

Comparison of the Performances of Regression-Based Machine Learning Algorithms in Die Design Cost Estimation

Rukiye TİPİ^{1*}, Hasan ŞAHİN², Şeyma DOĞRU³, Gül Çiçek ZENGİN BİNTAŞ⁴

¹Department of Smart Systems Engineering, Bursa Technical University, Bursa, Turkey

²Department of Industrial Engineering, Bursa Technical University, Bursa, Turkey

³Department of Computer Engineering, Bursa Technical University, Bursa, Turkey

⁴Graduate Education Institute, Kocaeli Health and Technology University, Kocaeli, Turkey

Abstract: This study aims to comparatively evaluate the performance of machine learning regression algorithms used in predicting mold costs in the automotive industry. Different types of regression algorithms, including Linear, Ridge, Lasso, K-Nearest Neighbors (KNN), Decision Tree, Random Forest, Support Vector Machine (SVM), Extreme Gradient Boosting, and Light Gradient Boosting Machines, were considered, and their performances on predicting mold costs and error rates were compared. The Random Forest Regression yielded the highest prediction accuracy at 98.197%.

Keywords: Machine Learning, Regression, Automotive, Die Design Cost, Prediction

Regresyon Tabanlı Makine Öğrenmesi Algoritmalarının Kalıp Tasarımı Maliyet Tahminindeki Performanslarının Karşılaştırılması

Özet: Bu çalışma, otomotiv endüstrisinde kalıp maliyetlerinin tahmin edilmesinde kullanılan makine öğrenmesi regresyon algoritmalarının performansını karşılaştırmalı bir şekilde değerlendirmeyi hedeflemektedir. Makine öğrenmesi regresyon algoritmalarından Lineer, Ridge, Lasso, K-En Yakın Komşu (KNN), Karar Ağacı, Rastgele Orman, Destek Vektör Makinesi (SVM), Aşırı Gradyan Artırma ve Hafif Gradyan Hızlandırma Makineleri gibi farklı regresyon çeşitleri üzerinde durulmuş ve bu algoritmaların kalıp maliyetlerinin tahmin başarısı ve hata oranları üzerindeki performansları karşılaştırılmıştır. Tahmini başarısını en iyi yapan regresyon ise Rastgele Orman Regresyonu %98.197 oranıyla elde edilmiştir.

Anahtar Kelimeler: Makine Öğrenmesi, Regresyon, Otomotiv, Kalıp Maliyet, Tahmin

Reference to this paper should be made as follows (bu makaleye aşağıdaki şekilde atıfta bulunulmalı): Tipi, R., Şahin, H., Doğru, Ş., Zengin Bintaş G. Ç., 'Comparison of the Performances of Regression-Based Machine Learning Algorithms in Die Design Cost Estimation', Elec Lett Sci Eng, vol. 19(2), (2023), 48-62.

1. GİRİŞ

Kalıp sektörü proje bazlı üretim yapan sektörlerden biridir. Otomotiv, beyaz eşya, makine sektörü vb. birçok sektörde sac metal kalıpları kullanılmaktadır. Bu kalıpların en önemli özelliği her kalıptan bir tane üretiliyor olmasıdır. Bu bağlamda tekli üretim olarak da adlandırılabilmektedir. Sac kalıpları otomotivde kullanılan sac parçaların, sac levhalardan basılması için kullanılır. Parçaya istenen formu verebilmek için çoğunlukla 3-5 kalıptan oluşan bir set üretilir. Bir otomobilde ortalama 300-350 sac parça bulunmaktadır. Bu parçaları basmak için ortalama 1200 civarında kalıp üretilmektedir.

*Corresponding author; rukiyepi@gmail.com

Kalıplar uygun preslere (tabla büyüklüğü, tonajı vb.) takılarak, üretim bandı mantığında kullanılır. Bandın bir ucundan giren sac levha, diğer ucundan otomobile takılmaya hazırlanacak bir sac parça olarak çıkar. Otomobilde kullanılan sac parçalar dışında, plastik parçalar için de kalıplarla üretilmektedir.

Kalıpçılık sektöründe maliyet analizi çok önemli aşamalardan birisidir. Müşterilerin birkaç gün gibi çok kısa bir süre içinde teklif talebinde bulunması nedeniyle kalıp firmalarının çoğunun detaylı analiz yapacak zamanları yoktur. Üstelik yapılan kalıp tektir ve eşsizdir. Belki benzeri daha önce yapılmıştır ama aynısı hiç yapılmamıştır. Tasarım süreci, üretimin her aşaması ve kullanılan malzemeler her kalıp için farklılık gösterir. İşte bu zaman baskısı içinde hazırlanan tekliflerde firmalar çoğu zaman ya bir maliyet faktörünü unuturlar ya da çok temkinli davranıp yüksek fiyat vererek potansiyel işi kaçırmaları. Bu nedenle daha maliyet analizinden itibaren proje çok iyi etüt edilip, üretim boyunca da hiç hata yapmadan bitirilmesi gerekmektedir. Bir kalıbın üretiminde hemen hemen tüm imalat yöntemleri kullanılır. Talaşlı ve talaşsız imalat işlemlerinde birçok farklı takım tezgahını kullanmak mümkündür.

Kalıp maliyeti otomotiv sektöründeki üretim sürecinde önemli bir unsur olarak kabul edilir. Bu nedenle, kalıp maliyetinin doğru bir şekilde tahmin edilmesi, üretim sürecinin daha verimli hale getirilmesine yardımcı olur. Maliyet hesaplamak için yazılım ürün kullanmayan firmalar Excel yardımıyla, kendi kurdukları formülleri kullanarak tahmin yapmaktadır. Çoğunlukla sezgilere ve deneyimlere dayanarak, kalıp projesinin başında maliyetleri tahmin etmeye çalışmaktadırlar. Bu durum kullanıcı kaynaklı maliyetlerin doğru bir şekilde hesaplanamamasına neden olmaktadır.

Kalıp maliyet hesabı yapılırken teklif verme sürelerinin uzaması ve sunulan tekliflerin gerçekleri göstermemesi, verilen tekliflerin işe dönüşmemesine neden olmaktadır. Hesaplamanın doğru yapılmaması ya da uzun sürmesi teklif verme sürecini ve güvenilirliğini azaltmaktadır. Ülkemizde kalıpçılık sektörünün hızla geliştiği dikkate alınır, dijital teknolojilerin kullanılarak tekliflerde doğru, hatasız, güvenilir ve rekabet edebilir bir maliyetleme politikası bu sürecin hızla geliştirilmesine ihtiyaç olduğunu göstermektedir. Sac salıp sektöründe yer alan şirketlerin üretim maliyet öngörü süreçlerinde birçok hata yaptıklarını ve bu süreçte çok zaman kaybettikleri belirlenmiştir. Kalıp maliyetlerini hesaplayan bazı programlar mevcuttur. Bu programların ortak noktası kural tabanlı yaklaşımlarla ve formüllerle maliyetleri tahmin etmeye çalışmalarıdır. Bu yazılımlar her kalıba ait esnek bir yapıda değildir. Yazılım kullanmayan firmalar da Excel yardımıyla, kendi kurdukları formülleri kullanarak tahmin yapmaktadır.

