makine öğrenmesi algoritması denendi. Farklı makine öğrenmesi algoritmalarının tercih edilme temel nedeni, her bir algoritmanın kredi riski değerlendirme sürecine katkı sağlayabilecek farklı özelliklere sahip olmasıdır. Kredi skortlama hem doğruluğu yüksek hem de yorumlanabilir modellerin kullanılmasını gerektiren karmaşık bir problemdir. Bu nedenle, birden fazla makine öğrenmesi algoritmasının uygulanması model performansının karşılaştırılmasını ve en uygun yöntemin seçilmesini sağlar.

LightGBM, XGBoost ve CatBoost, gradient boosting tabanlı algoritmalar güçlü modelleme kapasitelerine sahip olmaları ve yüksek doğruluklarıyla bilinmektedir. Bu algoritmalar, özellikle büyük veri setlerinde ve karmaşık

ilişkilerin olduğu durumlarda üstün performans göstermektedir (Chen ve Guestrin. 2016). Özellikle LightGBM ve CatBoost, bellek kullanımı ve işlem süresi açısından optimize edilmiştir. Bu da büyük ölçekli kredi skorlama çalışmalarında önemli bir avantaj sağlamaktadır (Prokhorenkova vd., 2018).

LightGBM, kayıp fonksiyonunu en aza indirmek için zayıf öğrenenlerin sırayla eklendiği gradyan artırıcı karar ağaçları yöntemini kullanır. Her yeni model, önceki modellerin yaptığı hataları düzeltmeye odaklanmaktadır. Denklem 3.1’de gösterildiği üzere,

$$L(y, F(x)) = \sum_{i=1}^n l(y_i, F_m - 1(x_i) + f_m(x_i)) + \Omega(f_m) \quad (3.1)$$

Burada $L(y, F(x))$ genel kayıp fonksiyonunu temsil eder, $l(y_i, F(x_i))$ belirli bir veri noktası için kaybı ifade eder, $F_{m-1}(x)$ önceki iterasyonlardaki modeli, $f_m(x)$ yeni eklenen zayıf öğreniciyi temsil eder, $\Omega(f_m)$ ise model karmaşıklık cezasını içerir.

LightGBM algoritması, konvansiyonel seviye temelli büyüme yaklaşımından ayrılarak yaprak odaklı bir strateji benimsemektedir. Bu metodolojide, optimal kazancı temin eden yaprak segmentasyonu gerçekleştirilerek ağacın asimetrik gelişimine olanak tanınmaktadır. Söz konusu yaklaşım, bilhassa hassasiyet ve etkinlik perspektifinden kayda değer avantajlar sunmaktadır. Bu algoritma, veri noktalarını histogram bazlı bir yaklaşımla decompose etmekte ve bu histogramlar aracılığıyla optimal bölünme noktalarını tespit etmektedir. Bu prosedür, hafıza kullanımını optimize ederken işlem sürecini de hızlandırmaktadır. Bununla birlikte, gradyan magnitudlerine dayalı veri örnekleme tekniği uygulanarak model eğitim süreci akselere edilmektedir. Yüksek gradyan değerlerine sahip örnekler, daha kritik olarak değerlendirilmekte ve eğitim sürecinde ağırlıklı olarak kullanılmaktadır (Ke vd., 2017).

XGBoost (Extreme Gradient Boosting) hız ve performans odaklı, gradyan artırma algoritmasının popüler ve verimli bir uygulamasıdır. Hem sınıflandırma hem de regresyon problemlerinde yüksek doğruluk sağlamak için yaygın olarak

kullanılır. XGBoost, temel olarak bir dizi karar ağacını sırayla eğiterek çalışır ve her ağaç önceki ağaçların hatalarını düzeltmeye odaklanır. Algoritma, her adımda bir önceki adımın tahminlerinin kalitesini değerlendirir ve daha iyi tahminler üretmek için ağırlıkları ayarlar. Algoritmanın temel amacı, zayıf tahminleyicilerin bir araya getirilmesiyle güçlü bir model oluşturmaktır. Bu sürecin temeli denklem 3.2’de ifade edilebilir,

$$f(x) = \sum_{t=1}^T \gamma_t h_t(x) \quad (3.2)$$

Burada $h_t(x)$, t. aşamada elde edilen zayıf öğrenici, γ_t ise bu öğrenicinin ağırlığıdır. Toplamda T zayıf öğrenici kullanılır.

XGBoost, aşırı öğrenmeyi önlemek için düzenleme teknikleri kullanır ve bu da onu karmaşık veri kümeleri için uygun hale getirir. Ayrıca, paralel işlem yetenekleri sayesinde büyük veri kümelerinde hızlı ve verimli bir şekilde çalışabilir. XGBoost'un başarısının arkasındaki temel faktörlerden biri, ikinci dereceden türev bilgilerini kullanarak kayıp fonksiyonunu optimize etmesidir. Bu da daha hızlı yakınsama ve daha iyi doğruluk sağlar (Chen ve Guestrin. 2016).

CatBoost (Category Boosting), Yandex tarafından geliştirilen, gradyan artırma tabanlı bir makine öğrenmesi algoritmasıdır. CatBoost, özellikle kategorik özellikleri işlemek ve aşırı öğrenmeyi (overfitting) önlemek konusunda etkilidir. Algoritma, sıralı artırma (ordered boosting) adı verilen bir teknik kullanarak her adımda bir önceki adımın hatalarını düzelteren yeni bir karar ağacı oluşturur. CatBoost, kategorik özellikleri işlemek için her bir özelliğin değerlerini hedefe göre sıralayan ve bu sıralamalara dayalı olarak sayısal değerler atayan bir yöntem kullanır. Bu yöntem, kategorik özellikleri etkili bir şekilde kullanarak modelin doğruluğunu artırır (Prokhorenkova vd., 2018).

CatBoost’taki amaç fonksiyonu, bir kayıp fonksiyonunun en aza indirilmesine dayanmaktadır. L tahmin edilen değerlere bağlıdır y_i ve gerçek etiketler y . Denklem 3.3’te gösterildiği üzere,

$$L = \sum_{i=1}^N (y_i + y)^2 \quad (3.3)$$

CatBoost ayrıca aşırı öğrenmeyi önlemek için simetrik ağaç yapısı ve rastgele permütasyonlar gibi teknikler kullanır. Simetrik ağaç yapısı tüm ağaçlarda aynı bölme koşullarını kullanarak modelin varyansını azaltır. Rastgele permütasyonlar ise her bir ağaç için eğitim verilerinin sırasını rastgele değiştirerek modelin önyargısını azaltır. CatBoost yüksek doğruluk, hızlı eğitim süresi ve kategorik özellikleri otomatik olarak işleme yeteneği gibi avantajlar sunar (Dorogush 2018).

Random Forest ve Decision Tree algoritmaları ise veri setinin özelliklerini anlamada ve modelin yorumlanabilirliğini artırmada etkili araçlardır. Decision Tree, karar kurallarını açık bir şekilde göstererek, hangi faktörlerin kredi riskini daha fazla etkilediğini belirlemeye olanak tanır. Random Forest ise birden fazla karar ağacını birleştirerek, modelin genelleme yeteneğini artırır ve aşırı öğrenme (overfitting) sorununu azaltır (Breiman 2001).

Lojistik Regresyon, kredi skortlama çalışmalarında sıkça kullanılan bir yöntemdir ve kredi riski ile bağımlı değişken arasındaki ilişkiyi anlamada temel bir rol oynar. Bu algoritma, modelin yorumlanabilirliğini korurken, doğrusal ilişkilerin varlığı durumunda etkili sonuçlar üretir (Hosmer Jr 2013). Nearest Neighbors ve Gaussian Naive Bayes algoritmaları, basit fakat güçlü yaklaşımlar sunarak, özellikle veri setinin yapısına göre iyi sonuçlar verebilir. Nearest Neighbors algoritması, örneklerin komşularına olan benzerliklerini temel alarak kredi skorunu tahmin ederken, Gaussian Naive Bayes, verinin dağılımını temel alarak hızlı ve etkili bir sınıflandırma sağlar (Murphy 2012).

Son olarak, Stacking Classification ve Soft Voting gibi topluluk yöntemleri (ensemble methods), birden fazla modelin tahminlerini birleştirerek, her bir modelin güçlü yönlerinden faydalanmayı ve genel model performansını artırmayı amaçlar. Bu yöntemler, kredi skortlama gibi yüksek hassasiyet gerektiren uygulamalarda, çeşitli algoritmaların kombinasyonu ile elde edilen tahminlerin doğruluğunu artırabilir (Wolpert 1992).

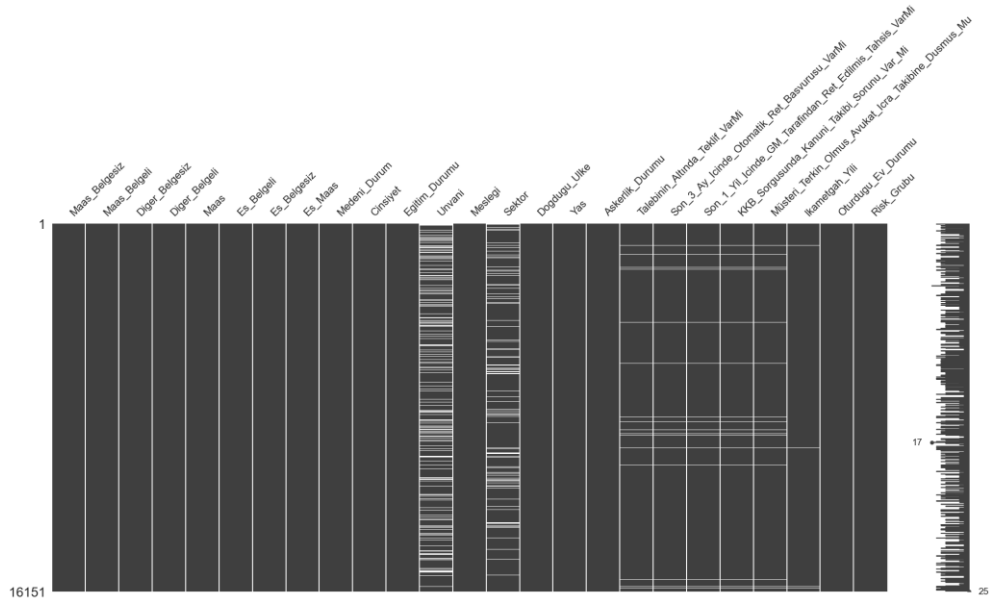
3.4 Uygulama

3.4.1 Veri ön işleme

Çalışmanın en uzun süren kısmı ön işleme alanıydı. Verilerin ağırlıklı olarak kategorisel veri olması ve birçok boş satır bulunması çalışmanın ön işleme aşamasına bariz etkisi oldu. Başlangıçta verilerin 9 tanesi float, 1 tanesi int ve 15 tanesi object veri tipindeydi. Öncelikle object veri türünde olan Medeni_Durum, Cinsiyet, Egitim_Durumu, Unvani, Meslegi, Sektor, Dogdugu_Ulke ve Askerlik_Durumu alanları kategori türüne çevrildi. Sonrasında veri setinde boş olan kısımlar tespit edildi. Çizelge 3.2’de boş satırları bulunan alanlar ve satır sayıları görülmektedir.

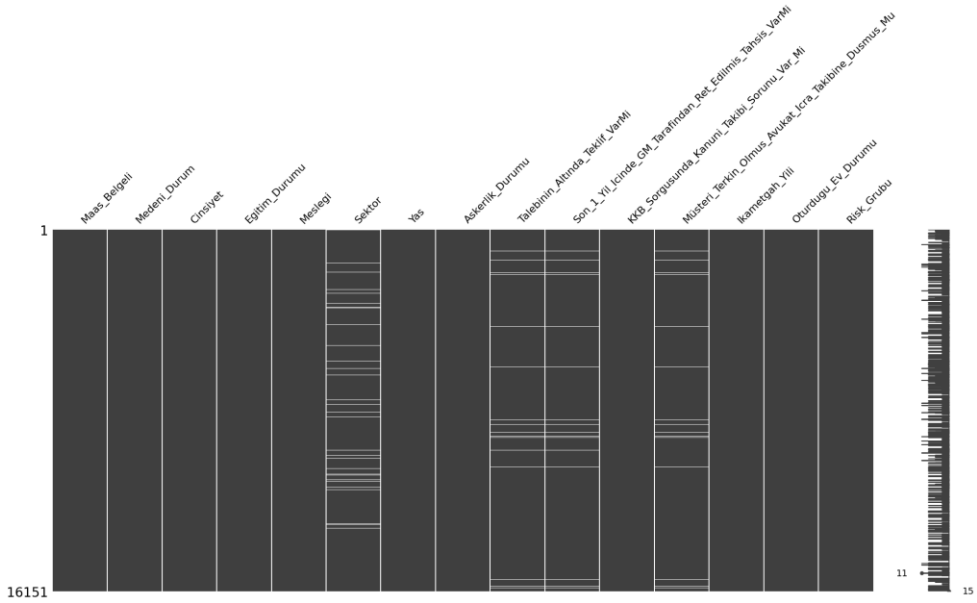
Çizelge 3.2 Veri setinde boş satırları bulunan alanlar

Alan	Boş Satır
Unvani	2525
Sektor	2082
Müsteri_Terkin_Olmus_Avukat_Icra_Takibine_Dusmus_Mu	446
KKB_Sorgusunda_Kanuni_Takibi_Sorunu_Var_Mi	446
Son_1_Yil_Icinde_GM_Tarafından_Ret_Edilmiş_Tahsis_VarMi	446
Son_3_Ay_Icinde_Otomatik_Ret_Basvurusu_VarMi	446
Talebinin_Altında_Teklif_VarMi	446
Ikametgah_Yili	98



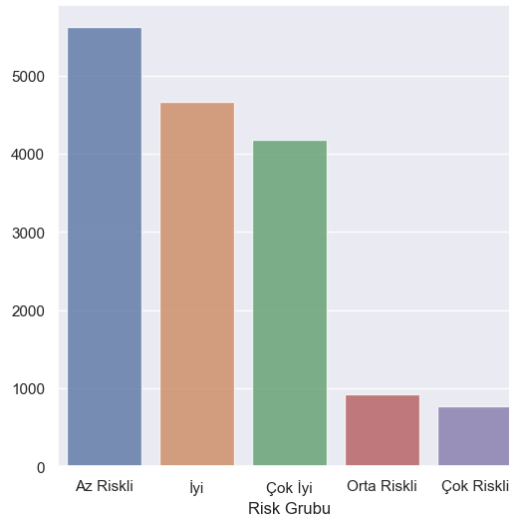
Şekil 3.2 Eksik alanların grafiksel gösterimi

Şekil 3.2’te görüldüğü üzere Unvani ve Sektor alanları en fazla boş satır bulunan alanlardır. Bunların haricinde Unvani alanı içinde kategorisel olarak 7413 satır “Ünvani yok” şeklinde etiketlenmişti. Böyle olduğunda yaklaşık 9938 satır boş olmaktaydı. “Maas_Belgesiz”, “Diger_Belgesiz”, “Diger_Belgeli”, “Unvani”, “Son_3_Ayda_Olumsuz_Istihbarattan_Dolayi_Ret_Edilmis_Teklifi_VarMi”, “Es_Belgesiz”, “Es_Belgeli”, “Dogdugu_Ulke”, “Yazilmis_Cek_Kontrolu_UygunMu” ‘Son_3_Ay_Icinde_Otomatik_Ret_Basvurusu_VarMi’ alanları tekil tip veri veya boş olmaları sebebiyle kaldırıldı. Kaldırılan veriler sonrasında çalışmaya devam edilecek alanlar Şekil 3.3’te gösterilmektedir.



Şekil 3.3 Eksik ve tekil alanların kaldırılması sonrası veri setinin gösterimi

Bir kişiye kredi verilme durumunun iyi veya kötü sonuca sahip olup olmayacağına 5 kategoriye ayırdığımız risk grubu alanından anlamaktayız. Şekil 3.4'te tüm risk grubu kategorileri veri setindeki sayıları ile verilmektedir.

