



Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü

İşletme Anabilim Dalı

**ATÖLYE TİPİ ÜRETİM İÇİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ  
İLE ÜRETİM SAATİ TAHMİNİ: HAVACILIK VE SAVUNMA  
SANAYİİ UYGULAMASI**

Barış ÖZKAYA

Doktora Tezi

Ankara, 2021



ATÖLYE TİPİ ÜRETİM İÇİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE ÜRETİM  
SAATİ TAHMİNİ: HAVACILIK VE SAVUNMA SANAYİİ UYGULAMASI

Barış ÖZKAYA

Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü

İşletme Anabilim Dalı

Doktora Tezi

Ankara, 2021

## KABUL VE ONAY

Barış ÖZKAYA tarafından hazırlanan “Atölye Tipi Üretim İçin Makine Öğrenmesi Yöntemleri İle Üretim Saati Tahmini: Havacılık Ve Savunma Sanayii Uygulaması” başlıklı bu çalışma, 16/06/2021 tarihinde yapılan savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından Doktora Tezi olarak kabul edilmiştir.

---

Prof. Dr. Cevriye GENCER (Başkan)

---

Prof. Dr. Aydın ULUCAN (Danışman)

---

Prof. Dr. Yetkin ÇINAR

---

Doç. Dr. Mehmet SOYSAL

---

Doç. Dr. Kazım Barış ATICI

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylıyorum.

Prof. Dr. Uğur ÖMÜRGÖNÜLŞEN

Enstitü Müdürü

## YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinleri yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan **“Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge”** kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü / Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir. <sup>(1)</sup>
- Enstitü / Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ..... ay ertelenmiştir. <sup>(2)</sup>
- Tezimle ilgili gizlilik kararı verilmiştir. <sup>(3)</sup>

16/06/2021

**Barış ÖZKAYA**

<sup>1</sup>“Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge”

- (1) Madde 6. 1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6. 2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç imkanı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.
- (3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir \*. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir. Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir.

\* Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir.

## ETİK BEYAN

Bu alıřmadaki bütn bilgi ve belgeleri akademik kurallar erevesinde elde ettiđimi, grsel, iřitsel ve yazılı tm bilgi ve sonuları bilimsel ahlak kurallarına uygun olarak sunduđumu, kullandıđım verilerde herhangi bir tahrifat yapmadıđımı, yararlandıđım kaynaklara bilimsel normlara uygun olarak atıfta bulunduđumu, tezimin kaynak gsterilen durumlar dıřında zgn olduđunu, **Prof. Dr. Aydın ULUCAN** danıřmanlıđında tarafımdan retilildiđini ve Hacettepe niversitesi Sosyal Bilimler Enstits Tez Yazım Ynergesine gre yazıldıđını beyan ederim.

**Barıř ZKAYA**

## TEŞEKKÜR

Her zaman olduđu gibi bu zorlu süreçte de yanımda olan sevgili eşime, çocuklarıma, anneme, babama, kardeşime sonsuz sevgilerimle.

Çalışma ve aile hayatımı göz ardı etmeyerek bu tez çalışması süresince benden destek ve anlayışlılığını esirgemeyen değerli danışman hocam Prof. Dr. Aydın Ulucan'a ve doktora boyunca bilgi ve deneyimlerinden faydalanma fırsatı bulduğum tüm hocalarıma,

Kıymetli görüşleriyle zor günlerimde bana destek olan ve yol gösteren değerli arkadaşlarım ve hocalarım Mehmet Soysal, Mustafa Çimen, Barış Atıcı ve Bülent Çekiç'e,

Değerli katkılarından dolayı Prof. Dr. Cevriye Gencer, Prof. Dr. Yetkin Çınar ve Prof. Dr. Umut Tosun'a

Bana olan inançları ve destekleri için çalışma arkadaşlarım ve yöneticilerime sevgi, saygı ve teşekkürlerimi sunarım.

## ÖZET

ÖZKAYA, Barış. *Atölye Tipi Üretim İçin Makine Öğrenmesi Yöntemleri İle Üretim Saati Tahmini: Havacılık Ve Savunma Sanayii Uygulaması*, Doktora Tezi, Ankara, 2021.

Özellikle havacılık ve savunma sanayiinde, teklif verme süreçlerinde firmaların maliyet tahminlerini doğru olarak yapmalarının yanı sıra hızlı yapmaları, firmaların rekabetçiliği ve sürdürülebilirlikleri açısından oldukça önemlidir. İçinde buldukları koşullara göre firmalar kabiliyetlerindeki farklı tahmin yöntemlerinden duruma en uygununu seçmek zorundadırlar. Bu çalışma kapsamında firmaların hız ve hassasiyet dengesini gözeterek maliyet tahmini yapabilme becerileri ile ilgili bazı araştırma sorularına cevaplar aranmıştır. Bu araştırma sorularına cevap bulabilmek için bir havacılık ve uzay sanayi firması tarafından üretilen bazı ürünlere ait gerçek veri seti kullanılarak ürünlerin üretim saatlerinin tahmini için 3 farklı detayda tahmin yaklaşımları oluşturulmuştur. Her tahmin yaklaşımı kapsamında regresyon, yapay sinir ağları ve rastgele orman metotları kullanılarak modellerin tahmin performansları kıyaslanmıştır. Her tahmin yaklaşımı farklı detay seviyesinde olduğundan tahmin oluşturma süreleri ve tahmin performansları birbirlerinden farklılıklar göstermektedir. Bu çalışmada elde edilen sonuçların karar vericiye, tahmin yaklaşımı ve kullanılacak metotlarla ilgili kısıtlar ile teklife özgü diğer kısıtları bir arada değerlendirerek hangi tahmin yaklaşımını ve metodunu seçmesi gerektiği konusunda yol göstermesi amaçlanmıştır. Ayrıca bu çalışma kapsamında oluşturulan modellerdeki bağımlı değişken olan üretim saatlerine öğrenme eğrisine göre veri dönüşümü uygulanmasının model performansı üzerindeki etkisi araştırılmıştır.

### **Anahtar Sözcükler**

Maliyet Tahmini, Makine Öğrenmesi, Öğrenme Eğrisi



## ABSTRACT

ÖZKAYA, Barış. *Estimating Manufacturing Hours In A Job Shop Manufacturing Environment: Aerospace and Defense Industry Application*, Ph.D. Dissertation, Ankara, 2021

Especially for the aerospace production companies, besides making accurate cost estimations in the bidding process, it is also important to make cost estimations fast enough. This is important for the companies to maintain their competitiveness and sustainability. In such an environment, companies should choose the best of their cost estimating methodologies they are capable of regarding to the constraints of the proposal and the company. In this dissertation, some research questions arise about the ability of the company for making cost estimations accurate and fast enough. To find the answers for these research questions, a real data set belonging to an aerospace company is used in the models. Three different cost estimating approaches are built according to their level of detail. Artificial neural networks, random forest and linear regression methods are used in each approach and their estimating performances are compared to each other. Since each three cost estimating approaches are different in level of detail, time needed for making the estimation and the accuracy of each approach is also different from each other. Findings of this dissertation is aimed to help the decision maker to choose the right estimating approach and the right estimating method subject to the related constraints. Also, the impact of the data transformation of the manufacturing hours according to the learning curve on the model performance is investigated in this dissertation.

### **Anahtar Sözcükler**

Cost Estimation, Machine Learning, Learning Curve

## İÇİNDEKİLER

KABUL VE ONAY .....	i
YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI .....	ii
ETİK BEYAN .....	iii
TEŞEKKÜR .....	iv
ÖZET .....	v
ABSTRACT .....	vi
İÇİNDEKİLER .....	vii
TABLolar DİZİNİ.....	x
ŞEKİLLER DİZİNİ.....	xii
GİRİŞ .....	1
1. BÖLÜM: MALİYET TAHMİN YÖNTEMLERİ.....	5
1.1. MALİYET TAHMİN YÖNTEMLERİ SINIFLANDIRMA ÇALIŞMALARI.....	6
1.2. TEMEL MALİYET TAHMİN YÖNTEMLERİ .....	7
1.2.1. Analoji Yöntemi.....	7
1.2.2. Parametrik Yöntem.....	8
1.2.3. Analitik Yöntem.....	10
1.2.4. Temel Maliyet Tahmin Yöntemlerinin Karşılaştırılması.....	10
1.3. DİĞER MALİYET TAHMİN YÖNTEMLERİ.....	11
1.4. HAVACILIK VE UZAY SANAYİNDE MALİYET TAHMİN ÇALIŞMALARI.....	15
1.5. MOTİVASYON .....	17
2. BÖLÜM: KULLANILAN METOTLAR.....	19
2.1. YAPAY SİNİR AĞLARI .....	20
2.1.1. Yapay Sinir Ağları Modeli .....	20

2.1.2.	Parametre Seçimi.....	24
2.1.3.	Yapay Sinir Ağları Geçmiş Uygulamalar.....	26
<b>2.2.</b>	<b>RASTGELE ORMAN.....</b>	<b>27</b>
2.2.1.	Rastgele Orman Modeli.....	27
2.2.2.	Parametre Seçimi.....	28
2.2.3.	Rastgele Orman Geçmiş Uygulamalar .....	29
<b>2.3.</b>	<b>DOĞRUSAL REGRESYON .....</b>	<b>30</b>
<b>2.4.</b>	<b>PERFORMANS DEĞERLERNDİRME KRİTERLERİ .....</b>	<b>30</b>
<b>3.</b>	<b>BÖLÜM: UYGULAMA.....</b>	<b>32</b>
<b>3.1.</b>	<b>PROBLEM TANIMI.....</b>	<b>32</b>
3.1.1.	YSA Gizli Nöron Sayılarının belirlenmesi .....	35
3.1.2.	Model Performans Sonuçlarının Kararlılığı.....	38
<b>3.2.</b>	<b>VERİ SETİ .....</b>	<b>39</b>
<b>3.3.</b>	<b>SON ÜRÜN TAHMİN MODELİ (TY1-S-SÜ).....</b>	<b>45</b>
<b>3.4.</b>	<b>DETAY PARÇA SEVİYESİ TAHMİN MODELLERİ .....</b>	<b>48</b>
3.4.1.	Birleşik Veri Tabanı Yaklaşımı (TY2-D-BV).....	49
3.4.1.1.	Detay Parça Bazında Model Performansları .....	50
3.4.1.2.	Son Ürün Toplamları Bazında Model Performansları .....	51
3.4.2.	Proses Veri Tabanları Yaklaşımı (TY3-D-PV) .....	53
3.4.2.1.	Proses veri tabanları Model Performansları .....	60
3.4.3.	Birleşik Veri Tabanı (TY2-D-BV) ve Proses Veri Tabanları (TY3-D-PV) Modellerinin Performans Kıyaslamaları .....	63
3.4.3.1.	Detay Parça Bazında Kıyaslama.....	63
3.4.3.2.	Son Ürün Toplamları Bazında Kıyaslama .....	64
<b>3.5.</b>	<b>ÖĞRENME EĞRİSİNE GÖRE VERİ DÖNÜŞÜMÜ VE MODEL PERFORMANSINA ETKİSİ .....</b>	<b>65</b>
3.5.1.	Detay Parça Bazında.....	68

3.5.2. Son Ürün Toplamları Bazında .....	69
<b>3.6. MODELLERİN KARŞILAŞTIRILMASI .....</b>	<b>71</b>
3.6.1. Regresyon Modelleri.....	72
3.6.2. Yapay Sinir Ağları Modelleri .....	74
3.6.3. Rastgele Orman Modelleri .....	75
<b>3.7. KARŞILAŞTIRMA ÖZETİ VE YORUMLAR.....</b>	<b>78</b>
<b>SONUÇ .....</b>	<b>83</b>
<b>KAYNAKÇA .....</b>	<b>87</b>
<b>EK 1. ETİK KURUL MUAFİYET FORMU .....</b>	<b>94</b>
<b>EK 2. ORJİNALLİK RAPORU .....</b>	<b>95</b>

## TABLolar DİZİNİ

Tablo 1 - Maliyet Tahmin Yöntemlerinin Karşılaştırması, Curran ve diğer. (2004) .....	11
Tablo 2 - Maliyet Tahmin Yöntemleri Detaylı Sınıflandırması – Karşılaştırma, Niazi ve diğer. (2006).....	13
Tablo 3 - Maliyet tahmin metotları sınıflandırması (Zhao ve diğer, 2015) .....	14
Tablo 4 - Maliyet tahmin literatürü çalışmalarında çoklukla kullanılan özellikler	18
Tablo 5 - Optimum gizli nöron sayıları .....	25
Tablo 6 - Tahmin Yaklaşımları Kodları .....	34
Tablo 7 - Parametre sayısına göre gizli nöron sayısı .....	36
Tablo 8 - Gizli nöron sayısı ve MAPE ilişkisi .....	37
Tablo 9 - Son Ürünlere Ait Detay Parça Özeti .....	41
Tablo 10 - Üretim Saatleri Frekans Değerleri.....	43
Tablo 11 - Detay Parçalara Ait Tanımlayıcı Özet İstatistik Bilgiler.....	44
Tablo 12 - Son Ürün Verileri .....	45
Tablo 13 - Bootstrap Örneklem.....	46
Tablo 14 - Eğitim ve Test Verileri.....	47
Tablo 15 - Son Ürün Tahmin Performans Sonuçları .....	47
Tablo 16 - Birleşik Veri Tabanındaki Proses Tiplerine göre Parça Sayıları.....	49
Tablo 17 - Birleşik Veri Tabanı Eğitim ve Test Verileri .....	50
Tablo 18 - Birleşik Veri Tabanı Modellerinin Karşılaştırması.....	50
Tablo 19 - Birleşik Veri Tabanı Modelleri Sonuçlarının Son Ürün Bazında Karşılaştırması .....	51
Tablo 20 - Son Ürün Bazında Tahmin Performansı .....	52
Tablo 21 - Modellerin Performans Karşılaştırması - Büküm.....	54
Tablo 22 - Modellerin Performans Karşılaştırması - Ekstrüzyon basit.....	54
Tablo 23 - Modellerin Performans Karşılaştırması - Ekstrüzyon çekme.....	56
Tablo 24 - Modellerin Performans Karşılaştırması - Çekme .....	56
Tablo 25 - Modellerin Performans Karşılaştırması – hidroform büküm .....	57
Tablo 26 – Modellerin Performans Karşılaştırması – Talaşlı İmalat.....	58

Tablo 27 – Modellerin Performans Karşılaştırması – Roll .....	59
Tablo 28 – Modellerin Performans Karşılaştırması – Basit Kesme .....	60
Tablo 29 - Proses veri tabanları Modelleri Sonuçlarının Son Ürün Bazında Karşılaştırması .....	62
Tablo 30 – Proses veri tabanları Son Ürün Bazında Tahmin Performansı.....	62
Tablo 31 - Detay Parça Bazında Kıyaslama - Eğitim .....	63
Tablo 32 - Detay Parça Bazında Kıyaslama - Test .....	64
Tablo 33 - Son Ürün Bazında Kıyaslama - Eğitim.....	64
Tablo 34 - Son Ürün Bazında Kıyaslama - Test.....	65
Tablo 35 - Öğrenme Eğrisine Göre Gerçekleşme Dönüşümü.....	67
Tablo 36 - Öğrenme Eğrisi ile dönüştürülmüş birleşik veri modeli.....	68
Tablo 37 - Öğrenme Eğrisi Veri Dönüşümünün Model Performansına Etkisi - Eğitim.....	69
Tablo 38 - Öğrenme Eğrisi Veri Dönüşümünün Model Performansına Etkisi – Test.....	69
Tablo 39 - Öğrenme Eğrisi İle Dönüştürülmüş Modellerin Son Ürün Toplamları Bazında Tahmin Sonuçları.....	70
Tablo 40 - Öğrenme Eğrisi İle Dönüştürülmüş Modellerin Son Ürün Toplamları Bazında Tahmin Performansları .....	70
Tablo 41 - Öğrenme Eğrisi Veri Dönüşümünün Model Performansına Etkisi – Eğitim.....	71
Tablo 42 - Öğrenme Eğrisi Veri Dönüşümünün Model Performansına Etkisi - Test.....	71
Tablo 43 - Regresyon Modelleri Son Ürün Toplamları Bazında Performans Verileri.....	73
Tablo 44 - YSA Modelleri Son Ürün Toplamları Bazında Performans Verileri ..	74
Tablo 45 - RO Modelleri Son Ürün Toplamları Bazında Performans Verileri ....	76
Tablo 46 - Son Ürün ve Detay Seviye Tahmin Performansları Özeti (Test Verileri İçin MAPE).....	78
Tablo 47 - Maliyet Tahmin Sürecinde Olası Kısıtlar .....	82

## ŞEKİLLER DİZİNİ

Şekil 1 - Yıllara göre Maliyet Mühendisliği Konulu Yayın Sayıları - Mörtl ve Schmied (2016).....	6
Şekil 2 - Parametrik Tahmin Süreci.....	9
Şekil 3 - Maliyet Tahmin Yöntemleri Sınıflandırması, Niazi ve diğer. (2006).....	12
Şekil 4 - Yapay sinir ağları yapısı.....	21
Şekil 5 - YSA Transfer Fonksiyonları .....	22
Şekil 6 – Rastgele Orman Yapısı .....	28
Şekil 7 - Model Performans Sonuçları Karşılaştırma.....	38
Şekil 8 - Son Ürün 1 için Basitleştirilmiş Örnek Ürün Ağacı .....	40
Şekil 9 - Son ürün test verileri .....	48
Şekil 10 - Birleşik Veri Tabanı Son Ürün Toplamı - Test.....	52
Şekil 11 - Proses veri tabanları Modelleri Sonuçlarının Son Ürün Bazında Karşılaştırması - Test.....	62
Şekil 12 -Örnek Öğrenme Eğrisi ve Veri Raporlama Aralığı .....	68
Şekil 13 - Regresyon modelleri son ürün toplamları - Test .....	72
Şekil 14 - Son ürün toplamları bazında test verileri performans değerleri.....	73
Şekil 15 - YSA Modelleri Son Ürün Toplamları - Test .....	74
Şekil 16 - Son ürün toplamları bazında test verileri performans değerleri.....	75
Şekil 17 - RO modelleri son ürün toplamları - Test .....	76
Şekil 18 - Son Ürün Toplamları Bazında Test Verileri Performans Değerleri....	77
Şekil 19 - Teklif Hazırlama Zaman Skalası .....	79
Şekil 20 – Son Ürün Seviyesi İçin MAPE - Tahmin Süresi İlişkisi .....	80
Şekil 21 – Detay Parça Seviyesi İçin MAPE - Tahmin Süresi İlişkisi.....	81

## GİRİŞ

Günümüzde giderek sertleşen rekabet koşulları, çoğu sektör gibi havacılık ve savunma sanayii de büyük ölçüde etkilemektedir. Bu durumda firmalar, ürünlerini hem kaliteli hem de maliyet etkin olarak üretmek zorunda ve kar marjlarını çok iyi kontrol ederek pazardaki fiyatlarını daima rekabetçi tutmak zorundadırlar. Böyle bir ortamda maliyet tahmininin önemi de giderek artmaktadır.

Maliyet tahmini, karar verme kriteri olarak erken tasarım aşamasında, teklif verme sürecinde, alternatif üretim yöntemleri arasında karar vermede etkin olarak kullanılmaktadır. Maliyet tahmininin doğru olarak yapılmasının yanı sıra hızlı yapılması da oldukça önemlidir. Özellikle teklif verme süreçlerinde, bazı durumlarda firmalar tedarikçi seçimlerini hızlıca tamamlayabilmek için teklife çağrı dokümanlarına çok hızlı bir şekilde cevap verilmesini beklerler. Maliyet tahmini için literatürde birçok farklı yöntem bulunmasına rağmen, makine öğrenmesi yöntemleri, hızlı ve hassas maliyet tahmini gerektiren durumlarda başarılı bir şekilde kullanılmaktadır. Makine öğrenmesi yöntemleri ile yüksek açıklama kabiliyetli doğrusal olmayan modeller oluşturularak doğru ve hızlı tahminler yapılabilmektedir.

Tez kapsamında üç araştırma sorusuna cevap aranmıştır:

1. Bir teklif verme sürecinde içinde bulunulan koşullara göre tahmin süresi ile tahmin hassasiyeti dengesini dikkate alarak nasıl bir tahmin yaklaşımı seçilmesi gerekir?
2. Makine öğrenmesi yöntemlerinin tahmin performansı geleneksel yöntemlere göre üstün müdür?
3. Bir işin tekrarlanması sonucu oluşan öğrenme eğrisine göre üretim saatlerine veri dönüşümü uygulanmasının model performansına etkisi var mıdır?

Bu araştırma sorularına cevap bulabilmek için bir havacılık ve uzay sanayi firması tarafından üretilen bazı ürünlere ait gerçek veri seti kullanılarak ürünlerin



retim saatlerinin tahmini iin farklı yaklaşımlar kapsamında eşitli modeller oluşturulmuş ve bu modellerin tahmin performansları incelenmiştir. Her yaklaşım kapsamında regresyon, yapay sinir ađları ve rastgele orman yöntemleri kullanılarak performans kıyaslamaları yapılmıştır.

Birinci ve ikinci araştırma sorularına cevap bulabilmek için üç ana yaklaşım kapsamında farklı modeller oluşturulacaktır. Üüncü araştırma sorusu için de örnek olarak ikinci yaklaşım ele alınarak öğrenme eğrisinin etkisi araştırılacaktır.

Birinci yaklaşımda son ürün toplam üretim saati tahmini için regresyon, yapay sinir ađları ve rastgele orman modelleri oluşturulacaktır. Bu yaklaşımda model kapsamındaki veriler son ürün bazında olduğundan veri ve parametre sayısı kısıtlı kalmakta, bu da tahmin sürecini kısaltmaktadır. Bu yaklaşımdaki amaç az veri ile hızlı bir tahminde bulunmak gerektiğinde oluşturulacak modellerin tahmin performansını ortaya koyabilmektir.

İkinci ve üçüncü yaklaşım kapsamında son ürünleri oluşturan detay paralar bazında tahmin modelleri oluşturulacaktır. İkinci yaklaşım olarak birleşik veri tabanı modelleri kapsamında son ürünlerdeki detay paralara ait veriler ile farklı imalat proseslerini içeren tek bir veri tabanı oluşturulacak ve bu veri tabanı kullanılarak yapay sinir ađları, rastgele orman ve geleneksel çoklu doğrusal regresyon yöntemleri ile tahmin modelleri (birleşik veri tabanı modeli) oluşturulacaktır. Üüncü yaklaşımda, detay para veri tabanı, farklı imalat proseslerine göre 8 alt veri tabanına bölünerek her bir alt veri tabanında farklı tahmin modelleri (Proses veri tabanları modelleri) oluşturulacaktır. Proses veri tabanları ile oluşturulan modellerin tahmin gücünün birleşik veri tabanı modellerine göre daha iyi olması beklenirken, 8 farklı model kurulması gerektiđi için daha uzun zaman alacaktır. Tahmin için yeterli zaman olmadığı durumlarda, birleşik veri tabanı modellerinin kullanılıp kullanılmaması, tahmin sonuçları incelenerek tartışılacaktır.

Elde edilen tahmin sonuçlarına göre ayrıca, regresyon ile makine öğrenmesi yöntemlerinin farklı prosesler içeren daha dađınık veri setleri için tahmin performanslarının nasıl deđiştirdi de gözlenecektir. Bu alışmanın sonuçlarına

göre, çok farklı prosesler içeren büyük veri setleri için hangi tahmin yönteminin daha başarılı olabileceği öngörülerek, sonraki çalışmalara ışık tutulması amaçlanmaktadır. Böylece, teklif çalışmaları kapsamında, farklı tipte ve çok sayıda parça içeren, teklif süresi kısıtlı durumlarda kullanılabilecek yöntemlerin belirlenmesine katkı sağlanmış olacaktır.

Üçüncü araştırma sorusuna cevap bulabilmek için örnek olarak ikinci yaklaşım kullanılarak tahmin modellerindeki bağımlı değişken olan üretim sürelerine, öğrenme eğrisine göre veri dönüşümü uygulandığında, tahmin modeli sonuçlarının iyileşip iyileşmediği araştırılacaktır. Burada bahsedilen öğrenme eğrisi, yapılan bir işin tekrar sayısı ile o işin süresi arasındaki ilişki olarak tanımlanmaktadır.

Bu çalışmanın karar vericiye, teklife özgü kısıtlar ile tahmin yaklaşımı ve kullanılacak metotlar ile ilgili kısıtları bir arada değerlendirerek hangi tahmin yaklaşımı ve metodunu seçmesi gerektiği konusunda yol göstermesi amaçlanmıştır.

Literatür incelendiğinde farklı makine öğrenmesi yöntemlerinin performanslarının kıyaslandığı birçok çalışma bulunabilir. Fakat bu çalışmalar genelde tek tip imalat prosesi içeren veri setleri üzerinde yapılmıştır. Bu çalışmanın literatüre katkılarında ilki, farklı imalat prosesleri içeren daha kapsamlı veri setleri için makine öğrenmesi yöntemlerinin performanslarının nasıl değiştiğini ortaya koymak olacaktır. Çalışmanın literatüre ikinci katkısı, öğrenme eğrisinin, maliyet tahmin modelleri ile entegre olarak kullanıldığı ilk çalışma olmasıdır.

Tezin birinci bölümünde maliyet tahmin yöntemleri ile ilgili tanımlamalar ve sınıflandırma çalışmaları hakkında bilgi verilerek havacılık ve savunma sanayii için örnek maliyet tahmin çalışmalarından bahsedilecektir. İkinci bölümde, tez çalışması kapsamında kullanılan metotlar hakkında detaylı bilgiler verilerek geçmiş uygulamalardan örnekler verilecektir. Üçüncü bölümde öncelikle problem tanımı ortaya konularak araştırma soruları belirtilecektir. Daha sonra veri seti incelenerek çalışma kapsamında gerçekleştirilen uygulamalar detaylı olarak anlatılacaktır. Bu bölümün sonunda yapılan uygulamaların sonuçları birbirleri ile

kıyaslanarak sonuçlar yorumlanacaktır. Son bölümde ise sonuç ve önerilere yer verilecektir.

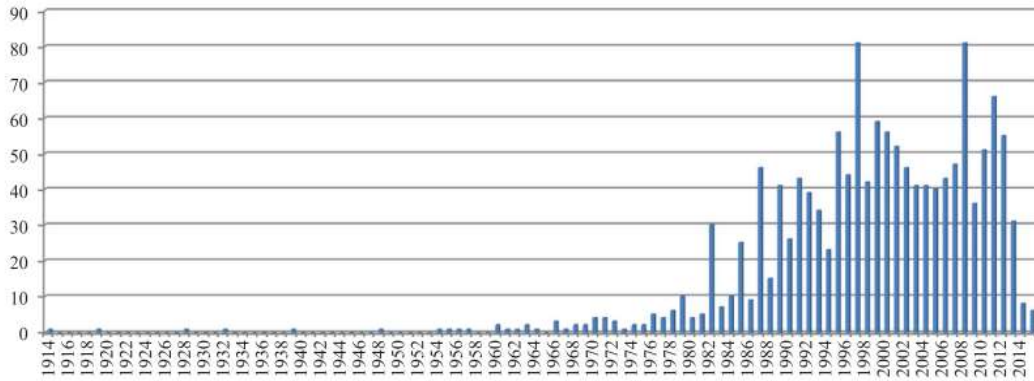
## 1. BÖLÜM

### MALİYET TAHMİN YÖNTEMLERİ

Doğru maliyet tahmini yapmak kurumlar için çok önemlidir; özellikle teklif çalışmalarını kapsamında maliyeti olması gerekenden yüksek tahmin etmek, kurumun yeni iş alamamasına, düşük tahmin etmek ise kurumun zarar etmesine neden olabilir. Bu nedenle özellikle artan rekabet koşullarında firmalar maliyet tahminine giderek daha fazla önem vermektedirler. Teknolojinin gelişmesi, pazardaki kar marjlarının azalması, küreselleşmenin de etkisiyle rekabet koşullarının artması gibi nedenlerle bu durum akademik çalışmalara da yansımış olup üniversite–sanayi işbirliğinin önemi de giderek artmaktadır.