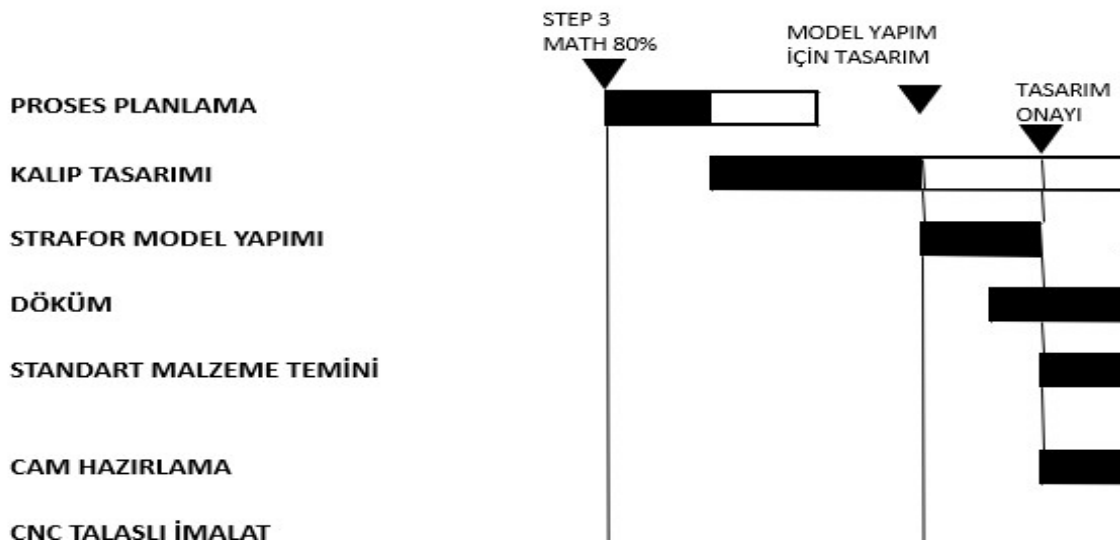
Çoğunlukla sezgilere, deneyimlere ve geleneksel yöntemlere dayanarak kalıp projesinin başında maliyetler tahmin edilmeye çalışılmaktadır. Bu durum kullanıcı kaynaklı maliyetlerin doğru bir şekilde hesaplanamamasına neden olmaktadır.

Otomotiv endüstrisinde, kalıp maliyetlerinin doğru bir şekilde tahmin edilmesi, işletmelerin rekabet avantajını koruması ve maliyet yönetimini etkin bir şekilde yapabilmesi için büyük önem taşır. Geleneksel yöntemlerle kalıp maliyet tahmini yapmak, zaman alıcı ve hata yapmaya elverişlidir. Bu nedenle, makine öğrenmesi teknikleri ve özellikle regresyon algoritmaları, kalıp maliyet tahmininde yenilikçi bir yaklaşım sunmaktadır. Bu çalışma, otomotiv sektöründe maliyet tahminini gerçekleştirmek için çeşitli makine öğrenmesi regresyon algoritmalarının farklı türlerini uygulayarak karşılaştırmalar yapmış ve sektöre yeni bir bakış açısı sunmuştur. Araştırma, geleneksel hesaplama yöntemlerinin yanı sıra farklı yöntemlerin de olduğu ve doğru tahmin başarısının elde edilebileceğini vurgulayarak literatüre katkı sağlamaktadır.

2. KALIP MALİYET HESABI VE KALIP ÜRETİM SÜRECİ

Otomotiv sektöründe kalıp üretim süreci, birçok aşamadan oluşur. Öncelikle, üretim sürecinin amacına ve üretilmek istenen parçaların tasarımına göre kalıp tasarımı gerçekleştirilir. Bu tasarım aşamasında, kalıp için gerekli malzemeler, boyutlar ve diğer özellikler belirlenir. Daha sonra, kalıp tasarımı gerçekleştirilen malzemeler kullanılarak kalıp üretimi gerçekleştirilir. Bu süreçte, kalıp için gerekli malzemeler kesilir, form verilir, işlenir ve montajı gerçekleştirilir.

Kalıp üretim sürecinin son aşaması olan deneme-devreye alma sürecinde, kalıp üretilen parçalar ile test edilir ve gerekli düzeltmeler yapılır. Bu aşamada, kalıp üretilen parçaların doğruluğu, kalitesi ve gerekli özellikleri kontrol edilir. Eğer gerekli düzeltmeler yapılmışsa, kalıp üretilen parçalar seri üretime hazır hale getirilir. Kalıp üretim sürecinin Şekil 1'deki gibi bir planlaması mevcuttur.



Şekil 1. Kalıp Üretim Süreci

Bu süreçte kalıp maliyeti hesaplanırken, birçok faktör dikkate alınmalıdır. Bunlar arasında üretim sürecinde kullanılan malzeme ve işçilik giderleri, kalıp tasarımı, yapımı, üretim süresi, üretim hacmi, üretim hızı, üretim kalitesi, işçilik maliyeti ve işçilik saati gibi birçok girdisi bulunmaktadır.

Kalıp maliyetinin hesabında etkili olan önemli etkenlerden biri de kalıptaki fiziksel boyutlardır. Yükseklik, en ve boy gibi özellikler, kalıbın maliyetinin hesabını doğrudan etkilemektedir. Boyutsal olarak büyük bir kalıp, daha çok malzeme kullanımı demektir. Boyutu daha büyük kalıpların işlenmesi, daha çok zaman ve efor gerektirmekte ve işçilik maliyetlerini artırmaktadır. Kalıpta kullanılan malzeme miktarı ve kalıp tasarımı, kalıp ağırlığını belirler. Ağır bir kalıp, daha çok malzeme ve kalıp işçiliği gerektirir. Bunun yanı sıra, ağır kalıpların taşınması ve işlenmesi büyük bir efor gerektirir, bu da işçilik süresini artırır, maliyete etki etmektedir.

Harcanan toplam zaman bu süreçte maliyet üzerinde büyük değişikliklere sebep olan bir faktördür. Kalıbın tasarımı, malzeme kesimleri, kalıba form verme işlemi ve montaj gibi adımların her biri zaman gerektirir. Karmaşık kalıp tasarımları veya detaylı işlemler büyük zaman alabilir ve dolayısıyla maliyeti artırır. Aynı zamanda, işçilik maliyetleri süre ile doğrudan ilişkilidir, çünkü uzun süren işlem daha fazla işçilik saatini gerektirmektedir.

Kalıp maliyetinin hesaplaması, oldukça kompleks bir süreçtir ve etki eden birçok etkenin bir araya getirilmesini zorunlu kılar. Bu karmaşık yapı, geleneksel yöntemlerle sınırlandırıldığında doğru ve gerçek sonuçlar üretmekte zorluk yaşar. Fakat makine öğrenme regresyon yöntemleriyle yeni bir yaklaşım benimsenerek, bu etkenlerin birbiri ile aralarında ilişki kurulması ve kalıp maliyetinin doğru bir şekilde hesaplanmasını sağlar.

Bu yaklaşım, geniş bir veri kümesi üzerinde algoritmaların eğitilmesini sağlar. Makine öğrenmesi regresyon algoritmaları, kompleks bir veri analitiği ve istatistiksel modelleme işlemiyle, kalıp maliyetinin ortaya çıkmasındaki etkili olan faktörler arasındaki ilişkileri keşfetmeye yardımcı olur. Bu sayede kalıp ağırlığı, kalıp boyutları, kalıp üretim süresi, malzeme giderleri, işçilik giderleri ve diğer önemli parametreler arasındaki ilişkileri anlayabilir ve kalıp maliyetinin doğru bir şekilde tahmin edilebilmesini sağlar.