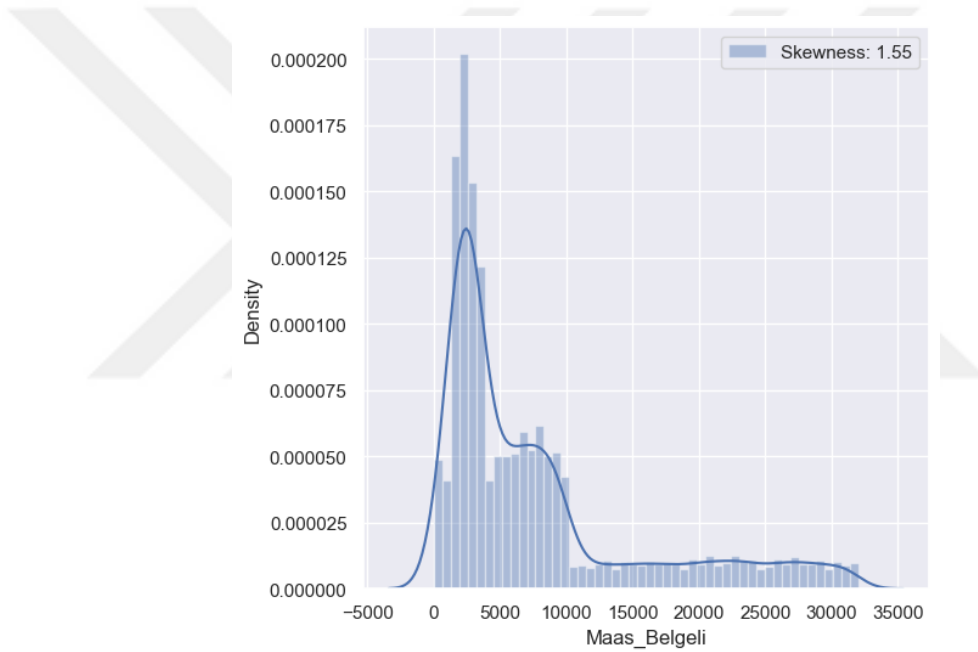


Şekil 3.4 Risk Grubu alanının toplam sayılarının grafiksel gösterimi

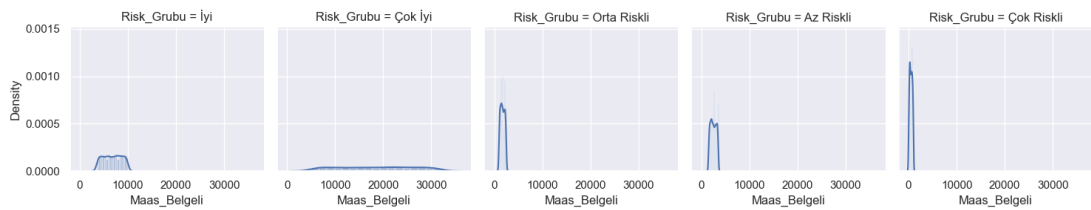
Risk grubuna göre bir kişi “Çok İyi ise” en iyi kredi kartı segmentine sahip bir kart verilebilir, ev ve araç kredisi kullanılır. Risk durumu “İyi” ise çoğu kredi kartı ve kredi verilebilir, ancak en iyi fırsatlar verilmeyebilir. “Az Riskli” olması durumunda kredi oranlarında bir ayrıcalık yapılmaz, belirlenecek tahsis limiti yüksek olmaz. “Orta Riskli” olması durumunda ise kredi oranı çok yüksek tutulur

ve çok düşük limitli kredi kartı verilebilir. Son olarak “Çok Riskli” durumunda çoğu kredi ve kredi kartı başvurusu reddedilir.

Maaş verileri ile ilgili Maas_Belgesiz, Maas_Belgeli, Diger_Belgesiz, Diger_Belgeli, Es_Belgeli ve Es_Belgesiz alanları bulunmaktadır. Bu alanların büyük çoğunluğu sıfır olarak yer almaktadır. Bunların haricinde alan boş görülmesin diye bir olarak girilen kayıtlar da bulunmaktadır. Aslında bunlar da sıfır olması gereken kayıtlardır. Tüm bu alanlar model için bir anlam taşımayacaklarından kaldırıldı. Maaş alanının şekil 3.5’te grafiksel olarak gösteriminde çarpıklık (skewness) değeri görülmektedir.

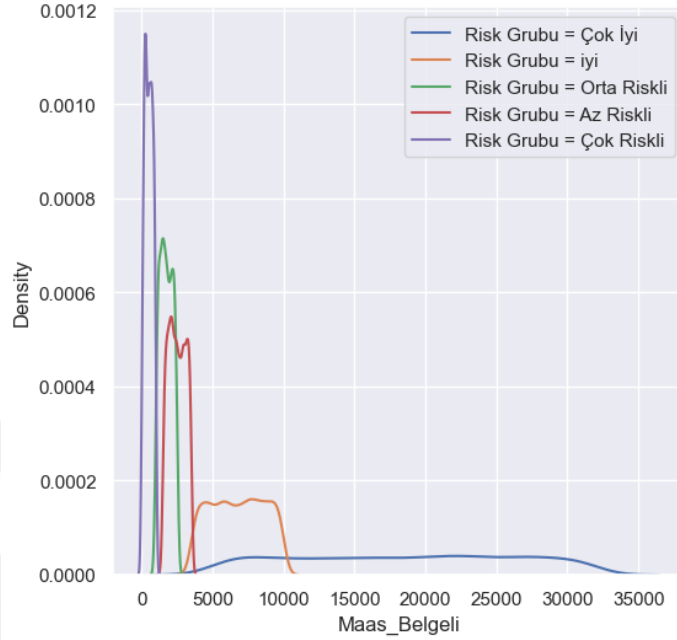


Şekil 3.5 Maaş alanının grafiksel gösterimi



Şekil 3.6 Maas alanının Risk Grubu ile ilişkili grafiksel gösterimi

Maaş alanına ait değerlerin Risk Grubu'na göre dağılımı şekil 3.6'da görülmektedir. Bu beş risk grubunun Maas alanına göre tek grafikte olduğu durumdaki gösterimi şekil 3.7'de verilmektedir.



Şekil 3.7 Maaş alanının Risk Grubu ile ilişkisinin tek grafiksel gösterimi

Sayısal değer olan Maaş, Yaş ve İkametgah_Yili alanlarında yer alan aykırı değerler(outliers) çıkarıldı. Bu şekilde bulunan 8 aykırı değer sonrası 16143 kayıt kalmaktadır. Bu üç alana ait çıkarılan aykırı değerler çizelge 3.3'te gösterilmektedir.

Çizelge 3.3 Veri setinden çıkarılan aykırı değerler

Kolon	['Maas_Belgeli', 'Yas', 'Ikametgah_Yili']
Değer	[873, 924, 1689, 2621, 3146, 4796, 6336, 7217]

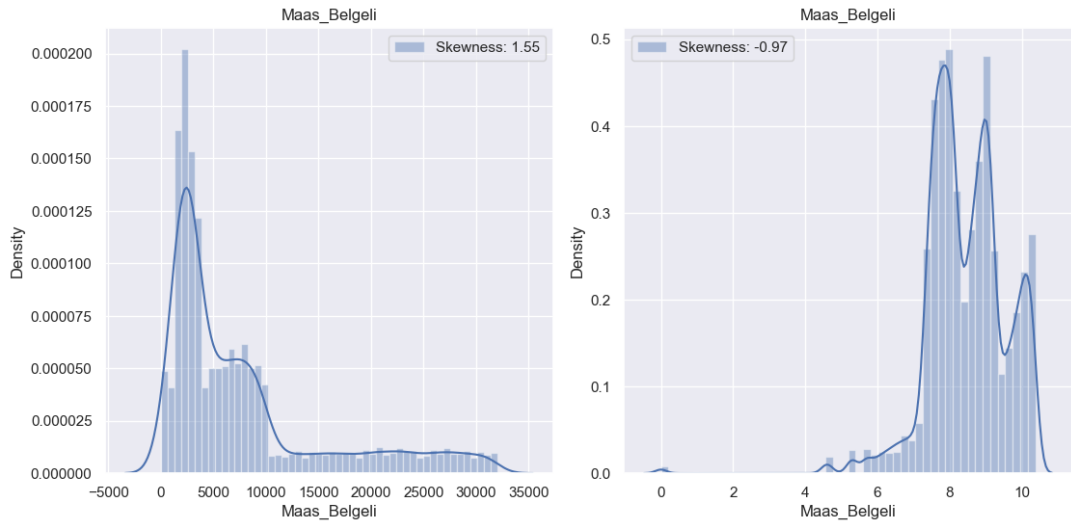
İkametgah_Yili alanı sayısal bir değer olması sebebiyle bu alanın ortalama değeri 5 olarak bulundu ve boş olan İkametgah_Yili alanları beş rakalı ile güncellendi.

Sektor alanına ait iki veri içeriği bulunmaktadır. Bunlar “Kamu Kurumu” ve “Özel Kurum”. Bu alana da one hot encoding yöntemi uygulanıp alan sayısı kolon sayısı

arttırıldı. Son olarak evet – hayır cevapları içeren “Son_1_Yil_Icinde_GM_Tarafindan_Ret_Edilmis_Tahsis_VarMi”, “Talebinin_Altinda_Teklif_VarMi” ve “Müsteri_Terkin_Olmus_Avukat_Icra_Takibine_Dusmus_Mu” alanlarına one hot encoding uygulandı. Son durumda kolon sayısı 20’ye çıkmış oldu.

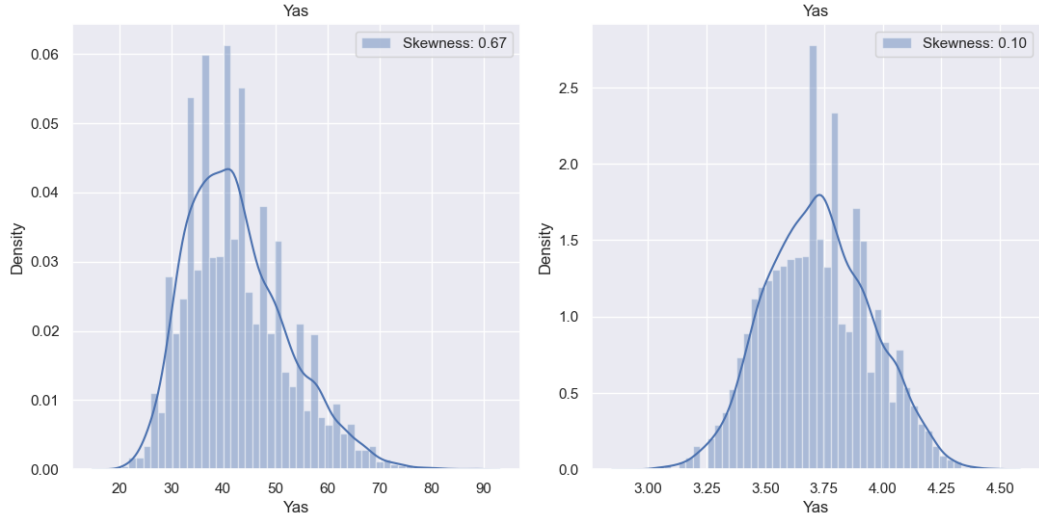
3.4.2 Veri dönüşümü

Veri dönüşümü işlemi sayısal olarak bulunan veriler üzerinde uygulandı. Şekil 3.8’de görüldüğü üzere Maaş alanının normal gösterimde çarpıklık değeri 1.55’ti. Maaş alanını logaritmik işleme tabi tuttuğumuzda çarpıklık değeri -0.97 oldu. Bu değer verinin işlenebilmesi adına daha sağlıklı bir aralığa yerleşti.



Şekil 3.8 Maas alanının normal ve logaritmik gösterimi

Şekil 3.9’da görüldüğü üzere Yas alanının normal gösterimde çarpıklık değeri 0.68’dir. Bu değer yüksek bir değer olmayıp bu haliyle değerlendirilebilmektedir. Yas alanını logaritmik işleme tabi tuttuğumuzda çarpıklık değeri 0.11’e düştü. Tüm satırların değerinin dolu olması ilk başta çarpıklık değerinin düşük olmasına sebep oldu.



Şekil 3.9 Yas alanının normal ve logaritmik gösterimi

Sayısal verilerin haricinde kalan kategori alanları "Medeni_Durum", "Cinsiyet", "Egitim_Durumu", "Meslegi", "Askerlik_Durumu", "Oturdugu_Ev_Durumu" ve "Risk_Grubu" Label Encoding işlemi ile sayısallaştırıldı. Tüm verileri sayısal hale getirdikten veriler %80 eğitim, %20 test verisi olacak şekilde ayrıldı. Sonrasında SMOTE (Bowyer vd., 2011) kütüphanesi kullanarak sentetik veri üretme işlemi ile veri seti 28,095 satıra ulaştı.

3.4.3 Makine öğrenmesi ile modelleme

Verilerimizi hazır hale getirdikten LightGBM, XGBoost, CatBoost, Random Forest, Decision Tree, Logistic Regression, Nearest Neighbors, Gaussian Naive Bayes, Stacking Classification ve Soft Voting modelleri çalıştırıldı. Çalışmaların sonuçları Accuracy, Precision ve Recall ve F1 Skor olarak gözlemlendi. Çizelge 3.4'te uygulanan makine öğrenmesi modelleri ve bu modellere ait sonuçlar görülmektedir.

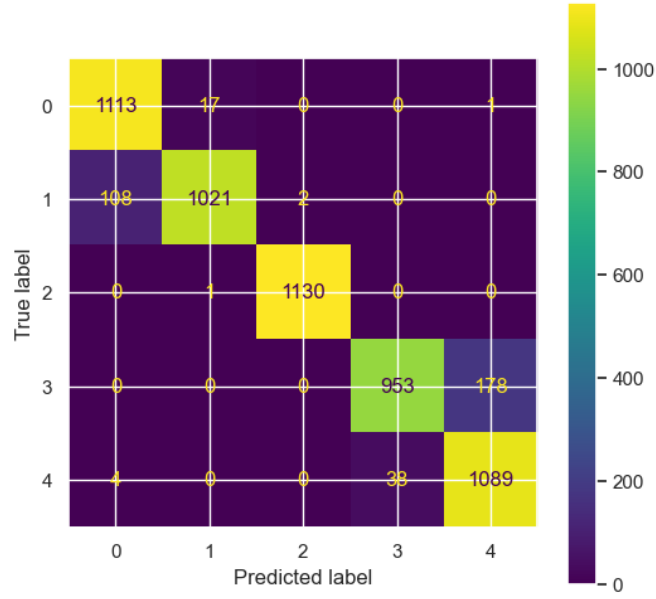
Çizelge 3.4 Model sonuçları

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	66.77	66.94	66.94	66.94
1 - Nearest Neighbors	59.77	63.09	59.77	61.39
3 - Nearest Neighbors	59.72	63.06	59.72	61.34
5 - Nearest Neighbors	59.22	61.86	59.22	60.51

7 - Nearest Neighbors	58.99	61.34	58.99	60.14
Gaussian Naive Bayes	51.72	70.12	51.72	59.53
Decision Tree	90.10	90.12	90.10	90.11
Random Forest (10 trees)	90.68	91.08	90.68	90.88
Random Forest (25 trees)	91.41	91.86	91.41	91.63
Random Forest (50 trees)	91.94	92.40	91.94	92.17
Random Forest (100 trees)	91.87	92.34	91.87	92.10
Stacking Classification	89.46	90.25	89.46	89.85
Soft Voting	91.09	91.19	91.09	91.14
CatBoost	<u>93.70</u>	<u>94.13</u>	<u>93.70</u>	<u>93.68</u>
XGBoost	93.39	93.67	93.39	93.37
LightGBM	<u>93.82</u>	<u>94.19</u>	<u>93.82</u>	<u>93.81</u>

Makine öğrenmesinde on farklı model denenmiştir. Bunlardan Nearest Neighbors 1,3,5 ve 7 neighbors parametresi ile, Random Forest ise 10,25,50 ve 100 trees parametreleri ile çalıştırılmıştır. Nearest Neighbors modelinde 1,3,5,7 neighborslar arasındaki ilişki doğrusal olarak gözlemlendi. Random Forest modelinde de benzer durum gözlemlenmiştir. Tüm modeller içinde en başarılı sonucu LightGBM verirken en düşük başarı oranı Gaussian Naive Bayes modelinde görüldü. LightGBM'den sonra ikinci sırada ise en başarılı model CatBoost oldu. Tüm on makine öğrenmesi algoritması ile gerçekleştirilen çalışmaya ait kod çıktılarına github üzerinden ulaşılmaktadır. (Karadag, Github.com 2024)

LightGBM modeline ait karışıklık matrisine göre True Positives değerleri 1113, 1021, 1130, 953, 1089 olarak ortaya çıkmaktadır. Bu değerler modelin her bir sınıf için doğru tahmin sayısını temsil etmektedir. Değerlerin yüksek olması modelin genel olarak yüksek doğrulukta olduğunu göstermektedir. False Positives ve False Negatives değerler ise modelin yanlış sınıflandırmalarını göstermektedir. Örneğin, gerçek sınıfı 0 olan ancak 1 olarak tahmin edilen 148 örnek bulunmaktadır. Bu tür False Positives ve False Negatives değerler modelin belirli sınıflar arasında karışıklık yaşadığını ifade eder. "Risk durumu" olarak "Çok Riskli" ile "Orta Riskli" ve "Çok İyi" ile "İyi" arasında büyük seviyede olmayan karışıklık yaşandığı görülmektedir. LightGBM'ye ait karışıklık matrisi Şekil 3.10'da verilmektedir.



