MAKİNE ÖĞRENMESİ İLE TALEP TAHMİNİ VE ENVANTER YÖNETİMİ

Amine BAYAR SERBEST

T.C.
BURSA ULUDAĞ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MAKİNE ÖĞRENMESİ İLE TALEP TAHMİNİ VE ENVANTER YÖNETİMİ

Amine BAYAR SERBEST
0009-0008-4673-3185

Doç. Dr. Aslı AKSOY
(Danışman)

YÜKSEK LİSANS
ENDÜSTRİ MÜHENDİSLİĞİ ANABİLİM DALI

BURSA – 2024
Her Hakkı Saklıdır

ÖZET

Yüksek Lisans Tezi

MAKİNE ÖĞRENMESİ İLE TALEP TAHMİNİ VE ENVANTER YÖNETİMİ

Amine BAYAR SERBEST

Bursa Uludağ Üniversitesi
Fen Bilimleri Enstitüsü
Endüstri Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Aslı AKSOY

Stok kontrolü her işletmenin göz önünde bulundurması gereken önemli bir husustur. Çok fazla stok, satış süresinin artmasına ve verimsizleşmesine, hatta kayıplara neden olabilir. Talebi ve envanteri tahmin etmek, fazla stoğu ve eksik stoğu en aza indirgemek için ve kayıp riskini azaltmak için çok gereklidir. İyi tahmin edilmiş talebin doğru şekilde yönetilmesi, bir dizi avantaj sağlar. Bunlar arasında stokta kalmama riskinin azaltılması, daha iyi müşteri deneyimi sağlanması, envanterin daha etkili bir şekilde planlanması ve daha az ürün iptalinin gerçekleşmesiyle maliyetlerin düşürülmesi gibi önemli faktörler bulunmaktadır. Yapay zekâ (YZ) çağında, makine öğrenmesi (MÖ) teknolojisinin yükselişi, birçok mühendislik uygulamasının karşılaştığı zorluklara hızlı ve etkili çözümler sunabilen önemli bir alandır. Makine öğrenimi, veri analizi, desen tanıma ve öngörülebilirlik gibi konularda çığır açan gelişmeler sunmaktadır. Ancak, bu alandaki ilerlemelere rağmen, hala birçok varsayım ve tartışma bulunmaktadır. Tekstil sektöründe makine öğrenmesi yöntemleri, geleneksel üretim süreçlerine yenilikçi ve verimli çözümler sunma potansiyeli taşımaktadır. Bu alandaki gelişmeler, işletmelerin üretim, stok yönetimi ve talep tahmini gibi kritik süreçlerde daha etkili ve optimize edilmiş kararlar almasına olanak tanımaktadır. Tekstil sektöründe bu konuda sınırlı literatür bulunması, bu çalışmanın işletmelere yol gösterici bir kaynak olması beklentisini artırmaktadır. Tekstil sektöründe, talep tahmini işlemi, üretim planlaması ve envanter yönetimi açısından kritik bir rol oynamaktadır. Bu çalışmada, makine öğrenmesi algoritmaları kullanılarak bir tekstil işletmesindeki gerçek veri seti üzerinde talep tahmini modeli geliştirilmiştir. Bu model, işletmenin gelecekteki taleplerini daha doğru bir şekilde tahmin etmeyi amaçlamaktadır.

Anahtar Kelimeler: Makine öğrenmesi, Talep tahmini, Tekstil Sektörü

2024, vii + 37 sayfa.

ABSTRACT

MSc Thesis

DEMAND FORECASTING AND INVENTORY MANAGEMENT WITH MACHINE LEARNING

Amine BAYAR SERBEST

Bursa Uludağ University
Graduate School of Natural and Applied Sciences
Department of Industrial Engineering

Supervisor: Doç. Dr. Aslı AKSOY

Inventory control is an important aspect that every business should consider. Excessive inventory can lead to increased sales cycle times, inefficiency, and even losses. Predicting demand and managing inventory are crucial to minimizing excess and shortage of stock and reducing the risk of loss. Properly managing forecasted demand provides several advantages, including reducing the risk of stockouts, providing a better customer experience, more effectively planning inventory, and reducing costs by minimizing product cancellations. In the era of Artificial Intelligence (AI), the rise of Machine Learning (ML) technology is a significant field that can provide fast and effective solutions to many challenges faced by engineering applications. Machine learning offers groundbreaking developments in areas such as data analysis, pattern recognition, and predictability. However, despite advancements in this field, there are still many assumptions and debates. Machine learning methods in the textile sector have the potential to offer innovative and efficient solutions to traditional production processes. These advancements allow businesses to make more effective and optimized decisions in critical processes such as production, inventory management, and demand forecasting. The limited literature in the textile sector on this topic increases the expectation that this study will be a guiding resource for businesses. Demand forecasting plays a critical role in production planning and inventory management in the textile sector. In this study, a demand forecasting model has been developed using machine learning algorithms on a real dataset from a textile business. This model aims to more accurately predict the company's future demands.

Key words: Machine learning, Demand forecasting, Textile Sector
2024, vii + 37 pages.

ÖNSÖZ ve/veya TEŞEKKÜR

Bu tez çalışmamın hazırlık sürecinde, rehberliği, yönlendirmesi ve samimi desteğiyle beni cesaretlendiren kıymetli danışmanım Doç. Dr. Aslı AKSOY ‘a içten teşekkürlerimi iletiyorum.
Kariyerimde ve bu akademik yolculuğumda beni sürekli teşvik eden, değerli önerileriyle aydınlatan ve her zaman desteklerini hissettiğim yöneticilerime ve iş arkadaşlarıma minnettarlığımı ifade ederim.
Hayatımın en önemli destekçileri olan aileme, her zaman yanımda oldukları için ve desteklerini esirgemedikleri için teşekkür ederim.

Amine BAYAR SERBEST
22/01/2024

İÇİNDEKİLER

ÖZET vi
ABSTRACT vii
ÖNSÖZ ve/veya TEŞEKKÜR viii
ŞEKİLLER DİZİNİ x
ÇİZELGELER DİZİNİ xi
1. GİRİŞ 1
2. KAYNAK ARAŞTIRMASI 4
3. MATERYAL ve YÖNTEM 9
3.1. Makine öğrenmesi 9
3.1.1. Makine Öğrenmesi Algoritmaları 9
3.2. Hata Ölçütleri 15
3.3. Hiperparametre Ayarlama 16
4. BULGULAR ve TARTIŞMA 18
4.1. Akış Haritası 18
4.1.1 Veri Seti Özellikleri 21
4.1.2. Veri Önişleme 22
4.1. Modelin Oluşturulması 25
5. SONUÇ 32
KAYNAKLAR 34
ÖZGEÇMİŞ 37

ŞEKİLLER DİZİNİ

Sayfa

Şekil 3.1 Rastgele orman algoritmasının işlem adımları 12
Şekil 4.1. Makine öğrenmesi uygulama haritası 19
Şekil 4.2. Rumuz, denye ve renk değişkenlerinin değişimi 22
Şekil 4.3. Miktar etiket değerinin dağılımı 23
Şekil 4.4. İlk oluşturulan korelasyon matrisi 24
Şekil 4.5. Son oluşturulan korelasyon matrisi 24
Şekil 4.6. Hazır hale gelen modelin test veri seti üzerindeki tahmin sonuçları 28
Şekil 4.7. CatBoost regresyon modelinin tahminlerinin dağılım grafiği 29
Şekil 4.8. Özellik önem sırası grafiği 30
Şekil 4.9. Makine öğrenmesi modelinin kullanılması 30

ÇİZELGELER DİZİNİ

		Sayfa
Çizelge 4.1.	Veri seti özellikleri……	21
Çizelge 4.2.	Makine öğrenmesi algoritmalarının başarı metriklerine etkisi	26
Çizelge 4.3.	Catboost algoritması için K katlamalı çapraz doğrulama model performansı	26
Çizelge 4.4.	CatBoost algoritması ve Ligthgbm algoritması için kullanılan parametreler …….	27

1. GİRİŞ

Değişkenlerin talep üzerindeki etkileri ve müşterilerin sürekli değişen ihtiyaçları göz önüne alındığında müşteri talebini tahmin etmek zordur. Operasyonel kararlar almanın zor olmasının yanında uygulama sırasında yanlış tahmin yapılması durumunda müşteri ve satış kaybının olması markanın prestijinin de kaybına sebep olabilir. Şirketler satış kayıplarını önlemek için kısa vadeli planlar yaparlar ve bu planın en önemli girdilerinden biri de talep tahminidir. Ancak, talebi etkileyen birçok değişkenin sektöre özgü farklılıklar göstermesi ve tedarikçi dışındaki etmenlerle şekillenmesi nedeniyle tahmin yapmak kolay değildir. Gelecek dönemlere ait satış miktarlarını ve ürün çeşitlerini doğru tahmin etmek firmalar için ciddi bir yatırımdır. Tahmin rakamlarına göre stok ve işçilik maliyeti iyileştirilebilir. Eksik veriler, özel günlerde dalgalanan talep, ekonomik kriz, tahmini etkileyen farklı değişkenler vb. talep tahmininin zorlukları arasında sayılabilir. İşletmelerin talep tahminlerini doğru yapmaları, stok yönetimini etkili bir şekilde planlamalarına ve operasyonel verimliliklerini artırmalarına olanak tanır. Bu doğru tahminler, müşteri memnuniyetini artırmanın yanı sıra, tedarik zincirindeki her bir halka için olumlu sonuçlar doğurabilir. Bu noktada, geleneksel yöntemlerin ötesine geçen ve büyük veri analizi ile öne çıkan bir yaklaşım, makine öğrenmesi, giderek daha fazla önem kazanmaktadır. Makine öğrenmesi, işletmelerin geçmiş satış verilerini inceleyerek ilişkileri tespit etmelerine, bu bilgileri kullanarak gelecekteki talebi tahmin etmelerine imkân tanır. Gelişmiş algoritmalar, büyük veri kümeleme teknikleri ve sofistike matematiksel modeller kullanarak, makine öğrenimi sayesinde daha önce elde edilemeyen derinlemesine anlayışlar geliştirebilir.

Talep tahmini için makine öğrenmesi, sadece mevcut verilere dayalı basit öngörülerden öteye geçer. Bu yöntem, dinamik ve değişken pazar koşullarını anlama yeteneği sayesinde, işletmelerin hızla değişen taleplerine daha duyarlı ve esnek bir şekilde tepki vermelerini sağlar.

Makine öğrenmesinin talep tahmini sürecine entegre edilmesi, işletmelerin stok seviyelerini iyileştirmelerine, müşteri taleplerine daha hızlı ve doğru bir şekilde yanıt

vermelerine olanak tanır. Bu da hem operasyonel maliyetleri düşürür hem de müşteri memnuniyetini artırır.

Tekstil ürünleri üzerinde karar verme süreçleri etrafında bir yapı oluşturmak, gelecekte oluşabilecek talebi doğru bir şekilde tahmin etmek için uygun bir tahmin modelinin oluşturulmasını gerektirir (Lorente-Leyva vd., 2021). Tekstil üretiminde, kolay bir işlem veya standart bir ürünün üretimi dahil çok miktarda veri üretimini ve depolanmasını gerektirir. Günümüzde firmalar için rekabetin hızla artması, pazarda güçlü kalmayı ve sürekliliği sürdürmeyi giderek zorlaştırmaktadır. Tekstil işletmeleri, rakiplerine üstünlük sağlamak ve sürekli gelişimi benimseyerek yeni müşteriler kazanmak için çeşitli yöntemlere başvurmalıdır. Türkiye'nin önde gelen istihdam ve ihracat sektörlerinden biri olan tekstil sektörü, yüksek rekabet oranına sahiptir. Yeni teknolojiler, firmaların rekabet avantajını elde etmelerinde kilit bir rol oynamaktadır. Makine öğrenimi yaklaşımları, bu sektörde benimsenen yeni teknolojiler arasında yer almaktadır. Başarıya ulaşabilmek için tekstil firmalarının gelecekle ilgili öngörülerde bulunma yeteneği de oldukça önemlidir. Gelecek hakkında doğru ve güvenilir bilgilere erişim, firmalara büyük fayda sağlayacaktır. Tekstil firmaları, bu teknolojileri kullanarak verileri doğru bir şekilde analiz edebilir, anlamlı yapıları çıkarabilir ve geleceğe yönelik planlamalar yapabilir, stratejik kararlar alabilirler (Ersöz ve Çınar, 2021).

Tekstil sektöründe müşteri beklentilerinin hızlı ve etkili şekilde karşılanma zorunluluğu, rekabet gücü gibi birçok kısıt, özellikle bu sektörde talep tahmin doğruluğunu kritik hale getirmektedir. Ayrıca, mevsimsel satışların hava şartlarına duyarlı olması, üretilen ürünlerin çeşitliliği, ürün tasarımlarının trendleri yakalaması gibi faktörler, talep tahminini zorlaştırmaktadır. Bu nedenle, sektörde faaliyet gösteren şirketlerin talep tahmin süreçlerini yönetirken, hızlı değişen talepleri, sezonluk değişimleri ve müşteri beklentilerini dikkate alarak esnek ve yenilikçi çözümler geliştirmeleri gerekmektedir (Öztürk, 2020).

