

**ÜRETİM YAPAN BİR İŞLETME İÇİN VERİ ODAKLI
KESTİRİMCİ BAKIM MODELLERİNİN
KARŞILAŞTIRILMASI**

**COMPARISON OF DATA DRIVEN PREDICTIVE
MAINTENANCE MODELS FOR A MANUFACTURING
COMPANY**

CEMAL AYDIN

DR. ÖĞR. ÜYESİ VOLKAN SÖNMEZ

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

Endüstri Mühendisliği Anabilim Dalı için Öngördüğü

YÜKSEK LİSANS TEZİ olarak hazırlanmıştır.

Kardeşlerime...

ÖZET

ÜRETİM YAPAN BİR İŞLETME İÇİN VERİ ODAKLI KESTİRİMCİ BAKIM MODELLERİNİN KARŞILAŞTIRILMASI

Cemal AYDIN

Yüksek Lisans, Endüstri Mühendisliği Bölümü

Tez Danışmanı: Dr. Öğr. Üyesi Volkan SÖNMEZ

Haziran 2021, 49 sayfa

Üretim yapan işletmeler için uygulanan bakım yöntemi en önemli problemlerden biridir. Endüstride uygulanan bakım stratejileri temel olarak iki grupta incelenmektedir. Bunlar düzeltici bakım ve önleyici bakımdır. Düzeltici bakım yöntemlerinde ekipman arızalandığında bakım yapılırken, önleyici bakım yöntemlerinde arıza meydana gelmeden önce bakım yapılmaktadır. Son yıllarda üretim sistemlerinde sensör teknolojilerinin gelişmesiyle, üretim ekipmanlarının basınç, sıcaklık, titreşim gibi değerleri sürekli olarak izlenebilmektedir. Bu da önleyici bakım yöntemlerinden biri olan kestirimci bakım tekniklerinin (ekipmanın durumuna göre bakım yapma/yapmama) endüstride daha yaygın kullanılmasını sağlamıştır. Ancak küçük ve orta boy işletmeler için üretim ekipmanlarının durumlarını sürekli olarak izleyebilecek ve kaydedebilecek bu tarz sensör sistemlerinin kurulması maliyetli olabilmektedir. Bu tez çalışması kapsamında da üretim ekipmanlarına yönelik bu tarz veriler olmaksızın, üretim hatlarına ait kayıp verilerinden yola çıkarak yine tahmine dayalı bakım modeli önerilmesi hedeflenmiştir.

Çalışma kapsamında 24 saat kesintisiz PVC profili üretimi yapan bir işletmeye ait vardiya/üretim hattı bazındaki kayıp verileri (kg), üretim ekipmanı çalışma hız farkları ve son bakım üzerinden geçen vardiya sayısı gibi veriler kullanılmıştır. Öncelikle planlanan kayıplardan (bakım kaynaklı kayıplardan) yola çıkarak ortalama bakım maliyetine göre bakım için bir eşik değeri (kg) belirlenmiştir. Daha sonra, üretim kayıp verilerinden, bir sonraki vardiyada çalıştırılması planlanan üretim hattı için muhtemel kayıp miktarını tahmin edecek modeller eğitilmiştir. Modellerin eğitilmesinde doğrusal regresyon, sinir ağları, rastgele orman ve gradyan artırma gibi veriden öğrenme algoritmaları kullanılmıştır. Eğitilen modellerin performansı karşılaştırıldığında en başarılı tahmin modelinin sinir ağları yöntemiyle eğitilen model olduğu görülmüştür.

Tez çalışmasının en son aşamasında çalıştırılması planlanan üretim hattı için bakım yapılıp yapılmayacağına kararının nasıl verilmesi gerektiği açıklanmıştır. Önerilen yöntemle göre çalıştırılması planlanan üretim hattındaki muhtemel kayıp (kg) hesaplanacak, kayıp (kg) değeri belirlenmiş olan eşik değeri ile karşılaştırılacaktır. Muhtemel kayıp (kg) eşik değerden büyükse bakım yapılacak, eşik değerden küçükse bakım yapılmayacaktır.

Anahtar Kelimeler: Kestirimci Bakım, İstatistiksel Öğrenme, Üretim Kayıpları, Tahmin Modelleri

ABSTRACT

COMPARISON OF DATA DRIVEN PREDICTIVE MAINTENANCE MODELS FOR A MANUFACTURING COMPANY

Cemal AYDIN

Master of Science, Department of Industrial Engineering

Supervisor: Asst. Prof. Volkan SÖNMEZ

June 2021, 49 pages

Maintenance method is one of the most important problems for manufacturing enterprises. Maintenance strategies applied in the industry are mainly examined in two groups. These are corrective maintenance and preventive maintenance. In corrective maintenance methods, maintenance is performed when equipment fails, while in preventive maintenance methods, maintenance is performed before failure occurs. With the development of sensor technologies in production systems in recent years, data of production equipment such as pressure, temperature, vibration can be monitored continuously. This has led to the more widespread use of predictive maintenance techniques (performing / not performing maintenance according to the condition of the equipment), which is one of the preventive maintenance methods. However, it can be costly for small and medium-sized enterprises to install such sensor systems that can continuously monitor and record the status of production equipment. Within the scope of this thesis, it is aimed to propose a predictive maintenance model based on the loss data of production lines without such recorded data for production equipment.

In the study, data, belonging to an enterprise that produces PVC profiles continuously for 24 hours, such as loss (kg) on the basis of shift / production line, production equipment operating speed differences and the number of shifts passed over the last maintenance, were used. First of all, a threshold value (kg) was determined for the maintenance according to the average maintenance cost based on the planned losses (losses due to maintenance). Then, models that will estimate the possible loss amount (kg) for the production line to be run in the next shift, were trained from the production loss data. Statistical learning algorithms such as linear regression, neural networks, random forest and gradient boosting were used to train the models. When the performance of the trained models was compared, it was seen that the most successful prediction model was the neural network.

In the final stage of the thesis, it is explained how to decide whether to perform maintenance or not for the production line to be operated. According to the proposed method, the possible loss (kg) in the related production line will be calculated and the estimated possible loss (kg) will be compared with the determined threshold value. If the possible loss (kg) is greater than the threshold value, maintenance will be performed, if it is less than the threshold value, no maintenance will be performed.

Keywords: Predictive Maintenance, Statistical Learning, Production Losses, Prediction Models

TEŞEKKÜR

Lisansüstü eğitimimde, tez konusu belirleme sürecindeki yardımı için ve tez çalışmam boyunca gerek akademik yönlendirme, gerekse motivasyon noktasında büyük bir sabırla sağladığı destek için, tez danışmanım ve değerli hocam Dr. Öğr. Üyesi Volkan SÖNMEZ' e

Lisans eğitimimden beri bende çok emeği olan, Hacettepe Endüstri Mühendisliği hocalarına, çalışanlarına ve bölüm arkadaşlarıma, kısaca bütün Hacettepe Endüstri Mühendisliği ailesine,

Çalışma hayatını çekilir kılan ve tez sürecinde beni sürekli motive etmeye çalışan TÜBİTAK Bilgi İşlem Daire Başkanlığı'ndan değerli mesai arkadaşlarıma,

En önemlisi hayatta her zaman ve her koşulda yanımda olan kardeşlerim Engin AYDIN, Eda AYDIN ve Şeyda AYDIN' a, bugünlere gelebilmemde büyük emeği olan anne ve babama,

Sonsuz teşekkürler...

Cemal AYDIN

Haziran 2021, Ankara

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT	iii
TEŞEKKÜR	v
İÇİNDEKİLER.....	vi
ŞEKİLLER DİZİNİ.....	viii
ÇİZELGELER DİZİNİ	ix
SİMGELER VE KISALTMALAR.....	x
1. GİRİŞ	1
1.1. Üretim Kayıpları ve Bakım Stratejileri	1
1.2. Problem Tanımı/ Hedef / Kapsam.....	2
1.3. Veriden Öğrenme Algoritmaları	2
1.4. Tez Organizasyonu.....	3
2. LİTERATÜR TARAMASI.....	4
3. METODOLOJİ.....	12
3.1. Temel Bileşen Analizi.....	12
3.2. Çoklu Doğrusal Regresyon	14
3.3. Yapay Sinir Ağları	16
3.4. Rastgele Orman	19
3.5. Gradyan Artırma	22
4. ANALİZ VE UYGULAMA	24
4.1. İşletme ve Tez Çalışmasında Kullanılan Veri.....	24
4.2. Veri Ön işleme Adımları	24
4.3. Genel Veri Analizi	26
4.4. Bakım Kararı İçin Eşik Değerin Belirlenmesi	28
4.5. TBA (Temel Bileşen Analizi) Uygulanması	29
4.5.1. Bütün Veri (1. Veri Seti) için Temel Bileşen Analizi	29
4.5.2. Sadece Belirlenen Hat (4.hat) ve Yıl (2012) için Temel Bileşen Analizi	31

4.6. Veriden Öğrenme Yöntemlerinin Uygulanması	33
4.6.1. Bütün Veri (1.Verdi Seti) Kullanılarak Modellerin Eğitilmesi	33
4.6.2. Sadece Belirlenen Hat (4.hat) ve Yıl (2012) için Modellerin Eğitilmesi	38
4.7. Modellerin Sonuçlarının Karşılaştırılması.....	42
4.8. Önerilen Bakım Modeli	44
4.9. İşletme için Önerilen Bakım Modelinin Genelleştirilmesi	46
5. SONUÇ VE ÖNERİLER.....	48
6. KAYNAKLAR	50
EKLER.....	57
EK 1- Tez Çalışması Orjinallik Raporu.....	57
ÖZGEÇMİŞ	59

ŞEKİLLER DİZİNİ

Şekil 1 – Bakım Stratejileri [4, 5]	1
Şekil 2 - TBA'nın iki boyutlu veri seti üzerinde uygulanması [44].....	12
Şekil 3 – YSA'daki bir nöronun yapısı	17
Şekil 4 – Sigmoid fonksiyon [51]	18
Şekil 5 – Çok Katmanlı İleri Beslemeli YSA yapısı	18
Şekil 6 – Rastgele Orman Modeli	21
Şekil 7 – Gradyan Artırma Algoritması	23
Şekil 8 – Bakım Yapılma Durumuna Göre Üretim Kayıp Büyüklükleri (1. Veri Seti)..	27
Şekil 9 – Bakım Yapılma Durumuna Göre Üretim Kayıp Büyüklükleri (2. Veri Seti)..	27
Şekil 10 – Kullanılan Sinir Ağı Topolojileri -1.....	34
Şekil 10 – Kullanılan Sinir Ağı Topolojileri - 2.....	35
Şekil 10 – Kullanılan Sinir Ağı Topolojileri - 3.....	35
Şekil 11 – Önerilen Bakım Modeli	45
Şekil 12 – Önerilen Bakım Modeli (Genelleştirilmiş)	47

ÇİZELGELER DİZİNİ

Çizelge 1 – Planlanan / Planlanmayan Kayıplar	24
Çizelge 2 – Hedef Değer ve Tahmin Ediciler.....	26
Çizelge 3 – TBA Sonrası Bileşenler (Bütün Veri)	29
Çizelge 4 – TBA Sonrası Özvektör Matrisi (Bütün Veri)	29
Çizelge 5 – Seçilen Bileşenlerden Oluşturulan Veri Seti Yapısı (Bütün Veri)	30
Çizelge 6 – TBA Sonrası Bileşenler (4. Hat ve 2012 yılı için)	31
Çizelge 7 – TBA Sonrası Özvektör Matrisi (4. Hat ve 2012 yılı için)	31
Çizelge 8 – Seçilen Bileşenlerden Oluşturulan Veri Seti Yapısı (4. Hat ve 2012 yılı için) .	32
Çizelge 9 – Doğrusal Regresyon Sonuçları (Bütün Veri)	33
Çizelge 10 – Yapay Sinir Ağları Sonuçları (Bütün Veri).....	35
Çizelge 11 – Rastgele Orman Modellerinin Sonuçları (Bütün Veri)	36
Çizelge 12 – Seçilen Modelin Sonuçları (Rastgele Orman - Bütün Veri).....	37
Çizelge 13 – Gradyan Artırma Modellerinin Sonuçları (Bütün Veri).....	37
Çizelge 14 – Seçilen Modelin Sonuçları (Gradyan Artırma - Bütün Veri).....	38
Çizelge 15 – Doğrusal Regresyon Sonuçları (4.hat 2012 Yılına Ait Veri)	39
Çizelge 16 – Yapay Sinir Ağı Sonuçları (4.hat 2012 Yılına Ait Veri).....	39
Çizelge 17 – Rastgele Orman Modellerinin Sonuçları (4.hat 2012 Yılına Ait Veri)	40
Çizelge 18 – Seçilen Modelin Sonuçları (Rastgele Orman - 4.hat 2012 Yılına Ait Veri) ...	40
Çizelge 19 – Gradyan Artırma Modellerinin Sonuçları (4.hat 2012 Yılına Ait Veri).....	41
Çizelge 20 – Seçilen Modelin Sonuçları (Gradyan Artırma - 4.hat 2012 Yılına Ait Veri)..	41
Çizelge 21 – Karşılaştırmalı Sonuçlar (Bütün Veri)	43
Çizelge 22 – Karşılaştırmalı Sonuçlar (4.hat 2012 Yılına Ait Veri)	43

SİMGELER VE KISALTMALAR

Simgeler

D	Veri Seti
D'	Standartlaştırılmış Veri Seti
D''	Temel Bileşen Analizi Sonrası Veri Seti
v	Özvektörler kümesi
V	Özvektör Matrisi
λ	Özdeğerler kümesi
β	Katsayı Vektörü
ε	Hata Terimi

Kısaltmalar

ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
IoT	Internet of Things
KOKH	Kök Ortalama Karesel Hata
KOKLH	Kök Ortalama Karesel Logaritmik Hata
OAS	Ortalama Artık Sapma
OKH	Ortalama Karesel Hata
OMH	Ortalama Mutlak Hata
SDF	Sembolik Dinamik Filtreleme
TBA	Temel Bileşen Analizi
TÜB	Toplam Üretken Bakım
YSA	Yapay Sinir Ağları

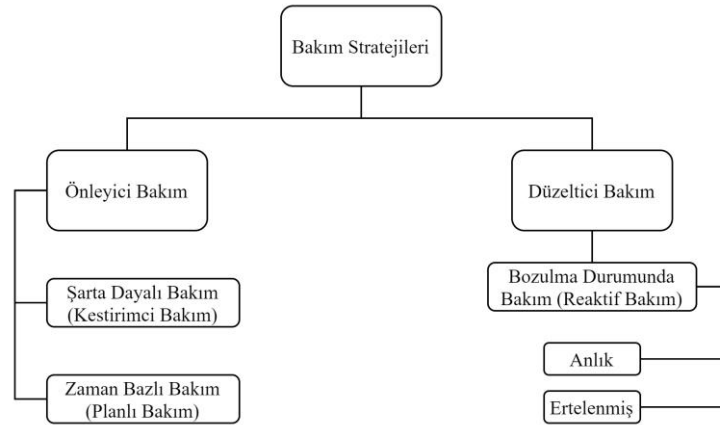
1. GİRİŞ

1.1. Üretim Kayıpları ve Bakım Stratejileri

Üretim yapan işletmeler için üretim kayıpları üretim verimliliği açısından ele alınan başlıca problemlerden biridir. Bu kayıplar, genellikle, ekipman arızaları, kurulum ve ayarlamalar, küçük-boş duruşlar, hız düşüşleri, kalite/yeniden işlemler, randıman kayıpları olmak üzere altı kategoride gruplandırılabilir [1].

Üretimde meydana gelen kayıpları en aza indirmenin yolu işletmede iyi bir bakım felsefesi uygulamaktır. Nakajima, Japon üretim felsefesinden doğan ve pek çok farklı sektörde uygulanan, arızayı oluşmadan önce önlemeyi amaçlayan, tüm çalışanları bakım sürecine dahil eden Toplam Üretken Bakım (TÜB) kavramını 1980'li yıllarda geliştirmiştir [2].

Bakım yöntemleri temel olarak düzeltici bakım ve önleyici bakım olmak üzere iki gruba ayrılmaktadır. Düzeltici bakımda bakım işlemi arıza meydana geldikten sonra yapılırken, önleyici bakımda arıza meydana gelmeden önce yapılmaktadır. Önleyici bakımın da kendi alt türleri bulunmaktadır. Bunlar şarta dayalı bakım ve zaman bazlı bakım olarak özetlenebilir [3].



Şekil 1 – Bakım Stratejileri [4, 5]

Son yıllarda üretim yapan işletmelerde gelişen sensör sistemleriyle birlikte üretim ekipmanlarına yönelik veri toplanması ve ekipman durumunun izlenmesi kolaylaşmış, kestirimci bakım yöntemleri diğer bakım yöntemlerine göre daha çok ön plana

çıkılmıştır. Kestirimci bakım teknikleri izlenen üretim sürecinin mevcut durumundan hareketle bakımın ne zaman yapılması gerektiğini tahmin etmeye dayalıdır. Bu yaklaşım rutin veya zamana dayalı diğer önleyici bakım tekniklerine göre bakım maliyetinde avantaj sağlamaktadır. Yalnızca belirli durumlar oluştuğunda bakım işlemi yapılacağından parçanın bozulma durumunun tahminine göre yapılan şarta dayalı bakım olarak da kabul edilebilir [6, 7] .

1.2. Problem Tanımı/ Hedef / Kapsam

Bu tezde, üretim yapan gerçek bir işletmenin üretim hattında geçmişte meydana gelen arıza/kayıp verilerinden yola çıkarak veri odaklı kestirimci bakım modelinin önerilmesi hedeflenmiştir.

Bir işletmede üretim yapan hat/makine/ekipmanlara ait sürekli izlenebilen sensör verileri olmaksızın, geçmiş kayıp büyüklüklerinden ve hatlardaki makine/ekipman çalışma hızlarından yararlanarak, kestirimci bakım modeli ya da bakım için karar destek modeli geliştirilip geliştirilemeyeceği sorusuna cevap aranmıştır.

Bu çalışmada gün – vardiya ve hat bazında, kg biriminden kaydedilen üretim kayıp verilerinden, son bakım üzerinden geçen süreden, son bakımdan sonra oluşan üretim kaybından ve hat bazında çalışan ekipman hızları farkından bir sonraki vardiyada oluşabilecek üretim kayıp büyüklüğünün tahmin edilmesi amaçlanmıştır. Tahmin edilen üretim kaybının büyüklüğüne göre bakım yapma ya da yapmama kararının verilebileceği düşünülmüştür.

Üretim kayıplarının büyüklüğünün tahmin edilmesinde önerilecek modellerin, veriden öğrenme algoritmaları kullanılarak oluşturulması planlanmıştır.

1.3. Veriden Öğrenme Algoritmaları

Veriden öğrenme algoritmaları genel olarak üç kategoride incelenmektedir [8]. Bunlar;

1. Denetimli Öğrenme
2. Denetimsiz Öğrenme
3. Pekiştirmeli Öğrenme

olarak sıralanabilir. Denetimli Öğrenme Algoritmaları hem girdi, hem çıktılarını içeren bir veri kümesinden matematiksel tahmin modeli oluştururken, Denetimsiz Öğrenme Algoritmaları yalnızca girdiler içeren bir veri kümesinden gruplama ve kümeleme gibi yöntemlerle verideki yapıyı çıkartır. Pekiştirmeli Öğrenme Algoritmaları ise temelde Markov Karar Süreci'ne dayanan, deneme yanılma yöntemiyle en yüksek ödül miktarına ulaşmayı hedefleyen hareket/davranış öğrenimi olarak tanımlanabilir.

