

## PERAKENDE SATIŞ ANALİZİ VE FİYAT OPTİMİZASYONU

Furkan KARAMAN

1. Sakarya Üniversitesi Bilgisayar Mühendisliği Bölümü, İstanbul Üniversitesi Yönetim Bilişim Sistemleri, İstanbul, furkan.karama22@gmail.com

### ÖZET

Bu makalede makine öğrenmesi teknikleri kullanılarak optimum fiyat tahmini yapma konusu incelenmiştir. Çalışma yapılırken bir toptancının 3 senelik satış veri kümesi incelenmiş ve gelecekteki 7 ayda optimum fiyatları tahmin edilmiştir. Bunu yaparken önce veri kümesinin optimizasyon için kullanılabilir hale gelmesi için veri kümesi veri düzenleme teknikleriyle düzenlendi, gelecekle ilgili tahmin yapmak için Autoregressive Integrated Moving Average (ARIMA) modeli, optimizasyon yapma kısmında ise parametrelerimizi rahat düzenleyebilmek adına Pulp kütüphanesi kullanılmıştır. ARIMA modeli kullanılmadan önce farklı zaman serisi modelleri de denenmiş ve incelenmiş ve performansı ve sağladığı avantajlardan dolayı ARIMA modeli üzerinde çalışılmak için tercih edilmiştir. Bu çalışmanın temel amacı firmalara satış fiyatı belirleme konusunda yardımcı olmaktır. Optimum fiyat tahminleri, müşterilerin ürünleri hak ettikleri fiyatlardan alabilmesini sağlarken, aynı zamanda firmaların maksimum kâr elde etmelerini ve ürünleri hak ettikleri değerde satış sunmalarını hedeflemektedir. Bu doğrultuda bu çalışma sürdürülebilir bir ekonomik büyümeyi desteklemeye yardımcı olmaktadır.

**Anahtar Kavramlar:** Fiyat Optimizasyonu, Veri Bilimi, Perakende

### ABSTRACT

The topic of anticipating optimal pricing using machine learning techniques is examined in this article. The study used an analysis of a wholesaler's 3-year sales record to predict the best cost for the following 7 months. To do this, the dataset was first optimized using data preparation methods. The Autoregressive Integrated Moving Average (ARIMA) model was employed to make predictions for the future, and the Pulp library was used to quickly change the parameters throughout the optimization process. Before using the ARIMA model, different time series models were tried and examined, and due to its performance and the advantages it provides, working on the ARIMA model has been preferred. This study's primary goal is to support companies in setting sales pricing. The aim of optimum pricing forecasts is to allow consumers to buy goods at fair prices while also increasing the profitability of companies and providing goods at their real value. This study can help support sustainable economic growth in this way.

**Keywords:** Price Optimization, Data Science, Retail

### 1.Giriş

Sürekli olarak büyük miktarda veri toplanan ve depolanan günümüz için bulunduğumuz çağın bir veri tufanı çağı olduğuna şüphe yok [1]. Şu anda mevcut olan büyük miktarda veriyle, hemen hemen her sektördeki şirketler, rekabet avantajı için verileri kullanmaya odaklanıyor [2]. Verilerin hacmi ve çeşitliliği, manuel analiz kapasitesini çok geride bıraktı ve bazı durumlarda geleneksel veri tabanlarının kapasitesini aştı. Aynı zamanda, bilgisayarlar çok daha güçlü hale geldi, ağ her yerde bulunuyor ve daha önce mümkün olandan daha geniş ve daha derin analizler sağlamak için veri kümelerini birbirine bağlayabilen algoritmalar geliştirildi ve veri bilimi tahmin edilenden çok daha fazla öneme sahip oldu [2]. Bununla birlikte veri biliminin alt dallarından olan makine öğrenmesi, finans sektöründe optimizasyon teknikleriyle birlikte kullanıldığında büyük avantajlar sağlayabilir duruma geldi. Optimizasyon, bir sistemde var olan kaynakların en verimli şekilde kullanarak belirli amaçlara ulaşmayı sağlayan bir teknoloji olarak tanımlanmaktadır [3]. Günümüzde bir ürünün fiyatını belirlemek firmaların kâr marjı adına hayati önem taşımaktadır çünkü bir ürünün fiyatı çok yüksek ise hedeflenen sayıda satış yapılamaz ve bir adet üründen

yüksek kâr edile dahi hedeflenen kâra ulaşlamamış olur, örneğin Webvan adlı online alışveriş hizmeti sunan bir şirket aşırı büyüme stratejisi, yüksek maliyetler ve düşük fiyatlı ürünler nedeniyle kar edememesi sonucunda 2001 yılında iflas etti [4]. Ters doğrusallıkta da aynı durum yaşanacaktır çünkü bir ürünün fiyatı çok düşük ise hedeflenenin üzerinde satış yapılır fakat her bir adet üründen çok düşük bir kâr elde edildiği için yine hedeflenen kâra ulaşlamamış olur , örneğin 1985 yılında kurulan BlockBuster, video kiralama sektöründe bir dev olarak biliniyordu fakat şirket kiralama işlemi için yüksek ücretler talep ediyordu ve internet üzerinden film ve dizi platformlarının da popülerleşmesi ve bu platformlarının ucuza hizmete sunulmasıyla şirketin sonu gelmiş oldu [5]. Bu fiyatı belirlerken de birçok etmeni göz önünde bulundurarak hareket etmemiz gerekiyor. Bu faktörler arasında arz ve talep, rekabet durumu, enflasyon, mevsimsellik gibi unsurların yer aldığı söylenebilir [6]. Fiyatın doğru belirlenmesi, bu faktörleri dikkate alarak hareket etmeyi gerektirir [6]. Arz ve talep faktörü, bir ürünün arzının ne kadar olduğunu ve tüketicilerin talep etme düzeyini ifade eder. Rekabetin yoğun olduğu bir pazarda fiyatlar daha rekabetçi olabilir ve düşebilir [6]. Genel fiyat seviyesindeki artışlar, üretim maliyetlerini etkileyebilir ve bu da fiyatları etkileyebilir [6].