Satışları tahmin etmek, günümüzde işletmeler için oldukça zorlu bir görev haline gelmiş ve bu bağlamda klasik yöntemler, çok boyutlu veri setlerinden kaynaklanan taleplere yeterince cevap verememekte veya özel durumları hesaplayamamaktadır. Makine

öğrenmesi, yapay zekâ çağının ortaya çıkardığı inovasyonlardan biri olarak, birçok mühendislik uygulamasının karşılaştığı zorluklara hızlı bir çözüm sunabilen önemli bir alandır. Tekstil sektöründe makine öğrenmesi yöntemleri, geleneksel üretim süreçlerine yenilikçi ve verimli çözümler sunma potansiyeli taşımaktadır. Bu alandaki gelişmeler, işletmelerin üretim, stok yönetimi ve talep tahmini gibi kritik süreçlerde daha etkili ve optimize edilmiş kararlar almasına olanak tanımaktadır.

Tekstil sektörü genellikle mevsimsel ve modaya bağlı olarak değişen taleplerle karşılaşabilir. Bu nedenle, taleplerin bu değişkenlikleri doğru bir şekilde öngörebilmek önemlidir. Bu sektördeki hızlı değişimler ve tüketicilerin tercihlerindeki ani değişiklikler, doğru tahminler yapmayı zorlaştırabilir. Çalışma yapılan şirket polyester iplik üreten bir tekstil firmasıdır. Ürünlerinin hammaddeleri genellikle başka ülkelerden tedarik edilir. Bu durum, tedarik zinciri üzerindeki belirsizliklerin artmasına ve yeni pazarlara açılma veya mevcut pazarlardaki daralmalar gibi faktörlerin, talep üzerindeki belirsizliğe sebep olmaktadır. Yanlış talep tahminleri, fazla stok veya talebi karşılayamama gibi durumları tetikleyerek şirketin maliyetlerini artırabilir. Doğru tahminler yapmak için geniş bir veri kapsamına sahip olmak ve bu verilerin güvenilirliğini sağlamak önemlidir. Bu çalışmanın konusu tekstil firmasının bu problem alanlarına çözüm bulmaktır. Çalışma yapılan şirketin tekstil ürünleri için daha doğru talep tahminleri yapabilmek adına makine öğrenmesi modelleri geliştirilmiştir.

2. KAYNAK ARAŞTIRMASI

Talep tahmini, sadece tedarik zincirinin her aşamasında değil, aynı zamanda işletmelerin birçok bölümünde etkili bir şekilde kullanılabilmektedir (Saatçioğlu, 2016). Gelecekteki talebi öngörebilmek için, geçmiş verilere, uzman görüşlerine, deneyimlere ve çıkarımlara dayanarak bilimsel ve ölçülebilir yöntemlerden faydalanılmaktadır. (Adalı, 2020).

Perakende satışlar genellikle belirgin bir eğilim ve mevsimsel farklılık gösterir. Bu farklılıkların en iyi nasıl modelleneceği ve tahmin edileceği, zaman serisi analizinde uzun süredir devam eden bir sorundur. Alon ve diğerleri (2001)' de yaptıkları çalışmada yapay sinir ağları ve geleneksel istatistik yöntemleri karşılaştırmış ve yapay sinir ağlarının daha iyi sonuçlar verdiğini göstermiştir. Liu ve diğerleri (2013), moda perakende talep tahminine yönelik yayınlar üzerinde kapsamlı bir inceleme gerçekleştirmiş ve farklı analitik yöntemlerin moda perakende talep tahminindeki avantajlarını ve dezavantajlarını keşfetmiştir. İnceledikleri literatür üzerinden, moda tahmin literatürünü saf istatistiksel modeller, saf yapay zekâ modelleri ve hibrit modeller açısından bir değerlendirme yapılmış ve saf istatistiksel yöntemlerin tek başlarına sofistike tahmin sonuçları üretmeye yeterli olmadığını göstermiştir. Yapay zekâ modellerinin daha başarılı sonuçlar verdiğinden bahsetmişlerdir.

Teknolojinin hızlı ilerlemesiyle birlikte yeni kavramlar ortaya çıkmakta ve gün geçtikçe artan veri miktarı ve çeşitliliği, veri yönetimini daha önemli hale getirmektedir. Özellikle üretim sektöründe artan rekabetle birlikte, kalite, maliyet ve verimlilik gibi unsurları yönetmenin yanı sıra, taleplere hızlı yanıt verebilmek de kritik bir öneme sahip hale gelmiştir. Bu hızlı yanıt verme ihtiyacını karşılamak için etkin bir planlama gereklidir ve etkin bir planlamanın temelini tahmin oluşturur (Ünlü, 2021). Geleneksel tahmin yöntemleri, eldeki geçmiş verilere dayanarak tahminler yapmada kullanılır. Ancak, günümüzde artan veri çeşitliliği ve miktarıyla birlikte, klasik tahmin yöntemleri yetersiz kalmaktadır. Bu noktada, yapay zekâ yöntemleri, özellikle makine öğrenimi ve derin öğrenme teknikleri, talep tahmininde daha etkili bir alternatif sunmaktadır. Bu yöntemler, büyük ve karmaşık veri setlerini işleyerek daha hassas tahminler yapabilir ve değişken koşullara daha hızlı adapte olabilir. Yapay zekâ destekli talep tahmin yöntemleri,

işletmelerin dinamik talep değişikliklerine daha hızlı tepki vermesine olanak tanır, böylece üretim, stok yönetimi ve tedarik zinciri planlamasında daha optimize edilmiş kararlar alabilirler. Bu nedenle, günümüzde işletmeler, yapay zekâ tabanlı tahmin yöntemlerini benimsemek suretiyle rekabet avantajı elde etmeyi hedeflemektedirler.

Kück ve Freitag (2021)' de tekstil sektöründeki talebin tahmin edilmesine odaklanan çalışmalara değinmektedir. Bu çalışmada talep tahmin yöntemi olarak K-en yakın komşu modelinin kullanımı araştırmışlardır. Teucke ve diğerleri (2014)'de bir tahmin modeli olarak destek vektör regresyonu algoritmasını kullanarak tekstil endüstrisinin bir örneğini ve belirli sayıda ürünün tahminini analiz etmektedir. Her iki çalışma da makine öğrenimi algoritmalarının istatistiksel modellerden daha iyi performans gösterdiğini ve nispeten düşük bir ortalama hata ürettiğini bulmuşlardır.

Genellikle, zamansal seri yöntemleri büyük veri setleriyle çalışırken, karmaşık parametreleri en uygun şekle sokma ve uzman deneyime ihtiyaç duymaları nedeniyle zorlayıcı olabilir. Ayrıca, bu yöntemler genellikle benzer bir yapı ile sınırlıdır ve bu nedenle tekstil ortamı gibi moda sektörlerinde kullanıma uygun olmayabilir. Bu geleneksel yöntemler, yapay zekâ (YZ) gibi daha modern yöntemlerle kıyaslandığında genellikle daha düşük performans gösterirler (Brahmadeep ve Thomassey, 2016).

Kang ve diğerleri (2017) çalışmalarında, bir bisiklet firmasının verilerini kullanarak belirlenen süre içindeki kiralama sayısını tahmin etmek istemişlerdir. Tahmin modelini oluşturmak için başlangıç zamanı, kullanıcı tipi, doğum yılı ve cinsiyet gibi değişkenler kullanılmıştır. Veriler, rassal bir şekilde %70 eğitim ve %30 test verisi olarak bölünmüştür. Performans kriteri olarak HKO kullanılmıştır. Bu değerlendirme sonuçlarına göre, üç algoritma arasında en iyi performansı gösteren algoritmanın rassal orman algoritması olduğu gözlemlenmiştir. Giri ve diğerleri (2019), kadın giyim ürünlerinin geçmiş satış verileri ve bunlarla ilişkilendirilmiş görüntüler baz alınarak, bir sonraki dönemdeki ürün talebini tahmin etmeyi hedeflemişlerdir. Bu araştırmada, genellikle öğrenme başarısı kanıtlanmış çok katmanlı algılayıcı modeli, regresyon modeli için kullanılmıştır. Performans değerlendirmesi için ortalama mutlak hata (OMH), hataların karelerinin ortalaması (HKO) ve hataların karelerinin ortalamasının karekökü

(HKOK) gibi kriterler kullanılmıştır. Sonrasında, belirlenmiş bir grup ürün satışının birkaç hafta boyunca oldukça doğru bir şekilde tahmin edildiği gözlemlenmiştir. Bu yöntem, geçmiş veriler ve görüntü analizi üzerine kurulu olarak, moda sektöründe talep tahmininde başarılı bir model sunma potansiyeline sahip olduğu sonucuna varmışlardır.

Loureiro ve diğerleri (2018) tarafından gerçekleştirilen çalışmada, derin öğrenme yaklaşımı ile elde edilen talep tahminleri, çeşitli tekniklerle karşılaştırılmıştır. Bu teknikler arasında karar ağaçları, rasgele orman, destek vektör regresyonu, yapay sinir ağları ve doğrusal regresyon yer almaktadır. Çalışmanın sonuçlarına göre kullanılan model, tekstil pazarındaki satışları tahmin etme konusunda başarılı olmuştur. Ancak, baz alınan performans metriklerinin bir kısmına göre, rastgele orman algoritmasının performansı daha iyi sonuçlar sergilemiştir. Shams ve diğerleri (2023), 10 özelliği içeren bir hedef olarak sıcaklığı tahmin etmek için makine öğrenimi (MÖ) regresyonlarını kullanarak dünya ortalaması sıcaklık ile bir dizi diğer değişken arasındaki bağlantıyı araştırmaya çalışmışlardır. Küresel sıcaklığı tahmin etmek için doğrusal regresyon (DR), rasgele orman (RO) regresyonu, karar ağaçları (KA) regresyonu, K-en yakın komşu regresyonu, destek vektör regresyonu (DVR) ve CatBoost regresyonu (CBR) gibi MÖ regresyonları kullanarak CBR'nin yakın zamandaki MÖ yaklaşımlarına göre yüksek sonuçlar elde ettiğini göstermiştir. Daldır (2021), havayolu ulaşım talebinin artması ve teknolojik gelişmelerin etkisiyle biriken büyük veri setlerinin makine öğrenimi yöntemleri kullanılarak analiz edilmesine odaklanmıştır. Uluslararası bir havayolu şirketinden alınan uçuş verileri ve çevrimiçi kaynaklardan elde edilen hava durumu verileri kullanılarak bir veri seti oluşturulmuştur. Oluşturulan veri seti, yapay sinir ağları, rastgele ormanlar, XGBoost, LightGBM ve CatBoost gibi makine öğrenimi modelleri ile incelenmiş ve performansları çeşitli hata matrisi temelli metriklerle değerlendirilmiştir.

Kharfan ve diğerleri moda perakendeciliğinde yeni başlatılan ürünler için talep tahmin sürecini optimize etmeyi amaçlayan bir veri odaklı metodoloji önermiştir. Üç adımlı bir yaklaşım kullanılarak, makine öğrenimi teknikleri olan kümeleme, sınıflandırma ve tahminleme yöntemleriyle talep tahmin süreci geliştirmişlerdir. Veri ön işleme süreci, tahmin doğruluğunu artırmak için yeni özellikler oluşturmayı ve önemli özellikleri seçmeyi içerir. Özellikle, kategorik özelliklerin işlenmesi ve tahminde kullanılması gibi

zorluklar ele alınmıştır. Comlan ve Koulo (2022) çalışmasında perakende bağlamında, satın alma miktarlarının doğru bir şekilde belirlenmesinin önemini vurgulamıştır. Yanlış belirlenen miktarların gereksiz depolama maliyetlerine veya satış kaybına yol açabileceğini, bu nedenle, etkin satın alma miktarlarının belirlenmesi için yapay zekâ yöntemlerinin kullanımı önermiştir.

Talep tahmini belirsiz ve karmaşık bir süreçtir ve birçok farklı faktör tahmin sonuçlarını etkileyebilir. Bunlar firmaya özgü, yani içsel ve dışsal parametreleri içerir. İçsel faktörler firmaya özgü iken, dışsal faktörler uluslararası ticaret yapan firmaların talep tahminlerini önemli ölçüde etkileyen makroekonomik göstergelerdir. Yasir ve diğerleri (2022), talep tahmininin içsel ve dışsal parametrelerin önemini araştırmışlardır. Bir tekstil hazır giyim firmasının günlük üretim verilerini kullanmışlardır. Talep tahmini için doğrusal regresyon (DR), destek vektör regresyon (DVR) ve uzun kısa süreli bellek modeli kullanılmıştır. Tahmini değerlendirmek için OMH, HKO ve HKOK hata metrikleri kullanılmıştır. Çeşitli operasyonel kararlarla olan yakın ilişkisi nedeniyle pazar talep tahmini, tüm organizasyonlarda temel faaliyetlerden biri olarak kabul edilmektedir. Ne yazık ki tekstil sektörü, kısa ürün yaşam döngüleri, özel etkinlikler ve trendlerin piyasada neden olduğu dalgalanma nedeniyle tahmin oluşturmada en çok zorluk çeken sektör olmuştur. Başlangıçtan beri tahminlerde geleneksel istatistiksel yöntemler kullanılmıştır. Medina ve diğerleri (2022), doğrusal regresyon, Ridge, Lasso, K-en yakın komşu, destek vektör Regresyon ve Rastgele Orman gibi regresyon odaklı algoritmalar uygulayarak tekstil endüstrisi için tahminler oluşturmaya yönelik bir araç olarak makine öğrenimini araştırmışlardır.