Eğitim sürecinde kullanılacak veri setleriyle algoritma, kalıp maliyeti üzerindeki etkili faktörleri tespit edebilir ve bu etkenler arasındaki ilişkiyi öğrenebilir. Böylece, gelecekte kalıp üretimleri için maliyet tahminleri daha doğru ve güvenilir bir hale gelmesini sağlayacaktır.

Bu yeni yaklaşım, kalıp üreticilerine daha doğru ve verimli bir karar verme süreci sağlar. Makine öğrenmesi regresyon algoritmalarıyla desteklenen bu yaklaşım, kalıp maliyetinin daha efektif bir şekilde hesaplanabilmesini ve devamlı olarak iyileştirilmesine olanak sağlar.

3. LİTERATÜR TARAMASI

Regresyon tabanlı makine öğrenme algoritmalarının kalıp tasarımı maliyet tahminindeki performanslarının karşılaştırılması, endüstriyel süreçlerin etkin yönetimi ve maliyet optimizasyonu için kritik bir öneme sahiptir. Regresyon tabanlı makine öğrenme algoritmalarının kalıp tasarımı maliyet tahminindeki performanslarını değerlendirmek amacıyla birçok önceki araştırma gerçekleştirilmiştir. Bu çalışmalar arasında şunlar yer almaktadır.

Vishal Naranje ve Shailendra Kumar, çalışmalarında bahsedilen uygulamada sistem modüllerini Windows işletim ortamında VB.6.0'da kodlanmıştır. Sistem, endüstrilerden alınan çeşitli tiplerde eksenel simetrik derin çekilmiş sac parçalar için test edilmiştir. Bulutta toplanan verilerin analiz edilerek ve yapay sinir ağının kural tabanlı yaklaşımı ile tahmin yapılması amaçlanmıştır. Uzman maliyet mühendislerine göre teklif süresinin dakikalara indiği ve yaptığı tekliflerle sistemin güven kazandığı görülmüştür. Sistem, çeşitli küçük ve orta ölçekli çok çeşitli derin çekilmiş parçalar için başarıyla doğrulanmıştır. Gelecekte, diğer sac levhaların maliyet tahmini için benzer türde bir yaklaşım kullanılabilir sonucuna varılmıştır [1].

Fangwei Ning, Yan Shi ve Maolin Cai, çalışmalarında derin öğrenme ile üretim sürecindeki parçanın maliyet tahmini için iki boyutlu (2B) ve üç boyutlu (3B) sinir ağı (CNN) görüntüleri daha önce geleneksel yöntemlerle hesaplanmış olan parçalar ile karşılaştırılarak tahmin öngörüsü yapmaya çalışılmıştır. CNN ile sistemin eğitiminin farklı çözünürlüklerde yapılması ile daha hassas bir hesaplama yapılabildiği sonucuna varılmıştır. Bunun yanında 2D CNN yerine, 3D CNN'nin kullanılması maliyet tahmininde mükemmel yakın performans gösterdiği ve yüksek bir uygulama değeri elde ettiği sonucuna ulaşılmıştır [2].

Zs. J.Viharos ve B. Mikó, çalışmalarında plastik enjeksiyon teknolojisinin rolünün otomotiv endüstrisinde giderek artan bir ivmeye sahip olduğunu belirtiyorlar. Kalıp üreticisi şirketler, yalnızca tasarım ve üretim verimliliği açısından değil, aynı zamanda artan teklif talebini karşılamaya çalışırken de önemli bir zorlukla karşı karşıyadır. Makale, paralel vaka tabanlı ve

buluşsal maliyet tahmin sistemini destekleyen maliyet tahmini için yapay bir sinir ağı uygulamasını sunmaktadır. Toplamda on dört maliyet türü, teklif hazırlayan bir uzman tarafından kolay ve hızlı bir şekilde belirleneceklerini gösteren, verilen CAD modeline dayalı olarak basitçe belirlenebilen on üç araç tanımlama parametresine dayanan umut verici bir sonuçla tahmin edilmiştir [3].

Karl Kuzman ile Blaž Florjanič, çalışmalarında proje süresi ve maliyet tahmini deneysel olarak şirket içi uzmanlar tarafından yürütülen bir süreç olduğunu belirtmektedirler. Bu süreçteki ana engellerden biri, ürün özellikleri ile kalıbı üretmek için gerekli işleme saatleri arasındaki ilişkinin doğru bir şekilde tanımlanmamasıdır. Bu makale, yapay sinir ağları kullanarak bu sorunun nasıl çözüleceğini amaçlamaktadır [4].

André Rodrigues ve arkadaşları, çalışmalarında kalıp parçalarının birbirinden farklı olduğu düşünüldüğünde, işleme sürelerinin doğru ve hızlı bir şekilde tahmin edilmesinin bir firmanın başarısı için büyük önem taşıdığını belirtmektedirler. Bu makale, standart enjeksiyon kalıbı parçaları için işleme süresi tahmininde yapay sinir ağlarını uygulamak için bir öneri sunmaktadır. Bu amaçla, yapay sinir modelini şekillendirecek geniş bir parça kümesi düşünülmüş ve yapay sinir ağlarını eğitmek için yeterli veriyi toplamak üzere işleme süreleri hesaplanmıştır. Ağ mimarisinin etkileri, giriş verileri ve ağı eğitimi için kullanılan değişkenler, en yüksek tahmin doğruluğuna sahip sinir ağını bulma amaçlanmıştır [5].

H.L. Cather ve K.H. Chan, çalışmalarında Computer Aided Calculation Estimation(CACE) adını verdikleri sistemde, ürün maliyet tahminleri yapabilmek için literatürde kullanılan yeni bir yazılım ailesini kullanmışlardır. CACE'nin literatüre en büyük katkısı tasarım ile üretimi entegre etmesi ve veri analiz yöntemleri ile eski maliyetler ile yeni teklifler arasındaki ilişkiye dair kullanıcıya öngörü sağlayabilmesidir. CACE farklı sektörler ve kullanıcılar için varyantlar geliştirmemiştir. Elde edilen sonuç ise gelecekte firmalar dijital becerilere sahip çalışanlara ihtiyaç duyacaktır. Modern maliyet tahmin araçlarını kullanacak mühendislerin de değerli becerilerle donatılması şirketlerin işgücü piyasasında başarılı olmalarına yardımcı olacaktır [6].

Federico Campi ve arkadaşları, çalışmalarında tahmin modeli ile, tüm süreç aşamaları ve hammadde ile ilgili ayrıntılı bir maliyet dökümü sağlayan tasarım aşamasında mevcut olan geometrik özelliklerin analiz yapmışlardır. Model üretim maliyetini tahmin etmeye, geometrik özellikleri ve maliyet kalemlerini birbirine bağlamaya olanak tanımaktadır. Model esas olarak tasarım ve maliyet mühendisleri için üretim sürecini ve tedarik zincirini geliştirmek için tasarlanmıştır. Yapılan tahmini maliyetler arasındaki mutlak ortalama sapma, hammadde için yaklaşık %4 ve proses için %21 toplam maliyet (hammadde ve üretim süreci) üzerindeki mutlak ortalama sapma yaklaşık %5 olarak hesaplanmıştır. Gerçek değerler ise maliyet mühendisleri tarafından sağlanmıştır [7].