Veri biliminin günümüzde kullanılmadığı alan neredeyse yok denebilir haliyle firmalarda ürün fiyatlarının optimum yani maksimum kâr elde ettirecek tutarını belirlerken bu bilim dalından yararlanmaktadır [2].

Optimizasyonun tarihçesi oldukça eskidir. Antik Yunan döneminde matematiksek problemleri çözmek için bazı yöntemler geliştirilmiştir [3]. Ancak, modern optimizasyon tekniklerinin gelişimi daha yakın zamanlara dayanmaktadır. Optimizasyonun modern tarihçesi ise 20. Yüzyılın başlarına kadar uzanır [3]. Lineer programlama, 1940'larda George Dantzig tarafından geliştirilen ve birçok uygulama alanında kullanılan bir optimizasyon yaklaşımıdır. Daha sonra, non-lineer programlama, tam sayılı programlama, dinamik programlama, stokastik programlama gibi farklı optimizasyon yaklaşımları geliştirilmiştir [3]. Bilgisayar teknolojisinin ilerlemesiyle birlikte, optimizasyon algoritmalarının uygulanması daha da yaygınlaştı [3]. Günümüzde ise kullanılan metotlara örnek olarak doğrusal programlama metotları, tamsayı-karışık doğrusal programlama metotları, doğrusal olmayan programlama metotları ve tamsayı-karışık doğrusal olmayan programlama metotları verilebilir [3]. Bilgisayar tabanlı optimizasyon yazılımları ve matematiksel modelleme araçları, karmaşık problemlerin çözümünde büyük bir rol oynamaktadır [3].

Messica ve Rokach(2020) Machine learning and operation research based method for promotion optimization of products with no price elasticity history çalışması fiyat optimizasyonu üzerine yapılmış bir çalışmadır bu çalışmada fiyat esnekliği geçmiş olmayan ürünler için fiyat optimizasyonu ile ilgili bir çalışma yapılmıştır ve bu makaledeki çalışmadan en bariz ayrıldıkları nokta budur geçmiş verilerden yararlanılmayarak yapılan bir çalışmadır ve bu ayrım yüzünden kullanılan yöntemlerde farklılık göstermektedir [7]. Örneğin bu çalışmada doğrusal olmayan programlama kullanılmışken bu makaledeki çalışmada doğrusal programlama yöntemi kullanılmıştır [7].

Rios ve Vera (2023) Dynamic pricing and inventory control for multiple products in a retail chain çalışması da fiyat optimizasyonu üzerine yapılmış bir çalışmadır bu çalışmada bir perakende zincirindeki birden fazla ürün için dinamik fiyatlandırma ve envanter kontrolü yapılmıştır [8].Bu çalışmanın bu makaledeki çalışmadan farklı yanı her Pazar için farklı bir fiyat tahmini yapıyor olması ve sadece elde edilen güncel veriler ile bunu yapıyor olmasıdır çünkü bu makaledeki çalışmada fiyat optimizasyonuna ek olarak ilerdeki 7 ay içinde tahminler üzerinden fiyat optimizasyonu yapılmıştır [8].

Bu çalışmaya benzer başka bir çalışmada ise Namli vd. (2019) yaptığı fiyat tahminlemede makine öğrenmesi teknikleri ve doğrusal regresyon yöntemlerinin kıyaslanması; Türkiye'de satılan ikinci el araç fiyatlarının tahminlemesine yönelik bir vaka çalışması ile ikinci el araç piyasasındaki sıfır araçtan daha pahalı ikinci el araç veya parametreler bazında daha kötü durumda olan aracın daha iyi durumda olan araca göre daha yüksek fiyata satılma problemlerine çözüm aranmıştır. Optimizasyon için doğrusal regresyon yöntemi kullanılmıştır. İkinci el araç için optimum fiyat tespiti yapılmıştır [9].

Fiyat optimizasyonu yaparken veri bilimi firmalara optimizasyon algoritmaları ve istatistiksel yöntemlerle birlikte karar aşamalarında büyük bir güven ve destek sağlamaktadır.[10] Bu güvene ve desteğe sahip olmak isteyen bütün firmalar bu disiplinden yararlanmak istemektedir [10].

Sonuç olarak, günümüzde veri bilimi, fiyat optimizasyonunda kullanılmadığı neredeyse hiçbir alan kalmamıştır.[2] Firmalar, veri bilimi ve optimizasyon algoritmalarını kullanarak ürün fiyatlarını optimum düzeyde belirleyerek maksimum kar elde etmeyi hedeflerken müşteriler ise aldıkları ürünleri adil fiyattan alma şansına sahip olurlar bütün bunların ışığında hem müşteri memnuniyeti olurken hem de müşteri memnuniyeti müşteri sadakatini peşinde getireceği için uzun vadeli kazanç elde edilmiş olur [11].

## 2.Verİ Kümesi ve Analizi

Kullanılacak veri kümesi bir Perakende Satış Veri kümesi (Retail Data Set) [9]. Bu veri kümesi bir toptancının 3 yıllık satışlarıyla ilgili verileri içerisinde barındıran bir veri kümesidir. Veriyi anlamak, analiz etmek ve düzenlemek optimizasyon modelini kurmadan önce kesinlikle yapılması gereken aşamalardır [12].

## 2.1. Veriyi Açıklama

Veri kümesinin sütunlarını inceleyip ne olduklarını açıklamak gerekirse ; InvoiceID adında bir sütun var bu sütun ilgili ürünün hangi fişe ait olduğunu belirlemek için kullanılıyor , Date adında diğer bir sütun hangi satışın hangi tarihte veya ilgili satır bilgilerinin hangi tarihe ait olduğunu belirten sütun diyebiliriz , ProductID adında diğer bir sütun her bir ürünün benzersiz id numaralarını içinde barındıran sütun diyebiliriz , diğer bir sütun adı ise TotalSales bu sütun , ilgili satırda ki toplam satışın fişe yansıyan miktarı ile ilgili bilgi veren sütundur diyebiliriz, Discount adında ki diğer bir sütun ise ilgili satırda ki toplam satış miktarına uygulanan toplam indirim tutarı hakkında bilgi vermektedir , CustomerID adında ki diğer bir sütun ise ilgili satırda yapılan satışın hangi müşteriye ait olduğu ile ilgili bilgi veren sütundur ve veri kümesine ait olan son sütunun adı ise Quantity, bu sütun ilgili satırda toplam ilgili üründen kaç adet satış yapıldığı hakkında bilgi veren bir sütundur [13]. Tablo 1’de veri kümesinin bütün kolonları ve seçilen ilk 5 satırı örnek olarak sunulmuştur.