Maliyetin, organizasyonlarda hassas ve kritik rol oynaması, maliyet tahmininin tasarım ve operasyonel stratejilerde önemli bir karar odağı olmasına neden olmaktadır. Optimum tasarım çözümleri yaratabilmek ve düşük maliyet, yüksek kalite ve zamanında teslimat anlamında müşteri memnuniyetinin en yüksek seviyede karşılayabilmek için şirketler tarafından doğru ve tutarlı maliyet tahminleri elde etmenin yolları araştırılmaktadır.

Maliyet her şirket için önemli olmasına rağmen bazı sektörlerde karmaşık ve daha hacimli üretim yapan şirketlerin daha çok odağında olmuştur. Maliyet tahmin literatürü incelendiğinde birkaç sektörün maliyet tahmin çalışmalarında ön plana çıktığı görülebilir. Yazılım geliştirme, inşaat ve üretim alanlarında maliyet tahmini literatürü oldukça gelişmiş olduğu görülmektedir. Üretim kapsamında havacılık ve uzay sanayi ile savunma sanayi, maliyet tahmin araştırmalarının sistematik ve kapsamlı olarak yapıldığı sektörlerdir. Bu durum literatüre de yansımış olup özellikle 1980'li yıllardan itibaren maliyet mühendisliği konulu yayınlarda gözle görülür bir artış olmuştur (Mörtl ve Schmied, 2016).



Şekil 1 - Yıllara göre Maliyet Mühendisliği Konulu Yayın Sayıları - Mörtl ve Schmied (2016)

### 1.1. MALİYET TAHMİN YÖNTEMLERİ SINIFLANDIRMA ÇALIŞMALARI

Ürün maliyet tahmini literatürü kapsamında basit standart ürünlerin yanında özelleştirilmiş montaj ürünlerine, proses bazında maliyet tahminlerinden genel giderler maliyet tahminine, ürünün konsept tasarım aşamasına özel maliyet tahmin tekniklerinden geç tasarım evresine genelleştirilmiş maliyet tahmin tekniklerine, geleneksel maliyet tahminlerinden özelleştirilmiş yeni maliyet tahmin tekniklerine kadar çok geniş bir yelpazede çalışmaların olduğu görülebilir.

Bazı araştırmacılar bu çalışmalar kapsamındaki maliyet tahmin tekniklerini sınıflandırmaya çalışmışlardır. Zhang ve diğer. (1996) yaptıkları çalışmada maliyet tahmin tekniklerini geleneksel detaylı kırımlı, basitleştirilmiş kırımlı, grup-teknoloji tabanlı, regresyon tabanlı ve aktivite tabanlı olmak üzere beş ana grupta toplamışlardır. Ben-Arieh ve Qian (2003), maliyet tahmin tekniklerini sezgisel, benzetim, parametrik ve analitik olarak gruplamışlardır. Shehab ve Abdalla (2002), maliyet tahmini yaklaşımlarını farklı bir açıdan ele alarak bilgi tabanlı, özellik tabanlı, işlevsel tabanlı ve operasyonel tabanlı olarak sınıflandırmışlardır. Cavalieri ve diğer. (2004) maliyet tahmin çalışmalarını benzetim, parametrik ve mühendislik yaklaşımı olmak üzere üç ana başlıkta toplamışlardır. Maliyet tahmin literatüründe birçok sınıflandırma çalışması

olmasına rağmen sistematik ve hiyerarşik bir çalışmanın eksikliğinden bahseden Niazi ve diğer. (2006), geniş kapsamlı bir sınıflandırma çalışması yapmışlardır.

Yapılan maliyet tahmin yöntemleri sınıflandırma çalışmalarının neredeyse hepsinde üç temel yöntem yer almaktadır. Bu yöntemler benzetim, parametrik ve analitik yöntemler olarak çalışmanın bağımsız bir elemanı ya da birden çok eleman grubunun toplandığı bir başlık olarak görülebilir. Bu yöntemler bu çalışma kapsamında temel maliyet tahmin yöntemleri başlığı altında incelenecektir.

## **1.2. TEMEL MALİYET TAHMİN YÖNTEMLERİ**

Basit sınıflandırmalar kapsamında maliyet tahmin yöntemleri genellikle analogi (benzetim), parametrik ve analitik olmak üzere üç ana kategoriye ayrılır. En geniş çaplı sınıflandırma Niazi ve diğer. (2006) tarafından yapılmış olup aslında bu çalışmada belirtilen her bir yöntem temel olarak bu üç kategoriden birine girer.

Analoji yönteminde, ürünün boyutları, karmaşıklığı gibi ana özellikler temel alınarak örnek parçanın maliyeti, maliyet tahmini yapılacak olan parçaya uyarlanır. Parametrik yöntemde ürünün maliyeti etkileyebilecek bazı tasarımsal özellikleri (boyutları, malzeme tipi, ağırlığı gibi) ile maliyeti arasında parametrik ilişkiler, regresyon teknikleri kullanılarak ortaya konur. Analitik yöntemde ise ürün, maliyet (malzeme, işçilik...) ya da yapısal (detay parçalar, montaj, boya...) alt bileşenlerine ayrılır, her bir alt bileşenin maliyeti tahmin edilir ve bunların toplamı ürünün maliyetini oluşturur.

### **1.2.1. Analoji Yöntemi**

Analoji yönteminde temel unsur, bir ürünün maliyetini, hedef ürüne, benzerliklerini gözeterek uyarlamaktır. Temel varsayım, benzer ürünlerin benzer maliyete sahip olmasıdır. Bu yöntemin başarılı olabilmesi için iki ürün arasındaki benzerliklerin ve farklılıkların doğru bir şekilde ortaya konulması ve bunların maliyete olan etkilerinin doğru bir şekilde yansıtılması gerekir. Bu yöntem ile tek bir örnek geçmiş veri ile yeni ürünün maliyet tahminini yapabilmek mümkündür. Fakat

burada maliyet tahmini uzmanının uzmanlığı gerçekten çok önemlidir. Bu yöntemde öncelikle ürünün tahmine baz olacak özellikleri belirlenir. Bu özellikler boyut, malzeme gibi fiziksel ya da teknoloji seviyesi, sağlamlık gibi farklı parametreler de içerebilir. Daha sonra bu özelliklere uygun olarak geçmiş veriler araştırılarak derlenir. Elde edilen geçmiş veriler gerekirse bugünkü zamana ve yeni ürünün şartlarına göre aylık ve toplam üretim adetleri de göz önüne alınarak normalize edilir. Uygulamada genellikle ilk ürünün maliyeti tahmin edilerek üretim takvimi ve toplam miktara göre öğrenme eğrisi uygulanarak toplam maliyet tahmin edilir.

Rush ve Roy (2001), çalışmalarında bu yöntemi detaylı bir şekilde irdeleyerek bu yöntemin kullanılması için detaylı bir prosedür ortaya koymuşlar ve bu prosedürün geçerliliğini havacılık, savunma ve otomotiv sektörlerinde yaptıkları anket çalışması ile doğrulamışlardır. Yazarlar analoji yöntemi kullanımında öncelikle nedensellik ilişkilerinin öneminin kavranması gerektiğini vurgulamışlardır. Duverlie ve Castelain (1999), çalışmalarında tasarım aşamasında parametrik ve analoji tahmin yöntemlerini karşılaştırmışlardır. Angelis ve Stamelos (2000), yazılım geliştirme efor tahmini için analoji yöntemini temel alan bir simülasyon çalışması yapmışlardır. Bootstrap metodu kullanarak geliştirdikleri istatistiksel simülasyon aracını aynı zamanda tahmin güven aralıklarını belirlemek için de kullanmışlardır.

### **1.2.2. Parametrik Yöntem**

Eck ve diğer. (2009) tanımına göre parametrik tahmin, maliyet tahmin ilişkileri ve matematiksel algoritmalar kullanarak yapılan maliyet tahminidir. Maliyeti etkileyen parametreler belirlenerek genellikle regresyon modelleri ile maliyet ile aralarındaki ilişkiler ortaya konur. Bağımlı değişken olarak maliyet, bağımsız değişken olarak parametreler kullanılarak elde edilen bu ilişkiler maliyet tahmin ilişkileri (MTİ) olarak adlandırılır. Parametreler, tahmin edilecek ürünün teknik, fiziksel ya da herhangi ölçülebilir farklı bir özelliği ile ilgili olabilir. Elde edilen maliyet tahmin ilişkileri kullanılarak hızlıca maliyet tahminleri yapılabilir.

Hızlı maliyet tahmini ihtiyacı ikinci dünya savaşı sırasındaki uçak üretimi esnasında ortaya çıktı. Savaş sonrası dönemde ABD’de oluşturulan çeşitli kurumların çalışmaları ile parametrik tahmin çalışmalarının temelleri atıldı. Bu çalışmalarda üretim maliyetlerinin yanı sıra tasarım ve tekrar etmeyen üretim maliyetleri için de parametrik çalışmalar yapıldı.

Parametrik tahmin süreci Şekil 2’de özetlenmiştir.



Şekil 2 - Parametrik Tahmin Süreci

Parametrik yöntemde dikkat edilmesi gereken hususlardan birisi parametrelerin kullanım aralığıdır. Parametreler, elde edildikleri geçmiş verilerin dışında bir aralıkta kullanıldığında hatalı tahmin ihtimali artar. Mesela 1 ile 3 metre arasındaki parçalardan oluşan bir veri seti ile oluşturulan parametrik denklemler, 3 metreden daha uzun parçaların maliyet tahmini için uygun olmayabilir. Dikkat edilecek bir başka husus da şirket dışında oluşturulmuş genel geçer parametrik denklemlerin kullanılmasında ortaya çıkar. Bu durumda maliyet tahmin ilişkilerinin kullanılacak olan şirkete göre uyarlanarak kalibre edilmesi gerekir.

Duverlie ve Castelain (1999), çalışmalarında tasarım aşamasında parametrik ve analogi tahmin yöntemlerini karşılaştırmışlardır. Watson ve diğer. (2006) çalışmalarında, havacılık ve uzay sanayiinde kullanılmak üzere, parça gruplarına göre farklı parametrik denklemler üreten excell tabanlı otomatik bir ara yüz geliştirmişlerdir. Cavalieri ve diğer. (2004) çalışmalarında otomotiv sanayinde bir üretim örneği için parametrik ve yapay sinir ağları modellerini karşılaştırmışlardır.



Xiaonan ve diğ er. (2019) uç ak üretim projeleri maliyet tahmini için genel bir parametrik modelleme çalışması yapmışlardır.

### **1.2.3. Analitik Yöntem**

Bu yöntemde tahmin edilecek ürün alt parçalara ya da alt işlemlere bölünür, bu alt bileşenler ayrı ayrı tahmin edilerek toplanmak suretiyle ürünün maliyet tahmini yapılır. Dolayısıyla bu yöntemde ürünün alt bileşenleri ile ilgili detaylı bilgiye sahip olunması gerekir. Alt bileşenler tasarım, mühendislik, malzeme, üretim maliyeti gibi veya daha detaylı unsurlar olabilir. Bu yöntemin en önemli dezavantajı tahmin edilecek ürün ile ilgili çok detaylı bilgiye sahip olunması gerekliliği ve tahmin için büyük çaba gerektirmesidir. Mandolini ve diğ er. (2019) çalışmalarında, saç metal parçaların üretim maliyeti tahmini için analitik yöntem kullanmışlardır. Üç boyutlu modellerden parçaların geometrik özelliklerini saptayarak aktivite ve özellik temelli bir analitik model geliştirmişlerdir. Lin ve diğ er. (2012) de tasarım aşamasındayken helikopter pervanelerinin maliyet tahmini amacıyla benzer bir çalışma yürütmüşlerdir. Schubel (2012) büyük rüzgar türbini kanatlarını üretim prosesleri detayında inceleyerek analitik bir maliyet tahmin modeli geliştirmiştir.

### **1.2.4. Temel Maliyet Tahmin Yöntemlerinin Karşılaştırılması**

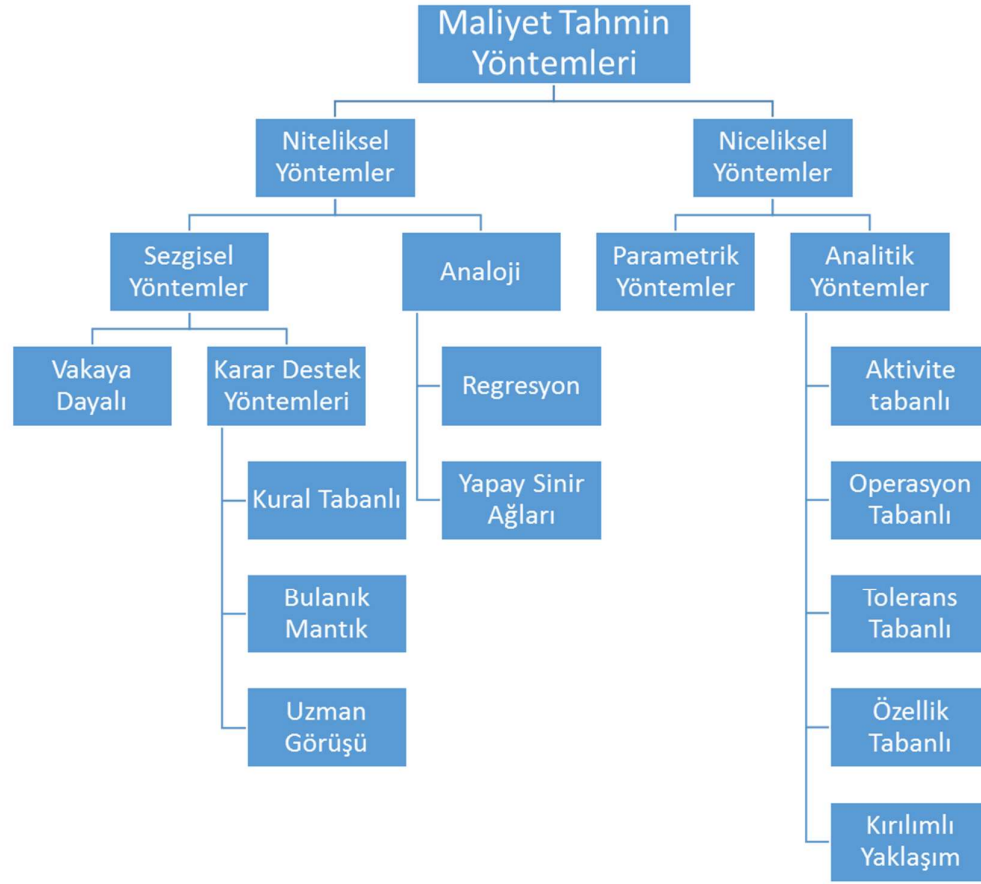
Sonuç olarak, üç ana yöntemin birbirlerine göre olumlu ve olumsuz yönleri mevcuttur. Tahmin için hangi yöntemin kullanılacağı, tahmin edilecek ürüne, tahmin edecek uzmana ve diğ er bazı parametrelere bağlı olarak değişecektir. Üç ana yöntemin özet olarak olumlu ve olumsuz yönleri Tablo 1'de görülebilir.

Yöntem	Olumlu	Olumsuz
Analitik	<ul style="list-style-type: none"> <li>• Neden-sonuç ilişkisi anlaşılır</li> <li>• Detaylı tahmin içerir</li> </ul>	<ul style="list-style-type: none"> <li>• Oluşturulması ve kullanılması zor</li> <li>• Uzmanlık gerektirir</li> </ul>
Analoji	<ul style="list-style-type: none"> <li>• Neden-sonuç ilişkisi anlaşılır</li> <li>• Analitik yöntemle göre daha kolay kullanım</li> </ul>	<ul style="list-style-type: none"> <li>• Uygun geçmiş veri gerekir</li> <li>• Uzmanlık gerektirir</li> </ul>
Parametrik	<ul style="list-style-type: none"> <li>• Kullanım kolaylığı</li> <li>• Kullanımda uzmanlık gerektirmez</li> <li>• Risk ve belirsizliği ortaya koyabilir</li> </ul>	<ul style="list-style-type: none"> <li>• Oluşturması zor olabilir</li> <li>• Neden-sonuç ilişkisi belirsiz olabilir</li> <li>• Teknoloji farklılığında kullanılamayabilir</li> </ul>

Tablo 1 - Maliyet Tahmin Yöntemlerinin Karşılaştırması, Curran ve diğer. (2004)

### 1.3. DİĞER MALİYET TAHMİN YÖNTEMLERİ

Maliyet tahmin literatürü kapsamında en geniş çaplı sınıflandırma çalışması Niazi ve diğer. (2006) tarafından yapılmış olup maliyet tahmin yöntemleri niceliksel ve niteliksel olarak iki ana grup altında toplanmıştır (Şekil 2). Niteliksel tahmin yöntemlerinin temelinde, yeni ürün ile eski ürünün benzerlik ve farklılıklarını ortaya koymak adına yapılan karşılaştırma analizleri vardır. Niceliksel yöntemlerin temelinde ise ürünün tasarımı, çeşitli özellikleri ya da üretim prosesleri detayında analiz edilmesi yer alır. Şekil 3'de bu sınıflandırma detaylı bir şekilde görülebilir. Tablo 2'de ise bu yöntemlerin avantaj ve dezavantajları özetlenmiştir.



Şekil 3 - Maliyet Tahmin Yöntemleri Sınıflandırması, Niazi ve diğer. (2006)

Maliyet Tahmin Yöntemleri			Olumlu Yönler	Olumsuz Yönler	
Niteliksel Yöntemler	Sezgisel Yöntemler	Vakaya dayalı	Yaratıcı yaklaşım	Geçmiş vakalara bağımlı	
		Karar Destek Yöntemleri	Kural Tabanlı	Optimum sonuçlar verebilir	Zaman alıcı
			Bulanık Mantık	Belirsizlik ortamında kullanılabilir	Karmaşık özellikli ürünler için zorlayıcı
			Uzman Görüşü	Daha hızlı ve tutarlıdır	Karmaşık programlama ihtiyacı
	Analoji	Regresyon	Daha basittir	Doğrusallık problemi	
		Yapay Sinir Ağları	Belirsiz ve doğrusal olmayan problemlere uygun	Veriye çok bağımlı	
Niceliksel Yöntemler	Parametrik Yöntemler		Maliyet unsurlarını etkin kullanır	Maliyet unsurları belirlenemediğinde kullanılamaz	
	Analitik Yöntemler	Operasyon Tabanlı	Optimum sonuçlar elde edebilmek için alternatif operasyonlar değerlendirilebilir	Zaman alıcı, detaylı proses bilgisi gerektirir	
		Kısımlı Yaklaşım	Daha basittir	Kaynaklar için detaylı maliyet bilgisi gerektirir	
		Tolerans Tabanlı	Maliyet etkin tasarım toleransları belirlenebilir	Detaylı tasarım bilgisi gerektirir	
		Özellik Tabanlı	Yüksek maliyetli özellikler belirlenebilir	Küçük ve karmaşık özellikler için maliyet belirlemek zordur	
		Aktivite tabanlı	Basit aktivite maliyetleri kullanıldığından basit ve etkindir	Erken tasarım aşamasında aktivite sürelerine ihtiyaç duyar	

Tablo 2 - Maliyet Tahmin Yöntemleri Detaylı Sınıflandırması – Karşılaştırma, Niazi ve diğer. (2006)

Niazi ve diğer. (2006)'in kapsamlı sınıflandırma çalışması oldukça ilgi görmesine rağmen Zhao ve diğer. (2015) konuyu farklı bir açıdan ele almışlardır. Daha önceki maliyet tahmin tekniklerini ve sınıflandırmalarını gözden geçirmiş ve maliyet tahminlerinin dört ana unsurdan oluştuğunu tespit etmişlerdir. Bu unsurlar “tahmin prosesi”, “tahmin modeli”, “hesaplama metodu” ve “yardımcı teknikler” olarak dörde ayrılmıştır. “Tahmin prosesi”, iş dağılım ağacı ve maliyet dağılım

ağacı üzerinde maliyet tahmin aktivitelerinin yönünü belirleyen unsurdur. Alttan-yukarıya prosesi, iş dağılım ağacındaki en küçük parçadan son ürüne doğru ve malzeme ile işçilik maliyet bazından maliyet dağılım ağacındaki son ürün toplam maliyetine doğru bir akış izler. Bunun tersine, yukarıdan-alta prosesi son ürün seviyesinden detay parça seviyesine ve toplam maliyet seviyesinden daha detaylı maliyet unsurlarına doğru bir yol izler. “Tahmin modeli” malzeme, parça geometrisi, üretim prosesi gibi bağımsız değişkenler ile bağımlı değişken yani maliyet arasındaki ilişkileri ortaya koyan temel hipotezi tanımlar. Maliyeti oluşturan unsurları temel alır. Özellik temelli model, aktivite temelli model, proses temelli model örnek olarak verilebilir. Model kapsamındaki maliyet parametrelerinin hangi mantık kapsamında seçileceğini belirler. “Hesaplama metodu” bağımlı ve bağımsız değişkenler arasındaki matematiksel ilişkiyi belirleyen unsurdur. Parametrik, benzetim, bulanık mantık, yapay sinir ağları ve muhasebesel metot olarak gruplanabilir. Son olarak “yardımcı teknikler” tahmin modelinin desteklenmesi ve matematiksel ilişkinin kurulması için kullanılır. Zincir aktivite analizi, regresyon, öğrenme algoritmaları örnek olarak verilebilir. Bahsedilen bu dört unsurun (Tablo 3) her birinden bir eleman seçilerek uygun şekilde bütüncül olarak kullanılmasıyla kapsamlı maliyet tahmin modelleri oluşturulabilir.

Tahmin Prosesi	Tahmin Modeli	Hesaplama Metodu	Yardımcı Teknikler
<ul style="list-style-type: none"> <li>• Alttan-yukarıya</li> <li>• Yukarıdan-alta</li> </ul>	<ul style="list-style-type: none"> <li>• Özellik temelli model</li> <li>• Aktivite temelli model</li> <li>• Vaka temelli model</li> <li>• Proses temelli model</li> <li>• Operasyon temelli model</li> <li>• Genetik Nedensel model</li> <li>• Tolerans Temelli Model</li> <li>• Belirsizlik içerikli model</li> <li>• Mali model</li> </ul>	<ul style="list-style-type: none"> <li>• Parametrik metod</li> <li>• Analoji metodu</li> <li>• Bulanık mantık metodu</li> <li>• Yapay sinir ağları metodu</li> <li>• Muhasebesel metod</li> </ul>	<ul style="list-style-type: none"> <li>• Maliyet tahmin modeline özel teknikler <ul style="list-style-type: none"> <li>- CAD/CAE teknikleri</li> <li>- Aktivite zincir analizi</li> <li>- Zaman etüdü</li> <li>- Proses-ölü zaman analizi</li> <li>- Nedensellik ilişkisi analizi</li> <li>- Maliyet tolerans ilişkisi analizi</li> <li>- Belirsizlik analizi</li> </ul> </li> <li>• Maliyet hesaplama metoduna özel teknikler <ul style="list-style-type: none"> <li>- Regresyon teknikleri</li> <li>- Maliyet zaman analizi teknikleri</li> <li>- Birleştirme teknikleri</li> <li>- Eğitim/öğrenme algoritmaları</li> </ul> </li> <li>• Genel geçer teknikler <ul style="list-style-type: none"> <li>- Veri madenciliği ve bilgi yönetimi</li> <li>- Uzmanlık tabanlı mühendislik teknikleri</li> <li>- İstatistiksel analizler</li> </ul> </li> </ul>

Tablo 3 - Maliyet tahmin metotları sınıflandırması (Zhao ve diğer, 2015)

#### 1.4. HAVACILIK VE UZAY SANAYİNDE MALİYET TAHMİN ÇALIŞMALARI

Havacılık ve uzay sanayiindeki şirketler için maliyet, çizelge ve kalite hedeflerini tutturmak büyük önem taşır. İkinci dünya savaşı ile birlikte savunma sanayinde sistemlerin ve büyük projelerin etkin olarak başarıya ulaşabilmesi için maliyet ve maliyet tahminin önemi anlaşıldı. Takip eden yıllarda sistem tasarımlarında maliyet de önemli bir etken olarak aktif rol oynadı. İkinci dünya savaşı sırasında ağırlığın maliyet tahmini için birincil parametre olduğu düşünülüyordu. Fakat savaş sonrası hızla gelişen teknoloji ile parametrelerin de değiştirilmesi gerektiği anlaşıldı. 1950'li yıllarda büyük endüstriler ürünlerinin maliyet tahmini için çeşitli metotlar geliştirme yollarını aramaya başladılar. Amerika'daki büyük havacılık şirketlerinin öncülüğünde geliştirilen bu metotlar tamamen şirketlere özel olarak tasarlanmıştı. 1970'lerden itibaren maliyet tahmini için PRICE ve SEER gibi ticari yazılımlar ortaya çıkmaya başladı. 1990'larda Amerikan hükümeti ve öncü sanayi kuruluşları işbirliği yaparak ortak bir girişim kapsamında "parametrik tahmin el kitabı"nı hazırladılar. Böylece hükümet ve sanayinin maliyet tahmini anlamında ortak bir dille konuşabilmeleri amaçlanmıştı. Bu aynı zamanda maliyet tahmin metotlarının geliştirilmesi ve iyileştirilmesi için de ortak bir zemin hazırlayarak daha hızlı ve güvenilir tahminlerin yapılabilmesini sağlamıştı. Böylece teklif hazırlama, değerlendirme ve müzakereler kapsamında çevrim süresi ve maliyet yüzde 80 düşürülmüştü.

Havacılık ve uzay sanayinde maliyet tahmin çalışmalarının öneminin giderek daha çok anlaşılması, bu konuda yapılan akademik çalışmalara da yansımıştır. Bu akademik çalışmalar spesifik bir detay parçanın üretim maliyeti tahmini ya da bir hava aracı ya da sisteminin ürün devri yaşam maliyeti tahmini gibi çok geniş bir yelpazede olabilmektedir. Bu bölümün devamında havacılık ve uzay sanayi maliyet tahmin literatürü kapsamında en çok kullanılan ana başlıklar ve bu başlıklara ait örnek çalışmalara değinilecektir.

**Sistem/uçak seviyesi maliyet tahmini:** Finger ve diğer. (2019), hibrit-elektrikli uçakların tasarım ve tedarik maliyetlerini tahmin etmek için bir model

geliştirmişlerdir. Chen ve diğer (2020), literatüre açık 22 adet uçak verisini kullanarak uçak maliyet tahmin modeli geliştirmişlerdir. Derledikleri veriler kapsamında 10 adet bağımsız parametre olup, YSA ve çoklu regresyon çalışmaları yapmışlardır. Çalışmalarında parametre seçimi için temel bileşenler analizi tekniği kullandıklarında modellerin tahmin performanslarının iyileştiğini göstermişlerdir.

