



# Bir Savunma Sanayi Firmasında Depo Bölgesi Atama Sisteminin Veri Madenciliği ve Makine Öğrenme Yaklaşımlarıyla İyileştirilmesi

Mehmet Ertem<sup>1\*</sup>

<sup>1\*</sup> Eskişehir Osmangazi Üniversitesi, Mühendislik-Mimarlık Fakültesi, Endüstri Mühendisliği Bölümü, Eskişehir, Türkiye, (ORCID: 0000-0001-5363-3619), [mertem@ogu.edu.tr](mailto:mertem@ogu.edu.tr)

(2nd International Conference on Applied Engineering and Natural Sciences ICAENS 2022, March 10-13, 2022)

(DOI: 10.31590/ejosat.1083226)

**ATIF/REFERENCE:** Ertem, M. (2022). Bir Savunma Sanayi Firmasında Depo Bölgesi Atama Sisteminin Veri Madenciliği ve Makine Öğrenme Yaklaşımlarıyla İyileştirilmesi, (34), 501-506.

## Öz

Günümüzde envanter kontrolü ve depolama operasyonları firmaların masraflarının oldukça önemli bir bölümünü oluşturmaktadır. Bu nedenle pek çok firma envanter kontrolü ve depolama operasyonları üzerindeki etkinliklerini arttırmayı amaçlamaktadır. Firmalar bu doğrultuda çözüm üretebilmek amacıyla, bazen depolama alanlarının sıfırdan tasarımını tartışırken bazen de maliyetler göz önüne alınarak var olan çalışma alanlarını optimize etmeyi amaçlamaktadır. Depo sistemlerinin etkinliğinin artırılması amacıyla kullanılan pek çok alternatif yöntem vardır. Ancak depoyu bir veri ambarı, içindeki hammadde ve parçaları da veri olarak düşünecek olursak, veri madenciliğinin kullanımı kaçınılmaz olacaktır. Bu çalışmada, bir depolama yöntemi olan depo bölgesi atama işleminin iyileştirilmesi adına veri madenciliği teknikleri kullanılmıştır. Modelleme adımında ise bir makine öğrenme yaklaşımı olan çok sınıflı sınıflandırma algoritmalarından yararlanılmıştır. Bu bağlamda, literatürde yaygın olarak kullanılan çok sınıflı sınıflandırma algoritmaları karşılaştırılarak performansı en yüksek 8 tanesi seçilmiştir. Seçilen algoritmalar ile çalışılan probleme uygun bir model Python kodlama dili kullanılarak geliştirilmiş ve depolanan yedek parçaların özellik verileri kullanılarak hangi ambar bölümüne ait olduğu tahmin edilmiştir. Yapılan veri madenciliği ve makine öğrenmesi yöntemleri sonucunda gerçek ambar bölümünü en doğru tahmin eden algoritmanın, %72'lik bir başarı oranıyla Rastgele Orman algoritması olduğu görülmüştür. Genel anlamda bu çalışmada kullanılan bazı verilerin eksik veya yanlış kaydedildiği düşünüldüğünde daha fazla ve tutarlı veri kullanılarak bu başarı oranının çok daha yüksek oranlara çıkabileceği öngörülmektedir.

**Anahtar Kelimeler:** Makine öğrenme, Veri madenciliği, Yapay zeka, Depo bölgesi atama, Depo yönetimi.

## Improving the Warehouse Zoning System of a Defense Industry Company Using Data Mining and Machine Learning Approaches

### Abstract

Today, inventory control and storage operations constitute a very important part of the costs of companies. For this reason, many companies aim to increase their effectiveness in inventory control and storage operations. In order to produce solutions in this direction, companies sometimes discuss the design of storage areas from scratch, and sometimes they aim to optimize existing workspaces by considering costs. There are many alternative methods used to increase the efficiency of warehouse systems. However, if we consider the warehouse as a data warehouse and the raw materials and parts in it as data, the use of data mining will be inevitable. In this study, data mining techniques are used to improve the storage zone assignment process, which is a storage method. In the modeling step, multi-class classification algorithms, which is a machine learning approach, were used. In this context, by comparing the multi-class classification algorithms commonly used in the literature, the 8 with the highest performance were selected. A model suitable for the problem studied with the selected algorithms was developed using the Python coding language and it was estimated which warehouse section belonged to the stored spare parts using the property data. As a result of the data mining and machine learning methods, it has been seen that the algorithm that predicts the real warehouse section most accurately is the Random Forest algorithm with a success rate of 72%. In general, considering that some of the data used in this study were incomplete or incorrectly recorded, it is predicted that this success rate could reach much higher rates by using more consistent data.