Tablo 1. Açıklanan veri kümesinin bütün sütunlarını ve ilk 5 satırı ile son 5 satırını göstermektedir.

Invoice ID	Date	ProductID	TotalSales	Discount	CustomerID	Quantity
328	2019-12-27	1684	796.610169	143.389831	185	4
329	2019-12-27	524	355.932203	64.067797	185	2
330	2019-12-27	192	901.694915	162.305085	230	4
330	2019-12-27	218	182.754237	32.895763	230	1
330	2019-12-27	247	780.101695	140.418305	230	4

## 2.2. Veri Analizi ve Düzenleme

Veri analizi, veri içerisinden, gizli kalmış, değerli, kullanılabilir bilgileri açığa çıkarmak ve stratejik karar destek sağlamak amacıyla kullanılan bir tekniktir [12]. Veri analizi, istatistiksel yöntemler, veri madenciliği teknikleri ve görselleştirme araçlarını kullanarak verilerin anlamlı bilgilere dönüştürülmesini sağlar [12].

### 2.2.1 Kullanılabilir Veri Elde Etme

Kullanılan veri kümesinde modeli veya algoritmaları hatalı işlem yapmaya yönlendirecek verileri yani standart sapması genel olarak 5 ve 10 arasında olan verilerin arasından standart sapması 30-40 civarlarında yüksek standart sapması olan veya modellerin ve algoritmaların kullanılabilmesi için bir özelliği barındıran yetersiz sayıda veri olması yani toplam veri setinin en az %2’sinin bu özellikleri barındırmaması bu verileri bir şekilde düzenlenmesi anlamına geliyor [14]. Veriler birkaç farklı yolla düzelebiliyorlar; veri kümesinden verileri kaldırmak, makine öğrenmesi algoritmaları ile hatalı verileri düzeltmek gibi metotlar veri düzenlemede kullanılabilir [12].

Kullanılan veri kümesinde bazı ürünlerin satışı veri kümesinde sadece 3-4 satır kadardı yani çok az sayıda idi ve bu çok az sayıda satışı olan ürünler veri kümesinden çıkarıldı.

Yine kullanılan veri kümesinde bazı ürünlerin satış fiyatı 0 olarak gözüküyordu bildiğim kadarıyla verilerin sahibi firma ile görüşme yapıp soruldu ve bunların bazılarının müşterilere promosyon hediyesi olarak verildiği için bazılarının ise eklenip sonradan faturadan çıkarıldığı için satış fiyatı veri kümesinde 0 olarak gözüküyordu bu veriler de daha doğru sonuçlar alınabilmek için veri kümesinden çıkarıldı.

Veri kümesine indirim yüzdesi adında bir sütun eklenildi çünkü bu sütunda ki verilerin optimizasyon probleminde de kullanılabileceği düşünüldü çünkü optimizasyon problemlerinde matematik işlemleri yapılırken oransal bir ifadenin daha yardımcı olacağı tespit edildi hem de bazı ürün satışlarında normalden çok daha fazla indirim yapılma ihtimali düşünüldü bunun tespitinin ise sadece oransal olarak kontrol edilerek yapılacağından veri kümesine indirim yüzdesi sütunu eklendi ve düşünülen gibi bütün indirim yüzdeleri %15 civarlarındayken birkaç tane verinin %30-50 civarlarında olduğu tespit edildi ve bu veriler de veri kümesinden çıkarıldı .

Veri kümesindeki ürünlerin 2019 yılından 2023 yılına kadar satışı yapılmış olmasının doğru sonuca yaklaşma konusunda fayda sağlayacağı düşünüldü çünkü son yıllarda satışı yapılmamışsa eğer bu ürün raftan kalkmış olabilirdi veya veri kümesinin ilk yıllarında satışı yok ama son yıllarda var ise bu ürün içinde analiz yapmak ve algoritmalarımızda kullanmak veri yetersizliği sebebi ile çalışmayı yanlış sonuçlara yönlendirebilme tehlikesi içeriyordu [23] ve bazı ürünlerin 2020, 2021 ve 2022 yıllarından sonra satışı gözüküyordu bu yüzden bu ürünler veri kümesinden çıkarıldı.

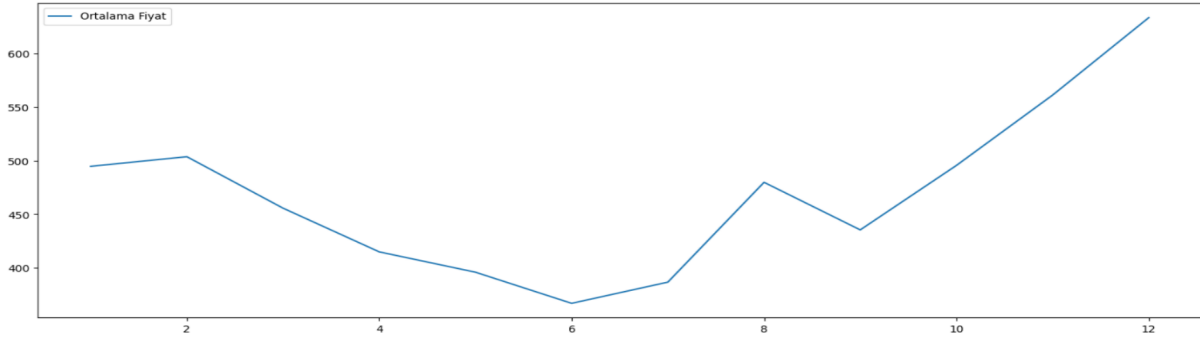
Optimizasyon işlemi uygulanırken çalışmanın asıl ulaşmak istediği veri olan ürünün bir adet fiyatı bilgisini veri kümesi içerisinde içermiyordu bu yüzden toplam satış tutarı haricinde sadece bir adet ürünün fiyat bilgisini veren grain\_price adında yeni bir sütun oluşturuldu.