Araştırılan kaynaklardan edinilen bilgiye göre tekstil sektöründeki makine öğrenimi yöntemleriyle yapılan tahminlerde, öncelikle sipariş ve talep verileri kullanılmıştır. Bu veriler, geçmiş talep trendlerini analiz ederek gelecekteki talepleri tahmin etmek için temel bir parametredir. Aynı zamanda stok seviyeleri, mevcut ürün özellikleri ve moda trendleri de dikkate alınır. Stok seviyeleri, talep karşılamak ve talep döngülerini öngörmek için önemlidir. Ürün özellikleri, özellikle moda tekstil sektöründe, tüketici taleplerini ve tercihlerini anlamak için kullanılır. Moda trendleri, mevcut pazar koşulları ve ekonomik faktörler de tahminleri etkileyen kritik parametrelerdir. Makine öğrenimi

algoritmalarının seçimi, kullanılan modelin başarısını belirler ve veri setinin doğru bir şekilde kullanılmasını gerektirir. Bu parametreler, tekstil sektöründe tahmin modellerinin oluşturulmasında ve doğru tahmin sonuçlarının elde edilmesinde kritik bir rol oynar. Bu nedenle, talep tahmini yaparken bu parametreleri dikkate almak, stok yönetimi, üretim planlaması ve pazar stratejilerinin etkili bir şekilde yönetilmesine katkı sağlar. Tekstil sektöründe bu konuda sınırlı literatür bulunması, bu çalışmanın işletmelere yol gösterici bir kaynak olması beklentisini artırmaktadır. Bu çalışma, karar ağacı algoritmaları ve doğrusal regresyon gibi regresyon odaklı algoritmaları kullanarak gerçek veri seti için bir talep tahmin modeli geliştirilmiştir.

Depo yönetimi, ürünlerin izlenmesi, ardından depolama ve kar getiren ürünlerin koordinasyon süreçlerini içerir ve sonunda ürünlerin doğru yerlere dağıtımını sağlar (Harper, 2010). Mal kabul sürecinden sonra, hammadde, yarı mamul ve ürünler belirli bir sistem içinde planlı ve organize bir biçimde ele alınır. Bu süreçte elleçleme, istifleme ve sevkiyat işlemleri gerçekleştirilirken, aynı zamanda ayırma, birleştirme, konsolidasyon, etiketleme, paketleme, paletleme gibi bir dizi lojistik aktivitenin potansiyeli bulunmaktadır.

Depo yönetimi, tedarik zinciri içindeki kritik bir unsurdur ve etkili bir depo yönetimi, iş süreçlerini optimize ederek verimliliği artırabilir. Bu nedenle, depo süreçlerinin etkili bir şekilde planlanması, organize edilmesi ve uygulanması, tüm lojistik operasyonlar için kritik öneme sahiptir (Görçün, 2017).

3. MATERYAL ve YÖNTEM

Bu bölümde, uygulamada kullanılan yöntemler açıklanmıştır. Makine öğrenmesi ve kullanılan algoritmalardan bahsedilmiştir.

3.1. Makine öğrenmesi

1959 yılında makine öğreniminin öncülerinden olan Arthur Samuel, makine öğrenmesini "bilgisayarlara açıkça programlamadan öğrenme yeteneği veren bir alan" olarak tanımlamıştır. Makine öğrenmesi basitçe, verilerin ayrıştırılması için algoritma kullanılması, öğrenme süreci ve ardından dünyadaki bir konu hakkında belirleme yapma veya tahmin uygulaması olarak açıklanmıştır. Bu yöntemde, manuel olarak programlama yerine model, mevcut bir veri seti ile eğitilir ve ardından yeni veriler üzerinde öğrenilmiş görevleri gerçekleştirir. Makine öğrenmesinde etkili bir öğrenme sağlayabilmek için örnek vakaların ve ilgili girdi parametrelerinin toplanması önemlidir. Bu özellik, makine öğrenme modelini eğitmek için kullanılacak verinin yeterli boyut ve kalitede olmasını gerektirir (Aylak vd., 2021).

Makine öğrenmesi uygulamak için kullanılan bir dizi program ve araç bulunmaktadır. Bu araçlar, makine öğrenimi algoritmalarını kullanarak verileri analiz edebilme ve tahmini sonuçlar elde edebilme yeteneği sunar. Makine öğrenimi sürecinde, programın öğrenmesi için kullanılacak veri seti genellikle büyük veri, eğitim verisi ve test verisi olmak üzere iki ana parçaya ayrılır. Eğitim verisi, modelin öğretildiği ve uyarlandığı veri setini içerir. Test verisi ise eğitim veri setinde oluşturulan modelin performansını değerlendirmek için kullanılır (Ersöz vd., 2021).

3.1.1. Makine Öğrenmesi Algoritmaları

Makine öğrenimi algoritmaları, büyük ve karmaşık veri setlerini keşfetmek, analiz etmek ve elde edilen sonuçları yorumlamak için kullanılan program parçacıklarıdır. Her bir algoritma, belirlenen amaca ulaşmak için uygulanan adımlar ve talimatlardan meydana gelir. Makine öğrenmesi modelleri, tahmin yapma ve verileri sınıflandırmayı amaçlar. Bu algoritmalar, büyük veri kümesini temsil eden ve eğitim verilerine dayanan parametreleri

kullanır. Eğitim verisinin boyutu arttıkça, makine öğrenmesi algoritmaları daha doğru sonuçlar verir. Makine öğrenimi alanındaki gelişmiş algoritmaların sayısı her geçen gün artmaktadır. Her algoritmanın amacının farklı olması kısıtları oluşturmada etkilidir. Farklı çeşitteki algoritmalar, verileri birbirinden farklı şekillerde analiz eder ve çözümler. Bu çalışmada kategorik verilerin çok olması, aşırı öğrenmenin önüne geçmek ve yüksek boyutlu veriyi verimli bir biçimde işleyebilmek için karar ağacı tabanlı algoritmalar kullanılmıştır. Doğrusal regresyon algoritmalarının da model üstünde performansı değerlendirilmiştir. Kullanılan makine öğrenmesi algoritmaları aşağıda açıklanmıştır:

Doğrusal Regresyon: Bu model, kendisinden türetilen farklı regresyon modellerinin temeli olarak kabul edilir. İşlemi iki değişken arasında doğrusal bir ilişki bulmaktan ibarettir. Bu değişkenler bağımsız değişken (X) ve bağımlı değişken (Y) olarak kabul edilir. Bu ilişki bağımlı değişkenin bağımsız değişkene dayalı olarak tahmin edilmesine olanak sağlar (Karaman ve Bektaş. 2023). Doğrusal regresyon modeli, mevcut değişkenlere bağlı olarak tek değişkenli veya çok değişkenli olabilir, yani yalnızca bir hedef değişken üzerinde gerçekleştiriliyorsa tek değişkenli olarak kabul edilir. Ancak çeşitli öngörücülerin kullanılması durumunda, model çok değişkenli hale gelir (Medina vd. 2022).

Doğrusal regresyon yöntemi bağımlı değişkeni tahmin etmek için bağımsız değişken ile bağımlı değişken arasında doğrusal bir ilişki kuran istatistiksel bir yöntemdir. Doğrusal regresyon eşitliği;

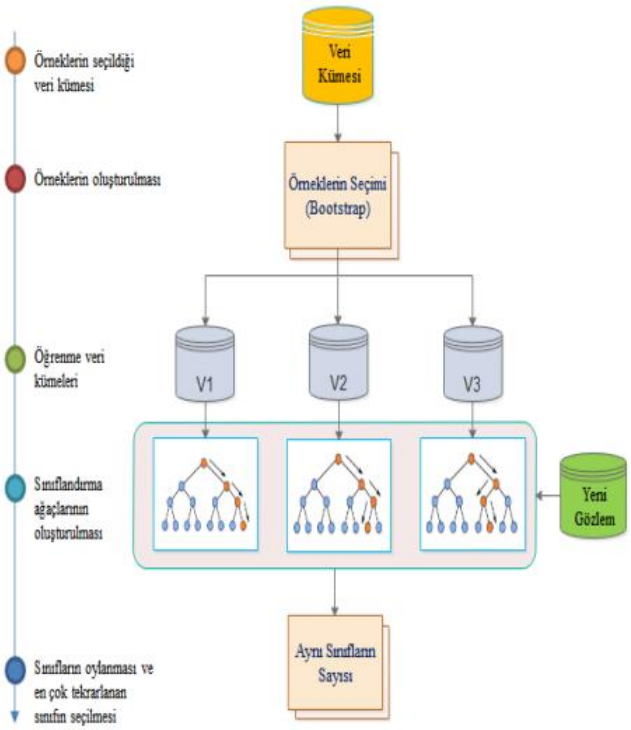
$$y=a_1x + b_1 \tag{3.1}$$

ile verilmektedir. Denklem (3.1) de x bir sonraki gün için fiyat tahmincisini, y cevap değişkenini, a_1 regresyon çizgisinin eğimini ve b_1 regresyon çizgisinin ordinat eksenini kestiği noktayı ifade etmektedir (Karasu vd. 2018).

Karar ağaçları, veri kümesindeki karmaşık ve doğrusal olmayan ilişkileri modelleme konusunda doğrusal regresyona göre daha esnektir ve kategorik değişkenlerle daha iyi çalışma eğilimindedir. Doğrusal regresyon genellikle bu tür değişkenleri dönüştürmek

veya kodlamak için ek işlemler gerektirir. Karar ağaçları, makine öğrenmesi tekniklerinin en yaygın ve bilinen yöntemlerinden biridir. Karar ağaçlarının kuruluş aşamasında en büyük dezavantajlarından biri, ağacın doğru bir şekilde oluşturulamamasıdır. Rastgele ormanlar, birçok karar ağacının bir araya gelmesi sonucu oluştuğu için, karar ağaçlarının bu dezavantajını ortadan kaldırmak amacıyla kullanılır. Adını, birçok karar ağacının bir araya gelmesinden kaynaklanan bu özelliğinden alan rastgele ormanlar, karar ağaçlarına özgü bu zorluğu aşmada etkili bir çözüm sunar (Güven, 2020).

Rastgele Orman Regresörü: Bu model, regresyona odaklanan karar ağaçlarından türetilmiştir. Ancak, bu model, bir problemi modellemek için çok sayıda ağaç uygulayarak çalışır. Her bir ağaç, verinin bir kısmıyla çalıştığından, bu yöntem karar ağaçları modellerinin eksikliklerini gidermeyi amaçlar (Loureiro vd. 2018). Ayrıca, her bir son düğüm tarafından oluşturulan ölçümdeki ortalamayı hesaplayarak sonuç elde eder. Bu prosedür, yalnızca tek bir ağaç kullanıldığında elde edilen doğruluğu artırmak amacıyla karar ağaçları modellerinde yapılan bir düzeltmedir. Modelin dezavantajları arasında, tek bir karar ağacında olduğu gibi sonucun ağaç yapısıyla görsel olarak sunulamaması ve modelin karmaşıklığı nedeniyle çok sayıda karar ağacının değerlendirilmesine ilişkin işlem adımlarının görüntülenememesi bulunmaktadır (Peker vd. 2017). Şekil 3.1'de Rastgele Orman Algoritmasının adımları gösterilmiştir.



Şekil 3.1 Rastgele orman algoritmasının işlem adımları

Gradyan arttırma modelleri rastgele orman algoritmasına göre daha karmaşık modeller oluşturarak tahmin performansını artırma ve belirli problemlerde daha iyi sonuçlar elde etme potansiyeli taşır. Gradyan arttırma yöntemi, karar ağaçları gibi zayıf öğrenicileri bir araya getirerek güçlü bir öğrenici oluşturmayı amaçlayan bir topluluk öğrenme yöntemidir. Bu algoritma, her adımda mevcut modelin hatalarına odaklanarak yeni bir karar ağacı ekler ve bu ağaçlar birleştirilerek tahmin hataları minimize edilir. Gradyan arttırma makineleri, zayıf modellerin hatalarını gradyanlar aracılığıyla belirler ve gradyan inişi kullanarak güçlü bir tahmin modeli oluşturur. Bu yöntem, genellikle regresyon ve sınıflandırma problemlerinde etkili bir şekilde kullanılır ve tahmin performansını artırmak için yaygın olarak tercih edilir (Singh vd, 2021).