Denetimli Öğrenme algoritmaları da kendi içerisinde iki grupta incelenmektedir. Bunlardan sınıflandırma algoritmaları, hedeflenen değişkenin sınıfını tahmin etmek amacıyla kullanılırken, regresyon algoritmaları hedeflenen değişkenin değerini tahmin etmek amacıyla kullanılmaktadır. Diğer bir ifadeyle çıktı değişkeni yalnızca belirli değerleri alabiliyorsa sınıflandırma algoritmaları, çıktı değişkeni belirli bir aralık içinde herhangi bir sayısal değer alabiliyorsa regresyon algoritmaları kullanılır [9].

Bu tez çalışması kapsamında, gün – vardiya ve hat bazında üretim kayıp miktarları (kg) değerleri veri setinde etiketli olarak bulunduğu ve bir sonraki vardiyada oluşabilecek kayıp büyüklüğü sayısal olarak tahmin edilmek istendiğinden yalnızca 'Regresyon için Denetimli Öğrenme Algoritmaları' kullanılacaktır. Denetimli Öğrenmede belirtilen regresyon problemi için Çoklu Doğrusal Regresyon, Yapay Sinir Ağları, Rastgele Orman ve Gradyan Artırma algoritmaları kullanılmıştır.

1.4. Tez Organizasyonu

Tezin ikinci bölümünde kestirimci bakım ile ilgili daha önce yapılmış çalışmalar kısaca özetlenmiştir.

Tezin üçüncü bölümünde, ilk olarak, veriden öğrenme öncesinde kullanılan temel bileşen analizi yöntemi anlatılmıştır. Sonrasında üretim kayıp (kg) miktarını tahmin edecek modellerin eğitilmesinde kullanılan veriden öğrenme algoritmaları anlatılmıştır.

Tezin dördüncü bölümünde sürekli üretim yapan bir işletmenin geçmişteki üretim kayıp verilerinden yola çıkarak, bir sonraki hat/ vardiya için üretim kayıp miktarını (kg) tahmin edecek modeller eğitilmiş, bu modellerin tahmin performansları karşılaştırılmıştır. En başarılı tahmin modeli ile bakım kararının nasıl verileceğine ilişkin bir çerçeve sunulmuş ve bu çerçeve genelleştirilmiştir.

Tezin sonuç bölümünde tezden elde edilen sonuçlar özetlenmiş, tezin sağlayabileceği katkı belirtilmiştir.

2. LİTERATÜR TARAMASI

Kestirimci bakım ile ilgili yapılan çalışmaların birçoğu veriden öğrenme yöntemlerinin kullanılmasıyla gerçekleştirilebilmiştir. Bu çalışmalarda genellikle üretim ekipmanının durumunu tahmin etmeyi amaçlayan modellerin oluşturulması hedeflenmiştir. Bu modellerin oluşturulmasındaki en büyük zorluk, üretim ekipmanlarına yönelik verilerin toplanması olmuştur. Endüstride son yıllarda özellikle IoT teknolojilerinin kullanımının artmasıyla makinelerin basınç, sıcaklık, titreşim gibi sensör verileri daha kolay ve daha çok toplanabilmiş bu da bakım planlamada veri odaklı kestirimci bakım tekniklerinin yaygınlaşmasını sağlamıştır [10].

Jay Lee [11, 12] çalışmalarında makine bozulmaları sırasındaki sensör verilerinden sinir ağı modellerinin eğitilebileceğini ve eğitilen modeller ile makine performanslarının izlenebileceğini belirtmiştir. Bilgisayarla entegre üretim sistemleri için makine arızalarını önceden tahmin edebilecek proaktif bir bakış açısı sunmuştur.

Richard C. M. Yam vd. [13] o güne kadar olan geleneksel şarta dayalı bakım yöntemine 'Kestirimci Akıllı Karar Destek Sistemi' adını verdikleri bir çerçeveye entegre etmişlerdir. Şarta dayalı akıllı arıza teşhisi ve ekipman bozulma eğiliminin tahminini sağlayan bir model ile şarta dayalı bakım yöntemini desteklemişlerdir. Ekipmandaki sıcaklık ve titreşim verilerine göre geri beslemeli yapay sinir ağı kullanarak geliştirdikleri modeli gerçek ekipman üzerinde test etmişlerdir. Aldıkları sonuçlara göre önerdikleri 'Kestirimci Akıllı Karar Destek Sistemi' modelinin ekipman bozulma eğiliminde güvenilir arıza teşhisi ve yüksek tahmin gücü sağladığını görmüşlerdir.

Kestirimci bakım ile ilgili çalışmalarda bakım kaynaklı maliyeti minimize etmek için optimizasyon yöntemleri de kullanılmıştır. Bu çalışmalarda ekipman için optimal bakım seviyesini belirleyecek optimizasyon modelleri oluşturulmuş ve çözülmüştür.

Marzio Marseguerra vd. [14] çalışmalarında, bakım işlemindeki kârı ve kullanılabilirliği maksimize edecek optimal bozulma seviyesini (bakım yapılması gereken seviye) belirlemeyi amaçlamışlardır. Bunun için çok kriterli optimizasyon problemi oluşturmuşlar ve modelledikleri probleme genetik optimizasyon ile çözüm sunmuşlardır. Makinedeki bozulma sürecini Monte Carlo simülasyonu ile gerçekleştirerek buldukları çözümün geçerliliğini test etmişlerdir.

Antoine Grall vd. [15] çalışmalarında kademeli olarak bozulan tek üniteli bir sistem için kestirimci bakım yapısı sunmuşlardır. Gereksiz bakım ya da ekipman arızası durumunda oluşan bakım kaynaklı maliyeti dengelemeyi amaçlamışlardır. Sistemin durumundan hareketle optimal ekipman kontrol ve ekipman değiştirme kararını sağlayacak bir model önermişlerdir. Modellerinde optimizasyon değişkenlerini 'Ekipman değişikliği eşik değeri' ve 'Periyodik Muayene Süresi' olarak belirlemişler ve değişkenlerin ortak optimizasyonu ile bakım maliyetinin en aza indirilebileceğini göstermişlerdir.

J.M. van Noortwijk [16] çalışmasında güvenilirlik alanında çokça araştırma konusu olmuş gamma prosesinin bakımda uygulanmasını araştırmıştır. Stokastik bozulmayı modelleyerek bakım maliyetini optimize etmek amacıyla kullanılan gamma prosesinin farklı bakım modellerinde kullanabileceğini göstermiştir. Gamma prosesinin bozulmadaki geçici değişkenliği iyi modellediğini belirtmiş, ekipman için optimal bakım kararını vermede kullanabileceğini göstermiştir.

Jay Lee vd. [17] çalışmalarında üretimdeki makinelerin ve diğer araçların performanslarındaki bozulma seviyesini nicel olarak değerlendirmek ve tahmin etmek amacıyla genel ve ölçeklenebilir bir akıllı bakım aracı geliştirmişlerdir. Doğrudan endüstride kullanılmak üzere geliştirdikleri bu araçta bakım kaynaklı üretim kayıplarını sifıra yakın seviyeye indirmeyi amaçlamışlardır. Geliştirdikleri bu araçta performans davranışını tahmin etmek için ARMA modelini kullanmışlardır.

Kestirimci bakım yönteminde önemli araştırma konularından bir tanesi de bakım yapılacak makine/araç/sistemin kalan kullanışlı ömür süresinin (Remaining Useful Life - RUL) tahmin edilmesidir. Tez kapsamında yapılan sınırlı literatür incelemesi değerlendirildiğinde, kestirimci bakım ile ilgili yapılan araştırmaların büyük bir bölümünün bu konuya odaklandığı görülmektedir. Kestirimci bakımda kalan kullanışlı ömür süresinin tahmin edilmesinde kullanılan yöntemlere bakıldığında ise model tabanlı yaklaşım ve veri odaklı yaklaşım olmak üzere iki yaklaşımın olduğu görülmektedir. Model tabanlı yaklaşımda tahmin modeli fiziksel yasalara ve matematiksel formüllerden çıkarılırken, veri odaklı yaklaşımda tahmin modeli geçmiş verilerden çeşitli öğrenme algoritmaları ile çıkarılmaktadır [18].

Ömer Faruk Eker vd. [19] literatürde kestirimci bakım konusunda çokça kullanılan deneme problemleri ve veri setleri üzerinde çalışmışlar ve bu problemlerin/veri

setlerinin model tabanlı yaklaşıma veya veri odaklı yaklaşıma uygunluğunu değerlendirmişlerdir.

Xiao-Sheng Si vd. [20] çalışmalarında kritik sistemler için kalan kullanışlı ömür süresinin tahmini için bozulma sırasında toplanan sinyal verilerine dayanan, doğrusal olmayan bir bozulma modeli oluşturulmuşlardır. Modelde bilinmeyen parametreleri maksimum olabilirlik yöntemine göre hesaplamışlardır. Araştırmalarında önerdikleri doğrusal olmayan modelin kestirimci bakımda, kalan kullanışlı ömür süresinin tahminindeki doğruluğu artırdığını gözlemlemişlerdir.

Kai Goebel vd. [21] çalışmalarında lityum iyon pillerinin kalan ömrünü tahmin etmek için model bazlı ve veri odaklı istatistiksel yöntemler kullanmışlardır. Lityum iyon pillerinin iç model verilerini kullanarak olasılıksal regresyon modeli ve parçacık filtresi modelini uygulamışlar ve bu modellerin tahmin başarısını karşılaştırmışlardır.

Bhaskar Saha vd. [22] çalışmalarında pillerin kalan ömrünün tahmininde bayes tabanlı olasılıksal öğrenme çatısı entegre etmiş, uygunluk vektör makineleri (relevance vector machines) ve parçacık filtresi modelini uygulamışlardır.

Bhaskar Saha vd. [23] diğer çalışmalarında aynı problem için ARIMA(1), genişletilmiş kalman Filtresi (2), parçacık filtresi-uygunluk vektör makineleri(3), algoritmalarını karşılaştırmışlar ve problem için en uygun çözümün parçacık filtresi-uygunluk vektör makineleri kombinasyon modeli olduğunu belirtmişlerdir.

Üretim süreçlerinde gerçekte makine/araç/sistemin bozulmasına genellikle mücade edilmeden, bakım yapılmaktadır. Bundan dolayı tahmin modelinin eğitileceği veride yeterince bozulmanın olmaması modelin başarısını oldukça etkilemektedir. Bu da kestirimci bakımda veriden öğrenme yöntemlerinin uygulanabilmesinin önündeki en büyük problemlerden biridir.

Abhinav Saxena vd. [24] çalışmalarında bu konuya değinerek, uçak gaz türbini motorunun bozulma sürecini simüle etmişlerdir. Bunun için C-MAPSS aracını kullanarak, simülasyon ortamının varsayımlarını ve parametrelerini ayarlamışlar, uçak gaz türbini motorunun bozulma sürecini birçok kez simüle etmişlerdir. Simülasyonda bozulma sırasındaki sensör verilerini zaman serisi şeklinde toplamışlardır. Uçak türbini simülasyonundan elde edilen sensör verileri, kestirimci bakım konusunda veri odaklı

çözümler sunmak amacıyla uluslararası prognostik konferansında gerçekleştirilen yarışmada ve daha sonraki akademik çalışmalarda kullanılmıştır.

Giuseppe Curcurù vd. [25] çalışmalarında global bakım maliyetini en aza indiren bakım süresinin hesaplanması için alternatif bir prosedür önermişlerdir. Bozulma süreci için stokastik bir model benimsemişler ve kusurlu ekipman izleme sisteminde önceki veriye dayanan bayesçi yaklaşım tabanlı kestirimci bakım yöntemi geliştirmişlerdir. Yönetimsel gereksinimleri ve izleme sistemindeki eksiklikleri de hesaba katan bu yöntemin klasik önleyici bakım yöntemine göre daha etkin olduğunu değerlendirmişlerdir.

Ming-Yi You vd. [26] çalışmalarında, bir dizi tekli sistem için istatistiksel olarak planlanmış ve tekil olarak geliştirilmiş bir kestirimci bakım politikası önermişlerdir. Uzun dönemli planlama için, klasik istatistiksel dağılım temelli önleyici bakım politikalarının yeteneği ile kestirimci bakımdaki tekil bir sistemin kalan ömrünü tahmin etme yeteneğini harmanlamışlardır. Önerdikleri yöntemi klasik önleyici bakım yöntemiyle karşılaştırmışlar ve daha iyi bir sonuç verdiğini görmüşlerdir.

Soumik Sarkar, vd. [27] çalışmalarında zaman serisi şeklindeki sensör verileri için sembolik dinamik filtreleme (SDF) bazlı anomali bulma algoritması önermişlerdir. Bu problemi çok sınıflı sınıflandırma problemi olarak ele almışlar, verideki gürültüyü azaltarak sınıflandırmadaki doğruluk oranını maksimize etmeye çalışmışlardır. Geliştirdikleri yöntemi daha önce yayınlanan C-MAPSS veri seti üzerinde uygulamış ve test etmişlerdir.

Rosmaini Ahmad vd. [28] çalışmalarında literatürde çokça tartışılan zamana dayalı bakım ve duruma dayalı bakım stratejilerini ayrı ayrı incelemişler, bu stratejilere göre nasıl bakım kararının verildiğini açıklamışlardır. Sonrasında bu iki temel stratejinin pratik olarak uygulanmasında karşılaşılan zorlukları karşılaştırmışlardır. Bu karşılaştırmayı bakımla ilgili gerekli verinin belirlenmesi, toplanması, toplanan verinin analizi, bakım modelinin oluşturulması ve bakım kararının verilmesi süreçlerine göre yapmışlardır. Duruma dayalı bakım stratejisini daha yararlı ve efektif bulduklarını belirtmişler, ancak bu stratejinin uygulanabilmesi için ekipmanlara yönelik sensör tabanlı teknolojilere ihtiyaç bulunduğunu belirtmişlerdir.

Kaibo Liu, vd. [29] çalışmalarında birden çok sensör verisi içeren sistemlerde, nitelik seviyesinde ve karar seviyesinde birleştirme yerine veri seviyesinde birleştirme modeli oluşturmuşlardır. Bunun için bozulma sürecine ait birden fazla zaman serisi verisini birleştirecek 'birleşik sağlık endeksi' modeli oluşturmuşlar, daha sonra oluşturdukları modeli daha önce yayınlanan C-MAPSS simülasyonu veri seti (uçak gaz türbini bozulma sürecine ait zaman serisi verisi) üzerinde test etmişlerdir.

Adriaan Van Horenbeek vd. [30] çoklu bileşen içeren ekipmanlar için bileşenlerin birbirine bağımlılıklarını (ekonomik, yapısal ve stokastik bağımlılıklar) da göz önüne alan, prognostik / tahmine dayalı, dinamik kestirimci bakım planlaması önermişlerdir. Uzun vadedeki birim zaman başına ortalama bakım maliyetini minimize etmeyi amaçlamışlardır. Sunulan bakım planlama stratejisinin performansını literatürdeki diğer beş geleneksel bakım politikaları ile karşılaştırmışlar, sunulan bakım planlama stratejisinin daha iyi performans sergilediğini belirtmişlerdir.

Hongfei Li vd. [31] çalışmalarında demiryolu ağ hızını etkileyen başlıca etkenlerden birinin, bakım çalışmalarından kaynaklanan hizmet kesintileri olduğunu belirtmişlerdir. Demiryollarındaki durumun izlenmesini sağlayan sıcaklık, gerilim, görüntü, kızılötesi, ağırlık, etki gibi tarihsel sensör verileri ile bozulma, bakım işlemi, hava sıcaklığı gibi verileri analiz etmişlerdir. İlgili verileri harmanlayarak dağıtık öğrenme ve analitik hiyerarşik öğrenmeye dayanan, hem tarihsel hem de gerçek zamanlı veriler üzerinden bozulmaları tahmin edecek modeller geliştirmişlerdir.

T. Praveenkumar vd. [32] otomobil viteslerinde oluşacak hataları önceden tahmin etmeyi sağlayacak bir model geliştirmeye çalışmışlardır. Bunun için titreşim verilerini kullanmışlar, Destek Vektör Makineleri yöntemiyle arıza teşhisini sağlayacak bir model eğitmişlerdir. Modeli iki farklı hız ve yüklenme koşullarında test etmişler ve sonuçları başarılı bulmuşlardır.

Kestirimci bakımdaki zorluklardan biri nicel göstergeleri belirlemek ve bunları çalıştırma maliyeti ve bozulma riski ile ilişkilendirmektir. Gian Antonio Susto vd. [33] bunun çözümüne yönelik bakım kararı için dinamik karar destek sistemi geliştirmişlerdir. Bakım kaynaklı beklenmeyen duruşların sıklığı ve atıl kullanım ömrü arasındaki denge açısından farklı performans çıktıları sağlayan, birden fazla sınıflandırıcıyı paralel olarak eğiten, sonrasında bu sınıflandırma çıktıları ile beklenen

maliyetleri dinamik olarak minimize eden bir metodoloji sunmuşlardır. Sınıflandırmada k en yakın komşu ve destek vektör makineleri algoritmalarını kullanmışlar, sonuçları akademik çalışmalarda karşılaştırma amacıyla kullanılan yarı iletken üretimindeki bakım verisi üzerinde test etmişlerdir.

Rune Prytz vd. [34] çalışmalarında taşımacılıkta kullanılan ticari kamyonlarda kritik bir parça olan hava kompresörlerinin bakım ihtiyacını denetimli makine öğrenmesi teknikleri kullanarak tahmin etmeye çalışmışlardır. Bunun için 2010 - 2013 yılları arasında üretilmiş 65000 Volvo marka kamyonun servis kayıtlarını ve servis sırasındaki araç değer verilerini kullanmışlardır. Başka amaçlar için kullanılan bu verilerden kestirimci modeller oluşturmaya çalışmışlardır. Bunun için tahmin edilen kalan kullanışlı ömür süresine göre t anında bakım yapıp yapmama kararını verecek modeli Rastgele Orman algoritması ile eğitmişlerdir. Gelişmiş sensör verileri olmaksızın servis kayıt verileri ve araç değer verilerini kullanarak bakım için pratik bir karar destek modeli önermişlerdir.

Koşul temelli bakım stratejisinde çok bileşenli sistemlere uygulanan bakım kararları sistem seviyesinde ya da bileşen seviyesinde uygulanmaktadır. Bu durum bakım kararında en iyi performansı vermeyebilir. Khac Tuan Huynh vd. [35] çalışmalarında bu durumu gidermek amacıyla bakım için çok seviyeli karar verme yaklaşımı sunmuşlardır. Bunun için bileşen seviyesinde ve sistem seviyesindeki karar verme yaklaşımlarını harmanlamışlar ve iki farklı kestirimci bakım stratejini karşılaştırmışlardır. Bakım kararı, ilkinde göre bileşenlerin kalan kullanışlı ömür süresine dayanırken, ikincisine göre ise hem bileşenlerin hem de sistemin kalan kullanışlı ömür süresine dayanmıştır. Karşılaştırma sonucunda çok bileşenli sistemler için, sundukları çok seviyeli bakım kararı yaklaşımının bileşen seviyesinde yapılan bakım kararına göre daha iyi sonuç verdiğini gözlemlemişlerdir.

Pin Lim, vd. [36] çalışmalarında çok katmanlı sinir ağları modelini kullanarak kalan kullanışlı ömür süresini hesaplamaya çalışmışlardır. Sundukları çatı sırasıyla; veri normalizasyonu, öznitelik seçimi, hareketli zaman penceresi, özellik çıkarma, çok katmanlı sinir ağı gibi aşamaları içermiştir. Kurdukları çatıyı C-MAPSS veri seti üzerinde uygulamış ve literatürdeki önceki modellerin sonuçlarından daha iyi bir sonuç elde etmişlerdir.