**Keywords:** Machine learning, Data mining, Artificial Intelligence, Warehouse zoning, Warehouse management.

\* Sorumlu Yazar: [mertem@ogu.edu.tr](mailto:mertem@ogu.edu.tr)

## 1. Giriş

Günümüzde pek çok saygın firma, giderek önemi artan depolama operasyonlarına daha fazla önem vermeye başlamıştır. Depolama operasyonları depolarda maliyet, etkinlik ve sürdürülebilirlik açısından kritik bir rol oynamaktadır. Operasyon maliyetlerin en küçülenmesi, etkinlik ve sürdürülebilirliğin de optimum düzeylerde olması beklenmektedir. Envanter kontrol ve stok taşımada etkinliği arttırabilmek amacıyla kullanılabilinecek yöntemlerden biri de literatürde “Warehouse Zoning / Put-away Zoning” yani depo bölgesi atama olarak geçen uygulamadır. Bu yöneme göre depolar, firmaların ihtiyaçlarına göre çeşitli bölümlere ayrılmaktadır. Bu sayede parçalar farklı özelliklerine göre bu bölümlerden birine konumlandırılabilir.

Bu çalışmada bir savunma sanayi firması, envanter kontrolündeki etkinliğini arttırmak ve stok birimlerini uygun bir şekilde depolayabilmek için, literatürde ‘depo bölgesi atama’ olarak geçen yöntemi kullanmaktadır. Parçaların, farklı depo bölgelerinden birine atanması gerektiğini savunan bu yöntem, firma çalışanlarının tecrübelerine ve o anki koşullara dayanmaktadır. Normalde bir parça veya hammaddenin depolanması için lokasyon ataması yapılmalıdır. Yani tam olarak hangi depo bölgesinin hangi rafına konacağı belirlenmelidir. Ancak bu her zaman mümkün olmayabilir. Bazı firmalarda lokasyon atamasından önce bölge ataması yapılır. Yani parça ambara girdiğinde zaten hangi depo bölgesine gideceği uzman görüşüne göre veya kullanılan başka sistemlere göre bellidir. Ancak o anda tam konumu bilinmemekle beraber bölgede boş olan yere ataması daha sonradan yapılır. Buna da depo bölgesi ataması denilmektedir.

Yeni bir projeye başlanıldığında daha önce hiç depolanmamış parça çeşitleriyle karşılaşmaktadır. Bu durum, ilgili firma çalışanlarını her bir parçanın nereye atanması gerektiğini belirleme zorluğuyla karşı karşıya bırakmaktadır. İlgili probleme daha iyi performanslı bir çözüm arayan firma, akıllı bir sistem tasarımına gereksinim duymuştur. Bu çalışma kapsamında böyle bir sistemin geliştirilmesi için veri madenciliği ve yapay zeka tekniklerinden biri olan makine öğrenmesi yöntemiyle çözüm sunulmuştur.

## 2. Materyal ve Metot

### 2.1. Literatür Taraması

Depolama sistemleri literatürde detaylı olarak çalışılmıştır. Depolama sistemleri ve yöntemleriyle ilgili yapılan çalışmalardan bazıları için (ör. [1], [2], [3], [4]) numaralı referanslara bakılabilir. Depolama sistemlerinin önemli bir parçası olan depo bölgesi atamasının yapılabilmesi için ilk etapta depolanacak ürünlerin sınıflandırılması gerekmektedir. Praveen vd. [5] literatürdeki tüm stok sınıflandırma yöntem ve yaklaşımlarını ele almıştır. Çalışmada yöntemler şu şekilde kategorize edilmiştir:

1. Tek Kriterli Sınıflandırma
2. Çok Kriterli Sınıflandırma
  - a) Model / İstatistik Temelli Yaklaşımlar
  - b) Yapay Zeka Temelli Yaklaşımlar

Çok kriterli sınıflandırma yöntemlerinden biri olan model/istatistik temelli yaklaşımlar, çoğu uygulamada tutarsız ve zaman alan yaklaşımlar olarak değerlendirilmiştir ([5]). Yapay zeka temelli yöntemler ise model/istatistik temelli yaklaşımlara göre daha tutarlı sonuçlar vermektedir. Ancak yapay zeka

tekniklerinin karmaşık olmasından dolayı envanter yöneticileri, uygulaması daha kolay olan geleneksel yöntemleri kullanmayı tercih etmektedirler. Yapay zeka temelli yaklaşımı kullanan bir depolama sistemi tasarımında bulanık mantıktan yararlanılabileceği savunulmuştur ([6]). Lam vd. [6] çalışmalarında OLAP tabanlı bir Bulanık Depolama Atama Sisteminin (Fuzzy Storage Assignment System-FSAS), depo yeri atama problemini çözmek için karar destek verilerinin kullanılabilirliğini arttırması önerilmiştir. FSAS; Veri Yakalama ve Analiz Modülü (DCAM) ve Depolama Yeri Atama Modülü (SLAM) olmak üzere iki modülden oluşturulmuştur. Bu iki ana modül veri kullanılabilirliğinin artırılması, mevcut verilerin tam olarak analiz edilmesi ve uzmanların bilgilerinin sisteme dönüştürülmesi amacıyla, araştırma hedeflerine ulaşmak için kullanılmıştır. DCAM; malın faaliyet bilgileri, talep korelasyonu ve depo bilgileri de dahil olmak üzere, depo mühendisi veya müdürü ile ilgili spesifik bilgileri sunan karar destek sistemi olarak çalıştırılmış ve bir sonraki modül olan SLAM için gereken parametreleri oluşturmuştur. Depo yöneticisi tarafından belirtilen kurallar ve parametreler yardımıyla ‘If-Then’ yapısını içeren modeller kurulmuştur. Tüm bunların uygulanması sonucunda oluşturulan bulanık sistemi çalıştırmak için, MATLAB-Bulanık Mantık Araç Kutusu kullanılmıştır. Sonuçta ürünler, en erişilebilir olan altın bölge (golden zone), orta erişilebilir olan gümüş bölge (silver zone) ve en az erişilebilir olan bronz bölgeye (bronze zone) yerleştirilmiştir.