Aydın ve diğerleri (2023), Gradyan arttırma yönteminin avantajlarından biri, heterojen özellikleri başarıyla işleyebilme yeteneği olduğundan bahsetmiştir. Bu özellik, farklı ölçeklerde ve farklı dağılımlara sahip özellikleri etkili bir şekilde ele alabilme kapasitesini ifade eder. Ayrıca, aşırı öğrenmeye karşı direnç gösterir ve yüksek boyutlu veriyi verimli

bir biçimde işleyebilir. XGBoost, LightGBM ve CatBoost gibi çeşitli gradyan arttırma algoritmaları mevcuttur.

$(x, y)_{i=1}^{N}$ arasındaki bir veri kümesi ele alınacak olursa , burada x=(x_1…..x_d) özellikleri temsil eder ve y hedef değere karşılık gelir. Gradyan arttırmanın amacı, belirli bir kayıp fonksiyonu $\psi(y, \mathrm{f})$ minimize ederek $\check{F}(\mathrm{x})$ ile x→ y işlevsel bağımlılığı bulmaktır. Gradyan arttırma $F^*(\mathrm{x})$ m fonksiyonunun ağırlıklı toplamı olarak hesaplar (Bentejac vd., 2021).

$$F_m(\mathrm{x}) = F_{m-1}(\mathrm{x}) + \rho_m h_m(\mathrm{x}) \tag{3.2}$$

Burada ρ_m , m^{th} fonksiyon $h_m x$ 'in ağırlığını ifade eder. Bu fonksiyonlar, tahmin modeli olarak hizmet eder ve $F^*(\mathrm{x})$'in iteratif olarak biriktirilen eklemeli bir yaklaşımıdır. İlk yaklaşım, Denklem (3.3)'de verilen gibi sabittir:

$$F_0(\mathrm{x}) = arg\ arg \sum_{i=1}^{N} L(y_i, \alpha) \tag{3.3}$$

Ardışık modeller, Denklem (3.4)'ü minimize etmektedir:

$$(\rho_m h_m(\mathrm{x})) = arg\ arg \sum_{i=1}^{N} L(y_i, F_{m-1})(x_i) + \rho h(x_i)) \tag{3.4}$$

XGBoost: Çok sayıda karar ağacı kullanarak sınıflandırma ve regresyon görevlerinde yaygın olarak kullanılan bir gelişmiş gradyan tabanlı karar ağacı algoritması olan XGBoost, ağacın sınıflandırma işlevini daha yeniden üretilebilir hale getirmek için bir düzenlileştirme yöntemi kullanır. Ayrıca, düzenleme, büyük veri durumlarında gereken özellik değeri tahminine de yardımcı olur. XGBoost'un amacı, kayıp terimi (L) ve düzenleme terimi (Ω) toplamını minimize etmektir (Bentejac vd., 2021).

$$L_{xgb} = \sum_{i=1}^{N} L(y_i, F(x_i)) + \sum_{m=1}^{M} \Omega\,(h_m) \tag{3.5}$$

$$\Omega\,(h_m) = \gamma T + \frac{1}{2}\lambda ||w||^2 \tag{3.6}$$

Burada T, ağacın yaprak sayısını ve w, yaprakların çıktı puanlarını ifade eder.

CatBoost Regressor: Karar ağaçları, çoğu gradyan artırma uygulamasında temel tahminleyiciler olarak kullanılır. Karar ağaçları sayısal özellikler için kullanışlı olabilir, ancak gerçekte birçok veri seti, tahmin için esaslı olan kategorik öğeler içerir. CatBoost kategorik parametrelere sahip, minimum değişken içeren ve üstün doğruluk sağlayan simetrik bir karar ağacı algoritma kuralına dayanmaktadır. Bu yöntemler, çeşitli hiperparametre ayarlama seçenekleri veya diğer kayıp fonksiyonları kullanılarak fonksiyonun uyumunu optimize etme konusunda büyük bir esneklik sunarlar. Veriyi önceden işleme koymaları gerekmez çünkü doğrudan girişle yani ham veri ile çalışabilirler (Shams vd. 2023). Bu algoritma kullanımdaki veri kümesinin rastgele bir permütasyonu gerçekleştirilir ve her örnek için ortalama bir etiket değeri hesaplanır. CatBoost model değerlendiricisi şu gerçeklerden yararlanır: İlk önce kullanılan tüm tek-sıcak kodlanmış özellikleri, kayan özellikleri ve istatistikleri 0-1 gibi ikiliye dönüştürür ve daha sonra bu ikili özellikleri model tahminlerini hesaplamak için kullanır (Prokhorenkova vd., 2018).

Bir veri kümesi üzerinden düşünüldüğünde $\{(x_k, y_k)\}_{k=1}^{n}$, x_k rastgele seçilen özelliklerin bir vektörü ve y_k'nin bir hedef olduğu durumda, CatBoost ilk başta veri kümesinin $(\sigma_0, \dots \sigma_s)$ bağımsız rastgele permütasyonunu oluşturur. Şimdi, bu permütasyonlar $(\sigma_0, \dots \sigma_s)$ bölünmelerin değerlendirilmesinde kullanılacak, σ_0 ise elde edilen ağaçların yaprak değerlerinin seçilmesi için kullanılacaktır. Sıralı güçlendirmede, M_1, ..., M_n gibi birçok destekleyici model bulunmaktadır. Her bir M_i modeli, σ permütasyonunda i numunesi için mevcut tahmin görevini yerine getirecektir. CatBoost, "tahmin kayması"nı önlemek için $\sigma_1 = \sigma_2$ gibi permütasyonlardan yararlanır. Bu teknik, modelin eğitilmesinde hedefin kullanılmasını engelleyerek, "Hedef İstatistiği" veya "gradyan tahmini" hesaplamalarında tahmin kayması sorununu ortadan kaldırır (Pujara vd. 2022)

LightGBM: Gradyan tabanlı karar ağaçlarının yeni bir türüdür ve çeşitli modelleme zorluklarına, sınıflandırma ve regresyon gibi görevlere yönelik bir çözüm sunar. LightGBM, birçok veri örneğini ve işlevi içermek için gradyan tabanlı tek taraflı örnekleme ve özel işlev gruplama gibi iki yeni strateji kullanır. Diğer gradyan artırma

teknikleri veya Aşırı Gradyan Yükseltme ile karşılaştırıldığında, LightGBM, karar ağacını dikey olarak artırır, bu da büyük hacimli verileri daha etkili bir şekilde işleyebilme yeteneğini geliştirir (Tekin ve Sarı, 2022).

Aşırı Ağaçlar Regresyonu: Extra-Tree olarak da bilinen bu yöntem, verilen sayısal özelliklere dayanarak en uygun kesme noktalarını rastgele seçerek ağaçlar oluşturur. Bu yöntem, özellikle yüksek boyutlu ve karmaşık problemlerde etkilidir. Ağaç oluşturma sürecinde, örnekleme (bootstrapping) kavramını kullanmaz ve kesme noktalarını rastgele seçer, bu da önyargısız bir yaklaşım sağlar. Bu yöntem genellikle rastgele ormanlardan daha esnek ve doğrusal olmayan modeller üretir (Polamuri vd., 2019).

3.2. Hata Ölçütleri

Hata ölçütleri, HKO, OMH, HKOK ve determinasyon katsayısını (R^2) içerir. OMH ve HKOK, tahmindeki pozitif ve negatif hataların karşılıklı olarak karşılaştırılmasından kaçınarak, sırasıyla tahmin edilen değerin gerçek değere yakınlığını ve tutarsızlığını değerlendirir (Denklem (3.7) ve Denklem (3.8)). HKO, tahmin edilen değerin gerçek değerden sapmasını yansıtır (Denklem 3.9). R^2 bağımlı değişkendeki değişikliklerin ne kadarının bağımsız değişkendeki değişikliklerden kaynaklandığını gösterir ve 1>R^2>0 arasında değer almalıdır. R^2 değeri 1'e ne kadar yakınsa, regresyon doğrusunun o kadar iyi uyduğu söylenmektedir (Denklem 3.10). Aşağıda sırasıyla R^2, HKO, OMH ve HKOK'a ait denklemler verilmektedir (Kaysal vd., 2022).

$$\text{OMH}=\frac{1}{N}\sum_{i=1}^{N}|x_i - x_i^*| \tag{3.7}$$

$$\text{HKOK}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - x_i^*)^2} \tag{3.8}$$

$$\text{HKO}=\frac{1}{N}\sum_{i=1}^{N}(x_i - x_i^*)^2 \tag{3.9}$$

$$R^2 = \frac{(\sum_{i=1}^{N}(x_i^* - \overline{x_i^*})(x_i - \overline{x_i}))^2}{\sum_{i=1}^{N}(x_i^* - \overline{x_i^*})^2 \sum_{i=1}^{N}(x_i - \overline{x_i})^2} \tag{3.10}$$

Denklemlerde x_i , x_i^* , $\overline{x_i}$, $\overline{x_i^*}$ sırasıyla gerçek değeri, tahmin edilen değeri, ortalama gerçek değeri, ortalama tahmin edilen değeri ve örneklem büyüklüğünü gösterir.

3.3. Hiperparametre Ayarlama

Makine öğrenmesi model tasarlarken kullanılan parametreler, iki gruba ayrılır: eğitim süreci sırasında doğrudan veriden elde edilebilen parametreler ve tasarımcı tarafından önceden belirlenen parametreler. Bunlar model parametreleri ve hiperparametre olarak adlandırılır. Model parametreleri genellikle veriden tahmin edilir. Tasarımcının bu parametreleri ayarlaması beklenmez. Bu parametreler öğrenilen modelin bir parçası olarak kaydedilir (Al-Khazraji vd. 2022).

Hiperparametre ayarlama süreci, her tahmin modelindeki optimal hiperparametre değerlerini bulmayı içerir. Bu süreç, aşırı uyumu veya yetersiz uyumu önlemek amacıyla bir veri kümesiyle eğitildiğinde gerçekleştirilen bir optimizasyona dayanır (Kaysal vd., 2023). Her model, sayılar veya karakterler gibi hiperparametrelerle karakterize edilir. En uygun hiperparametre değerlerini bulmanın temel ve kapsamlı bir yolu, her bir hiperparametre için bir değer aralığı belirlemek ve eğitim veri kümesi üzerinde model değerlendirmesini birleştirmektir. Bu görev, programlama ortamındaki özel kütüphaneleri kullanarak optimize edilebilir. Örneğin, Python'daki GridSearch adlı bir araç, her bir hiperparametre için belirli bir değer aralığını kullanarak eğitim verileri üzerinde kombinasyonlar gerçekleştirir ve modelin veri kümesine en iyi uyum sağlayan ideal hiperparametre kombinasyonunu belirler. Bu yöntem, kapsamlı bir arama yapar, kullanımı kolaydır ve paralelleştirilebilir yani farklı hiperparametre kombinasyonlarını bağımsız olarak değerlendirebilir (Medina vd. 2022).

Hiperparametre optimizasyonu işlemi, modeli eğitirken çapraz doğrulama yöntemi kullanılmaktadır. Klasik bir makine öğrenimi aşamalarında, veri seti genellikle eğitim ve test kümeleri olarak ikiye ayrılır. Modelin öğrenmesi için kullanılan eğitim veri seti ve modelin performansını değerlendirmek amacıyla kullanılan test verisi seti olarak ikiye ayrılır. Veri setini bu şekilde ikiye ayırma işlemi farklı test veri setlerinden kaynaklanan dağılım problemlerine yol açabilir. Bu sorunu aşmak için verilerin rastgele bölünmesi ve

her bir bölünmede farklı bir veri kümesinin test verisi olarak kullanılması olan K-Katlı Çapraz Doğrulama yöntemi kullanılır (Tekin ve Sarı, 2022). Bu şekilde, tüm verilerin eğitim ve test verisi olarak kullanılması, modelin gerçek dünya verilerindeki performansının daha doğru bir şekilde değerlendirilmesini sağlar. Özellikle veri setindeki gözlem birimi sayısının sınırlı olduğu durumlarda, eğitim ve test setinin bölünme sürecinde test setinin taşıdığı bilgilerin etkinliği şüphe uyandırabilir. İlk aşamada, veri seti rastgele bir şekilde %80 eğitim ve %20 test seti olarak ayrılabilir. Daha sonra, k-katlı çapraz doğrulama yöntemi kullanılarak eğitim seti, modelin performansının değerlendirilmesi için daha fazla rastgele alt kümelere (örneğin, 5 kat) ayrılır (Aydın, 2018). Her bir kat için model, eğitim verileri üzerine kurulur ve test verileri üzerinde değerlendirilir. Ardından, elde edilen performans ölçütleri (R^2, HKO, OMH ve HKOK) kullanılarak ortalama performans değerleri hesaplanır. Bu ortalama değerler modelin, problemi ne kadar etkili bir şekilde genelleştirdiğine dair önemli bilgiler sunar (Karasu vd., 2018).