### 2.2.2. Veri Görselleştirme ve Zaman Serisi Analizleri

İnsanın algılama sistemi yalnızca 3 boyutlu olduğu için çok boyutlu veriler üzerine düşünmesi gerektiğinde zorlanabiliyor bunun için çok boyutlu verileri görselleştirerek çok boyutlu veriyi aralarında bulunan ilişkiyi kaybetmeden 2 veya 3 boyuta indirmeye işlemine veri görselleştirme deniyor [1].

#### 2.2.2.1. Mevsimsellik

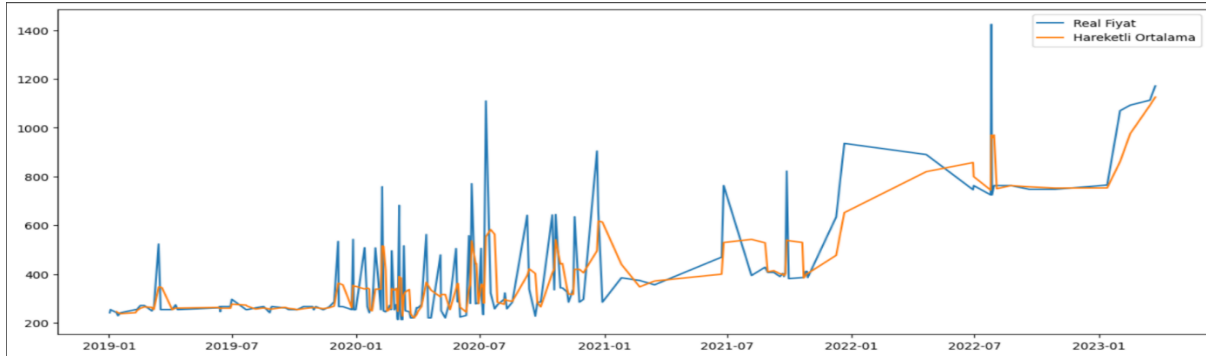
Kışın kullanılan mont, bot, eldiven gibi ürünlerin kışın daha pahalıya satılması yazın kullanılan güneş kremi, şort gibi ürünlerin yazın daha pahalı olması mevsimsellik ile açıklanır [15]. Bunun analizini yapabilmek adına 252 id numaralı ürünün bütün yıllardaki aylık fiyat ortalamaları bir grafik üzerinde gösterilmeye çalışıldı. Ortaya çıkan grafiği yorumladığımızda en yüksek fiyatların ekim, kasım, aralık, ocak, şubat aylarında olduğu tespit edildi. İlgili grafik şekil 1'de sunuldu.



Şekil 1. 252 id numaralı ürün için 4 yıllık mevsimsellik ile ilgili bilgi grafiği

#### 2.2.2.2. Hareketli Ortalama

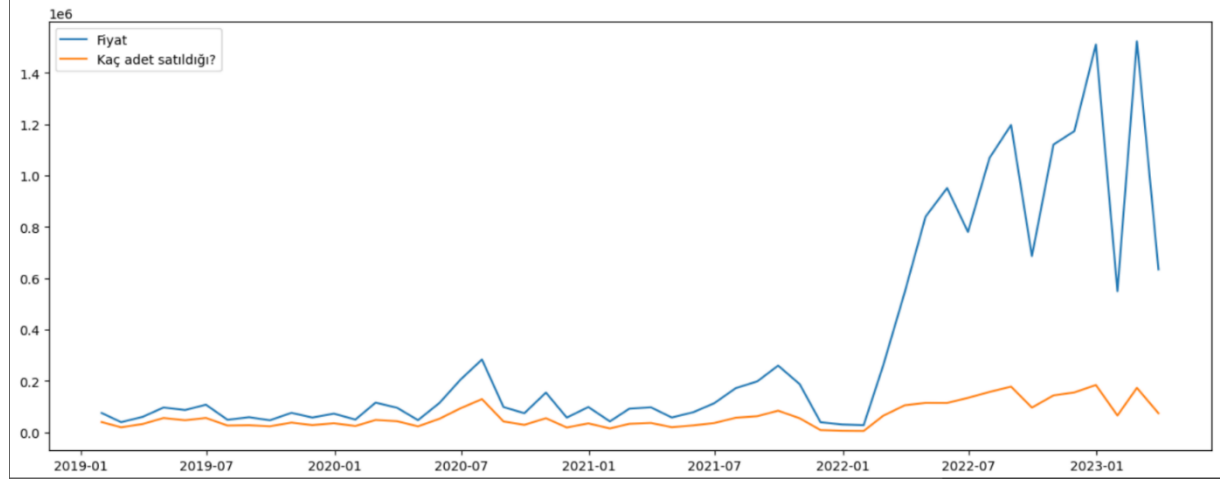
Veri kümesindeki ürünlerin fiyat istikrarının olup olmadığını yani kısa zaman önce satılan ürün fiyatı ile şimdiki fiyatı arasındaki benzerliği görebilmek için hareketli ortalama analizi tekniği kullanılabilir [16] ve bu tekniği Python dilinde koda dökmek için Rolling fonksiyonundan yardım alınabilir [16]. Bu teknik kullanıldığında ise bu teknik bir grafik üzerinde yorumlanmıştır ve bu grafiği yorumlarken gerçek fiyatı temsil eden çizgi ile hareketli ortalamayı temsil eden çizgi arasında en çok benzerlik olan noktalar fiyat istikrarının en yüksek olduğu; gerçek fiyatı temsil eden çizgi ile hareketli ortalamayı temsil eden çizgi arasında benzerliğin en az olduğu noktalar ise fiyat istikrarının da en az olduğu noktalar olarak tespit edilmiştir. İlgili grafik şekil 2'de sunulmuştur.



Şekil 2. 252 id numaralı ürün için 4 yıllık hareketli ortalama ile ilgili bilgi grafiği

#### 2.2.2.3. Enflasyon Analizi

Enflasyon , genel olarak mal ve hizmetlerin genel fiyat düzeyinde ki sürekli artışı anlamına gelir yani belirli bir zaman diliminde fiyat düzeyinde ki artış olarak kısaca açıklayabiliriz [17] fakat her ürün eşit ve düzenli şekilde enflasyona uğramaz yine bunu analiz edebilmek için aylık olarak toplam satış tutarını ve toplam ürün satış adetini ayrı ayrı toplayıp bir grafik çizdirilmiştir ve bu grafik yorumlandığında 2019'dan 2022 başlarına kadar dalgalanmalı ve düşük düzeyde enflasyon olurken 2022 yılı içerisinde enflasyon oranının çok yüksek noktalara geldiği tespit edilmiştir. İlgili grafik şekil 3'te sunulmuştur.