**Yaşam devri maliyet tahmini:** Zhaodong ve diğer (2015) havacılık sanayi kapsamında donanım tasarım ve üretim maliyeti tahmini için tasarım evresinin başındaki belirsizliği de dikkate alan bir model geliştirmişlerdir. DAPCA IV modeli kapsamında kullanılan hız, ağırlık gibi parametrelere ek olarak donanım etkinlik endeksini de maliyet tahmin modellerine dahil etmişlerdir. Ön tasarım evresi normalde performans gereksinimleri tarafından şekillendirilir. Fakat kullanıcılar sadece performans gereksinimlerinin karşılandığı tasarımlar değil, aynı zamanda mümkün olan en düşük maliyetli ürün tasarımını isterler. Bu optimizasyonu sağlayabilmek adına Boer ve Stevens (2006), ürün yaşam devri maliyet tahmin modeli geliştirerek ön tasarım evresinde optimizasyon sağlayan bir metod üzerine çalışmışlardır. Bu kapsamda satın alma, bakım, operasyon ve servis gibi ürün yaşam döngüsünün çeşitli evreleri için maliyet tahmini oluşturulabilen bir yazılım geliştirmişlerdir. Banazadeh ve Jafari, (2013), havacılık sistemlerinin yaşam devri maliyet tahminine farklı bir açıdan bakarak karmaşıklık endeksi teorisini, tahmin modelleri kapsamında kullanmışlardır. İlk uçuş yılı, ağırlık, motor tipi, iletişim mesafesi gibi çeşitli tasarım parametrelerini kullanarak karmaşıklık endeksi hesaplamışlar ve tahmin modellerini bu endeks bazında oluşturmuşlardır. Modellerini bir vaka çalışmasında denemişler ve tahmin performanslarının, klasik yöntemlere göre daha başarılı olduğunu belirtmişlerdir.

**Erken tasarım evresi maliyet tahmini:** Erken tasarım evresinde sahip olunan maliyet bilgisi, parçanın üretim maliyeti ve performansı arasındaki dengenin sağlanması açısından tasarımcılar için önemli bir parametredir. Bu problemin çözümü için Lin ve diğer. (2011) konsept tasarım aşamasında üretim maliyeti ve sistem performansının entegre edildiği bir model geliştirmişlerdir. Üretim için tasarım konseptini kullanan bu model kapsamında aktivite temelli maliyet tahmin

yöntemleri kullanmışlardır. Geliştirdikleri modeli vaka çalışması olarak bir helikopter pervanesi projesi kapsamında kullanarak alternatif tasarımların maliyete ve sistem performansına etkilerini göstermişlerdir.

**Üretim maliyet tahmini:** Collopy ve Eames (2001), havacılık sanayinde üretim parçalarının maliyetlerini tahmin edebilmek için bilgi teorisini baz alarak bir metod geliştirmişlerdir. Çalışmaya göre bir parçanın tasarımı kapsamındaki bilgi büyüklüğü arttıkça parçanın üretim maliyeti de artmaktadır. Bu çalışma ile ağırlık ve malzeme bilgilerini temel parametre olarak kullanan modellerinin maliyeti açıklamada bazı durumlarda yetersiz olabildiklerini vurgulamışlardır. Hagnell ve Akermo (2015), havacılık sanayinde kullanılan kompozit parçaların üretim maliyet tahmini için çalışmalar yapmışlardır. Farklı tipte üretim metotları için farklı parametrik modeller üreterek hangi durumda hangi üretim metodunun daha maliyet etkin olduğunu ortaya koymuşlardır. Parçanın boyutsal bilgilerine ek olarak parçanın geometrik karmaşıklığını da modellerinde parametre olarak kullanmışlardır. Konsept tasarım aşamasındaki önemli aşamalardan biri de üretim maliyetlerini tahmin ederek maliyetlerin düşürülmesini sağlamaktır. Lukic ve diğer. (2016) bu amaçla bir parçanın dijital tasarım bilgisini kullanarak parçanın üretim maliyetini tahmin eden bir yazılım geliştirmişlerdir. Bu yazılım kapsamında maliyet tahmini için aktivite ve özellik temelli tahmin modelleri kullanmışlardır.

## 1.5. MOTİVASYON

Firmalar çeşitli sebeplerle ve çeşitli koşullarda maliyet tahmini yapmak durumunda olabilirler. Genel olarak maliyet modelinin oluşturulması için harcanan çaba arttıkça tahmin başarısı da artmakta fakat bunun yanında tahmin süreci de uzamaktadır. Karar verici içinde bulunduğu koşullara uygun olarak maliyet tahmin sürecinin süresi ve tahminin hassasiyet dengesi için bir karar vermek durumunda kalabilir. Bu çalışmada farklı yaklaşımlar kullanılarak maliyet tahmin modelleri oluşturulacak ve bu modellerin tahmin performansları incelenecektir. Bu çalışmanın asıl amacı farklı yaklaşımlarla ortaya konulan modellerin hangisinin en iyi sonuç verdiği değil, hangi yaklaşımın hangi durumda kullanılması gerektiğine dair karar vericiye teknik bir girdi sağlayabilmektir.



Çalışma kapsamında tahmin modelleri oluşturulurken regresyon, yapay sinir ağları ve rastgele orman metotları kullanılarak tahmin performansları karşılaştırılacaktır. Böylece bu metotların farklı durumlarda tahmin performansının nasıl değiştiği ve makine öğrenmesi yöntemlerinin geleneksel yöntemlere kıyasla başarılı olup olmadığı gözlemlenebilecektir.

Literatür incelendiğinde farklı makine öğrenmesi yöntemlerinin performanslarının kıyaslandığı birçok çalışma bulunabilir. Fakat bu çalışmalar genelde tek tip imalat prosesi içeren veri setleri üzerinde yapılmıştır. Bu çalışmanın literatüre katkılarında ilki, farklı imalat prosesleri içeren daha kapsamlı veri setleri için makine öğrenmesi yöntemlerinin performanslarının nasıl değiştiğini ortaya koymak olacaktır. Çalışmanın literatüre ikinci katkısı, öğrenme eğrisinin, maliyet tahmin modelleri ile entegre olarak kullanıldığı ilk çalışma olmasıdır. Maliyet tahmin literatürü incelendiğinde Tablo 4'te belirtilen özelliklere sahip farklı çalışmalar bulunmasına rağmen, bu tez çalışması kapsamında bu özelliklerin tümüne yer verilmektedir.

Referans	Farklı Bilgi Seviyesine Göre Farklı Model	Altın-Üste Maliyetleme	Üretim Maliyeti Tahmin	Farklı Parça Tipleri İçin Tek Model	Makine Öğrenmesi	Veri Dönüşümü	YSA Parametre Optimizasyonu
Deng ve Yeh (2010)		X			X		X
Collopy ve Eames (2001)			X	X			
Molcho ve diğer (2014)	X		X				
Watson ve diğer. (2006)	X		X				
Deng ve Yeh (2012)			X		X	X	
Zhao ve diğer. (2015)		X	X				
Liu ve diğer. (2018)			X		X		X
Özcan ve diğer. (2018)			X		X		X
Castagne ve diğer. (2008)		X	X				
Wang ve Stockton (2001)			X		X		X

Tablo 4 - Maliyet tahmin literatürü çalışmalarında çoklukla kullanılan özellikler

## 2. BÖLÜM

### KULLANILAN METOTLAR

Regresyon analizi, 1970'li yıllardan beri maliyet tahmin çalışmalarında kullanılmaktadır (Deng ve Yeh, 2011). Regresyonun en önemli avantajı, bağımlı ve bağımsız değişken arasındaki ilişkileri doğrusal olarak açıklayabilmesidir. Bir kullanıcı için bunun anlamı, hangi parametrenin sonuç değerlerini nasıl etkilediğini açıkça görebiliyor olmasıdır. Bu da özellikle maliyet tahmin çalışmalarında neden-sonuç ilişkisinin doğru ve açık bir şekilde ortaya konulabilmesini sağlar. Bu olumlu yönlerinin yanında regresyon analizinin olumsuz yönleri de vardır. Garza and Rouhana (1995) regresyon analizinin çok sayıda karmaşık parametreler içeren problemler için uygun olmadığını ortaya koymuşlardır.

Wang (2007), makine öğrenmesini gelişmiş bir maliyet tahmin metodu olarak tanımlamıştır. Makine öğrenmesi metotları kullanılarak, daha az veri ile daha esnek ve daha doğru modeller kurulabileceğini belirtmiştir. Literatürde, makine öğrenmesi yöntemlerinin, regresyon analizine göre daha başarılı tahmin performansı sergilediğini ortaya koyan bir çok çalışma bulunmaktadır, (Deng ve Yeh, 2011; Ciurana ve diğer, 2008; Verlinden ve diğer. (2008); Xie ve diğer, 2017).

Maliyet tahmin literatüründe en çok kullanılan makine öğrenmesi yöntemlerinden birisi de yapay sinir ağlarıdır. Çok başarılı sonuçlar verebilen yapay sinir ağları çeşitli yöntemlerin kıyaslandığı çoğu çalışmada temel yöntem olarak kullanılmaktadır. Bu çalışma kapsamında da YSA, temel başvuru yöntemi olarak kullanılmaktadır. Bu çalışmada kullanılan diğer bir makine öğrenmesi yöntemi de rastgele orman yöntemidir. RO, hem sınıflandırma hem de regresyon amaçlı kullanılabilir. Nispeten yeni olan bu algoritmanın akademik araştırmalar kapsamında kullanımı giderek yaygınlaşmasına rağmen üretim maliyet tahmini çalışmaları kapsamında çok az çalışma bulunmaktadır. RO yöntemi, üretim maliyet tahmin çalışmaları kapsamında YSA ve regresyona göre daha başarılı

sonular verip veremeyeceđini ortaya koyabilmek adına bu tez alıřması kapsamına dahil edilmiřtir.

## **2.1. YAPAY SİNİR AĐLARI**

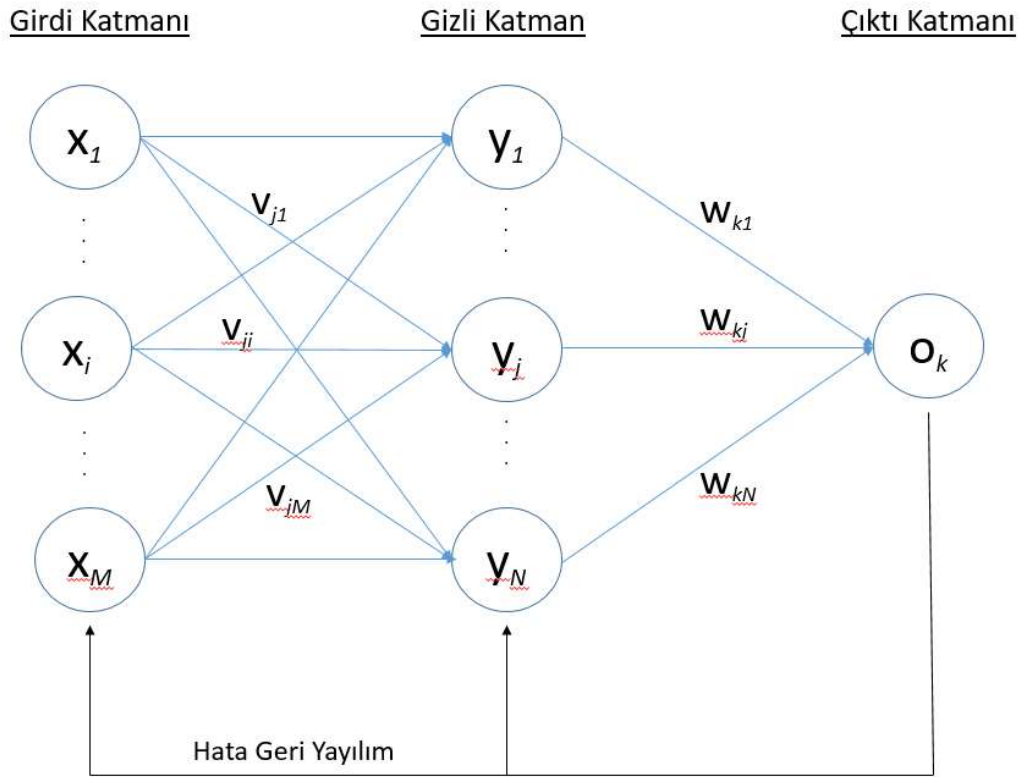
Yapay sinir ađları ile ileri derece dođrusal olmayan fonksiyonlar yardımıyla karmařık problemler iin bařarılı tahmin modelleri kurulabilir ve kullanılacak parametre sayısında herhangi bir kısıtlama yoktur (Deng ve Yeh, 2010). YSA'larının kullanımı iin herhangi bir varsayıma gerek olmaması eřitli problem tiplerinde kullanılmasına olanak sađlar. En uygun ađ yapısının belirlenmesi iin genel geer bir kural yoktur ve YSA'larının sonuları en iyi özüm garantisi vermez.

YSA ok eřitli alanlarda ve eřitli problemlerde uygun bir mimari ile oluřturulduđunda ok bařarılı sonular verebilse de algoritmanın dođası geređi bazı zayıflıkları da vardır. Algoritması geređi, kullanılan girdilerin YSA'nın ürettiđi sonulara ne kadar etki ettiđi bilinemez. Kullanılan gizli nöron sayısı ve bu nöronların ađırlıkları ile dođrusal olmayan bir algoritma oluřturulduđundan YSA bir kara kutu olarak tanımlanmaktadır. Ayrıca en iyi sonuları üretecek bir YSA mimarisi ve parametrelerinin seimi iin genel geer bir kural yoktur. Bu kapsamda özellikle gizli nöron sayısı, ilk ađırlıkların seimi ve kullanılan öđrenme algoritmasının seimi algoritmanın tahmin performansını olduka etkilemektedir. Parametrelerin seimi genellikle deneme-yanılma yöntemi ile yapılmakta olup en uygun parametre kombinasyonunun seimi önemli bir aba gerektirmektedir.

### **2.1.1. Yapay Sinir Ađları Modeli**

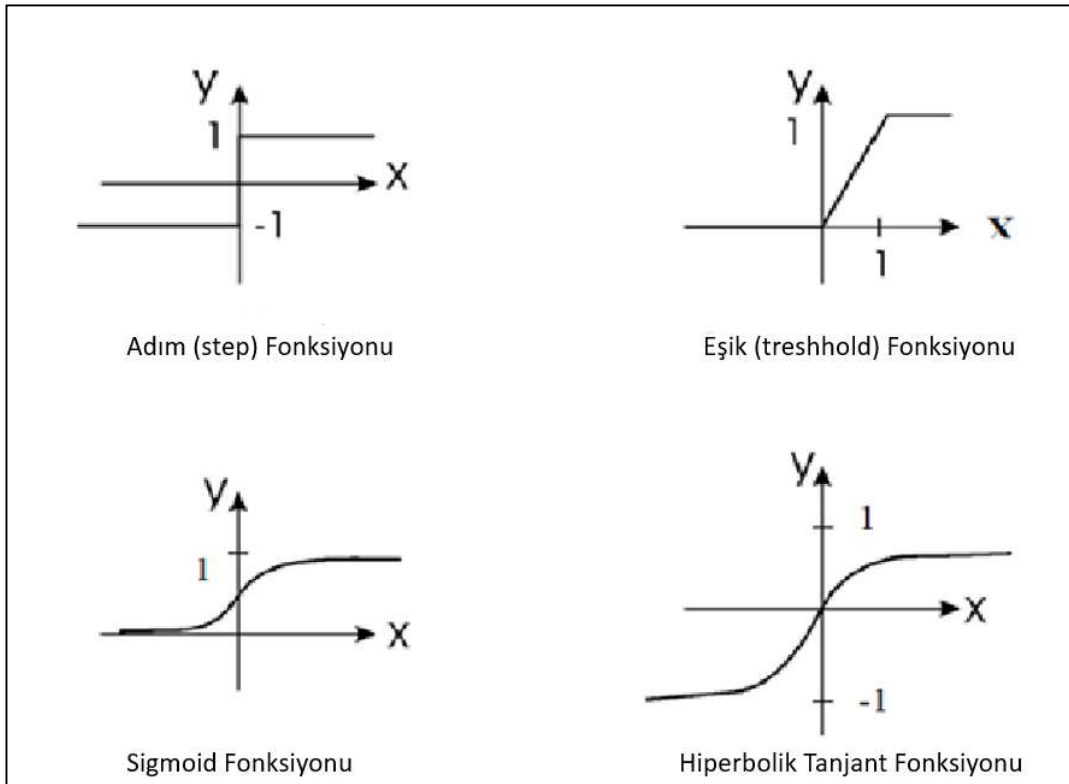
Yapay sinir ađları (YSA), insan beynindeki hücrelerin alıřma prensibini taklit ederek alıřan bir yöntemdir. YSA'nın gemiş verilerden öđrenme kapasitesi vardır. Bir veri setindeki karmařık dođrusal olmayan iliřkileri ortaya koymakta bařarılı bir yöntemdir.

YSA, eğitim verileri ile ağ çıktılarını karşılaştırarak hataları en aza indirmeye çalışır. Bunu yaparken hata geri yayılımı algoritması kullanarak ağ ağırlıklarını sürekli olarak değiştirir ve istenen minimum hata düzeyine ulaşmaya çalışır. Yapay sinir ağlarının ana yapısı Şekil 4'te verilmiştir. Girdi katmanı, ara katman ve çıktı katmanı olmak üzere üç ana katmandan oluşur. Girdi katmanı, dışarıdan girdi alan hücreleri yani parametre değerlerini içerir. Ara katmandaki her bir hücre, girdi katmanından ağırlıklar ile çarpılmış olarak gelen toplamı alır, belirlenmiş transfer fonksiyonundan geçirir ve çıktı katmanına gönderir. Çıktı katmanındaki hücre veya hücreler, ara katmandan ağırlıklar ile çarpılmış olarak gelen toplamı alır, belirlenmiş transfer fonksiyonundan geçirerek ağın çıktısını oluşturur.



Şekil 4 - Yapay sinir ağları yapısı

Transfer fonksiyonu, bir önceki katmandaki gelen değerleri dönüştürerek hücrenin çıktısını belirleyen fonksiyondur. YSA çalışma prensibine göre, hücre çıktı değerlerinin normalize olunması istenir, bu yüzden transfer fonksiyonu çıktıları genellikle  $[0,1]$  ya da  $[-1,1]$  değerleri arasındadır. En çok kullanılan transfer fonksiyonları arasında basamak fonksiyonu, eşik fonksiyonu, sigmoid fonksiyonu ve hiperbolik tanjant fonksiyonu sayılabilir (Şekil 5).



Şekil 5 - YSA Transfer Fonksiyonları

Gizli katmanda olması gereken hücre sayısı için kesinleşmiş genel geçer bir kural yoktur. Optimum hücre sayısı, problemin yapısı, parametre sayısı, verilerin durumu gibi probleme özel olarak değişebilir. Literatürde optimum ara hücre sayısını belirlemek için bir çok çalışma bulunabilir. Bu çalışmalarda genellikle deneme yanılma, çeşitli sezgisel ve meta sezgisel yaklaşımlar denenmektedir.

Ara katmanda gereğinden az ya da fazla hücre kullanmak, tahmin performansını olumsuz etkiler.

Hata geri yayılımı kullanan bir YSA algoritması şu aşamalardan oluşur:

**İleri doğru hesaplama:** Bu aşamada ara katmandaki her  $j$  hücresi, bir  $y_j$  çıktısı üretir. Çıktı katmandaki her  $k$  hücresi, bir  $O_k$  çıktısı üretir.  $v_{ji}$ , girdi katmanı ile ara katmandaki hücreleri birbirine bağlayan ağırlıkları temsil eder.  $w_{kj}$ , ara katman ile çıktı katmanındaki hücreleri birbirine bağlayan ağırlıkları temsil eder.  $f$  ise transfer fonksiyonudur.

$$y_i = f(\text{net}_j = \sum_{i=1}^M v_{ij} x_i)$$

$$O_k = f(\text{net}_k = \sum_{j=1}^N w_{kj} y_j)$$

**Hata hesaplama:** Bu aşamada altta belirtilen denklem kullanılarak hata hesaplaması yapılır. Olması gereken çıktı değeri  $d_k$  ile, ağın çıktı değeri ise  $O_k$  ile gösterilir.

$$E = \frac{1}{2} \sum_{k=1}^L (d_k - O_k)^2$$

**Hata geri yayılımı:** Hata geri yayılım aşamasında arzu edilen ağ çıktı seviyesine yaklaşıncaya kadar gradient descent ağırlıkları revize edilir. Alttaaki denklemler yardımıyla her iterasyonda ağ içerisindeki ağırlıklar yeniden hesaplanır. Belirlenen bir iterasyon sayısından sonra ya da istenen hata değerine ulaşıldığında iterasyonlara son verilir.

$$w_{kj} = w_{kj} + \Delta w_{kj} = w_{kj} - n \frac{\partial E}{\partial w_{kj}}$$

$$v_{ji} = v_{ji} + \Delta v_{ji} = v_{ji} - n \frac{\partial E}{\partial v_{ji}}$$

### 2.1.2. Parametre Seçimi

Yapay sinir ağları modeli kapsamındaki parametrelerin doğru seçimi, modelin performansını doğrudan etkilemektedir. Gizli nöron sayısı, eğitim fonksiyonu, ara nöron transfer fonksiyonu, algoritma performans kriteri gibi çeşitli parametreler literatürdeki çalışmalarda irdelenerek en iyi parametre kombinasyonu seçimi için çeşitli çözüm alternatifleri geliştirilmiş, fakat hiçbir çalışmada genel geçer bir kural ortaya konulamamıştır. Bu çalışmaların çoğunda YSA modelinin performansını en çok etkileyen parametrenin gizli nöron sayısı olduğu belirtilmiştir. Şekeroğlu ve Dimililer, (2020), çalışmalarında optimum gizli nöron sayısını hesaplayan çeşitli çalışmaları derlemişlerdir. Bu çalışmalarda optimum gizli nöron sayısının tamamen modele ve veriye özgün olarak belirlenebileceği ortaya konulmuştur. Literatürdeki bazı çalışmalar ve belirlenen optimum gizli nöron sayıları Tablo 5'te örnek olarak verilmiştir.

Referans	Optimum Gizli Nöron Sayısı
Li ve diğer. (2018)	33
Liu ve diğer. (2017)	20
Kumar ve Singh. (2018)	7
Dimililer ve Zarrouk. (2017)	40
Yuan ve Yu. (2013)	5

Tablo 5 - Optimum gizli nöron sayıları

Bu tez çalışması kapsamında oluşturulan YSA modelleri için de deneysel olarak en uygun gizli nöron sayısı bulunmaya çalışılmıştır. Tez çalışması kapsamında oluşturulan her YSA modeli için 5 farklı gizli nöron sayısına sahip model denenerek en iyi sonucu veren model değerlendirmeye alınmıştır. Gizli nöron sayıları, literatürde genel kabul görmüş kurallardan biri olan alttaki formül (Vujicic ve diğer., 2016) baz alınarak belirlenmiştir. Kullanılan gizli nöron sayıları ve ilgili modellere ait performans sonuçları bölüm 3.1.1’de detaylı olarak belirtilmiştir.

$$\text{Gizli nöron sayısı} = (\text{girdi sayısı} + \text{çıkıktı sayısı}) \times 2 / 3$$

Bu çalışma kapsamındaki YSA modelleri kapsamında kullanılan parametreler aşağıda belirtilmiştir. Gizli nöron sayısının dışındaki parametreler için, MATLAB programının standart varsayılan parametreleri kullanılmıştır.

Gizli nöron sayısı: deneme-yanılma (bölüm 3.1.1’de detaylı olarak belirtilmiştir)

Eğitim fonksiyonu: Levenberg-Marquardt

Ara nöron transfer fonksiyonu: sigmoid

Performans kriteri: Mean squared error



### 2.1.3. Yapay Sinir Ağları Geçmiş Uygulamalar

Temel olarak yapay sinir ağlarının ilk geliştiricisi McCulloch and Pitts (1943) olarak sayılabilir. Hebb (1949), bu konsepti geliştirerek yapay sinir ağlarındaki ağırlıkları değiştirerek bir öğrenmenin olabileceğini ortaya koymuştur. Daha sonra bu konseptin de geliştirilmesi ve bilgisayar teknolojisindeki gelişmelerle birlikte YSA'nın, yapay zeka kapsamında kullanım alanı oldukça yaygınlaşmıştır. Deng ve Yeh (2011), uçak yapılarında kullanılan talaşlı imalat parçalarının üretim maliyeti tahmini için bir YSA geliştirerek sonuçları çoklu regresyon ve SVM tahmin sonuçları ile karşılaştırmışlardır. Verlinden ve diğer. (2008) yaptıkları çalışmada sac metal parçaların üretim maliyeti tahmini için ağırlık, uzunluk, delik sayısı, büküm sayısı gibi parça özelliklerini parametre olarak kullanarak regresyon ve YSA modelleri oluşturmuşlardır. Kurdukları YSA modellerinde deneme yanılma yöntemi ile gizli katmandaki hücre sayısını 13 olarak belirlemişlerdir. Atia ve diğer. (2017) talaşlı imalat parçalarının işlem süresi tahmini üzerinde yaptıkları çalışmada parametrik bir yaklaşımla YSA modeli kurmuşlardır. Parça saatini etkileyen parametreler olarak hacim, sertlik, yüzey pürüzlüğü ve takım malzemesini kullanmışlardır. Modellerinde aktivasyon fonksiyon çeşidini ve hücre sayısını deneme yanılma yoluyla belirlemişlerdir. Liu ve diğer. (2018) internet ara yüzü üzerinden yaklaşık fiyat teklifi verebilen bir sistem geliştirmişlerdir. Parçanın üç boyutlu modeli ara yüz üzerinden sisteme yüklenir, sistem parça kategorisini tanıyarak ilgili algoritmaya göre fiyat teklifi verir. Veri tabanında uygun örneklem yoksa fiyat teklifini literatürdeki en uygun yöntemle göre hazırlar. Tahmin kapsamında kullandıkları YSA modelinde parametre optimizasyonu için genetik algoritma kullanmışlardır. Loyer ve diğer. (2016) jet motorundaki kompresör pervaneleri üretim maliyeti tahmini için YSA da dahil olmak üzere bir çok model kurarak tahmin performanslarını karşılaştırmışlardır. Çoklu doğrusal regresyon, genelleştirilmiş katkı modeli, YSA, destek vektör regresyonu ve gradyan destekli ağaçlar modelleri arasından destek vektör regresyonu ve gradyan destekli ağaçlar modellerinin diğerlerine göre daha başarılı performans gösterdiklerini ortaya koymuşlardır. Wang ve Stockton (2001) torna parçaları maliyet tahmini için kurdukları YSA modelinde gizli katman sayısı ve gizli katmandaki hücre

sayılarını belirlemek taguchi metodu uygulamışlardır. Gizli kat sayısını 1-3 arası, gizli katmandaki hücre sayısını 1-10 arası değiştirerek tahmin performansındaki değişimi grafiksel olarak incelemişlerdir.