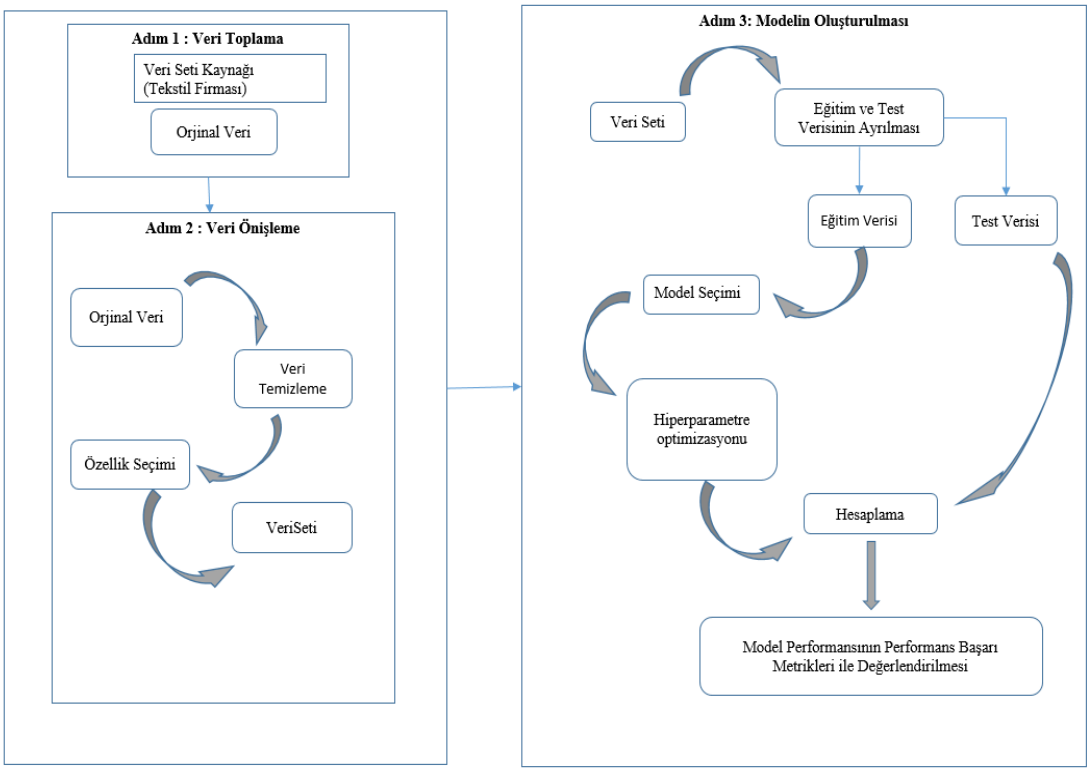
4. BULGULAR ve TARTIŞMA

Bu çalışma, polyester iplik üreten bir tekstil firmasının tedarik zinciri belirsizlikleri, pazar değişimleri ve stok problemleri üzerinde odaklanarak, daha doğru tahminler yapabilmek için makine öğrenmesi modelleri geliştirmeyi amaçlamaktadır. Çalışmanın kapsamında, polyester iplik üreten bir tekstil firmasındaki 2 yıllık satış verileri üzerinde detaylı bir analiz gerçekleştirilmiştir. Bu analiz adım adım ilerleyerek başlamıştır. İlk aşamada, çözümleme için gerekli olan veriler SAP ve firmada kullanılan sistemler üzerinden toplanmıştır. İkinci aşamada, toplanan veriler özenle ön işlemden geçirilmiş ve temizlenmiştir. Son aşama ise, çeşitli modellerin oluşturulması ve bu modellerin performanslarının detaylı bir biçimde değerlendirilmesiyle tamamlanmıştır.

Bu bölümde, kullanılan analiz yöntemleri adım adım bir akış haritası çerçevesinde Şekil 4.1'de açıklanmıştır. Her bir aşama, elde edilen bulguların güvenilirliği ve analiz sürecinin bütünlüğü açısından özenle tasarlanmıştır.

4.1. Akış Haritası

Bu çalışmada Şekil 4.1'de gösterilen akış haritası izlenmiştir. Bu plan, makine öğrenimi sürecinde izlenen tüm adımları detaylı bir şekilde göstermektedir. İlk aşamada, bir tekstil firmasının 2021-2022 yılları arasındaki gerçek satış verileri kullanılmıştır.



Şekil 4.1. Makine öğrenmesi uygulama haritası

İkinci aşamada, satış veri setindeki bilgiler analiz edilmiş ve bazı verilerin dönüşümü ile çalışma için daha iyi sonuçlar elde edebileceğini düşündüğümüz yeni öznitelikler oluşturulmuştur. Yeni özniteliklerin eklenmesi ve bazı özniteliklerin birleştirilmesiyle elde edilen genişletilmiş veri seti üzerine, çeşitli veri ön işleme adımları uygulanmıştır. Bu kapsamlı veri ön işleme sürecinde, veri setindeki öznitelikler detaylı bir istatistiksel analize tabi tutulmuştur. Nümerik değişkenler içindeki potansiyel aykırı değerler titizlikle belirlenmiş ve gerekli görüldüğünde düzeltilmiştir. Bu adımlar, elde edilen veri setinin daha güvenilir, tutarlı ve analiz için uygun hale getirilmesini sağlamak amacıyla gerçekleştirilmiştir. Kategorik değişkenler içinde, istatistiksel olarak anlamsız olan kategorilerin temizlenmesi amacıyla, öznitelik bazında sıfıra eşdeğer olan yüzdelik dilimdeki kategoriler çıkarılmıştır. Veri temizleme adımından sonra, istatistiksel filtreleme yöntemleri kullanılarak her girdi değişkeninin hedef değişkenle olan ilişkisi değerlendirilmiş ve en belirgin korelasyona sahip girdi değişkenleri seçilmiştir. Veri önişleme adımı tamamlandıktan sonra son aşama olan model oluşturma aşamasına geçilmiştir. Eğitim ve test verileri olarak iki kategoriye ayrılan veri seti, bu aşamada

incelenmiştir. Modelin öğrenme süreci için kullanılan eğitim seti, analiz amacına ve veri setinin özelliklerine uygun modellerle donatılmıştır. Seçilen bu modeller, belirlenen performans metriklerine en uygun sonuçları elde etmek amacıyla grid arama fonksiyonu kullanılarak optimize edilmiştir. Bu kapsamlı süreçte, her bir model için en etkili hiperparametre kombinasyonları titizlikle belirlenmiş ve modelin genel performansını artırmak üzere uyarlanmıştır. Bu adımlar, modelin daha yüksek hassasiyet, doğruluk ve genel başarı elde etmesini sağlamak için gerçekleştirilmiştir. Bu aşamada, modelin güvenilirliğini artırmak amacıyla 10 katlı çapraz doğrulama prosedürü detaylı bir şekilde uygulanmıştır. Bu prosedür, farklı test veri setleri üzerinde yapılan doğruluk değerlendirmelerinde ortaya çıkabilecek büyük farklılıklara karşı etkili bir önlem olarak ele alınmıştır. Her bir çapraz doğrulama katında elde edilen sonuçlar kapsamlı bir şekilde analiz edilerek, modelin genel performansındaki istikrarı artırmak için gerekli düzeltmeler yapılmıştır. Bu sayede, modelin çeşitli veri kesitlerine karşı daha tutarlı ve güvenilir bir performans sergilemesi sağlanmıştır. Makine öğrenimi modelleri, uygun hale getirildikten sonra eğitim veri seti üzerinde eğitilmiştir. Eğitim sürecinin tamamlanmasının ardından, her bir model performansları ayrıntılı bir şekilde karşılaştırılmıştır. Bu aşamada, modellerin genel başarıları, R^2, HKO, OMH ve HKOK gibi çeşitli performans metrikleri üzerinden değerlendirilmiştir. Elde edilen sonuçlar, her bir modelin belirlenen hedeflere ne kadar uygun olduğunu belirlemede kullanılmış ve en etkili modelin seçilmesine yönelik bilgiler sunmuştur. Bu kıyaslama süreci, modellerin gerçek dünya verilerindeki performanslarını anlamak ve en etkili modelin seçilmesini sağlamak için önemli bir aşama oluşturmuştur. Python 3 kullanılarak, veri ön işleme aşamalarından biri olan öznitelik seçimi, kategorik verilerin sayısal formata dönüştürülmesi, standartlaştırma, veri dengesizliği ile başa çıkma, hiperparametre optimizasyonu ve model performansının değerlendirilmesi gibi adımlar titizlikle uygulanmıştır. Bu süreçler, veri setinin kalitesini artırmak, modelin daha güçlü ve genelleştirilebilir hale gelmesini sağlamak ve nihayetinde analiz amacına uygun sonuçlar elde etmek için gerçekleştirilmiştir. Python 3 platformu, bu adımların etkin bir şekilde uygulanmasına olanak tanıyan geniş bir kütüphane ekosistemine sahip olmasıyla, veri bilimi ve makine öğrenimi çalışmalarında yaygın olarak tercih edilen bir dil olarak öne çıkmaktadır.

4.1.1 Veri Seti Özellikleri

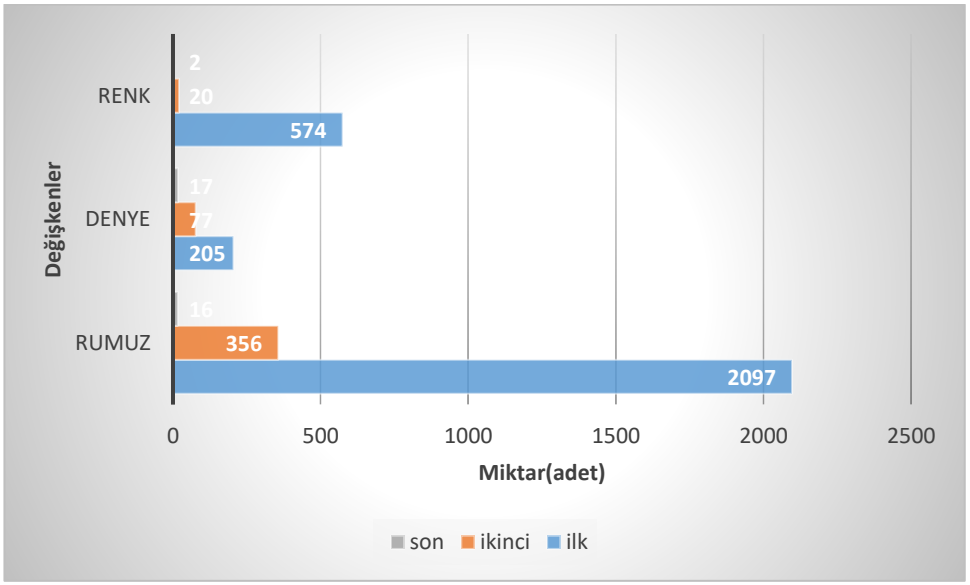
Bu çalışmada kullanılan veri seti, polyester iplik üreten bir tekstil firmasının 2021-2022 satış verilerinden oluşmaktadır. Bu veri seti 8 farklı özelliği içermektedir. Özelliklerin ismi, tanımı, tipi (kategorik/nümerik) ve kategorik olan özelliklerin sahip olduğu farklı ifadelerin miktarı Çizelge 4.1' de gösterilmektedir. Nümerik olan özellikler yalnızca gerçek sayılar içerdiğinden farklı ifade miktarı aralık olarak gösterilmiştir. Çalışma kapsamında, 2021-2022 yılları arasında tekstil sektöründe gerçekleşen 3039 satış verisi analiz edilmiştir. Rumuz, hafta, denye, renk, dolar, fiyat (dolar), ürün cinsi olarak toplam 7 adet bağımsız değişken içermektedir. Amaç talebi tahmin etmek olduğu için bağımlı değişken olarak miktar verisi etiket olarak seçilmiştir.