Makine öğrenmesi regresyon algoritmaları, kalıp maliyetlerinin tahmin edilmesi gibi gerçek değerleri tahmin etmek için kullanılan etkili araçlardır. Bu algoritmalar, veriler arasındaki ilişkileri modellemek ve tahmin yapmak için istatistiksel yöntemler ve matematiksel formüller kullanır. Çeşitli regresyon algoritmaları, kalıp maliyetlerinin tahmin edilmesi için kullanılabilir. Bu algoritmalar arasında lineer regresyon, çoklu regresyon, destek vektör regresyonu (SVR), karar ağaçları ve rastgele ormanlar gibi popüler olanları bulunur. Her bir algoritmanın farklı avantajları ve dezavantajları vardır ve performansları veri seti ve kullanım senaryosuna bağlı olarak değişebilir. Karşılaştırmalı bir değerlendirme yaparken, farklı regresyon algoritmalarının performansını ölçmek için bazı metrikler kullanılabilir. Bu metrikler arasında ortalama karesel

hata (MSE), ortalama mutlak hata (MAE), belirli bir eşiğe göre doğru tahmin yüzdesi gibi ölçümler bulunur.

Aşağıda, verilen referanslardan elde edilen bilgilere dayanarak bir tablo oluşturulmuştur. Tabloda, ulaşılan referansları amacı, yayın yılı, kullanılan yöntem ve elde edilen sonuçlar özetlenmiştir.

Tablo 1. Literatür Özeti

Referans	Amaç	Yıl	Yöntem	Sonuçlar
[1]	Derin çekilmiş parçaların maliyet tahminini yapmak için bir bilgi tabanlı sistem oluşturmak.	2014	Bilgi tabanlı sistem	Gerçek veriler üzerinde doğruluk oranı ve tahmin hataları değerlendirilmiştir.
[2]	Derin öğrenme yöntemine dayalı bir imalat maliyeti tahmin yöntemi geliştirmek.	2020	Derin öğrenme	Tahmin doğruluğu ve hesaplama hızı üzerinde değerlendirme yapılmıştır.
[3]	Enjeksiyon kalıp maliyet tahmininde yapay sinir ağı yaklaşımını kullanmak.	2011	Yapay sinir ağı	Tahmin hataları ve doğruluk oranı değerlendirilmiştir.
[4]	Enjeksiyon kalıplarının imalat süresini tahmin etmek için yapay sinir ağı tabanlı bir model geliştirmek.	2012	Yapay sinir ağı	Tahmin doğruluğu ve hesaplama süresi üzerinde değerlendirme yapılmıştır.
[5]	İşleme süresini tahmin etmek için yapay sinir ağı yaklaşımını kullanmak.	2022	Yapay sinir ağı	Tahmin hataları ve doğruluk oranı üzerinde değerlendirme yapılmıştır.
[6]	Bilgisayar destekli maliyet tahmininde entegre bir yaklaşım geliştirmek.	2014	Entegre yaklaşım	Veri erişimi, analizi ve maliyet tahmini üzerine odaklanan bir yaklaşım sunulmuştur.
[7]	Açık kalıp dövme teknolojisiyle simetrik bileşenlerin tasarımı için analitik bir maliyet tahmin modeli oluşturmak.	2020	Analitik model	Tasarım parametreleri ve dövme işlemi özelliklerine dayalı olarak maliyet tahmini yapılmıştır.
[8]	Makine öğrenmesi hakkında genel bir bilgi sunmak.	2023	Bilgi sunumu	Makine öğrenmesinin tanımı, kullanım alanları ve temel kavramları hakkında bilgi verilmiştir.
[11]	Lasso regresyonu hakkında genel bir bilgi sunmak.	2023	Bilgi sunumu	Lasso regresyonunun tanımı, kullanım alanları ve temel kavramları hakkında bilgi verilmiştir.
[14]	Eksik verilerle kalıp sınıflandırma yapmak için bir inceleme sunmak.	2010	İnceleme makalesi	Eksik verilerin olduğu durumlarda kalıp sınıflandırma teknikleri ve yöntemlerine dair bir inceleme sunulmuştur.
[15]	Veri madenciliği modelleri kullanarak potansiyel hava kirliliği emisyonlarını tahmin etmek için karar ağacı algoritmalarını karşılaştırmak.	2011	Karşılaştırmalı çalışma	Farklı karar ağacı algoritmalarının potansiyel hava kirliliği emisyonlarını tahmin etmedeki performansları karşılaştırılmıştır.

Bu tablo, farklı referanslarda yapılan çalışmaların amacını, yayın yılını, kullanılan yöntemi ve elde edilen sonuçları özetlemektedir. Her bir çalışma, kalıp maliyetlerinin tahmin edilmesi veya imalat sürelerinin tahmin edilmesi üzerinde farklı yöntemler ve yaklaşımlar kullanmıştır. Sonuçlar, tahmin doğruluğu, tahmin hataları ve hesaplama süreleri gibi metrikler kullanılarak değerlendirilmiştir. Tablo 1 farklı makine öğrenimi regresyon algoritmalarının kalıp

maliyetlerini tahmin etme başarısını karşılaştırarak nasıl bir değerlendirme yapabileceğinizi açık bir şekilde göstermektedir.

Maliyet tahmini alanında yapılan literatür taramasından farklı olarak makine öğrenme regresyon algoritmaları kullanılarak sonuçlar edilmiştir. Geleneksel yaklaşımlarda, kalıp projelerinin maliyet tahminleri genellikle kullanıcıların sezgilerine ve deneyimlerine dayanır, bu da kullanıcı kaynaklı maliyetlerin doğru bir şekilde hesaplanmasını zorlaştırır. Geleneksel yöntemler yerine veriye dayalı bir modelleme yaklaşımı sunan bu yeni yöntem, daha doğru ve hassas tahminler yapmayı hedeflemektedir. Ayrıca, veri seti ve özelliklerin seçimi konusunda da özgün bir yaklaşım sunulmaktadır. Yeni yöntem, daha kapsamlı veya daha özelleştirilmiş bir veri seti kullanarak tahmin doğruluğunu artırma potansiyeline sahiptir. Farklı algoritmaların etkisi ve parametre ayarlamalarıyla tahmin doğruluğu artırılmaya çalışılmıştır. Farklı endüstrilere veya kalıp türlerine uyarlanabilir olması ve büyük ölçekli veri setleriyle başa çıkabilmesi, işletmelere daha etkili maliyet tahminleri yapma imkânı sunmaktadır. Bu makale, maliyet tahmini alanında yapılan literatür taramasından farklı bir yaklaşım sunarak, daha doğru ve yenilikçi bir tahmin modelinin potansiyelini ortaya koymaktadır.

4. MATERYAL VE METOT

4.1. Makine Öğrenmesi

Makine öğrenimi, yapay zeka alt kümesi olarak, kendini öğrenme yeteneği sayesinde kararların doğruluğunu artırabilme kapasitesine sahiptir. Günümüzde, makine öğrenimi algoritmaları sonuçlar çıkarmak ve veri setlerini tahmin etme yeteneğine sahiptir ve bu yöntemlerin önde gelenlerindedir [21]. Doğrudan programlamak yerine, bilgisayarlara verilerden öğrenmeyi ve deneyimle gelişmeyi öğretmeye odaklanır. Makine öğreniminde algoritmalar, büyük veri kümelerindeki kalıpları ve korelasyonları bulmak ve bu analize dayalı olarak en iyi kararları ve tahminleri yapmak üzere eğitilir. Makine öğrenimi uygulamaları kullanıldıkça gelişir ve daha fazla veriye eriştikçe daha doğru hale gelir [8]. Makine öğrenme algoritmaları, geleneksel istatistiksel analiz yöntemlerinin aksine, değişkenlerin verileri arasındaki ilişkileri kullanarak tahmin verileri sağlar. Bazı makine öğrenme algoritmaları, tahmin modellerinin oluşturulmasında ve değişkenler arasındaki ilişkilerin hesaplanmasında regresyon analizi yöntemlerinin altyapısını kullanarak önemli bir rol oynar [20].