## 2.2. RASTGELE ORMAN

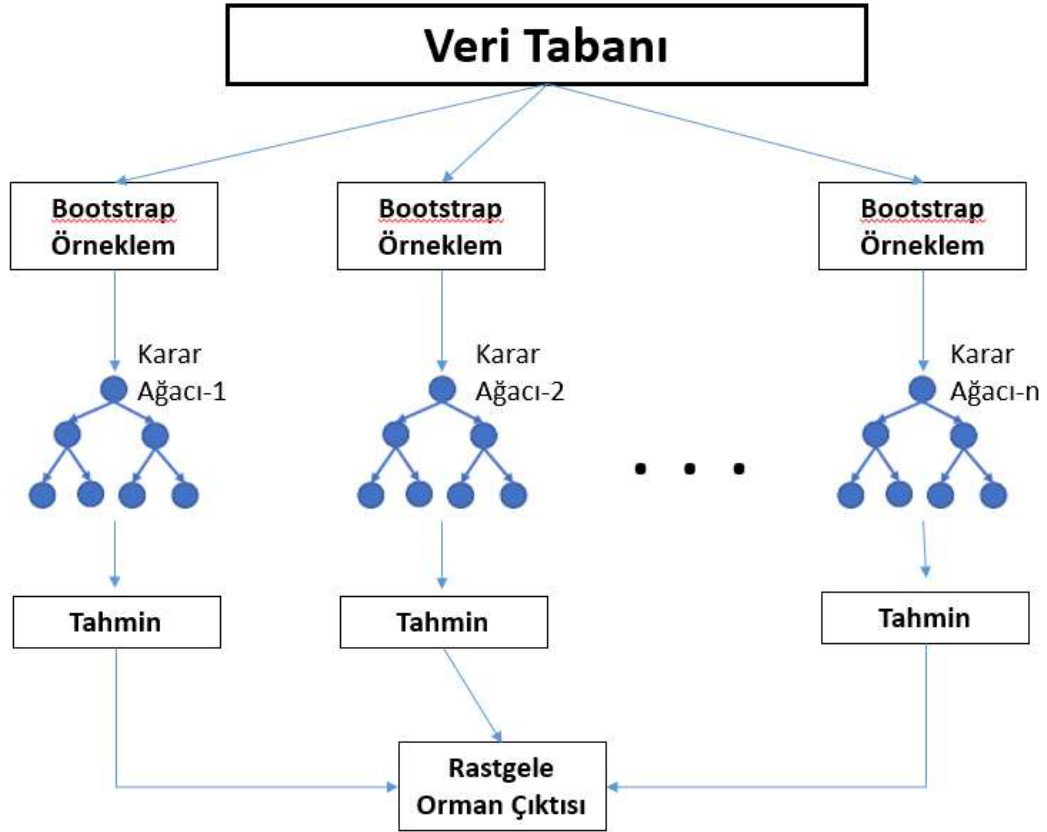
Rastgele orman (RO), Breiman (2001) tarafından oluşturulmuş bir topluluk öğrenmesi yöntemidir. Hem sınıflandırma hem de regresyona uyarlanabilir. RO kapsamında her biri ayrı bir regresyon fonksiyonu olarak çalışan birçok karar ağacı üretilir ve bunların sonuçlarının ortalaması, RO çıktısı olarak kullanılır.

### 2.2.1. Rastgele Orman Modeli

RO algoritmasının temel taşları karar ağaçları olmasına rağmen, karar ağaçlarından daha iyi sonuçlar üretebilir. Algoritma kapsamında birbirinden bağımsız karar ağaçları üretilir. Her ağaç oluşumunda bağımsız değişkenler kümesinin rastgele seçilmiş bir alt kümesi kullanılır. RO, bir topluluk öğrenme yöntemidir ve Breiman (1996) tarafından önerilen torbalama (bagging: bootstrap aggregation) yöntemini kullanır. Torbalama yöntemi sadece karar ağaçları ile değil, diğer regresyon yöntemleri ile de kullanılabilir. Bu yöntem ile varyans düşürülerek tahmin performansının iyileştirilmesi hedeflenir. RO için kullanıcı tarafından iki parametre belirlenir. Bunlardan birincisi düğümlerde kullanılacak değişken sayısı, ikincisi geliştirilecek ağaçların sayısıdır.

1. Veri setinden bootstrap yöntemi kullanılarak örneklem seçilir. Bootstrap örnekleme veri setinin yaklaşık 2/3'ünden oluşur.
2. Bağımsız değişkenler kümesinden  $k$  elemanlı rastgele bir alt küme seçilerek ağaçlar oluşturulur. Yani her ağaç veri setinin belli özellikleriyle eğitilir. Dallara ayrılma işlemine bilgi kazancına bakılarak karar verilir.
3. Oluşturulan ağaçların tahmin ortalaması, RO modelinin tahmin çıktısıdır.

RO hem sınıflandırma hem de regresyon amaçlı kullanılabilir. Nispeten yeni olan bu algoritmanın akademik arařtırmalar kapsamında kullanımı giderek yaygınlařmasına rađmen üretim maliyet tahmini alıřmaları kapsamında ok az alıřma bulunmaktadır.



Őekil 6 – Rastgele Orman Yapısı

### 2.2.2. Parametre Seimi

Rastgele orman algoritması kapsamında bazı parametrelerin seimi, modelin performansını etkileyebilmektedir. Bu parametrelerden bařlıcaları oluřturulacak ađa sayısı, her yapraktaki minimum gzlem sayısı, dđmlerde kullanılan bađımsız deđiřken sayısı olarak sıralanabilir. Fernndez-Delgado ve diđer.,

(2014)'e göre RO, standart varsayılan parametre değerleri ile oldukça başarılı performans sergileyen bir algoritmadır. Probst ve diğer., (2019) çalışmalarında, RO algoritmasındaki parametrelerin model ve veri setine özgün olarak dikkatli bir şekilde ayarlanması ile modelin performansında iyileşme olabildiğini fakat bu iyileşmenin çok az olduğunu göstermişlerdir. Bu nedenle bu tez çalışması kapsamında oluşturulan RO modelleri için MATLAB programındaki altta belirtilen standart parametre değerleri kullanılmıştır.

Oluşturulacak ağaç sayısı: 100

Her yapraktaki minimum gözlem sayısı: 5

Düğümelerde kullanılan bağımsız değişken sayısı: bağımsız değişken sayısının 1/3'ü

### **2.2.3. Rastgele Orman Geçmiş Uygulamalar**

Görece olarak yeni bir makine öğrenme algoritması olan rastgele ormanın popüleritesi gittikçe artmasına rağmen henüz literatürde kullanım alanları sınırlıdır. Çeşitli konularda tahmin ya da sınıflandırma algoritması olarak kullanılsa da maliyet tahmin literatüründe henüz kullanımı çok yaygınlaşmamıştır.

Yapılan maliyet tahmin literatürü araştırmasında yazılım geliştirme sektöründe daha çok çalışmanın yapıldığı gözlemlenmiştir. Mustapha ve Abdelwahed, (2019) yazılım geliştirme efor tahmini çalışmalarında rastgele orman modelini sorgulamışlardır. Çalışmalarında önce ağaç sayısı ve ağaçları geliştirmede kullanılan parametre sayılarının değiştirilmesinin model performansı üzerindeki etkileri araştırılmıştır. Aynı çalışmada daha sonra rastgele orman ile regresyon modellerini COCOMO, ISBGS ve Tuketuku veri tabanları üzerindeki tahmin performanslarını karşılaştırarak rastgele orman modelinin daha başarılı olduğunu ortaya koymuşlardır. Sataphathy ve diğer. (2016) çalışmalarında yazılım geliştirme efor tahmini için rastgele orman yöntemi ile diğer popüler makine öğrenmesi yöntemlerini kıyaslamışlardır. Radyal tabanlı fonksiyon ağları, yapay sinir ağları ve logaritmik doğrusal regresyon yöntemleri performansları ile

kıyaslandığında rastgele orman modelinin daha tutarlı sonuçlar verdiğini göstermişlerdir.

Üretim maliyet tahmin çalışmaları kapsamında yayınlanmış bir makaleye rastlanılamamış olsa da BOEING firması tarafından bazı çalışmalar yapılmakta olup Proje bütçe tahmini, lojistik maliyet tahmini gibi projeler maliyet tahmin konferanslarında sunulmuştur (Mourikas ve Nelson, 2017). Bu anlamda bu çalışmanın üretim maliyet tahmin literatürüne katkıda bulunması beklenmektedir.

### 2.3. DOĞRUSAL REGRESYON

Doğrusal regresyon istatistiğin temel yöntemlerinden biri olup çoğu tahmin çalışmasında performans karşılaştırması için temel oluşturur. Bir parametre ayarlaması gerektirmeyen basit yapısı itibarıyla anlaşılması ve uygulaması kolay olup, tahmin modeli için kullanıcıya iç görü sağlar. Veriye uygun olarak doğrusal bir model oluşturulur ve bu modele göre tahmin yapılır. Bağımsız değişkenler (  $x_1, x_2, \dots, x_n$  ) ile bağımlı değişken (  $y$  ) arasında ilişkiyi açıklar. Doğrusal model, tahmin değerleri ile gerçek değerler arasındaki hata karelerinin toplamını minimize ederek oluşturulur. Çoğu tahmin modeline göre oluşturulması ve kullanılması daha kolay olmasına rağmen değişkenler arasında doğrusal olmayan ilişkilerin varlığında düşük performans gösterebilir.

### 2.4. PERFORMANS DEĞERLENDİRME KRİTERLERİ

Bu çalışma kapsamında oluşturulan modellerin tahmin performanslarını değerlendirmek için üç kriter kullanılmıştır:

**R<sup>2</sup> (korelasyon katsayısı):** Tahmin modelinin veri kümesini ne kadar açıkladığının bir göstergesidir. [0 – 1] aralığında olup değeri 1'e ne kadar yakınsa modelin açıklama gücü o kadar yüksektir.

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

**MSE (hata kareleri ortalaması):** Modelin tahmin hassaslığının bir göstergesidir. Hataların kareleri alındığından tahminlerde büyük sapmalar olduğunda abartılı sonuçlar verebilir.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

**MAPE (ortalama mutlak yüzde hata):** Yüzdesel bir oran olduğundan, özellikle farklı veri setleri üzerindeki tahmin performanslarını kıyaslarken faydalı bir göstergedir.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100$$

## 3. BÖLÜM

### UYGULAMA

Bu bölümde öncelikle problem tanımı ortaya konularak bu çerçevede cevapları aranacak araştırma soruları tanımlanacaktır. Daha sonra veri seti incelenerek çalışma kapsamında gerçekleştirilen uygulamalar detaylı olarak anlatılacaktır. Bu bölümün sonunda ise yapılan uygulamaların sonuçları birbirleri ile kıyaslanarak sonuçlar yorumlanacaktır.

#### 3.1. PROBLEM TANIMI

Maliyet tahmini, özellikle havacılık ve uzay sanayiinde birçok farklı alanda kullanılan ve rekabet koşullarının zorlaşmasıyla önemi de giderek artan bir kavramdır. Yeni iş alımları için rekabetçi teklif vermek zorunda olan firmalar, maliyet tahminine giderek daha fazla önem vermektedir. Firmalar maliyet tahminlerinde hem yeterince doğru hem de yeterince hızlı olmalıdırlar. Bazı ihalelerde firmalar ROM (yaklaşık) fiyat talebinde bulunabilir ve bu taleplerine çok kısa bir süre içinde cevap bekleyebilirler. Bazı durumlarda ise firmalar detaylı fiyat taleplerine rağmen çok kısa süre içinde cevap talep edebilirler. Bazı durumlarda teklif veren firmanın elinde belirli bir süre içinde cevap verilmesi gereken o kadar çok ihale olur ki, mevcut işgücü her ihaleye gereken zamanı ayıramayabilir. Bu gibi durumlarda teklif veren firmalar kabiliyetlerindeki farklı tahmin yöntemlerinden duruma en uygununu seçmek zorundadır. Bu seçimi yapabilmek için ise firmanın hangi yöntemin ne kadar süreceğini ve tahmin performansının hangi düzeyde olduğunu bilmesi gerekmektedir.

Tanımlanan bu problem çerçevesinde tez kapsamında üç araştırma sorusuna cevap aranmıştır:

1. Araştırma Sorusu: Bir teklif verme sürecinde içinde bulunulan koşullara göre tahmin süresi ile tahmin hassasiyeti dengesini dikkate alarak nasıl bir tahmin yaklaşımı seçmek gerekir?

Birinci araştırma sorusuna cevap verebilmek için üç farklı yaklaşım kapsamında tahmin modelleri geliştirilmiştir. Birinci yaklaşımda son ürün bazında toplam üretim saati tahmini için modeller oluşturulacaktır. Bu yaklaşımda model kapsamındaki veriler son ürün bazında olduğundan veri ve parametre sayısı kısıtlı kalmakta, bu da tahmin sürecini kısaltırken tahmin hassasiyetini de azaltmaktadır. Bu yaklaşımdaki amaç az veri ile hızlı bir tahminde bulunmak gerektiğinde oluşturulacak modellerin tahmin performansını ortaya koyabilmektir. Çalışma kapsamında imalat ve montaj yöntemleri açısından benzer nitelikte olan 12 adet son ürüne ait veriler kullanılarak tahmin modelleri oluşturulacaktır.

İkinci ve üçüncü yaklaşımda son ürünlerin ürün ağacındaki detay parçaların üretim saatlerinin tahmini için modeller oluşturulacaktır. Detay parça bazında yapılan tahminler toplanarak son ürünlerin toplam üretim saatleri tahmin edilecektir. Çalışmada 12 adet son ürünün ürün ağacındaki toplam 1099 detay parçaya ait veriler kullanılarak tahmin modelleri oluşturulacaktır.

İkinci yaklaşım kapsamında detay parçalara ait verilerin tümünü içeren tek bir veri tabanı (birleşik veri tabanı) oluşturularak farklı proses tipleri için Tablo 9 ve Tablo 11'de belirtilen bağımsız değişkenler kullanılarak ortak tek bir model oluşturulacaktır.

Üçüncü yaklaşımda, detay parça veri tabanı, Tablo 9'da belirtilen farklı imalat proseslerine göre 8 alt veri tabanına bölünerek her bir proses alt veri tabanı için farklı tahmin modelleri (Proses veri tabanları modelleri) oluşturulacaktır.

Proses veri tabanları ile oluşturulan modellerin tahmin gücünün birleşik veri tabanı modellerine göre daha iyi olması beklenirken, 8 farklı model kurulması gerektiği için daha uzun zaman alacaktır.

Çalışma kapsamında farklı yaklaşımlar altında oluşturulan modellerin okuyucu tarafından daha rahat takip edilebilmesi amacıyla yaklaşımlar için bir kodlama sistemi oluşturulmuştur (Tablo 6).



Tahmin Yaklaşımı Sırası	Tahmin Yaklaşımı Adı	Tahmin Seviyesi	Tahmin Yaklaşımı Kodu
Tahmin Yaklaşımı 1	Son Ürün	Son Ürün	TY1-S-SÜ
Tahmin Yaklaşımı 2	Birleşik Veri Tabanı	Detay Parça	TY2-D-BV
Tahmin Yaklaşımı 3	Proses veri tabanları	Detay Parça	TY3-D-PV

Tablo 6 - Tahmin Yaklaşımları Kodları

2. Araştırma Sorusu: Makine öğrenmesi yöntemlerinin tahmin performansı geleneksel yöntemlere göre üstün müdür?

Bu soruya cevap bulabilmek için her tahmin yaklaşımı kapsamında ayrı ayrı regresyon, YSA ve RO modelleri oluşturularak tahmin performansları birbirleri ile kıyaslanacaktır. Günümüzde makine öğrenme yöntemleri, tahmin modellerinde sıklıkla kullanılmakta ve başarılı sonuçlar vermektedir. Regresyon analizi, bağımlı ve bağımsız değişkenler arasındaki ilişkileri açık bir şekilde ortaya koyabilmesine rağmen bazı durumlarda açıklama gücü yeterince iyi olmayabilmektedir. YSA ve RO gibi makine öğrenmesi yöntemleri değişkenler arasındaki karmaşık ilişkileri başarılı bir şekilde ortaya koyabilmelerine rağmen doğrusal olmayan yapılarından dolayı her bir değişkenin sonucu nasıl etkilediğini açıklayamazlar. Bir çok çalışmada başarılı sonuçlar veren YSA, bu çalışmada da temel başvuru yöntemi olarak kullanılmaktadır. Nispeten yeni olan RO yönteminin akademik araştırmalar kapsamında kullanımı giderek yaygınlaşmasına rağmen üretim maliyet tahmini çalışmaları kapsamında çok az çalışma bulunmaktadır. Bu çalışma kapsamında RO yönteminin regresyon ve YSA yöntemlerine göre başarılı olup olmadığı araştırılacaktır.

3. Araştırma Sorusu: Bir işin tekrarlanması sonucu oluşan öğrenme eğrisine göre üretim saatlerine veri dönüşümü uygulanmasının model performansına etkisi var mıdır?

Üçüncü araştırma sorusuna cevap bulabilmek için tahmin modellerindeki bağımlı değişken olan üretim sürelerinin, öğrenme eğrisine göre veri dönüşümü yapılarak

modellendiğinde, tahmin modeli sonuçlarının iyileşip iyileşmediği araştırılacaktır. Bu amaçla birleşik veri tabanı yaklaşımı kapsamında oluşturulan modeller, bağımlı değişkenin öğrenme eğrisine göre dönüştürülmüş haliyle tekrarlanacaktır. Bu şekilde oluşturulan regresyon, YSA ve RO modellerinin performans sonuçları, öğrenme eğrisi kullanılmayan model performansları ile kıyaslanacaktır.

Her model için ilgili verilerin yaklaşık %75'i eğitim verisi olarak kullanılarak modeller oluşturulacak, verilerin kalanı ise oluşturulan modellerin doğruluğunu ve geçerliliğini sınamak için kullanılacaktır. Daha sonra modellerden elde edilen tahmin değerleri ile gerçek değerler karşılaştırılarak modellerin performansları değerlendirilecektir.

Tez kapsamında oluşturulan regresyon modelleri MS Excel ile oluşturulmuş olup, yapay sinir ağları ve rastgele orman modelleri MATLAB programı aracılığıyla oluşturulmuştur.

Oluşturulan modellerin performanslarını ölçmek ve birbirleriyle kıyaslayabilmek için performans parametreleri olarak  $R^2$ , MSE (mean squared error) ve MAPE (mean absolute percentage error) kullanılmıştır.

Bir son ürünün işçilik maliyeti detay parçaların üretimi ve montaj olmak üzere iki ana kalemden oluşur. Bu çalışma kapsamında son ürünlerin montaj işçiliği hariç toplam üretim saatleri, yani son ürünleri oluşturan detay parçaların toplam üretim saatleri için tahmin modelleri oluşturulacaktır. Son ürün üretim saati anılan yerlerde bu kapsam düşünülmelidir. Havacılık sanayinde montaj işçiliği, kullanılan bağlayıcıların çeşidi ve miktar parametrelerine bağlı olup basit matematiksel hesaplarla kolaylıkla tahmin edilebilmektedir. Bu nedenle montaj işçiliği tahmini bu tez çalışması kapsamının dışında tutulmuştur.

### **3.1.1. YSA Gizli Nöron Sayılarının belirlenmesi**

Yapay sinir ağları modelleri için gizli nöron sayısının doğru seçimi modelin performansını doğrudan etkilemektedir. Bu konu literatürdeki çeşitli çalışmalarda

irdelenmiş olup hiçbir çalışmada genel geçer bir kural ortaya konulamamış, optimum gizli nöron sayısının modele ve veri setine bağlı olarak değiştiği belirtilmiştir (Şekeroğlu ve Dimililer, 2020).

Bu tez çalışması kapsamında oluşturulan YSA modelleri için de deneysel olarak en uygun gizli nöron sayısı bulunmaya çalışılmıştır. Tez çalışması kapsamında oluşturulan her YSA modeli için 5 farklı gizli nöron sayısına sahip model denenerek en iyi sonucu veren model değerlendirmeye alınmıştır. Gizli nöron sayıları, literatürde genel kabul görmüş kurallardan biri olan alttaki formül (Vujicic ve diğer., 2016) baz alınarak belirlenmiştir.

$$\text{Gizli nöron sayısı} = (\text{girdi sayısı} + \text{çıkıktı sayısı}) \times 2 / 3$$

Bu formüle göre hesaplanan gizli nöron sayıları Tablo 7'de belirtilmiştir. Buradan elde edilen gizli nöron sayısı baz alınarak Tablo 8'de belirtilen gizli nöron sayılarına göre 5 farklı model denenerek en iyi sonucu veren model değerlendirmeye alınmıştır.

	Parametre Sayısı	Genel Kurala Göre Gizli Nöron Sayısı
Son Ürün	2	2
Birleşik Veri Tabanı	16	11
Büküm Prosesi	2	2
Ekstrüzyon-Basit Prosesi	2	2
Ekstrüzyon-Çekme Prosesi	2	2
Çekme Prosesi	1	1
Hidroform-Büküm Prosesi	2	2
Talaşlı İmalat Prosesi	2	2
roll prosesi	1	1
basit kesme prosesi	2	2

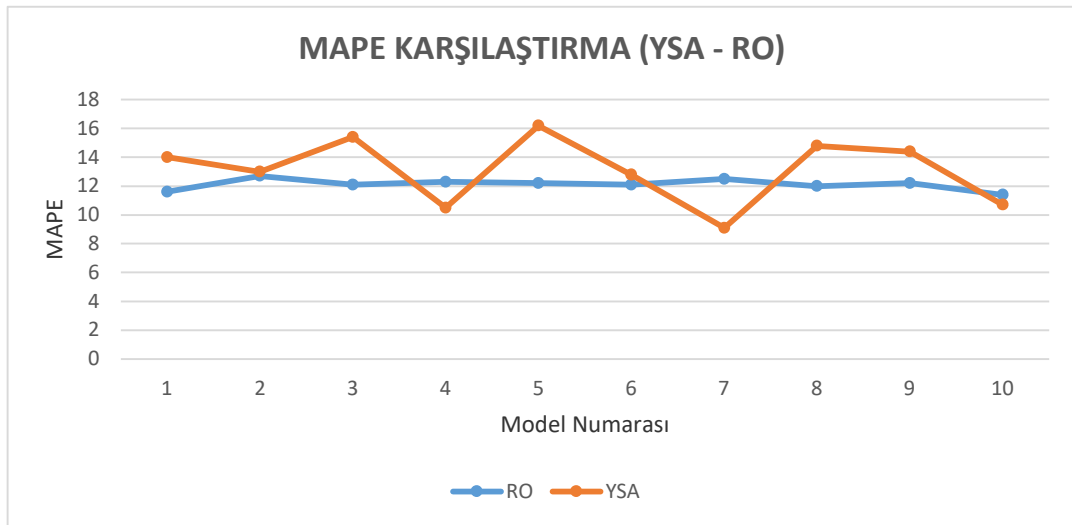
Tablo 7 - Parametre sayısına göre gizli nöron sayısı

		Gizli Nöron Sayısı				
		1. deneme	2. deneme	3. deneme	4. deneme	5. deneme
Son Ürün	Gizli Nöron Sayısı	2	6	10	14	18
	MAPE	9,1	11	11	11	14
Birleşik Veri Tabanı	Gizli Nöron Sayısı	3	7	11	15	19
	MAPE	82	74	86	145	145
Birleşik Veri Tabanı - Öğrenme Eğrisi	Gizli Nöron Sayısı	3	7	11	15	19
	MAPE	63	129	100	184	161
Büküm Prosesi	Gizli Nöron Sayısı	2	6	10	14	18
	MAPE	16	19	35	29	19
Ekstrüzyon-Basit Prosesi	Gizli Nöron Sayısı	2	6	10	14	18
	MAPE	20	21	28	32	51
Ekstrüzyon-Çekme Prosesi	Gizli Nöron Sayısı	2	6	10	14	18
	MAPE	2,9	2,2	1,9	1,5	2
Çekme Prosesi	Gizli Nöron Sayısı	2	6	10	14	18
	MAPE	9,2	8,1	8,7	12	7,4
Hidroform-Büküm Prosesi	Gizli Nöron Sayısı	2	6	10	14	18
	MAPE	49	44	55	59	61
Talaşlı İmalat Prosesi	Gizli Nöron Sayısı	2	6	10	14	18
	MAPE	21	20	17	22	19
Roll Prosesi	Gizli Nöron Sayısı	2	6	10	14	18
	MAPE	100	18	16,9	17,2	20
Basit Kesme Prosesi	Gizli Nöron Sayısı	2	6	10	14	18
	MAPE	11,1	11,3	10,8	11,2	12

Tablo 8 - Gizli nöron sayısı ve MAPE ilişkisi

### 3.1.2. Model Performans Sonuçlarının Kararlılığı

Algoritmaları gereği YSA ve RO, aynı eğitim verisini kullansalar bile farklı tahmin çıktıları üreten farklı modeller oluştururlar. Bunun nedeni algoritmaları kapsamında rastgeleselliği kullanmalarıdır. Örneğin YSA algoritmasında ilk ağırlıklar rastgele olarak belirlenir. RO algoritmasında ise her ağaç oluşumunda bağımsız değişkenler kümesinin rastgele seçilmiş bir alt kümesi kullanılır. Bunun sonucu olarak bu algoritmalar her eğitildiklerinde farklı tahmin performansları sergileyen farklı modeller oluştururlar. Bu çalışma kapsamında tahmin performansı değişkenliğini ortaya koymak için son ürün tahmin modeli için YSA ve RO algoritmaları 10'ar kez çalıştırılmış ve test verileri için MAPE değerleri Şekil 7'de karşılaştırılmıştır.



Şekil 7 - Model Performans Sonuçları Karşılaştırma

Grafikte görüldüğü üzere, oluşturulan 10 farklı YSA modelinin tahmin performansları birbirlerine göre farklılık gösterirken, RO modellerinin tahmin performansı oldukça stabildir. Bir algoritmanın tahmin performansı değerlendirilirken bu oynaklık akılda tutulmakla birlikte, algoritmanın en iyi performansı değerlendirmeye alınmalıdır. Bu nedenle bu çalışma kapsamında YSA algoritması kullanılan her aşamada 10 farklı YSA modeli üretilmiş ve bu 10

modelden en iyi tahmin performansı sonucu veren model değerlendirmeye alınmıştır.

### 3.2. VERİ SETİ

Çalışma kapsamında bir havacılık ve uzay sanayi firması tarafından üretimi yapılan uçak alt-montaj parçalarına ait gerçek veriler kullanılmıştır. Verilere ait özet bilgileri aşağıda yer almakta olup firmanın gizlilik prensibine istinaden daha fazla detay tez kapsamında paylaşılmamaktadır.

Veri seti, benzer nitelikte 12 adet son ürünü (satışı yapılan) oluşturan toplam 1099 adet detay parça bilgilerini kapsamaktadır. Bu çalışma kapsamında oluşturulan modeller, son ürünlerin ve son ürünleri oluşturan detay parçaların işçilik sürelerini tahmin etmek üzere oluşturulmuştur. Maliyet tahmin literatürü incelendiğinde, bazı çalışmalarda toplam maliyet (Watson ve diğerleri, 2006), bazı çalışmalarda ise işçilik süreleri (Deng ve Yeh, 2011) için tahmin modelleri oluşturdukları görülebilir.

$$\Sigma \text{ maliyet} = \Sigma (\text{işçilik maliyeti}) + \Sigma (\text{malzeme maliyeti}) + \Sigma (\text{diğer maliyetler})$$

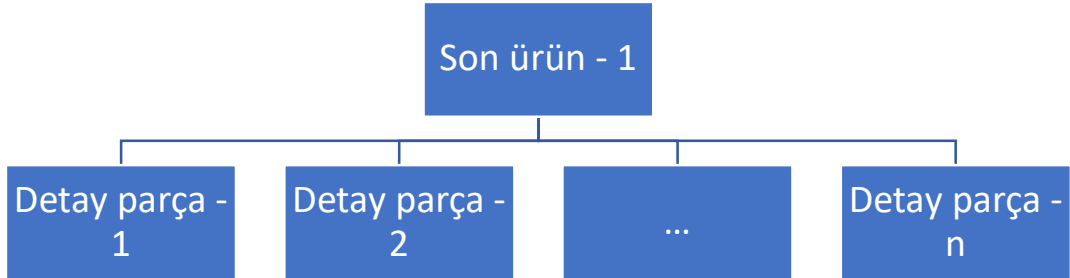
Toplam maliyeti oluşturan üç ana unsur yukarıdaki formülde belirtilmiştir. Bu üç unsurun toplam maliyet içindeki yüzdeleri işin tipine göre değişim göstermekle birlikte ortalama olarak sırasıyla %45, %45, %10 olarak varsayılabilir. Malzeme maliyeti kullanılacak malzemenin tipine ve boyutlarına göre piyasa fiyatları kullanılarak kolaylıkla tahmin edilebileceğinden asıl önemli olan işçilik maliyetinin doğru tahminidir. Diğer maliyetlerin hem toplam maliyet içindeki payı az olduğundan hem de asıl olarak genel giderleri içerdiğinden ürün bazında tahmini yapılmayıp firma için toplamı önemlidir.