Şekil 3.10 LightGBM modeline ait karışıklık matrisi

3.5 Kredi Skor Tespiti Önyüz Çalışmaları

Şekil 3.11’de bir müşteriye Kredi Tahsis Modülü’nün önyüz tasarımı gösterilmektedir. Toplam 14 alana veri girişi yapılarak müşterinin kredibilitesi tespit edilir. Bu alanlardan üçü sayısal değerlerdir. Bunlar; Maaş, Yaş ve İkametgah Yılı’dır. Diğerleri ise kategorik değerler olarak girilmektedir. Tüm veriler girildikten sonra “Skoru Getir” butonuna basarak uygulamanın arka tarafındaki Makine öğrenmesi modeli çalışır ve sonuç döndürülür. Sorgu sonucu ortaya çıkan sonuçlar Çizelge 3.5’te gösterilmektedir.

Çizelge 3.5 Skor sonuçları

Kredi Skoru	Kredi Durumu	Kredi Açıklaması
A	Mükemmel	En iyi kredi kartı segmentine sahip bir kart verilebilir, ev ve araç kredisi kullandırılır.
B	İyi	Çoğu kredi kartı ve kredi verilebilir, ancak en iyi fırsatlar verilmeyebilir.
C	Orta	Kredi oranlarında bir ayrıcalık yapılmaz, belirlenecek tahsis limiti yüksek olmaz.
D	Zayıf	Kredi oranı çok yüksek tutulur ve çok düşük limitli kredi kartı verilebilir.
E	Çok Zayıf	Çoğu kredi ve kredi kartı başvurusu reddedilir.

Kredi Skor Tahmini

Maaş :

Medeni Durum :

Cinsiyet :

Sektör :

Eğitim Durumu :

Mesleği :

Yaş :

Askerlik Durumu :

Talebinin Altında Teklif Var mı? :

Son 1 Yıl İçinde Ret Edilmiş Tahsis Var mı? :

Kanuni Takibi Sorunu Var mı? :

İcra Takibine Düşmüş mü? :

İkametgah Yılı :

Oturdugu Ev Durumu :

Skoru Getir

Müşterinin Kredi Skoru:A
Müşterinin Kredi Durumu:Çok İyi
Skor Açıklaması:En iyi kredi kartı segmentine sahip bir kart verilebilir, ev ve araç kredisi kullandırılır.

25000

Evli

Erkek

Kamu Kurumu

Fakülte/Lisans

Bankacı

40

Yapıldı

Hayır

Hayır

Hayır

Hayır

10

Kendine Ait

Şekil 3.11 Kredi skor çalışmasının önyüzü

3.6 Sonuçlar ve Tartışma

Bu çalışma, finansal kuruluşların kredi riskini değerlendirme ve kredi tahsis süreçlerini optimize etme yolunda makine öğrenimi tekniklerinin potansiyelini ele almaktadır. Geleneksel istatistiksel yöntemlere kıyasla, makine öğrenimi modelleri, büyük ve karmaşık veri kümelerindeki doğrusal olmayan ilişkileri

ortaya çıkararak daha doğru ve ayrıştırıcı tahminler sağlayabilir. Bu çalışmada da vurgulandığı gibi, müşterilerin demografik ve geçmiş finansal veri kaynaklarından yararlanma yeteneği, makine öğrenimi tabanlı kredi puanlama modellerini özellikle değerli kılmaktadır. Çalışma, 16151 satırlık gerçek sentetik veri üzerinde yapılan uygulamalı analizlerle farklı makine öğrenmesi modellerine ait çıktılar gözlemlenmiştir. Ayrıca SMOTE tekniği ile artırılan veri seti, bu tür algoritmaların gerçek dünya uygulamaları için ölçeklenebilirliğini ve potansiyelini göstermektedir.

Çalışma sonucunda, test edilen on farklı makine öğrenimi algoritması arasında LightGBM algoritmasının en yüksek başarı oranını elde ettiği görülmüştür. Bu bulgu, LightGBM'in kategorik verileri etkili bir şekilde işleyebilme ve aşırı öğrenmeyi önleme yetenekleri sayesinde, kredi riski değerlendirmesi gibi karmaşık bir görev için uygun bir araç olduğunu göstermektedir. Sonuç olarak, bu çalışma, finansal kuruluşların kredi tahsis süreçlerini iyileştirmek, riskleri en aza indirmek ve daha bilinçli kararlar almak için makine öğrenimi tekniklerinden nasıl yararlanabilecekleri konusunda önemli bilgiler sunmaktadır.

4. SOSYAL MEDYA ANALİZİ

Son yıllarda finans sektörü, kişisel kredi skorlaması için sosyal medya verilerinden yararlanmaya artan bir ilgi göstermiştir. Geleneksel finansal verilere kıyasla, sosyal medya verileri bu amaç için hem fırsatlar hem de zorluklar sunmaktadır (Guangming Guo 2016). Sosyal medya verilerini kullanmanın temel avantajlarından biri, bir borçlunun profilini daha kapsamlı bir şekilde değerlendirme potansiyeli sunmasıdır (Firmansyah ve Ramdani 2018) .

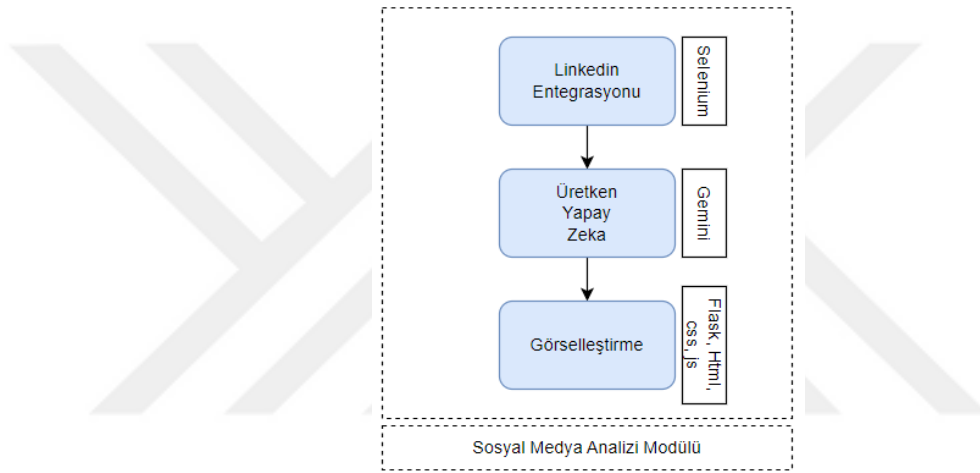
Geleneksel kredi değerlendirmesi genellikle müşterilerle ait geçmiş veya güncel finansal ölçütlere dayanır, ancak bu veriler özellikle sınırlı kredi geçmişine sahip bireyler için eksik veya güvenilir olamaz. Bu gibi durumlarda, sosyal medya verileri bir müşterinin mesleği, davranışı ve sosyal bağlantıları hakkında değerli bilgiler sunabilir. Bu da finansal bilgileri tamamlayabilir ve genel kredi karar verme sürecini iyileştirebilir (Óskarsdóttir vd., 2019).

Ancak, kredi analizi için sosyal medya verilerinin etkili kullanımı aynı zamanda zorluklar da sunmaktadır. Sosyal medya verilerinin heterojen ve yapılandırılmamış doğası, anlamlı içgörüler elde etmek için gelişmiş veri işleme ve analitik yetenekleri gerektirir. Ayrıca kredi kararları için kişisel sosyal medya bilgilerinin kullanılması konusunda dikkatle ele alınması gereken etik ve gizlilik endişeleri bulunmaktadır. Bu zorluklara rağmen, sosyal medya verilerini kredi skorlama modellerine dahil etmenin potansiyel faydaları, bu alanı bankacılık ve finansal hizmetler sektöründe artan ilgi ve araştırma alanı haline getirmektedir.

Çalışmanın bu kısmında, finansal verilerden bir müşterinin kredi skoru oluştuktan sonra meslek bilgilerini referans alarak yapılan sosyal medya analizi ile kredi kararı daha kararlı hale getirilmektedir. Müşterinin teyit ettiği meslek ve çalıştığı kurum bir otomasyon aracı olan Selenium ile alınarak daha sonra üretken yapay zeka ile mesleği ve çalıştığı iş yeri durumuna göre bir analiz çıkarılır. Bu analiz çıktısı kredi verilmesine karar verecek sisteme destek niteliğindedir.

4.1 Mimari Topolojisi

Bu araştırmada banka müşterilerinin teyit ettiği meslek bilgisi üzerinden bir çalışma yürütülmektedir. Sosyal medya hesabı LinkedIn üzerinden kişinin bilgileri Selenium otomasyon uygulaması ile sorgulanır. LinkedIn'den alınan meslek ve iş yeri bilgileri metin mühendisliği (prompt engineering) yapılarak Gemini isimli üretken yapay zeka kısmında sonuç üretmesi beklenir. Son olarak Python flask ile bir arayüz üzerinden görselleştirme yapılır. Tüm bu adımlara ait mimari topolojiye Şekil 4.1'de yer verilmektedir.



Şekil 4.1 Sosyal medya analiz modülünün mimari topolojisi

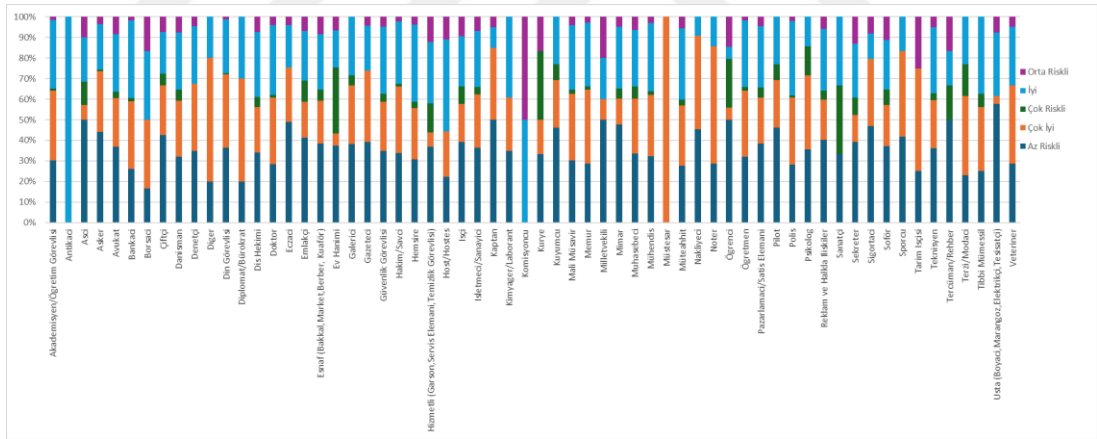
4.2 Veri Kaynağı

Araştırmanın bu aşamasında veri kaynağı olarak LinkedIn kullanılmaktadır. LinkedIn özellikle beyaz yaka çalışanların profillerinin daha doğru kabul edildiği, mesleklerinin ve iş yerlerine ait detaylı bilgilere ulaşılabilen bir sosyal medya platformudur. Bir banka müşterisinin bildirdiği meslek ve iş yeri bilgisinin teyitinin yapılacağı en doğru sosyal medya platformu LinkedIn'dir. LinkedIn üzerinden verileri almak için Selenium Web Driver kullanılmaktadır. Selenium WebDriver, açık kaynaklı bir araçtır ve programatik olarak web tarayıcılarını kontrol etmesini sağlar, web tabanlı görevlerin otomatikleştirilmesini mümkün kılar (Team 2019). Uygulama, esnek ve genişletilebilir bir çerçeve sunar, bu çerçevede yazılmış testlerin web sayfaları ile etkileşime geçmesini, form

doldurulmasını, buton tıklanmasını ve web sayfalarının içeriğini doğrulamanını mümkün kılar. Bunun yanında Selenium, kullanıcı etkileşimleri simüle etmenin kabiliyetine sahiptir.

4.3 Üretken Yapay Zeka

LinkedIn üzerinden Selenium otomasyonu ile alınan müşteriye ait meslek ve çalışma yeri bilgileri Şekil 4.2’de görülen verilerle birlikte Gemini üretken yapay zeka modeli ile yorumlaması istenmektedir. Gemini, Google DeepMind tarafından geliştirilen bir açık kaynaklı dil modelidir. Transformer mimarisine dayanarak, geniş bir veri kümesi üzerinde eğitilir ve metin üretme, çeviri, özetleme ve kod oluşturma gibi çeşitli NLP görevlerinde üstün performans gösterir. Gemini’nin açık kaynaklı doğası, araştırmacıların ve geliştiricilerin modelin yapısını inceleyerek, özelleştirerek ve yeni uygulamalar geliştirerek modelin potansiyelini daha da genişletmelerini sağlar (Deepmind 2024).



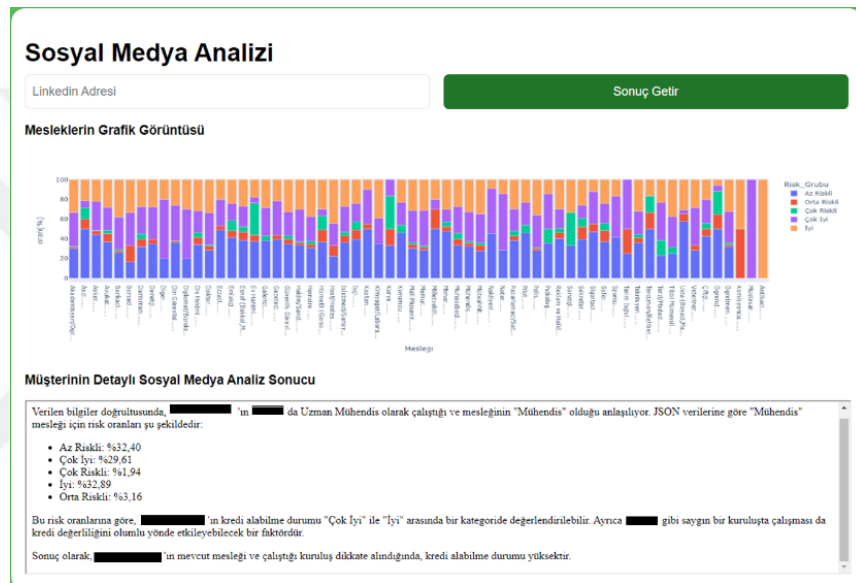
Şekil 4.2 Mesleklerin risk grubuna göre oranları

Şekil 4.2’de mesleklerin risk grubuna göre oranları verisi kredi skoru hesaplama aşamasında kullanılan veri setinin içinde yer alan mesleklerin verileridir. Bu veriler json formatına getirilerek Gemini modele girdi olarak verilmektedir. Girdi aşamasında Çizelge 4.1’de prompt verilerek sonuç beklenmektedir. Prompt, Gemini gibi dil modeline verilen spesifik metinleri ifade etmektedir.