Şekil 3. 252 id numaralı ürün için 4 yıllık enflasyon analizi ile ilgili bilgi grafiği

### 3.VERİ KÜMESİ İÇİN SİMÜLAYON UYGULAMA

Simülasyon, bir araç, bir makine, bir sistem, bir olguya özgü işleyiş biçiminin incelenme, gösterilme ya da açıklanma amacıyla bir maket ya da bir bilgisayar programı aracılığıyla yapay bir şekilde yeniden üretilmesidir [18]. Simüle etmek, gerçek olmayan bir şeyi gerçekmiş gibi sunmak göstermeye çalışmaktır [18].

Veri kümesindeki veriler ile veri kümesinin bittiği noktadan sonraki 7 aylık veri simüle edilerek edinilen yeni veriler ile bu 7 aylık fiyat optimizasyonu yapıldı.

#### 3.1.Veri Kümesi Simülasyonunda Kullanılan Yöntem

Bu simülasyonu yapmak için autoregressive integrated moving average (ARIMA) modeli kullanıldı. Bu model kullanılmadan önce Seasonal autoregressive integrated moving average with eXogenous factors(SARIMAX), Long short-term memory(LSTM) ve Probabilistic forecasting(Prophet) yöntemleri de denendi ve bu yöntemler denendikten sonra bu modellerin performanslarını ölçmek için Mean absolute error(MAE), Mean squared error(MSE), Root mean square error(RMSE) ve Coefficient of Determination(R2 Score) performans ölçütleri kullanıldı. Bunlardan alınan ölçütlerde hepsinin skorları birbirine çok yakındı bu yüzden hangi modelin seçileceğine karar verilirken farklı kısıtlarda göz önünde bulunduruldu. ARIMA modeli ele alınan herhangi bir seri durağan olsun ya da olmasın mevsimsel unsur içersin ya da içermesin bilgisayar paket programlarıyla bir çözüm olanağı sağlamaktadır [19] ayrıca ARIMA modellemesi analiz aşaması için de zaman serilerindeki trendler, dönemsel desenler ve dalgalanmalar gibi yapıları tanımlama konusunda, zaman serisindeki beklenmeyen veya aykırı değerleri tespit etme konusunda çok yardımcı olmaktadır [20]. Bu sebeplerden dolayı bu çalışma yapılırken ARIMA modeli seçildi. ARIMA modelini kullanırken tarih verisi index olarak alındı.

##### 3.1.1.ARIMA MODELİ

ARIMA modeli temel olarak geçmiş değerler ve hata terimleri arasındaki ilişkiyi temsil eder ve bu ilişkileri kullanarak gelecekteki değerleri tahmin etmeye çalışır [21,15].

ARIMA modelinin çalışma mantığı üç bileşenden oluşur: otoregresif (AR) bileşen, entegrasyon(I) bileşeni ve hareketli ortalama (MA) bileşeni [16].

Otoregresif (AR) bileşen kısaca mevcut değeri geçmiş değerlerle ilişkilendirir denilebilir [16].

Entegrasyon (I) bileşen kısaca zaman serisindeki durağanlığı sağlamak için uygulanan bir fark alma işlemidir [16].

Hareketli ortalama (MA) bileşen kısaca adından anlaşılacağı üzere hareketli ortalama bileşendir, hata terimlerini kullanarak mevcut değeri açıklar [16].

##### 3.1.2.ARIMA Modelini Kullanılışı

ARIMA modeli ile önce toplam satış miktarının tahmini yapılmıştır. Bunu yapmak için sadece 252 id numaralı ürünün “Date” ve “TotalSales” sütunlarını alınarak yeni bir veri kümesi üretilmiş ve bu veri kümesinin index’i “Date” sütunu olarak belirlenmiştir ve veri kümesinin 7 ay sonrasına kadar tahminde bulunulmuştur. Bu üretilen verileri optimizasyon problemi çözülürken kullanılacağı için ve optimizasyon işleminin daha uygun sonuçlar çıkarabilmesi adına gün gün olan veriler aylık

veriye çevrilerek 7 satırlık her ay için olan bütün verilerin ortalamalarını barındıran bir veri kümesi üretilmiştir ve sonuçlar tablo 2’de sunulmuştur.

Tablo 2. Veri kümesinin bittiği tarihten sonraki 7 ay ile ilgili toplam satış verisinin tahmin

Date	TotalSales
2023-03-31	4300.940435
2023-04-30	5278.587595
2023-05-31	6428.445919
2023-06-30	8092.460855
2023-07-31	9290.194664
2023-08-31	10834.074673
2023-09-30	12007.343402

Toplam satış miktarı ile ilgili yapılan işlemlerin aynıları da satış adeti sayısı için yapılmıştır yani veri kümesinden sonraki 7 ayda her ayın bir gününde ortalama kaç adet ürün satılacağını tahmin edilmiştir ve sonuçlar tablo 3’te sunulmuştur.

Tablo 3. Veri kümesinin bittiği tarihten sonraki 7 ay ile ilgili satış adetinin tahmin tablosu

Date	TotalSales
2023-03-31	4.749361
2023-04-30	5.105629
2023-05-31	4.963596
2023-06-30	5.042195
2023-07-31	5.005385
2023-08-31	4.991656
2023-09-30	5.046099

Burada satış adeti sayısının küsuratlı sayılar çıkması ne kadar anlamsız olsa da ortalama alındığı için böyle olması gayet normaldir eğer yuvarlama işlemi veya küsuratı yok etme işlemi gibi bir işlem yapılmış olsaydı veri kaybı olacağından dolayı bu yapılmak istenilmemiştir [22] zaten optimizasyon problemimizde kullanılacak veriler oldukları için bu durum bir sorun teşkil etmemiştir hatta veri kaybı olmadığı için daha optimum bir sonuca yaklaştırmıştır.