Montaj yapılan ürünler için de detay parça üretim ve montaj işçiliği olarak iki işçilik unsuru vardır. Detay parçaların üretim yöntemi ve montaj planlama ve

teknolojisine göre deęişmekle birlikte toplam işçilik maliyeti detay üretim için %70 ve montaj için %30 olarak varsayılabılır. Montaj işçilięi tahmini, kullanılacak bağlayıcı sayı ve tiplerine göre ve kullanılacak montaj teknolojisine göre (otomatik, yarı otomatik, elle montaj...) kolaylıkla hesaplanabilir. Burada da asıl önemli olan detay parçaların işçilik sürelerini tahmin edebilmektir. Dolayısıyla tez kapsamında son ürünler kapsamındaki detay parçaların işçilik sürelerinin tahmini ele alınacaktır.

$$\Sigma (\text{işçilik maliyeti}) = \Sigma (\text{detay parça üretim maliyeti}) + \Sigma (\text{montaj maliyeti})$$

Şekil 8'de örnek olarak "Son Ürün -1" için basitleştirilmiş bir ürün ağacı yapısı gösterilmiştir.



Şekil 8 - Son Ürün 1 için Basitleştirilmiş Örnek Ürün Ağacı

Her son ürün, farklı tipte ve sayıda detay parçaların birleşmesinden oluşur. Son ürün bazında hangi proses tipinden kaçar adet detay parça olduğu bilgisi Tablo 9'da özetlenmiştir.

<b>Proses Tiplerine Göre Detay Parça Sayıları</b>									
	<b>Büküm</b>	<b>Extrüzyon - Basit</b>	<b>Extrüzyon - Çekme</b>	<b>Hidroform Büküm</b>	<b>Talaşlı İmalat</b>	<b>Roll</b>	<b>Basit Kesme</b>	<b>Çekme</b>	<b>Toplam Parça Sayısı</b>
<b>Son Ürün -1</b>	12	15	1	28	5	2	4	2	<b>69</b>
<b>Son Ürün -2</b>	19	15	1	33	10	2	7	2	<b>89</b>
<b>Son Ürün -3</b>	80	39		113	52	32	4	2	<b>322</b>
<b>Son Ürün -4</b>	12	15	1	28	5	2	4	2	<b>69</b>
<b>Son Ürün -5</b>	22	16	1	33	13	2	1	2	<b>90</b>
<b>Son Ürün -6</b>	15	11		58	21	2	8	1	<b>116</b>
<b>Son Ürün -7</b>	56	18			2	3	1	1	<b>81</b>
<b>Son Ürün -8</b>	57	11	1		1		1	2	<b>73</b>
<b>Son Ürün -9</b>	37	6			1		1	2	<b>47</b>
<b>Son Ürün -10</b>	26	23				3		1	<b>53</b>
<b>Son Ürün -11</b>	25	13	2		1	1	1	2	<b>45</b>
<b>Son Ürün -12</b>	25	14	1		1	1	1	2	<b>45</b>
<b>TOPLAM</b>	<b>386</b>	<b>196</b>	<b>8</b>	<b>293</b>	<b>112</b>	<b>50</b>	<b>33</b>	<b>21</b>	<b>1099</b>

Tablo 9 - Son Ürünlere Ait Detay Parça Özeti



Proses tanımları kısaca aşağıda belirtilmiştir:

**Büküm:** Sac metalin düz bir eksen çevresinde gerilerek kalıcı bir büküm oluşturması

**Extrüzyon – basit:** Extrüzyon malzemenin form verilmeden ve et kalınlığı azaltılmadan işlendiği parçalardır. Malzemenin boyunun kısaltılması, delik delinmesi ve extrüzyon bacalarında düz kesim yapılması işlemlerini kapsar.

**Extrüzyon – çekme:** Extrüzyon malzemenin bir kalıp yardımıyla gerdirilerek veya çekilerek form verilmesidir.

**Talaşlı imalat:** Kesilen stok malzemenin istenilen ölçü, şekil, geometrik tolerans ve yüzey kalitesinde iş parçası üretimi için kesici takım ve mekanik enerji kullanılarak yapılan talaş kaldırma işlemidir.

**Roll:** Sac metal levhanın silindirler arasından geçirilmesi ile levhaya yuvarlak bir form verilmesidir.

**Basit kesme:** Sac metal levhanın iki eksenli olarak parça geometrisinde kesilmesidir.

**Çekme:** Sac metal levhanın kenarlardan çekilmek suretiyle bir kalıp etrafında gerdirilerek form verilmesidir.

Parametrelere ilişkin tanımlamalar aşağıda verilmiştir:

**İşemri başına ortalama miktar:** Detay parçaların atölye içindeki üretimleri, işemirleri ile takip edilir. İşemirlerinde parçaların hangi tezgahlarda üretileceği ve üretim parti büyüklüğü yer alır. Bu parametre bir parça için ortalama üretim parti büyüklüğüdür. Bir parça için, üretim parti büyüklüğü genelde benzer olmakla birlikte aynı olmak zorunda değildir. Üretim parti büyüklüğü, parça üretimi esnasındaki takım hazırlama gibi tekrar etmeyen işler için harcanan zamanın parçanın birim işçilik süresine dağıtılmasından dolayı, parçanın ortalama üretim zamanı üzerinde etkisi vardır.

**Aylık ihtiyaç miktarı:** Son ürünün aylık ortalama ihtiyaç miktarı ile detay parçanın bir adet son ürünün ürün ağacındaki miktarının çarpımıdır.

**Toplam üretilen miktar:** Detay parçanın firma içinde o güne kadar toplam kaç adet üretildiği bilgisidir.

**Alan:** Parçanın üretildiği ham malzemenin uzunluk ve genişliğinin çarpımıdır.

**Hacim:** Parçanın üretildiği ham malzemenin hacmidir. (Uzunluk X genişlik X kalınlık)

Veri setini oluşturan detay parçalara ait parametreler ve ilgili özet istatistik bilgisi Tablo 11'de verilmiştir.

Detay parçaların üretim saatlerine göre yüzdesel dağılımı Tablo 10'da belirtilmiştir.

<b>Değer Aralığı (Üretim Saati)</b>	<b>Frekans (Parça Sayısı)</b>	<b>Yüzdesel Oran</b>
0 - 0,1	426	39%
0,1 - 0,5	450	41%
0,5 - 1	109	10%
1 - 2	37	3%
2 - 3	27	2%
3 - 4	11	1%
4 - 5	4	0%
5- 10	27	2%
10- 20	8	1%
<b>TOPLAM</b>	<b>1099</b>	<b>100%</b>

Tablo 10 - Üretim Saatleri Frekans Değerleri

	<b>İş emri Başına Ortalama Miktar</b>	<b>Aylık İhtiyaç Miktarı</b>	<b>Toplam Üretilen Miktar</b>	<b>Alan (mm<sup>2</sup>)</b>	<b>Hacim (mm<sup>3</sup>)</b>	<b>Kalınlık (mm)</b>	<b>Uzunluk (mm)</b>	<b>Genişlik (mm)</b>
<b>Ortalama</b>	40	54	3429	166357	1197166	5	633	149
<b>Standart Hata</b>	1	1	80	28895	209409	0	26	8
<b>Ortanca</b>	33	65	4128	22200	31395	1	225	115
<b>Mod</b>	60	65	5000	0	0	1	200	0
<b>Standart Sapma</b>	29	30	2645	957919	6942162	14	867	257
<b>Örnek Varyans</b>	819	917	6993634	9,18E+11	4,82E+13	185	751698	66051
<b>Basıklık</b>	14	21	16	78	79	18	6	40
<b>Çarpıklık</b>	3	3	3	9	9	4	2	6
<b>En Küçük</b>	1	20	151	0	0	0	19	0
<b>En Büyük</b>	240	325	25325	10080000	82500000	110	4950	2390

Tablo 11 - Detay Parçalara Ait Tanımlayıcı Özet İstatistik Bilgiler

### 3.3. SON ÜRÜN TAHMİN MODELİ (TY1-S-SÜ)

Bu bölümde son ürünün üretim saatinin belirlenmesinde etkili olan parametreler kullanılarak son ürün bazında tahmin modelleri oluşturulmuştur. Model kapsamında bağımsız değişkenler olarak son ürün bazında “alan” ve “toplam parça sayısı”, bağımlı değişken olarak da son ürünün toplam üretim saati kullanılmıştır. Proses tiplerine göre detay parça sayıları, aylık ihtiyaç miktarı ve toplam üretilen miktar parametrelerinin de bağımsız değişken olarak kullanıldığı modeller de denenmiş olup, anlamlı sonuçlar alınamamıştır.

SON ÜRÜN KODU	ALAN (m2)	TOPLAM PARÇA SAYISI
Son Ürün -1	1,08	69
Son Ürün -2	1,07	89
Son Ürün -3	1,61	322
Son Ürün -4	1,08	69
Son Ürün -5	0,98	90
Son Ürün -6	0,85	116
Son Ürün -7	0,80	81
Son Ürün -8	1,01	73
Son Ürün -9	0,75	47
Son Ürün -10	0,80	53
Son Ürün -11	0,82	45
Son Ürün -12	0,82	45

Tablo 12 - Son Ürün Verileri

Veri sayısının azlığından dolayı daha sağlıklı tahmin modelleri oluşturabilmek için bootstrap yöntemi ile veriler çoğaltılmıştır. Bu yöntem ile yerine koymalı olarak rastgele seçilen 2 son ürün içeren 100 örneklem üretilmiştir. Bağımlı ve bağımsız değişken verileri olarak her örneklemden 2 son ürüne ait verinin ortalamaları kullanılmıştır.

Tablo 13'te bootstrap yöntemi ile rastgele oluşturulan örneklemelerin bir kısmı görülmektedir.

Örneklem Numarası	1. Son ürün numarası	2. Son ürün numarası
1	10	9
2	4	11
3	12	10
4	6	11
5	4	6
6	1	2
7	4	5
8	7	12
9	7	3
...	...	...
100	12	5

Tablo 13 - Bootstrap Örneklem

Bootstrap yönteminde bir veri setinden rastgele olarak yerine koymalı şekilde veriler seçilerek yeni bir veri seti yaratılır. Makine öğrenmesi yöntemleri kapsamında veri az olduğu ya da dağınık olduğu durumlarda aşırı uyumu engellemek için sıklıkla başvuru olan bir yöntemdir. 12 adet son ürün rastgele bir şekilde ikiye bölünmüş ve Tablo 14'te belirtildiği şekilde 8 tanesi, bootstrap yöntemi ile üretilmiş olan 100 adet veri ile birlikte eğitim verisi olarak ve kalan 4 tanesi test verisi olarak kullanılmıştır.

Eğitim Verileri	Test Verileri
Son Ürün -1	Son Ürün -2
Son Ürün -3	Son Ürün -5
Son Ürün -4	Son Ürün -9
Son Ürün -6	Son Ürün -12
Son Ürün -7	
Son Ürün -8	
Son Ürün -10	
Son Ürün -11	
Bootstrap örnek -1	
Bootstrap örnek -2	
Bootstrap örnek -3	
---	
Bootstrap örnek -100	

Tablo 14 - Eğitim ve Test Verileri

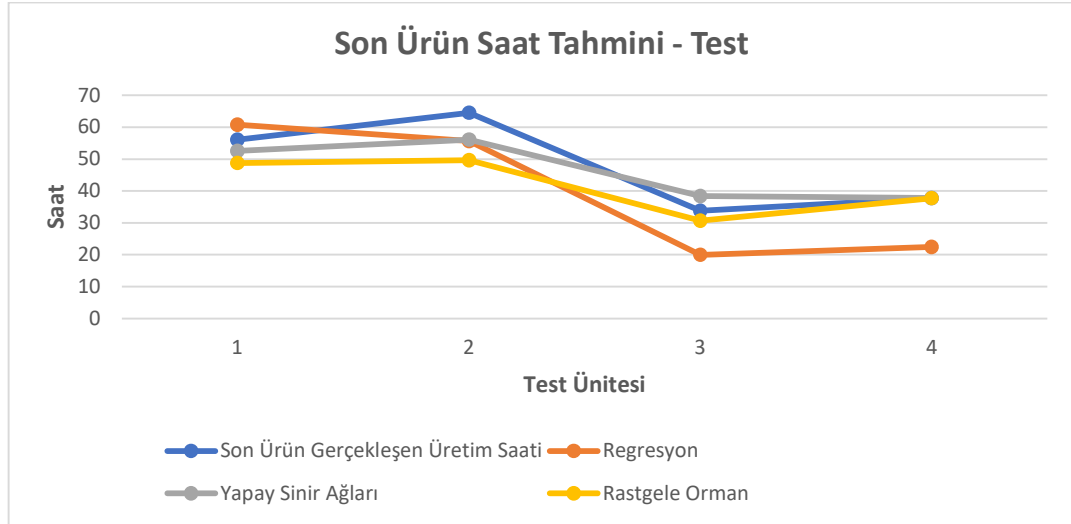
Yukarıda belirtilen veriler kullanılarak regresyon, yapay sinir ağları ve rastgele orman modelleri oluşturulmuştur. Oluşturulan modellerin eğitim ve test verileri için tahmin performans değerleri Tablo 15'te görülmektedir.

	EĞİTİM			TEST		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
<b>R<sup>2</sup></b>	0,97	0,98	0,97	0,89	0,81	0,92
<b>MSE</b>	54	36	51	131	58	71
<b>MAPE</b>	14	10	4	26	9	11

Tablo 15 - Son Ürün Tahmin Performans Sonuçları

Veri setine bağlı olarak değişmekle birlikte, R<sup>2</sup>'nin 0,70 üzerinde olduğu durumlarda genellikle modelin açıklama gücünün yeterli olduğunu varsayılır. Bu bölümde oluşturulan tüm modellerin R<sup>2</sup> değerleri hem eğitim verileri hem de test verileri için 0,80 üzerinde olduğundan, modellerin geçerliliği doğrulanmıştır.

Oluşturulan modeller ile Tablo 14'te belirtilen test ünitelerinin tahmin edilen toplam üretim saatleri Şekil 9'da görülmektedir.



Şekil 9 - Son ürün test verileri

Performans değerleri ve grafik incelendiğinde oluşturulan üç modelin de başarılı performans göstermesine rağmen MSE ve MAPE değerleri baz alındığında makine öğrenmesi yöntemlerinin (YSA ve RO), regresyona göre daha iyi sonuç verdiği görülmektedir. MSE ve MAPE değerlerine göre en iyi performansı YSA modeli göstermiştir.

### 3.4. DETAY PARÇA SEVİYESİ TAHMİN MODELLERİ

Uygulamanın ikinci aşamasında detay parça seviğinde modeller oluşturularak son ürün üretim saatleri tahmin edilecektir. Oluşturulan modeller ile detay parçaların saatleri tahmin edilecek, son ürünlerin ürün ağaçlarındaki detay parçaların tahmin süreleri toplanarak son ürünlerin toplam üretim saati tahmin edilecektir. Regresyon, yapay sinir ağları ve rastgele orman teknikleri kullanılarak modeller oluşturulacak ve bu modellerin performansları birbirleri ile kıyaslanacaktır. Bu aşamada “Birleşik veri tabanı modeli” ve “Proses veri tabanları modeli” yaklaşımı olmak üzere iki farklı model yaklaşımı uygulanacaktır. Birleşik veri tabanı yaklaşımında veri tabanındaki tüm parçalar için ortak tek bir tahmin modeli oluşturulacaktır. Proses veri tabanları yaklaşımında ise veri tabanındaki parçalar proseslerine göre gruplanarak her bir grup için ayrı olmak üzere toplam 8 tahmin modeli oluşturulacaktır.

### 3.4.1. Birleşik Veri Tabanı Yaklaşımı (TY2-D-BV)

12 adet son ürün kapsamındaki tüm detay parçalar ile tek bir veri tabanı oluşturularak tek bir tahmin modeli kurulacaktır. Oluşturulan veri tabanında toplam 1099 adet parça olup proses tiplerine göre miktarları Tablo 16'da özetlenmiştir.

Proses Tipi	Parça Sayısı
Büküm	386
Extrüzyon - Basit	196
Extrüzyon - Çekme	8
Hidroform Büküm	293
Talaşlı İmalat	112
Roll	50
Basit Kesme	33
Çekme	21
<b>Toplam Parça Sayısı</b>	<b>1099</b>

Tablo 16 - Birleşik Veri Tabanındaki Proses Tiplerine göre Parça Sayıları

Veri seti, eğitim ve test verisi olarak iki bölüme ayrılmıştır. Proses tiplerine göre parçaların dengeli dağılımı gözetilerek 8 son ürüne ait veriler eğitim, 4 son ürüne ait veriler test verisi olarak kullanılmıştır. Bu dağılıma göre 2., 5., 9. ve 12. son ürünlere ait detay parçalar test, diğerleri ise eğitim verisi olarak kullanılacaktır. Böylece 828 parça yani toplam detay parçaların %75'i eğitim, 271 parça yani toplamın %25'i ise test verisi olarak ayrılmıştır (Tablo 17).

<b>Eğitim Verileri</b>	Son Ürün -1'e ait detay parçalar
	Son Ürün -3'e ait detay parçalar



	Son Ürün -4'e ait detay parçalar
	Son Ürün -6'e ait detay parçalar
	Son Ürün -7'e ait detay parçalar
	Son Ürün -8'e ait detay parçalar
	Son Ürün -10'e ait detay parçalar
	Son Ürün -11'e ait detay parçalar
<b>Test Verileri</b>	Son Ürün -2'e ait detay parçalar
	Son Ürün -5'e ait detay parçalar
	Son Ürün -9'e ait detay parçalar
	Son Ürün -12'e ait detay parçalar

Tablo 17 - Birleşik Veri Tabanı Eğitim ve Test Verileri

Bu veri tabanı ile oluşturulacak modeller için Tablo 11'de belirtilen parametrelere ilaveten parçaların proses tipleri kukla değişken olarak tanımlanmıştır.

### 3.4.1.1. Detay Parça Bazında Model Performansları

Regresyon, yapay sinir ağları ve rastgele orman yöntemleri kullanılarak modeller oluşturulmuş ve bu modellerin performansları birbirleri ile kıyaslanmıştır. Üretim saati tahmini için oluşturulan modellerin detay parça bazında performans parametreleri alttaki tabloda özetlenmiştir.

	EĞİTİM			TEST		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
<b>R<sup>2</sup></b>	0,85	0,96	0,93	0,88	0,98	0,89
<b>MSE</b>	0,39	0,09	0,19	0,52	0,09	0,49
<b>MAPE</b>	131	42	26	137	74	40

Tablo 18 - Birleşik Veri Tabanı Modellerinin Karşılaştırması

Birleşik veri tabanı kapsamında oluşturulan modellerin performans özetleri Tablo 18'da karşılaştırılmıştır. Test verileri değerlendirildiğinde R<sup>2</sup> ve MSE kriterlerine göre yapay sinir ağları daha iyi performans göstermiş olmasına rağmen MAPE kriterine göre rastgele orman modeli daha iyi performans göstermiştir. Üç model için de R<sup>2</sup> değerleri değerlendirildiğinde oluşturulan

modellerin açıklama güçlerinin gayet başarılı olduğu görülmektedir. Modeller için MSE değerlerinin küçük olmasına rağmen MAPE değerlerinin büyük olduğu dikkat çekmektedir. Bunun nedeni Tablo 10'da belirtildiği üzere toplam 1099 adet detay parçanın yüzde 80'inin üretim saatlerinin 0,5'den küçük olmasıdır. Tüm parçalar ortak tek bir model ile tahmin edilmeye çalışıldığında tahmin değerlerinin gerçekleştirmelerden sapmaları yüzdesel olarak daha büyük olmaktadır.

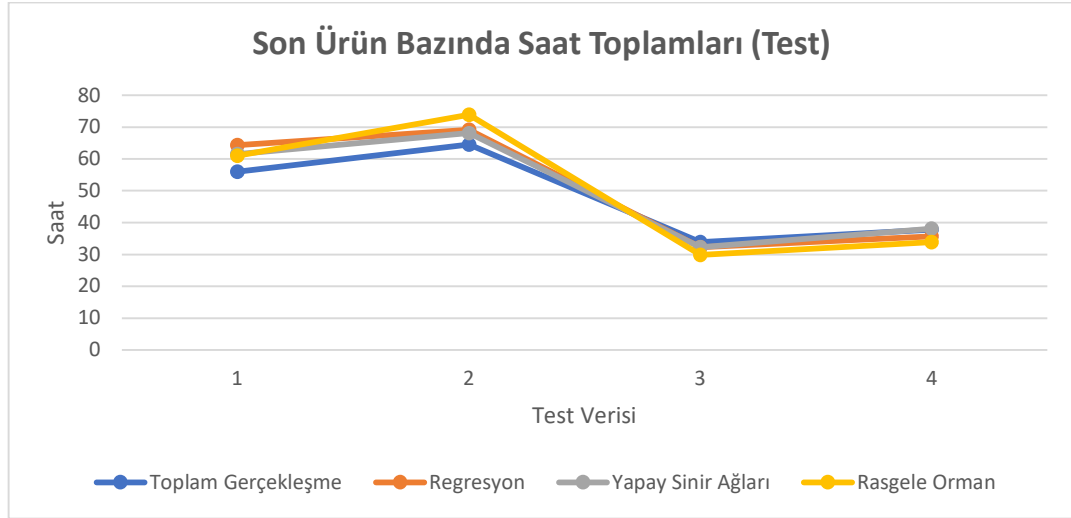
### 3.4.1.2. Son Ürün Toplamları Bazında Model Performansları

Birleşik veri tabanı yaklaşımı ile oluşturulan modeller ile detay parça bazında yapılan tahminler, ürün ağaçlarına göre son ürünler bazında toplanarak Tablo 19'de karşılaştırılmıştır.

Eğitim / Test	Son Ürün No	Toplam Gerçekleşme	Birleşik Veri Tabanı ile Tahmin Toplamı		
			Regresyon	Yapay Sinir Ağları	Rastgele Orman
Eğitim	Son Ürün -1	41	45	43	42
	Son Ürün -3	216	191	217	217
	Son Ürün -4	41	45	43	42
	Son Ürün -6	46	63	48	50
	Son Ürün -7	27	31	27	28
	Son Ürün -8	42	43	42	37
	Son Ürün -10	25	21	25	25
	Son Ürün -11	37	36	39	34
<b>Eğitim</b>	<b>TOPLAM</b>	<b>475</b>	<b>475</b>	<b>484</b>	<b>474</b>
Test	Son Ürün -2	56	64	62	59
	Son Ürün -5	65	69	75	73
	Son Ürün -9	34	32	34	29
	Son Ürün -12	38	36	38	34
<b>Test</b>	<b>TOPLAM</b>	<b>192</b>	<b>201</b>	<b>209</b>	<b>195</b>

Tablo 19 - Birleşik Veri Tabanı Modelleri Sonuçlarının Son Ürün Bazında Karşılaştırması

Test üniteleri için gerçekleşme ve tahmin değerleri Şekil 10'de görülmektedir.



Şekil 10 - Birleşik Veri Tabanı Son Ürün Toplamı - Test

Detay bazda yapılan tahminlerin toplamı ile elde edilen son ürün üretim saatleri tahminleri için performans parametreleri Tablo 20'de özetlenmiştir. Modellerin açıklama güçlerinin hem eğitim hem de test verileri için oldukça yüksektir. Regresyon modelinde MSE değerinin eğitim ve test verisi için oldukça farklı olduğu görülmektedir. Tablo 19 incelendiğinde bunun nedeni son ürün-3'e ait tahmin hatasının oransal olarak olmasa da sayısal olarak yüksek olmasıdır.

Modellerin MAPE performansları birbirine yakın olmakla birlikte, test verilerine göre YSA modelinin daha başarılı olduğu söylenebilir.

Birleşik Veri Tabanı ile Son Ürün Toplam Saati Tahmin Performansı						
	Eğitim			Test		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
<b>R<sup>2</sup></b>	0,98	1,00	1,00	0,98	1,00	1,00
<b>MSE</b>	120	2	8	24	35	29
<b>MAPE</b>	13	3	5	8	7	11

Tablo 20 - Son Ürün Bazında Tahmin Performansı

### 3.4.2. Proses Veri Tabanları Yaklaşımı (TY3-D-PV)

Bu aşamada birleşik veri tabanındaki parçalar, Tablo 16'da belirtilen prosesler bazında bölünerek 8 ayrı veri tabanı oluşturulmuş ve her bir proses için ayrı modeller kurulmuştur. Regresyon, yapay sinir ağları ve rastgele orman yöntemleri kullanılarak modeller oluşturulmuş ve bu modellerin detay parça üretim saat tahminleri için performansları birbirleri ile kıyaslanmıştır. Bağımsız değişkenler belirlenirken korelasyon matrisi incelenerek proses tipine göre üretim saati üzerinde etkili olduğu düşünülen değişkenler ile regresyon modelleri kurulmuş, p-değerinin anlamlı olduğu ve en yüksek uyarlanmış  $R^2$  değerine sahip olan modeldeki parametreler dikkate alınarak regresyon, YSA ve RO modelleri oluşturulmuştur.

#### Büküm Prosesi

Bağımsız değişkenler: aylık ihtiyaç miktarı, uzunluk

Eğitim Veri Sayısı	283
Test Veri Sayısı	103
Eğitim Oranı	73%

En düşük üretim saati: 0,01

En yüksek üretim saati: 0,65

Yukarıda tabloda belirtildiği gibi büküm prosesi veri tabanındaki 283 adet eğitim verisi kullanılarak regresyon, YSA ve RO modelleri oluşturulmuş ve oluşturulan bu modeller 103 adet test verisi ile test edilmiştir. Büküm parçaları veri tabanı ile oluşturulan modellerin performans özetleri Tablo 21'de özetlenmiştir.

	EĞİTİM			TEST		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
R <sup>2</sup>	0,83	0,91	0,92	0,88	0,95	0,95
MSE	0,002	0,001	0,001	0,003	0,001	0,002
MAPE	52	19	22	59	16	19

Tablo 21 - Modellerin Performans Karşılaştırması - Büküm

Bu parçaların üretim saatleri 0,01 – 0,65 saat aralığında yani oldukça küçük olduğu için modellerin MSE değerleri de çok düşüktür. Modelin açıklama gücü, MSE ve MAPE kriterleri dikkate alındığında YSA modelinin daha başarılı olduğu söylenebilir.

### Ekstrüzyon-basit

Bağımsız değişkenler: iş emri başına ortalama miktar, uzunluk

Eğitim Veri Sayısı	145
Test Veri Sayısı	51
Eğitim Oranı	74%

En düşük üretim saati: 0,02

En yüksek üretim saati: 0,80

Ekstrüzyon-basit parçalar veri tabanı ile oluşturulan modellerin performans özetleri Tablo 22'de özetlenmiştir.

	EĞİTİM			TEST		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
R <sup>2</sup>	0,48	0,61	0,81	0,55	0,70	0,56
MSE	0,01	0,01	0,01	0,05	0,01	0,01
MAPE	31	25	21	201	20	61

Tablo 22 - Modellerin Performans Karşılaştırması - Ekstrüzyon basit

$R^2$  dikkate alındığında üç modelin de açıklama güçlerinin yeterince iyi performans sergilemediği, yani bu parametreler ile bu tipteki parçaların üretim saatlerinin yeterince iyi açıklanamadığı görülmektedir. Muhtemelen bu tip parçaların teknik verileri daha detaylı bir şekilde incelenerek üretim saatini etkileyen başka parametreler bulunabilecektir fakat bu detayda bir çalışma bu tez kapsamının dışındadır. MSE kriterine göre YSA ve RO modellerinin tahmin performansları birbirine yakın olmakla birlikte MAPE dikkate alındığında YSA modelinin en iyi performansı gösterdiği görülmektedir. Bu proses kapsamındaki parçaların üretim saatleri düşük olduğu için modellerin MSE değerlerinin de çok düşük olduğu görülmektedir.