Xiang Li, vd. [37] çalışmalarında uzmanlık bilgisi gerektiren model bazlı kestirimci bakım yöntemleri yerine salt veri odaklı tekniklerin uygulanmasının daha pratik olduğunu öne sürmüşler, C-MAPSS veri seti için derin evrimsel sinir ağları kullanarak alternatif bir model eğitmişler ve kestirimci bakım için kalan kullanışlı ömür süresini tahmin etmişlerdir.

Bo Luo, vd. [38] zamanla değişen üretim koşulları altında erken arıza tespiti sağlayan bir metot geliştirmişlerdir. Çalışmalarında, üretim yapan CNC tezgahının 288 günlük titreşim sinyallerinden, itki yanıtlarını otomatik olarak seçebilecek derin öğrenme modeli oluşturmuşlardır. Sonrasında bu itki yanıtlarından mekanik arızayı erken tespit edecek dinamik özellikleri tanımlamışlar ve özellikler arasındaki benzerliğe göre durum indeksi belirlemişlerdir. Belirledikleri durum indeksinden yola çıkarak CNC tezgahının ne zaman arıza vereceğini tahmin eden model önermişler ve modeli test ederek başarılı olduğunu görmüşlerdir. Böylelikle zamanla değişen koşullar altında, değişkenlikten etkilenmeyecek şekilde erken arıza tespiti yapılabileceğini göstermişlerdir.

Bin Zhang, vd. [39] makinelerdeki rulman performansındaki düşüşü değerlendirecek bir metot geliştirmişlerdir. Tekrarlayan sinir ağı (RNN) tabanlı bir model oluşturarak rulmanın arıza yayılım bilgisini kapsamlı bir şekilde kullanmayı amaçlamışlardır. Titreşim tepki mekanizmasına dayalı bozulma simülasyonu gerçekleştirerek, dalga biçimi entopisi (WFE) adında yeni bir gösterge kullanmışlardır. Daha sonra bu gösterge ve geleneksel diğer göstergeleri tekrarlayan sinir ağı (RNN) modeline girdi olarak sunmuşlar, rulmanın bozulma durumunu ve kalan kullanım ömrünü başarılı bir şekilde tahmin ettiğini görmüşlerdir.

Alberto Diez-Olivan, vd. [40] endüstriyel bakımda veri birleştirme ve makine öğrenmesi konularıyla ilgili yapılan güncel çalışmaları toparlamışlar, araştırma konularının hangi alanlarda yoğunlaştığını belirtmişlerdir. Bu amaçla literatürde kullanılan öznitelik çıkarma ve makine öğrenmesi yöntemlerini; arızaya sebep olan durumu anlamaya yönelik yöntemler (tanımlayıcı), arızanın ne zaman olabileceğini tahmin etmeyi amaçlayan yöntemler (tahmin edici), oluşması muhtemel arızanın etkisini minimize etmeyi amaçlayan yöntemler olmak üzere üç kategoride sınıflandırmışlardır.

Radhya Sahal, vd. [41] büyük veri ve büyük verinin işlenmesi için mevcut açık kaynak teknolojilerinin güçlü ve zayıf yanlarını gözden geçirmişler, endüstri 4.0 problemlerinde

kullanımlarını belirlemek için sistematik bir metodoloji kullanmışlardır. Demiryolu taşımacılığı ve rüzgar enerjisi alanlarında kestirimci bakımın seçilmiş iki kullanım durumu için gereksinimleri belirlemişler, kullanım senaryoları için büyük veri teknolojilerinin bazı optimal kombinasyonlarını önermişlerdir.

Biao Wang, vd. [42] çalışmalarında makine rulmanlarının kalan kullanışlı ömür süresini tahmin etmek için kendi geliştirdikleri hibrit bir yaklaşım önermişlerdir. Rulmanların bozulma verisinden yola çıkarak Frechet mesafesi ile birleştirilmiş üstel bozulma modelleri kullanarak tahmin modeli oluşturmuşlar, modeli PRONOSTIA rulman veri seti üzerinde denemişlerdir. Geliştirdikleri yöntemin tahmin doğruluğunu artırdığını belirtmişlerdir.

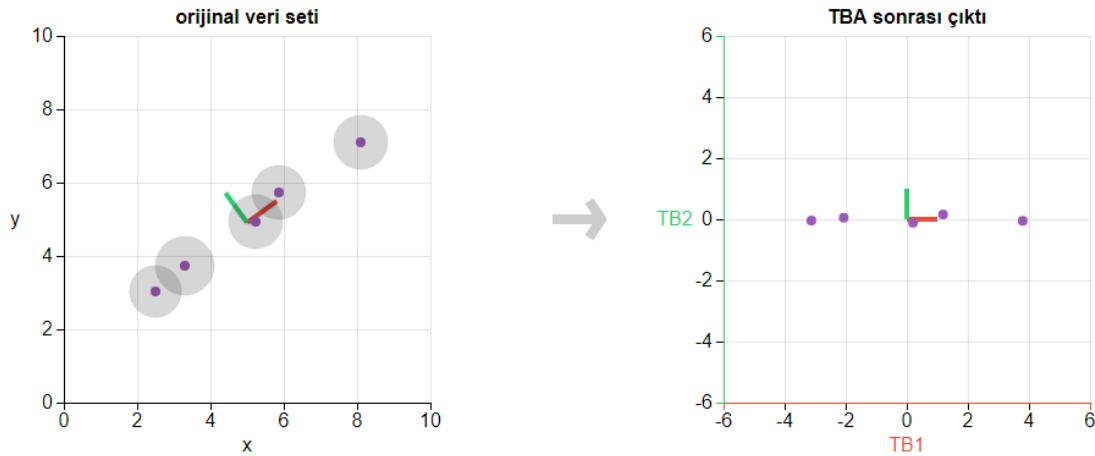
Kestirimci bakım ile ilgili yapılan çalışmalar genel olarak değerlendirildiğinde, çalışmaların büyük bölümünün, ekipman durumunu veya kalan kullanışlı ömür süresini tahmin etmeye yönelik olduğu görülmektedir. Bu tahminler genel anlamda model tabanlı yaklaşım ve veri odaklı yaklaşım olmak üzere iki tip yaklaşımla gerçekleştirilmiştir. Model tabanlı yaklaşımda, ekipmanın fiziksel durumunun formüle edilerek modellenmesi ve bu modele göre tahmin yapılması gerekmektedir. Bunun için söz konusu ekipman hakkında yüksek düzeyde teknik bilgi ve uzman bilgisine ihtiyaç bulunmaktadır. Bu durumdan dolayı çalışmalar büyük ölçüde veri odaklı yaklaşımlar üzerine yoğunlaşmıştır. Veri odaklı yaklaşımda tahmin modelleri veriden öğrenme yöntemleri ile oluşturulduğundan, bu modellerin eğitilmesi için ilgili ekipmana yönelik sıcaklık, basınç, titreşim gibi verilere ihtiyaç duyulmuştur. Bu veriler ekipmanın çalışmasından bozulmasına kadar geçen sürede çeşitli sensör sistemlerinden toplanmış ya da gerçek ortamdaki bozulma simülasyon yoluyla gerçekleştirilerek elde edilmiştir. Tez çalışması kapsamında yapılan literatür taramasında, veri odaklı kestirimci bakım yöntemleri ile ilgili çalışmaların büyük çoğunluğunun ekipmanın durumunu anlatan sensör verilerine dayandığı görülmüştür. Ancak bu verileri elde etmek için kurulması ve izlenmesi gereken sensör sistemleri oldukça maliyetlidir. Üretim ekipmanına yönelik bu tarz sensör verileri olmaksızın, üretim hattındaki üretim kayıp verilerinden yola çıkarak oluşabilecek muhtemel üretim kaybına göre bakım yöntemi öneren bir çalışmaya yapılan literatür taramasında rastlanmamıştır. Bu çalışma kapsamında da böyle bir bakım yöntemi önerilmesi hedeflenmiştir.

3. METODOLOJİ

3.1. Temel Bileşen Analizi

Temel bileşen analizi; değişken setindeki varyansın büyük çoğunluğunu açıklayabilen daha az sayıda ortogonal değişken (aralarında korelasyon olmayan) kullanarak değişken setinin boyutunu düşürmeyi amaçlayan bir prosedürdür. Bu işlemde birbiriyle büyük oranda ilişkili değişkenler temel bileşen olarak adlandırılan doğrusal bağımsız değişkenlere dönüştürülür [43].

TBA iki boyutlu bir veri setinde aşağıdaki gibi görselleştirilebilir [44]. İlgili veri setinde x ve y eksenleri uzunluk ve ağırlığı ifade etmektedir. Orijinal veri seti varyans açısından incelendiğinde varyansın büyük ölçüde tek yönlü olduğu görülmektedir. TBA sonrasında her noktanın yeni bir (x,y) değerine sahip olduğu yeni bir koordinat sistemi yaratılmış olur. Yeni koordinat sisteminde eksenler fiziksel olarak anlamlı değişkenleri ifade etmemekte; orijinal veri setindeki uzunluk ve ağırlık değişkenlerinin birer kombinasyonunu belirtmektedir. Kombinasyonla yaratılmış, ortogonal olan (aralarında korelasyon olmayan) yeni değişkenlerin her biri temel bileşen olarak adlandırılır. TBA sonrasındaki koordinat sistemi incelendiğinde verideki değişkenliğin TB1 (1. temel bileşen) ile büyük ölçüde açıklanabildiği ve TB2'nin (2. temel bileşen) göz ardı edilebileceği belirtilmektedir.



Şekil 2 - TBA'nın iki boyutlu veri seti üzerinde uygulanması [44]

TBA algoritması adımları şu şekilde özetlenebilir [45] .

a- Veri setini standardize et:

N boyutlu orijinal veri setindeki (D) her bir değişken kendi içinde $\mu = 0$ ve $\sigma = 1$ olacak şekilde standartlaştırılır.

$$x_{\text{standart}} = \frac{x - \mu}{\sigma}$$

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1N} \\ d_{21} & d_{22} & \dots & d_{2N} \\ d_{31} & d_{32} & \dots & d_{3N} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \Rightarrow D' = \begin{bmatrix} d'_{11} & d'_{12} & \dots & d'_{1N} \\ d'_{21} & d'_{22} & \dots & d'_{2N} \\ d'_{31} & d'_{32} & \dots & d'_{3N} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

b- Veri setindeki değişkenler arasındaki kovaryans matrisini hesapla:

Standardize edilmiş veri seti üzerinden (D') bütün değişkenlerin birbiri arasındaki ilişkiyi gösteren kovaryans matrisi (A) hesaplanır. $(a_1, a_2, a_3, \dots, a_N)$ veri setini oluşturan N tane değişken olmak üzere

$$\text{Cov}(x, y) = \frac{\sum (x_i - \bar{x}) * (y_i - \bar{y})}{N - 1}$$

$$A = \begin{bmatrix} \text{cov}(a_1, a_1) & \text{cov}(a_1, a_2) & \dots & \text{cov}(a_1, a_N) \\ \text{cov}(a_2, a_1) & \text{cov}(a_2, a_2) & \dots & \text{cov}(a_2, a_N) \\ \vdots & \vdots & \dots & \vdots \\ \text{cov}(a_N, a_1) & \text{cov}(a_N, a_2) & \dots & \text{cov}(a_N, a_N) \end{bmatrix}$$

c- Kovaryans matrisi için özdeğerleri ve özvektörleri hesapla:

Kovaryans matrisi hesaplandıktan sonra (A kovaryans matrisi $N \times N$ olmak üzere), bu matrise ait özvektörler (v) ve özdeğerler (λ) hesaplanır.

$$Av = \lambda v \Rightarrow Av - \lambda v = 0$$

$$(A - \lambda I) v = 0 \Rightarrow \det(A - \lambda I) = 0$$

d- Özdeğerleri ve karşılık gelen özvektörleri sırala:

Hesaplanan özdeğerler ve karşılık gelen özvektörler sıralanır. λ büyükten küçüğe doğru sıralı bir küme ve v bu kümenin elemanlarına karşılık gelen özvektörler olmak üzere,

$$(\lambda_{\max}, \lambda_{\min}) = (\lambda_1, \lambda_n) \Rightarrow \lambda = (\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n)$$

$$v = (v_1, v_2, v_3, \dots, v_n)$$

e- k tane özdeğer seç ve bir özvektör matrisi oluştur:

Büyükten küçüğe sıralanan özdeğerler doğrultusunda özdeğerlere karşılık gelen özvektörler (bileşenler) ile verideki açıklanabilen varyans yüzdesi hesaplanabilir. Örneğin ilk üç bileşenin açıklayabildiği varyans yüzdesi şu şekilde hesaplanabilir.

$$\frac{\lambda_1 + \lambda_2 + \lambda_3}{\lambda_1 + \lambda_2 + \lambda_3 + \dots + \lambda_n} \times 100$$

Bu doğrultuda hedeflenen açıklanabilen varyans yüzdesi elde edilecek şekilde en büyük özdeğere sahip k tane bileşen seçilir. Seçilen k tane bileşenin açıklayabildiği varyans yüzdesi yine aynı şekilde hesaplanır.

$$\frac{\lambda_1 + \lambda_2 + \lambda_3 + \dots + \lambda_k}{\lambda_1 + \lambda_2 + \lambda_3 + \dots + \lambda_n} \times 100$$

Açıklanabilen varyans yüzdesi doğrultusunda seçilen k tane özvektör birleştirilerek özvektör matrisi oluşturulur.

$$V = \begin{bmatrix} v_{11} & v_{21} & \dots & v_{k1} \\ v_{12} & v_{22} & \dots & v_{k2} \\ \vdots & \vdots & \dots & \vdots \\ v_{1N} & v_{2N} & \dots & v_{kN} \end{bmatrix}$$

f- Veri seti matrisini dönüştür:

Standardize edilmiş veri seti (D') özvektör matrisi (V) ile çarpıldığında birbiri ile ortogonal k tane temel bileşenden oluşan yeni bir veri seti (D'') elde edilir.

$$D'' = D' \times V$$

$$D'' = \begin{bmatrix} d'_{11} & d'_{12} & \dots & d'_{1N} \\ d'_{21} & d'_{22} & \dots & d'_{2N} \\ d'_{31} & d'_{32} & \dots & d'_{3N} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \times \begin{bmatrix} v_{11} & v_{21} & \dots & v_{k1} \\ v_{12} & v_{22} & \dots & v_{k2} \\ \vdots & \vdots & \dots & \vdots \\ v_{1N} & v_{2N} & \dots & v_{kN} \end{bmatrix} = \begin{bmatrix} d''_{11} & d'_{12} & \dots & d'_{1k} \\ d''_{21} & d'_{22} & \dots & d'_{2k} \\ d''_{31} & d'_{32} & \dots & d'_{3k} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

3.2. Çoklu Doğrusal Regresyon

Çoklu doğrusal regresyon modeli, eldeki veriyi doğrusal bir denkleme uydurarak, iki veya daha fazla bağımsız değişken ile bir tane bağımlı değişkenin ilişkisini açıklamaya çalışır [46]. Çoklu doğrusal regresyon modelinde X_i bağımsız değişkenlerinin her biri Y bağımlı değişkeni ile ilişkilidir. N tane bağımlı değişkenden oluşan çoklu doğrusal regresyon modeli şu şekilde ifade edilebilir.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \dots + \beta_n x_{in} + \varepsilon$$

Denklemden y_i bağımlı değişkeni, x_{ij} bağımsız değişkenleri, β_0 sabit terimi, β_j her bir bağımlı değişken için eğim katsayısını, ε ise modelin hata terimini ifade etmektedir.

Y bağımlı değişkenlerin vektörü, X bağımsız değişkenlerin matrisi ve β ise regresyon modeli katsayı vektörü olarak ele alındığında, regresyon modeli matris formunda şu şekilde ifade edilebilir.

$$Y = X \beta + \varepsilon$$

Regresyon modelinde amaç, bağımsız değişken değerlerinden bağımlı değişkenin değerinin tahmin edilmesini sağlayacak en iyi doğrusal eğriyi bulmaktır. Bu da ortalama karesel hatayı minimum yapacak β değerini bulmak anlamına gelmektedir. Ortalama karesel hata şu şekilde hesaplanmaktadır.

$$OKH = \frac{1}{n} \sum_{i=1}^{i=n} \varepsilon_i^2 = \frac{1}{n} (\varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_n^2)$$

Matris biçiminde gösterildiğinde [47];

$$= [\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_n] \times \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \varepsilon^T \times \varepsilon$$

ε yerine $Y - X \beta$ konulduğunda;

$$\begin{aligned} OKH &= \frac{1}{n} (Y - X \beta)^T (Y - X \beta) \\ &= \frac{1}{n} (Y^T - \beta^T X^T) (Y - X \beta) = \frac{1}{n} (Y^T Y - \beta^T X^T Y - Y^T X \beta + \beta^T X^T X \beta) \end{aligned}$$

$\beta^T X^T Y = (Y^T X \beta)^T$ olduğundan ve $Y^T X \beta$ skaler bir sayı olduğu için $\beta^T X^T Y = Y^T X \beta$ denilebilir. $Y^T X \beta$ yerine $\beta^T X^T Y$ konularak denklem düzeltildiğinde şu son şekli alır.

$$OKH = \frac{1}{n} (Y^T Y - 2 \beta^T X^T Y + \beta^T X^T X \beta)$$

Ortalama karesel hatayı minimum yapacak β değerini bulmak, bir optimizasyon problemidir. Dolayısıyla OKH fonksiyonunun β ya göre türevi alınır.

$$\nabla OKH = \frac{1}{n} (\nabla Y^T Y - 2 \nabla \beta^T X^T Y + \nabla \beta^T X^T X \beta) = \frac{1}{n} (0 - 2 X^T Y + 2 X^T X \beta)$$

Elde edilen sonuç bize Jakobi matrisini vermektedir.

$$J(\beta) = \frac{2}{n} (X^T X \beta - X^T Y)$$

Jakobi matrisi kullanılarak başlangıç olarak bir β katsayı vektörü alınır. β katsayı vektöründen Gradyan İniş Algoritması kullanılarak, k öğrenme hızıyla belirli bir iterasyon sonunda OKH'yi minimum yapacak optimum katsayı vektörü bulunur.

$$\beta_{\text{yeni}} = \beta_{\text{eski}} - J(\beta) \times k$$

OKH'ı minimum yapan optimum katsayı vektörü Y değerinin tahminlemede kullanılır.

$$\hat{Y} = X \beta_{\text{optimum}}$$

Çoklu doğrusal regresyon modelinin dayandığı varsayımlar şunlardır [48]:

- Bağılı değişken ve bağımsız değişkenler arasında doğrusal bir ilişki vardır.
- Bağımsız değişkenler birbiri ile çok fazla ilişkili değildir.
- Hatalar $\mu = 0$ ortalamalı ve σ varyanslı olarak normal dağılımdadır.
- Hata değerleri birbirinden bağımsızdır.
- Hata değerleri bağımsız değişkenlerin büyüklüğüne bağlı olmaksızın sabit varyans gösterir.