4.2. Makine Öğrenme Regresyon Algoritmaları

Makine öğrenmesi denetimli öğrenme türünden olan regresyon algoritmaları tanımları aşağıda açıklanmıştır.

Lineer Regresyon, Lineer regresyon analizi, bir değişkenin değerini başka bir değişkenin değerine göre tahmin etmek için kullanılır. Tahmin etmek istediğiniz değişkene bağımlı değişken denir. Diğer değişkenin değerini tahmin etmek için kullandığınız değişkene bağımsız değişken denir. Bu analiz biçimi, bağımlı değişkenin değerini en iyi tahmin eden bir veya daha fazla bağımsız değişkeni içeren doğrusal denklemin katsayılarını tahmin eder. Doğrusal regresyon, tahmin edilen ve gerçek çıktı değerleri arasındaki tutarsızlıkları en aza indiren düz bir çizgiye veya yüzeye uyar. Bir dizi eşleştirilmiş veri için en uygun doğruyu bulmak için "en küçük kareler" yöntemini kullanan basit doğrusal regresyon hesaplayıcıları vardır. Daha sonra X'in (bağımlı değişken) değerini Y'den (bağımsız değişken) tahmin edersiniz [9].

Ridge Regresyon, bağımsız değişkenlerin yüksek oranda ilişkili olduğu senaryolarda çoklu regresyon modellerinin katsayılarını tahmin etme yöntemidir. Ekonometri, kimya ve mühendislik gibi birçok alanda kullanılmıştır. Çok sayıda parametreye sahip modellerde yaygın olarak ortaya

çıkan doğrusal regresyonda çoklu bağlantı sorununu azaltmak için özellikle yararlıdır. Genel olarak, yöntem, kabul edilebilir bir sapma miktarı karşılığında parametre tahmin problemlerinde gelişmiş verimlilik sağlar [10].

Lasso Regresyon, Lasso regresyonu bir düzenleme tekniğidir. Daha doğru bir tahmin için regresyon yöntemleri yerine kullanılır. Bu model büzülme kullanır. Büzülme, veri değerlerinin ortalama olarak merkezi bir noktaya doğru küçüldüğü yerdir. Kement prosedürü, basit, seyrek modelleri (yani daha az parametrelili modeller) teşvik eder. Bu özel regresyon türü, yüksek düzeyde çoklu bağlantı gösteren modeller için veya değişken seçimi/parametre eleme gibi model seçiminin belirli kısımlarını otomatikleştirmek istediğinizde çok uygundur [11].

KNN Regresyonu, KNN algoritması önce örüntü tanıma olmak üzere yoğun olarak yapay zekâ alanında kullanılan uzaklık tabanlı bir sınıflayıcı algoritmasıdır. KNN sınıflandırma tekniğinde test verileri ile eğitim verileri arasındaki uzaklıklar hesaplanarak test örneğine en yakın K tane örnek seçilmektedir. Daha sonra seçilen K tane örneğin sınıf bilgilerinden yararlanarak çoğunluk oylaması ile test örneklerinin sınıfı elde edilmektedir [12].

Karar Ağacı Regresyonu, Karar ağacı metodu, tahmin etmede ve sınıflandırmada kullanılan önemli veri madenciliği metodları arasında yer alır. Karar ağacı algoritması, girdi verisi olmayan kök düğüm ve birer girdi alan her biri, iç düğümlerden meydana gelen yönlü bir ağaçtır. Çıktı verileri başka bir düğüm tarafından girdi verisi olarak alınan düğümler test veya iç düğümü, çıktı verileri başka bir düğüme girdi olmayan düğümler de yaprak düğümler olarak isimlendirilir. Karar ağacının metodunda her bir iç düğümün verisi, örnek uzayına girdi öznelik değerinin belirli bir fonksiyona bağlı kalmasına dayalı olarak iki veya daha fazla parçaya bölünmektedir [13].

Sınıflandırma işleminde ve tahmin etmede karar ağacının kullanılması, eğitim örneklerinden modelin ortaya çıkarılması ve bu modelin, test verisi örnekleri kullanılarak uygun test etme ölçütleri yoluyla değerlendirilmesinin yapılması ve bu modelin gelecek zamanlardaki değerlerinin tahmin edilme işlemlerinde kullanılması ile işlemektedir [14-15].

Veri setlerinden otomatik bir şekilde karar ağacı yapısını elde etmek amacıyla daha önceden geliştirilmiş birden çok karar ağacı algoritması vardır. Karar ağacı algoritmaları genel olarak genelleştirme hatasını en az duruma düşüren ve karar ağacı yapısını en uygun bir şekilde ortaya çıkarmayı hedeflemek ile ortalama derinlik, düğüm sayısı veya başka hedef fonksiyonlarını en az duruma düşürmeyi hedeflemek de mümkün olacaktır [12].

Rastgele Orman Regresyonu, Karar ağacı metodunun n kez çalıştırılması ile daha iyi tahmin edebilmeyi sağlamaktadır. Veri setinden rastgele olarak seçilmiş karar ağaçları alt kümelerdir. Bir araya gelen karar ağaçları, karar ormanlarını oluşturmaktadırlar. Kayıp veri setlerinde ya da dengesiz veri setlerinde daha iyi sonuçlar üretmektedirler [16].

Destek Vektör Makinesi (SVM), İstatistiksel bir sınıflandırıcı olan destek vektör makineleri, örüntülerin değerlerinin bilinmediği değişkenlerin arasındaki veri setlerinde çalıştırılmak koşuluyla önerilen bir sınıflandırma çeşitidir. Sınıflandırmayı bir düzlem üzerinde olan veriler arasında bir sınır oluşturarak iki farklı gruba dağıtır. Oluşturulan sınır birçok verinin birbirini arasındaki mesafesi en fazla olan yer olarak sağlar [17].

XGBoost Regresyonu, Aşırı Gradyan Artırma karar ağacı algoritmasını baz alan, hızı fazla performansı başarılı bir metottur. XGBoost tekniği, karar ağacını ortaya çıkarırken paralel çalışma tekniğini kullandığı için işlemler hızlı olarak ortaya çıkar. Bu metod ilk olarak

“max_depth” değerini kullanarak ağaçtaki derinlik hesabını bulmaktadır. Ağacı aşağıya doğru yönde fazla derin ise geriye doğru olarak budama işlemi yaparak çalışmasını kaldığı yerden sürdürür. Bu yöntem izlenerek modelin veriyi ezberleme durumunun önüne geçerek fazla derecede öğrenme problemi için çözüm aramaktadır. En bilinen ve önemli yeteneklerinden bir tanesi de verileri birçok ağaca sınıflandırırken uygun noktayı bulmak üzere veri setindeki gözlem noktasını ağırlık değerlerine göre kullanabilior olmasıdır [18].