ARIMA modeli bu noktada farklı bir amaç için kullanılmıştır ve optimizasyon işlemi için üretilen veri kümesine Enflasyon Oranı adında yeni bir sütun eklenmiştir. Bu sütunun eklenmesinin en büyük sebebi artan enflasyonun bu ürünü ne kadar etkileyeceğini tahmin etmek ve enflasyon sütunu sayesinde optimizasyon problemimizde mevsimselliği kullanabilmektir. Bunu yapabilmek için yine veri kümesinden sadece 252 id numaralı ürünün tarih ve grain\_price yani adet fiyatı sütunları alınarak yeni bir veri kümesi üretildi. Daha sonrasında bu veri kümesi Price\_Change adında enflasyon oranlarının hesaplanıp eklendiği yeni bir sütuna eklendi.

Burada yapılan işlem bir satırdaki adet fiyatının bir önceki satırdaki veya bir önceki gündeki de denilebilir çıkarıldıktan sonra çıkan sonucu yeniden üzerinde çalışılan satırın bir önceki satırına bölmek ve çıkan sonucu yüzdelik bir oran elde edebilmek adına 100 ile çarpmak bu işlemler sonucunda enflasyon yüzdeleri hakkında bilgi veren bir satır elde edilmiştir.

Sonrasında yapılan işlemler yine toplam satış ve adet sayısında yapılan işlemler gibi veri kümesinin bittiği noktadan 7 ay sonrasına kadar olan veri kümesinin tahmin edilmesi ve tahmin işlemi yapıldıktan sonra çıkan sonuçlar da tablo 4’ de sunulmuştur.

Tablo 4. Veri kümesinin bittiği tarihten sonraki 7 ay ile ilgili enflasyon tahmin tablosu

ENFLASYON
16.788506
11.158916
11.158916
11.158916
11.158916
11.158916
11.158916
11.158916

#### 4.OPTİMİZASYON

Optimizasyon, belirli hedeflere ulaşmak için en iyi ve en uygun çözümü bulmayı amaçlayan bir matematiksel disiplindir [22]. Optimizasyon, sınırlı kaynakların verimli kullanılması, maliyetlerin azaltılması, karın maksimize edilmesi (Bu çalışmada çözüm bulunmak istenilen problem) veya performansın en üst düzeye çıkarılması gibi çeşitli problemlerin çözümünde kullanılır [22].

##### 4.1.Optimizasyon Problemlerinde Çoklu Ekstremler Problemi

Optimizasyon problemlerinde çoğu zaman birden fazla maksimum noktası olur fakat bunlardan hepsi en iyi çözümü vermez çünkü bunların içinde yerel maksimum noktaları da vardır [23] örnek ile anlatılmak istenirse bir dağa tırmanırken etrafında bulunan küçük tepelere de tırmanmış oluruz fakat tırmandığımız tepeler dağın zirvesi değildir fakat bazen algoritmalarımız bu noktaları zirve zannedip bir probleme yol açabiliyorlar.

##### 4.2.Problem İçin Uygulanan Optimizasyon Modeli

Lineer programlama (LP) yöntemi kısıtla tanımlanan bir matematiksel modeli çözmek için kullanılan bir yöntemdir [22]. LP için modelleme kolaylığı vardır [22]. Kar maksimizasyonu çalışmalarında genelde LP kullanılır [24]. LP aynı zamanda çeşitli kısıtları aynı anda yönetme ve istenilen matematiği mükemmel şekilde gerçekleştirme yeteneklerine sahiptir [22,24]. Bu bilgiler doğrultusunda bu çalışmada LP kullanılmaya karar verilmiştir. Lineer programlama (LP) yaklaşımı, doğrusal yapıdaki kısıtları ihlal etmeden, doğrusal formdaki amaç fonksiyonunu en iyilemeyi sağlayan ve bu en iyileme sonucunda karar değişkenlerinin aldıkları değerleri bulan bir yaklaşımdır. Her lineer programlama modelinin üç temel bileşeni vardır: Karar değişkenleri, amaç fonksiyonu ve kısıtlar. Lineer programlama (LP) problemi çözülürken kolay ve esnek bir yol sağlaması için Pulp kütüphanesi kullanıldı.

Pulp, Python programlama dili için açık kaynaklı bir LP kütüphanesidir [25].

Pulp, LP modelleri oluşturmayı, kısıtları tanımlamayı, değişkenleri belirlemeyi sağlayan bir dizi fonksiyon ve sınıf içerir [25]. Pulp'ın sunduğu ana sınıflardan biri 'LpProblem' sınıfıdır bu sınıf LP problemi için temel bir yapı oluşturur [25].

Burada problemin bir maksimum bulma problemi olduğu belirtildi.

Ardından, 'LpVariable' sınıfını kullanarak değişkenler tanımlandı;

Burada id olarak df\_monthly\_predSalesTest veri kümesinin Date sütununu alındı bu sayede her tarih için ayrı ayrı rahat işlem yapılabilir ve buna ek olarak da maksimize etmek istenilen verinin de adet fiyatı yani grain\_price sütunu olduğunu belirtildi.

Ve yine ardından, '+' operatörünü kullanarak kısıtlar ve objektif fonksiyonu eklenildi.

Daha önceden adet fiyatının maksimize etmek için uğraşıldığı modele belirtilmişti fakat asıl amaç bunu maksimize ederken aslında toplam satış tutarının en yüksek halini bulmaktır bu yüzden amaç fonksiyonu olarak veya literatürdeki diğer ismi ile objektif fonksiyonu olarak adet fiyatı ve adet sayısının çarpımının maksimum olması gerektiği belirtiliyor.