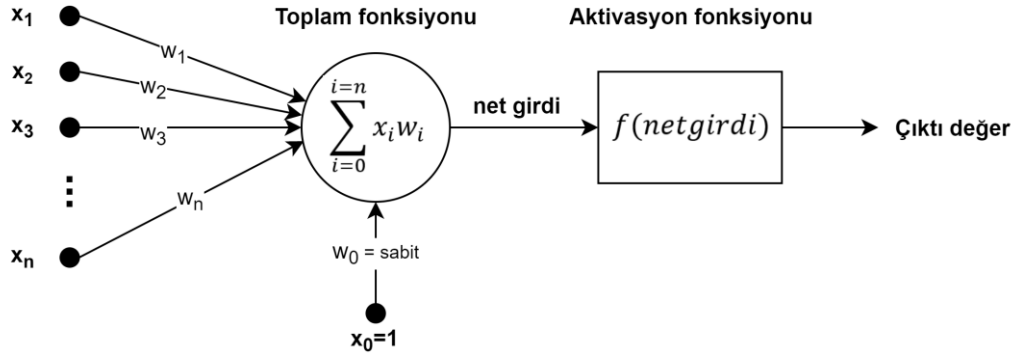
3.3. Yapay Sinir Ağları

Yapay sinir ağları, beynimizi oluşturan biyolojik sinir hücrelerinden esinlenen hesaplama sistemleridir. Bu tür sistemler belirli kurallar ile programlanmak yerine, işledikleri verilerden otomatik olarak belirleyici özellikleri çıkararak birtakım işlemler yapmayı öğrenirler [49].

Bir YSA, biyolojik beyindeki nöronları modelleyen ve yapay nöronlar adı verilen birbirine bağlı düğümlerden oluşur. Biyolojik beyin hücrelerindeki sinapslardaki gibi her bağlantı üzerinden diğer nöronlara bir sinyal iletilir. YSA uygulamasında bir bağlantıdaki "sinyal" sayıya denk gelmektedir ve herhangi bir nörondaki çıktı değeri, o nörondaki girdi değerlerinin toplamı üzerinden doğrusal olmayan bir fonksiyonla

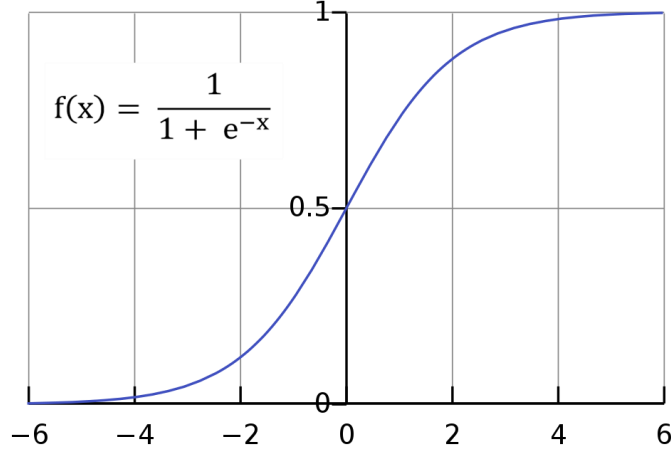
hesaplanır. Bu fonksiyona aktivasyon fonksiyonu adı verilir. Nöronları bağlayan bağlantılara kenar adı verilir. Nöronlar ve kenarlar öğrenme sırasında dinamik olarak ayarlanan ağırlığa sahiptirler. Bu ağırlıklar iletilmekte olan sinyalin gücünü attırır ya da azaltır. Tipik olarak nöronlar katmanlar halinde bulunurlar. Sinyaller genelde giriş katmanından çıkış katmanına doğru ilerler ve her bir katmanda girdi değerleri için farklı dönüşümler gerçekleştirilebilir.

Özetle tipik bir YSA nöronlar, bu nöronları bağlayan kenarlar, ağırlıklar ve nörondaki girdi değerini çıktı değerine dönüştüren aktivasyon fonksiyonundan oluşur. Bir YSA' daki nöronun yapısı aşağıdaki şekilde gösterilmiştir. [50]



Şekil 3 – YSA'daki bir nöronun yapısı

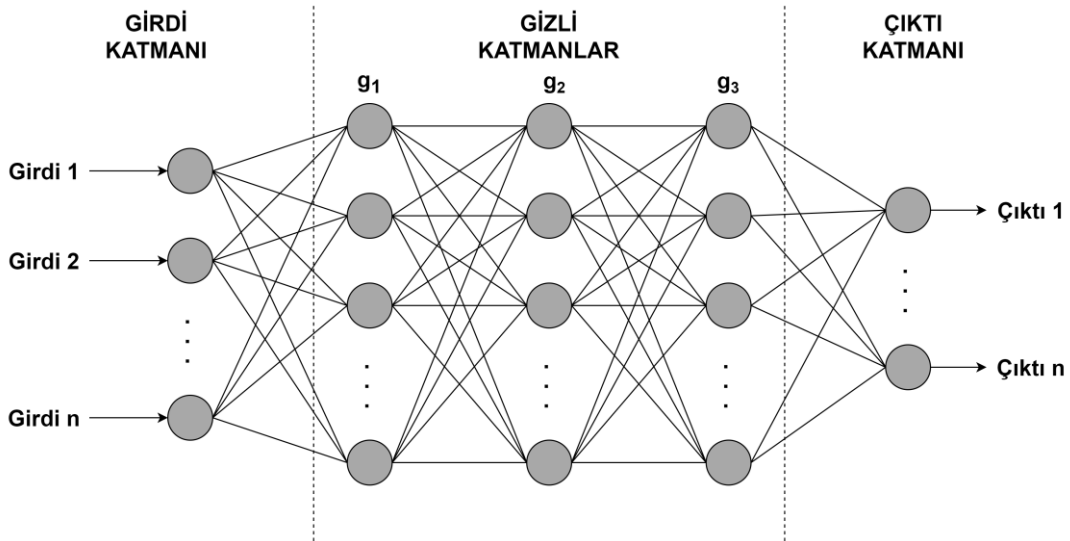
YSA modelinde çeşitli aktivasyon fonksiyonları kullanılabilir. Karmaşık sistemleri daha iyi modelleme yetenekleri olduğu için genellikle doğrusal olmayan aktivasyon fonksiyonları kullanılır. Bunlardan en yaygın kullanılanları sigmoid/lojistik, tanH, ReLU olarak belirtilebilir.



Şekil 4 – Sigmoid fonksiyon [51]

Sinir ağı modellerinde doğrusal olmayan karar yüzeylerini daha iyi ifade edebilmek için birden fazla katman kullanılabilir. Bu tarz modellerde ham girdiden daha yüksek ve karmaşık seviyede nitelikler öğrenilebilir. Böylece her katmanda girdi verilerini daha soyut ve birleşik bir şekilde temsil eden bir model elde edilebilir.

İleri Beslemeli Sinir Ağları, Radyal Tabanlı Sinir Ağları, Tekrarlayan Sinir Ağları, Evrimsel Sinir Ağları, Modüler Sinir Ağları gibi YSA çeşitleri bulunmaktadır. Bu çalışma kapsamında birçok problem için yaygın olarak uygulanan çok katmanlı ileri beslemeli yapay sinir ağları kullanılacaktır. Aşağıda çok katmanlı ileri beslemeli bir YSA yapısı gösterilmiştir [52].



Şekil 5 – Çok Katmanlı İleri Beslemeli YSA yapısı

İleri beslemeli YSA'ları eğitmede kullanılan en temel yöntem geri yayılım algoritmasıdır. Çalışma kapsamında da sinir ağlarını eğitmek için stokastik gradyan azaltma yöntemini kullanan geri yayılım algoritması kullanılacaktır [53].

Geri yayılım algoritması adımları şu şekilde özetlenebilir.

- n_i sayıda girdi, n_{gizli} sayıda gizli, $n_{çikti}$ sayıda çıktı nöronu olan ileri beslemeli sinir ağı yarat.
- Sinir ağındaki her bir ağırlığı w_i rastgele küçük bir sayı ile başlat.
- Sonlandırma koşulu yerine getirilene kadar aşağıdaki işlemleri tekrarla,
 - Her bir eğitim verisi için $\langle (x_1, x_2, \dots, x_n), t \rangle$

// Girdiyi ağ üzerinde ileriye doğru yay.

1. Girdi verisini sinir ağına sok ve her bir nöron için çıktıyı o_k hesapla

// Hatayı ağ üzerinde geriye doğru yay.

2. Çıktı katman seviyesindeki her bir nöron için hata terimini δ_k hesapla

$$\delta_k = o_k(1 - o_k)(t_k - o_k)$$

3. Gizli katman seviyesindeki her bir nöron için hata terimini δ_h hesapla

$$\delta_h = o_h(1 - o_h) \left(\sum_k w_{h,k} \delta_k \right)$$

4. Her bir ağırlığın ağırlığını belirlenmiş öğrenme hızına η göre yeniden hesapla

$$w_{i,j} = w_{i,j} + \Delta w_{i,j} \quad \Delta w_{i,j} = \eta \delta_j x_{i,j}$$

3.4. Rastgele Orman

Rastgele Ormanlar, sınıflandırma, regresyon veya diğer problemlerin çözümünde kullanılan kolektif öğrenme yöntemlerinden biridir. Bu yöntemde öğrenme sırasında çok sayıda karar ağacı oluşturulur ve tahminde ise bu karar ağaçlarının her birinin modu (sınıflandırma problemi için) veya ortalaması (regresyon problemi için) alınır [54].

Kolektif öğrenme yöntemleri temel olarak boosting ve bagging (Bootstrap aggregating) olmak üzere ikiye ayrılmaktadır. Her iki yöntemde de orijinal veri setinden, yerine koyma ile rastgele örneklemeler alınır. Bundan dolayı yaratılan veri setlerinde bazı veriler tekrar edebilir. Bu şekilde orijinal veri seti ile aynı boyutta yeni veri setleri yaratılmış olur.

Bagging [55] yönteminde yaratılan veri setlerinde her bir verinin bulunma olasılığı eşittir. Bu yöntemde yaratılan N tane veri seti paralel olarak eğitilerek birbirinden farklı N tane model elde edilmiş olur. Tahminde ise bu N tane modelin tahmin sonuçlarının modu veya ortalaması alınır. Tahminlemede birden fazla modelin kullanılmasıyla varyansın azaltılması ve tahminin eğitim verisine aşırı uyum göstermesi durumu azaltılmaya çalışılır.

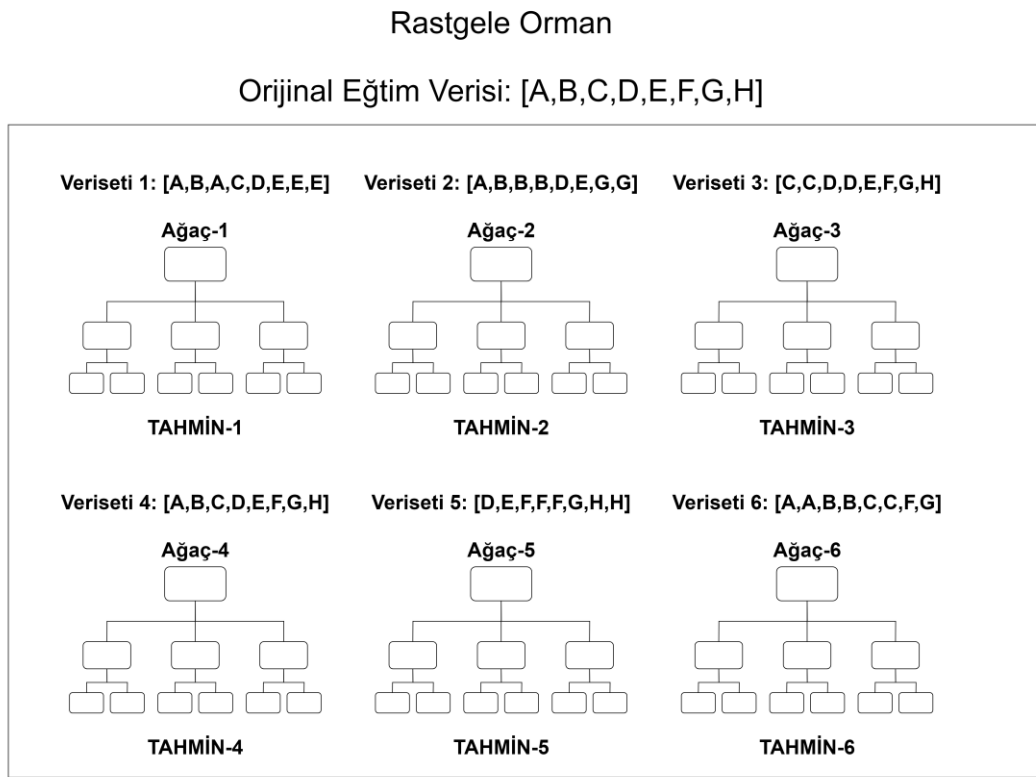
Boosting [56] yönteminde ise eğitim paralel olarak değil, sıralı olarak gerçekleştirilir. Her bir model bir önceki modelin başarısız tahminlerini azaltmayı amaçlar. Bunu başarısız tahmin verileri üzerindeki ağırlığı artırarak yapar. Sonuçta bir önceki modelin başarısız tahminlerini azaltan N tane model eğitilmiş olur. Tahminde ise bu N tane modelin aritmetik ortalaması alınmaz. Eğitim verisi üzerinde başarılı olmuş modellerin ağırlığı daha fazla olacak şekilde ağırlıklı ortalaması alınır. Boosting yöntemi yanlılığı azaltırken, eğitim verisine aşırı uyum göstermesi durumunu artırabilir.

Rastgele Orman bir bagging yöntemidir. Orijinal eğitim veri setinden yerine koyma ile rastgele alınan örneklemelerden çok sayıda eğitim veri seti oluşturulur. Her bir veri setinden bir karar ağacı oluşturulur. Karar ağaçlarının birlikte oluşturduğu yapı Rastgele Orman olarak adlandırılır. Tahminlemede ise bu karar ağaçlarının her birinin verdiği sonucun modu veya ortalaması alınır.

Rastgele Orman bireysel eğitilmiş karar ağaçlarının tahmin başarısından daha iyi performans göstermektedir. Karar Ağaçları topluluk oluşturarak, birbirlerini, ağaç bazında oluşan bireysel tahmin hatalarından korurlar. Rastgele Orman modelini oluşturan her bir karar ağacı arasında düşük korelasyon bulunduğundan, karar ağaçlarının bireysel tahminlerinden daha doğru olan topluluk tahmini (ortalama tahmin/çoğunluk tahmini) rastgele orman yöntemiyle elde edilebilir. Rastgele Orman Yönteminin tahminlemede iyi bir performans gösterebilmesi için şu iki ön koşul gereklidir.

- 1- Niteliklerin (veri setindeki deęişkenlerin) hedef deęer ile aralarında anlamlı bir ilişki olmalıdır, böylece bu nitelikleri kullanarak oluşturulan modeller rastgele tahmin etmekten daha iyi sonuç verir.
- 2- Her bir karar ağacı tarafından yapılan tahminlerin (ve dolayısıyla hataların) arasında düşük korelasyon olmalıdır.

Aşağıda altı ağaçtan oluşan örnek bir rastgele orman modeli gösterilmiştir. Modele göre her bir ağaç orijinal eğitim veri setinden rastgele alınmış örneklerle eğitilmiştir. Tahminlemede ise bu altı ağacın verdiği sonuçların ortalaması (regresyon problemi için) veya modu (sınıflandırma problemi için) alınacaktır.



Şekil 6 – Rastgele Orman Modeli

Rastgele Orman Modelinin sağladığı iki temel avantaj bulunmaktadır.

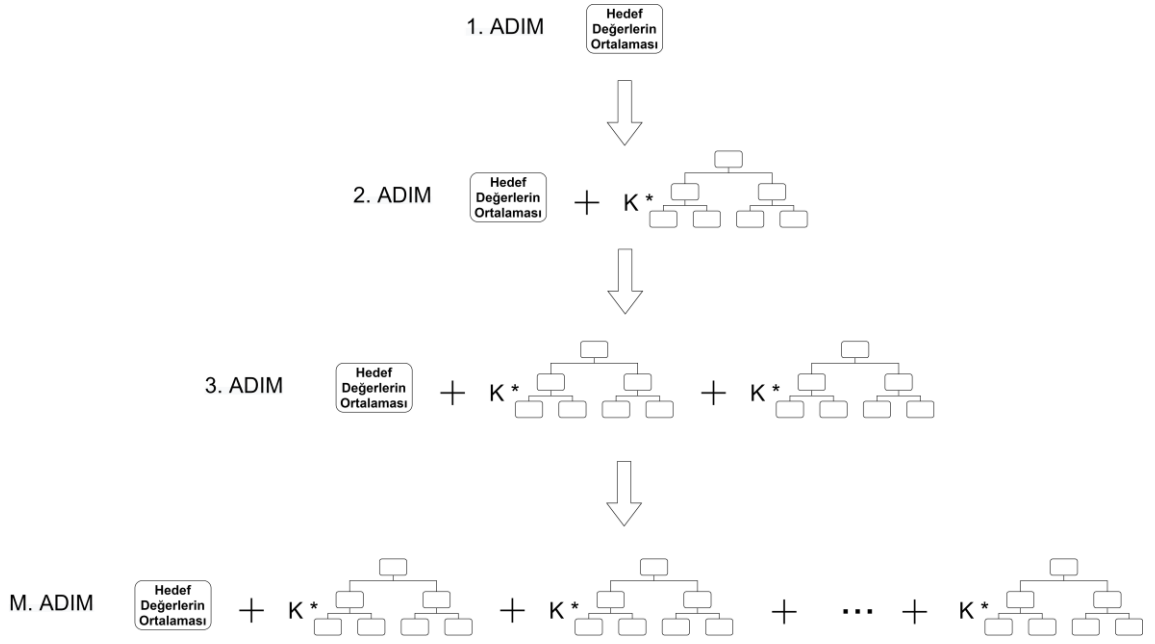
- 1- Karar Ağaçlarının tipik sorunu olan modelin eğitim verisine aşırı uyma problemini azaltır.
- 2- Karar ağaçları arasında ağaç düğümlerini oluşturan nitelik rastgeleliğini sağlayarak, ağaçlar arasındaki korelasyonu azaltır, böylece daha başarılı tahminler elde edilir.

3.5. Gradyan Artırma

Gradyan Artırma algoritması da Rastgele Ormanlar gibi sınıflandırma, regresyon veya diğer problemlerin çözümünde kullanılan kolektif öğrenme yöntemlerinden biridir. Rastgele Orman algoritmasında olduğu gibi birden fazla karar ağacının eğitilmesine dayanır. Gradyan Artırma algoritması Rastgele Orman algoritmasından farklı olarak kolektif öğrenmedeki boosting yöntemini kullanır. Yani karar ağaçları paralel değil, sıralı olarak eğitilir.

Regresyon problemi için kullanılan Gradyan Artırma Algoritması en basit haliyle kabaca şu şekildedir.

- 1- Veri setindeki hedef değerlerinin ortalaması alınır. Böylece tek bir karar ağacının tek bir yaprağından oluşan ve bütün durumlar için ortalama değeri tahminleyen başlangıç modeli oluşturulmuş olur.
- 2- Önceden belirlenmiş regresyon ağacı sayısına (M) ulaşana kadar şu adımlar tekrar edilir.
 - a. Modelin kaba hataları (gözlemlenen hedef değeri – tahmin edilen değeri) bütün veriler için hesaplanır.
 - b. Sonrasında hataları tahmin edecek yeni bir regresyon ağacı eğitilir. Regresyon ağacı eğitilirken ağacın büyüklüğü, yani ağaçtaki yaprak sayısı parametre olarak genellikle 8-32 arasında belirlenir. Regresyon ağacında yaprak sayısı sınırlı olduğundan aynı yaprağa düşen farklı hedef değerler varsa bunların ortalaması alınarak yaprağa tek bir değer yazılır.
 - c. Eğitilen yeni regresyon ağacı bir önceki modele eklenir. Her bir ağaç modele bir öğrenme oranı ile çarpılarak eklenir. Böylece aşırı uyma sorunu yaşanmamış olur.(Bu öğrenme oranı parametre olarak kullanıcı tarafından önceden belirlenir.)
- 3- M tane regresyon ağacından kolektif model eğitilmiş olur.