### **Ekstrüzyon-çekme**

Bu proses kapsamındaki veri sayısının azlığından dolayı daha sağlıklı tahmin modelleri kurabilmek için bootstrap yöntemi ile veriler çoğaltılmıştır. Bu yöntem ile rastgele seçilen 2 son ürün içeren 100 örneklem üretilmiştir. Bağımlı ve bağımsız değişken verileri olarak her örneklemdeki 2 son ürüne ait verinin ortalamaları kullanılmıştır.

Bağımsız değişkenler: iş emri başına ortalama miktar, parça uzunluğu

	Örneklem sayısı	Bootstrap ile çoğaltılmış örneklem sayısı
Eğitim Veri Sayısı	5	85
Test Veri Sayısı	3	23
Eğitim Oranı	63%	78%

En düşük üretim saati: 0,79

En yüksek üretim saati: 2,35

Ekstrüzyon-çekme parçalar veri tabanı ile oluşturulan modellerin performans özetleri Tablo 23'te özetlenmiştir.

	EĞİTİM			TEST		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
<b>R<sup>2</sup></b>	0,97	0,99	0,99	0,96	0,98	0,97
<b>MSE</b>	0,005	0,001	0,002	0,01	0,001	0,003
<b>MAPE</b>	4	1	3	5	2	4

Tablo 23 - Modellerin Performans Karşılaştırması - Ekstrüzyon çekme

Bu proses için oluşturulan modellerin her üçü de gayet başarılı olarak birbirine yakın performanslar göstermişlerdir. Yine de bir karşılaştırma yapmak gerekirse YSA modelinin diğerlerine göre daha başarılı olduğu söylenebilir.

### Çekme

Bağımsız değişkenler: Alan

Eğitim Veri Sayısı	13
Test Veri Sayısı	8
Eğitim Oranı	62%

En düşük üretim saati: 1,90

En yüksek üretim saati: 7,50

Çekme Prosesi parçaları veri tabanı ile oluşturulan modellerin performans özetleri Tablo 24'te özetlenmiştir.

	EĞİTİM			TEST		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
<b>R<sup>2</sup></b>	0,62	0,84	0,64	0,81	0,98	0,60
<b>MSE</b>	1,40	0,6	1,61	0,68	0,1	1,55
<b>MAPE</b>	21	12	31	13	7	34

Tablo 24 - Modellerin Performans Karşılaştırması - Çekme

Regresyon ve RO modellerinin  $R^2$  deęerleri modellerin açıklama gücünün yeterince iyi olmadığını göstermektedir. Üç performans kriterine göre de YSA modelinin daha başarılı olduęu söylenebilir.

### Hidroform Büküm

Bağımsız deęişkenler: iş emri başına ortalama miktar, alan

Eđitim Veri Sayısı	227
Test Veri Sayısı	66
Eđitim Oranı	77%

En düşük üretim saati: 0,02

En yüksek üretim saati: 1,22

Hidroform Büküm parçaları veri tabanı ile oluşturulan modellerin performans özetleri Tablo 25'de özetlenmiştir.

	EĐTİM			TEST		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
$R^2$	0,71	0,84	0,91	0,77	0,68	0,58
MSE	0,01	0,01	0,00	0,01	0,03	0,04
MAPE	69	58	33	35	44	58

Tablo 25 - Modellerin Performans Karşılaştırması – hidroform büküm

Test verileri performans deęerlerine göre regresyon modelinin dięerlerine göre daha başarılı olduęu görülmektedir.

### Talaşlı İmalat



Bağımsız değişkenler: iş emri başına ortalama miktar, hacim

Eğitim Veri Sayısı	87
Test Veri Sayısı	25
Eğitim Oranı	78%

En düşük üretim saati: 0,24

En yüksek üretim saati: 16,20

Talaşlı İmalat parçaları veri tabanı ile oluşturulan modellerin performans özetleri Tablo 26'da özetlenmiştir.

	EĞİTİM			TEST		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
<b>R<sup>2</sup></b>	0,76	0,81	0,87	0,91	0,98	0,94
<b>MSE</b>	2,93	2,29	2,07	2,49	0,34	2,91
<b>MAPE</b>	78	35	35	73	17	37

Tablo 26 – Modellerin Performans Karşılaştırması – Talaşlı İmalat

Test verileri performans değerlerine göre yapay sinir ağları modelinin diğerlerine göre belirgin şekilde daha başarılı olduğu söylenebilir.

## Roll

Bağımsız değişkenler: uzunluk

Eğitim Veri Sayısı	45
Test Veri Sayısı	5
Eğitim Oranı	90%

En düşük üretim saati: 0,02

En yüksek üretim saati: 2,66

Roll parçaları veri tabanı ile oluşturulan modellerin performans özetleri Tablo 27'de özetlenmiştir

	EĞİTİM			TEST		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
<b>R<sup>2</sup></b>	0,92	0,99	0,64	0,01	1,00	0,24
<b>MSE</b>	0,02	0,02	0,11	0,20	0,001	0,26
<b>MAPE</b>	77	20	68	325	17	350

Tablo 27 – Modellerin Performans Karşılaştırması – Roll

Performans değerlerine göre yapay sinir ağları modelinin diğerlerine göre belirgin şekilde daha başarılı olduğu söylenebilir.

### Basit Kesme

Bağımsız değişkenler: aylık ihtiyaç miktarı, alan

Eğitim Veri Sayısı	23
Test Veri Sayısı	10
Eğitim Oranı	70%

En düşük üretim saati: 0,01

En yüksek üretim saati: 0,08

Basit Kesme parçaları veri tabanı ile oluşturulan modellerin performans özetleri Tablo 28'de özetlenmiştir

	EĞİTİM			TEST		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
<b>R<sup>2</sup></b>	0,63	0,94	0,54	0,98	0,98	0,63
<b>MSE</b>	0,0001	0,0002	0,0002	0,0001	0,0000	0,0002
<b>MAPE</b>	50	21	52	11	10	48

Tablo 28 – Modellerin Performans Karşılaştırması – Basit Kesme

Performans değerlerine göre yapay sinir ağları modelinin diğerlerine göre daha başarılı olduğu söylenebilir.

#### 3.4.2.1. Proses veri tabanları Model Performansları

Bir önceki bölümde her bir proses grubu için ayrı modeller kurulmuş, böylece 8 farklı proses grubu için 8 farklı model elde edilmiş ve her bir modelin tahmin performansı ayrı ayrı incelenmişti. Bir önceki bölümdeki amaç, farklı proses tiplerindeki parçalar için tahmin modelleri oluşturarak bu modellerin geçerliliğini sınamaktı. Bu tez kapsamındaki asıl amaç ise detay parçalardan oluşan son ürünlerin toplam saatlerini tahmin etmek olduğundan Proses veri tabanları yaklaşımının performansını belirleyebilmek için son ürün toplamlarında tahmin performansını incelemek gerekmektedir. Bu nedenle bu bölümde toplam 1099 parçanın her biri için ilgili proses modeli kullanılarak yapılan tahminlerin bütüncül performansı sorgulanacaktır. Bir önceki bölümde 8 son ürüne ait toplam 828 parça eğitim, 4 son ürüne ait 271 parça eğitim verisi olarak kullanılmıştır. Bu bölümde eğitim ve test verileri önce detay parça bazında, sonra da son ürün toplamında performans değerleri hesaplanacaktır.

### 3.4.2.1.1. Detay Parça Bazında

Proses veri tabanları yaklaşımı ile tahmin edilen detay parçaların, 828 eğitim verisi ve 271 test verisi için parça bazında model performansları alttaki tabloda görülebilir.

	EĞİTİM			TEST		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
<b>R<sup>2</sup></b>	0,87	0,90	0,91	0,94	0,99	0,93
<b>MSE</b>	0,34	0,25	0,25	0,27	0,04	0,33
<b>MAPE</b>	56	32	29	83	23	45

Tüm performans kriterleri değerlendirildiğinde yapay sinir ağları modelinin diğerlerine göre daha başarılı olduğu söylenebilir.

### 3.4.2.1.2. Son Ürün Seviyesi Toplamı Bazında

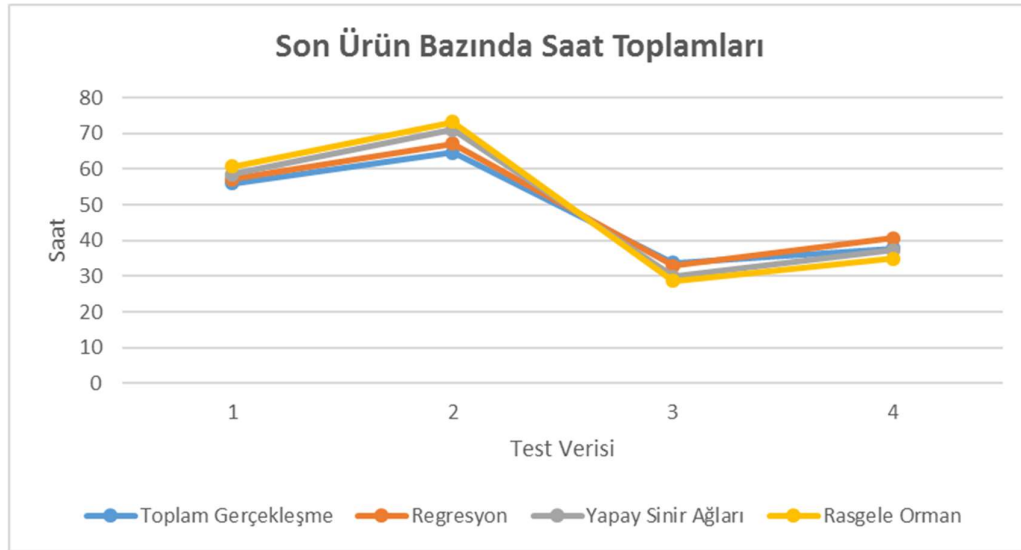
Her bir parçanın saati, ilgili proses modeli kullanılarak tahmin edilmiş ve ürün ağacına göre son ürün bazında saat toplamları alttaki tablodaki verilmiştir.

Eğitim / Test	Son Ürün No	Toplam Gerçekleşme	Proses veri tabanları ile Tahmin Toplamı		
			Regresyon	Yapay Sinir Ağları	Rastgele Orman
Eğitim	Son Ürün -1	41	44	45	44
	Son Ürün -3	216	201	190	208
	Son Ürün -4	41	44	45	44
	Son Ürün -6	46	53	48	54
	Son Ürün -7	27	27	27	27
	Son Ürün -8	42	41	41	38
	Son Ürün -10	25	24	26	21
	Son Ürün -11	37	41	38	35
<b>Eğitim</b>	<b>TOPLAM</b>	<b>475</b>	<b>475</b>	<b>459</b>	<b>471</b>
Test	Son Ürün -2	56	57	56	61

	Son Ürün -5	65	67	66	73
	Son Ürün -9	34	33	33	29
	Son Ürün -12	38	41	37	35
<b>Test</b>	<b>TOPLAM</b>	<b>192</b>	<b>198</b>	<b>192</b>	<b>198</b>

Tablo 29 - Proses veri tabanları Modelleri Sonuçlarının Son Ürün Bazında Karşılaştırması

Test üniteleri için son ürün toplamları bazında gerçekleşme ve tahmin değerleri Şekil 11'de görülmektedir



Şekil 11 - Proses veri tabanları Modelleri Sonuçlarının Son Ürün Bazında Karşılaştırması - Test

Son ürün bazında elde edilen model performansları alttaki tabloda özetlenmiştir.

	Proses veri tabanları ile Son Ürün Toplam Saati Tahmin Performansı					
	Eğitim			Test		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
<b>R<sup>2</sup></b>	1,00	1,00	0,99	0,99	1,00	1,00
<b>MSE</b>	37	85	22	4	1	33
<b>MAPE</b>	7	5	9	4	2	11

Tablo 30 – Proses veri tabanları Son Ürün Bazında Tahmin Performansı

$R^2$  değerlerine göre üç modelin de gayet başarılı olduğu görülmektedir. MSE ve MAPE kriterlerine göre YSA modeli, diğerlerine göre daha başarılı olmuştur.

### 3.4.3. Birleşik Veri Tabanı (TY2-D-BV) ve Proses Veri Tabanları (TY3-D-PV) Modellerinin Performans Kıyaslamaları

Bazen firmalar sadece detay parçalardan oluşan teklife çağrı dokümanlarına cevap vermek durumunda olabilirler. Bu durumda firmalar hızlı bir tahminde bulunabilmek için son ürün tahmin yaklaşımı kullanamayacak ve dolayısıyla ellerinde sadece “birleşik veri tabanı” ve “Proses veri tabanları” olmak üzere iki tahmin yaklaşımı olacaktır. Bu gibi durumlarda firmanın hangi yaklaşımı seçeceğine karar vermesi için bu iki yaklaşımın süre – tahmin başarısı ilişkisini biliyor olması büyük önem taşır. Bu nedenle bu bölümde bu iki yaklaşımın tahmin performansları detaylı olarak incelenecektir.

#### 3.4.3.1. Detay Parça Bazında Kıyaslama

Bu bölümde, birleşik veri tabanı ve proses veri tabanları yaklaşımları ile oluşturulan modeller kullanılarak yapılan tahminlerin, detay parça bazında performansları değerlendirilmiştir.

Oluşturulan modellerin eğitim verileri için performans değerleri Tablo 31’de, test verileri için performans değerleri Tablo 32’de özetlenmiştir.

	Regresyon		Yapay Sinir Ağları		Rastgele Orman	
	Birleşik Veri Tabanı	Proses veri tabanları	Birleşik Veri Tabanı	Proses veri tabanları	Birleşik Veri Tabanı	Proses veri tabanları
$R^2$	0,85	0,87	0,96	0,90	0,93	0,91
MSE	0,39	0,34	0,09	0,26	0,19	0,25
MAPE	131	56	42	32	26	29

Tablo 31 - Detay Parça Bazında Kıyaslama - Eğitim

	Regresyon		Yapay Sinir Ağları		Rastgele Orman	
	Birleşik Veri Tabanı	Proses veri tabanları	Birleşik Veri Tabanı	Proses veri tabanları	Birleşik Veri Tabanı	Proses veri tabanları
<b>R<sup>2</sup></b>	0,88	0,94	0,98	0,99	0,89	0,93
<b>MSE</b>	0,52	0,27	0,09	0,04	0,49	0,33
<b>MAPE</b>	137	83	74	23	40	45

Tablo 32 - Detay Parça Bazında Kıyaslama - Test

Modellerin performansları değerlendirildiğinde, regresyon ve yapay sinir ağları modelleri için Proses veri tabanları modellerinin, birleşik veri tabanı modellerine göre tahmin performanslarının daha iyi olduğu görülmektedir. Rastgele orman yöntemi için ise iki veri tabanı için de performansının benzer olduğu görülmektedir.

### 3.4.3.2. Son Ürün Toplamları Bazında Kıyaslama

Bu bölümde, birleşik veri tabanı ve proses veri tabanları ile oluşturulan modeller kullanılarak detay parça bazında yapılan tahminler ürün ağacına göre son ürün bazında toplanmış, elde edilen son ürün toplamlarının toplam saat gerçekleştirmelerine göre performansları değerlendirilmiştir.

Oluşturulan modellerin eğitim verileri için performans değerleri Tablo 33'de, test verileri için performans değerleri Tablo 34'te özetlenmiştir.

	Regresyon		Yapay Sinir Ağları		Rastgele Orman	
	Birleşik Veri Tabanı	Proses veri tabanları	Birleşik Veri Tabanı	Proses veri tabanları	Birleşik Veri Tabanı	Proses veri tabanları
<b>R<sup>2</sup></b>	0,98	1,00	1,00	1,00	1,00	0,99
<b>MSE</b>	120	37	22	85	7	22
<b>MAPE</b>	13	7	3	5	5	9

Tablo 33 - Son Ürün Bazında Kıyaslama - Eğitim

	Regresyon		Yapay Sinir Ağları		Rastgele Orman	
	Birleşik Veri Tabanı	Proses veri tabanları	Birleşik Veri Tabanı	Proses veri tabanları	Birleşik Veri Tabanı	Proses veri tabanları
<b>R<sup>2</sup></b>	0,98	0,99	1,00	1,00	1,00	1,00
<b>MSE</b>	24	4	35	1	29	33
<b>MAPE</b>	8	4	7	2	11	11

Tablo 34 - Son Ürün Bazında Kıyaslama - Test

Performans verileri incelendiğinde regresyon ve yapay sinir ağları yöntemleri ile oluşturulan Proses veri tabanları modellerinin, birleşik veri tabanı modellerine göre daha başarılı olduğu görülmektedir. Rastgele orman yöntemi kullanıldığında ise, büyük bir fark olmamasına rağmen birleşik veri tabanı modelinin daha başarılı olduğu görülmektedir.

### 3.5. ÖĞRENME EĞRİSİNE GÖRE VERİ DÖNÜŞÜMÜ VE MODEL PERFORMANSINA ETKİSİ

Bir işçi, yaptığı bir işi tekrarladıkça o işteki tecrübesi ve bilgi birikimi artacak ve bu da o işteki performansının iyileşmesine neden olacaktır. Öğrenme eğrisi, tekrar edilen bir işte, bir işçinin öğrenme performansını ortaya koyan matematiksel bir modeldir (Wright, 1936). Literatürde öğrenme eğrisini farklı matematiksel modeller olarak ortaya koyan birçok çalışma bulunmasına rağmen Wright (1936) tarafından ortaya konulan model genel olarak kabul görmektedir. Bu modele göre bir işteki tekrar sayısı ikiye katlandıkça, o işin süresi belirli bir yüzdesel oranda azalmaktadır. Bu çalışmada öğrenme eğrisi olarak Wright (1936) modeli kullanılmıştır.

Bu bölümde tahmin modellerindeki bağımlı değişken olan üretim sürelerinin, öğrenme eğrisine göre veri dönüşümü yapılarak modellenmesinin, model tahmin performansına etkisi olup olmadığı araştırılacaktır. Bu amaçla birleşik veri tabanı yaklaşımı kapsamında oluşturulan modeller, bağımlı değişkenin öğrenme eğrisine göre dönüştürülmüş haliyle tekrarlanacaktır. Bu şekilde oluşturulan



regresyon, YSA ve RO modellerinin performans sonuçları, öğrenme eğrisi kullanılmayan model performansları ile kıyaslanacaktır.

Normal bir tahmin sürecinde modelin oluşturulacağı örneklem belirlenerek, bu örneklem için şirket veri tabanından gerçekleşen üretim süreleri raporlanır. Örneklem verilerine hızlı erişebilmek ve verinin büyüklüğünden dolayı sorguların hata vermeden çalışabilmesi için genellikle sorgulara bir tarih sınırı getirilerek belirlenen bir tarih aralığındaki en güncel veriler raporlanır. Bu durumda bir parçaya ait geçmiş verilerin sınırlı bir kısmı analiz edilebileceğinden, analiz sonuçları kullanım durumuna göre yanıltıcı olabilecektir. Parçanın geçmiş üretimlerine ait üretim saatlerini daha doğru analiz ederek daha anlamlı model sonuçları elde edebilmek için öğrenme eğrisi dikkate alınmalıdır. Maliyet tahmin literatüründe öğrenme eğrisinin model performansları üzerindeki etkisini ele alan bir çalışma bulunamamıştır. Bu anlamda bu bölümde yapılan çalışmaların literatüre katkıda bulunması beklenmektedir.

Hazırlanan gerçekleştirme raporunda, her bir parçaya ait gerçekleştirmeler, parçaların farklı üretim ünitelerini kapsıyor olabilir. Bu durumda her parçayı aynı modelde değerlendirebilmek için gerçekleştirmelerini ortak bir paydada toplamak, modelin daha sağlıklı olması için önemlidir. Konunun daha iyi anlaşılması için Şekil 12’de iki örnek parça gösterilmiştir. Varsayalım ki raporda örnek “A” parçası için 60. ve 80. üniteler arası için gerçekleştirmeler gelmiştir. Örnek öğrenme eğrisine göre bu üniteler arası ortalaması 28,8 birimdir. Raporda örnek “B” parçası için 100. ve 120. üniteler arası için gerçekleştirmeler gelmiştir. Örnek öğrenme eğrisine göre bu üniteler arası ortalaması 25,2 birimdir. Bu iki parçayı aynı modelde değerlendirebilmek için aynı üretim üniteleri paydasında toplamak gerekir. Ortak paydayı 150. üretim ünitesi olarak belirlediğimizi varsayarsak, iki parçanın da gerçekleştirmelerini öğrenim eğrisindeki 150. üniteye göre dönüştürmek gerekecektir. Tablo 35’de görüleceği üzere, öğrenim eğrisinin 150. ünitesi 23 birimdir. Parçaların gerçekleştirmelerini tablodaki öğrenim eğrisi birim değerlerine göre 150. üniteye dönüştürdüğümüzde A parçası için 9,6 saat B parçası için 16 saat değerleri elde edilir.

	60 - 80 üniteler arası ortalama	100 - 120 üniteler arası ortalama	150. ünite değeri
<b>Öğrenme eğrisi birim değeri</b>	28,8	25,2	23
<b>A parçası gerçekleşmesi</b>	12		
<b>B parçası gerçekleşmesi</b>		20	
<b>A parçası gerçekleşmesinin 150. üniteye veri dönüşümü</b>			9,6
<b>B parçası gerçekleşmesinin 150. üniteye veri dönüşümü</b>			16,0

Tablo 35 - Öğrenme Eğrisine Göre Gerçekleşme Dönüşümü

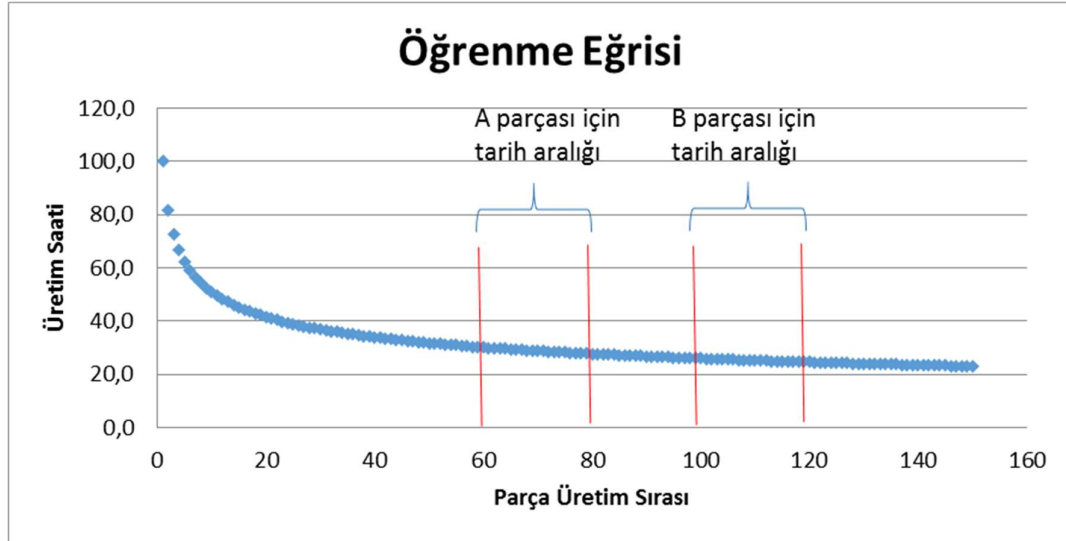
Öğrenme eğrisi genel formülü:  $y_x = y_1 \times x^{-b}$

$y_x$  =  $x$ . sıradaki ünitenin üretim saati

$x$  = üretim sırası

$b$  = öğrenim eğrisi katsayısı

Örnek bir öğrenme eğrisi Şekil 12'de verilmiştir.



Şekil 12 -Örnek Öğrenme Eğrisi ve Veri Raporlama Aralığı

Bu bölümde, ikinci aşamada oluşturulan birleşik veri tabanı modeli, öğrenme eğrisi de hesaba katılarak tekrarlanmıştır. Tahmin modellerindeki bağımlı değişken olan üretim sürelerinin, öğrenme eğrisine göre veri dönüşümü yapılarak modellendiğinde, tahmin modeli sonuçlarının iyileşip iyileşmediği araştırılacaktır.

### 3.5.1. Detay Parça Bazında

Oluşturulan modellerin detay parça bazında performansı, Tablo 36'de özetlenmiştir.

	EĞİTİM			TEST		
	Regresyon	YSA	RO	Regresyon	YSA	RO
<b>R<sup>2</sup></b>	0,84	0,93	0,94	0,87	0,98	0,88
<b>MSE</b>	0,44	0,19	0,18	0,59	0,10	0,59
<b>MAPE</b>	129	95	26	137	63	40

Tablo 36 - Öğrenme Eğrisi ile dönüştürülmüş birleşik veri modeli

Öğrenme eğrili modelin, “birleşik veri tabanı yaklaşımı” bölümünde oluşturulan model ile performans karşılaştırması eğitim verileri için Tablo 37’de test verileri için Tablo 38’de özetlenmiştir.

	Regresyon		Yapay Sinir Ağları		Rastgele Orman	
	Birleşik Veri Tabanı	Birleşik Veri Tabanı (ÖE)	Birleşik Veri Tabanı	Birleşik Veri Tabanı (ÖE)	Birleşik Veri Tabanı	Birleşik Veri Tabanı (ÖE)
<b>R<sup>2</sup></b>	0,85	0,84	0,94	0,93	0,93	0,94
<b>MSE</b>	0,39	0,44	0,15	0,19	0,19	0,18
<b>MAPE</b>	131	129	98	95	27	26

Tablo 37 - Öğrenme Eğrisi Veri Dönüşümünün Model Performansına Etkisi - Eğitim

	Regresyon		Yapay Sinir Ağları		Rastgele Orman	
	Birleşik Veri Tabanı	Birleşik Veri Tabanı (ÖE)	Birleşik Veri Tabanı	Birleşik Veri Tabanı (ÖE)	Birleşik Veri Tabanı	Birleşik Veri Tabanı (ÖE)
<b>R<sup>2</sup></b>	0,88	0,87	0,98	0,98	0,89	0,88
<b>MSE</b>	0,52	0,59	0,10	0,10	0,48	0,59
<b>MAPE</b>	137	137	96	63	42	40

Tablo 38 - Öğrenme Eğrisi Veri Dönüşümünün Model Performansına Etkisi – Test

Öğrenme eğrisine göre veri dönüşümü yapılarak oluşturulan modellerden yapay sinir ağları modelinde gözle görülür bir iyileşme olmuş, fakat diğer modellerde sonuçlar birbirlerine çok yakın olsa da bir iyileşme gözlemlenmemiştir.

### 3.5.2. Son Ürün Toplamları Bazında

Öğrenme eğrisine göre veri dönüşümü yapılmış haliyle birleşik veri tabanı modelleri için son ürün bazında tahmin toplamaları Tablo 39’da verilmiştir.

Eğitim / Test	Son Ürün No	Toplam Gerçekleşme (öğrenme eğrisi dönüşümlü)	Birleşik Veri Tabanı ile Son Ürün Toplam Saati Tahmin		
			Regresyon	Yapay Sinir Ağları	Rastgele Orman
Eğitim	Son Ürün -1	44	46	41	42
	Son Ürün -3	229	205	211	230
	Son Ürün -4	43	47	41	43
	Son Ürün -6	49	66	53	53
	Son Ürün -7	27	32	30	29
	Son Ürün -8	43	44	44	38
	Son Ürün -10	25	22	24	25
	Son Ürün -11	38	37	39	35
Eğitim	<b>TOPLAM</b>	<b>498</b>	<b>498</b>	<b>482</b>	<b>495</b>
Test	Son Ürün -2	59	67	58	61
	Son Ürün -5	68	73	63	75
	Son Ürün -9	34	33	35	31
	Son Ürün -12	38	36	38	35
Test	<b>TOPLAM</b>	<b>200</b>	<b>209</b>	<b>194</b>	<b>202</b>

Tablo 39 - Öğrenme Eğrisi İle Dönüştürülmüş Modellerin Son Ürün Toplamları Bazında Tahmin Sonuçları

Oluşturulan modellerin son ürün bazında toplam tahmin performansları Tablo 40'da verilmiştir.