Çizelge 4.1 Üretken yapay zeka modeline verilen metin

Prompt
Bilgiler:' + person_tr + title_tr + location_tr + 'Elimde bir veri seti var. Bu veriyi Json olarak sana veriyorum. Bu verilerde mesleğe göre risk oranı bulunuyor. Bu verileri ve bir önceki adımdaki kişinin çalıştığı yer ve mesleğini kullanarak kredi alabilme durumunu yorumlayabilir misin? Json Veri Seti:{"Meslek":"Akademisyen/Öğretim Görevlisi","Az Riskli":"30.14%","Çok İyi":"34.04%","Çok Riskli":"1.06%","İyi":"33.33%",.....}

4.4 Uygulama



Şekil 4.3 Sosyal medya analiz uygulamasının çıktısı

Sosyal medya analizine ait tüm geliştirme içerikleri github üzerinde paylaşılmaktadır (Karadag, Github.com 2024). Uygulama Selenium otomasyonu ve üretken yapay zeka promptu içermektedir. Bir banka müşterisinin kredi skoru tespiti yapıldıktan sonra LinkedIn hesabı üzerindeki meslek ve işyeri bilgileri Selenium uygulaması ile alınarak, üretken yapay zekaya girilmektedir. Sonuç olarak banka müşterisinin kredi verilip verilmeme durumu ve teyit ettiği meslek bilgisi için ek bir adım ortaya koyulmaktadır. Uygulama bir önyüz üzerinden Şekil 4.3'te görüldüğü üzere sonuç çıkarmaktadır. Sonuç, müşterinin meslek bilgisine ve çalıştığı yere göre bir yorumda bulunmaktadır. Ayrıca bu yorumu oluştururken json formatında promptta ek olarak verilen meslek – risk grubu

arasındaki ilişkiyi de hesaba katmaktadır. Bu sonuç banka personelinin yorumuna göre kredinin verilip verilmemesine katkı sağlamaktadır.

4.5 Sosyal Medya Analiz Sonuçları ve Kazanımlar

Günümüzde sosyal medya analizi konusu önem arz etse de veri toplama ve karar alma süreçlerinde hala istenilen seviyede görülmemektedir. Bir kurumsal firmanın sosyal medya verisi veya o firma için yapılan yorumlar daha sağlıklı alınabilirken, şahıslarda bu durum oldukça zordur. Birbirine benzer isim ve soyisim kombinasyonu ile aynı kişilere ulaşılabilinmektedir. Bu durum şahısların sosyal medya verilerini doğru bir şekilde almayı zorlaştırmaktadır. Bu araştırmada sosyal medya verilerinin alınması LinkedIn üzerinden gerçekleştirilmektedir. LinkedIn bir iş ve kariyer platformu olduğu için kişi profilleri diğer sosyal medya araçlarına göre daha doğru sonuç vermektedir.

Çalışma sonucu direk karar veren bir sistem yerine yardımcı bir sistem olarak tasarlanmaktadır. Bir önceki adımda kredi skoru tespit edildikten sonra büyük oranda müşteriye kredi verilip verilmeyeceğinin kararı belli olmaktadır. Ancak müşterinin beyan ettiği mesleği, çalıştığı yer ve kendi veri setimizde bulunan verilerle üretken yapay zeka aracılığı ile destek bir sistem oluşturulmaktadır. Bu sistem bugün için otomatik karar alma sürecine dahil edilmezken, ilerleyen çalışmalarda olması muhtemeldir. Çalışmada ayrıca üretken yapay zekaya yer verilerek finans alanında bir uygulama yöntemi de ortaya koyulmaktadır. Çalışma sonuçlarının doğru ve fayda sağlayan bir destek unsuru olduğunu kabul edildiğinde bu yöntemden yola çıkılarak finans alanında farklı araştırmaların da olması beklenmektedir.

5. BLOKZİNCİR TABANLI FİNANSAL VERİ PAYLAŞIMI

Günümüzde bankacılığın en temel konularından biri kredilerin tahsilatıdır. Kredilerin gecikmeye düşmeden tahsil olmasını sağlamanın ilk şartı ise, doğru müşteriye doğru kredinin kullandırılmasıdır. Kredi değerlendirme aşamasında, tahsis mercilerine ışık tutacak en kritik veriler istihbarat sorgulamaları ile elde edilmektedir. İstihbarat verileri genellikle merkezi kurumlardan elde edilmektedir. Merkezi kredi kayıt kurumları ihtiyaç duyulan tüm verileri sağlayamazlar ve gerçekleşen işlemde belli bir süre sonra bu verileri bankalar ile paylaşırlar. Ayrıca bu verilerin merkezi tek bir kuruluştaki bulunması manipülasyona açıktır.

Müşteriler ihtiyaç duydukları kredi ihtiyaçlarını bankalar vasıtasıyla karşılarlar. Bankalar, kredi onay işlemleri gerçekleşeceği zaman müşterilere ait istihbarat bilgilerine ihtiyaç duyarlar. Bu veriler Merkezi Kredi Kayıt Bürosu veya Risk Merkezi tarafından bankalardan toplanılır, konsolide edilir ve bankalar ile tekrardan paylaşılır. Bunun haricinde elde edilen verilerden, bankacılık işlemlerinde kullanılmak üzere her müşteriye özgü bir Findex (kredi) skoru üretilir. Bu skor Türkiye’de tek bir kuruluş tarafından üretilirken Amerika Birleşik Devletleri’nde FICO, Vantage, CE, Educational credit score gibi farklı türde ve farklı kuruluşlarca üretilen skorlar vardır.

Blokszincir, merkezi olmayan bir ağda birçok cihaz veya “düğüm” tarafından paylaşılan dağıtılmış bir defter olduğundan, paylaşılan bir kayıt tutma türüdür. Bu düğümlerin her biri, tüm kayıt blokszincirinin bir kopyasını tutar (Lemieux 2016). Ademi merkeziyetçilik, şeffaflık, sağlamlık, denetlenebilirlik ve güvenlik gibi ayırt edici özellikleri nedeniyle, blokszincirinin birçok geleneksel iş yapış şeklini etkilediği görülmektedir (Hoonsopon 2019). Blokszincir teknolojisinin bilinirliği Bitcoin ile başlamaktadır. Bitcoin’in teknik dokümanı, Satoshi Nakamoto adlı bir kişi veya grup tarafından 2008 yılında yazılan “Bitcoin: A Peer-to-Peer Electronic Cash System” başlıklı makaledir (Nakamoto 2008). Makale, bir finansal aracı kullanmadan blokszincir teknolojisi kullanılarak dijital varlıkların transferini öne sürmektedir. Elektronik para transferlerinin en ciddi

sorunlarından biri olan çifte harcama sorununu bu transferleri gerçekleştirirken nasıl çözdüğü de yazıda detaylı olarak anlatılıyor. Bu transfer sisteminin altyapısı, dağıtılmış bir defter yapısı sağlayan blokzincir teknolojisi üzerine kuruludur. Her ağ katılımcısı, blokzincirinde bir düğüm olarak kabul edilir ve düğümler tarafından yapılan işlemler, bloklar halinde birbirine bağlanır. Satoshi Nakamoto, Bitcoin teknik incelemesini 2008’de ilk kez yayınladığından beri, üç nesil blokzinciri kabul edildi. Blokzincir 1.0 kripto para işlemleri için, Blokzincir 2.0 finansal uygulamalar için ve Blokzincir 3.0 finans dışındaki hükümet, sağlık ve bilim gibi alanlardaki uygulamalar için kullanılıyor (Perera vd., 2020).

Blokzincir benzeri bir sistem ilk olarak kriptograf David Chaum tarafından 1982 tarihli “Karşılıklı Şüpheli Gruplar Tarafından Kurulan, Korunan ve Güvenilen Bilgisayar Sistemleri” tezinde önerildi (Sherman vd., 2019). Blokzincir, Bitcoin’den önce önerilmiş olsa da merkezi olmayan blokzincir kavramı ilk olarak Bitcoin ile kavramsallaştırıldı. 2013 yılında Ethereum’un piyasaya sürülmesiyle Vitalik Buterin, yalnızca bir kripto para birimi değil, aynı zamanda geliştiricilerin kendi akıllı sözleşmelerini programlayabilecekleri açık kaynaklı, merkezi olmayan bir blokzincir ağının altyapısını da önerdi. Ethereum blokzincir ağı artık programlanabilir ve uygulanabilen bir akıllı sözleşme yapısına sahiptir.

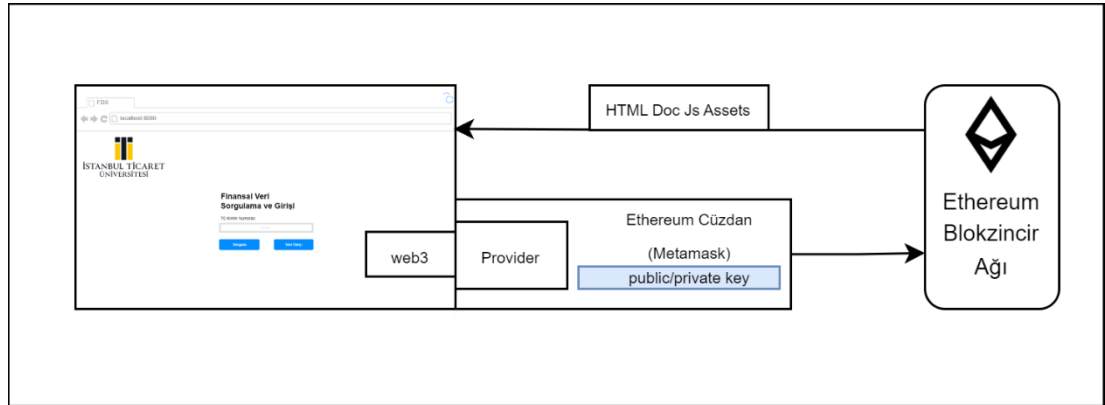
Bu çalışmada; finans kuruluşlarının kredi tahsisi yaptığı müşteriye ait limit, risk ve teminat bilgileri detaylı olarak blokzinciri altyapısı yoluyla eş zamanlı olarak paydaş olan tüm kuruluşlarla paylaşılır. İstihbarat birimleri bu bilgilere ulaşmak istediğinde ise hızlı bir şekilde kimlik numarası üzerinden sorgulayarak finans verilerine ulaşır. Blokzincir altyapısı üzerine tasarlanmış bir çalışma olması sebebiyle güvenilir, şifrelenmiş bir veriye çok hızlı bir şekilde ulaşılmış olur.

Finans kuruluşları arasında istihbarat verilerinin paylaşılması için kurulacak altyapıda, her bir düğümü bir finans kuruluşu bir düğüm olarak düşünülmektedir. Bir finans kuruluşu tarafından yapılacak finans veri girişi, ağa bağlı tüm diğer kuruluşlarda eş zamanlı olarak görülebilmektedir. Daha sonra ağa dahil edilecek yeni bir finans kuruluşu olursa, o kuruluş da geriye dönük tüm

blokları kendi tarafına çekebilmektedir. İlerleyen zamanda ağdan bir finans kuruluşu ayrılabilir bile mevcut işleyişte bir sorun olmayacaktır.

5.1 Mimari Topolojisi

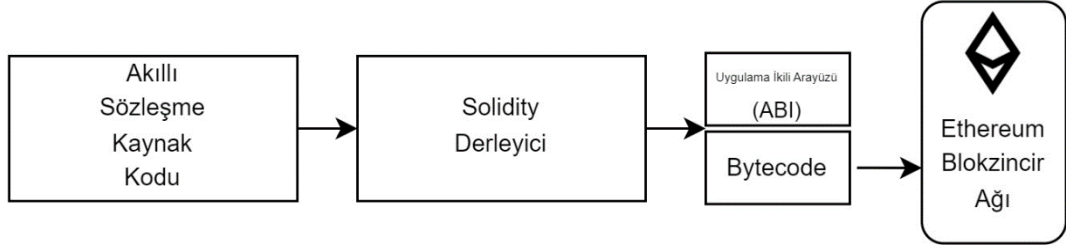
Sistemin mimari topolojisi Şekil 5.1’de görüldüğü üzere farklı finans kuruluşlarının bir web arayüzü üzerinden blokzincir ağına bağlanmasını göstermektedir. Sistem, ethereum ağı üzerinde kurgulanmaktadır. Hem test ağının hem canlı ortamının kolay uyarlanabilir, dokümantasyonunun detaylı ve anlaşılır, daha önce deneyimlenmiş birçok projenin olması ethereum platformunun seçiminde en önemli etkenlerdendir. Uygulama arayüzü merkeziyetsiz yapılara ulaşması için web3 altyapısını kullanmaktadır. Ethereum’un kurucu ortaklarından Gavin Wood, Web 3.0’dan ilk olarak kripto para biriminin 2014’teki piyasaya sürülmesinden kısa bir süre sonra bahsetmektedir. Web3 Ethereum’un dijital cüzdanı Metamask üzerinden ağa yazabilmektedir. Ethereum ağından okuma işleminde ise veriler HTML üzerinde sunulur.



Şekil 5.1 Blokzincir tabanlı finansal veri girişi ve sorgulama topolojisi

Yazılan bir akıllı sözleşmenin Ethereum ağında yaygınlaştırılması şekil 5.2’de gösterilmektedir. Ethereum’un programlama dili Solidity’de yazılan kod öncelikle derlenir ve sonrasında hem ABI hem de bytecode oluşur. ABI, json formatında olup Solidity dilindeki atama yapılan nesneleri gösterir. Ethereum, ağın temel bileşeni olarak EVM kullandığından, yüksek seviyeli dillerde yazılmış

akıllı sözleşme kodunun çalıştırılabilmesi için EVM bayt kodunda derlenmesi gerekir. EMV Baytcodu, EVM üzerinde çalıştırılabilir bir koddur ve akıllı sözleşme ABI, EVM bayt kodu ile etkileşim için bir arayüzdür.



Şekil 5.2 Blokzincir ağında akıllı sözleşmenin yaygınlaştırılması

5.2 Blokzincir Teknolojisi

5.2.1 Blokzincir kalite özellikleri

Yapılan çalışmada hangi blokzincir platformunun seçileceğine dair özellikler ve açıklamaları Çizelge 5.1’de gösterilmektedir. Çizelgede hem ürün özellikleri hem de tercih edilmesini sağlayan özellikleri açısından bir araştırma listesi oluşturuldu. Örneğin, izin türü, değerlendirilen blokzincir platformunun ürün özelliği iken yaygınlaştırma sıklığı ve dokümantasyonunun yeterli olup olmaması platformu tercih edilmesini sağlayan özellikler olarak ortaya çıkmaktadır. Bu özellikler değerlendirmeye tabi tutulan dokuz blokzincir platformu için incelenmektedir.