Light GBM Regresyonu, Hafif Gradyan Hızlandırma Makineleri tekniği, histogram tabanlı bir alt yapıya sahiptir. CatBoost metodunun çalışma şekline benzer bir şekilde öznelik isimlerinin girdi verisi olmak üzere çekerek sınıflayıcı özneliklerini konu edinir. Light GBM metodunda karar ağacının eğitilme zamanı, sonucu çıkarılan hesaplama ve devamında bölünme ile ters olamayan bir orantıya sahiptir. Bu yeteneği sonucunda hem eğitime zamanının uzun olmaması hem de kaynak kullanımını azaltmaktadır. Karar ağacı tabanlı olan bu teknik, ağacı en optimize bir şekilde yaprak bilgilerine ayırırken, kalan yükseltme teknikleri ağacı yaprak olarak değil derinlik olarak veya seviye olarak ayırmaktadır. Bu durumun sonucunda, Light GBM metodu aynı yaprak üstünde geliştiğinden yaprak-bazlı teknik ve seviye-bazlı teknikten daha çok kaybı düşürmektedir. Bu şekilde var olan yükseltme metotlarından herhangi biri tarafından sonuç alınabilecek çok daha uygun doğruluğa sahip olunur [19].

Tablo 2. Regresyon Algoritmalarının Benzer ve Farklı Yönleri

Algoritma (Regresyon Türü)	Temel İlke	Farklılık	Benzerlik
Linear	Doğrusal ilişkiyi modelleme	Yüksek boyutlu ve basit verilere uygun anlaşılır model yapısı	Sadece doğrusal ilişkileri modelleyebilir ve kalan regresyon modellerinin temeli
Ridge	Doğrusal ilişkiyi modelleme, L2 düzenleme	L2 düzenlemesi ile aşırı uydurmaya karşı dirençli	Linear regresyon çeşidi baz alınmıştır
Lasso	Doğrusal ilişkiyi modelleme, L1 düzenleme	L1 düzenlemesi ile özellik seçimi yapar	Linear regresyon çeşidi baz alınmıştır
K-En Yakın Komşu	Veri noktalarının komşuluklarına bağlı tahmin	Komşuluk bilgisine bağlı esnek modelleme	Parametre olan k değerinin seçimi önemli ve veriye göre benzerliklere odaklanır
Karar Ağacı	Karar ağaçlarını kullanarak tahmin	Ağaç yapısı kullanılarak karmaşıklık ele alınır	Ağaç yapısı baz alınır
Rastgele Orman	Birden çok karar ağacının birleşmesi sonucu yapılan tahmin	Ağaçların birleşmesi ile varyans azaltılır	Ağaç yapısını baz alan toplu modelleme
SVM	Optimizasyon ile destek vektörleri kullanılarak yapılan tahmin	Kernel fonksiyonlarıyla birlikte çok boyutlu uzaylarda çalışabilir	Temelinde destek vektör makinaları
XGBoost	Zayıf öğrenebilen modelleri birleştirerek tahmin	Zamanla zayıf modelleri güçlendirmek için kullanılır	Daha fazla ayar gerektirebilir ve karar ağaçları gibi toplu modelleme yapar
Light GBM	Aşırı gradyan temelli hafif modelleme	Daha düşük kaynak kullanımı ve daha hızlı eğitim süresi	Kategorik özellikleri etkili bir şekilde işler ve aşırı gradyan artırma baz alınır

5. ANALİZ VE BULGULAR

Uygulama aşamasında dört adım şeklinde gerçekleşmiştir. Birinci adımda veri kümesinin oluşturulması, ikinci adımda modele atılmadan önce veriler üzerinde bir ön işlemenin uygulanması, üçüncü adımda makine öğrenmesi regresyon algoritmaları kullanılarak model oluşturulması ve son adımda ise algoritma sonuçlarının karşılaştırılması yapılmıştır.

5.1. Veri Kümesi

Toplam özellik sayısı 8 ve 52 satır verinin sentetik olarak üretilmesiyle birlikte toplamda 2092 satır veriye ulaşılmıştır. Bu çalışmadaki veri kayıtları Mubitek Tasarım ve Bilişim Mak. San. ve Tic. Ltd. Şti den alınmıştır.

Girdi parametreleri aşağıdaki gibidir.

- Operasyon Numarası: Üretilen kalıbı türü ve zorluğu hakkında maliyete etki eden bir metriktir.
- Operasyon: Kalıbın türü/adı olarak ifade edilir.
- En: Kalıbın enini ifade eder.
- Boy: Kalıbın boyunu ifade eder.
- Yükseklik: Kalıbın yükseklik bilgisidir.
- Saat: İşçilik süresini ifade eder.
- Kilogram: Kalıbın toplam ağırlığıdır.

Çıktı olarak ise toplam maliyet yer almaktadır.

Maliyet hesaplama süreçlerinde başka girdiler de bulunmaktadır. Yukarıda kullanılan girdiler en yaygın olan girdilerdendir.

5.2. Veri Kümesi Ön İşlem

Veri kümesi içindeki birbirine benzememesi gereken satırların kontrolü gerçekleştiğinde 1851 satır bulunmaktadır. Algoritmaların ezberleme yapmaması için kopya olan satırlar veri setinden çıkarılmıştır. Tipi nesne olan sütundaki tüm veriler nümerik formata dönüştürülmüştür. Örnek olarak operasyon sütununda 4 farklı tip bulunmaktadır. Nümerik olarak etiketlenmiştir.

5.3. Veri Kümesi Modellemesi

İkinci adımda regresyon algoritmaları, Python programlama dilini kullanarak farklı teknik yöntemleriyle eğitilmiştir. Veri seti, algoritmaların performansını ölçmek ve değerlendirmek amacıyla kullanılmıştır. Veri ön işleme adımında, veri kümesindeki eksik ve anlamsız özelliklerin çıkarılması, sayısal özelliklerin işlenmesi ve veri normalizasyonu gibi önemli aşamalar gerçekleştirilerek kullanılacak modellerin doğruluğu optimize edilmeye çalışılmıştır. Ridge, Lasso, Lineer, K-En Yakın Komşu, Karar Ağacı, Rastgele Orman, Destek Vektör Makinesi, Aşırı Gradyan Artırma ve Hafif Gradyan Hızlandırma Makineleri Regresyonları gibi türlü algoritmaları kullanılmıştır ve bu algoritmaların eğitimi sırasında, her bir algoritmanın özel gereksinimlerine uygun olarak veri kümesi üzerinde gerekli optimizasyonlar sağlanmıştır. Bu adımlar sayesinde regresyon modellerinin doğruluğu ve genel performansını ve başarısını artırmak amaçlanmıştır. Python dilinin gücü kullanılarak regresyon algoritmalarının verimli bir

şekilde eğitilmesi ve buna bağlı olarak veri kümesinde tahmin başarısı yüksek bir performans elde edilmesi amaçlanmıştır.

Ön işlemesi yapılan veri kümesi %20 test verisi ve %80 eğitim verisi olarak rastgele olarak dağıtım yapılmıştır.

Regresyon yöntemlerinde kullanılan metrikler r-kare, ortalama mutlak hata (MAE), ortalama kare hata(MSE), kök ortalama kare hata (RMSE), ortalama mutlak yüzde hata (MAPE) ve ortalama mutlak yüzde hata değerinden elde edilen tahmin başarısıdır.