Çizelge 4.1. Veri seti özellikleri

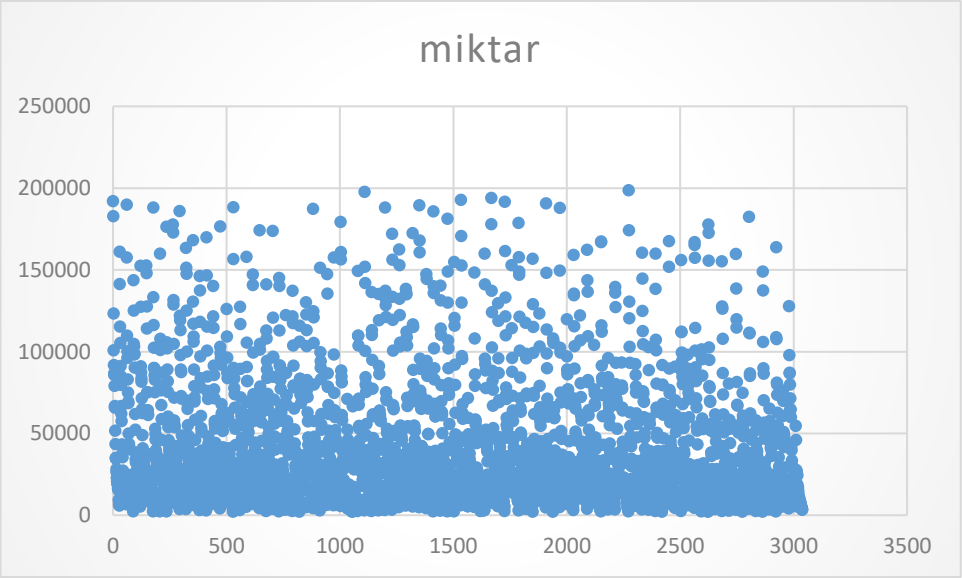
Özellik ismi	Tanım	Özellik Tipi	Miktar
Rumuz	Satışı yapılan polyester iplik malzemelerinin ismi	Kategorik	16
Hafta	Polyester ipliklerin yıl içinde satıldığı hafta	Kategorik	52
Denye	Polyester iplik kalınlığı	Kategorik	17
Renk	Polyester ipliğin rengi	Kategorik	2
Dolar	Satış yapılan haftanın dolar kuru	Nümerik	6,97-18,67
Fiyat	Satışı yapılan polyester ipliğin hafta fiyatı	Nümerik	0,07-6,76
Ürün cinsi	Satışı yapılan polyester ipliğin özelliği	Kategorik	11

4.1.2. Veri Önişleme

Veri dönüşüm adımları, tamamıyla Excel programı üzerinde gerçekleştirilmiştir. Veri setinin dönüşümü sırasında, uygulama amacına uygun olmayan değişkenler ve istatistiksel analiz yapılamayacak olan boş hücreler belirlenmiş ve bu unsurlar veri setinden çıkarılmıştır. Bu sayede, başlangıçta 18361 olan ürün sayısı, uygunsuz verilerin temizlenmesiyle birlikte 13168'e indirilmiştir. Bu dönüşüm adımlarıyla veri setinin kalitesi artırılmış ve analiz için daha uygun bir form kazandırılmıştır. Rumuz verisi parçalara ayrılıp renk, ürün cinsi ve denye olarak yeni öznitelikler oluşturulmuştur. Daha sonra Abc analizi yapılarak en çok satılan ve toplam stoğun %70'ini oluşturan A grubu baz alınmış ve veri sayısı 3039'a düşürülmüştür. En çok satılan grubun baz alınmasıyla özellikle rumuz, denye ve renk değişkenlerinin sayısında ciddi bir azalma gerçekleşmiştir. Rumuz, denye ve renk değişkenlerinin değişimi Şekil 4.2'de gösterilmiştir. İlk başta rumuz sayısı 2097 adet iken veri setinin son halinde 16 adete düşmüştür. Aynı şekilde renk 574 adetten 2 adete, denye 205 adetten 17 adete düşürülmüştür. Son 2 yılda çok az satış gören ürünler çıkarılmış ve etiket değeri olan miktar verisi için aykırı ve anlamsız değerler temizlenmiştir. Miktar etiket değerinin dağılımı Şekil 4.3' de gösterilmiştir.

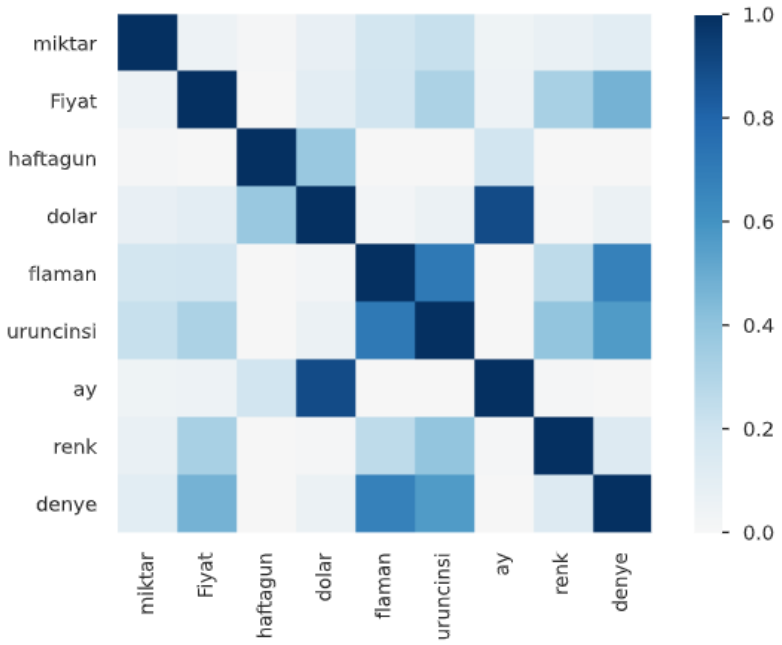


Şekil 4.2. Rumuz, denye ve renk değişkenlerinin değişimi

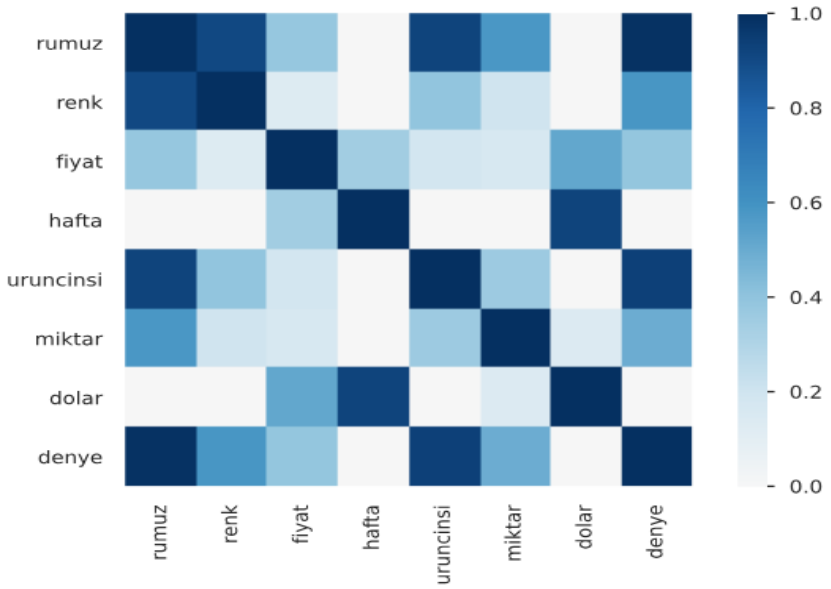


Şekil 4.3. Miktar etiket değerinin dağılımı

Tarih verisi olarak başta gün, hafta ve ay verisi değerlendirilip gün ve ay verisinin sonuca çok etki etmediği tespit edildiğinden hafta verisi kullanılmıştır. İlk olarak kategorik verilerde eksik değerler mode, nümerik verilerde eksik değerler mean ile doldurulmuştur. Kategorik verilere one-hot encoding (0-1) işlemi uygulanarak sayısal verilere dönüştürülmüştür. İlk oluşturulan veri seti setindeki korelasyon matrisi Şekil 4.4 de gösterilmiştir. Haftanın günlerini ifade eden haftagün ve ay değerlerinin miktar etiketine etkisi az olduğu için Şekil 4.5' de görüldüğü gibi yılın haftalarını temsil eden hafta verisi olarak değiştirilmiş ve miktar etiketine daha çok etki ettiği görülmüştür. İlk korelasyon matrisinde denye ve flaman verisi ayrı ayrı miktar etiketiyle daha az korelasyon oluştururken iki veriyi birleştirip denye olarak kullanılınca miktar etiketine daha çok etki ettiği görülmüştür. Miktar etiketine etki eden öznitelikler ile veri seti son haline getirilmiştir. Rumuz, denye, fiyat ve ürün cinsi değişkenlerinin miktar etiketiyle korelasyonu yüksek çıkmıştır.



Şekil 4.4. İlk oluşturulan korelasyon matrisi



Şekil 4.5. Son oluşturulan korelasyon matrisi

4.1. Modelin Oluşturulması

Satış verilerinin doğru tahmin edilebilmesi için öncelikle, bir tekstil firmasına ait 2 yıllık satış veri seti üzerinde çeşitli makine öğrenmesi yöntemleri uygulanarak veri seti uygun hale getirilmiştir. Veri seti, rastgele bir şekilde eğitim ve test veri seti olarak ayrılmıştır. Daha sonra, girdi değişkenlerini normalleştirmek amacıyla Min-Max normalleştirme yöntemi kullanılmıştır. Bu teknik, veri özelliklerini belirli bir aralığa sıkıştırarak özelliklerin dağılımını belirli bir aralıkta tutmayı amaçlar ve algoritmanın özellikler arasında dengeli bir şekilde işlemesine olanak tanır. Bu ölçeklendirme yönteminde, her özellik değeri, o özelliğin minimum ve maksimum değerleri arasında bir orana çevrilmektedir. Bu, her özelliğin aynı ölçekte olmasını sağlar.

$$x' = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{4.1}$$

Denklem (4.1)' de x' normalleştirilmiş veriyi, x_i girdi verisini, x_{min} girdi veri seti içindeki en küçük değeri, x_{max} ise girdi veri seti içindeki en büyük değeri ifade etmektedir (Kaysal vd., 2022).

Probleme ve veri setine uygun materyal ve yöntem bölümünde anlatılan regresyon modelleri Python 3 platformunun Pycaret kütüphanesinde uygulanmıştır. Eğitim veri setine uygulanan modellerin performansının değerlendirilmesinde sırasıyla R^2, HKO, OMH ve HKOK esas alınmıştır ve Çizelge 4.2' de gösterilmiştir. CatBoost regresyon ve LigthGBM algoritması en yüksek R^2 ve en düşük HKO, OMH ve HKOK oranlarına sahiptir.

Çizelge 4.2. Makine öğrenmesi algoritmalarının başarı metriklerine etkisi

Model	OMH	HKO	HKOK	R2
Catboost	11.550	262.821.209	16.186	0.8197
Lightgbm	11.647	265.103.430	16.256	0.8177
Rastgele Orman	11.781	287.071.773	16.907	0.8034
Aşırı Ağaçlar Regresyonu	11.971	300.952.271	17.311	0.7932
Xgboost	12.422	302.741.452	17.352	0.7921
Gradyan Artırma Regressor	14.529	414.351.797	20.326	0.7165
Karar Ağaçları	15.314	482.014.533	21.876	0.6702
Doğrusal Regresyon	17.556	542.562.764	23.269	0.6275

Eğitim veri seti, modelin öğrenmesi için kullanılırken, performansını değerlendirmek amacıyla test veri seti kullanılmıştır. Ancak, bu yaklaşımın getirdiği farklı test veri setleriyle kaynaklanan dağılım problemlerini ele almak için K-Katlı Çapraz Doğrulama yöntemi benimsenmiştir. Bu süreçte, veri seti çok daha fazla grubu değerlendirmek adına 10 farklı gruba bölünmüştür. Algoritma, 10 kez eğitilip test edilmiştir. Her iterasyonda (fold) modelin performansı kaydedilmiş ve toplam performansın ortalaması alınmıştır. K-Katlı Çapraz Doğrulama, hiperparametre optimizasyonunda daha iyi bir rehberlik sağlar. Çünkü farklı hiperparametre ayarlarının performansı farklı alt kümeler üzerinde değerlendirilir ve en iyi hiperparametrelerin seçilmesine yardımcı olur. Catboost modelinin performansının daha güvenilir bir şekilde değerlendirildiği, Çizelge 4.3'de gösterilmiştir.

Çizelge 4.3. Catboost algoritması için K katlamalı çapraz doğrulama model performansı

Fold	OMH	HKO	HKOK	R2
0	12.286	271.494.207	16.477	0.8216
1	11.329	242.930.770	15.586	0.8149
2	11.088	231.799.055	15.225	0.8638
3	11.450	249.844.381	15.806	0.8067
4	12.489	280.595.457	16.751	0.7937
5	11.229	240.435.672	15.506	0.8231
6	11.331	266.513.309	16.325	0.8121
7	11.925	281.109.117	16.766	0.8119
8	12.818	354.765.848	18.835	0.7938
9	11.639	262.390.354	16.198	0.8186
Mean	11.758	268.187.817	16.348	0.8160
Std	563	33.064.687	970	0.0187

Değişkenler, makine öğrenimi algoritmalarının çıktı etiketlerini tahmin etmek için veri setinde kullanılan belirli yapılandırmalardır. Modellerin etkili bir şekilde çalışabilmesi ve başarılı sonuçlar elde edebilmesi için, seçilen algoritmaların optimal performansını sağlayacak hiperparametre değerleri belirlenmelidir. Bu amacı gerçekleştirmek üzere, tüm olası kombinasyonları denemek için GridSearch yöntemi uygulanmıştır. Gerçekte kayıp fonksiyonu veya temel gösterge seçimi araştırmacının hedefine, modeli nasıl kullanmak istediğine göre değişmektedir. Hiperparametre seçimi R^2 ve HKOK performans ölçütleri optimize edilecek şekilde belirlenmiştir. HKOK, türev alınabilir olduğu için kolayca minimum veya maksimum değerleri belirlemeye yardımcı olur. Ayrıca, simetrik ve karesel bir kayıp fonksiyonu olması nedeniyle tercih edilmiştir (Yalçınkaya, 2022). Optimizasyon sonuçları birbirine çok yakın çıksa da HKOK performans ölçütünün optimizasyonu daha iyi sonuçlar vermiştir.