Çizelge 5.1 Blokzincir platformlarının kalite özellikleri

Özellik	Açıklama
Kripto Birimi (Cryptocurrency)	Kuruluş aynı zamanda bir kripto para birimine de sahip mi? Ethereum, Ether ve Ripple ise aynı isimde kripto para birimine sahiptir.
Akıllı Kontrat (Smart Contract)	Belirli şartlar sağlandığında gerçekleştirilen ve anonim taraflar arasında güvenilirliği sağlayan akıllı kontrat altyapısına sahip mi?
İzin (Permission)	Kullanıcı rollerini ve izinlerini tanımlama ilkesi (permissioned, permissionless, ikisi de)
Mutabakat (Consensus)	Madencilere atanacak uzlaşma metodu (POW, POS, POI, BFT)

Kullanılan Programlama Dili	Akıllı kontrat tanımlarken kullanılan programlama dili
Maliyet	Blokszincir platformunda akıllı kontratı çalıştırırken istenilen ücret
Ölçeklenebilirlik	Performansı arttırmak için kaynak ekleme kolaylığı
Performans	Saniyedeki işlem sayısı, geçerli bir işlemi yürütmek için ortalama süre
Bekleme Süresi	Akıllı kontratların yürütülmesi için beklemek gerekli mi?
Sanallaştırma	Platform tarafından kullanılan izolasyon sisteminin türü
Tokenizasyon	Token bazlı atama ilkesi
Yönetim	Platformu yöneten ve yürüten organizasyon
Topluluk Aktivitesi	Platformu geliştiren topluluk aktivitesinin büyüklüğü ve sıklığı
Destek	Kurulum ve platformun kullanımı konusunda destek veriliyor mu?
Lansman Tarihi	Platformun ilk kararlı versiyonunun yayınlanma tarihi
Sermaye	Platformun yönettiği dijitalleştirilmiş varlıkların değeri (para birimleri, kripto para birimleri, değerli metaller, emtialar, malzemeler olarak ifade edilir)
Yaygınlaştırma Sıklığı	Yeni sürümlerin yaygınlaştırma sıklığı
Dokümantasyon	Kullanıcı dokümanının varlığı ve teknik kapasitesi
İş/Sektör Odağı	İş odağı hangi uygulama alanında

Finans alanında kullanılabilecek blokszincir platformlarına ait özelliklerin değerlendirme karşılaştırması Çizelge 5.2, 5.3 ve 5.4'te verilmektedir. Ethereum, Corda, Hyperledger Fabric, Nem, Stellar, Ripple, Cardano, Eros ve Tron blokszincir platformlarına ait özellikler sıralanmaktadır.

Çizelge 5.2 Ethereum, Corda ve Hyperledger Fabric değerlendirme sonuçları

Özellik	Ethereum	Corda	Hyperledger Fabric
Kripto Birimi (Cryptocurrency)	Ether	Yok	Yok
Akıllı Kontrat (Smart Contract)	Evet	Evet	Evet
İzin (Permission)	Permissionless	Permissioned	Permissioned
Mutabakat (Consensus)	Proof of Work	valid and unique	BFT

Kullanılan Programlama Dili	Solidity, LLL, Serpent	Kotlin	Go, Node.js, Java
Maliyet	Evet	Hayır	Hayır
Ölçeklenebilirlik	Hayır	Evet	Evet
Performans (TPS)	14-27	*617	3.5
Bekleme Süresi	Evet	Hayır	Hayır
Sanallaştırma	VM	Docker	Docker
Tokenizasyon	Evet	Evet	Hayır
Yönetim	Ethereum	R3 Company	Linux Foundation
Topluluk Aktivitesi	Evet	Evet	Evet
Destek	Evet	Evet	Evet
Lansman Tarihi	Temmuz 2015	Ekim 2017	Mart 2017
Sermaye	Yüksek	Yüksek	Yüksek
Yaygınlaştırma Sıklığı	Yüksek	Yüksek	Yüksek
Dokümantasyon	Detaylı	Detaylı	Detaylı
İş/Sektör Odağı	Farklı	Finans	Farklı

*Corda Enterprise 4 (Issue +Repeated Pay)

Çizelge 5.3 Nem, Stellar ve Ripple değerlendirme sonuçları

Özellik	Nem	Stellar	Ripple
Kripto Birimi (Cryptocurrency)	Xem	Stellar	XRP
Akıllı Kontrat (Smart Contract)	Evet	Evet	Evet
İzin (Permission)	Permissionless & Permissioned	Permissionless & Permissioned	Permissionless
Uzlaşma (Consensus)	Proof of Importance	SCP	RPCA
Kullanılan Programlama Dili	Java	JavaScript, Java, Go	JavaScript, Java, Python
Maliyet	Evet	Evet	Evet
Ölçeklenebilirlik	Evet	Evet	Evet
Performans (TPS)	100	1	1500
Bekleme Süresi	Hayır	Hayır	Evet
Sanallaştırma	Docker	Docker	Docker
Tokenizasyon	Evet	Evet	Evet
Yönetim	Nem	Stellar	XRPL Foundation
Topluluk Aktivitesi	Düşük	Hayır	Evet
Destek	Hayır	Hayır	Evet
Lansman Tarihi	Mart 2015	Temmuz 2014	Haziran 2012
Sermaye	Düşük	Düşük	Yüksek
Yaygınlaştırma Sıklığı	Yüksek	Yüksek	Yüksek

Dokümantasyon	Detaylı	Detaylı	Detaylı
İş/Sektör Odağı	Farklı	Finans	Farklı

Çizelge 5.4 Cardano, Eos ve Tron değerlendirme sonuçları

Özellik	Cardano	Eos	Tron
Kripto Birimi (Cryptocurrency)	Ada	Eos	Tron
Akıllı Kontrat (Smart Contract)	Evet	Evet	Evet
İzin (Permission)	Permissionless	Permissioned	Permissionless
Uzlaşma (Consensus)	Proof of Stake	Delegated Proof of Stake	Delegated Proof of Stake
Kullanılan Programlama Dili	Marlowe, Plutus	C++	Solidity
Maliyet	Evet	Hayır	Evet
Ölçeklenebilirlik	Evet	Evet	Evet
Performans (TPS)	15-20	4000	2000
Bekleme Süresi	Evet	Hayır	Evet
Sanallaştırma	Docker	Docker	VM
Tokenizasyon	Evet	Evet	Evet
Yönetim	Cardano Foundation	Black.one	Tron Network
Topluluk Aktivitesi	Evet	Evet	Evet
Destek	Evet	Evet	Evet
Lansman Tarihi	Eylül 2017	Haziran 2018	Mayıs 2018
Sermaye	Yüksek	Orta	Orta
Yaygınlaştırma Sıklığı	Yüksek	Düşük	Düşük
Dokümantasyon	Detaylı	Detaylı	Detaylı
İş/Sektör Odağı	Finans	Farklı	Farklı

Farklı blokzincir altyapıları incelendi ve çalışmada kullanılacak Ethereum altyapısı şu kriterlere göre seçildi.

- Ethereum diğer altyapılara göre daha kapsamlı bir dokümana sahiptir. Ayrıca bir problem ile karşılaşılması durumunda dokümantasyon harici başka kaynaklardan da bilgi edinilebilmektedir.
- Ethereum'un diğer altyapılara göre daha yaygın kullanımı bulunması.
- Akıllı sözleşmelerin programlanması amacıyla basit ve kullanışlı Solidity diline sahip olması.
- Arkasında güçlü bir ekosisteminin bulunması.
- Finans alanında geliştirilmiş çalışmaların varolması.

5.2.2 Blokzincir türleri

Blokzincir uygulaması oluşturmanın ilk adımı doğru altyapıyı seçmektir. Dört ana tip blokzincir ağı mevcuttur (Wegrzyn ve Wang, 2021). Bankalar arasında Özel blokzincir ağı veya Konsorsiyum blokzincir ağı kullanılması uygun görülmektedir. Şekil 5.3 blokzincir türlerini ve bunların birbirleriyle ilişkilerini göstermektedir.

5.2.2.1 Halka açık Blokzincir ağı

Bitcoin bu tür merkezi olmayan ağın en bilinen örneğidir. Halka açık blok zinciri, paydaşlar arası çalışan, herkesin ağa katılmasına, ledger (muasebe defterine) erişmesine ve işlemleri doğrulamasına olanak tanıyan merkezi olmayan bir ağıdır. Bu tür ağlara merkezi olmayan ağ denir, yani merkezi bir otorite yoktur. Tüm katılımcılar (paydaşlar) eşit haklara sahiptir ve yeni veri blokları oluşturup doğrulayarak ağa katkıda bulunabilirler. Tüm işlem verileri tüm katılımcılar tarafından görülebilen genel bir muasebe defterine kaydedildiğinden, şeffaflığı ve açıklığı ile karakterize edilir. Bu özellik, verilerin ağdan fikir birliği olmadan değiştirilememesini sağlayarak sistemin güvenliğini ve güvenilirliğini artırır (Nakamoto 2008). Dahası, genel blok zincirleri, işlemleri doğrulamak için PoW veya PoS gibi fikir birliği mekanizmaları kullanır ve bu da herhangi bir tek varlığın ağı kontrol etmesini önleyerek ek bir güvenlik katmanı sağlar (Buterin 2014). Halka Açık blok zincirlerinin merkezi olmayan yapısı, bunları finansal işlemler, tedarik zinciri yönetimi ve dijital kimlik doğrulaması gibi şeffaflığın, güvenliğin ve güvenin en önemli olduğu uygulamalar için özellikle uygun hale getirir. Ancak bu avantajlar, ölçeklenebilirlik zorlukları ve özellikle PoW tabanlı sistemlerde yüksek enerji tüketimi gibi dezavantajları da beraberinde getiriyor (Croman vd., 2016).

5.2.2.2 Özel Blokzincir ağı

Özel blokzincir ağı, diğer adıyla izinli blokzincir, katılımın yalnızca belirli katılımcılara izin verildiği bir tür dağıtık defter teknolojisidir. Bu tür ağlarda,

halka açık blokzincir ağlarının aksine, yalnızca yetki verilmiş katılımcılar işlem doğrulama ve blok oluşturma işlemlerini gerçekleştirebilir. Özel blokzincir ağları, genellikle bir organizasyon veya konsorsiyum tarafından yönetilir ve bu yönetim modeli, ağın gizliliği, ölçeklenebilirliği ve verimliliği üzerinde olumlu etkiler sağlar. Bu tür ağlar, uzlaşma mekanizmalarının performans odaklı optimize edilmesine olanak tanır ve katılımcıların ihtiyaçlarına göre özel kurallar ve izinler uygulamaya elverişlidir. Finans sektörü gibi veri gizliliği ve düzenleyici çerçevelere uyum gerektiren alanlarda özel blokzincir ağları, verilerin gizliliğini korumak ve kamuya açık olmasını engellemek açısından avantajlıdır. Özel blokzincir ağları, veri gizliliğini ve ağ üzerindeki kontrolü artırarak, kamuya açık maruziyetin engellenmesi gereken kurumsal uygulamalar için uygun bir seçenek sunar. Bununla birlikte, bu avantajların karşılığında, şeffaflığın azalması ve ağın yönetiminden sorumlu merkezi otoriteye olan güvenin artması, blokzincir teknolojisinin temelini oluşturan bazı merkeziyetsizlik ilkelerinin zayıflamasına neden olabilir (Zheng vd., 2017).

5.2.2.3 Hibrit Blokzincir ağı

Hibrit blokzincir ağı hem özel hem de halka açık blokzincir ağlarının özelliklerini birleştiren bir yapı sunar. Bu tür ağlar, katılımcıların belirli kısımlara erişimini kontrol eden özel blokzincir'in gizliliğini ve güvenliğini, halka açık blokzincir'in şeffaflığı ve dağıtılık yapısıyla birleştirir. Hibrit blokzincir, özellikle karmaşık iş süreçlerinde ve veri gizliliği ile kamusal doğrulama gerektiren durumlarda ideal bir çözüm olarak öne çıkar. Örneğin, bir şirketin hassas verilerini özel bir blokzincir 'de tutarken, bu verilerin doğruluğunu halka açık bir blokzincir aracılığıyla umuma kanıtlaması mümkündür. Hibrit blokzincir ağları, farklı kullanım senaryoları için esneklik sağlayarak hem ölçeklenebilirlik hem de güvenlik ihtiyaçlarını karşılamada etkili bir model sunmaktadır. Bununla birlikte, hibrit yapıların karmaşıklığı, ağın yönetim ve bakım süreçlerinde ek zorluklar yaratabilir, bu da iyi tanımlanmış bir yönetim modelinin gerekliliğini ortaya koyar (Xu vd., 2017). IBM'in tedarik zincirlerinde kullanılan “Food Trust” projesi bu tür ağlara güzel bir örnektir.

5.2.2.4 Konsorsiyum Blokzincir ağı

Blokzincir teknolojisi alanında konsorsiyum ağları, özel ve halka açık ağlara kıyasla farklı bir yapı sunar. Tek bir düğüm tarafından kontrol edilen özel blok zincirlerin aksine, konsorsiyum blokzincirleri önceden seçilmiş bir grup kuruluş tarafından yönetilir. Konsorsiyum blokzincir ağı, yarı-özel bir blokzincir türü olarak tanımlanır ve bu ağlarda birden fazla kuruluşun katılımı söz konusudur. Bu tür ağlarda, bir konsorsiyum tarafından belirlenen bir grup önceden yetkilendirilmiş katılımcı, blok oluşturma ve işlem doğrulama süreçlerinden sorumludur. Konsorsiyum blokzincir'leri, hem özel hem de halka açık blokzincir'lerinin avantajlarını birleştirerek, güvenilirliğin yüksek olduğu, ancak tamamen merkezi olmayan bir yapı gerektirmeyen durumlar için ideal bir çözüm sunar. Örneğin, bankacılık, sigortacılık, tedarik zinciri yönetimi gibi alanlarda, birden fazla bağımsız kuruluşun veri paylaşımına ihtiyaç duyduğu durumlarda kullanılabilir. Bu ağlar, veri güvenliği, işlem hızları ve ölçeklenebilirlik açısından avantajlar sunarken, aynı zamanda birden fazla kuruluşun işbirliğini ve mutabakatını gerektirir. Konsorsiyum blokzincir ağlarının, kurumsal işbirliğinin ve süreçlerin etkin bir şekilde yönetilmesi için güvenilir bir altyapı sağladığını, ancak katılımcıların güvenlik ve gizlilik ihtiyaçlarının dikkatle ele alınması gerektiğini belirtmektedirler (Jacob Eberhardt 2017) .



Şekil 5.3 Blokzincir türleri

5.2.3 Akıllı kontrat

İlk olarak 1994 yılında Nick Szabo tarafından ortaya atılan akıllı sözleşmeler kavramı, bir blokzincir ağında konuşlandırılan, kendi kendini doğrulayan ve kendi kendini çalıştıran bilgisayar programlarını ifade eder. Esas olarak önceden tanımlanmış kurallar dizisinden oluşan bu programlar, merkezi olmayan bir şekilde çalışır ve blokzincir'in doğasında olan değişmezlik nedeniyle değiştirilmeye karşı dayanıklıdır. Akıllı sözleşmeler, ağ genelinde çoğaltma ve eşler arası gözetim yoluyla, anlaşmaların güvenli ve şeffaf bir şekilde yürütülmesini sağlar (Szabo 1997)

Merkezi olmayan defter teknolojilerine, özellikle blokzincir'e yapılan büyük miktardaki yatırım ve ilgili elektronik defterlerdeki işlemleri kaydetmek, paylaşmak ve senkronize etmek için yeni bir çözüm olarak kabul edilmeye yönelik hızlı gelişimiyle, Ethereum akıllı sözleşmeleri tercih edilen altyapı haline gelmektedir. Ethereum, finansal araçlar da dahil olmak üzere bankacılık sektöründe geniş bir uygulama yelpazesini hayata geçirmektedir (Laarabi ve Maach. 2020). Ethereum akıllı sözleşmelerinin bankacılık sektöründe kullanımı büyük ilgi ve ilgi görmüştür. Bankalar, Ethereum akıllı sözleşmelerini operasyonlarına dahil etme potansiyelini giderek daha fazla araştırmaktadır. Akıllı sözleşme, sözleşme koşullarının doğrudan kod satırlarına yazıldığı, kendi kendini yürüten bir sözleşmedir. Bu sözleşmeler, araçlara veya üçüncü kişilere ihtiyaç duymadan, önceden belirlenen koşullar sağlandığında otomatik olarak yürürlüğe girer. Ethereum Ağı bağlamında, yürütülecek koşulları ve eylemleri içeren kod yazılarak akıllı bir sözleşme oluşturulabilir. Bu kod daha sonra değişmez ve şeffaf bir halde Ethereum Ağı'na dağıtılır. Bir kez dağıtıldığında, Ethereum Ağı üzerindeki akıllı sözleşmeyle birden fazla taraf etkileşime geçebilir.