Bu çalışmada amaç, hali hazırda belli olan toplama bölgelerine stok birimi ataması yapabilecek bir model ortaya koymaktır. Literatürdeki stok sınıflandırma yaklaşımları genellikle stokları belirli özelliklerine göre kümelemeyi amaçlamaktadır. Dolayısıyla bu yaklaşımlar ya depo tasarlarken ya da tasarlanmış farklı depo bölgelerini kendi içinde düzenlerken kullanılmaktadır. Bahsedilen yöntem, depo bölgesi ataması gerçekleştirildikten sonra bölge içi atamaya katkı sağlamaktadır. Dolayısıyla bu çalışma için uygun bir yöntem değildir. Bu bilgiler doğrultusunda çalışmadaki amacımıza yönelik çözüm üretebilecek farklı yaklaşımlar ele alınmıştır. Elde bulunan veriler analiz edildiğinde ve değerlendirildiğinde veri madenciliği ile bir sınıflandırma yaklaşımının problemimize çözüm getireceği görülmüştür.

### 2.2. Problem Tanımı

Çalışma kapsamında uygulama yapılan firma, ihtiyaca özel, etkin, güvenilir ve yenilikçi paletli ve tekerlekli zırhlı araçlar tasarlayan, üreten ve alanında dünya liderleri arasında kabul edilen bir kara savunma sistemleri firmasıdır. Firmanın ana deposu tamamen kendisi tarafından işletilmektedir ve depoda hammadde, ara mamul ile bazı son mamuller stoklanmaktadır. Bu depo kendi başına bir bütün halinde düşünülebileceği gibi içindeki farklı özelliklere sahip bölgeler dikkate alındığında birden fazla bölüm halinde de ele alınabilir. Depo içerisinde farklı özelliklere sahip 7 stoklama bölgesi bulunmaktadır. Bu bölgelerin yapıları ve genel durumları Tablo 1’de verilmiştir.

Depodaki hareketler, depo çalışanları ile planlama, envanter kontrol bölümü tarafından kontrol edilmektedir. Literatürde depo bölgesi atama olarak adlandırılan fonksiyon, stok birimlerinin kabul edilmesi (receiving) ve stoklanması (stocking) operasyonlarından oluşmaktadır. Firma, envanter kontrolünde etkinliğini arttırmak amacıyla, depolarda bir bölümlendirme uygulaması olan depo bölgesi atama sisteminden yararlanmaktadır. Bir parça deponun kabul bölümüne geldiğinde parçanın hangi bölgeye gideceğine bu 7 ambar bölgesinden biri seçilerek karar verilmektedir. Bölge seçimi yapıldıktan sonra

operatör, o bölge içerisinde deneyim ve tecrübesine dayanarak stok birimini depolamaktadır.