Şekil 7 – Gradyan Artırma Algoritması

Regresyon problemlerinde kullanılan Gradyan Artırma Algoritması matematiksel olarak şu şekilde özetlenebilir.

$\{(x_i, y_i)\}_{i=1}^n$ eğitim veri seti, $L(y_i, F(x))$ türevlenebilir kayıp fonksiyonu olmak üzere

- 1- Sabit bir değer ile başlangıç modeli oluştur: $F_0(x) = \operatorname{argmin}_{\gamma} \sum_{i=1}^n L(y_i, \gamma)$
- 2- $m=1$ 'den M ' e kadar tekrarla (M önceden belirlenmiş karar ağacı sayısı olmak üzere)

- a. Her bir $i = 1, 2, \dots, n$ için hesapla: $r_{im} = - \left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)}$
- b. Bütün r_{im} değerlerine göre regresyon ağacı oluştur, her bir $j=1 \dots J_m$ veri çıktısı için R_{jm} çıktılarını (ağaç çıktı değerleri) içeren regresyon ağacı oluştur.
- c. Regresyon ağacındaki her bir $j=1 \dots J_m$ çıktıdan bütün yaprakların net değerini $\gamma_{jm} = \operatorname{argmin}_{\gamma} \sum_{x_i \in R_{ij}} L(y_i, F_{m-1}(x_i) + \gamma)$ hesapla.
- d. Bir önceki modele göre yeni regresyon ağacını ekleyerek modeli güncelle $F_m(x) = F_{m-1}(x) + \nu \sum_{j=1}^{J_m} \gamma_{jm} I(x \in R_{jm})$

- 3- Nihai Model: $F_M(x)$

4. ANALİZ VE UYGULAMA

4.1. İşletme ve Tez Çalışmasında Kullanılan Veri

Bu tez çalışması kapsamında PVC profili üreten bir işletmenin 01.12.2010 - 22.11.2013 tarihleri arasındaki üretim kayıp verileri kullanılmıştır [57]. İşletme 24 saat kesintisiz üretim yapan bir işletme olup, her bir üretim vardiyası 12 saatten oluşmaktadır. İşletmede üretim 27 tane üretim hattı üzerinden sağlanmaktadır. Ancak her bir vardiyada 27 üretim hattı birden tam kapasite çalışmamaktadır. Kullanılan veride, tarih ve vardiya bazında hangi üretim hatlarının çalıştığı, ilgili vardiyada ne kadar kayıp olduğu ve kaybın oluşma nedeni bilgisi bulunmaktadır. Kaybın oluşma türüne göre (planlanan ve planlanmayan kayıp türü) ilgili üretim hattında ve vardiyada bakım yapılıp yapılmadığı bilgisi de çıkarılabilmektedir. Ayrıca ilgili veride her bir çalışan üretim hattında PVC ekstrüzyon işlemini yapan ekipmanların planlanan çalışma hızı ve üretim hattındaki gerçek çalışma hızı bilgisi de bulunmaktadır.

Eldeki veriden yola çıkarak işletmede ilgili üretim hattı için muhtemel kaybı tahmin edebilecek modellerin eğitilmesi hedeflenmiş ve bu doğrultuda bakım kararı için model önerilmesi hedeflenmiştir.

4.2. Veri Ön İşleme Adımları

İşletmedeki kayıp miktarlarını tahmin edebilecek modellerinin eğitilebilmesi için öncelikle veri üzerinde çeşitli düzenlemeler yapılmıştır. Veri ön işleme aşamasında 01.12.2010 - 22.11.2013 tarihleri arasında işletmede üretim yapan (günde iki vardiya olmak üzere) 27 hatta ait kayıp verileri baz alınmıştır. Hatlarda oluşan kayıplar planlanan ve planlanmayan kayıplar olarak aşağıdaki şekilde ayrılmıştır.

Çizelge 1 – Planlanan / Planlanmayan Kayıplar

PLANLANMAYAN KAYIP TÜRLERİ	PLANLANAN KAYIP TÜRLERİ
Kalibratör Kirliliği (dk.)	Başlatma ve Kapatma Kayıpları (kg)
Operasyonel İşlem Problemleri (kg)	Tırtıl Bakımı (dk.)
Tırtıl Hatası (dk.)	Değişim (dk.)
Spesifikasyon Dışı Geometri (kg)	Kompresör Bakımı (dk.)
Yüzey Hataları (kg)	Konveyör Bakımı (dk.)
Kompresör Arızası (dk.)	Kalıp & Kalibratör Temizleme (dk.)
Konveyör Arızası (dk.)	Kalıp & Kalibratör Revizyonları (dk.)
Testere Arızası (dk.)	Testle İlgili Kayıplar (kg)
Soğutma Sistemi Arızası (dk.)	Akım Yönü Bakımı (dk.)
Boşaltma Arızası (dk.)	Ekstrüder Bakımı (dk.)
Kalıp Kirliliği (dk.)	Başka bir tarifeye gitme (kg)

Çizelge 1 – Planlanan / Planlanmayan Kayıplar (devamı)

PLANLANMAYAN KAYIP TÜRLERİ	PLANLANAN KAYIP TÜRLERİ
Akım Yönü Arızası (dk.)	Testere Bakımı (dk.)
Ekstrüder Arızası (dk.)	Talep Eksikliği (dk.)
Filtrasyon Sistemi Arızası ARIZASI (dk.)	Tesis Bakımı (dk.)
Yanlış Formülasyon (dk.)	Vakum / Su pompası Bakımı (dk.)
Spesifikasyon Dışı Renk/Parlaklık (kg)	Su tankının yıkanması (dk.)
Personel Eksikliği(dk.)	Boşaltma Pompası Bakımı (dk.)
Yedek Parça Eksikliği (dk.)	Soğutma Sistemi Bakımı (dk.)
Düşük Dereceli Kuru Karışım (dk.)	
Operatör Hatası (dk.)	
Diğer Kayıplar (dk.)	
Spesifikasyon Dışı Fiziksel Test Sonuçları (kg)	
Güç kaynağı (dk.)	
Ham madde kıtlığı (dk.)	
Spirofleks Arızası (dk)	
Vakum / Su pompası Arızası (dk.)	
Kalıp & Kalibratör Kirliliği (dk.)	

Planlanan kayıpların hepsi bakım olarak kabul edilmiştir. Yani ilgili hatta belirtilen gün ve vardiyada planlanan bir kayıp varsa bakım yapıldığı kabul edilmiştir.

Planlanmayan kayıplar ise kendi içerisinde zaman bakımından kayıplar (dk.) (toplamda 1.561 kayıp verisi) ve miktar bakımından kayıplar (kg) (toplamda 38.797 kayıp verisi) olmak üzere ayrılmıştır. Sonrasında planlanmayan kayıp verilerinden, her bir hat için gün ve vardiya bazında kayıp miktarları (miktar ve zaman bakımından) ayrı ayrı hesaplanmış, planlanan kayıp verilerinden ise bakım yapıp yapılmadığı bilgisi çıkarılmıştır. (İlgili hat belirtilen gün ve vardiyada çalışmıyorsa kayıp miktarı ve bakım bilgisi hesaplanmamıştır.)

Tahmin modelleri oluşturulurken zaman bakımından kayıplar (dk.) göz ardı edilerek yalnızca miktar (kg) bazındaki kayıplar dikkate alınmış, eğitilen modellerin yalnızca kg bazındaki planlanmayan kayıp miktarını tahmin edebilmesi hedeflenmiştir.

Kg bazındaki kayıp miktarlarını tahmin edebilmek için modellerin eğitiminde kullanılacak tahmin ediciler ve hedef değer şu şekilde belirlenmiştir.

Çizelge 2 – Hedef Değer ve Tahmin Ediciler

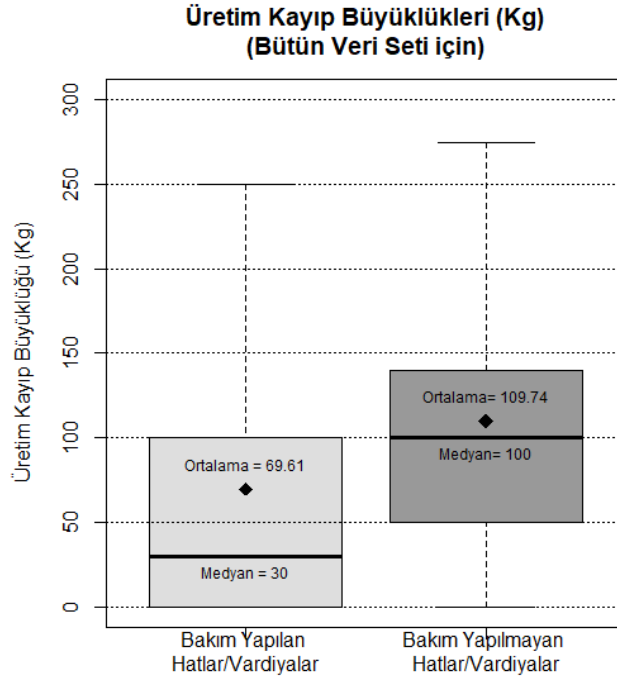
Değişken Türü	Değişken İsmi	Değişken açıklama
Hedef Değer	KAYIP_KG	Kayıp (Kg)
1. Tahmin Edici	SON_HAFTA TOPLAM_KAYIP_KG	İlgili hat için son haftadaki toplam kayıp büyüklüğü (Kg)
2. Tahmin Edici	SON_HAFTA KAYIP_SAYISI	İlgili hat için son haftadaki kayıp sayısı (Adet)
3. Tahmin Edici	SON_AY TOPLAM_KAYIP_KG	İlgili hat için son aydaki toplam kayıp büyüklüğü (Kg)
4. Tahmin Edici	SON_AY_KAYIP SAYISI	İlgili hat için son aydaki kayıp sayısı (Adet)
5. Tahmin Edici	SON_BKMDAN_GECEN VARDIYA_SAYISI	İlgili hat için son yapılan bakım üzerinden geçen vardiya sayısı (Adet)
6. Tahmin Edici	SON_BKMDAN_SONRA TOPLAM_KAYIP_KG	İlgili hat için son bakımdan itibaren gerçekleşen toplam kayıp büyüklüğü (Kg)
7. Tahmin Edici	MAKINE_CALISMA HIZ_FARKI	Makinenin [Planlanan Çalışma hızı - Gerçek Çalışma hızı]

Veri ön işleme sonrasında bütün hatlar (27 hat) ve bütün yıllar/vardiyalar için elde edilen veri seti yanında, spesifik bir hat bazında tahmin performanslarının değişip değişmeyeceğini görmek amacıyla yalnızca 4 numaralı üretim hattının 2012 yılına ait ikinci bir veri seti de oluşturulmuştur.

4.3. Genel Veri Analizi

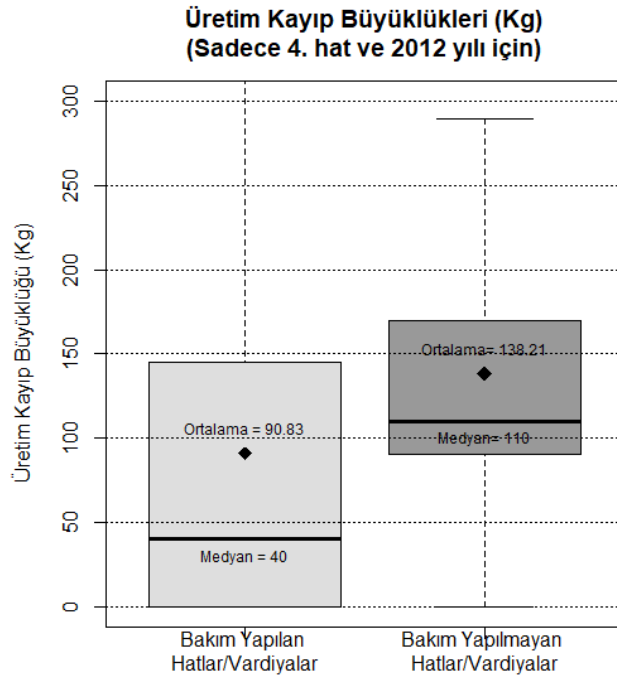
Ön işleme sonrası elde edilen, gün, vardiya ve hat bazında üretim kayıp verilerini içeren 22179 satırlık veri seti incelendiğinde, 1759 hat/vardiyada bakım yapıldığı, 20420 hat/vardiyada ise bakım yapılmadığı görülmüştür. Önceki bölümde elde edilen her iki veri seti genel olarak analiz edildiğinde, bakım yapılan hat/vardiya kayıtlarındaki üretim kayıp büyüklüğünün (kg), bakım yapılmayan hat/vardiya kayıtlarındaki üretim kayıp büyüklüğüne (kg) göre daha düşük olduğu görülmüştür.

İlk veri setindeki (bütün veri seti) kayıp büyüklüklerine bakıldığında bakım yapılan hat/vardiyalarda ortalama üretim kayıp büyüklüğünün 69,61 olduğu, bakım yapılmayan hat/vardiyalarda ise 109,74 olduğu gözlenmiştir. Yine ilk veri setinde bakım yapılan hat/vardiyalarda ortanca (medyan) değer 30 olduğu, bakım yapılmayan hat/vardiyalarda ise 100 olduğu gözlenmiştir.



Şekil 8 – Bakım Yapılma Durumuna Göre Üretim Kayıp Büyüklükleri (1. Veri Seti)

İkinci veri setindeki (Sadece 4. Hat ve 2012 yılına ait veri seti) kayıp büyüklüklerine bakıldığında ise bakım yapılan hat/vardiyalarda ortalama üretim kayıp büyüklüğünün 90,83 olduğu, bakım yapılmayan hat/vardiyalarda ise 138,21 olduğu gözlenmiştir. Yine ikinci veri setinde bakım yapılan hat/vardiyalarda ortanca (medyan) değer 40 olduğu, bakım yapılmayan hat/vardiyalarda ise 110 olduğu gözlenmiştir.



Şekil 9 – Bakım Yapılma Durumuna Göre Üretim Kayıp Büyüklükleri (2. Veri Seti)

İki veri setine ait grafiğe bakıldığında bakım yapmanın, üretim sırasında oluşacak planlanmayan kayıp miktarını (kg) düşürdüğü açıkça görülmektedir. Bu durumda her vardiyada üretime başlamadan önce bakım yapılmasının uygun olacağı düşünülebilir. Ancak böyle bir yöntemin planlanmayan kayıp miktarını (kg) azaltırken, bakım kaynaklı planlanan kayıp miktarını (dk) artıracığı unutulmamalıdır. Dolayısıyla her vardiya başında bakım yapmak yerine belirli durumlar oluştuğunda bakım yapmak daha doğru gözükmemektedir.

4.4. Bakım Kararı İçin Eşik Değerin Belirlenmesi

Tez çalışmasında önerilen bakım modeline göre, ilgili vardiyada bakım yapılıp yapılmayacağına kararının verilebilmesi için, muhtemel üretim kaybı için bir eşik değerin belirlenmesi önerilmektedir. Bu doğrultuda, bakım modeline göre çalıştırılması planlanan hat ve vardiya için öncelikle muhtemel üretim kaybı (kg) tahmin edilecek daha sonra tahmin edilen üretim kaybı, belirlenen eşik değerden büyükse bakım yapılacak, değilse bakım yapılmayacaktır.

Eşik değeri belirlenirken başlatma ve kapatma kayıpları, kalıp & kalibratör, talep eksikliği dışındaki planlanan kayıplar (bakım kaynaklı kayıpların) dikkate alınmıştır. Belirtilen bakım kaynaklı kayıpların ortalaması miktar cinsinden (kg) 214,95 kg olarak hesaplanmıştır. Buradan hareketle bakım kararı için eşik değeri 215 kg olarak belirlenmiştir.

Sonuç olarak, tahmin edilen muhtemel üretim kaybı, ortalama bakım maliyeti olarak belirlenen eşik değerden (215 kg) büyükse bakım yapılacak, küçükse bakım yapılmayacaktır.

4.5. TBA (Temel Bileşen Analizi) Uygulanması

4.5.1. Bütün Veri (1. Veri Seti) için Temel Bileşen Analizi

Tahmin modellerinin veriden öğrenme algoritmalarıyla eğitilmesi öncesinde tahmin değişkenleri üzerinde Temel Bileşen Analizi (TBA) uygulanmıştır.

TBA için 'h2o' [58] paketi kullanılmıştır. Kg cinsinden kayıpları tahmin etmede kullanılan yedi tahmin değişkeninden TBA yöntemiyle birbirinden bağımsız yine yedi tane yeni bileşen yaratılmıştır. TBA yönteminde bileşenlerin doğru bir şekilde oluşabilmesi için veri [0,1] aralığında normalize edilmiştir. Yaratılan bileşenler önem sırasına göre şu şekildedir.

Çizelge 3 – TBA Sonrası Bileşenler (Bütün Veri)

Bileşen	Standart sapma	Açıklanabilen Varyans	Açıklanabilen Var. (Kümülatif)
pc1	0,347	0,611	0,611
pc2	0,175	0,155	0,766
pc3	0,162	0,133	0,899
pc4	0,127	0,082	0,981
pc5	0,043	0,009	0,99
pc6	0,038	0,007	0,997
pc7	0,009	0,003	1

TBA'nın sonucunda oluşan bileşenler incelendiğinde en önemli beş bileşenin varyansın 99%' unu açıklayabildiği görülmüştür. Özvektör matrisi ise aşağıdaki gibi oluşmuştur.

Çizelge 4 – TBA Sonrası Özvektör Matrisi (Bütün Veri)

	pc1	pc2	pc3	pc4	pc5
SON_HAFTA TOPLAM_KAYIP_KG	0,264	-0,351	-0,451	0,423	0,637
SON_HAFTA KAYIP_SAYISI	0,578	0,286	0,151	0,637	-0,39
SON_AY TOPLAM_KAYIP_KG	0,362	-0,391	-0,536	-0,348	-0,54
SON_AY_KAYIP SAYISI	0,651	0,329	0,151	-0,54	0,386
SON_BKMDAN_GECEN VARDIYA_SAYISI	0,128	-0,458	0,591	-0,018	-0,008
SON_BKMDAN_SONRA TOPLAM_KAYIP_KG	0,153	-0,567	0,335	0,0008	0,012
MAKINE_CALISMA HIZ_FARKI	0,0006	-0,0003	-0,0002	-0,0003	0,001

Özvektör matrisi, TBA sonrası oluşan temel bileşenlerin yorumlanabilmesine olanak sağlamaktadır. Temel bileşenlerin hangi değişkenlerin hangi katsayı kombinasyonları ile oluştuğu özvektör matrisinden çıkarılabilmektedir. Örneğin 1.temel bileşen, değişkenlerin şu şekilde ortogonal kombinasyonundan oluşmaktadır.

$$\begin{aligned}
 pc1 = & (0,264 \times SON_HAFTA_TOPAM_KAYIP_KG) + (0,578 \\
 & \times SON_HAFTA_KAYIP_SAYISI) + (0,362 \\
 & \times SON_AY_TOPLAM_KAYIP_KG) + (0,651 \\
 & \times SON_AY_KAYIP_SAYISI) + (0,128 \\
 & \times SON_BKMDAN_GECEN_VARDIYA_SAYISI) + (0,153 \\
 & \times SON_BAKIMDAN_SONRA_TOPLAM_KAYIP_KG) + (0,0006 \\
 & \times MAKINE_CALISMA_HIZ_FARKI)
 \end{aligned}$$

Bu bağlamda özvektör matrisi incelendiğinde farklı değişkenlerin farklı temel bileşenler üzerinde farklı etkilerinin olduğu görülmektedir. Örneğin son bakımdan sonra oluşan toplam kayıp miktarının 2. Temel bileşen üzerindeki etkisi fazla iken 4. temel bileşen üzerindeki etkisi fazla değildir. Ayrıca makine çalışma hız farkının da temel değişkenler üzerindeki etkisinin görece daha az olduğu söylenebilir.