Tablo 3. Metrik ve Formüller

Metrik	Formül
R-Kare	$1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$
MAE	$\frac{1}{N} \sum y_i - \hat{y}_i $
MSE	$\frac{1}{N} \sum (y_i - \hat{y}_i)^2$
RMSE	$\sqrt{MSE} = \sqrt{\frac{1}{N} \sum (y_i - \hat{y}_i)^2}$
MAPE	$\frac{100\%}{n} \sum \frac{ y_i - \hat{y}_i }{ y_i }$

y_i gerçek değeri, \hat{y}_i tahmin edilen değeri, \bar{y} ortalama değeri ve n iste toplam veri adetini temsil eder.

5.4. Bulgular ve Sonuçların Karşılaştırılması

Tablo 4’te “Tahmin Başarısı” başlığı ise ortalama mutlak yüzde hata değerinden elde edilmiştir. Formülü ise 1-MAPE olarak hesaplanmıştır. 9 tip Regresyon yöntemi verilere uygulanmıştır. Bu yöntemlere ait sonuçlar Tablo 4 de ayrıntılı olarak gösterilmiştir.

Tablo 4. Hata Değerleri ve Tahmin Başarı Sonuçları

Regresyon Yöntemi	R-Kare	MAE	MSE	RMSE	MAPE	Tahmin Başarısı
Linear Regresyon	0.88	5431.27	47417580.95	6886.04	%4.67	%95.32
Ridge Regresyon	0.88	5445.70	47716104.20	6907.68	%4.68	%95.31
Lasso Regresyon	0.88	5445.68	47715310.30	6907.62	%4.68	%95.31
KNN Regresyonu	0.91	5032.26	41147485.97	6414.63	%4.29	%95.70
Karar Ağacı Regresyonu	0.91	4835.78	40778322.75	6385.79	%4.09	%95.90
Rastgele Orman Regresyonu	0.96	3396.13	19927875.25	4464.06	%2.86	%97.13

SVM Regresyonu	0.91	4921.5	38867600.44	6234.39	%4.21	%95.78
XGBoost Regresyonu	0.91	5115.95	40519544.29	6365.5	%4.46	%95.53
Light GBM Regresyonu	0.79	8182.95	91325080.58	9556.42	%8.08	%91.91

Tablo, farklı regresyon algoritmalarının bir regresyon problemindeki performansını ölçen türlü metrikleri içermektedir.

R-Kare: Rastgele Orman Regresyonu, en yüksek R-Kare değeri %96 ile en iyi performansı gösteriyor. Modelin bağımsız değişkenler ile açıklanan varyansın yüksek bir yüzdesini açıkladığını göstermektedir.

Linear, Ridge ve Lasso Regresyonları arasında benzer R-Kare değerleri bulunmaktadır, bu durum bu modellerin benzer ölçüde açıklama yeteneğine sahip olduğunu gösterir. Light GBM, diğer modellere göre en düşük R-Kare değeri ile daha zayıf bir performans sergilemektedir.

Hata Metrikleri (MAE, MSE, RMSE): Rastgele Orman Regresyonu, MAE, MSE ve RMSE ölçütlerinde diğer modellere göre daha düşük değerlere sahiptir. Bu durum modelin tahmin sonuçlarının gerçek değerlere daha yakın olduğunu gösterir.

Linear, Ridge ve Lasso Regresyonları arasında benzer hata ölçütleri gözlemlenmektedir. Light GBM, MAE, MSE ve RMSE metriklerinde diğer modellere göre daha yüksek değerlere sahiptir. Bu durum daha düşük tahmin başarısına işaret etmektedir.

MAPE (Ortalama Mutlak Yüzde Hata): Rastgele Orman Regresyonu, %2.86 ile en düşük ortalama mutlak yüzde hata değerini gösterir, bu sonuç tahmin başarısının gerçek değerlerden %2.86 kadar sapma gösterdiğini ifade eder.

Light GBM ise %8.08 oranı ile en yüksek MAPE değerini işaret eder, yüksek oranda hata içerdiğini gösterir.

Tahmin Başarısı: Rastgele Orman Regresyonu, %97.13 ile en yüksek tahmin başarısına sahiptir, yani modelin çoğu durumunu başarıyla tahmin ettiğini gösterir.

Light GBM Regresyonu, %91.91 ile en düşük tahmin başarısına sahiptir, bu da modelin genelde daha düşük bir doğruluk seviyesine sahip olduğunu gösterir.

Open-die forging teknolojisi kullanılarak üretilen aksisimetrik parçaların maliyet tahmini üzerine bir analitik model geliştirilmiş ve tasarım aşamasındaki geometrik özelliklerin detaylı bir maliyet ayrışımı sağlanmıştır [7]. Bu çalışmada otomotiv endüstrisinde kullanılan kalıp maliyetlerini tahmin etmek için çeşitli makine öğrenimi regresyon algoritmalarını değerlendirmiş ve Random Forest Regression'ın en yüksek tahmin başarısını elde ettiğini belirtmiştir. Her iki çalışma da maliyet tahmininde analitik veya makine öğrenimi modellerinin etkili olabileceğini vurgulamaktadır. Makine öğrenimi yöntemlerini kullanarak belirli endüstri uygulamalarında tahmin modelleri geliştirilmesi amaçlanmaktadır. İnanç ve ark.[20] ile bu çalışmada ML algoritmaları, teslimat süresi ve enerji maliyeti tahminleri ile kalıp maliyetleri öngörme amacı taşımaktadır. Her iki araştırma da endüstri uygulamalarında maliyet etkinliği ve verimliliği

artırmaya yönelik uygulamalı sonuçlar elde etmektedir. Ayrıca, her iki çalışmanın kullanılan ML algoritmaları, veri giriş değişkenleri ve performans metrikleri açısından farklılıklar göstermektedir. İnanç ve ark [20] gradient boosting algoritması belirgin bir başarı elde ederken, bu çalışmada Random Forest Regression en yüksek tahmin doğruluğunu sağlamaktadır. Sonuç olarak, her iki çalışma da endüstri spesifik uygulamalarda ML kullanımının potansiyelini vurgulayarak, farklı sektörlerdeki maliyet tahminleme problemlerine odaklanarak benzerlik ve farklılıklarını ortaya koymaktadır. Atalan ve Ark. [21] çalışmasında, acil servis sağlık kaynaklarının tahsisinde kullanılan AdaBoost (AB) algoritmasının, hastaların tedavi sayısı ve bekleme süresi tahmininde diğer algoritmaları aşarak en iyi performansı sergilediği belirlenmiş olup bu çalışma ise farklı maliyet katsayılarına dayalı senaryoların acil servis verimliliği üzerindeki etkilerini değerlendirmeyi amaçlamıştır. Ayrıca otomotiv endüstrisi çalışmasında ise Rastgele Orman Regresyonu'nun kalıp maliyetlerini tahminlemede %98.197'lik yüksek bir doğruluk oranıyla öne çıktığı tespit edilmiştir. Eser [22], üretim ve hizmet sistemlerinde karar almak için veri madenciliği temelli bir metodoloji sunarak, ağaç tabanlı ve komşuluk tabanlı yöntemlerle regresyon modelleri geliştirmeyi önermiştir. Aykırı değer analiziyle doğruluğu artırılan bu modeller, bir sac metal kalıp imalat firmasında üretim sürelerinin tahmin edilmesi üzerine uygulanarak başarılı sonuçlar elde etmiştir. Bu çalışma, otomotiv endüstrisinde kalıp maliyetlerini tahmin etmek için kullanılan çeşitli makine öğrenmesi regresyon algoritmalarını karşılaştırmaktadır. Rastgele Orman Regresyonu'nun %98.197'lik başarı oranıyla öne çıktığı belirtiliyor, bu da kalıp maliyet tahminlerinde yüksek doğruluk sağladığını gösteriyor.