Tablo 1. Uygulama firmasının depo bölümleri

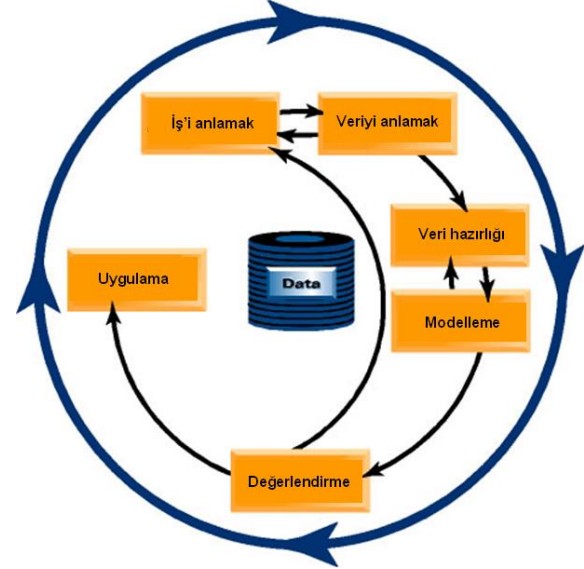
Silo Ambarı	Büyük malzemelerin tutulduğu 17 metre yüksekliğindeki depodur. Oldukça büyük ve görece değerli parçalar burada stoklanmaktadır.
Ana Ambar Ağır Yük Raflar	67 cm 7 kat ve 83 cm 6 kat gibi raf aralıkları bulunan ana raf bölgesidir. Stok biriminin depolanması ve toplanması forklift ve hand truck gibi araçlarla yapılmaktadır.
Dar Koridor	40x60x28 cm ölçülerindeki yaklaşık 12.000 kutudan oluşmaktadır. 3 boyutta hareket edebilen ve manuel olarak çalıştırılan gelişmiş forkliftler kullanılmaktadır. Görece orta büyüklükteki stok birimlerinin depolandığı bölgedir
Ek-1 Ambar Rafları	Çalışma sistemi olarak ana ambar ile aynıdır fakat rafların travers aralıkları farklıdır
Ek-2 Ambar Rafları	Ana raf bölgesine ve Ek-1 raflarına göre daha uzun travers aralıklarının kullanıldığı raf bölgesidir. Boyut ölçüleri bozuk stok birimlerinin depolandığı bölgedir. Ayrıca Ek-2 ambarın içerisinde 3 ayaklı raflarda bulunmaktadır. Bu raflar daha fazla ağırlık taşıyabilmektedir.
Kardex	Dikey doğrultuda çalışan atlıkarınca depolama sistemidir. Vida gibi küçük stok birimlerinin depolanmasında kullanılır
Mezanin	Kaynak hattına giden küçük parçaların depolandığı lokasyondur. Katlı bir sistemdir. Herhangi bir ekipman kullanımı gerektirmez.

Bölge seçimini firmanın planlama ve envanter kontrol bölümünde çalışan endüstri mühendisleri yapmaktadır. Endüstri mühendisi bölge seçimi yaparken farklı değişkenleri kullandığı birtakım formüllerden ve tecrübelerinden yararlanmaktadır. Depoya gelen her yeni parça için bu işlemleri yapmak oldukça zahmetli ve zaman alıcı bir iştir. Firma bu sebeplerden ötürü, yeni gelen her parçayı otomatik olarak ilgili bölgeye atayabilecek bir sistem, model ya da algoritma geliştirmek istemektedir. Bu çalışma kapsamında da, parçanın atanması gereken depo bölgesinin bilgisini otomatik olarak verebilecek bir model geliştirilmiştir. Kullanılacak yöntem, veri madenciliği ve makine öğrenmesi yaklaşımı ile, parçanın depo bilgilerini kullanarak hangi bölgeye atanması gerektiğini belirleyen bir model olacaktır. Problemin bazı açılardan genele uymaması nedeniyle literatürde var olan stok sınıflandırma yöntemlerinin uygulanması çözüm getirmeyecektir. Elde bulunan verilerin durumları da göz önüne alındığında en uygun yöntem, veri madenciliği adımlarıyla veriyi anlamak, bilgi çıkartmak ve modele hazır hale getirmek ve makine öğrenmesi ile en uygun bölgeyi bulmak olacaktır.

## 2.3. Veri Madenciliği ve Makine Öğrenme

Veri madenciliği, eldeki verilerden bilgi çıkarımı yapmaya olanak sağlayan, çeşitli sektörlerde başarıyla kullanılan bir veri işleme yaklaşımıdır. Veri madenciliği kısaca; veri toplama, veriyi ön işleme, modelleme ve değerlendirme aşamaları olarak betimlenebilir (bakınız Şekil-1). Veri; veri ambarlarından, veri tabanlarından ve veri barındırabilen her ortamdan çıkartılabilmektedir. Veriyi toplama yöntemi veri madenciliği çalışmalarında son derece önemlidir. Çünkü çıkarılan veri, aranan bilgiyi türetebilecek yapıda olmalıdır. Veri çıkartma işleminin ardından veri ön işleme metotları gelmektedir. Eldeki veri şu anki haliyle kullanıma hazır değildir ve çeşitli işlemlerden geçmelidir.

Verinin doğruluğunun araştırılması, eksikliklerin giderilmesi, boyut indirilmesi, kullanılabilirliğinin değerlendirilmesi, tahmin edilmesi, veriyi olumsuz etkileyen uç noktaların silinmesi, sınıflandırılması, kümelenmesi gibi işlemler verinin ön işleme aşamasında gerçekleştirilmektedir. Daha sonra ise, işlenmiş verinin hangi metotlarla (modellerle) bilgiye dönüştürülebileceğine karar verilmesi gerekmektedir. İstenen sonuç ve eldeki veri bu metotların seçimini ciddi şekilde etkilemektedir. Elde edilen bilgi, probleme doğrudan çözüm olabileceği gibi, bilgiden yola çıkarak probleme çözüm geliştirilmesi de sağlanabilir.