TBA'dan hareketle en önemli beş bileşen veriden öğrenme algoritmalarında kullanılmak üzere seçilmiştir. Seçilen beş bileşen için ham veriden yeni veri seti oluşturulmuştur.

Çizelge 5 – Seçilen Bileşenlerden Oluşturulan Veri Seti Yapısı (Bütün Veri)

Değişken türü	Değişken ismi	Değişken açıklama
Hedef Değer	KAYIP_KG	Kayıp (Kg)
1. Tahmin Edici	pc1	1. temel bileşen
2. Tahmin Edici	pc2	2. temel bileşen
3. Tahmin Edici	pc3	3. temel bileşen
4. Tahmin Edici	pc4	4. temel bileşen
5. Tahmin Edici	pc5	5. temel bileşen

4.5.2. Sadece Belirlenen Hat (4.hat) ve Yıl (2012) için Temel Bileşen Analizi

Hat bazında tahmin modellerinin performansının da incelenebilmesi için veri ön işleme kısmında oluşturulan diğer veri seti (Yalnızca 4 numaralı hattın 2012 yılına ait oluşturulan veri seti) için de TBA uygulanmıştır. İlk veri setinde uygulandığı gibi, kg cinsinden kayıpları tahmin etmede kullanılan yedi tahmin değişkeninden birbirinden bağımsız yine yedi tane yeni bileşen yaratılmıştır. TBA yönteminde bileşenlerin doğru bir şekilde oluşabilmesi için veri [0,1] aralığında normalize edilmiştir. Yaratılan bileşenler önem sırasına göre şu şekildedir.

Çizelge 6 – TBA Sonrası Bileşenler (4. Hat ve 2012 yılı için)

Bileşen	Standart sapma	Açıklanabilen Varyans	Açıklanabilen Var. (Kümülatif)
pc1	0,376	0,577	0,577
pc2	0,252	0,259	0,836
pc3	0,163	0,109	0,945
pc4	0,091	0,034	0,979
pc5	0,048	0,010	0,989
pc6	0,045	0,008	0,997
pc7	0,018	0,003	1

TBA'nın sonucunda oluşan bileşenler incelendiğinde en önemli beş bileşenin varyansın 98,9%' unu açıklayabildiği görülmüştür. Özvektör matrisi ise aşağıdaki gibi oluşmuştur.

Çizelge 7 – TBA Sonrası Özvektör Matrisi (4. Hat ve 2012 yılı için)

	pc1	pc2	pc3	pc4	pc5
SON_HAFTA TOPLAM_KAYIP_KG	-0,284	-0,129	-0,675	0,593	-0,108
SON_HAFTA KAYIP_SAYISI	-0,349	-0,252	-0,526	-0,643	0,139
SON_AY TOPLAM_KAYIP_KG	-0,385	-0,422	0,346	0,424	0,205
SON_AY_KAYIP SAYISI	-0,432	-0,468	0,358	-0,225	-0,224
SON_BKMDAN_GECEN VARDIYA_SAYISI	-0,473	0,520	0,117	-0,048	0,007
SON_BKMDAN_SONRA TOPLAM_KAYIP_KG	-0,487	0,500	0,065	0,026	-0,013
MAKINE_CALISMA HIZ_FARKI	-0,003	0,005	0,009	0,018	0,935

Özvektör matrisi TBA sonrası oluşan temel bileşenlerin yorumlanabilmesine olanak sağlamaktadır. Temel bileşenlerin hangi değişkenlerin hangi katsayı kombinasyonları ile oluştuğu özvektör matrisinden çıkarılabilmektedir. Örneğin 2. temel bileşen değişkenlerin şu şekilde ortogonal kombinasyonundan oluşmaktadır.

$$pc2 = (-0,129 \times SON_HAFTA_TOPAM_KAYIP_KG) + (-0,252 \times SON_HAFTA_KAYIP_SAYISI) + (-0,422 \times SON_AY_TOPLAM_KAYIP_KG) + (-0,468 \times SON_AY_KAYIP_SAYISI) + (0,520 \times SON_BKMDAN_GECEN_VARDIYA_SAYISI) + (0,500 \times SON_BAKIMDAN_SONRA_TOPLAM_KAYIP_KG) + (0,005 \times MAKINE_CALISMA_HIZ_FARKI)$$

Bu bağlamda özvektör matrisi incelendiğinde farklı değişkenlerin farklı temel bileşenler üzerinde farklı etkilerinin olduğu görülmektedir.

TBA'dan hareketle en önemli beş bileşen veriden öğrenme algoritmalarında kullanılmak üzere seçilmiştir. Seçilen beş bileşen için ham veriden (Yalnızca 4 numaralı hattın 2012 yılına ait oluşturulan veri seti) yeni veri seti oluşturulmuştur. Böylece veri seti algoritmalar için hazır hale getirilmiştir.

Çizelge 8 – Seçilen Bileşenlerden Oluşturulan Veri Seti Yapısı (4. Hat ve 2012 yılı için)

Değişken türü	Değişken ismi	Değişken açıklama
Hedef Değer	KAYIP_KG	Kayıp (Kg)
1. Tahmin Edici	pc1	1. temel bileşen
2. Tahmin Edici	pc2	2. temel bileşen
3. Tahmin Edici	pc3	3. temel bileşen
4. Tahmin Edici	pc4	4. temel bileşen
5. Tahmin Edici	pc5	5. temel bileşen

TBA sonrası birbirine ortogonal beş temel bileşenden oluşturulan iki adet dönüştürülmüş veri seti (bütün veri ve 4. Hat / 2012 yılına ait veri seti), veriden öğrenme algoritmaları uygulanmak üzere hazır hale getirilmiştir.

4.6. Veriden Öğrenme Yöntemlerinin Uygulanması

TBA Uygulandıktan sonra elde edilen veri setlerinde veriden öğrenme algoritmalarıyla kg cinsinden kayıp miktarlarının tahmin edebilecek modeller eğitime çalışılmıştır. Bunun için 1. ve 2. veri setleri için ayrı ayrı veriden öğrenme algoritmaları uygulanmıştır. Belirtilen problem regresyon problemi olduğundan tahmin modellerinin eğitiminde denetimli regresyon algoritmaları (doğrusal regresyon, sinir ağları, rastgele orman, gradyan artırma) kullanılmıştır.

4.6.1. Bütün Veri (1. Veri Seti) Kullanılarak Modellerin Eğitilmesi

4.6.1.1 Doğrusal Regresyon

Doğrusal regresyon modelinin eğitilmesi için 'h2o' paketindeki geliştirilmiş doğrusal modeller fonksiyonu [59] kullanılmıştır. Tahmin edici değerler arasında yanlılık oluşmaması açısından değişkenler 0 ortalama ve 1 varyans ile standardize edilerek eğitilmiştir. Modelin performans göstergeleri k-kat çapraz doğrulama yöntemi (k=10) ile elde edilmiştir. Verinin uyum sağladığı doğrusal regresyon modeli hedef değerlerin değişkenliğini seçilen tahmin edicilerle açıklamada yetersiz gözükmiştir. R kare değeri 0,133 ortalama mutlak hata (OMH) değeri ise 54,601 çıkmıştır. Eğitim için kullanılan temel bileşenlerden (tahmin edicilerden) tümünün tahmin değişkenine etkisinin olduğu gözükmiştir.

Çizelge 9 – Doğrusal Regresyon Sonuçları (Bütün Veri)

Değişken	Katsayılar	Standart Hata	Z Değeri	P değeri
Sabit değişken	106,755	0,567	188,109	0
pc1	40,479	1,643	24,626	0
pc2	-105,760	3,247	-32,571	0
pc3	-145,278	3,514	-41,334	0
pc4	20,444	4,469	4,573	0
pc5	88,773	13,229	6,710	0

Hata Metriği	Hata Değeri
Ortalama Karesel Hata	7146,718
Kök Ortalama Karesel Hata	84,538
Ortalama Mutlak Hata	54,601
Kök Ortalama Karesel Logaritmik Hata	1,358
Ortalama Artık Sapma	7146,718
R kare	0,133

4.6.1.2 Sinir Ağları

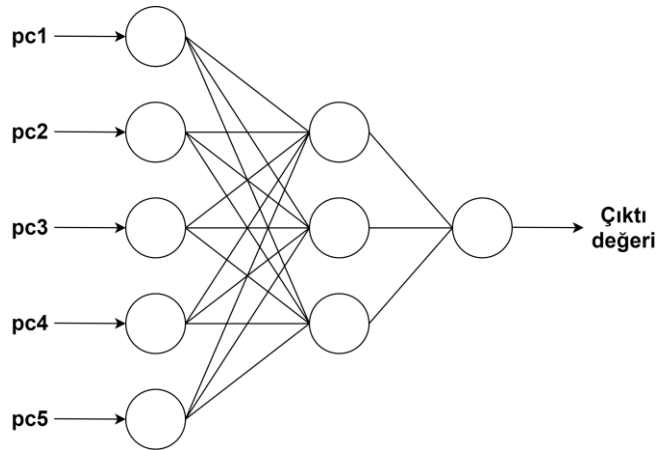
En iyi tahmin performansını sağlayacak sinir ağı modelinin tespiti için farklı sinir ağı topolojileri kullanılmıştır. Sinir ağları eğitilirken sinir ağı topolojisi ile ilgili olarak model parametrelerinde iki unsurun belirlenmesi gerekmektedir. Bunlar sinir ağına kullanılacak gizli katman sayısı ve bu katmanlarda bulunması gereken nöron (düğüm) sayısıdır.

Pratikte birçok problem için tek bir gizli katman kullanılması yeterlidir. İki tane gizli katman kullanılması gereken problem çok ender olmakla birlikte, kullanılması tercih edilebilmektedir. İki'den fazla gizli katman kullanılmasının ise teorik olarak bir açıklaması/faydası bulunmamaktadır [60] .

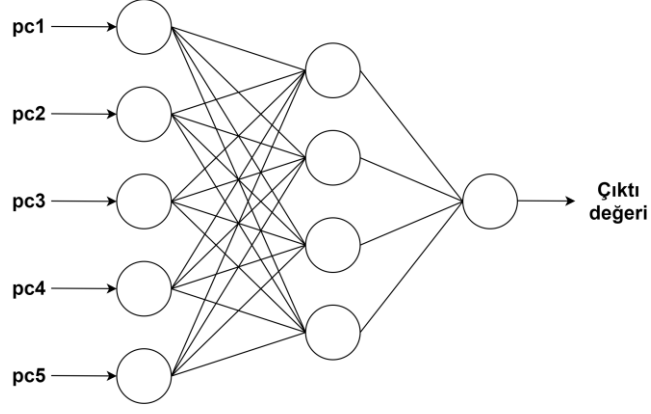
Sinir ağına gizli katmanlarda bulunması gereken düğüm sayısı için ise literatürde kesinleşmiş bir yöntem bulunmamakta, genellikle farklı sayıda düğüm içeren modeller denenmekte ve sonuçlar karşılaştırılmaktadır. Ancak gizli katmanlarda çok fazla nöron (düğüm) kullanılması modelde aşırı uyma sorununa yol açabilmektedir. Buna bağlı olarak doğru nöron sayısını belirlemek için aşağıdaki gibi tavsiye yöntemleri bulunmaktadır [60] .

- Gizli nöronların sayısı, girdi katmanının boyutu ile çıktı katmanının boyutu arasında olmalıdır.
- Gizli nöronların sayısı, girdi katmanının boyutunun $2/3$ 'ü ile çıktı katmanının boyutunun toplamı kadar olmalıdır.
- Gizli nöronların sayısı, giriş katmanının boyutunun iki katından az olmalıdır.

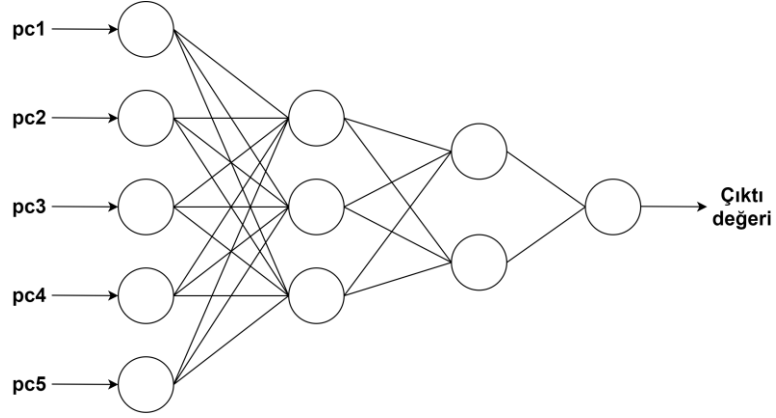
Bu doğrultuda, aşağıdaki gibi üç farklı topoloji belirlenmiştir.



Şekil 10 – Kullanılan Sinir Ağı Topolojileri -1



Şekil 10 – Kullanılan Sinir Ağı Topolojileri - 2



Şekil 10 – Kullanılan Sinir Ağı Topolojileri - 3

Sinir ağı modellerini oluşturmak için ‘h2o’ paketindeki sinir ağları [61] fonksiyonu kullanılmıştır. Tahmin edici değerler arasında yanlılık oluşmaması açısından değişkenler 0 ortalama ve 1 varyans ile standardize edilerek eğitilmiştir. Modellerin performans göstergeleri k-kat çapraz doğrulama yöntemi (k=10) ile elde edilmiştir.

Çizelge 10 – Yapay Sinir Ağları Sonuçları (Bütün Veri)

Hata Metriği	Hata Değeri (1. Sinir ağı)	Hata Değeri (2. Sinir ağı)	Hata Değeri (3. Sinir ağı)
Ortalama Karesel Hata	7167,004	7092,527	7007,608
Kök Ortalama Karesel Hata	84,658	84,217	83,711
Ortalama Mutlak Hata	54,953	53,532	52,479
Kök Ortalama Karesel Logaritmik Hata	1,370	1,348	1,325
Ortalama Artık Sapma	7167,004	7092,527	7007,608

Eđitilen sinir ađı modellerinin hata metrikleri karřılařtırıldıđında modellerin tahminleme bařarısında ok nemli bir fark olmadıđı grlmřtr. Bununla birlikte 3. sinir ađı modelinin (iki gizli katman, 3+2 gizli dđmden oluřan modelin) diđerlerine gre daha bařarılı olduđu grlmřtr. Bundan dolayı diđer algoritmalarla karřılařtırmak iin  model arasından 3. sinir ađı seilmiřtir.

4.6.1.3 Rastgele Orman

Rastgele orman algoritması genellikle sınıflandırma iin kullanılsa da regresyon problemlerinde tahmin iin de kullanılabilir. Rastgele orman ynteminde modelin eđitilebilmesi iin iki tane temel parametrenin belirlenmesi gerekmektedir. Bunlar modelde bulunacak ađa sayısı ile her bir ađacın maksimum derinliđidir. Leo Breiman [62] rastgele orman algoritmasında istenildiđi kadar ađa kullanılabileceđini belirtmiřse de bu durum modellerin eđitilme sresini artırabilmektedir. Bir diđer alıřmada ađa sayısının 68-128 arasında olması gerektiđi belirtilmiřtir [63] . Diđer parametre olan maksimum derinliđin ise yksek belirlenmesinin modelde ařırı uyma sorununa yol aabileceđi belirtilmiřtir. Bu bađlamda regresyon problemi iin, regresyon ađacı sayısının 70, 85 ve 100, maksimum derinliđin ise 10, 15 ve 20 olduđu dokuz farklı rastgele orman modeli eđitilmiřtir.

Modeller eđitilirken ‘h2o’ paketindeki Rastgele Orman Algoritması [64] uygulanmıřtır ve modeller k-kat apraz dođrulama yntemi (k=10) ile karřılařtırılmıřtır.

izelge 11 – Rastgele Orman Modellerinin Sonuları (Btn Veri)

		Regresyon Ađacı Sayısı		
		70	85	100
Maksimum Derinlik	10	53,335	53,403	53,263
	15	53,387	53,223	53,210
	20	53,281	53,255	53,041

Farklı parametrelere göre çalıştırılan rastgele orman modelleri ortalama mutlak hata (OMH) değerlerine göre karşılaştırıldığında modeller arasında önemli bir fark olmadığı görülmüştür. Rastgele orman modelinde daha fazla ağaç kullanmanın modelin performansına çok fazla etki etmediği görülmüş, maksimum derinliğin artırılmasının ise modellerin performansını çok az iyileştirdiği görülmüştür. Bu doğrultuda ağaç sayısı 70 derinliği ise 20 olan model diğer algoritmalarla karşılaştırmak üzere seçilmiştir. Seçilen modelin diğer hata metrikleri aşağıdaki gibidir.

Çizelge 12 – Seçilen Modelin Sonuçları (Rastgele Orman - Bütün Veri)

Hata Metriği	Hata Değeri
Ortalama Karesel Hata	7028,443
Kök Ortalama Karesel Hata	83,835
Ortalama Mutlak Hata	53,281
Kök Ortalama Karesel Logaritmik Hata	1,338
Ortalama Artık Sapma	7028,443

4.6.1.4 Gradyan Artırma

Gradyan artırma algoritmasında modelin eğitilebilmesi için, rastgele orman algoritmasında olduğu gibi, modelde bulunacak ağaç sayısı ile her bir ağacın maksimum derinliği olmak üzere iki tane temel parametrenin belirlenmesi gerekmektedir. Bu bağlamda gradyan artırma modelinde, rastgele orman algoritmasında olduğu gibi regresyon ağacı sayısının 70, 85 ve 100, maksimum derinliğin ise 10, 15 ve 20 olduğu dokuz farklı gradyan artırma modeli eğitilmiş ve karşılaştırılmıştır. Regresyon problemi için 'h2o' paketindeki Gradyan Artırma Algoritması [65] uygulanmış ve modeller yine k-kat çapraz doğrulama yöntemi (k=10) ile karşılaştırılmıştır.

Çizelge 13 – Gradyan Artırma Modellerinin Sonuçları (Bütün Veri)

		Regresyon Ağacı Sayısı		
		70	85	100
Maksimum Derinlik	10	53,381	53,586	53,480
	15	54,334	54,741	55,099
	20	55,553	55,593	55,826

Farklı parametrelere göre çalıştırılan gradyan artırma modelleri ortalama mutlak hata (OMH) değerlerine göre karşılaştırıldığında kullanılan ağaç sayısının önemli bir etkisi olmadığı görülmüştür. Bununla birlikte maksimum derinlik parametresinin arttıkça modellerin tahmin başarısının azaldığı görülmüştür. Yani modellerde derinliğin artmasıyla aşırı uyma sorunu görülmüş, bu da tahmin performansını olumsuz etkilemiştir. Bu doğrultuda ağaç sayısı 70 derinliği ise 10 olan model diğer algoritmalarla karşılaştırmak üzere seçilmiştir. Seçilen modelin diğer hata metrikleri aşağıdaki gibidir.

Çizelge 14 – Seçilen Modelin Sonuçları (Gradyan Artırma - Bütün Veri)

Hata Metriği	Hata Değeri
Ortalama Karesel Hata	7188,01
Kök Ortalama Karesel Hata	84,782
Ortalama Mutlak Hata	53,381
Kök Ortalama Karesel Logaritmik Hata	--
Ortalama Artık Sapma	7188,01

4.6.2. Sadece Belirlenen Hat (4.hat) ve Yıl (2012) için Modellerin Eğitilmesi

Spesifik bir hat bazında tahmin performanslarının değişip değişmeyeceğini görmek amacıyla sadece 4 numaralı üretim hattının 2012 yılına ait ikinci bir veri seti oluşturulmuş, bu ikinci veri setine de TBA uygulanmıştır. Bu bölümde hazırlanan ikinci veri seti üzerinde veriden öğrenme algoritmaları uygulanmıştır.