Varsayılan parametrelerle daha iyi çalışan ilk iki algoritma üstünde hiperparametre optimizasyonu yapılmıştır. Hiperparametre seçimi modülleri, HKOK değerini optimize etmek üzere tasarlanmıştır. CatBoost ve Lightgbm algoritmaları için varsayılan parametrelerle ve hiperparametre optimizasyonu yapılarak elde edilen sonuçlar Çizelge 4.4'te gösterilmiştir.

Çizelge 4.4. CatBoost algoritması ve Ligthgbm algoritması için kullanılan parametreler

Model	**catboost**		**lightgbm**	
	Varsayılan	**hiperparametre**	**Varsayılan**	**hiperparametre**
Parametreler	iterations=1000	iterations=1000	num_leaves: 29	num_leaves: 31
	learning_rate=0.03	learning_rate=0.0445	learning_rate: 0.1	learning_rate: 0.1
	depth=6	depth=6	max_depth: -1	max_depth: -1
	l2_leaf_reg=3	l2_leaf_reg=3	min_data_in_leaf: 20	min_data_in_leaf: 20
	model_size_reg=0.5	model_size_reg=0.5	subsample: 1	subsample: 1
	border_count=254	border_count=254	lambda_l1: 0	lambda_l1: 0
	thread_count=-1	thread_count=0	lambda_l2: 0	lambda_l2: 0
	random_seed=0	random_seed=7237	silent:true	silent:warn
			random state:none	random_state:7237
OMH	11.549	10.292	11.647	10.562
HKO	262.821.209	204.571.653	265.103.430	216.076.318
HKOK	16.185	14.303	16.256	14.700
R2	0,8197	0.8534	0.8177	0.8451

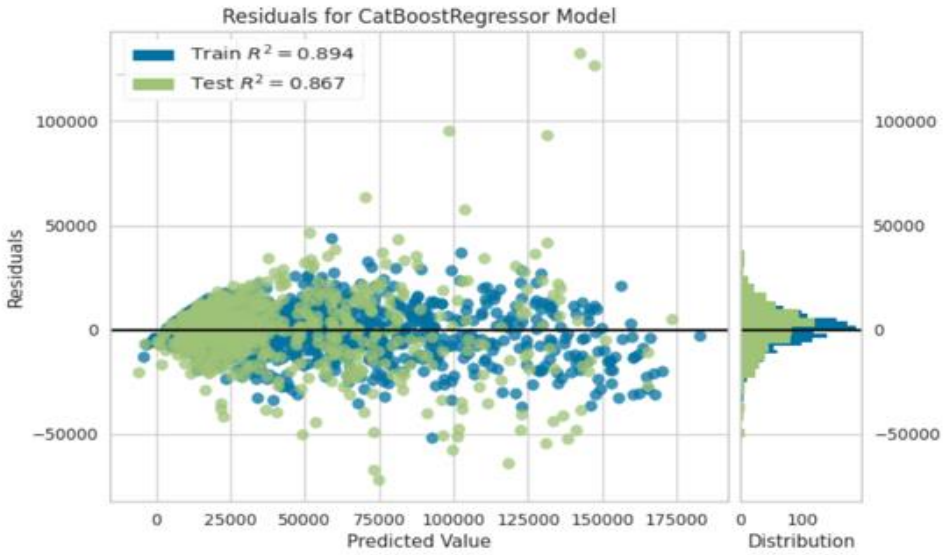
Hiperparametre optimizasyonu yapıldıktan sonra oluşturulan model ilk başta ayrılan test seti için çalıştırılmış ve tahmin sonuçları detaylı bir şekilde Şekil 4.6' da sunulmuştur.

	Model		MAE		MSE		RMSE		R2
	CatBoost Regressor		10292.2483		204571652.5368		14302.8547		0.8534
	rumuz	hafta	denye	renk	dolar	fiyat	uruncinsi	miktar	prediction_label
0	0150f048R	H1	D031	14	7.35390	1.82	TNRYO	91781	71645.655088
1	0100f036R	H1	D016	1	7.35390	1.69	TNRYO	85637	103047.158839
2	0100f096R	H1	D018	1	7.35390	2.04	TIMSFYO	43389	44854.423236
3	0075f072R	H1	D012	1	7.35390	2.04	TIMSFYO	34317	37475.587585
4	0150f096R	H1	D033	1	13.38846	2.72	TNRINYO	91154	68144.172866
...	...	...	...	...	...	...	...	...	...
603	0300f096T	H52	D057	1	18.67968	2.44	TASSLI	18968	14922.400335
604	0150f288R	H52	D089	1	18.67968	2.57	TICSLY	7624	11654.036592
605	0300f096T	H52	D057	1	18.67968	2.29	TNRIN	20217	24655.864757
606	0300f072R	H52	D056	1	18.67968	1.99	TNRSC	4912	3053.626163
607	0150f048T	H52	D031	1	18.67968	2.44	TNRIN	3132	11117.899972

Şekil 4.6. Hazır hale gelen modelin test veri seti üzerindeki tahmin sonuçları

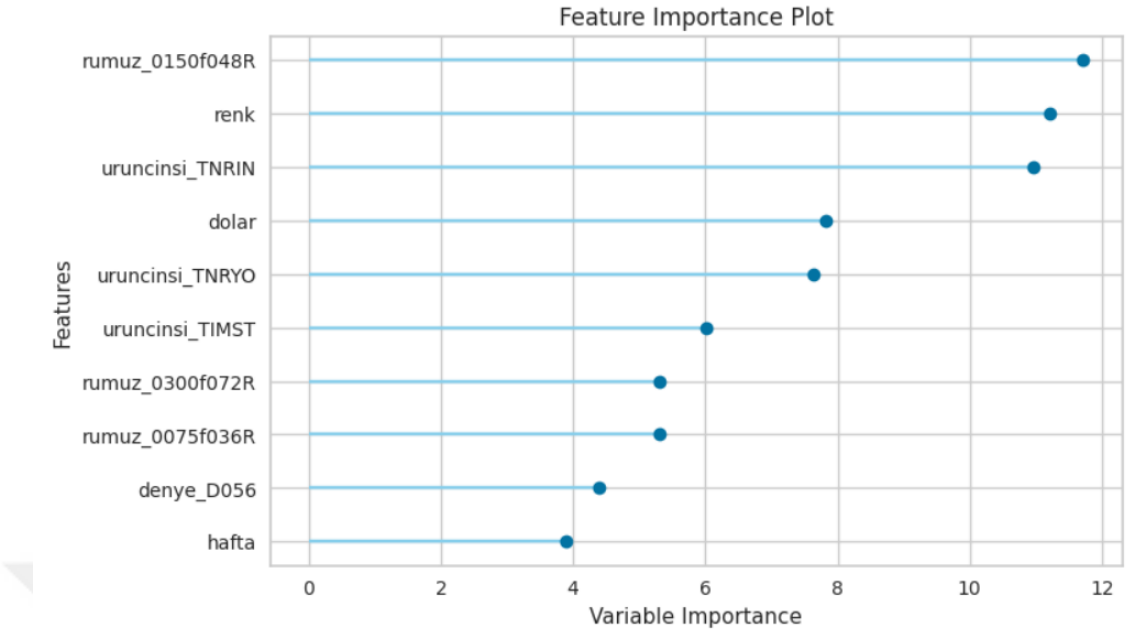
Doğru hiperparametre değerleri seçilerek modelin performansında artış, daha hızlı eğitim süreleri ve daha iyi genelleme yeteneği ile daha güvenilir ve tutarlı tahminler elde edilmiştir. CatBoost regresyon algoritmasının R^2 değeri 0.8201'den 0.8534 değerine yükselmiş ve HKOK değeri 16186'dan 14303' e düşmüştür. Diğer başarı metriklerinde de daha iyi sonuçlar elde edilmiştir.

CatBoost regresyon modelinin eğitim verisi ve test verisi için tahminlerinin dağılımı Şekil 4.7' da gösterilmiştir. Grafiğin x ekseninde, örneklerin endeksleri veya benzeri bir sıralama bulunurken, y ekseninde eğitim verisi ve test verisinin tahminleri arasındaki farklar (artıklar veya hata değerleri) bulunur. Eğer bir nokta sıfıra yakınsa, bu, modelin o belirli örneği test ve eğitim veri seti için doğru bir şekilde tahmin ettiği anlamına gelir. Ancak, eğer nokta sıfıra yakın değilse, bu, modelin o örneği yanlış tahmin ettiğini gösterir.



Şekil 4.7. CatBoost regresyon modelinin tahminlerinin dağılım grafiği

Bir özellik önem sırası grafiği (feature importance plot) Şekil 4.8'de gösterilen, bir makine öğrenimi modelinin, her bir özelliğin hedef değişken üzerindeki etkisini görselleştirmek için kullanılmıştır. Bu grafik, modelin hangi özelliklerin tahmin yeteneğinde daha büyük bir rol oynadığını anlamak için kullanılır. Rumuz olarak 0150f048R ve ürün cinsi olarak TNRIN değerlerinin final model tahmininde etkisi büyüktür. Fiyat ve dolar değişkenlerinin de hedef değişken üzerinde etkili olmuştur.



Şekil 4.8. Özellik önem sırası grafiği

Model oluşturulduktan sonra artık kullanıma hazır hale getirilmiştir. 10 farklı veriyle tahmin yapılmıştır ve sonuçları Şekil 4.9' da gösterilmiştir.

	rumuz	hafta	denye	renk	dolar	fiyat	uruncinsi	prediction_label
0	0300f072R	H1	D056	1	18.709999	1.99	TNRIN	89866.341560
1	0300f072R	H1	D056	1	18.709999	2.14	TIMST	58100.186653
2	0150f048R	H1	D031	1	18.709999	2.04	TIMSL	74821.043908
3	0300f096T	H2	D057	1	18.750000	2.24	TNRIN	27464.611506
4	0075f036R	H3	D011	14	18.780001	2.83	TNRYO	8701.040355
5	0150f144R	H3	D034	1	18.780001	2.29	TIMSFYO	21572.350781
6	0150f048T	H4	D031	1	18.780001	2.49	TNRIN	17904.894023
7	0150f048R	H4	D031	1	18.780001	1.89	TNRYO	87679.613953
8	0150f048R	H4	D031	14	18.780001	2.32	TNRYO	33324.243234
9	0300f096T	H1	D057	1	18.709999	2.54	TASSLI	21277.207400

Şekil 4.9. Makine öğrenmesi modelinin kullanılması

CatBoost regresyon algoritması, kategori değişkenleriyle başa çıkma yeteneği, hızlı ve etkili çalışma, dengesiz veri kümeleriyle başa çıkabilme yetisi sayesinde başarılı sonuçlar elde edilmiştir. Eğitilmiş CatBoost modeli, envanter yönetim sistemine entegre edilir. Bu, tahminlerin otomatik olarak alınıp kullanılabileceği bir arayüz sağlayabilir veya bir API üzerinden entegrasyon gerçekleştirilebilir. Model envanter yönetim sistemine haftalık güncel talep tahminlerini sağlar. Bu tahminler, stok seviyelerini belirlemek, sipariş miktarlarını optimize etmek ve envanter maliyetlerini düşürmek için kullanılabilir. Modelin performansını düzenli olarak gözden geçirip gerektiğinde güncellenir. Talep tahmininde kullanıldığında envanter yönetim süreçlerine daha iyi planlama, daha hızlı güncellemeler ve daha kesin tahminler yapma yeteneği kazandırabilir. Bu da envanter maliyetlerini azaltabilir, stok seviyelerini optimize edebilir ve müşteri taleplerine daha etkili bir şekilde yanıt verebilir.

5. SONUÇ

Tekstil sektörü, genellikle mevsimsel ve modaya bağlı talep değişimleriyle karşılaşır. Bu dinamik yapının yanı sıra, hammadde tedariki için uluslararası bağlantılar ve pazar belirsizlikleri, doğru talep tahminini zorlaştırabilir. Yeni pazarlara açılma veya daralmalar gibi faktörler de talep üzerinde belirsizlik yaratabilir. Bu durumda, yanlış tahminler fazla stok veya talebi karşılayamama gibi sorunlara yol açabilir, bu da şirketin maliyetlerini artırabilir. Çalışma, tekstil şirketinin ürün taleplerini daha doğru bir şekilde tahmin etmek amacıyla makine öğrenmesi algoritması kullanılarak talep tahmin modeli geliştirilmiştir.