Şekil 1. Veri Madenciliği Süreci

Freitas vd. [7], veri madenciliğinin hangi görevleri yerine getirmek üzere kullanılabileceğini yaptığı çalışmada belirtmiştir. Bu görevler izlenen şekildedir:

**Sınıflandırma (Classification)**, üzerinde en çok çalışılan veri madenciliği görevidir. Sınıflandırmada amaç, verilerin nitel etiketlerinin bulunduğu durumlarda, faktörlerin bu etiketleri nasıl etkilediğinin araştırılması ve etiketsiz verilerin isabetli olarak etiketlenmesidir.

**Tahmin (Prediction)**, sayısal verilerin tahmin edilmesi görevidir. X bağımsız faktörleriyle Y bağımlı (hedef) faktörünün tahmin edilmesi amaçlanmaktadır. Çözüme, regresyon yöntemleriyle ulaşılabilmektedir.

**Clustering (Kümeleme)**, sınıflandırma görevleriyle sıkça karıştırılmaktadır. Sınıflandırmada etiket (sınıf) değerleri belirli

iken, kümelemede veriden hareketle etiketler bulunur. Sınıflandırma görevleriyle beraber kullanımı yaygındır.

*İlişki kuralı çıkarma (Discovery of Association Rules)*, veri madenciliğinde en çok kullanılan görevlerin arasında yer almaktadır. Amaç, birbirinden farklı değişkenlerin, diğer değişkenlerle arasında nasıl bir ilişki olduğunu ortaya çıkarmaktır. Örneğin, bir süpermarkette x ve y gıdalarını alan bir kişinin z gıdasını da alıyor olması, x,y ve z gıdaları arasında bir ilişki olduğunu gösterebilir.

Makine öğrenimi (Machine Learning), bilgisayar sistemlerinden algılanan ya da veri tabanlarında tutulan verilerin üzerinde bir öğrenimin gerçekleşmesi ve bu öğrenmeye dayalı olarak hedef verinin türetilmesi/tahmin edilmesi kavramlarını araştıran bir bilim dalıdır. İnsan davranışlarını taklit etmeye çalışabileceği gibi insanın kompüsyonel olarak erişemeyeceği seviyedeki işlevleri de yerine getirebilecek şekilde kullanılmaktadır. Günümüzde ise yapay zeka yaklaşımlarının büyük bir kısmını makine öğrenmesi teknikleri oluşturmaktadır. Makine öğrenme teknikleri hedeflenen sonuçlara göre bazı sınıflara ayrılabilir:

- Gözetimli Öğrenme (Supervised Learning)
- Gözetimsiz Öğrenme (Unsupervised Learning)
- Pekiştirmeli Öğrenme (Reinforced Learning)

Gözetimli Öğrenme, verilen girdi kümesinden, istenen çıktı kümesinin elde edilebilmesi için, bir model eğiten algoritmaların genel ismidir. İstenen hedef çıktılar başlangıçta elde bulunmaktadır. Girdileri ve bu girdilerin sebep olduğu çıktılar bir öğrenci (training) modülüne sunularak, kullanılan ilgili algoritma ile öğrenci modülü eğitilir ve bir model geliştirilir. Oluşturulan model, hedef çıktıları bulunmayan girdiler verilerek test edilir. Hedef çıktıları doğru tahmin etme kapasitesine göre de modelin doğruluğu (accuracy) değerlendirilir [8]. Gözetimli öğrenmede temel olarak kabul edilen görevler ve görevlerin çözümünde kullanılan bazı temel algoritmalar aşağıdaki gibidir:

- Tahmin
  - Basit Doğrusal Regresyon
  - Çoklu Doğrusal Regresyon
  - Polinom Regresyon
  - Lojistik Regresyon
  - Destek Vektör Makineleri
  - Karar Ağacı
  - Rassal Ağaçlar
- Sınıflandırma
  - Lojistik Regresyon
  - K En Yakın Komşuluk
  - Destek Vektör Makineleri (SVM)
  - SVM ve Kernel Hilesi
  - Naive Bayes • Karar Ağacı
  - Rassal Orman

Gözetimsiz Öğrenme (Unsupervised Learning), gözetimli öğrenmeden farklı olarak, verileri sebepsonuç ya da giriş-çıkış şeklinde etiketlemeden, veri içerisinde var olan ilişkilerin ve yapıların öğrenilmesidir. Bu nedenle bu tekniklerin temel amacı, verinin altında yatan yapıyı ve dağılımı kestirerek bilgi üretmektir. Gözetimsiz öğrenmede kümeleme (clustering) ve ilişkilendirme (association) temel görevlerdir.