4.6.2.1 Doğrusal Regresyon

Doğrusal regresyon modelinin eğitilmesi için yine 'h2o' paketindeki genelleştirilmiş doğrusal modeller fonksiyonu kullanılmıştır. Tahmin edici değerler arasında yanlılık oluşmaması açısından değişkenler 0 ortalama ve 1 varyans ile standardize edilerek eğitilmiştir. Modelin performans göstergeleri k-kat çapraz doğrulama yöntemi (k=10) ile elde edilmiştir. Verinin uyum sağladığı doğrusal regresyon modeli hedef değerini değişkenliğini seçilen tahmin edicilerle açıklamada son derece başarısız gözükmiştir. R kare değeri -0.042, ortalama mutlak hata (OMH) değeri ise 65,69 çıkmıştır.

Çizelge 15 – Doğrusal Regresyon Sonuçları (4.hat 2012 Yılına Ait Veri)

Değişken	Katsayılar	Standart Hata	Z Değeri	P değeri
Sabit değişken	135,538	4,366	31,042	0
pc1	-20,848	11,614	-1,795	0,073
pc2	15,967	17,316	0,922	0,357
pc3	5,224	26,682	0,195	0,844
pc4	-21,687	47,767	-0,454	0,650
pc5	-15,004	90,545	-0,165	0,868

Hata Metriği	Hata Değeri
Ortalama Karesel Hata	8412,963
Kök Ortalama Karesel Hata	91,722
Ortalama Mutlak Hata	65,690
Kök Ortalama Karesel Logaritmik Hata	1,193
Ortalama Artık Sapma	8412,963
R kare	-0,042

4.6.2.2 Sinir Ağları

Sinir ağı modeli olarak ilk veri setinde kullanılan üç farklı sinir ağı topolojisi eğitilmiştir. (Tek gizli katman - 3 düğüm, tek gizli katman – 4 düğüm, iki gizli katman 3+2 düğüm) Değişkenler 0 ortalama ve 1 varyans ile standardize edilerek eğitilmiştir. Modelin performans göstergeleri k-kat çapraz doğrulama yöntemi (k=10) ile elde edilmiştir.

Çizelge 16 – Yapay Sinir Ağı Sonuçları (4.hat 2012 Yılına Ait Veri)

Hata Metriği	Hata Değeri (1. Sinir ağı)	Hata Değeri (2. Sinir ağı)	Hata Değeri (3. Sinir ağı)
Ortalama Karesel Hata	8273,028	9174,181	8043,42
Kök Ortalama Karesel Hata	90,956	95,781	89,683
Ortalama Mutlak Hata	64,610	67,583	64,064
Kök Ortalama Karesel Logaritmik Hata	1,194	1,203	1,182
Ortalama Artık Sapma	8273,028	9174,181	8043,42

Eğitilen sinir ağı modellerinin hata metrikleri karşılaştırıldığında 3. sinir ağı modelinin (iki gizli katman, 3+2 gizli düğümden oluşan modelin) diğerlerine göre biraz daha başarılı olduğu görülmüştür. Bundan dolayı diğer algoritmalarla karşılaştırmak için üç model arasından 3. sinir ağı seçilmiştir.

4.6.2.3 Rastgele Orman

Regresyon problemi için, ilk veri setinde (bütün veri seti) olduğu gibi, regresyon ağacı sayısının 70, 85 ve 100, maksimum derinliğin ise 10, 15 ve 20 olduğu dokuz farklı rastgele orman modeli eğitilmiştir. Yine 'h2o' paketindeki Rastgele Orman Algoritması uygulanmıştır. Modeller k-kat çapraz doğrulama yöntemi (k=10) ile karşılaştırılmıştır.

Çizelge 17 – Rastgele Orman Modellerinin Sonuçları (4.hat 2012 Yılına Ait Veri)

		Regresyon Ağacı Sayısı		
		70	85	100
Maksimum Derinlik	10	65,727	65,291	64,927
	15	64,505	65,082	64,470
	20	64,231	64,662	64,388

Farklı parametrelere göre çalıştırılan rastgele orman modelleri ortalama mutlak hata (OMH) değerlerine göre karşılaştırıldığında modeller arasında önemli bir fark olmadığı görülmüştür. Rastgele orman modelinde daha fazla ağaç kullanmanın modelin performansına çok fazla etki etmediği görülmüş, maksimum derinliğin artırılmasının ise modellerin performansını çok az iyileştirildiği görülmüştür. Bu doğrultuda ağaç sayısı 70 derinliği ise 20 olan model diğer algoritmalarla karşılaştırmak üzere seçilmiştir. Seçilen modelin diğer hata metrikleri aşağıdaki gibidir.

Çizelge 18 – Seçilen Modelin Sonuçları (Rastgele Orman - 4.hat 2012 Yılına Ait Veri)

Hata Metriği	Hata Değeri
Ortalama Karesel Hata	8059,285
Kök Ortalama Karesel Hata	89,773
Ortalama Mutlak Hata	64,231
Kök Ortalama Karesel Logaritmik Hata	1,193
Ortalama Artık Sapma	8059,285

4.6.2.4 Gradyan Artırma

Gradyan artırma modelinde, rastgele orman algoritmasında olduğu gibi regresyon ağacı sayısının 70, 85 ve 100, maksimum derinliğin ise 10, 15 ve 20 olduğu dokuz farklı gradyan artırma modeli eğitilmiş ve karşılaştırılmıştır. Regresyon problemi için yine ‘h2o’ paketindeki Gradyan Artırma Algoritması uygulanmıştır. Modeller yine k-kat çapraz doğrulama yöntemi (k=10) ile karşılaştırılmıştır.

Çizelge 19 – Gradyan Artırma Modellerinin Sonuçları (4.hat 2012 Yılına Ait Veri)

		Regresyon Ağacı Sayısı		
		70	85	100
Maksimum Derinlik	10	65,652	65,153	66,847
	15	66,121	66,848	67,346
	20	67,516	67,666	68,030

Farklı parametrelere göre çalıştırılan gradyan artırma modelleri ortalama mutlak hata (OMH) değerlerine göre karşılaştırıldığında ilk veri setinde karşılaşılan durumlar ikinci veri seti için de görülmüştür. Kullanılan ağaç sayısının model performansında önemli bir etkisi olmadığı görülmüştür. Bununla birlikte maksimum derinlik parametresinin arttıkça modellerin tahmin başarısının azaldığı görülmüştür. Yani modellerde derinliğin artmasıyla aşırı uyma sorunu görülmüş, bu da tahmin performansını olumsuz etkilemiştir. Bu doğrultuda, tıpkı ilk veri setinde olduğu gibi, ağaç sayısı 70 derinliği ise 10 olan model diğer algoritmalarla karşılaştırmak üzere seçilmiştir. Seçilen modelin diğer hata metrikleri aşağıdaki gibidir.

Çizelge 20 – Seçilen Modelin Sonuçları (Gradyan Artırma - 4.hat 2012 Yılına Ait Veri)

Hata Metriği	Hata Değeri
Ortalama Karesel Hata	8565,234
Kök Ortalama Karesel Hata	92,548
Ortalama Mutlak Hata	65,652
Kök Ortalama Karesel Logaritmik Hata	1,124
Ortalama Artık Sapma	8565,234

4.7. Modellerin Sonuçlarının Karşılaştırılması

Önceki bölümde, veri ön işleme sonrası, işletmeye ait bütün üretim kayıp verisi için ve sadece belirlenen hattın (4. hat) 2012 yılına ait kayıp verisi için veriden öğrenme algoritmaları uygulanmıştı. Bundaki amaç, işletmede çalıştırılacak üretim hatlarında bir sonraki vardiyada kg bazında oluşabilecek kayıp miktarını tahmin etmektir. Tahmin modellerinin eğitilmesinde literatürde sıklıkla kullanılan temel yöntemlerden çoklu doğrusal regresyon, yapay sinir ağları, rastgele orman, gradyan artırma algoritmaları uygulanmıştır. Bu bölümde her iki veri seti için uygulanan algoritmaların sonuçları karşılaştırılmıştır.

Veriden öğrenme yöntemleri karşılaştırılırken k-kat çapraz doğrulama yöntemi (k=10) ile eğitilen modellerin hata metriklerine bakılmıştır. Elde edilen sonuçlar incelendiğinde uygulanan tüm algoritmalar için birinci veri seti (bütün kayıp verileri) ile eğitilmiş modellerin tahminleme başarısının, ikinci veri seti (sadece belirlenen hattın 2012 yılına ait kayıp verisi) ile eğitilmiş modellerin tahminleme başarısından daha yüksek olduğu gözlemlenmiştir. Bunun ilk veri setinin hacminin ikinci veri setinin hacmine göre daha büyük olmasından kaynaklandığı düşünülmüştür.

Algoritma bazında bir karşılaştırma yapıldığında ise her iki veri seti için de yapay sinir ağları ile eğitilen modelin tahminlemede diğer yöntemlere göre daha başarılı sonuç verdiği görülmüştür. Yapay sinir ağlarını ise sırasıyla, rastgele orman, gradyan artırma ve doğrusal regresyon yöntemleri izlemiştir.

En başarılı tahminleme başarısını gösteren yapay sinir ağları yönteminin performans ölçülerine bakıldığında, ilk veri seti için ortalama mutlak hatanın (OMH) 52,479 ikinci veri seti için ise ortalama mutlak hatanın (OMH) 64,064 olduğu görülmektedir. Bu da bütün veri için eğitilen YSA modelinin üretim hattındaki kayıp tahmininde kullanıldığında ortalama olarak +/- 52 kg sapabileceğini göstermiştir.

Eđitilen modellerin hata metrikleri ařađıdaki tablolarda karřılařtırmalı olarak gsterilmiřtir.

izelge 21 – Karřılařtırmalı Sonular (Bütün Veri)

Hata Metriđi	Dođrusal Regresyon	Sinir Ađları	Rastgele Orman	Gradyan Artırma
Ortalama Karesel Hata	7146,718	7007,608	7028,443	7188,01
Kk Ortalama Karesel Hata	84,538	83,711	83,835	84,782
Ortalama Mutlak Hata	54,601	52,479	53,281	53,381
Kk Ortalama Karesel Logaritmik Hata	1,358	1,325	1,338	--
Ortalama Artık Sapma	7146,718	7007,608	7028,443	7188,01

izelge 22 – Karřılařtırmalı Sonular (4.hat 2012 Yılına Ait Veri)

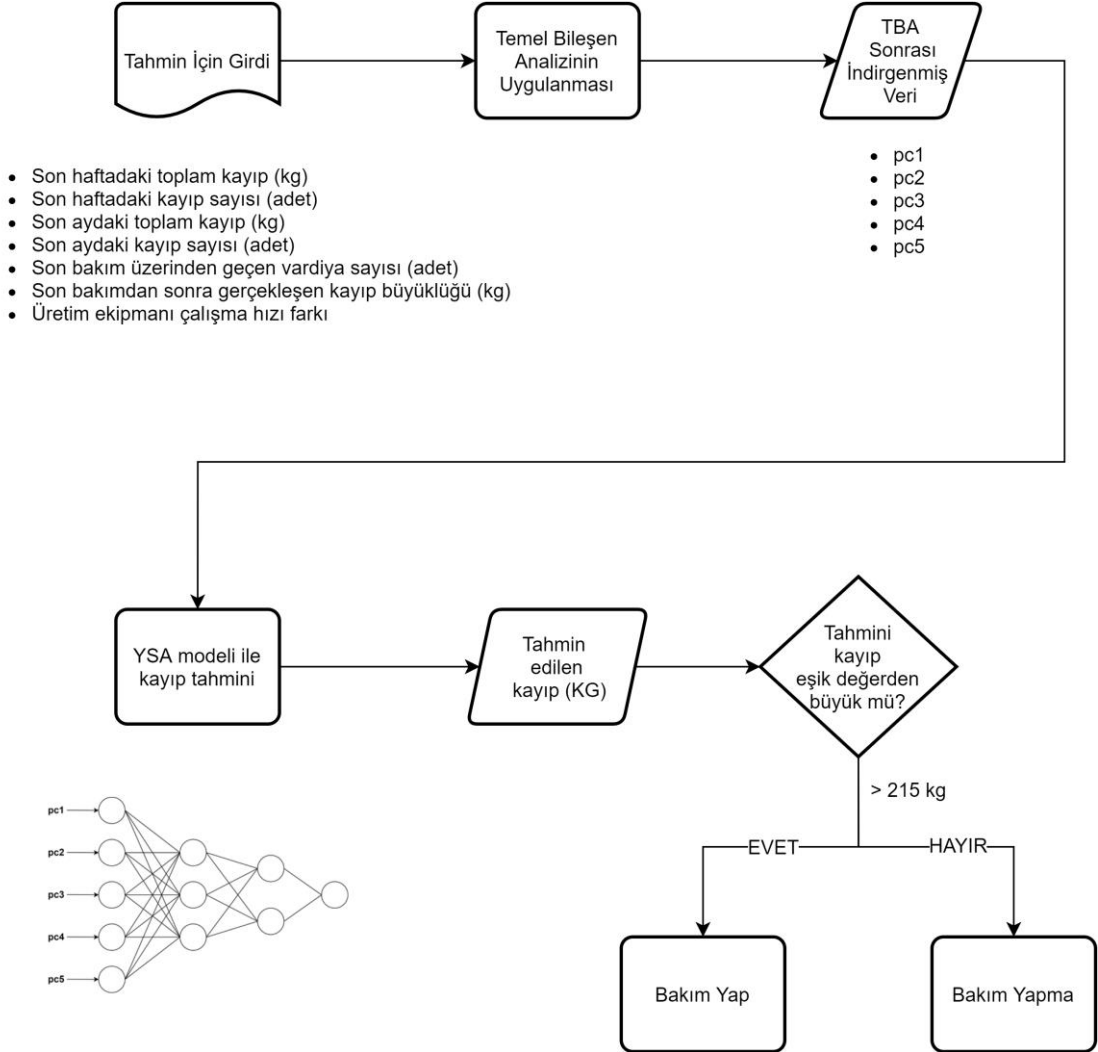
Hata Metriđi	Dođrusal Regresyon	Sinir Ađları	Rastgele Orman	Gradyan Artırma
Ortalama Karesel Hata	8412,963	8043,42	8059,285	9709,379
Kk Ortalama Karesel Hata	91,722	89,683	89,773	98,536
Ortalama Mutlak Hata	65,690	64,064	64,231	69,447
Kk Ortalama Karesel Logaritmik Hata	1,193	1,182	1,193	1,202
Ortalama Artık Sapma	8412,963	8043,42	8059,285	9709,379

4.8. Önerilen Bakım Modeli

İşletmede çalıştırılacak üretim hatlarında meydana gelecek kayıp miktarının tahmin edilmesinde en iyi sonuç veren yöntemin yapay sinir ağları yöntemi olduğu görülmüştü. Buradan hareketle işletmede çalıştırılan üretim hatları için bir bakım modeli önerilebileceği düşünülmüştür.

Önerilen bakım modeli kısaca şu adımlardan oluşmaktadır.

- 1- Tahmin için girdi verisi alınır.
 - a- Üretim için çalıştırılması planlanan üretim hattı için şu veriler raporlanarak alınır:
 - Son haftadaki toplam kayıp (kg)
 - Son haftadaki kayıp sayısı (adet)
 - Son aydaki toplam kayıp (kg)
 - Son aydaki kayıp sayısı (adet)
 - Son bakım üzerinden geçen vardiya sayısı (adet)
 - Son bakımdan sonra gerçekleşen toplam kayıp büyüklüğü (kg)
 - b- Üretim hattında çalıştırılacak üretim ekipmanının (makinenin) planlanan çalışma hızı ile gerçek çalışma hızı arasındaki fark hesaplanır. (planlanan çalışma hızı – gerçek çalışma hızı)
- 2- İlk maddedeki girdi verisi daha önce elde edilmiş olan TBA modeliyle dönüştürülür. Böylece birbirine ortogonal beş temel bileşenden oluşan kayıp tahmini için hazır bir veri satırı elde edilmiş olur.
- 3- Birbirine ortogonal beş temel bileşenden oluşan veri daha önce eğitilmiş YSA modeline girdi olarak sokulur, YSA modeli ile tahminleme yapılır, elde edilen değer ilgili hat için tahmini kayıp miktarını (kg bazında) verir.
- 4- Tahmini kayıp miktarının belirlenmiş olan eşik değerden (215 kg) büyük olup olmadığına bakılır. Eşik değerden büyükse ilgili vardiya için bakım planlanır, eşik değerden küçükse ilgili vardiya için bakım yapılmaz.

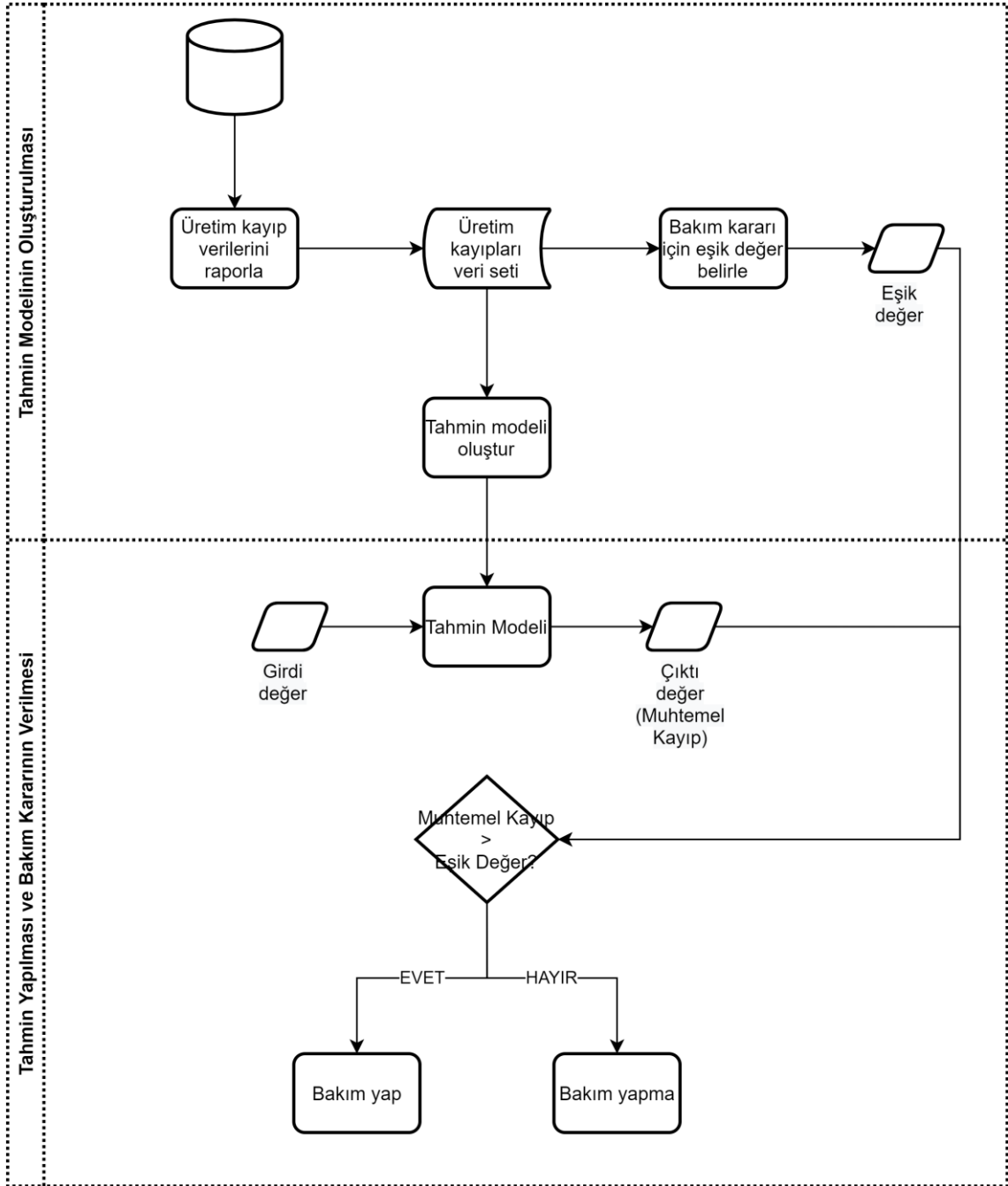


Şekil 11 – Önerilen Bakım Modeli

4.9. İşletme için Önerilen Bakım Modelinin Genelleştirilmesi

Bir önceki bölümde işletme için özetlenen bakım modeli, tez çalışmasına konu olan işletmeye özel bir çözüm olarak önerilmişti. Bu bölümde işletme için uygulanan yöntemden yola çıkarak, tahmini üretim kaybının büyüklüğüne dayanan bakım modeli için daha genel bir çerçeve sunulmak hedeflenmiştir.

