






PARÇA ISKARTALARININ MAKİNE ÖĞRENMESİ KULLANILARAK AZALTILMASI: OTOMOTİV SEKTÖRÜNDE BİR UYGULAMA

Emine EŞ YÜREK *
Betül YAĞMAHAN *
Burak Celal AKYÜZ *
Ebubekir Sıddık SAMAST *
Nezire Dilan ÇETREZ *

Alınma: 06.07.2021; kabul: 09.03.2022

Öz: Bu çalışmada, enjektör imal eden bir firmanın taşlama makinesindeki insan faktörüne bağlı ıskartaların azaltılması amaçlanmıştır. İnsan faktörüne bağlı ıskartalar, makinenin taş değişimi, arıza gibi çeşitli nedenlerle durmasından sonra makine operatörünün, makine ve ürün parametrelerinde yaptığı ayarlamalardan kaynaklanmaktadır. İskartaların azaltılması amacıyla iki aşamalı bir çözüm yaklaşımı önerilmiştir. İlk aşamada, makine öğrenmesi kullanılarak kalite tahminlenmiştir. Kalite tahminleme, bir sınıflandırma problemi olarak ele alınmıştır. Farklı sınıflandırma algoritmaları modellenerek en yüksek performansa sahip Destek Vektör Makineleri (DVM) algoritması seçilmiştir. İkinci aşamada ise, DVM kullanılarak kaliteli ürün ile sonuçlanması beklenen uygun parametre değerleri belirlenmiştir. Güncel veri dikkate alınarak parametre değerlerinin otomatik olarak revize edilmesi ve operatöre sunulması amacıyla bir öneri sistemi geliştirilmiştir. Bu öneri sistemi ile, taşlama işleminin insan etmenine olan bağlılığının ortadan kaldırılarak dijitalleşmesi amaçlanmıştır.

Anahtar Kelimeler: Iskarta azaltma, Kalite sınıflandırma, Makine öğrenmesi, Parametre ayarı, Dijitalleşme

Reducing Part Rejects Using Machine Learning: A Case Study In Automotive Industry

Abstract: This study aims to reduce the human-based rejects in the grinding machine of a company that manufactures injectors. Human-based rejects are caused by the operator's adjustments to the machine and product parameters after the engine stops due to some reasons such as stone change or breakdown. A two-stage solution approach is proposed to reduce rejects. In the first stage, quality is predicted using machine learning. Quality prediction is addressed as a classification problem. Various classification algorithms are modeled, and the outperforming Support Vector Machine (SVM) is selected. In the second stage, the proper parameter values expected to result in a quality product are determined using SVM. A system is developed to automatically revise the parameter values by considering the updated data and recommend them to the operator. This system aims to digitalize the grinding process by removing its dependence on the human factor.

Keywords: Reject reduction, Quality classification, Machine learning, Parameter adjustment, Digitalization

* Bursa Uludağ Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, 16059 Nilüfer Bursa

İletişim Yazarı: Emine Eş Yürek (esyurek@uludag.edu.tr)