*Pekiştirmeli Öğrenme (Reinforced Learning)*, amaca yönelik ne yapılması gerektiğini öğrenen bir makine öğrenmesi yaklaşımıdır. Pekiştirmeli öğrenmede ajan (agent) adı verilen

öğrenen makine, karşılaştığı durumlara bir tepki verir ve bunun karşılığında da sayısal bir ödül sinyali alır. Ajan/öğrenen makine aldığı bu ödül puanını maksimuma çıkartmak için çalışır. Bu şekilde çalışan deneme yanılma yöntemi, pekiştirmeli öğrenmenin en ayırt edici özelliğidir. Pekiştirmeli öğrenmede yapay sinir ağları, destek vektör makineleri, bayes ağları ve genetik algoritmalar yaygın olarak kullanılmaktadır.

## 2.4. Çok Sınıflı Sınıflandırma

Bu çalışmada kullanılacak verilerin yapısı ele alındığında bir sınıflandırma problemi söz konusudur. İlgili parçanın ilgili ambar bölümüne yerleştirilmesi için, parçanın ait olduğu bölüm tahmin edilmelidir ve bu da makine öğrenmesinde bir sınıflandırma görevidir. Hedef çıktı, yani ambar bölümleri bu modelde sınıfları oluşturmaktadır. İki den fazla ambar bölümü olduğu için ikiden fazla sınıf bulunmaktadır. Bu nedenle geleneksel sınıflandırma algoritmaları, parametre değişimi yapılmadan ya da ek kıstaslar eklenmeden kullanılamamaktadır. Literatürde “Çok Sınıflı Sınıflandırma (Multiclass Classification)” olarak geçen problemimiz için çeşitli yaklaşımlar bulunmaktadır. Bu yaklaşımları genel olarak 3 kategoride toplamak mümkündür. Bunlar:

- İkili Sisteme Çevirme Yaklaşımı
  - One vs rest
  - One vs one
- Hiyerarşik Sınıflandırma Yaklaşımı
- İkili Sistemin Uzantısı Olarak Belirtme Yaklaşımı
  - Yapay Sinir Ağları
  - Aşırı Öğrenme Makineleri
  - K En Yakın Komşuluk
  - Naive Bayes
  - Karar Ağaçları
  - Destek Vektör Makineleri

Bahsedilen yaklaşımlar kullanılarak problemi en iyi temsil edecek olan yaklaşım performans karşılaştırması yapılarak Bulgular kısmında tartışılmıştır.

## 3. Araştırma Sonuçları ve Tartışma

‘Python’ makine öğrenmesi çalışmalarında kullanılan en popüler yazılım dilidir. Bu doğrultuda Python’da rahat çalışma imkanı sağlayan Spyder IDE’si kullanılarak modellemeler geliştirilmiştir.

### 3.1. Verilerin Modellemeye Uygun Hale Getirilmesi

Modelleme çalışmaları Python programlama dili ile gerçekleştirilecektir. Dolayısıyla verilerin temel ön işlemleri yapıldıktan sonra Python’da çalışabilir hale getirilmeleri gerekmektedir. Python çalışma yapısı gereği virgülle ayrılmış formattaki verileri kabul etmektedir. Excel’de son tablo olarak düzenlenen veriler bu formata uygun şekilde ‘.csv’ dosyalarına çıkarılmış, ardından veriler Python çalışma ortamına yüklenmiştir.

Modelleme yapılırken kullanılacak verilerin numerik verilerden oluşması gerekmektedir. Kategorik olan veriler bu formata uymamaktadır ve dolayısıyla sistematik bir şekilde veri yapısını bozmadan sayısallaştırılması gerekmektedir. Kategorik verilerin numerik hale getirilmesi için 0- 1 tamsayı kodlama yaklaşımı kullanılmıştır. Normalde bir sütundan oluşan kategorik veri, 0-1 kodlama yaklaşımı ile kategorik faktörün içerdiği farklı kategori sayısı kadar yeni sütuna aktarılmaktadır. Bu doğrultuda



veriler Python üzerinde ‘OneHotEncoding’ modülü kullanılarak 0- 1 tamsayı hale dönüştürülmüştür.

### 3.2. Verilerin Eğitim ve Test Kümesine Ayrılması ve Normalizasyonu

Gözetimli makine öğrenmesi, ‘X’ girdi değişkenleri ve ‘Y’ hedef değişkenini kullanarak bir model geliştirme temelinden oluşmaktadır. Model öğrenme işlemini yaparken verilen bütün veriyi kullanarak bir algoritma belirlemektedir. Dolayısıyla öğrenmek için kullandığı verileri çok yüksek oranda tahmin edebilecektir. Bu durum modelin başarısını ölçerken zorluklar ortaya çıkarmaktadır. Verilerden öğrenilen modellerin başarılarının daha iyi ölçülebilmesi ve test edilebilmesi için veriler 2 parçaya bölünmektedir. Bölünen verilerin ilk kısmı modeli öğrenmek için kullanılırken kalan kısmı ise modelin doğruluğunu test etmek için kullanılmaktadır. Bölünme eşik değeri, öğrenme verisi için 0,25-0,4 arasında kabul edilmektedir. Bu çalışma kapsamında bölünme eşik değeri 0,33 olarak kabul edilmiştir ve verilerin rastgele 2/3’ü öğrenme modülü için ve kalan 1/3’ü oluşturulan modeli test etmek için ayrılmıştır.