Daha önceden '+' operatörü ile kısıtlarda koyulabilirdi söylenmişti. Bu kısıtların belirlenmesi çok fonksiyonel ve çok değişkenli bir durumdur [11,26]. Bu işlemin neden yapıldığı çok basit şekilde anlatılması istenirse eğer; örneğin evimize elektrik çekmek için bir kablo uzunluğu belirlememiz gerekiyor. Bu ipi biz çok kısa seçersek evimize ulaşmaz ve amacımıza ulaşamamış oluruz fakat kabloyu çok uzun seçersek bu seferde bizim için gereksiz bir maliyet oluşur bunun için en

şu kadar olmalı ve en çok şu kadar olmalı gibi kurallar koymamız gerekir ; örneğin en az trafı ile evimiz arasında ki mesafe ve binamızın yüksekliği kadar uzunlukta bir kablo diyerek bir kısıt koyabiliriz aynı şekilde problemimize bir de en çok ne kadar olmalı açısından da yaklaşmalı ve ikisinin arasında bir alan bırakıp optimum değeri hesaplamalıyız. Burada da problemi çözerken yapılan işlem aynen bu şekilde adet fiyatı veri kümesindeki en küçük değer ve hesaplanmaya çalışılan satırın enflasyon oranı ile çarpımından küçük olmamalı gibi kısıtlar konuldu.

Burada kısıt olarak veri kümesinin en düşük değerinin enflasyona göre oranı hesaplandığında çıkan sonuçtan adet fiyatının daha küçük olmaması gerektiği, veri kümesinin ortalama değeri ile enflasyon oranına göre hesaplama yapıldığında çıkan değerden küçük olması gerektiği, indirimli fiyat hesaplanıp bu indirim yüzdesi ile çarpıldığında çıkan sonuçtan büyük olması gerektiği ve yine indirimli fiyat hesaplanıp indirim yüzdesine bölüldüğünde çıkan sonuçtan küçük olması gerektiği belirtildi.

Bu şekilde model çözülmüş ve içerisinde sonuçları saklamaktadır.

## 5.DEĞERLENDİRME

Bu çalışmada, perakende ürün satan bir firmanın dört yıllık satış verilerinden oluşan veri kümesinin analizleri yapılmış ve gelecek 7 aydaki ürün fiyatlarını optimum kar edilebilecek şekilde dengeleyen bir makine öğrenmesi modeli çalışılmıştır. Veri kümesinin içinde modeli hata yapmaya yönlendirecek veya modelin çalışma prensibine uygun olmayan veriler ayıklanıp düzeltilmiştir. Kullanılabilir veri elde edildikten sonra optimizasyon işleminde kullanılması gereken veriler simülasyon işlemine tabi tutulmuştur. Bu simülasyon işlemi yaparken ARIMA modeli kullanılmıştır ve bu sonuçlar elde edildikten sonra lineer programlama yöntemi kullanılarak optimizasyon işlemi uygulanmış ve veri kümesinin sonlandığı tarihten sonraki 7 ay için optimum fiyat tahmini yapılmıştır. Bu çalışma firmaların piyasada belirli bir süre geçirmiş ürünleri üzerinde fiyat belirleme konusunda tavsiye almak için kullanılabilir.

Modelimize 252 id numaralı ürün ile ilgili verileri verildi ve bu veriler doğrultusunda veri kümesinin sonlandığı tarihten sonraki 7 ay için aylık ortalama optimum fiyatı tahmin etmesi istenildi. Model çözüldükten sonra sonuçların modelin içerisinde saklandığı söylenmişti bunlara erişebilmek için ise değerlerin bir döngü içerisinde ekrana yazdırılması gerekiyor. Aşağıda Tablo5'te sonuçlar sunulmuştur.

Tablo 5. 252 id numaralı ürünün veri setinin bittiği tarihten 7 ay sonrasına kadar aylık adet fiyatının optimizasyon modeli ile optimizasyon yapılmış sonuçları

TARİH	FİYAT
31.03.2023	1068.59
30.04.2023	1219.97
31.05.2023	1443.01
30.06.2023	1428.93
31.07.2023	1572.92
31.08.2023	1839.35
30.09.2023	2016.55

252 id numaralı ürün için sonuçlar 2023 yılının 3. ayı için 1068 TL, 4. ayı için 1219 TL, 5. ayı için 1443 TL, 6. ayı için 1428 TL, 7. ayı için 1572 TL, 8. ayı için 1839 TL, 9. ayı için 2016 TL olarak tespit edildi. Gerçek veri kümesinde de fiyatların en son 1000 lirayı geçtiği görülmüştü ve gerçek veri kümesinde de tahmin edilen gibi son yıllarda çok hızlı bir fiyat artışı vardı. Tahminde bu şekilde yine yüksek bir enflasyon olacağı tahmin edilmiş. Genelde görülen zaten enflasyon yüksek bir ivme aldığında düşse dahi çok hızlı bir düşüş olmuyor bu ivmede. Bu yüzden tahminlerin gerçeği yansıttığı düşünülüyor.

## 6.SONUÇ

Bu çalışma yapılırken veri kümesinin üzerinde çalışılmaya başlanmadan önce hangi sebeplerden dolayı veri analizi ve veri düzenlemesi yapılır öğrenildi. Optimizasyonun literatürdeki genel kullanımları araştırılıp kavrandı ve Python yazılım dilinde doğrusal optimizasyonun uygulanışı öğrenildi. Optimizasyon uygulaması yapılırken gelecek ile ilgili işlem yapabilmek için zaman serisi modelleri kullanımı ve hangi amaca hizmet ettikleri öğrenildi. Bu makalede veri analizi ve düzenlemesinin yapılma amacı ve yapılırken izlenmesi gereken yol detaylı bir şekilde anlatıldı. Zaman serisi analizleri özelinde mevsimsellik analizi, hareketli ortalama analizi ve enflasyon analizi grafikler ile detaylı bir şekilde anlatıldı. Zaman serisi analizi anlatıldıktan sonra veri kümesi için simülasyon işlemi uygulayabilmek adına içerisinde zaman serisi analizi matematiğini bulunduran ARIMA modeli anlatıldı. Optimizasyon işleminin bizim çalışmamız özelinde gerçekleştirimi detaylı bir şekilde olmak üzere