5.2.4 Uzlaşma

Ethereum, PoW konsensüs mekanizması kullanarak işlemleri doğrulamak ve blok zincirini güvenli hale getirmek için tasarlanmıştır. Bu mekanizmada, ağdaki düğümler, yeni blokları eklemek için karmaşık matematiksel problemleri çözmek

için rekabet eder. Bu problemleri çözen düğüm, yeni bloğu ekleme ve ödül olarak ETH kazanma hakkına sahip olur. PoW, merkezi olmayan bir ağ için güvenlik ve doğruluğu sağlamak için etkili bir yöntem olarak kabul edilir, ancak enerji tüketimi ve merkeziyetsizliğe olan potansiyel tehditleri nedeniyle eleştirilir (Buterin 2014). Bu tehditlerin en başında ise Sybil saldırıları gelmektedir.

Sybil saldırıları tehdidini azaltmak için çeşitli blokzincir projeleri, merkezi olmayan ve Sybil'e dayanıklı konsensüs mekanizmalarını uygulamaya koydu. Bu mekanizmalar, kimlik doğrulama için dağıtılmış bir meşruiyet kaynağı oluşturmak amacıyla oylama, kefil olma ve yorumlama gibi öznel girdilere dayanır. Bu yaklaşım, "doğrulamayı kim doğruluyor?" şeklindeki kritik soruyu kabul etmektedir.

Halka açık blokzincir ağları, açık doğaları nedeniyle doğası gereği Sybil saldırılarına karşı hassastır. Buna karşılık, sınırlı erişime sahip kapalı ekosistemler olarak işlev gören özel ağlar, bu tür saldırılara ilişkin önemli ölçüde daha düşük risk sunar. Konsensüs, düğümler arasında verilerin geçerliliğine ilişkin toplu sözleşmeyi ifade eder. Bu fikir birliğine, PoW veya PoS gibi kriptografik fikir birliği algoritmaları aracılığıyla ulaşılır. PoW ve PoS, blokzincir sistemlerindeki Sybil saldırılarına karşı dayanıklılıklarıyla tanınan yerleşik fikir birliği mekanizmalarıdır (Hu vd., 2018). Bu çalışma, güvenli bir özel Ethereum ağ ortamında PoS'un potansiyelini araştırıyor. Bulgular, özel bir ağ ve güçlü bir PoS fikir birliği mekanizmasından yararlanılarak, Sybil saldırı risklerini etkili bir şekilde azaltan güvenli bir mimarinin elde edilebileceğini göstermektedir.

5.3 Uygulama

Blokzincir teknolojisinin uygulanması, finans sektöründe kredi tahsis verilerinin güvenli ve şeffaf bir şekilde saklanması ve paylaşılması için cazip bir imkan sunmaktadır. Bu dağıtılmış defter sistemi, bankalar, borçlular ve diğer ilgili taraflar da dahil olmak üzere kredi tahsis sürecinde yer alan tüm paydaşlar arasında güven ve şeffaflığı mümkün kılmaktadır. Ayrıca blokzincir teknolojisi, kredi tahsis işlemlerinin verimliliğini önemli ölçüde arttırmaktadır. Bankalar,

kredi tahsis verileri için bu teknolojiden yararlanarak borçlu bilgilerinin doğrulanmasını ve onaylanmasını kolaylaştırır, bu da geleneksel süreçlerle yapılan işlemlere nazaran zaman ve maliyetlerin azalmasına sebep olmaktadır.

Müşteri kredi kullanmak istediğinde geçmiş ve güncel finansal bilgileri kullanılır. Müşteri krediyi kullandıktan sonra limit, risk ve teminat detayları çok önemlidir. Bankalar kendi müşterilerine ait bu verileri aylık olarak saklamaktadırlar. Aynı müşterinin farklı bir bankadan kredi kullanmak istemesi durumunda limit, risk ve teminat bilgilerinin değerlendirilebilmesi için bu verilere ihtiyaç duyulmaktadır. Merkezi bir sistemin mevcut olması halinde, başka bir bankanın erişim talebinde bulunması durumunda bu verilerin elde edilmesi için kullanılır. Bu tür tüketici veri işlemleri için de blokzincir temelli modeller kullanılmaktadır. Blokzincir teknolojisi üzerine uygulanan model anında, şeffaf ve değişmez bir yaklaşım sunmaktadır.

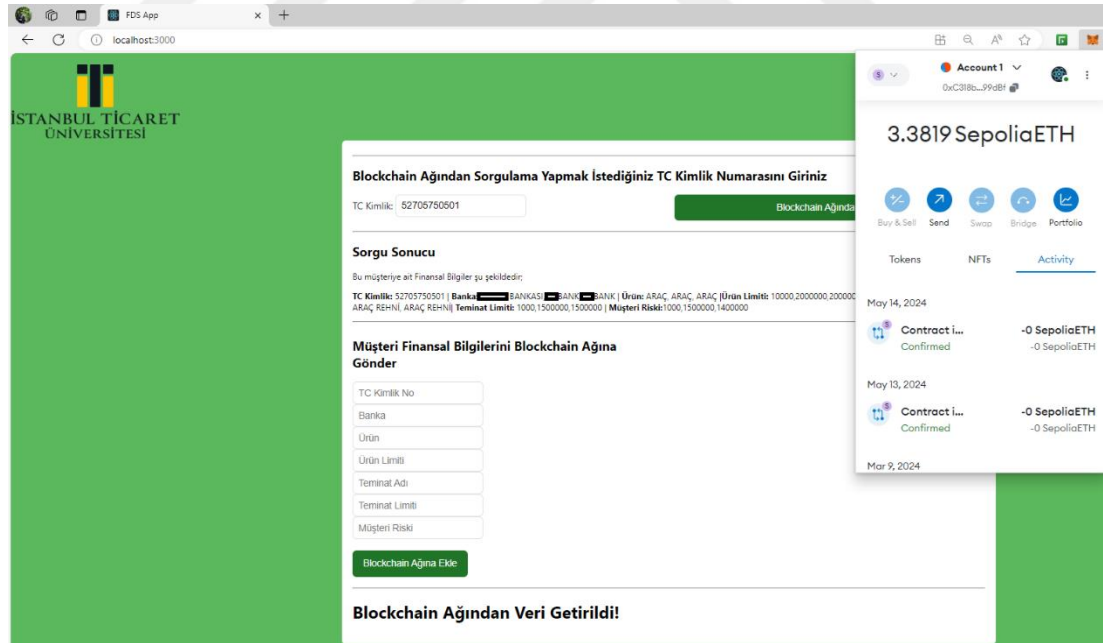
Bu araştırmada öncelikle blokzincir tabanlı model için akıllı sözleşme geliştirildi. Akıllı sözleşme Solidity'de geliştirildi ve Ethereum ağında çalışacak şekilde düzenlendi. Bu akıllı sözleşme ile müşterinin risk limiti ve teminat bilgileri blokzincir ağına kaydedilmektedir. Kayıtlara ulaşılması gerektiğinde kimlik numarası üzerinden sonuç alınabilmektedir (Karadag, Github.com 2024). Solidity dilinde yazılan akıllı sözleşme önce derlenir, ardından Uygulama ABI ve Bayt Kodu oluşturulur. Bayt kodu Ethereum ağında kullanılır. ABI, akıllı sözleşmeyi arayüz programından çağırmak için kullanılan JSON tabanlı bir kod dizisidir. Şekil 5.2, akıllı sözleşmenin Ethereum ağında uygulanmasına yönelik süreçleri göstermektedir.

Şekil 5.1'de gösterildiği gibi, Ethereum ağında konuşlandırılan bir akıllı sözleşmeye erişim, bir ön uç uygulaması, bir Ethereum cüzdanı ve bir sağlayıcı arasındaki etkileşimi gerektirir. Infura gibi sağlayıcılar bu etkileşimi kolaylaştırmaktadır. Ön uç geliştirme sırasında akıllı sözleşme adresi ve ona karşılık gelen ABI kodu çok önemli unsurlardır. Bu bileşenlerden yararlanan ön uç uygulama, seçilen sağlayıcı aracılığıyla kullanıcının Ethereum cüzdanıyla bağlantı kurar. Özellikle, verilerin blokzincirine yazılması (örneğin finansal

bilgilerin kaydedilmesi) bir gas ücretine tabi olurken, mevcut verilerin okunması genellikle ücretsizdir. Ethereum blok zincirinde işlem ücreti, 5.1'de olduğu gibi, hesaplama birimi maliyeti (gas fiyatı) ve kullanıcı tarafından belirlenen maksimum tahsis edilen hesaplama birimleri (gas limit) tarafından belirlenir.

$$\text{İşlem Maliyeti} = \text{Gaz Fiyatı} * \text{Gaz Limiti} \quad (5.1)$$

Araştırma sonucuna ait uygulamanın önyüzü Şekil 5.4'te gösterilmektedir. Ethereum ağı üzerinde tasarlanmış ve Metamask cüzdan aracılığı ile işlem yapılan sistem bir TCKN üzerinden sorgu işlemini gerçekleştirir. Kredi tahsisi gerçekleşen müşteriye ait TCKN, banka, limit, ürün, ürün limiti, teminat adı, teminat limiti ve risk bilgileri blokzincir ağına kaydedilir. Kaydedilirken işlemin Ethereum cinsinden bir işlem ücreti bulunur. Ağa kaydedilen finansal bilgiler bir başka node (banka) tarafından müşteriye ait TCKN üzerinden sorgulanır. Sorgulama sonucu bir işlem ücreti gerektirmez.



Şekil 5.4 Blokzincir tabanlı finansal veri paylaşım önyüzü

5.4 Sonular ve Tartışma

Blokzincir tabanlı finansal veri paylaşımı modelini temel alan bu alıřma, kredi tahsis edilen banka mřterilerinin kredi tahsis verilerinin paylaşımını ortaya koymaktadır. Bankalar arası veri paylaşımı, Ethereum ağı zerinde Solidity diliyle yazılmış bir akıllı sözleşme ile mümkün olmaktadır. Hazırlanan akıllı sözleşmenin Ethereum ağına konuřlandırılmasından ve ardından bu ağ zerinden verinin nasıl yazıp okunacağından bahsedilmektedir. Ayrıca özel blokzincir ağı zerinde yapısı ve PoS konsenss yntemi sayesinde Sybil saldırılarına karřı gvenli bir ortam saėlamaktadır. Blokzincir tabanlı finansal veri paylaşımı modeli, özel Ethereum ağı ve PoS konsenss mekanizması dikkate alınarak tasarlandı. Bu řekilde blokzincir teknolojisi, bankacılıkta geleneksel finansal veri paylaşımı eksikliklerine dnřtrc bir czm sunmaktadır. Paylaşılan, deėiřmez bir defter, tahsis srecini kolaylařtırır, veri gvenliėini artırır ve gerek zamanlı risk deėerlendirmesine katkı saėlar.

Kresel verilerin katlanarak bymesi, paydařlar arasında gvenli depolama ve verimli paylaşımı gerektirmektedir. Blokzincir teknolojisi, gvenli ve řeffaf veri alışveriřini kolaylařtırarak bu alanda nc olarak ortaya çıkmaktadır. Bu zelliėin nmzdeki yıllarda finans sektrnde daha fazla benimsenmesine yol aması muhtemeldir. Ancak dzenleme ve uyumluluk engelleri devam etmektedir. Bu zorlukların stesinden gelmek, finansal kurumlar iin ok sayıda kullanım senaryosunun kilidini aacaktır. Byle bir senaryo, bankacılıkta yaygın bir ara olan Teminat mektuplarının NFT'ler aracılığı ile saklanmasına olanak vermektedir. Teminat mektuplarını NFT'lerle tokenize eden teknoloji, doėası gereėi kopyalanmayı ve sahteciliėi nleyerek bankaların ilgili riskleri daha etkili bir řekilde ynetmesine olanak tanıyor. NFT'lerin finansal ekosisteme entegrasyonu, geliřmiř gvenlik, verimlilik ve iř birliėini teřvik etme konusunda umut vadetmektedir.

6. SİSTEMİN GENEL İŞLEYİŞİ

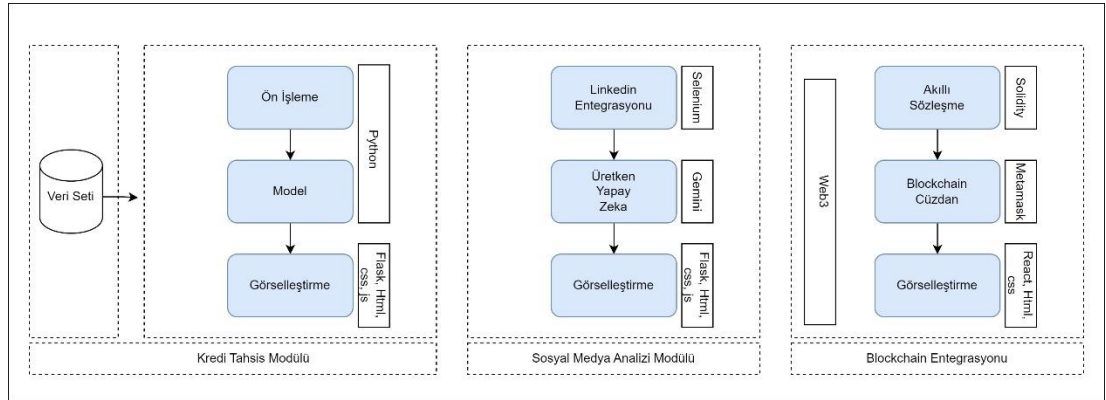
Bu araştırma, kredi değerlendirme süreçlerini iyileştirmek ve finansal istihbarat paylaşımını optimize etmek amacıyla üç aşamalı bir model önermektedir. İlk aşamada, makine öğrenmesi algoritmaları kullanılarak banka müşterilerinin sentetik finansal-demografik ve sosyal medya verileri temel alınarak kapsamlı bir kredi riski değerlendirmesi gerçekleştirilir. Bu analiz, müşteriye özel kredi risk profillerinin oluşturulmasını, skorlamasını ve hızlı ve etkili kredi tahsis süreçlerinin hayata geçirilmesini sağlar. İkinci aşamada, müşterinin beyan ettiği meslek ve işyeri bilgilerini doğrulamak ve kredi kararlarını desteklemek amacıyla sosyal medya analizi gerçekleştirilir. Bu analiz, Selenium gibi otomasyon araçları ve üretken yapay zeka teknolojileri kullanılarak gerçekleştirilir. Üçüncü aşamada ise, onaylanan kredilere ilişkin limit, risk ve teminat gibi kritik istihbarat bilgileri, blokzinciri teknolojisi kullanılarak eş zamanlı ve güvenli bir şekilde paydaş finans kuruluşlarıyla paylaşılır. Bu merkezi olmayan ve şifrelenmiş yapı, istihbarat birimlerine kimlik numarası üzerinden hızlı ve güvenilir veri erişimi imkanı sunarken, finansal veri paylaşımında şeffaflığı ve veri güvenliğini artırır.

6.1 Mimari Topolojisi

Araştırma üç aşamadan oluşmaktadır. Bu üç aşamaya ait topoloji Şekil 6.1'de gösterilmektedir. İlk aşama olan kredi riski değerlendirmesi için, 16151 satırlık senetetik seti kullanılmıştır. Bu veri seti, ön işleme aşamasında eksik ve aykırı değerlerin ele alınması ve normalizasyon işlemleri ile optimize edilmiştir. One Hot Encoding yöntemi ile yeni özellikler türetilmiş ve SMOTE tekniği ile veri seti 28.095 satıra genişletilmiştir. Modelleme aşamasında, on farklı makine öğrenmesi algoritması karşılaştırılmış ve en yüksek performansı gösteren LightGBM algoritması seçilerek .pkl uzantılı bir model oluşturulmuştur. Bu model, web tabanlı bir arayüz üzerinden görselleştirilmek üzere Python'da Flask kütüphanesi kullanılarak entegre edilmiştir.