Veriler kullanılacak algoritmalarca aynı ağırlıklarda anlaşılmalıdır. Verileri normalize etmek, verileri sayısal değerlerinin verdikleri ağırlıklardan kurtararak aynı uzay düzleminde temsil edebilmek amacını taşımaktadır. Modelde kullanılacak algoritmalar çoğunlukla verilerin n boyutlu uzaydaki konumlarına bakacaktır. Dolayısıyla sayıca büyük veya küçük olan veriler modelin doğruluğunu etkileyecektir. Bu durumdan kurtulmak amacıyla veriler normalizasyon tekniklerinin uygulanması gerekmektedir. Veriler Python’da “Scikit.learn” kütüphanesinin ‘StandartScaler’ modülü kullanılarak bu verilere göre normalleştirilmiştir. Bu yöntemin seçilmesinin sebebi sıradışı verilere karşı daha sağlam bir veri yapısı oluşturabilmektir.

### 3.3. Bulgular

Python’da, daha önceki bölümlerde verilen veri hazırlık çalışmalarından sonra modeller uygulanmaya başlanmıştır. Her bir farklı model algoritması için verilerin bir öğrenci modülü üzerinden eğitilmesi söz konusudur. Her algoritma kendi çalışma prensibine göre öğrenci modülünden desenler öğrenmiş ve sonuçları test modüllerinde başarı ölçümüne tabi tutulmuştur. Tablo-2’de kullanılan makine öğrenme algoritmalarının performans sonuçları özetlenmiştir. Tablo-2’nin ilk sütununda verilen kriterler makine öğrenme algoritmalarının performanslarını değişik açılardan ölçen ve yaygın kullanılan performans kriterleridir. Ayrıca Tablo-2’nin daha rahat okunabilmesi için renk matrisi yaklaşımı kullanılmıştır. Burada kırmızıdan yeşile doğru renk sırasıyla en düşük performanstan en yüksek performansa doğru değişmektedir.

### 3.4. Tartışma

Tablo 2’de özet şekilde verilen bulgulara göre makine öğrenme algoritmalarının depo bölgesi atama problemi için performansları karşılaştırılmıştır. Burada Lojistik regresyonla sınıflandırma (LOGR) modeli, verilerin %66,3 lük bir kısmını doğru tahmin etmektedir. MMC değeri de bu tahminin rassal olmadığını gerçekten veriler içinde desenlerin yakalandığını temsil etmektedir. KNN Modeli, verilerin %65,5’lik bir kısmını doğru tahmin etmektedir. MMC değeri de bu tahminin lojistik regresyon sınıflandırıcısına göre daha başarılı bir şekilde ambarları tahmin ettiğini göstermektedir. SVM modeli, ambar bölümlerini %66,3 oranında doğru tahmin etmektedir. Naive

Bayes sınıflandırma algoritması ciddi anlamda düşük skorlar elde etmiştir. %11,5’lik bir başarı ile değerlendirilmesi mantıklı bulunmamıştır. Bu başarısızlığın sebebi algoritmanın faktör sayısı arttıkça zayıflamasından kaynaklanmaktadır. Kullanılan DTC Algoritması, %59,6’lık bir doğruluk oranı vermiştir. DTC algoritması şu ana kadar değerlendirmeye alınan en zayıf algoritma olarak belirlenmiştir. RFC algoritması %72,5’lik bir başarı ile ambar bölümlerini tahmin etmiştir. MMC değerinin %64,4 olması da algoritmanın sınıflandırmada diğerlerine göre çok daha başarılı olduğunu göstermektedir. RFC algoritması standart algoritmalarla karşılaştırıldığında çok daha güçlü bir algoritmadır. Gücünü karar ağacı temelinden alım ensemble algoritmalarla harmanlayarak elde etmektedir. Bu durum ensemble metodlarının daha başarılı sonuçlar getirebileceği fikrini doğurmuştur. Renk matrisinde de (Tablo-2) görüldüğü gibi (koyu yeşiller en iyi, koyu kırmızılar en kötü) en başarılı algoritma rastgele orman olarak bulunmuştur. Her değerlendirme ölçütüne gerek en büyük başarıyı sergileyen algoritmadır. Dolayısıyla ambar bölümlerinin tahmin edilmesinde kullanılmak üzere seçilmiştir.