- Söz konusu üretim hattı/hatları için üretim kayıp verilerini raporla. Üretim hattına ait üretim kayıp verisi tutulmuyorsa öncelikle ilgili hat/hatları izle, veri topla.
- Bakım kararını üretim kayıpları ile ilişkilendirebilmek için, veriden ya da uzman bilgisinden yararlanarak eşik (sınır) değeri belirle.
- Üretim hat/hatlarında muhtemel kayıp tahmin edebilecek tahmin modelleri için girdi parametrelerini belirle, tahmin modellerini oluştur.
- Tahmin modellerini karşılaştır, muhtemel üretim kaybını tahminleyecek en iyi tahmin modelini seç.
- İlgili üretim hat/hatları çalıştırılacağı zaman muhtemel üretim kaybını tahmin et.
- Muhtemel üretim kaybını eşik değerler karşılaştır. Eşik değerin üstündeyse bakım yap, değilse bakım yapma.



Şekil 12 – Önerilen Bakım Modeli (Genelleştirilmiş)

5. SONUÇ VE ÖNERİLER

Bu tez çalışması kapsamında kesintisiz PVC profili üretimi yapan bir işletmenin, hat ve vardiya bazında tutulmuş geçmişe yönelik üretim kayıp verilerinden, üretim ekipmanı çalışma hızlarından, son bakım üzerinden geçen vardiya sayısından ve son bakımdan sonra oluşan kayıp miktarından yola çıkarak ilgili hat ve vardiya için kg bazındaki muhtemel kayıp miktarını tahmin edecek modeller eğitilmiş ve karşılaştırılmıştır. İlgili hat için muhtemel kayıp miktarının tahmin edilebileceği öngörülmüş, böylece tahmin edilen değerin, belirlenen eşik değerle karşılaştırılarak bakım kararının bu doğrultuda verilebileceği düşünülmüştür.

İlki verinin tamamını içeren, ikincisi ise sadece belirli bir hattın kayıp verilerini içeren iki ayrı veri seti oluşturulmuştur. Daha sonra bu iki veri seti üzerinde temel bileşen analizi uygulanarak birbirleri arasında korelasyon olmayan 5 temel bileşenden ve bir hedef değerden oluşan veri setleri elde edilmiştir. Veri setleri ile üretim hattı için kg bazında kayıp miktarını tahmin edecek modeller eğitilmiştir. Modellerin eğitilmesinde doğrusal regresyon, sinir ağları, rastgele orman ve gradyan artırma algoritmaları kullanılmıştır. Algoritmaların başarısı karşılaştırıldığında kg bazında oluşabilecek üretim kaybının tahmin edilmesinde, her iki veri seti için de yapay sinir ağlarının daha başarılı olduğu görülmüştür.

Çalışmanın sonraki adımında ilgili vardiyada çalıştırılması planlanan üretim hattı için bakım kararının nasıl verilmesi gerektiği modellenmiştir. Önerilen modele göre, çalıştırılacak üretim hattına karar verildiğinde; öncelikle yapay sinir ağları modeli ile muhtemel kayıp miktarı tahmin edilecektir. Sonrasında tahmin edilen kayıp miktarı daha önceden belirlenmiş olan eşik değer ile karşılaştırılacak ve ilgili vardiya için bakım planlanıp planlanmayacağına kararı verilecektir.

Çalışmanın en son adımında ise işletme için önerilen model, üretim kayıplarını izleyen ve raporlayan diğer işletmeler için genel çerçeve olarak sunulmuştur.

Önceki çalışmalar incelendiğinde, kestirimci bakım modellerinin çoğunlukla üretim ekipmanlarına ait basınç, sıcaklık, hız, titreşim gibi verilerden yola çıkarak oluşturulduğu görülebilir. Son yıllarda, özellikle endüstride kullanılan sensör teknolojisinin gelişmesiyle, ekipmanlara yönelik bu tarz verilerin sürekli olarak izlenmesi ve toplanması kolaylaşmıştır. Bu da bakım planlamada veri odaklı kestirimci bakım yönteminin daha çok ön plana çıkmasını sağlamıştır. Ancak küçük ve orta boyuttaki işletmeler için üretim ekipmanının durumunu izleyecek bu tarz pahalı sensör sistemlerinin kurulması mümkün olmayabilmektedir. Bu çalışma kapsamında da, üretim ekipmanlarına yönelik bu tarz sensör verileri olmaksızın, üretim hatlarının geçmişteki kayıp verilerinden yola çıkarak, vardiya boyunca meydana gelecek muhtemel kayıp büyüklüğünü tahmin edecek modeller eğitilmiştir. Tahmin edilen muhtemel kayıp büyüklüğüne göre de bakım kararını belirleyecek bir model önerilmiştir.

Çalışma kapsamında, yüksek maliyetli sensör sistemlerine sahip olmayan, ancak üretim kayıplarını raporlayan ve izleyen küçük ve orta boydaki işletmeler için de, elde bulunan veriden yola çıkarak ve temel yöntemler kullanarak bakım modelleri oluşturulabileceği gösterilmiştir.

6. KAYNAKLAR

- [1] V. Sönmez, M. C. Testik ve Ö. M. Testik, «Overall equipment effectiveness when production speeds and stoppage durations are uncertain,» *The International Journal of Advanced Manufacturing Technology*, cilt 95, no. 1-4, pp. 121-130, 2018.
- [2] S. Nakajima, TPM Development Program: Implementing Total Productive Maintenance, Productivity press, 1989.
- [3] E. Hupjé, «9 Types Of Maintenance: How To Choose The Right Maintenance Strategy,» 21 Haziran 2018. [Çevrimiçi]. Available: <https://www.roadtoreliability.com/types-of-maintenance>.
- [4] P. P. Ruiz, B. K. Foguem ve B. Grabot, «Generating knowledge in maintenance from Experience Feedback,» *Knowledge - Based Systems*, cilt 68, pp. 4-20, 2014.
- [5] «EN13306:2001, Maintenance Terminology – European Standard,» CEN (European Committee for Standardization), Brussels, 2001.
- [6] R. Gouriveau, K. Medjaher ve N. Zerhouni, From Prognostics and Health Systems Management to Predictive Maintenance 1: Monitoring and Prognostics, ISTE Ltd and John Wiley & Sons, Inc., 2016.
- [7] R. K. Mobley, An Introduction to Predictive Maintenance, Elsevier, 2002.
- [8] E. Alpaydm, Introduction to Machine Learning, MIT press, 2020.
- [9] G. James, D. Witten, T. Hastie ve R. Tibshirani, An Introduction to Statistical Learning, New York: Springer, 2013.
- [10] J. Lee, E. Lapira, B. Bagheri ve H.-a. Kao, «Recent advances and trends in predictive manufacturing systems in big data environment,» *Manufacturing Letters*, cilt 1, no. 1, pp. 38-41, 2013.
- [11] J. Lee, «Machine performance monitoring and proactive maintenance in computer-integrated manufacturing: review and perspective,» *International Journal of Computer Integrated Manufacturing*, pp. 370-380, 2007.

- [12] J. Lee, «Measurement of machine performance degradation using a neural network mode,» *Computers in Industry*, cilt 30, no. 3, pp. 193-209, 1996.
- [13] R. C. M. Yam, P. W. Tse, L. Li ve P. Tu, «Intelligent Predictive Decision Support System for Condition-Based Maintenance,» *The International Journal of Advanced Manufacturing Technology*, cilt 17, no. 5, pp. 383-391, 2001.
- [14] M. Marseguerra, E. Zio ve L. Podofillini, «Condition-based maintenance optimization by means of genetic algorithms and Monte Carlo simulation,» *Reliability Engineering & System Safety*, cilt 77, no. 2, pp. 151-165, 2002.
- [15] A. Grall, L. Dieulle, C. Bérenguer ve M. Roussignol, «Continuous-time predictive-maintenance scheduling for a deteriorating system,» *IEEE Transactions on Reliability*, cilt 51, no. 2, pp. 141-150, 2002.
- [16] J. v. Noortwijk, «A survey of the application of gamma processes in maintenance,» *Reliability Engineering & System Safety*, cilt 94, no. 1, pp. 2-21, 2009.
- [17] J. Lee, J. Ni, D. Djurdjanovic, H. Qiu ve H. Liao, «Intelligent prognostics tools and e-maintenance,» *Computers in Industry*, cilt 57, no. 6, pp. 476-489, 2006.
- [18] M. Daigle, «Model-Based Prognostics,» 10 1 2014. [Çevrimiçi]. Available: https://www.phmsociety.org/sites/phmsociety.org/files/Daigle-ModelBasedPrognostics-Tutorial-PHM2014_1.pdf.
- [19] Ö. F. Eker, F. Çamcı ve I. Jennions, «Major Challenges in Prognostics: Study on Benchmarking,» 2012.
- [20] X. S. Si, W. Wang, C. Hu ve D. H. Zhou, «Remaining Useful Life Estimation Based on a Nonlinear Diffusion Degradation Process,» *IEEE Transactions on Reliability*, cilt 61, no. 1, pp. 50-67, 2012.
- [21] K. Goebel, B. Saha, A. Saxena, J. R. Celaya ve J. P. Christophersen, «Prognostics in battery health management,» *IEEE instrumentation & measurement magazine*, cilt 11, no. 4, pp. 33-40, 2008.
- [22] B. Saha, K. Goebel, S. Poll ve J. Christophersen, «Prognostics Methods for Battery Health Monitoring Using a Bayesian Framework,» *IEEE Transactions on Instrumentation and Measurement*, cilt 58, no. 2, pp. 291 - 296, 2009.

- [23] B. Saha, K. Goebel ve J. Christophersen, «Comparison of prognostic algorithms for estimating remaining useful life of batteries,» *Transactions of the Institute of Measurement and Control*, cilt 31, no. 3-4, pp. 293-308, 2009.
- [24] A. Saxena, K. Goebel, D. Simon ve N. Eklund, «Damage propagation modeling for aircraft engine run-to-failure simulation,» *2008 International Conference on Prognostics and Health Management*, 2008.
- [25] G. Curcurù, G. Galante ve A. Lombardo, «A predictive maintenance policy with imperfect monitoring,» *Reliability Engineering & System Safety*, cilt 95, no. 9, pp. 989-997, 2010.
- [26] M.-Y. You, F. Liu, W. Wang ve G. Meng, «Statistically Planned and Individually Improved Predictive Maintenance Management for Continuously Monitored Degrading Systems,» *IEEE Transactions on Reliability*, cilt 59, no. 4, pp. 744-753, 2010.
- [27] S. Sarkar, X. Jin ve A. Ray, «Data-driven fault detection in aircraft engines with noisy sensor measurements,» *Journal of Engineering for Gas Turbines and Power*, cilt 133, no. 8, 2011.
- [28] R. Ahmad ve S. Kamaruddin, «An overview of time-based and condition-based maintenance in industrial application,» *Computers & Industrial Engineering*, cilt 63, no. 1, pp. 135-149, 2012.
- [29] K. Liu, N. Gebraeel ve J. Shi, «A Data-Level Fusion Model for Developing Composite Health Indices for Degradation Modeling and Prognostic Analysis,» *IEEE Transactions on Automation Science and Engineering*, cilt 10, no. 3, pp. 652-664, 2013.
- [30] A. V. Horenbeek ve L. Pintelon, «A dynamic predictive maintenance policy for complex multi-component systems,» *Reliability Engineering & System Safety*, cilt 120, no. SI, pp. 39-50, 2013.
- [31] H. Li, D. Parikh, Q. He, B. Qian, Z. Li, D. Fang ve A. Hampapur, «Improving rail network velocity: A machine learning approach to predictive maintenance,» *Transportation Research Part C: Emerging Technologies*, cilt 45, no. SI, pp. 17-

- 26, 2014.
- [32] T. Praveenkumar, M. Saimurugan, P. Krishnakumar ve K. I. Ramachandran, «Fault diagnosis of automobile gearbox based on machine learning,» *12th GLOBAL CONGRESS ON MANUFACTURING AND MANAGEMENT*, Vellore, INDIA, 2014.
- [33] G. A. Susto, A. Schirru, S. Pampuri, S. McLoone ve A. Beghi, «Machine Learning for Predictive Maintenance: A Multiple Classifier Approach,» *IEEE Transactions on Industrial Informatics*, cilt 11, no. 3, pp. 812-820, 2015.
- [34] R. Prytz, S. Nowaczyk, T. Rögnvaldsson ve S. Byttner, «Predicting the need for vehicle compressor repairs using maintenance records and logged vehicle data,» *Engineering Applications of Artificial Intelligence*, cilt 41, pp. 139-150, 2015.
- [35] K. T. Huynh, A. Barros ve C. Bérenguer, «Multi-Level Decision-Making for The Predictive Maintenance of k -Out-of-n :F Deteriorating Systems,» *IEEE Transactions on Reliability*, cilt 64, no. 1, pp. 94-117, 2015.
- [36] P. Lim, C. K. Goh ve K. C. Tan, «A time window neural network based framework for Remaining Useful Life estimation,» *International Joint Conference on Neural Networks (IJCNN)*, 2016.
- [37] X. Li, Q. Ding ve J.-Q. Sun, «Remaining useful life estimation in prognostics using deep convolution neural networks,» *Reliability Engineering and System Safety*, cilt 172, pp. 1-11, 2018.
- [38] B. Luo, H. Wang, H. Liu, B. Li ve F. Peng, «Early Fault Detection of Machine Tools Based on Deep Learning and Dynamic Identification,» *IEEE Transactions on Industrial Electronics*, cilt 66, no. 1, pp. 509-518, Ocak 2019.
- [39] B. Zhang, S. Zhang ve W. Li, «Bearing performance degradation assessment using long short-term memory recurrent network,» *Computers in Industry*, cilt 106, pp. 14-29, Nisan 2019.
- [40] A. Diez-Olivan, J. D. Ser, D. Galar ve B. Sierra, «Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0,» *Information Fusion*, cilt 50, pp. 92-111, Ekim 2019.

- [41] R. Sahal, J. G. Breslin ve M. I. Ali, «Big data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case,» *Journal of Manufacturing Systems*, cilt 54, pp. 138-151, Ocak 2020.
- [42] B. Wang, Y. Lei, N. Li ve N. Li, «A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings,» *IEEE Transactions on Reliability*, cilt 69, no. 1, pp. 401-412, Mart 2020.
- [43] T. Kourti, «Multivariate Statistical Process Control and Process Control, Using Latent Variables,» *Comprehensive Chemometrics Chemical and Biochemical Data Analysis*, Elsevier, 2009, pp. 21-54.
- [44] V. Powell, «Principal Component Analysis Explained Visually,» 12 Şubat 2015. [Çevrimiçi]. Available: <https://setosa.io/ev/principal-component-analysis/>.
- [45] G. Singh, «Understanding Principle Component Analysis(PCA) step by step,» 7 Ocak 2020. [Çevrimiçi]. Available: <https://medium.com/analytics-vidhya/understanding-principle-component-analysis-pca-step-by-step-e7a4bb4031d9>.
- [46] M. Lacey, *Statistics 101-103, Introduction to Statistics - Lecture Notes, Multiple Linear Regression*, 1997.
- [47] C. Shalizi, 36-401, *Modern Regression Lecture Notes, Lecture 13: Simple Linear Regression in Matrix Format*, 2015.
- [48] J. W. Osborne ve E. Waters, «Four assumptions of multiple regression that researchers should always test,» *Practical Assessment, Research, and Evaluation*, cilt 8, no. 2, Ocak 2002.
- [49] Y.-Y. Chen, Y.-H. Lin, C.-C. Kung, M.-H. Chung ve I.-H. Yen, «Design and Implementation of Cloud Analytics-Assisted Smart Power Meters Considering Advanced Artificial Intelligence as Edge Analytics in Demand-Side Management for Smart Homes,» *Sensors*, cilt 19, no. 9, 2 Mayıs 2019.
- [50] T. M. Mitchell, *Machine Learning*, 1997.
- [51] J. Han ve C. Moraga, «The influence of the sigmoid function parameters on the speed of backpropagation learning,» *From Natural to Artificial Neural*

- Computation .IWANN 1995. Lecture Notes in Computer Science*, cilt 930, Berlin, Springer, 1995, pp. 195-201.
- [52] F. Bre, J. M. Gimenez ve V. D. Fachinotti, «Prediction of wind pressure coefficients on building surfaces using Artificial Neural Networks,» *Energy and Buildings*, cilt 158, pp. 1429-1441, 1 Ocak 2018.
- [53] «H2o Tutorial - Deep Learning (Neural Networks),» 2020 Nisan 14. [Çevrimiçi]. Available: <https://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/deep-learning.html>.
- [54] T. K. Ho, «Random decision forests,» *Proceedings of 3rd International Conference on Document Analysis and Recognition*, cilt 1, pp. 278-282, 1995.
- [55] L. Breiman, «Bagging predictors,» *Machine Learning*, cilt 24, no. 2, pp. 123-140, 1996.
- [56] L. Breiman, «Bias, variance, and arcing classifiers,» Statistics Department, University of California, Berkeley, CA, USA, 1996.
- [57] V. Sönmez ve M. C. Testik, «Using accurately measured production amounts to obtain calibration curve corrections of production line speed and stoppage duration consisting of measurement errors,» *The International Journal of Advanced Manufacturing Technology*, cilt 88, no. 9, p. 3257–3263, 17 Haziran 2016.
- [58] «H2o Tutorial - Principal Component Analysis (PCA),» 19 Kasım 2020. [Çevrimiçi]. Available: <https://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/pca.html>.
- [59] «H2o Tutorial - Generalized Linear Model (GLM),» 2 Eylül 2020. [Çevrimiçi]. Available: <https://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/glm.html>.
- [60] J. T. Heaton, *Introduction to Neural Networks for Java*, 2005.
- [61] «H2o tutorial - Deep Learning (Neural Networks),» 7 Ekim 2020. [Çevrimiçi]. Available: <https://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/deep-learning.html>.
- [62] L. Breiman ve A. Cutler, «Random Forests,» [Çevrimiçi]. Available:

https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm#remarks.

[63] T. M. Oshiro, P. S. Perez ve J. A. Baranauskas, «How Many Trees in a Random Forest?,» *8th International Conference on Machine Learning and Data Mining in Pattern Recognition*.

[64] «H2o Tutorial - Distributed Random Forest (DRF),» 10 Aralık 2020. [Çevrimiçi]. Available: <https://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/drfs.html>.

[65] «H2o tutorial - Gradient Boosting Machine (GBM),» 11 Kasım 2020. [Çevrimiçi]. Available: <https://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/gbm.html>.