Bu araştırma, 2021-2022 yıllarında bir tekstil sektörü firmasına ait satış verilerini kullanarak, makine öğrenimi yöntemini uygulayarak değişkenleri analiz etmeyi ve talep tahmin modeli için temel oluşturacak model performanslarını değerlendirmeyi amaçlamaktadır. Çalışma kapsamında, firmanın 2 yıllık satış verileri kullanılarak bir veri kümesi oluşturulmuş ve bu veriler Excel üzerinde analiz edilerek, makine öğrenimi için daha anlamlı bir forma dönüştürülmüştür. Sonrasında veri kümesine ve çalışmanın amacına en uygun algoritmalar seçilmiştir. En iyi sonuçları veren algoritmalar tespit edilmiştir. Bunlar CatBoost, LightGBM, XGboost, Aşırı Ağaçlar Regresyonu, Rastgele orman algoritmaları olmuştur. Hiperparametre optimizasyonu sayesinde, makine öğrenimi modellerinin veri kümesini daha etkili bir şekilde değerlendirip yorumlaması sağlanmıştır. Böylece; en iyi sonucu veren CatBoost regresyon algoritması R^2 değeri 0.8201'den 0.8534 değerine yükselmiştir.

Bu çalışma bağlamında gözlemlenen kaynaklar, benzer problemler ve veri setleri üzerinde yapılan incelemelere göre, CatBoost algoritmasının diğer algoritmalara kıyasla daha etkili bir performans sergilediğini göstermektedir. Ek olarak, dengesiz dağılıma sahip veri setleri için tahmin başarısını artırmak amacıyla kullanılan veri dengeleme ve hiperparametre optimizasyonu yöntemlerinin önemli olduğu tespit edilmiştir. Sonuç olarak, makine öğrenmesi yöntemleri ile oluşturulan talep tahmin modelinin, tekstil sektöründeki talebi tahmin etmede kullanılabileceği anlaşılmıştır. Tekstil sektöründe makine öğrenmesi yöntemleri, geleneksel üretim süreçlerine yenilikçi ve verimli

çözümler sunma potansiyeli taşımaktadır. Bu çalışma, işletmelerin üretim, stok yönetimi ve talep tahmini gibi kritik süreçlerde daha etkili ve optimize edilmiş kararlar almasına olanak sağlamaktadır. Gelecekteki çalışmalarda, araştırmacılar, evrensel verilerle veri setini genişleterek daha kapsamlı bir talep tahmin modeli geliştirmek ve bu çalışmanın bulgularını daha da güçlendirmek için çaba sarf edebilirler.

KAYNAKLAR

Adalı, E. (2020). Makina İmalat Sanayiinde Talep Tahmini: Elektromekanik Sanayiinde Bir Uygulama. Yüksek Lisans Tezi. T.C. Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı.

Al-Khazraji H., Raoof Nasser A. ve Khlil S. (2022). An intelligent demand forecasting model using a hybrid of metaheuristic optimization and deep learning algorithm for predicting concrete block production. IAES International Journal of Artificial Intelligence (IJ-AI). 11. 649. 10.11591/ijai. v11.i2.pp649-657.

Alon I., Qi M. ve Sadowski R. (2001). Forecasting aggregate retail sales: A comparison of artifcial neural networks and traditional methods. Journal of Retailing and Consumer Services. 8. 147-156. 10.1016/S0969-6989(00)00011-4.

Aydın, C. (2018). Makine öğrenmesi algoritmaları kullanılarak itfaiye istasyonu ihtiyacının sınıflandırılması. *Avrupa Bilim ve Teknoloji Dergisi*, (14), 169-175.

Aylak, B. L, Oral, O. Yazıcı, K. (2021), "Yapay Zekâ ve Makine Öğrenmesi Tekniklerinin Lojistik Sektöründe Kullanımı" El-Cezerî Fen ve Mühendislik Dergisi, 8 (1); 74-93.

Bentejac C., Csörgö A. ve Martinez-Munoz G. (2021). A comparative analysis of gradient boosting algorithms. Artificial Intelligence Review, 54, 1937-1967

Brahmadeep, & Thomassey, S. (2016). 8 - Intelligent demand forecasting systems for fast fashion. Information Systems for the Fashion and Apparel Industry, 145–161.

Chaudhuri, T. D., Ghosh, I., & Singh, P. (2017). Application of Machine Learning Tools in Predictive Modeling of Pairs Trade in Indian Stock Market. IUP Journal of Applied Finance, 23(1), 5–25.

Comlan, M., & Koulo, E. (2022). Sales Forecast and Design Generation for Textile Products Using Machine Learning. In Future of Information and Communication Conference (pp. 183-197). Cham: Springer International Publishing.

Çelik S., Çetinkaya Bozkurt Ö. & Ekşili N. (2022). Çalışan Performansı Ölçeğindeki İfadelerin Karar Ağacı Algoritması ile Belirlenmesi. *Mehmet Akif Ersoy Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 9(1), 561-584. doi: 10.30798/makuiibf.913478

Daldır I. (2021). Havacılık Sektöründe Uçuş Gecikmelerinin Makine Öğrenmesi Temelli Analizi Ve Tahmini. Akdeniz Üniversitesi Sosyal Bilimler Enstitüsü

Ergül Aydın Z., İçmen Erdem B. ve Erzurum Çiçek Z. (2023). Gradyan artırma yöntemleriyle bisiklet paylaşım talebini tahmin etme. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 29(8), 824-832.

Ersöz F., Çınar Y. (2021). Veri Madenciliği ve Makine Öğrenimi Yaklaşımlarının Karşılaştırılması: Tekstil Sektöründe bir Uygulama. Avrupa Bilim ve Teknoloji Dergisi, (29), 397-414.

Giri C., Thomassey S., Balkow J. ve Zeng X. (2019). Forecasting New Apparel Sales Using Deep Learning and Nonlinear Neural Network Regression. 2019 International Conference on Engineering, Science, and Industrial 90 Applications (ICESI), Engineering, Science, and Industrial Applications (ICESI), 2019 International Conference On, 1–6.

Görçün, Ömer Faruk (2017). Depo ve Envanter Yönetimi. İstanbul: Beta.

Güven, İ. (2020). Perakende Hazır Giyim Endüstrisinde Yapay Zeka Yöntemleri İle Talep Tahmini. Yüksek Lisans Tezi, Karabük Üniversitesi.

Harper, Ralph L. (2010). Warehouse Technology in the Supply Chain Management Systems. Institute of Electrical and Electronics Engineers, 6-10.

Kang Z., Zuo Y., Huang Z., Zhou F. ve Chen P. (2017). Research on the Forecast of Shared Bicycle Rental Demand Based on Spark Machine Learning Framework. 2017 16th International Symposium on Distributed Computing and Applications to Business, Engineering and Science (DCABES), Distributed Computing and Applications to Business, Engineering and Science (DCABES), 2017 16th International Symposium on, DCABES, 219–222.

Karaman Ö. A., Bektaş Y. (2023). Makine Öğrenmesi ve Optimizasyon Yöntemleri ile Uzun Dönem Elektrik Enerjisi Tahmini: Türkiye Örneği, Müh. Bil. ve Araş. Dergisi, 5(2) 285-292

Karasu S., Altan A., Saraç Z. ve Hacıoğlu R. (2018). "Bitcoin fiyatlarının zaman serisi verileri kullanılarak makine öğrenmesi yöntemleriyle tahmin edilmesi". 26. Sinyal İşleme ve İletişim Uygulamaları Konferansı (SIU). İzmir, Türkiye. s. .1-4, doi: 10.1109/SIU.2018.8404760.

Kaysal, K., Akarslan, E., & Hocaoğlu, F. O. (2022). Türkiye Kısa Dönem Elektrik Yük Talep Tahmininde Makine Öğrenmesi Yöntemlerinin Karşılaştırılması. Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Dergisi, 9(2), 693-702. https://doi.org/10.35193/bseufbd.1004827

Kaysal, K., Yurttakal, A. H., & Hocaoğlu, F. O. (2023). Hibrit derin öğrenme yöntemi kullanılarak hiperparametre optimizasyonu ile yenilenebilir elektrik enerjisi tahmini. *Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, *12*(3).

Kharfan, M., Chan, V. W. K., & Firdolas Efendigil, T. (2021). A data-driven forecasting approach for newly launched seasonal products by leveraging machine-learning approaches. Annals of Operations Research, 303(1-2), 159-174.

Kück M., Freitag M. (2021), Forecasting of customer demands for production planning by local k-nearest neighbor models. Int. J. Prod. Econ. 231, 107837. https://doi.org/10.1016/j.ijpe.2020.107837

Liu N., Ren S., Choi T. M., Hui C. L. ve Ng S. F. (2013). Sales forecasting for fashion retailing service industry: a review. Mathematical Problems in Engineering.

Lorente-Leyva, L. L., Alemany, M. M. E., Peluffo-Ordóñez, D. H., & Araujo, R. A. (2021). Demand forecasting for textile products using statistical analysis and machine learning algorithms. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12672 LNAI, 181–194. https://doi.org/10.1007/978-3-030-73280-6_15

Loureiro A.L.D., Miguéis V.L., da Silva L.F.M. (2018). Exploring the use of deep neural networks for sales forecasting in fashion retail. Decis. Support Syst. **114**, 81–93 . https://doi.org/10.1016/j.dss.2018.08.010

Medina H., Peña M., Siguenza-Guzman L., Guamán R. (2022). Demand Forecasting for Textile Products Using Machine Learning Methods. In: Botto-Tobar, M., Montes León, S., Torres-Carrión, P., Zambrano Vizuete, M., Durakovic, B. (eds) Applied Technologies. ICAT 2021. Communications in Computer and Information Science, vol 1535. Springer, Cham. https://doi.org/10.1007/978-3-031-03884-6_23

Öztürk S. (2020). Hızlı Moda Sektöründe Makine Öğrenmesi Yöntemleri İle Satış Miktarlarının Tahmin Edilmesi. Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü

Peker M., Özkaraca O., & Kesimal B. (2017)). Enerji Tasarruflu Bina tasarımı İçin Isıtma ve Soğutma Yüklerini Regresyon Tabanlı Makine Öğrenmesi Algortimaları ile Modelleme. Bilişim Teknolojileri Dergisi.10(4),443-449. https://doi.org/10.17671/gazibtd.310154

Polamuri, S. R., Srinivas, K., & Krishna Mohan, A. (2019). "Stock Market Prices Prediction using Random Forest and Extra Tree Regression.". International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878

Prokhorenkova L., Gusev G., Vorobev A., Dorogush A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. In *Advances in neural information processing systems* (pp. 6638–6648).

Pujara A., Pattabiraman V., Parvath R. (2022). Food Demand Forecast for Online Food Delivery Service Using CatBoost Model. In: Haldorai, A., Ramu, A., Mohanram, S., Lu, J. (eds) 3rd EAI International Conference on Big Data Innovation for Sustainable Cognitive Computing. EAI/Springer Innovations in Communication and Computing. Springer, Cham. https://doi.org/10.1007/978-3-030-78750-9_9

Saatçioğlu, D. (2016). Aralıklı Talep Yapısına Sahip Ürünlerin Talep Tahmininde Makine Öğrenme Yöntemlerinin Uygulanması, Doktora Tezi, T.C. İstanbul Üniversitesi Sosyal Bilimler Enstitüsü İşletme Ana Bilim Dalı Üretim Yönetimi Bilim Dalı.

Shams M.Y., Tarek Z., Elshewey A.M., Hany M., Darwish A., Hassanien A.E. (2023). A Machine Learning-Based Model for Predicting Temperature Under the Effects of Climate Change. In: Hassanien, A.E., Darwish, A. (eds) The Power of Data: Driving Climate Change with Data Science and Artificial Intelligence Innovations. Studies in Big Data, vol 118. Springer, Cham. https://doi.org/10.1007/978-3-031-22456-0_4

Singh U., Rizwan M., Alaraj M., Alsaidan I. (2021). A Machine Learning-Based Gradient Boosting Regression Approach forWind Power Production Forecasting: A Step towards Smart Grid Environments. Energies. 14, 5196. https://doi.org/10.3390/en14165196

Tekin A.T, Sarı C. (2022). Mağaza Tabanlı Talep Tahmini: Topluluk Öğrenmesi Yaklaşımları ile Zaman Serisi Yaklaşımı Karşılaştırılması. International Engineering and Technology Management Summit 2022

Teucke M., Ait Alla A., El-Berishy N., Beheshti-Kashi S., Lütjen M. L. (2014). Forecating of Seasonal Apparel Products. https://doi.org/10.1007/978-3-319-23512-7_63

Ünlü V. (2021). Demand Forecastıng Wıth Machıne Learnıng Algorıthm: Implementatıon In A Productıon Facılıty. Ankara Yıldırım Beyazıt Unıversıty, Ankara.

Yalçınkaya, A., E. (2022). Tedarik Zinciri Yönetiminde Etik Ticaret Platformlarının Ve Global Optimizasyon Sürecinin Sürdürülebilirlik Açısından Etkileri. Doktora Tezi, Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü.

Yasir M, Ansari Y, Latif K, Maqsood H, Habib A, Moon J & Rho S (2022). Machine learning–assisted efficient demand forecasting using endogenous and exogenous indicators for the textile industry, International Journal of Logistics Research and Applications, DOI: 10.1080/13675567.2022.2100334