Tablo 2. Makine öğrenme algoritmaları performans değerleri (LOGR: Lojistik regresyon sınıflandırma algoritması, KNN: K-En Yakın Komşuluk Algoritması, SVM: Destek Vektör Makineleri Algoritması, NB: Naive Bayes sınıflandırma algoritması, DTC: Karar Ağacı Algoritması, RFC: Rastgele Orman Algoritması)

Renk Matrisi	LOGR	KNN	SVM	NB	DTC	RFC
Accuracy	0.66	0.66	0.66	0.12	0.60	0.72
MMC	0.54	0.56	0.54	0.08	0.51	0.64
Balanced Accuracy	0.53	0.57	0.53	0.31	0.55	0.62
F1 Score (Micro avg)	0.66	0.68	0.67	0.10	0.63	0.74
F1 Score (Macro avg)	0.56	0.60	0.56	0.10	0.54	0.65
F1 Score (Weighted avg)	0.66	0.68	0.66	0.06	0.64	0.74
Precision (Micro avg)	0.66	0.68	0.67	0.10	0.63	0.74
Precision (Macro avg)	0.61	0.68	0.66	0.38	0.54	0.72
Precision (Weighted avg)	0.66	0.69	0.67	0.56	0.64	0.74
Recall (Micro avg)	0.66	0.68	0.67	0.10	0.63	0.74
Recall (Macro avg)	0.53	0.57	0.53	0.31	0.55	0.62
Recall (Weighted avg)	0.66	0.68	0.67	0.10	0.63	0.74

## 4. Sonuç

Günümüzde depo sistemlerinin etkinliğinin artırılması amacıyla kullanılan pek çok alternatif yöntem vardır. Ancak depoyu bir veri ambarı, içindeki hammadde ve parçaları da veri olarak düşünecek olursak, veri madenciliğinin kullanımı kaçınılmaz olacaktır.

Bu çalışmada, bir depolama yöntemi olan depo bölgesi atama işleminin etkinleştirilmesi adına veri madenciliği teknikleri kullanılmıştır. Modelleme adımında da makine öğrenmesi içinde

çok sınıflı sınıflandırma algoritmalarından yararlanılmıştır. Oluşturulan model ile parçanın özellik verileri kullanılarak hangi ambar bölümüne ait olduğu tahmin edilmeye çalışılmıştır.

Yapılan veri madenciliği ve makine öğrenmesi yöntemleri sonucunda gerçek ambar bölümünü en doğru tahmin eden algoritmanın, %72'lik bir başarı oranıyla Rastgele Orman algoritması olduğu görülmüştür. Algoritmanın parametreleri bulgular bölümünde belirtilmiştir.

Sonuçlar daha detaylı incelendiğinde, genel anlamda veride bir dengesizlik söz konusudur ve bu çalışma kapsamında verideki dengesizliğin üstesinden verilerin yapısı nedeniyle gelinebilmiştir. Dengesizliğin sonucunda bazı ambar bölümlerinin doğru tahmin edilme oranı diğerlerine göre daha başarılı olurken bazıları da daha başarısız olmaktadır. Bizim görüşümüze göre bu durumun sebebi, verilerin yetersiz olmasından kaynaklanmaktadır. Buna rağmen %72'lik başarı oranı azımsanacak bir değer değildir ve çalışmalarla geliştirilebilir. Bu bilginin kullanılabilmesi için bahsi geçen firmada planlama ve envanter kontrol bölümünün, ambar bölümlerin doluluk oranları, MRP kayıtları ve planlanan parça gelişleri verilerini kullanarak atama yapacak bir sistem geliştirmesi gerekmektedir. Bu sayede daha etkin bir stoklama sistemine sahip olunacaktır.

## **5. Teşekkür**

Bu çalışmanın tamamlanmasında bana destek olan öğrencilerim Oğuzhan Alptekin ve Cansu Demirel'e teşekkürlerimi sunarım.

## **Kaynakça**

- [1] Palamutçuoğlu, B. T. Lojistik Yönetimi Ders Notları. Manisa.J, 2012.
- [2] Tompkins J. A., White J. A., Bozer Y. A., Frazelle E. H., Tanchoco J. M. A: "Facilities Planning", NJ: John Wiley and Sons, 2003.
- [3] Zhou, Xiao Guang, Na Mou, and Ling Yun Wei. "A Study for Storage Allocation in Synchronized Zones Based on the Association Analysis of Goods." In Applied Mechanics and Materials, vol. 687, pp. 4658-4665. Trans Tech Publications Ltd, 2014.
- [4] Derickx, K., & Raa, B. (2011-2012). A comparative study of different storage.
- [5] Ozkaya, U., Melgani, F., Bejiga, M. B., Seyfi, L., & Donelli, M. (2020). GPR B scan image analysis with deep learning methods. Measurement, 165, 107770.
- [6] Lam, C., Lee, C., Ho, G., & Yip, T. (2009). Development of an OLAP Based Fuzzy Logic System for Supporting Put Away Decision. International Journal of Engineering Business Management, 7-12.
- [7] Freitas, A. A. (2003). A Survey of Evolutionary Algorithms for Data Mining and Knowledge Discovery.
- [8] Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of Classification. Informatica 31 (249-268).
- [9] Mehrotra, D. (2018). Multiclass classification of mobile applications as per energy consumption. Journal of King Saud University –Computer and Information Sciences.