	Öğrenme Eğrili Birleşik Veri Tabanı ile Son Ürün Toplam Saati Tahmin Performansı					
	Eğitim			Test		
	Regresyon	Yapay Sinir Ağları	Rastgele Orman	Regresyon	Yapay Sinir Ağları	Rastgele Orman
<b>R<sup>2</sup></b>	0,99	1,00	1,00	0,98	1,00	1,00
<b>MSE</b>	114	44	8	22	8	20
<b>MAPE</b>	12	6	5	7	3	8

Tablo 40 - Öğrenme Eğrisi İle Dönüştürülmüş Modellerin Son Ürün Toplamları Bazında Tahmin Performansları

Öğrenme eğrili modelin birleşik veri tabanı modeli ile son ürün toplamları bazında performans karşılaştırması eğitim verileri için Tablo 41’de test verileri için Tablo 42’de özetlenmiştir.

	Regresyon		Yapay Sinir Ağları		Rasgele Orman	
	Birleşik Veri Tabanı	Birleşik Veri Tabanı (ÖE)	Birleşik Veri Tabanı	Birleşik Veri Tabanı (ÖE)	Birleşik Veri Tabanı	Birleşik Veri Tabanı (ÖE)
<b>R<sup>2</sup></b>	0,98	0,99	1,00	1,00	1,00	1,00
<b>MSE</b>	120	114	32	44	8	8
<b>MAPE</b>	13	12	9	6	6	5

Tablo 41 - Öğrenme Eğrisi Veri Dönüşümünün Model Performansına Etkisi – Eğitim

	Regresyon		Yapay Sinir Ağları		Rasgele Orman	
	Birleşik Veri Tabanı	Birleşik Veri Tabanı (ÖE)	Birleşik Veri Tabanı	Birleşik Veri Tabanı (ÖE)	Birleşik Veri Tabanı	Birleşik Veri Tabanı (ÖE)
<b>R<sup>2</sup></b>	0,98	0,98	0,99	1,00	1,00	1,00
<b>MSE</b>	24	22	12	8	36	20
<b>MAPE</b>	8	7	5	3	11	8

Tablo 42 - Öğrenme Eğrisi Veri Dönüşümünün Model Performansına Etkisi - Test

Öğrenme eğrisi ile dönüştürülmüş modellerin son ürün toplamları bazında tahmin performansları değerlendirildiğinde, YSA ve RO modellerinde fark edilebilir bir iyileşme görülürken regresyon modelinde çok küçük bir iyileşme görülmektedir.

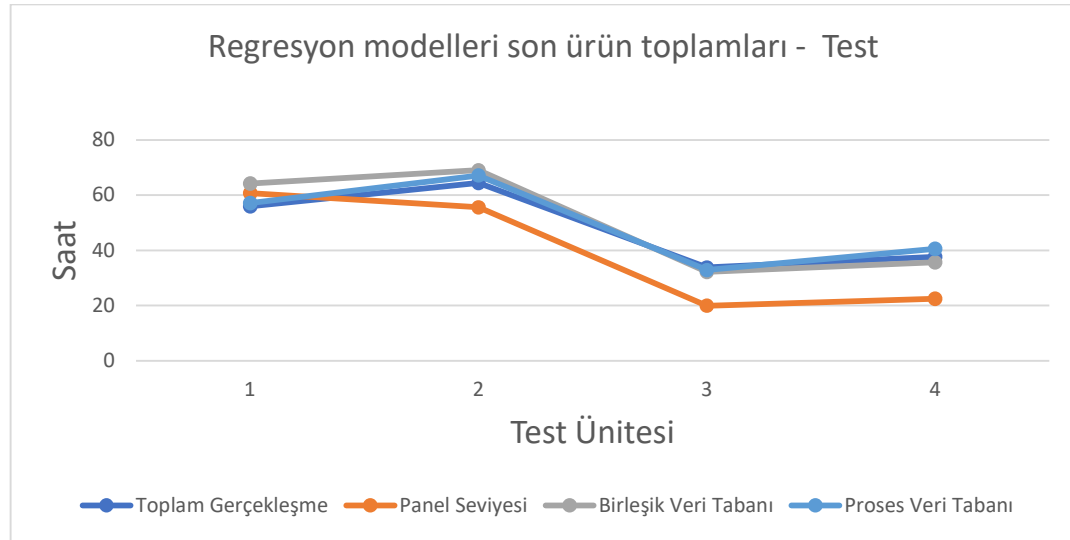
### 3.6. MODELLERİN KARŞILAŞTIRILMASI

Bu bölümde “son ürün”, “birleşik veri tabanı” ve “proses ver tabanı” yaklaşımları kapsamında oluşturulan modellerin performansları karşılaştırılacaktır. Model tahminleri son ürün bazında toplanarak aynı bazdaki gerçekleştirmeler ile tahmin performansları özetlenecektir. Bu performanslar hem eğitim hem de test verileri için değerlendirilecektir. Böylelikle modellerin kendi içinde eğitim ve test

performansları da değerlendirilebilecektir. Modellerin tahmin performansları için  $R^2$ , MAPE ve MSE verileri tablolarda raporlanacaktır. Modellerin birbirleri ile karşılaştırılabilmesi için performans kriteri olarak MAPE kullanılacaktır. Her tahmin modeli için, test verilerine ait performans kriterleri daha kolay yorumlanabilmesi için bir grafikte özetlenecektir. Bu grafiklerde  $R^2$  değerleri sağ ekseninde, diğer performans kriter değerleri sol ekseninde gösterilecek olup, sol eksen logaritmik, sağ eksen doğrusal ölçekli olacaktır.

### 3.6.1. Regresyon Modelleri

Tez kapsamında oluşturulan regresyon modellerinin test verileri için son ürün toplamları bazında gerçekleşme ve tahmin değerleri karşılaştırması Şekil 13'de görülebilir.



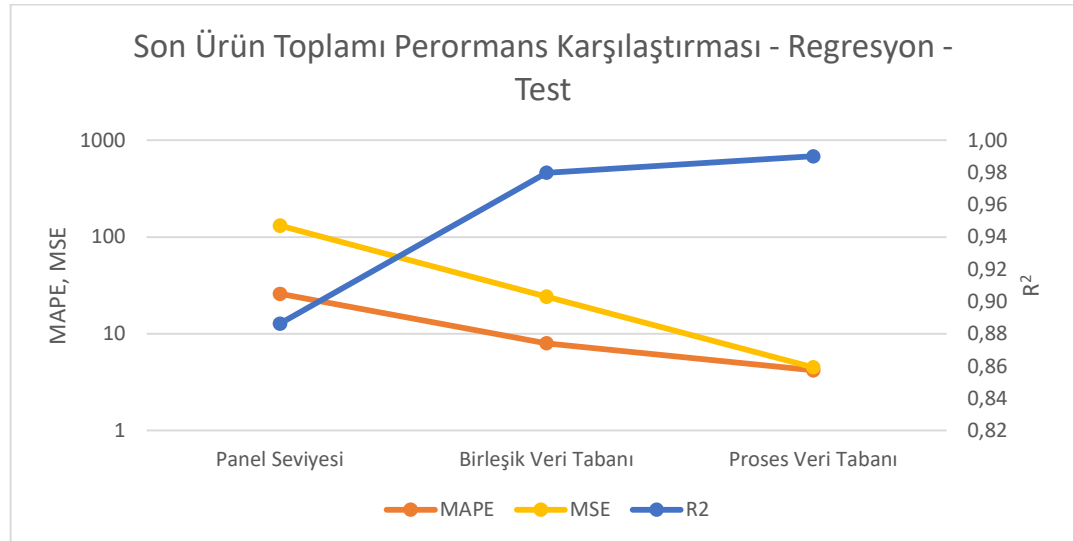
Şekil 13 - Regresyon modelleri son ürün toplamları - Test

Tez kapsamında oluşturulan regresyon modellerinin son ürün toplamları bazında performans değerleri Tablo 43'te görülebilir.

	Eğitim			Test		
	Son ürün (TY1-S-SÜ)	Birleşik Veri Tabanı (TY2-D-BV)	Proses veri tabanları (TY3-D-PV)	Son ürün (TY1-S-SÜ)	Birleşik Veri Tabanı (TY2-D-BV)	Proses veri tabanları (TY3-D-PV)
<b>R2</b>	0,98	0,98	1,00	0,89	0,98	0,99
<b>MSE</b>	112	120	37	131	24	4
<b>MAPE</b>	24	13	7	26	8	4

Tablo 43 - Regresyon Modelleri Son Ürün Toplamları Bazında Performans Verileri

Son ürün toplamaları bazında test verileri için performans değerleri Şekil 14'te özetlenmiştir.



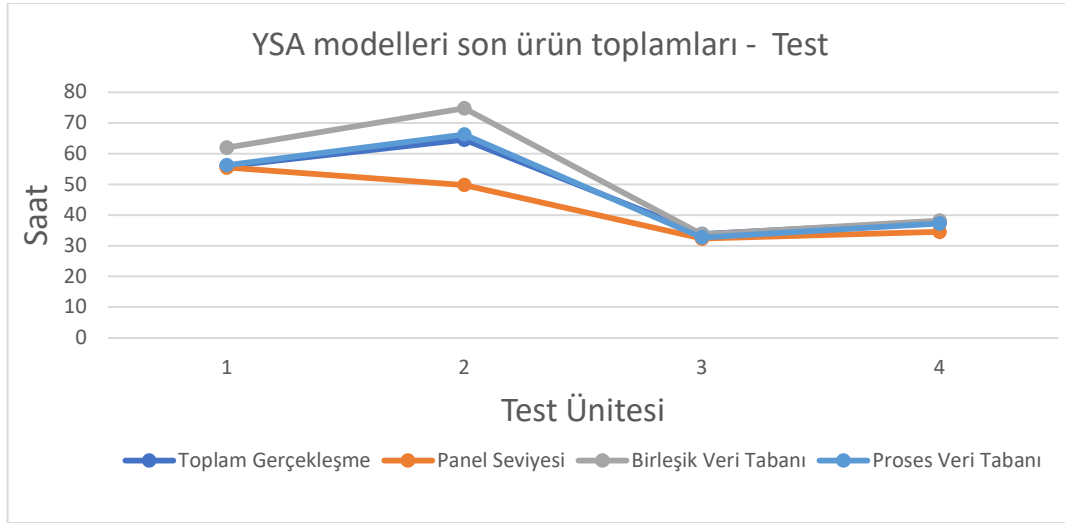
Şekil 14 - Son ürün toplamaları bazında test verileri performans değerleri

Grafikte sağa doğru gidildiğinde tahmin performansının iyileşmesi beklentisine paralel sonuçlar alındığı Şekil 14'te kolaylıkla gözlemlenebilir. Sonuç olarak regresyon modellerinde, modelin detay seviyesi arttıkça tahmin performansı da artmaktadır.



### 3.6.2. Yapay Sinir Ağları Modelleri

Tez kapsamında oluşturulan YSA modellerinin test verileri için son ürün toplamları bazında gerçekleşme ve tahmin değerleri karşılaştırması Şekil 15'te görülebilir.



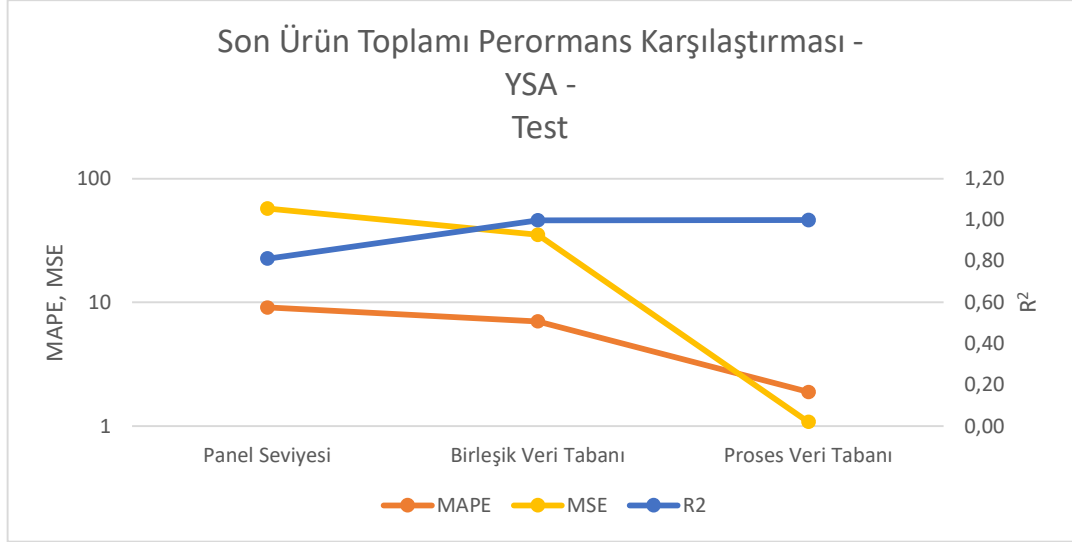
Şekil 15 - YSA Modelleri Son Ürün Toplamları - Test

Tez kapsamında oluşturulan YSA modellerinin son ürün toplamları bazında performans değerleri Tablo 44'te görülebilir.

	Eğitim			Test		
	Son ürün (TY1-S-SÜ)	Birleşik Veri Tabanı (TY2-D-BV)	Proses veri tabanları (TY3-D-PV)	Son ürün (TY1-S-SÜ)	Birleşik Veri Tabanı (TY2-D-BV)	Proses veri tabanları (TY3-D-PV)
<b>R2</b>	1,00	1,00	1,00	0,81	1,00	1,00
<b>MSE</b>	50	2	85	57	35	1
<b>MAPE</b>	18	3	5	9	7	2

Tablo 44 - YSA Modelleri Son Ürün Toplamları Bazında Performans Verileri

Son ürün toplamları bazında test verileri için performans değerleri Şekil 16'da özetlenmiştir.

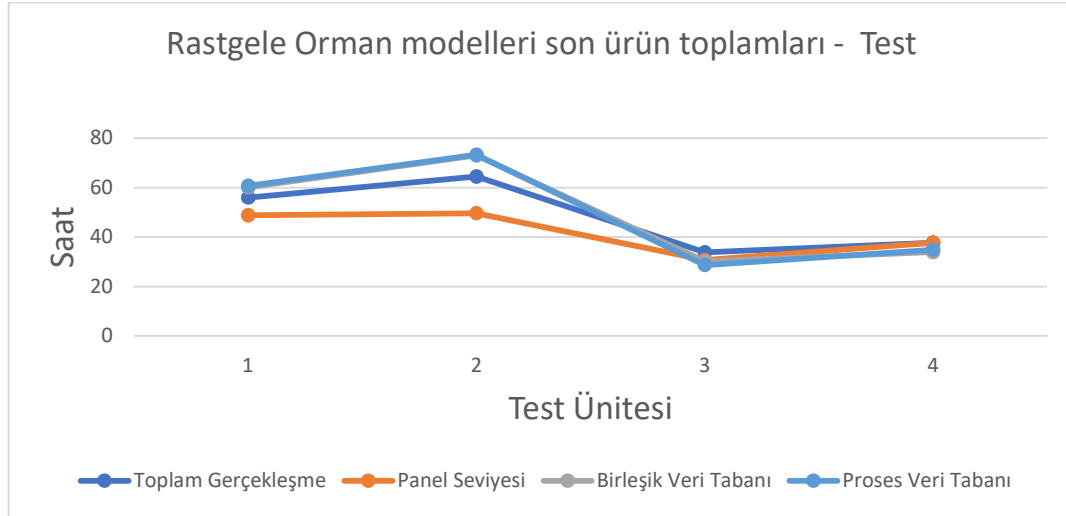


Şekil 16 - Son ürün toplamları bazında test verileri performans değerleri

Grafikte sağa doğru gidildiğinde tahmin performansının iyileşmesi görülmektedir. Tahmin modellerinin detaylandırılması ile daha iyi tahmin performans elde edilebildiği görülmektedir.

### 3.6.3. Rastgele Orman Modelleri

Tez kapsamında oluşturulan RO modellerinin test verileri için son ürün toplamları bazında gerçekleşme ve tahmin değerleri karşılaştırması Şekil 17'de görülebilir.



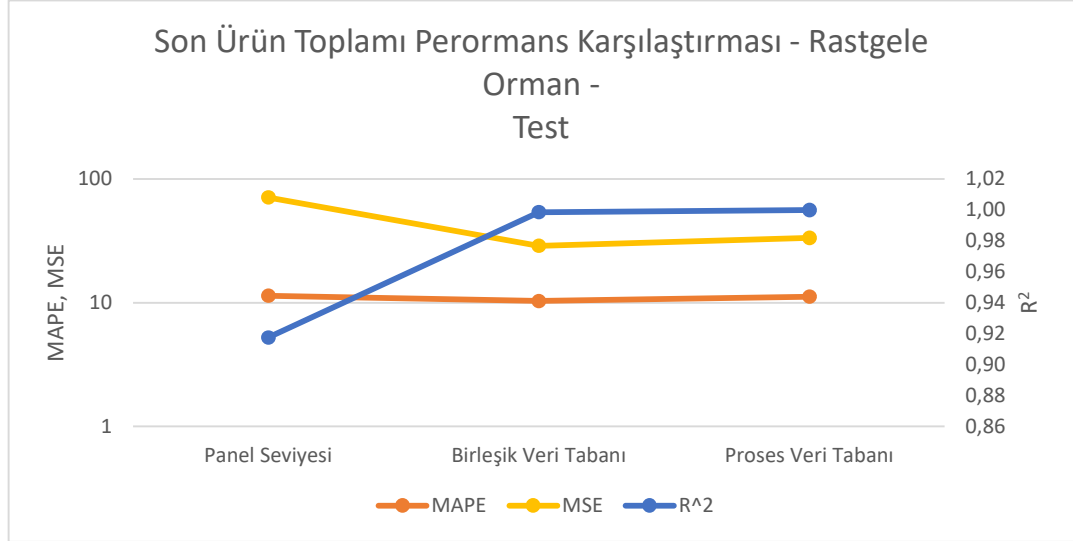
Şekil 17 - RO modelleri son ürün toplamları - Test

Tez kapsamında oluşturulan RO modellerinin son ürün toplamları bazında performans değerleri Tablo 45'te görülebilir.

	Eğitim			Test		
	Son ürün (TY1-S-SÜ)	Birleşik Veri Tabanı (TY2-D-BV)	Proses veri tabanları (TY3-D-PV)	Son ürün (TY1-S-SÜ)	Birleşik Veri Tabanı (TY2-D-BV)	Proses veri tabanları (TY3-D-PV)
<b>R<sup>2</sup></b>	1,00	1,00	0,99	0,92	1,00	1,00
<b>MSE</b>	556	8	22	71	29	33
<b>MAPE</b>	9	5	9	11	11	11

Tablo 45 - RO Modelleri Son Ürün Toplamları Bazında Performans Verileri

Son ürün toplamları bazında test verileri için performans değerleri Şekil 18'de özetlenmiştir.



Şekil 18 - Son Ürün Toplamları Bazında Test Verileri Performans Değerleri

Her üç yaklaşım için MAPE değerleri birbirlerine yakın olmakla birlikte R<sup>2</sup> ve MSE değerlerine göre birleşik veri tabanı ve Proses veri tabanları modelleri, son ürün modeline göre daha başarılı olmuşlardır.

### 3.7. KARŞILAŞTIRMA ÖZETİ VE YORUMLAR

Son ürün, birleşik veri tabanı ve proses veri tabanları yaklaşımlarının test verileri için performans sonuçları Tablo 46'da özetlenmiştir. Farklı seviyedeki ve farklı büyüklükteki hataları yüzdesel olarak karşılaştırmaya olanak sağladığı ve aynı zamanda maliyet tahmin literatürü kapsamında en çok kullanılan performans ölçme ölçütü olduğundan özet için MAPE değerleri kullanılmıştır. "Birleşik veri tabanı" ve "proses veri tabanları" modelleri ile detay parça seviyesinde üretim süreleri tahminleri yapılmış ve bu süreler ürün ağacına göre son ürün bazında toplanarak son ürünlerin toplam üretim saatleri tahmin edilmiştir. Bu süreler ile son ürünlerin toplam üretim süreleri gerçekleştirmeleri kullanılarak MAPE değerleri hesaplanmış ve böylece modellerin performansları son ürün bazında ölçülmüştür. Son ürün tahmin yaklaşımı ile detay parça tahmin yaklaşımlarının karşılaştırılabilmesi için bu MAPE değerleri tabloda "son ürün seviyesi" satırlarında belirtilmiştir.

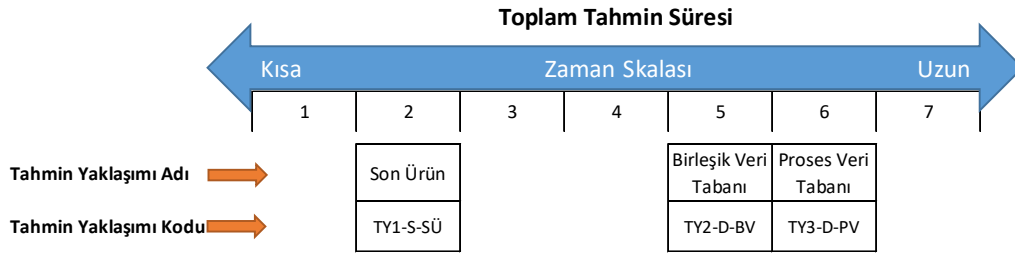
		Son ürün (TY1-S-SÜ)	Detay parça seviyesi	
			Birleşik veri tabanı (TY2-D-BV)	Proses veri tabanları (TY3-D-PV)
Regresyon	Son ürün seviyesi	26	8	4
	Detay parça seviyesi		137	83
Yapay Sinir Ağları	Son ürün seviyesi	9	7	2
	Detay parça seviyesi		74	23
Rastgele Orman	Son ürün seviyesi	11	11	11
	Detay parça seviyesi		40	45

Tablo 46 - Son Ürün ve Detay Seviye Tahmin Performansları Özeti (Test Verileri İçin MAPE)

Son ürün seviyesinde değerlendirildiğinde son ürün, birleşik veri tabanı ve proses veri tabanları yaklaşımlarının her üçü için de en iyi tahmin performansını YSA göstermiştir. Regresyon ve YSA metotları için tahmin yaklaşımı detaylandıkça (tabloda sağa doğru gidildikçe) performansın da fark edilir şekilde iyileştiği gözlemlenmektedir. RO için ise bu durum geçerli olmayıp tahmin performansında anlamlı bir değişiklik olmamaktadır.

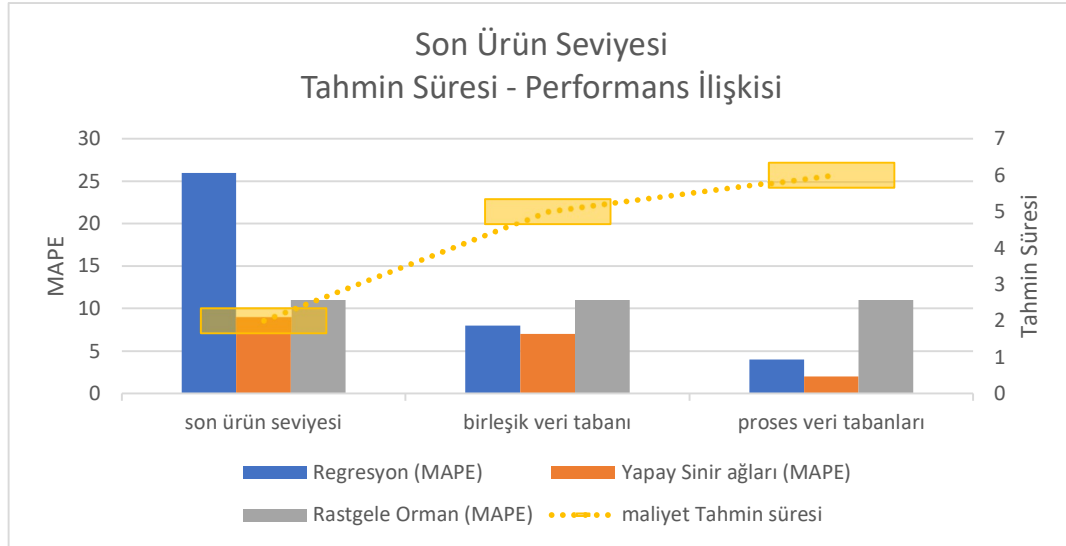
Detay parça seviyesinde ise birleşik veri tabanı yaklaşımı için RO, proses veri tabanları yaklaşımı için ise YSA modeli daha başarılı olmuştur.

İçinde buldukları koşullara göre firmalar, tahmin hazırlama süresi ve tahmin hassasiyeti dengesini gözeterek hangi tahmin yaklaşımını seçmeleri gerektiğine karar vermeleri gerekir. Burada basitçe iki parametreye indirgenmiş karar probleminin hassasiyet parametresi için Tablo 46'da yer alan verilerden yararlanılabilir. Süre parametresi için de örnek teşkil etmesi adına Şekil 19 oluşturulmuştur. Bu şekilde her üç tahmin yaklaşımı için tahmin oluşturma süresi 7'li bir zaman skalası üzerinde örnek olarak belirtilmiştir. Gerçekte bu süreler teklif kapsamının büyüklüğü, veri setinin içeriği ve detay seviyesi gibi birçok parametreye bağlı olarak değişebilmektedir.



Şekil 19 - Teklif Hazırlama Zaman Skalası

Bu tez çalışması kapsamında oluşturulan üç tahmin yaklaşımı için son ürün seviyesinde MAPE değerleri ve bu yaklaşımların tahmin skalasındaki süreleri Şekil 20'de gösterilmektedir.

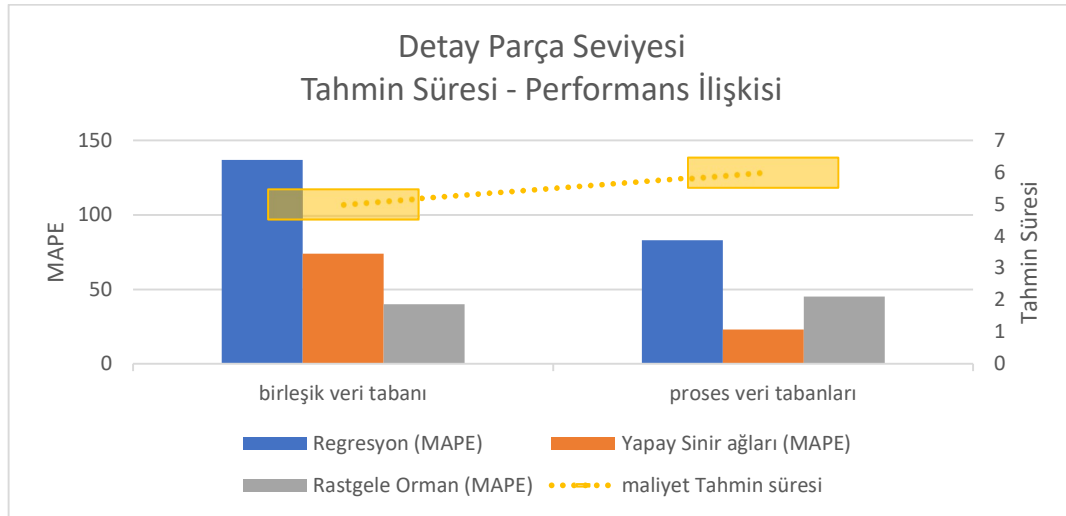


Şekil 20 – Son Ürün Seviyesi İçin MAPE - Tahmin Süresi İlişkisi

Firmalar, çalışılacak teklife özgün diğer kısıtları da göz önüne alarak Şekil 20'de özetlenen süre ve hassasiyet kısıtlarına göre hangi tahmin yaklaşımını ve metodu kullanacağına karar verebilirler.