1. GİRİŞ

Bilgi teknolojisindeki gelişmeler, işletmelerin bu teknolojiyi geleneksel imalat uygulamalarına entegre etmesini sağlayarak çok yönlü bir değişim sürecini başlatmıştır. Dijitalleşmenin imalat sistemlerine entegrasyonu; işletmelere ürün, makine, süreç, operatör, hammadde, çevre şartları gibi pek çok veriyi toplama ve depolama imkanı sağlamıştır. İşletmeler açısından veri toplama ve depolama kolaylaşsa da çeşitli faktörler arasındaki ilişkilerin anlaşılması ve bu ilişkilerin etkilerinin belirlenmesi daha karmaşık hale gelmiştir (Doğan ve Birant, 2021). Geleneksel yöntemlerle çözümlenemeyecek büyüklükteki bu karmaşık veriyi işletmeler açısından faydaya dönüştürme noktasında ise makine öğrenmesi önemli bir role sahip olmaktadır. Karmaşık ilişkileri anlama ve tanımadaki başarısı, makine öğrenmesine ilgiyi arttırmış ve çok geniş bir alanda kullanılmasını sağlamıştır. Doğan ve Birant (2021), makine öğrenmesi konusunda son 20 yılda gerçekleştirilmiş çalışmaları inceleyerek bunları çizergeleme, görüntüleme, kalite ve arıza olmak üzere dört başlık altında toplamış ve çalışmaların ağırlıklı olarak kalite konusunda yoğunlaştığını belirlemiştir. Köksal ve diğ. (2011), kalite konusunu; ürün/süreç kalitesini tanımlama (kalite karakteristiklerini belirleme), kalitenin tahmini, kalitenin sınıflandırılması ve parametre optimizasyonu olmak üzere dört faaliyet grubu altında incelemiştir. Söz konusu çalışma, özellikle bu dört kalite iyileştirme konusunun hangi sektörlerde ne sıklıkta çalışıldığı ve hangi veri madenciliği yöntemlerinin uygulandığı sorularına cevap bulmaya odaklanmıştır.

Ürün kalitesini ölçmek için kullanılan geleneksel yöntem istatistiksel kalite kontroldür. Bu yöntemde, periyodik olarak alınan numunelerin kalitesi kontrol edilir ve örneklemden yararlanılarak ana kütle kalitesine dair çıkarımda bulunulur. Fakat istatistiksel kalite kontrol, ürünlerin kalite spesifikasyonları içinde olmasına çalışsa bile zaman zaman spesifikasyonların dışında ürünler çıkabilir. Bu ürünler iadelere, dolayısıyla kalitesizlik maliyetlerine neden olmaktadır. Ürün kalitesini ölçmek için bir diğer alternatif özellikle seri imalat sistemlerinde, hat sonunda her ürün için kalite kontrolün gerçekleştirilmesidir. Bu yöntem hem zaman kaybına hem de yüksek operasyonel maliyetlere neden olmaktadır. Özellikle yüksek maliyetli imalat işlemlerinin gerçekleştirildiği sektörlerde, süreç içerisinde oluşan ıskartaların hat sonundaki kalite kontrol aşamasına kadar işlenerek gelmesi yerine kalitenin süreç boyunca izlenmesi ve gerekli tedbirlerin alınması önem arz etmektedir. Kalitenin takibi ise zahmetli ve maliyetli olabilmektedir. Makine öğrenmesi, ürün kalitesini ölçerek izlemek yerine, çeşitli faktörlerin ürün kalitesi üzerindeki etkilerini anlayarak kalite tahminlemede önemli bir role sahip olmaktadır. Chou ve diğ. (2010), yarı iletken sektöründe kaliteyi tahmin etmek amacıyla dört adımdan oluşan bir sistem önermektedirler. İlk olarak, sensörler yardımıyla anlık veri toplanmakta ve kritik öneme sahip öznitelikler belirlenmektedir. İkinci adımda, ürünün kalitesinde kritik öneme sahip faktörler ile ürün kalitesi arasındaki ilişki destek vektör makinesi (DVM) ile modellenmektedir. Üçüncü adımda, destek vektör sınıflandırıcısının parametre değerleri genetik algoritma ile belirlenmektedir. Dördüncü adımda ise, çeşitli performans kriterlerine göre tahminin güvenilirliği düştüğünde artımlı öğrenme devreye girerek DVM parametreleri yeniden ayarlanmaktadır. Schorr ve diğ. (2020), hidrolik valf üreten bir seri imalat sistemindeki matkap makinesinde işlenen parçalara ait veriyi kullanarak rassal orman ile tolerans aralığının son derece dar olduğu delme işleminin kalitesini tahmin etmişlerdir. Garcia ve diğ. (2019), ekstrüzyon yönteminin kalitesini k -en yakın komşuluk (k -EYK), doğrusal regresyon ve DVM kullanarak tahmin etmişlerdir. Tahminleme sonucunda, mesafe ağırlıklı k -EYK yöntemi ile doğrusal ve yarıçap temelli fonksiyon tabanlı DVM'nin en iyi sonuçları verdiği gösterilmiştir. Bai ve diğ. (2018) ise, kalite tahmininde yüzeysel ve derin öğrenme yöntemlerini tahmin performansı açısından karşılaştırmışlardır. İki farklı büyüklükteki veri seti üzerinde yapılan çalışma, tahmin hatası bakımından derin öğrenmenin yüzeysel öğrenmeden daha iyi sonuç verdiğini göstermektedir.