Sonuç olarak, Rastgele Orman Regresyonu diğer modellere göre daha iyi bir performans sergilemektedir. Yüksek r-kare değeri düşük hata değerleri ve yüksek tahmin başarısı ile en doğru tahminleri yapan regresyondur. Diğer modellere göre daha düşük bir performansı ise Light GBM Regresyonu göstermiştir. Model seçimi ise kullanılan ölçütlere ve problemin ihtiyaçlarına bağlı olarak değişiklik göstermektedir.

6. SONUÇ

Bu çalışma regresyon tabanlı makine öğrenme algoritmalarının kalıp tasarımı maliyet tahminindeki performanslarını karşılaştırarak önemli bilgiler sunmuştur. Elde edilen sonuçlar, farklı regresyon algoritmalarının kalıp tasarımı maliyet tahmininde kullanılabilirliği ve doğruluk düzeyini karşılaştırmıştır. Bulgularımız, tahmini başarısını en iyi yapan regresyon ise Rastgele Orman Regresyonu %98.197 oranıyla elde edilmiş regresyon algoritmalarının diğerlerine göre daha iyi performans gösterdiğini göstermiştir.

Sonuç olarak, otomotiv sektöründe kalıp maliyetlerinin tahmininde Rastgele Orman Regresyonu gibi makine öğrenmesi algoritmalarının kullanılması önerilebilir. Bu algoritmalar, sezgilere ve deneyimlere dayalı tahminlerden daha doğru ve güvenilir sonuçlar elde etmeyi sağlayarak üretim sürecinin daha verimli hale gelmesine yardımcı olacaktır. Regresyon tabanlı makine öğrenme algoritmaları, kalıp tasarımı maliyet tahmini konusunda önemli bir araç olabilir ve bu algoritmaların seçimi sürecinde dikkatlice incelenmelidir. Ayrıca endüstriyel süreçlerin daha iyi planlanması ve maliyet optimizasyonu için bilimsel bir temel sağlama amacını taşımaktadır.

Bu sonuçlar, endüstriyel süreçlerde maliyet optimizasyonu ve etkin yönetim konularında uygulama potansiyeli taşımaktadır. Ayrıca, tedarik zinciri planlaması, üretim süreçleri ve proje yönetimi gibi farklı alanlarda da kullanılabilirler. Gelecekteki çalışmalarda, daha fazla veri kaynağı ve geliştirilmiş özellik mühendisliği teknikleri kullanarak tahmin performansını artırmak için bu alanda daha fazla araştırma yapılması önerilmektedir.

Kaynakça

- [1] Naranje, V., Kumar, S., & Hussein, H. M. A. (2014). A knowledge based system for cost estimation of deep drawn parts. *Procedia Engineering*, 97, 2313-2322.
- [2] Ning, F., Shi, Y., Cai, M., Xu, W., & Zhang, X. (2020). Manufacturing cost estimation based on a deep-learning method. *Journal of Manufacturing Systems*, 54, 186-195.
- [3] Viharos, Zs. J., & Mikó, B., (2011). Artificial neural network approach for injection mould cost estimation. *Proceedings of the 44th CIRP Conference on Manufacturing Systems, New Worlds of Manufacturing, Madison, WI, USA* (Vol. 13).
- [4] Florjanič, B., & Kuzman, K., (2012). Estimation of Time for Manufacturing of Injection Moulds Using Artificial Neural Networks-based Model, Preliminary Communication. *Polimeri: časopis za plastiku i gumu*, 33(1), 12-21.
- [5] Rodrigues, A., Silva, F. J., Sousa, V. F., Pinto, A. G., Ferreira, L. P., & Pereira, T. (2022). Using an Artificial Neural Network Approach to Predict Machining Time. *Metals*, 12(10), 1709.
- [6] Cather, H. L., & Chan, K. H., (2014). Integrated Approach to Computer Aided Estimating. *Fourth International Conference on Factory 2000-Advanced Factory Automation* (pp. 349-355). IET.
- [7] Campi, F., Mandolini, M., Favi, C., Checcacci, E., & Germani, M. (2020). An analytical cost estimation model for the design of axisymmetric components with open-die forging technology. *The International Journal of Advanced Manufacturing Technology*, 110, 1869-1892.
- [8] SAP, (2023). *What is Machine Learning?*, Accessed: May. 21,2023, <https://www.sap.com/hk/products/artificial-intelligence/what-is-machine-learning.html>
- [9] IBM, (s.a.). *What is linear regression?*, Accessed: May. 19,2023, <https://www.ibm.com/topics/linear-regression>
- [10] Wikipedia, (s.a.). *Ridge regression*, Accessed: May. 19, 2023, https://en.wikipedia.org/wiki/Ridge_regression
- [11] Greet Learning Team, (2023). *What is lasso regression?*, Accessed: May. 19,2023, <https://www.mygreatlearning.com/blog/understanding-of-lasso-regression/>
- [12] Bramer, M. A. (2017). Principles of data mining, *Springer*, London.
- [13] Maimon, O., & Rokach, L. (2010). Introduction to knowledge discovery and data mining. In *Data mining and knowledge discovery handbook* (pp. 1-15). Boston, MA: Springer Us.
- [14] García-Laencina, P. J., Sancho-Gómez, J. L., Figueiras-Vidal, A. R. (2010). Pattern Classification with Missing Data: A Review, *Neural Computing and Applications*, 19(2), 263-282.
- [15] Birant, D. (2011). Comparison of Decision Tree Algorithms for Predicting Potential Air Pollutant Emissions with Data Mining Models, *Journal of Environmental Informatics*, 17(1), 46-53.
- [16] Ayyadevara, V. K. (2018). Random Forest. In: *Pro Machine Learning Algorithms*, Springer, 105–116. https://doi.org/10.1007/978-1-4842-3564-5_5
- [17] Wang, P. W., Lin, C. J. (2014). Support vector machines. In *Data Classification: Algorithms and Applications*, 187–204. <https://doi.org/10.1201/b17320>
- [18] Muratlar, E. R., (2020). *XGBoost nasıl çalışır? Neden iyi performans gösterir?*, Accessed: Feb. 5,2023, <https://www.veribilimiokulu.com/xgboost-nasil-calisir/>
- [19] Khandelval P., (2017). *Which algorithm takes the crown: Light GBM vs XGBoost?*, Accessed: Feb. 5,2023, <https://www.analyticsvidhya.com/blog/2017/06/whichalgorithm-takes-the-crown-light-gbm-vs-xgboost/>

- [20] İnaç, H.; Ayözen, Y.E.; Atalan, A.; Dönmez, C.Ç. (2022). Estimation of Postal Service Delivery Time and Energy Cost with E-Scooter by Machine Learning Algorithms. *Appl. Sci.*, 12, 12266.
- [21] Atalan, A., Şahin, H., & Atalan, Y. A. (2022). Integration of Machine Learning Algorithms and Discrete-Event Simulation for the Cost of Healthcare Resources. In *Healthcare*, Vol. 10, No. 10, p. 1920.
- [22] Eser, G. (2021). Tahmin Problemleri İçin Regresyon Ağacı Ve Komşuluk Tabanlı Yöntemler Geliştirilmesi: Kalıpcılık Sektöründe Bir Uygulama. Bursa Uludağ Üniversitesi Fen Bilimleri Enstitüsü, Yayınlanmamış Yüksek Lisans Tezi, Bursa.