Bazen teklif kapsamı montaj halindeki son ürün seviyesinde olmayıp sadece çeşitli detay parçalardan oluşabilirler. Bu durumda son ürün tahmin yaklaşımı kullanılamayacağından birleşik veri tabanı ya da proses veri tabanları yaklaşımlarından birini seçmek gerekecektir. Bu gibi durumlar için karar vericinin teklife özgün diğer durumları da değerlendirerek Şekil 21'de özet olarak belirtilen tahmin süresi ve tahmin hassasiyetine göre bir karar vermesi gerekir. Bu durumda model ve yaklaşımlar için detay parça seviyesindeki MAPE değerlerini kullanmak gerekecektir. Şekil 21'e göre regresyon ve YSA için Proses veri tabanları kullanımında, birleşik veri tabanına göre tahmin süresi uzamasına rağmen tahmin hassasiyeti artmaktadır. RO için ise bu durum geçerli olmayıp

tahmin hassasiyetinde anlamlı bir deęişiklik olmamaktadır. Buna göre Proses veri tabanları yaklaşımı uygulayacak kadar vakit olmadığında birleşik veri tabanı yaklaşımı ile RO metodunu kullanmak daha uygun görünmektedir. Yeterli süre olduğu durumda ise Proses veri tabanları yaklaşımı ile YSA metodunu kullanmak daha uygun görünmektedir.



Şekil 21 – Detay Parça Seviyesi İçin MAPE - Tahmin Süresi İlişkisi

Yukarıda yaklaşım ve tahmin metodunun seçimi için sadece süre ve hassasiyet olarak iki temel kriterden bahsedilmesine rağmen pratikte durum biraz daha karmaşık olabilmektedir. Karar verme sürecince bir çok kriter göz önüne alınmak durumunda kalınabilir. Bu kısıtlardan bazıları Tablo 47’de özetlenmiştir. Bu kısıtlardan “veri kısıtı” ve “teklif isterleri kısıtı”, yaklaşım seviyesinin son ürün ya da detay parça seviyesinde olmasının belirlenmesinde dikkate alınacak öncelikli kısıtlardır. “Zaman kısıtı” hız parametresini etkilemekte olup “teklifin gerçekleşme potansiyeli” ve “teklifin stratejik önemi” kısıtları doğruluk parametresi için önemlidir. “Teklifin değerlendirilmesi” kısıtı ise teklif kapsamında kullanılacak tahmin metodunun seçimini öncelikli olarak etkilemektedir.



<b>Kısıt</b>	<b>Açıklama</b>
Veri Kısıtı	Teklif kapsamındaki veriler detay parça ya da son ürün bazında olabilir.
Zaman Kısıtı	Teklif içeriğinin büyüklüğü ve miadı, mevcut diğer tekliflerin iş yoğunluğu
Teklif İsterleri Kısıtı	Teklif isteyen firmanın detay parça bazında ya da son ürün bazında maliyet bilgisi talebi
Teklifin Gerçekleşme Potansiyeli	Bazı teklifler pazar araştırması amaçlı olabilir, firmalar arası stratejik ilişki nedeniyle olabilir.
Teklifin Stratejik Önemi	Teklif kapsamındaki iş kapsamının, firmanın stratejik planına uygun olup olmaması
Teklifin Değerlendirilmesi	Üst yönetim ya da müşteri tarafından tahmin ilişkilerinin (parametrik formüllerin) değerlendirilmek istenmesi

Tablo 47 - Maliyet Tahmin Sürecinde Olası Kısıtlar

Bu çalışmanın karar vericiye teklif çalışmasına özgü kısıtlar (Tablo 47) ile tahmin yaklaşımı ve tahmin metotları (Tablo 46 ve Şekil 19) ile ilgili kısıtları bir arada değerlendirerek hangi tahmin yaklaşımı ve metodunu seçmesi gerektiği konusunda yol göstermesi amaçlanmıştır.

## SONUÇ

Özellikle havacılık ve uzay sanayiinde, firmaların teklif verme süreçlerinde hızlı ve doğru maliyet tahmini yapabilmeleri, firmaların rekabetçiliği ve sürdürülebilirlikleri açısından oldukça önemlidir. Bu durumda firmalar kabiliyetlerindeki farklı tahmin yöntemlerinden duruma en uygununu seçmek zorundadır. Bu çalışma kapsamında firmaların hız ve hassasiyet dengesini gözleterek maliyet tahmini yapabilme becerileri ile ilgili olarak üç araştırma sorusuna cevap aranmıştır:

1. Araştırma Sorusu: Bir teklif verme sürecinde içinde bulunulan koşullara göre tahmin süresi ile tahmin hassasiyeti dengesini dikkate alarak nasıl bir tahmin yaklaşımı seçmek gerekir,
2. Araştırma Sorusu: Makine öğrenmesi yöntemlerinin tahmin performansı geleneksel yöntemlere göre üstün müdür,
3. Araştırma Sorusu: Öğrenme eğrisine göre üretim saatlerine veri dönüşümü uygulanmasının, tahmin modeli performansına etkisi var mıdır.

Bu araştırma sorularına cevap bulabilmek için “son ürün”, “birleşik veri tabanı” ve “Proses veri tabanları” yaklaşımları olmak üzere 3 farklı detayda tahmin yaklaşımları oluşturulmuştur. Her tahmin yaklaşımı kapsamında regresyon, yapay sinir ağları ve rastgele orman metotları kullanılarak modellerin tahmin performansları kıyaslanmıştır.

Son ürün tahmin yaklaşımında son ürünler bazında toplam üretim saati tahmini için modeller oluşturulmuştur. Bu yaklaşımda model kapsamındaki veriler son ürün seviyesinde olduğundan veri ve parametre sayısı kısıtlı kalmakta, bu da tahmin sürecini kısaltırken tahmin hassasiyetini de azaltmaktadır. Birleşik veri tabanı yaklaşımında detay parçalara ait verilerin tümünü içeren tek bir veri tabanı oluşturularak farklı proses tipleri için ortak tek bir model oluşturulmuştur. Bu yaklaşım için, son ürün yaklaşımından daha uzun bir süreye gerek duyulmaktadır (Şekil 19). Proses veri tabanları yaklaşımı kapsamında ise detay parça veri

tabanı, farklı imalat proseslerine göre 8 alt veri tabanına bölünerek her bir proses alt veri tabanı için farklı tahmin modelleri oluşturulmuştur. Bu yaklaşım için de birleşik veri tabanı yaklaşımına göre daha uzun bir süreye gerek duyulmaktadır.

Regresyon ve YSA metotları için tahmin yaklaşımı detaylandıkça yani tahmin için gereken süre uzadıkça tahmin performansının da fark edilir şekilde iyileştiği gözlemlenmektedir. RO için ise bu durum geçerli olmayıp tahmin performansında anlamlı bir değişiklik olmamaktadır.

Oluşturulan modeller son ürün seviyesinde değerlendirildiğinde YSA'nın her yaklaşım için regresyon ve RO metotlarına göre daha başarılı olduğu görülmektedir (Tablo 46). Regresyon, son ürün yaklaşımında RO'ya göre daha kötü performans göstermesine rağmen birleşik veri tabanı ve Proses veri tabanları yaklaşımlarında daha başarılı tahmin performansı göstermektedir. Modeller detay parça seviyesinde değerlendirildiğinde ise birleşik veri tabanı yaklaşımında RO'nun, Proses veri tabanları yaklaşımında ise YSA'nın daha başarılı olduğu görülmektedir. Sonuç olarak makine öğrenmesi yöntemlerinin, regresyona göre genel olarak daha başarılı olduğu görülmektedir.

Üçüncü araştırma sorusuna cevap bulabilmek için örnek olarak birleşik veri tabanı modelleri, üretim sürelerinin öğrenme eğrisine göre veri dönüşümü yapılmış haliyle tekrarlanmış ve tahmin performansları ilk haliyle kıyaslanmıştır. Öğrenme eğrisine göre veri dönüşümü yapılarak oluşturulan modellerden yapay sinir ağları modelinde gözle görülür bir iyileşme olurken diğer modellerde kayda değer bir performans değişikliği olmamıştır. Bu durum aynı zamanda YSA'nın, verilere ne kadar duyarlı bir metot olduğunu ortaya koymaktadır. Öğrenme eğrisine göre veri dönüşümünün etkisinin bu çalışmada sınırlı kalmasının nedeni çalışma kapsamındaki verinin dağılımı ile ilgilidir. Parçaların üretim süreleri incelendiğinde parçaların %80'inin 0,5 saatten daha az, kalan %20'sinin ise 0,5 ile 17 saat aralığında olduğu görülmektedir. Çok küçük parçaların gerçekleşen üretim saatlerinin belirlenmesinde yapılan ölçüm hataları yüzdesel olarak büyük olabilmektedir. Aynı zamanda öğrenme eğrisi etkisinin, el işçiliğine daha çok bağımlı olan parçalarda daha çok ortaya çıkması beklenir. El işçiliği yüksek parçaların üretim saatleri de görece olarak daha çok olmaktadır. Bu tür parçalar

veri setinde azınlıkta olduğundan bu çalışma kapsamında öğrenme eğrisi etkisinin de sınırlı kaldığı değerlendirilmektedir.

Yukarıda özetlenen sonuçlar incelendiğinde YSA metodunun, diğer metotlarla karşılaştırıldığında genel olarak daha iyi tahmin performansı gösterdiği söylenebilir. Buna rağmen YSA performansı, veriye ve algoritmanın parametrelerine göre çok değişkenlik gösterebilmektedir. RO metodu bu çalışma kapsamında iyi bir tahmin performansı sergilemekle birlikte bu çalışmada ortaya çıkan en önemli özelliği tahmin performansındaki tutarlılığıdır. Veri sayısının ve bağımsız değişken sayısının değişmesi ile RO tahmin performansını çok fazla değiştirmemektedir. Regresyonun tahmin performansı da veriye bağlı olarak büyük değişim gösterebilmektedir. Regresyonun, makine öğrenmesi yöntemlerine göre en önemli üstünlüğü girdiler ile tahmin sonuçlarının nasıl değiştiğinin izlenebilir olmasıdır. Makine öğrenmesi yöntemlerinde ise bu mümkün olmadığından bu yöntemler “kara kutu” olarak tanımlanmaktadır.

Sonuç olarak bu çalışmanın karar vericiye teklife özgü kısıtlar ile tahmin yaklaşımı ve kullanılacak metotlar ile ilgili kısıtları bir arada değerlendirerek hangi tahmin yaklaşımı ve metodunu seçmesi gerektiği konusunda yol göstermesi amaçlanmıştır.

Bu çalışmada elde edilen sonuçlar ile teklif çalışmasına özgü kısıtları (Tablo 47) da içeren bir karar destek sistemi tasarlanarak bu çalışma genişletilebilir.

Öğrenme eğrisi, el işçiliğinin daha yoğun olduğu üretimlerde etkisini daha çok göstermektedir. Öğrenme eğrisinin model performansı üzerindeki etkisini daha iyi gözlemleyebilmek için kompozit üretimi gibi el işçiliği daha yoğun parçaların çoğunlukta olduğu ürünler ile bu çalışma tekrarlanarak öğrenme eğrisinin etkisi incelenebilir.

Bu çalışmanın amaçlarından birisi de makine öğrenmesi yöntemlerinin geleneksel yöntemlere göre daha başarılı olup olmadığının araştırılmasıdır. Bu amaçla literatürdeki temel başvuru yöntemlerinden birisi olan yapay sinir ağları ve nispeten yeni bir yöntem olan ve maliyet tahmin literatüründe yeterince yer verilmeyen yöntemlerden birisi olan rastgele orman seçilmiştir. Bu çalışma

kapsamı daha farklı makine öğrenmesi yöntemleri ile tekrar edilerek farklı makine öğrenmesi yöntemlerinin hangi durumlarda daha başarılı olduğu ortaya konulabilir.

Makine öğrenmesi yöntemleri kapsamında veri az olduğu ya da dağınık olduğu durumlarda oluşturulan tahmin modellerinde aşırı uyumu engellemek için sıklıkla başvurulan yöntemlerden birisi de bootstrap yöntemidir. Bu çalışma kapsamında gerekli görülen yerlerde bu yöntem ile veriler çoğaltılarak modeller oluşturulmuştur. Bu amaçla başka yöntemlerin kullanılmasının model performansını etkileyip etkilemediği sonraki benzer çalışmalarda incelenebilir.

## KAYNAKÇA

- Angelis, L., Stamelos, I., (2000). A simulation tool for efficient analogy based cost estimation. *Empirical Software Engineering*, 5(1), 35-68
- Atia, M. R., Khalil, J., & Mokhtar, M. (2017). A Cost Estimation Model for Machining Operations; an YSA Parametric Approach. *Journal Of Al Azhar University Engineering Sector*, 12(44), 878-885.
- Banazadeh, A., & Jafari, M. H. (2013). A heuristic complexity-based method for cost estimation of aerospace systems. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 227(11), 1685-1700.
- Ben-Arieh, D., & Qian, L. (2003). Activity-based cost management for design and development stage. *International Journal of Production Economics*, 83(2), 169-183.
- Breiman L., (2001). Bagging predictors. *Machine Learning*, 24(2),123–140.
- Breiman, L., (2001). Random Forests. *Machine Learning*, 45(1), 5 – 32.
- Boer, J. F., & Stevens, J. (2006). Helicopter life cycle cost reduction through pre-design optimisation. *Proceedings of the 25th International Congress of the Aeronautical Sciences, ICAS*
- Castagne, S., Curran, R., Rothwell, A., Price, M., Bénard, E., & Raghunathan, S. (2008). A generic tool for cost estimating in aircraft design. *Research in Engineering Design*, 18(4), 149-162.
- Cavalieri, S., Maccarrone, P., & Pinto, R. (2004). Parametric vs. neural network models for the estimation of production costs: A case study in the automotive industry. *International Journal of Production Economics*, 91(2), 165-177.

- Chen, X., Yi, M., & Huang, J. (2020). Application of a PCA-ANN Based Cost Prediction Model for General Aviation Aircraft. *IEEE Access*, 8, 130124-130135.
- Collopy, P. D., & Eames, D. J. (2001). Aerospace manufacturing cost prediction from a measure of part definition information (No. 2001-01-3004). *SAE Technical Paper*.
- Ciurana, J., Quintana, G., & Garcia-Romeu, M. L. (2008). Estimating the cost of vertical high-speed machining centres, a comparison between multiple regression analysis and the neural networks approach. *International Journal of Production Economics*, 115(1), 171-178.
- Curran R., Raghunathan, S., Price, M. (2004). Review of aerospace engineering cost modelling: The genetic causal approach. *Progress in Aerospace Sciences* 40(8), 487–534.
- Deng, S., & Yeh, T. H. (2010). Applying least squares support vector machines to the airframe wing-box structural design cost estimation. *Expert Systems with Applications*, 37(12), 8417-8423.
- Deng, S., & Yeh, T. H. (2011). Using least squares support vector machines for the airframe structures manufacturing cost estimation. *International Journal of Production Economics*, 131(2), 701-708.
- Dimililer, K., & Zarrouk, S. (2017). ICSPi: Intelligent classification system of pest insects based on image processing and neural arbitration. *Applied Engineering in Agriculture*, 33(4), 453.
- Duverlie, P., Castelain, J.M. (1999). Cost estimation during design step: parametric method versus case based reasoning method. *International Journal of Advanced Manufacturing Technology*, 15(12), 895–906.
- Eck, D., Brundick, B., Fettig, T., Dechoretz, J. (2009). Parametric estimating handbook. *The International Society of Parametric Analysis (ISPA)*.

- Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems?. *The Journal of Machine Learning Research*, 15(1), 3133-3181.
- Finger, D. F., Götten, F., Braun, C., & Bil, C. (2019). Cost estimation methods for hybrid-electric general aviation aircraft. *Asia-Pacific International Symposium on Aerospace Technology (APISAT 2019)* (pp. 265-277).
- Garza, J., Rouhana, K. (1995). Neural network versus parameter-based application in cost estimating. *Cost Engineering* 37 (2), 14–18.
- Hagnell, M. K., & Åkermo, M. (2015). A composite cost model for the aeronautical industry: *Methodology and case study*. *Composites Part B: Engineering*, 79, 254-261.
- Hebb, D.O., (1949). *The Organization of Behavior: A Neuropsychological Theory*. Wiley, New York.
- Kumar, J., & Singh, A. K. (2018). Workload prediction in cloud using artificial neural network and adaptive differential evolution. *Future Generation Computer Systems*, 81, 41-52.
- Li, Y., Dong, W., Yang, Q., Zhao, J., Liu, L., & Feng, S. (2018). An automatic impedance matching method based on the feedforward-backpropagation neural network for a WPT system. *IEEE Transactions on Industrial Electronics*, 66(5), 3963-3972.
- Lin, T., Lee, J. W., & Bohez, E. L. J. (2012). New integrated model to estimate the manufacturing cost and production system performance at the conceptual design stage of helicopter blade assembly. *International Journal of Production Research*, 50(24), 7210-7228.
- Liu, H., Guo, K., Zhang, Z., Yu, D., Zhang, J., Ning, P., Niu, P. (2017). High-power LED Photoelectrothermal analysis based on backpropagation artificial neural networks. *IEEE Transactions on Electron Devices*, 64(7), 2867-2873.




- Liu, W., Yang, C., & Zhou, X. (2018). A network quotation framework for customised parts through rough requests. *International Journal of Computer Integrated Manufacturing*, 31(12), 1220-1234.
- Loyer, J. L., Henriques, E., Fontul, M., & Wiseall, S. (2016). Comparison of Machine Learning methods applied to the estimation of manufacturing cost of jet engine components. *International Journal of Production Economics*, 178, 109-119.
- LUKIĆ, D., ĐURĐEV, M. M. S. B. M., & ANTIĆ, J. V. A. (2016). Manufacturing cost estimation in the conceptual process planning. *Machine Design*, 8(3), 83-90
- Mandolini, M., Favi, C., Germani, M., Marconi, M., & Raffaelli, R. (2019). An Analytical Cost Estimation Approach for Generic Sheet Metal 3D Models. *Computer-Aided Design & Applications*, 16(5), 936-950
- Molcho, G., Cristal, A., & Shpitalni, M. (2014). Part cost estimation at early design phase. *CIRP Annals*, 63(1), 153-156.
- Mourikas, K., King, J., Nelson, D., (2017). Machine Learning Approach to Cost Analysis. *Proceedings of the 2017 ICEAA Conference*.
- Niazi, A., Dai, J. S., Balabani, S., & Seneviratne, L. (2006). Product cost estimation: Technique classification and methodology review. *Journal of Manufacturing Science and Engineering, Transactions of the ASME* 128, 563–575
- McCulloch, W.S., Pitts, W.A., (1943). A logical calculus of the ideas imminent in nervous activity. *Bulletin of Mathematics and Biophysics* 5, 115–133.
- Mörtl, M., & Schmied, C. (2016). Design for Cost—A Review of Methods, Tools and Research Directions. *Journal of the Indian Institute of Science*, 95(4), 379-404.
- Mustapha, H., & Abdelwahed, N. (2019). Investigating the use of random forest in software effort estimation. *Procedia Computer Science*, 148, 343-352.

- Özcan, B., Kumru, P. Y., & Fiğlalı, A. (2018). Forecasting operation times by using Artificial Intelligence. *International Advanced Researches and Engineering Journal*, 2(2), 109-116.
- Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(3), e1301.
- Rush, C., Roy, R. (2001). Expert judgment in cost estimating: modelling the reasoning process. *Concurrent Engineering*, 9(4), 271-284.
- Satapathy, S. M., Acharya, B. P., & Rath, S. K. (2016). Early stage software effort estimation using random forest technique based on use case points. *IET Software*, 10(1),10-17.
- Schubel, P. J. (2012). Cost modelling in polymer composite applications: case study – analysis of existing and automated manufacturing processes for a large wind turbine blade. *Composites Part B: Engineering* 43(3), 953– 960.
- Sekeroglu, B., & Dimililer, K. (2020). Review and analysis of hidden neuron number effect of shallow backpropagation neural networks. *Neural Network World*, 30(2), 97-112.
- Shehab, E. M., & Abdalla, H. S. (2002). A design to cost system for innovative product development. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 216(7), 999-1019.
- Vujicic, T., Matijevic, T., Ljucovic, J., Balota, A., & Sevarac, Z. (2016). Comparative analysis of methods for determining number of hidden neurons in artificial neural network. *In Central European Conference on Information and Intelligent Systems*, 219-223.
- Wang, Q., & Stockton, D. (2001). Cost model development using artificial neural networks. *Aircraft Engineering And Aerospace Technology*, 73(6), 536-541.

- Wang, H.S. (2007). Application of BPN with feature-based models on cost estimation of plastic injection products. *Computers and Industrial Engineering* 53, 79–94.
- Watson, P., Curran, R., Murphy, A., & Cowan, S. (2006). Cost estimation of machined parts within an aerospace supply chain. *Concurrent engineering*, 14(1), 17-26.
- Xiaonan, C. H. E. N., HUANG, J., & Mingxu, Y. I. (2019). Development cost prediction of general aviation aircraft projects with parametric modeling. *Chinese Journal of Aeronautics*. 32(6), 1465-1471.
- Xie, N. M., Yin, S. M., & Hu, C. Z. (2017). Estimating a civil aircraft's development cost with a GM (1, N) model and an MLP neural network. *Grey Systems: Theory and Application*. 7(1), 2-18
- Wright, T. P. (1936). Factors affecting the cost of airplanes. *Journal Of The Aeronautical Sciences*, 3(4), 122-128.
- Verlinden, B., Duflou, J. R., Collin, P., & Cattrysse, D. (2008). Cost estimation for sheet metal parts using multiple regression and artificial neural networks: A case study. *International Journal of Production Economics*, 111(2), 484-492..
- Yuan, J., & Yu, S. (2013). Privacy preserving back-propagation neural network learning made practical with cloud computing. *IEEE Transactions on Parallel and Distributed Systems*, 25(1), 212-221.
- Zhaodong, H., Rongxuan, L., & Jing, J. (2015). Development and Production Costs Estimating For Aviation Equipment Based on Uncertainty Design. *Procedia Engineering*, 99, 143-149.
- Zhang, Y. F., Fuh, J. Y., & Chan, W. T. (1996). Feature-based cost estimation for packaging products using neural networks. *Computers in Industry*, 32(1), 95-113.

Zhao, X., Verhagen, W. J., & Curran, R. (2015). Estimation of aircraft component production cost using knowledge based engineering techniques. *Advanced Engineering Informatics*, 29(3), 616-63

## EK 1. Etik Kurul Muafiyet Formu

 <p><b>HACETTEPE ÜNİVERSİTESİ</b> <b>SOSYAL BİLİMLER ENSTİTÜSÜ</b> <b>TEZ ÇALIŞMASI ETİK KOMİSYON MUAFİYETİ FORMU</b></p>
<p><b>HACETTEPE ÜNİVERSİTESİ</b> <b>SOSYAL BİLİMLER ENSTİTÜSÜ</b> <b>İŞLETME ANABİLİM DALI BAŞKANLIĞI'NA</b></p> <p style="text-align: right;">Tarih: 16/06/2021</p> <p>Tez Başlığı: Atölye Tipi Üretim İçin Makine Öğrenmesi Yöntemleri İle Üretim Saati Tahmini: Havacılık Ve Savunma Sanayii Uygulaması</p> <p>Yukarıda başlığı gösterilen tez çalışmam:</p> <ol style="list-style-type: none"> <li>1. İnsan ve hayvan üzerinde deney niteliği taşımamaktadır,</li> <li>2. Biyolojik materyal (kan, idrar vb. biyolojik sıvılar ve numuneler) kullanılmasını gerektirmemektedir.</li> <li>3. Beden bütünlüğüne müdahale içermemektedir.</li> <li>4. Gözlemsel ve betimsel araştırma (anket, mülakat, ölçek/skala çalışmaları, dosya taramaları, veri kaynakları taraması, sistem-model geliştirme çalışmaları) niteliğinde değildir.</li> </ol> <p>Hacettepe Üniversitesi Etik Kurullar ve Komisyonlarının Yönergelerini inceledim ve bunlara göre tez çalışmamın yürütülebilmesi için herhangi bir Etik Kurul/Komisyon'dan izin alınmasına gerek olmadığını; aksi durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.</p> <p>Gereğini saygılarımla arz ederim.</p> <p style="text-align: right;">16.06.2021</p> <p><b>Adı Soyadı:</b> Barış Özkaya</p> <p><b>Öğrenci No:</b> N12247826</p> <p><b>Anabilim Dalı:</b> İşletme</p> <p><b>Programı:</b> Doktora</p> <p><b>Statüsü:</b> <input type="checkbox"/> Yüksek Lisans <input checked="" type="checkbox"/> Doktora <input type="checkbox"/> Bütünleşik Doktora</p>
<p><b><u>DANISMAN GÖRÜŞÜ VE ONAYI</u></b></p> <p style="text-align: center;">_____ Prof. Dr. Aydın ULUCAN</p> <p><b>Detaylı Bilgi:</b> <a href="http://www.sosyalbilimler.hacettepe.edu.tr">http://www.sosyalbilimler.hacettepe.edu.tr</a> Telefon: 0-312-2976860 Faks: 0-3122992147 E-posta: <a href="mailto:sosyalbilimler@hacettepe.edu.tr">sosyalbilimler@hacettepe.edu.tr</a></p>

## EK 2. Orijinallik Raporu

 <p><b>HACETTEPE ÜNİVERSİTESİ</b> <b>SOSYAL BİLİMLER ENSTİTÜSÜ</b> <b>DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPORU</b></p>
<p><b>HACETTEPE ÜNİVERSİTESİ</b> <b>SOSYAL BİLİMLER ENSTİTÜSÜ</b> <b>İŞLETME ANABİLİM DALI BAŞKANLIĞINA</b></p> <p style="text-align: right;">Tarih: 03/07/2021</p> <p>Tez Başlığı : ATÖLYE TİPİ ÜRETİM İÇİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE ÜRETİM SAATİ TAHMİNİ: HAVACILIK VE SAVUNMA SANAYİİ UYGULAMASI</p> <p>Yukarıda başlığı gösterilen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler ve d) Sonuç kısımlarından oluşan toplam 104 sayfalık kısmına ilişkin, 03/07/2021 tarihinde tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda işaretlenmiş filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 2 'dir.</p> <p>Uygulanan filtrelemeler:</p> <ol style="list-style-type: none"> <li>1- <input checked="" type="checkbox"/> Kabul/Onay ve Bildirim sayfaları hariç</li> <li>2- <input checked="" type="checkbox"/> Kaynakça hariç</li> <li>3- <input checked="" type="checkbox"/> Alıntılar hariç</li> <li>4- <input type="checkbox"/> Alıntılar dâhil</li> <li>5- <input checked="" type="checkbox"/> 5 kelimedenden daha az örtüşme içeren metin kısımları hariç</li> </ol> <p>Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve bu Uygulama Esasları'nda belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.</p> <p>Gereğini saygılarımla arz ederim.</p> <p style="text-align: right;">Tarih ve İmza</p> <p><b>Adı Soyadı:</b> Barış ÖZKAYA</p> <p><b>Öğrenci No:</b> N12247826</p> <p><b>Anabilim Dalı:</b> İşletme</p> <p><b>Programı:</b> İşletme</p> <p><b>Statüsü:</b> <input checked="" type="checkbox"/> Doktora <input type="checkbox"/> Bütünleşik Dr.</p>
<p><b><u>DANIŞMAN ONAYI</u></b></p> <p>UYGUNDUR.</p> <p>_____ Prof. Dr. Aydın ULUCAN</p>