Kalite tahmini, kalitenin nümerik değerlerle karakterize edildiği durumlarda; kalite sınıflandırması ise kalitenin nominal veya ordinal değerlerle karakterize edildiği durumlarda

uygulanabilmektedir (Köksal ve diğ., 2011). Ramulu ve diğ. (2019) yukarıdaki çalışmalardan farklı olarak, ekstrüzyon işlemi sonucundaki ürünün kalitesini sınıflandırmışlardır. Sınıflandırıcılardan k -EYK, karar ağacı, kural çıkarsama ve oylama modelleri uygulanarak yapılan karşılaştırmada en iyi sonucun k -EYK ile elde edildiği gösterilmektedir. Mohammadi ve Wang (2016), aşınmaya dayanıklı malzeme imalatı yapan bir firmada ürünün kalitesini DVM kullanarak sınıflandırmışlardır. DVM, doğrusal ve yarıçap temelli olmak üzere iki tip çekirdek fonksiyon kullanılarak uygulanmıştır. Yapılan karşılaştırmada, yarıçap temelli çekirdek fonksiyonun performansının doğrusal çekirdek fonksiyona göre daha yüksek olduğu gösterilmektedir.

İmalat süreçlerinde, süreç parametreleri ile ürün kalitesi arasında çoğu zaman karmaşık bir ilişki vardır. Üretim başlamadan, süreç parametrelerine ait değerlerin seçimi için sıklıkla simülasyona başvurulur (Sapounas ve diğ., 2020). Fakat çeşitli etkenlerden dolayı zaman içinde operatörler parametre değerleri üzerinde değişiklik yapmak durumunda kalırlar. Bu durumda ise parametre değerlerinin seçimi genellikle kural-tabanlı, deneyime dayalı veya deneme yanılma yöntemi ile gerçekleştirilmektedir (Strasser ve diğ., 2018). Bu durum, hem bu konuda deneyimli operatör gerektirdiği hem de hataya çok açık olduğu için zor bir iştir. Ayrıca, süreç parametreleri geleneksel yöntemlerle belirlendiğinde, çoğu zaman bütün bir partinin üretimi süresince aynı değerler kullanılır (Strasser ve diğ., 2018). Oysaki ürün kalitesi; makine, süreç ve ürüne bağlı olarak değişebilir. Bu da işlenecek her parçada, makine, ürün ve sürecin o anki durumunu gösteren parametrelerin dikkate alınmasını gerektirir. Adesanya ve diğ. (2011), termoplastik kabloların üretildiği ekstrüzyon işlemine ait süreç parametrelerini makine öğrenmesi ile belirlemişlerdir. Girdi ve çıktı parametreleri arasındaki ilişkinin yapay sinir ağları (YSA) ile modellendiği çalışmada, hedef kalite değerine ulaşmayı sağlayacak girdi parametrelerini belirlemek için regresyon modeli elde edilmiştir. Gao ve diğ. (2018) ise, plastik enjeksiyon işlemindeki parametreleri optimize etmek amacıyla bir hatanın, bu hata ile ilişkili parametrelerin iki zıt değeri arasında sıkıştığı düşüncesinden yola çıkarak sınıflandırma yapmaktadırlar. Jurkovic ve diğ. (2018), yüksek hızlı tornalama işleminde üç farklı kesme parametresini belirlemek için destek vektör regresyonu, polinom regresyon ve YSA kullanmışlardır. Parametrelerin ikisinde polinom regresyon, diğerinde YSA daha yüksek performans göstermiştir. Teixidor ve diğ. (2015), mikro lazer işlemede süreç parametrelerini belirlemek için k -EYK, YSA, karar ağacı ve doğrusal regresyon modellerini kullanmışlardır. Yapılan çalışma, en iyi performansın YSA ve karar ağacı ile elde edildiğini göstermektedir.