İkinci aşamada, finansal verilere dayalı kredi skoru hesaplamasının ardından, müşterinin beyan ettiği meslek ve işyeri bilgilerini doğrulamak ve kredi kararlarını daha güvenilir hale getirmek için sosyal medya analizi gerçekleştirilir. Müşteri tarafından sağlanan meslek ve işyeri bilgileri, Selenium kullanılarak otomatik olarak toplanır ve daha sonra üretken yapay zeka ile analiz edilerek, kredi değerlendirme sürecine destekleyici bilgiler üretilir.

Son aşamada, kredi tahsis sürecinin ardından, finans kuruluşları tarafından müşteriye sağlanan limit, risk ve teminat gibi istihbarat bilgileri, blokzinciri altyapısı aracılığıyla eş zamanlı olarak paydaş tüm kuruluşlarla paylaşılır. Her bir finans kuruluşunun bir düğüm olarak temsil edildiği bu dağıtık yapı, veri güvenliğini ve şeffaflığı en üst düzeye çıkarır. İstihbarat birimleri, kimlik numarası sorguları aracılığıyla bu verilere hızlı ve güvenli bir şekilde erişebilirken, sisteme yeni katılan finans kuruluşları da geçmişe dönük tüm bloklara erişim sağlayabilir. Bu merkezi olmayan yapı, ağdan bir kuruluşun ayrılmasının bile sistemin işleyişini etkilememesini sağlayarak sağlamlık ve süreklilik sunar.

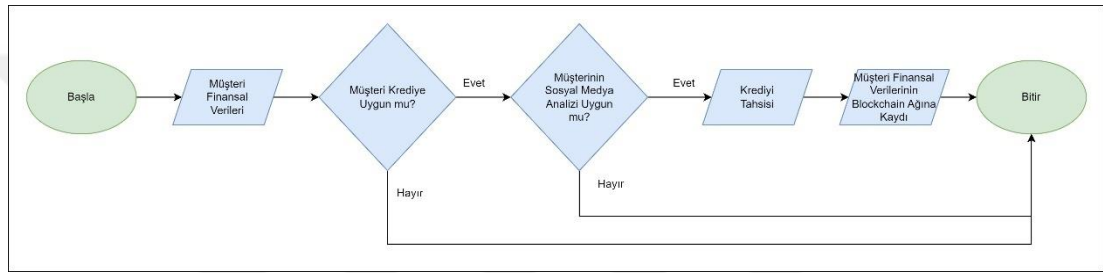


Şekil 6.1 Sistemin genel topolojisi

6.2 İş Akışı

Kredi başvuru süreci müşterinin bankaya talepte bulunmasıyla başlamaktadır. Müşteri bir kredi kartı veya kredi talebinde bulunur. Banka müşterinin demografik, geçmiş ve şimdiki finansal verileri analiz edilerek kredi itibarını

değerlendirir. Bu işlem bir makine öğrenmesi üzerinden müşteri verileri girilerek gerçekleştirilir. Müşteri uygun değil ise süreç sonlandırılır. Eğer uygun ise bir sonraki aşama olan sosyal medya analizi gerçekleştirilir. Sosyal medya analizi müşterinin beyan ettiği meslek ve işyeri bilgisi üzerinden yapılır. Sosyal medya analizi olumlu olan müşteri bir sonraki aşamaya geçer. Olumsuz ise süreç sonlandırılır. Son olarak kredi tahsis onayının ardından müşterinin kredi limiti, ilgili risk profili ve gerekli teminat bilgileri blokzincir tabanlı ağa yüklenir. Diğer finans kurumları daha sonra müşterinin TCKN'sini kullanarak sistemi sorgulayarak bu verilere erişebilir. Tüm bu iş akışı Şekil 6.2'de gösterilmektedir.



Şekil 6.2 Sistemin genel iş akışı

6.3 Uygulama

Araştırmanın üç fazına ait uygulama sonuçları şu şekildedir;

Faz I: Şekil 6.3, Kredi Tahsis Modülü'nün kullanıcı arayüzü tasarımını sunmaktadır. Müşteri kredibilitesini değerlendirmek için 14 veri alanı kullanılmaktadır. Bu alanlardan Maaş, Yaş ve İkametgah Yılı nicel verileri temsil ederken, geri kalanlar kategorik veriler olarak işlenmektedir. Tüm verilerin girilmesinin ardından, "Skoru Getir" düğmesine basılmasıyla uygulama arka planında çalışan makine öğrenmesi modeli tetiklenir ve sonuç üretilir.

Kredi Skor Tahmini

Maaş :	25000
Medeni Durum :	Evli
Cinsiyet :	Erkek
Sektör :	Kamu Kurumu
Eğitim Durumu :	Fakülte/Lisans
Mesleği :	Bankacı
Yaş :	40
Askerlik Durumu :	Yapıldı
Talebinin Altında Teklif Var mı? :	Hayır
Son 1 Yıl İçinde Ret Edilmiş Tahsis Var mı? :	Hayır
Kanuni Takibi Sorunu Var mı? :	Hayır
İcra Takibine Düşmüş mü? :	Hayır
İkametgah Yılı :	10
Otundugu Ev Durumu :	Kendine Ait

[Skoru Getir](#)

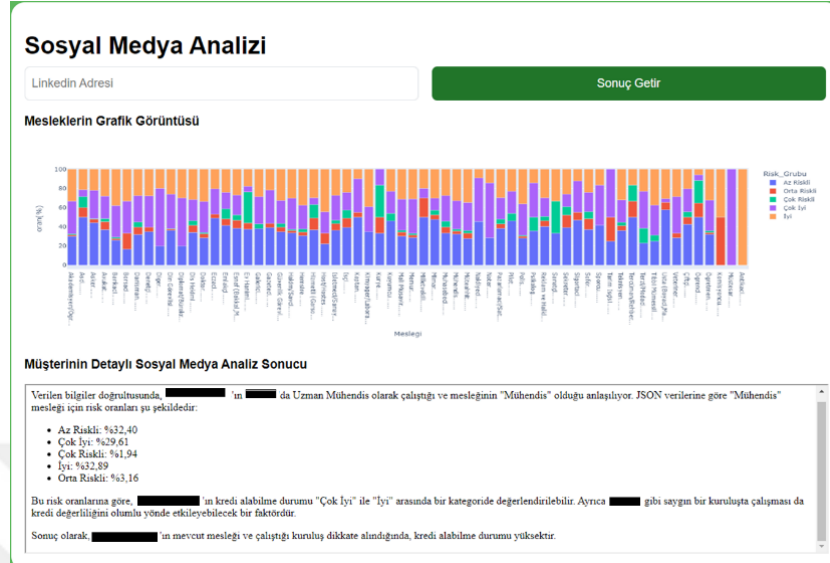
Müşterinin Kredi Skoru A
Müşterinin Kredi Durumu Çok İyi
Skor Açıklaması: En iyi kredi kartı segmentine sahip bir kart verilebilir, ev ve araç kredisi kullanılır.

Şekil 6.3 Kredi skor tahmini

Faz II: Bu kısımda banka müşterilerinin kredi değerlendirme süreçlerini iyileştirmeyi amaçlayan, Selenium tabanlı web otomasyonu ve üretken yapay zeka entegrasyonlu bir sistem sunulmaktadır. Sistem, ilk olarak müşterinin kredi puanı analizini takiben, Selenium aracılığıyla LinkedIn profilinden meslek ve işyeri bilgilerini elde etmektedir. Elde edilen bu veriler, daha sonra üretken yapay zeka modeline giriş olarak sunulmaktadır. Model, müşteriye kredi verilip verilmeyeceği konusunda ön değerlendirme yapmakta ve teyit edilen meslek bilgisi için ek bir doğrulama adımı sağlamaktadır. Şekil 6.4'te gösterilen kullanıcı arayüzü üzerinden sunulan sistem çıktısı, müşterinin meslek ve işyeri bilgilerine dayalı bir yorum içermektedir. Üretken yapay zeka modeli, bu yorumu oluştururken, JSON formatında sağlanan meslek-risk grubu ilişkisini de dikkate almaktadır. Sonuç olarak, sistemin sunduğu bu değerlendirme ve yorumlar, banka personelinin nihai kredi verme kararını destekleyici bir kaynak olarak işlev görmektedir.

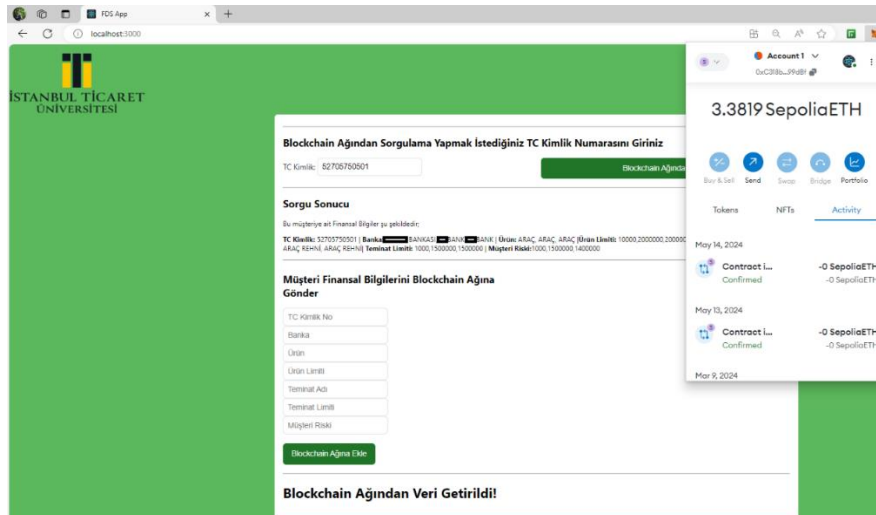
Faz III: Araştırmanın bu kısmında geliştirilen uygulamanın kullanıcı arayüzü Şekil 6.5'te sunulmaktadır. Ethereum blokzinciri platformu üzerinde tasarlanan ve Metamask cüzdan entegrasyonu ile işlem gerçekleştirme imkanı sunan sistem, TCKN bazlı sorgulama işlevselliğine sahiptir. Kredi tahsisi gerçekleşen

müşterilere ait TCKN, banka bilgileri, ürün bilgileri, teminat bilgileri ve risk tutarı blokzincire kaydedilmektedir.



Şekil 6.4 Sosyal medya analizi

Bu kayıt işlemi, Ethereum ağına ödenmesi gereken bir işlem ücreti gerektirmektedir. Blokzincire kaydedilen finansal verilere, farklı bir banka tarafından müşteriye ait TCKN kullanılarak erişilebilir ve sorgulanabilir. Sorgulama işlemi herhangi bir ek işlem ücreti gerektirmemektedir.



Şekil 6.5 Blokzincir tabanlı finansal veri paylaşımı

7. SONUÇ VE ÖNERİLER

Bu bölümde araştırmanın sonuçlarını sunmanın yanı sıra gelecekteki çalışmalara katkı sağlayacak önerilere yer verilmektedir.

Araştırmanın ilk aşamasında banka müşterilerine ait kredi skor tahmini yapmak için sentetik veri seti üzerinden çalışma gerçekleştirilmiştir. Bu veri seti uzman görüşü ile hazırlanmış finansal ve demografik bilgiler içeren 16151 satır veriden oluşmaktadır. Araştırma kapsamında on farklı makine öğrenmesi yöntemleri uygulanmıştır. Tüm bu çalışmaların sonucunda LightGBM algoritmasının %93.82 doğruluk ile en yüksek başarı oranını elde ettiği görülmüştür. Bu bulgu, LightGBM'nin kategorik verileri etkili bir şekilde işleyebilme ve aşırı öğrenmeyi önleme yetenekleri sayesinde, kredi riski değerlendirmesi gibi karmaşık bir görev için uygun bir araç olduğunu göstermektedir.

Kredi tahsis süreçlerinde demografik ve finansal verilerin yanında sosyal medya verileri de önem arz etmektedir. Kurumsal ve Kobi firmaların sosyal medya analizleri daha kolay yapılabilirken bireysel müşterilerin sosyal medya analizleri hala istenilen sonuçları ortaya koyamamaktadır. Bu çalışmada müşterinin beyan ettiği meslek ve işyeri bilgisi bir sosyal medya platformu üzerinden alınarak üretken yapay zeka ile yorumlanması sağlanmıştır. Kredi tahsis değerlendirmelerinde üretken yapay zeka modellerinden faydalanabileceği ortaya konulmuştur.

Araştırmanın kapsamında bir diğer önemli konu ise hem geleneksel hem de blokzincir tabanlı finansal veri paylaşım modellerinin incelenmesi olmuştur. Blokzincir tabanlı modeller, teknoloji ve güvenlik gibi özellikler açısından üstünken, regülasyonlar açısından geleneksel yöntemler daha güçlü görülmektedir. Çalışma sonuçlarına dayalı bir karşılaştırma tablosu Çizelge 7.1'de görülebilir. Regülasyon konuları devam ederken, blokzincir tabanlı sistemlerin potansiyel faydaları yadsınamaz. Blokzincir tabanlı sistemler verimli, bankalar arasında gelişmiş işbirliği, güvenli ve şeffaf bir çerçeve içerisinde gelişmiş risk yönetimi sağlayabilmektedir. Blokzincir teknolojisi olgunlaştıkça ve

düzenleyici ortamlar geliştikçe, bu sistemler devrim yaratma potansiyeline sahip olup, bankacılık sektöründe güvenli ve verimli finansal veri paylaşımının önünü açabilmesi muhtemeldir.

Çizelge 7.1 Geleneksel ve blokzincir tabanlı finansal veri paylaşım sistemlerinin değerlendirilmesi

Özellik	Geleneksel	Blokzincir tabanlı
Merkezileşme	Merkezileştirilmiş	Merkezi olmayan
Şeffaflık	Gerçek zamanlı veri paylaşımı sağlar	Gerçek zamanlı çoklu veri paylaşımı sağlar
Güvenlik	Veriler şirket çalışanları tarafından manipüle edilebilir	Blokzincir'in değişmezliği, verilerin güncellenmesine karşı korunma sağlar
Risk değerlendirmesi	Risk yalnızca bankanın kendi verileriyle hesaplanır	Diğer bankalardan alınan verilerle kapsamlı risk değerlendirmesi yapılır
Yeterlik	Bankalar arasında tekrarlanan doğrulama süreçleri	Tekrarlanan süreçlerin ortadan kaldırılması
Düzenleme	Hükümetler merkezi sisteme izin veriyor	Düzenlemenin iyileştirilmesi gerekiyor

7.1 Öneriler

İlerleyen araştırmalarda bu çalışmada kullanılan veri seti üzerinden daha farklı makine öğrenmesi algoritmaları denenerek daha yüksek başarı oranına ait sonuçlar elde edilebilir. Bunun yanında üretken yapay zeka modelleri ile daha gelişmiş kredi tahsis süreçleri oluşturulabilir. Bu araştırmalar sadece kredi tahsis çalışmalarında değil farklı bankacılık alanlarında da ortaya koyulabilir.

Dünya üzerindeki küresel verilerin katlanarak büyümesi, paydaşlar arasında güvenli veri depolama ve verimli paylaşımı gerektirmektedir. Blokzincir teknolojisi, güvenli ve şeffaf veri alışverişini kolaylaştırarak bu alanda öncü olarak ortaya çıkmaktadır. Bu özelliğin önümüzdeki yıllarda finans sektöründe daha fazla benimsenmesine yol açması muhtemeldir. Ancak düzenleme ve uyumluluk engelleri devam etmektedir. Bu zorlukların üstesinden gelmek, finansal kurumlar için çok sayıda kullanım senaryosunun kilidini açacaktır. Bankacılıkta finansal verilerin paylaşımı haricinde kıymetli evraklar da

blokzincir tabanlı sistemler ile paylaşılabilir. Örneğin bu kıymetli evrakların başında gelen teminat mektupları NFT'lerle tokenize edilip kopyalanmayı ve sahteciliği önlenerek bankaların ilgili riskleri daha etkili bir şekilde yönetmesine olanak tanıyabilir.