optimizasyon işlemi genel olarak anlatıldı. Bu makalede bahsedilen çalışmanın amacı geçmiş verilerden yola çıkarak bir firmanın ürünlerini satması gereken fiyatın optimum noktasını tespit etmek ve gelecekte olması gereken optimum fiyatlarla ilgili tavsiye vermektir. Bu çalışmanın fayda sağlayacağı kişi veya kuruluşlardan bahsedilirse müşterisi olan her kuruluş denilebilir çünkü sadece somut bir ürün için fiyat optimizasyonu değil geçmiş verileri olan her alanda kullanılabilir bir çalışmadır. Halihazırda bu yapılan çalışmanın çok daha gelişmiş hali olan ürünlerde firmaların kullanımına hazır bir şekilde yazılım sektöründe bulunmaktadır örnek vermek gerekirse optiwisdom şirketine ait olan kolay.ai ürünü fiyat belirleme konusu ile ilgili yardımcı olmasının yanında müşteri analitiği ile ilgili, gelir ve gider analitiği ile ilgili detaylı analiz ve müşteri segmentasyonu gibi rekabetçi piyasada yardımcı olacak birçok yeteneği içerisinde barındırmaktadır [27]. Bu yapılan çalışmaya devam edildiği takdirde fiyat optimizasyonunun biraz daha geliştirilmesi ve fiyat optimizasyonunun yanına müşteri segmentasyonu özelliği de eklenerek müşteri bazında fiyat optimizasyonu yapılması hedeflenmektedir.

## 7.KAYNAKÇA

- [1] Anitha P, Malini M. Patil, "A Review on Data Analytics for Supply Chain Management: A Case study", 24 Mayıs 2018 10.5815/ijieeb.2018.05.05
- [2] Foster Provost and Tom Fawcett, BIG DATA, "Data Science and its Relationship to Big Data and Data-Driven Decision Making", 13 Feb 2013, /10.1089/big.2013.1508
- [3] Doç.Dr. Metin TÜRKAY, Koç Üniversitesi, Endüstri Mühendisliği Bölümü, "OPTİMİZASYON MODELLERİ VE ÇÖZÜM METODLARI", 2011
- [4] CNN Money (2001) "Webvan shuts down.", 9 July 2001, Accessed 6 June 2013
- [5] CNBC NEWS, "Netflix didn't kill Blockbuster — how Netflix almost lost the movie rental wars", Sep 22.2020
- [6] Philip Kotler ve Kevin Lane Keller "Marketing Management", 2011, 8-9 syf.
- [7] Asnat Greenstein-Messica, Lior Rokach, "Machine learning and operation research based method for promotion optimization of products with no price elasticity history", Mart-Nisan 2020, /10.1016/j.elerap.2019.100914
- [8] John H. Rios, c, Jorge R. Vera, "Dynamic pricing and inventory control for multiple products in a retail chain", Mart 2023, /10.1016/j.cie.2023.109065
- [9] Ersin Namlı, Ramazan Ünlü, Ecem Gül, "FİYAT TAHMİNLEMESİNDE MAKİNE ÖĞRENMESİ TEKNİKLERİ VE DOĞRUSAL REGRESYON YÖNTEMLERİNİN KIYASLANMASI; TÜRKİYE'DE SATILAN İKİNCİ EL ARAÇ FİYATLARININ TAHMİNLENMESİNE YÖNELİK BİR VAKA ÇALIŞMASI", 23 Mayıs 2019, /10.36306/konjes.654952
- [10] Bazarara , M.S.,Jarvis, J.J. ve Sherali , H.D., "Linear Proramming and Network Flows.", 2011, 1-32 syf.
- [11] İnternet: <https://marketbusinessnews.com/financial-glossary/optimal-price-definition-meaning/>, 2023
- [12] Ali KOYUNCUGİL, Nermin ÖZGÜLBAŞ, Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları, 22 Temmuz 2010
- [13] Şadi Evren ŞEKER, "Retail Data Set", Mart 2023, 10.34740/kaggle/ds/3067824
- [14] Adem Tekerek, Akademik Bilişim Konferansı Bildirileri, "Veri Madenciliği Süreçleri ve Açık Kaynak Kodlu Veri Madenciliği Araçları", 4 Şubat 2011
- [15] Abdur Rahman, Md Mahmudul Hasan "Modeling and Forecasting of Carbon Dioxide Emissions in Bangladesh Using Autoregressive Integrated Moving Average (ARIMA) Models.", Ağustos 2017, /10.4236/ojs.2017.74038
- [16] Jose Marcial Portilla , Using Python and Auto ARIMA to Forecast Seasonal Time Series, 26 Mart,2018,
- [17] Fisher, I., "The Theory of Interest: As Determined by Impatience to Spend Income and Opportunity to Invest It",2014, 3-15 syf.
- [18] Jean Baudrillard, "Simülakrlar ve Simülasyon" 1981, syf .25-55
- [19] Ahmet TORTUM, Oğuzhan GÖZCÜ, Muhammed Yasin ÇODUR, İğdir Üniversitesi Fen Bilimleri Enstitüsü Dergisi "Türkiye'de Hava Ulaşım Talebinin Arıma Modelleri ile Tahmin Edilmesi", 2014
- [20] Jeffrey E Jarrett, Ph.D., and Eric Kyper, Ph.D., "ARIMA Modeling With Intervention to Forecast and Analyze Chinese Stock Prices", 30 Nisan 2011, <https://doi.org/10.5772/50938>
- [21] İbrahim Delibaşoğlu blog.Python – Zaman Serisi İnceleme, Temmuz 2017
- [22] Mitchell, S.,Stump and E.(2019)."Pulp: A Python Library for Linear Optimization." Journal of Open Source Software
- [23] Cansın KAYA, MİMAR SİNAN GÜZEL SANATLAR ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ, Yüksek Lisans Tezi, "DOĞRUSAL OLMAYAN PROGRAMLAMA ile PORTFÖY ANALİZİ", Ocak 2012
- [24] Hakkı KUL, Yüksek Lisans Tezi, "İNŞAAT PROJELERİNDE DOĞRUSAL PROGRAMLAMA YÖNTEMİYLE SÜRE-KALİTE-MALİYET OPTİMİZASYONU, Mayıs 2010
- [25] İnternet :<https://pypi.org/project/PuLP/>, 2 Haziran 2023
- [26] İdil İsmigüzel , How to Develop Optimization Models in Python, 27 Haziran 2020
- [27] Şadi Evren ŞEKER, KOBİ'ler için Basit Yapay Zeka Çözümü : Kolay.AI, YBS Ansiklopedi, v.11, is 1, pp. 24-46, 2023