Bu araştırmanın gelecekteki yönelimleri, kredi skorlama süreçlerinde daha kapsamlı ve heterojen veri kaynaklarının kullanımını ve çeşitli modellerin optimizasyonunu hedeflemektedir. Mevcut sosyal medya verilerine ek olarak, bireylerin dijital ayak izleri, e-ticaret davranışları, mobil uygulama kullanım örüntüleri ve diğer dijital etkileşimlerinden elde edilecek verilerin kredi risk değerlendirme yöntemlerine entegrasyonu planlanmaktadır. Bu zenginleştirilmiş veri setlerinin, daha hassas ve kişiselleştirilmiş kredi skorlarının oluşturulmasına katkıda bulunması öngörülmektedir. Ayrıca, bu tür verilerin kullanımı, özellikle yeni müşteri segmentlerinin kredi değerliliğinin değerlendirilmesinde etkili olabilir ve finansal kapsayıcılığı artırabilir.

Gelecekteki bir diğer kritik araştırma alanı, blokzincir teknolojisinin kredi süreçlerine entegrasyonunun derinlemesine incelenmesi olacaktır. Blokzincir tabanlı finansal veri paylaşım modellerinin geliştirilmesi, bankalar arasında güvenli ve şeffaf bir veri alışverişi sağlamanın yanı sıra, bu süreçlerin otomasyonu ve hızlandırılması için de önemli fırsatlar sunmaktadır. Bununla birlikte, bu sistemlerin düzenleyici çerçevelerle uyumlu hale getirilmesi için gerekli yasal ve teknik altyapının incelenmesi de büyük önem taşımaktadır. Blokzincir teknolojisinin yaygınlaşması ve olgunlaşmasıyla, bu alanda gerçekleştirilecek araştırmalar, bankacılık sektöründe daha güvenli, verimli ve hızlı kredi süreçlerinin geliştirilmesine yönelik önemli katkılar sağlama potansiyeline sahiptir.

KAYNAKLAR

- Bowyer, K. W., Chawla, N. V., Hall, L. O., Kegelmeyer, W. P., 2011. SMOTE: Synthetic Minority Over-sampling Technique. CoRR abs.
- Breiman, L., 2001. Random forests. Machine learning,, 5-32.
- Buterin, V., 2014. A next-generation smart contract and decentralized application platform. whitepaper, 3(37), 2-1.
- Byström, H., 2019. Blockchains, Real-time Accounting and the Future of Credit Risk Modeling. Ledger Journal, 4, 40-47.
- Chakraborty, S., Aich, S., Seong, S. J., Kim, H. C., 2019. A Blockchain based Credit Analysis Framework for Efficient Financial Systems. PyeongChang, Korea (South): 2019 21st International Conference on Advanced Communication Technology (ICACT).
- Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, (785-794).
- Croman, K., Decker, C., Eyal, I., Gencer, A. E., Juels, A., Ahmed Kosba, A. M., Wattenhofer, R., 2016. On Scaling Decentralized Blockchains: (A Position Paper). International conference on financial cryptography and data security (106-125). Springer Berlin Heidelberg.
- Dastile, X., Celik, T., Potsane, M., 2020. Statistical and machine learning models in credit scoring: A systematic literature survey. Applied Soft Computing, 91.
- Deepmind., 2024. Gemini. <https://deepmind.google/technologies/gemini/> adresinden alındı
- Diaconescu, P., Neagoe, V.-E., 2020. Credit Scoring Using Deep Learning Driven by Optimization Algorithms. 2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), 1-6.
- Dietterich, T. G., 2000. Ensemble methods in machine learning. International workshop on multiple classifier systems (1-15). Berlin: Springer Berlin Heidelberg.
- Dorogush, A. V., 2018. CatBoost: gradient boosting with categorical features support. arXiv preprint.
- Eliuz, E., 2009. Tüketici Kredilerinde Risk Yönetimi ve Bir Skorkart Modeli Önerisi. Yüksek Lisans Tezi.

- Emekter, R., Tu, Y., Jirasakuldech, B., Lu., M., 2015. Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending. *Applied Economics*, 47(1), 54-70.
- Express, A., 2022. AMEX: Credit Score Model. 03 31, 2023 tarihinde <https://www.kaggle.com/code/gopidurgaprasad/amex-credit-score-model> adresinden alındı
- Firmansyah, H. B., Ramdani, A. L., 2018. The Role of Islamic Financial Technology (FinTech) Start-Up in Improving Financial Inclusion in Indonesia Case: Angsur. 3rd International Conference of Integrated Intellectual Community (ICONIC).
- Goel, D. K., 2021. Bank Customer's Credit Score Prediction Using Feature Selection and Data Mining Algorithm. *Progress in Advanced Computing and Intelligent Engineering. Advances in Intelligent Systems and Computing*, 1299, 889-897.
- Goodfellow, I. B., 2016. *Deep Learning*. MIT press.
- Guangming Guo, F. Z., 2016. From Footprint to Evidence: An Exploratory Study of Mining Social Data for Credit Scoring. *CM Trans. Web*, 1 - 38.
- Gunnarsson, B. R., Broucke, S. V., Baesens, B., Óskarsdóttir, M., Lemahieu, W., 2021. Deep learning for credit scoring: Do or don't? *European Journal of Operational Research*, 295(1), 292-305.
- Hart, T. C., 1967. Nearest neighbor pattern classification. 12(1).
- Hayashi, Y., 2016. Application of a rule extraction algorithm family based on the Re-RX algorithm to financial credit risk assessment from a Pareto optimal perspective. *Operations Research Perspectives*, 3(2214-7160), 32-42.
- Hayashi, Y., 2022. Emerging Trends in Deep Learning for Credit Scoring: A Review. *Electronics* 2022, 11(19).
- Hochreiter, S., 1997. Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- Hofmann, P. D., 1994. Statlog (German Credit Data) Data Set. Hamburg: Institut für Statistik und "Okonometrie.
- Hoonsopon, W. V., 2019. Blockchain characteristics and consensus in modern business processes. *Journal of Industrial Information Integration*, 13, 32-39.
- Hosmer Jr, D. W., 2013. *Applied logistic regression*. New York: John Wiley & Sons.

- Hu, S., Hou, L., Chen, G., Weng, J., Li, J., 2018. Reputation-based distributed knowledge sharing system in blockchain. Proc. 15th EAI Int. Conf. Mobile Ubiquitous Syst. Comput. Netw. Services, 476-481.
- J. Lohokare, R. D., 2017. Automated data collection for credit score calculation based on financial transactions and social media. 2017 International Conference on Emerging Trends & Innovation in ICT (ICEI), 134-138.
- Jacob Eberhardt, S. T., 2017. On or Off the Blockchain? Insights on Off-Chaining Computation and Data. European Conference on Service-Oriented and Cloud Computing (3-15). Springer International Publishing.
- James, G. M., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning. New York: Springer.
- Kaggle., 2023. Kaggle. 04 12, 2023 tarihinde <https://www.kaggle.com/datasets/parisrohan/credit-score-classification> adresinden alındı
- Karadag, B. (2024, 07 01). Github.com. <https://github.com/BulutKaradag/Credit-Score-with-Machine-Learning> adresinden alındı
- Karadag, B. (2024, 07 14). Github.com. <https://github.com/BulutKaradag/Social-Media-Analysis> adresinden alındı
- Karadag, B. (2024, 07 20). Github.com. <https://github.com/BulutKaradag/Blockchain-based-KYC-Model/blob/2fa00d0461189d98d94f11b3d9907ac7028833d0/KYC%20Smart%20Contract.sol> adresinden alındı
- Karadag, B. (2024, 5 19). <https://github.com/>. 6 20, 2024 tarihinde <https://github.com/BulutKaradag/Credit-Score-with-Machine-Learning/blob/main/CustomerCreditDataSet.csv> adresinden alındı
- Karayılan, H., 2019. Blockchain and its applications for financial technology solutions. İstanbul: Yüksek Öğretim Dergisi.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... Liu, T.-Y., 2017. Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems, s. 3146-3154.
- Kredi Kayıt Bürosu. (2024, 08 08). KKB. <https://www.kkb.com.tr/> adresinden alındı
- Kulkarni, S. V., Dhage, S. N., 2019. Advanced credit score calculation using social media and machine learning. Journal of Intelligent & Fuzzy Systems, 36(3), 2373-2380.

- Kumar, A., Shanthi, D., Bhattacharya, P., 2021. Credit Score Prediction System using Deep Learning and K-Means Algorithms. Hyderabad, India: IOP Publishing.
- Kuppili, V., Tripathi, D., Edla, D. R., 2022. Credit score classification using spiking extreme learning machine. *Computational Intelligence*, 36(2), 402-426.
- Laarabi, M., Maach, A., 2020. Understanding risk assessment in the context of fractional ownership using Ethereum smart contract. *Adv. Sci. Technol. Eng. Syst. J.*, 5(5), 1028-1035.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *nature*, 521(7553), 436-444.
- Lemieux, V. L., 2016. Trusting records: is Blockchain technology the answer? *Records Management Journal*, 26(2), 110-139.
- Luo, C., Wu, D., Wu, D., 2017. A deep learning approach for credit scoring using credit default swaps. *Engineering Applications of Artificial Intelligence*, 65, 465-470.
- Murphy, K. P., 2012. *Machine Learning A Probabilistic Perspective*. MIT press.
- Nakamoto, S., 2008. Bitcoin: A Peer-to-Peer Electronic Cash System. 04 18, 2023 tarihinde <https://bitcoin.org/bitcoin.pdf> adresinden alındı
- Óskarsdóttir, M., Bravo, C., Sarraute, C., Vanthienen, J., Baesens, B., 2019. The value of big data for credit scoring: Enhancing financial inclusion using mobile phone data and social network analytics. *Applied Soft Computing*, 74, 26-39.
- Patel, S. B., Bhattacharya, P., Tanwar, S., Kumar, N., 2021. KiRTi: A Blockchain-Based Credit Recommender System for Financial Institutions. *IEEE Transactions on Network Science and Engineering*, 8(2), 1044-1054.
- Perera, S., Nanayakkara, S., Rodrigo, M. N., Senaratne, S., Weinand, R., 2020. Blockchain Technology: Is it Hype or Real in the Construction Industry. *Journal of Industrial Information Integration*, 17, 1-20.
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., Gulin, A., 2018. CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems*.
- Roman, D., Stefano, G., 2016. Towards a Reference Architecture for Trusted Data Marketplaces: The Credit Scoring Perspective. Vienna, Austria: 2nd International Conference on Open and Big Data (OBD).
- Sabato, G., 2010. *Credit Risk Scoring Models*.

- Sherman, A. T., Javani, F., Zhang, H., Golaszewski, E., 2019. On the Origins and Variations of Blockchain Technologies. in IEEE Security & Privacy, 17(1), 72-77.
- Singh, D. K., Goel, N., 2020. Analysing Data Mining Techniques on Bank Customers for Credit Score. Noida, India: 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO).
- Stefania, A., Vamosy, D. F., 2019. Predicting consumer default: A deep learning approach. (w26165).
- Szabo, N., 1997. Formalizing and Securing Relationships on Public Networks. First Monday, 2(9). doi:10.5210/fm.v2i9.548
- Team, S. P., 2019. Selenium. Selenium WebDriver Documentation: https://www.selenium.dev/documentation/webdriver_getting_started/ adresinden alındı
- Tsai, C.-F., Wu, J.-W., 2008. Using neural network ensembles for bankruptcy prediction and credit scoring. Expert Systems with Applications, 34(4), 2639-2649.
- Turan, M. E., 2019. Doğrusal ve Lojistik Regresyon Yöntemleri ile Kredi Skorlama Sistemleri ve Bir Uygulama. Yüksek Lisans Tezi.
- Wang, Y., Zhang, Y., Lu, Y., Yu, X., 2020. A Comparative Assessment of Credit Risk Model Based on Machine Learning - a case study of bank loan data. Procedia Computer Science, 174(1877-0509), 141-149.
- Wegrzyn, K. E., Wang, E. (2021, 03). Foley. <https://bit.ly/3BbLQTF> adresinden alındı
- Wolpert, D. H., 1992. Stacked generalization. Neural networks, 241-259.
- Xu, X., Weber, I., Staples, M., Zhu, L., Bosch, J., Bass, L., Rimba, P., 2017. A taxonomy of blockchain-based systems for architecture design. 2017 IEEE international conference on software architecture (ICSA) (243-252). IEEE.
- Yang, F., Qiao, Y., Huang, C., Wang, S., Wang, X., 2021. An Automatic Credit Scoring Strategy (ACSS) using memetic evolutionary algorithm and neural architecture search. Applied Soft Computing, 113(A).
- Yang, F., Qiao, Y., Qi, Y., Bo, J., Wang, X., 2022. BACS: blockchain and AutoML-based technology for efficient credit scoring classification. Annals of Operations Research.
- Zhang, Z., Niu, K., Liu, Y., 2020. A Deep Learning Based Online Credit Scoring Model for P2P Lending. IEEE Access, 8, 177307-177317.

Zheng, Z., Xie, S., Dai, H., Chen, X., Wang, H., 2017. An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. 2017 IEEE International Congress on Big Data (BigData Congress) (557-564). Honolulu: IEEE.

Zhu, B., Yang, W., Wang, H., Yuan, Y., 2018. A hybrid deep learning model for consumer credit scoring. Chengdu, China: International Conference on Artificial Intelligence and Big Data (ICAIBD).



ÖZGEÇMİŞ

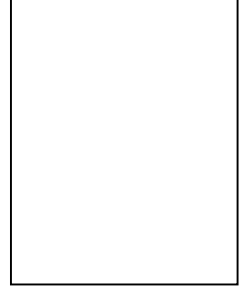
Adı Soyadı : Bulut Karadağ

Doğum Yeri ve Yılı :

Medeni Hali :

Yabancı Dili : İngilizce

E-posta :



Eğitim Durumu

Lise :

Lisans :

Yüksek Lisans :

Mesleki Deneyim

Piworks	2010-2013
Kuvey Türk Katılım Bankası	2013-2016
Vakıf Katılım Bankası	2016-2022
Neova Katılım Sigorta	2022- Devam ediyor

Yayınları

Gökçen, A., Karadağ, B., Riva, C., Boyacı, A., 2021. Artificial intelligence-based COVID-19 detection using cough records.

Karadağ, B., Akbulut, A., Zaim, A. H., 2023. Cyberbullying detection through deep learning: A case study of Turkish celebrities on Twitter. In Web Intelligence 21(1), 61-70. IOS Press.

- Karadağ, B., Akbulut, A., Zaim, A. H., 2022. A review on blockchain applications in fintech ecosystem. In 2022 International Conference on Advanced Creative Networks and Intelligent Systems (ICACNIS) (1-5). IEEE.
- Karadağ, B., Arıkan, E., 2021. Choosing the Appropriate Branch for Participation Banks Through Machine Learning. In Proceedings of 6th International Conference on Harmony Search, Soft Computing and Applications: ICHSA 2020, Istanbul, 269-279. Springer Singapore.
- Karadağ, B., Bölükbaş, O., Ünal, M. A., 2020. Makine Öğrenmesi İle Bireysel Müşteriler İçin Finansman Ürün Önerilmesi. Academic Perspective Procedia, 3(1), 438-444.
- Karadağ, B., Erbaşı, Z., Karagöz, K. T., Özmen, S., 2022. Digital Academy for Mentally Disabled. Orclever Proceedings of Research and Development, 1(1), 58-74.
- Karadağ, B., Zaim, A. H., Akbulut, A., 2024. Blockchain Based KYC Model for Credit Allocation in Banking. IEEE Access.