Weichert ve diğ. (2019), makine öğrenmesi ile kalite iyileştirmesini, üretim sürecinde parametre ayarlama kriterine göre iki başlık altında sınıflandırmışlardır. İlk sınıfta yer alan ve üretim parametrelerini optimize etmeyen çalışmalar kaliteyi tahminlemede ve bu tahmini kalitesiz ürünle sonuçlanacak bir süreci erken sonlandırarak zaman ve maliyet tasarrufu sağlamakta kullanılmaktadır. İkinci sınıf kapsamına giren çalışmalarda ise ilk olarak makine öğrenmesi ile ürün kalitesi tahmin edilmekte, daha sonra çeşitli optimizasyon yöntemleri ile üretim parametreleri optimize edilmektedir. Silva ve diğ. (2014), frezeleme işleminin parametrelerini optimize etmek amacıyla YSA ve optimizasyon tekniklerini bir arada kullanmışlardır. Bunun nedeni, çıkan ürünün kalitesinin kesici takımların durumundan bağımsız olmamasıdır. Söz konusu çalışma, her işlemde sonra sensörler yardımıyla toplanan kesici takım ölçümlerini dikkate alarak takım aşınımı, yüzey pürüzlülüğü ve takımların kalan ömrünü geri yayılım algoritması ile tahmin etmiştir. Daha sonra tahmin edilen bu değişkenlerin yer aldığı tahmini üretim maliyeti, melez bir genetik algoritma ile minimize edilmiştir. Ciurana ve diğ. (2009), lazer mikro işlemede süreç parametreleri ile kalite karakteristikleri arasındaki ilişkiyi YSA ile modelledikten sonra minimum yüzey pürüzlülüğü ve minimum hacimsel hatayı verecek parametre seçimini parçacık sürü optimizasyon algoritması ile gerçekleştirmişlerdir.

Kalitenin, makine öğrenmesi kullanılarak iyileştirildiği çalışmalar çoğunlukla tek bir süreç üzerinde yoğunlaşmaktadır (Arif ve diğ., 2013). Oysaki çoğu zaman bir hammadde, nihai ürün haline gelene kadar ardışık birçok işlemde geçmektedir. Dolayısıyla nihai ürüne gelene kadar

her bir iş istasyonundaki yarı mamulün kalitesi sadece bulunduğu istasyondaki işlemde değil, önceki istasyonlardaki faktörlerden de etkilenmektedir. Arif ve diğ. (2013) çok aşamalı bir imalat sürecinde, öncelikle temel bileşenler analizi ile ürün karakteristikleri ve süreç değişkenleri arasındaki ilişkiyi ortaya koymuş, daha sonra farklı karar ağacı algoritmaları ile nihai ürünün kalitesini tahmin etmişlerdir. Bu yöntem ile çok aşamalı üretim sistemlerinde tek aşamalı sistemlerde uygulanan tahmin modellerine göre daha iyi sonuç alındığı gösterilmiştir. Strasser ve diğ. (2018), çok adımlı bir metal işleme sürecinde parametre optimizasyonu için regresyonu kullanmışlardır. Girdi ve çıktı değerleri arasındaki ilişki regresyon ile modellendikten sonra elde edilen fonksiyon parametre tahmininde kullanılmıştır.

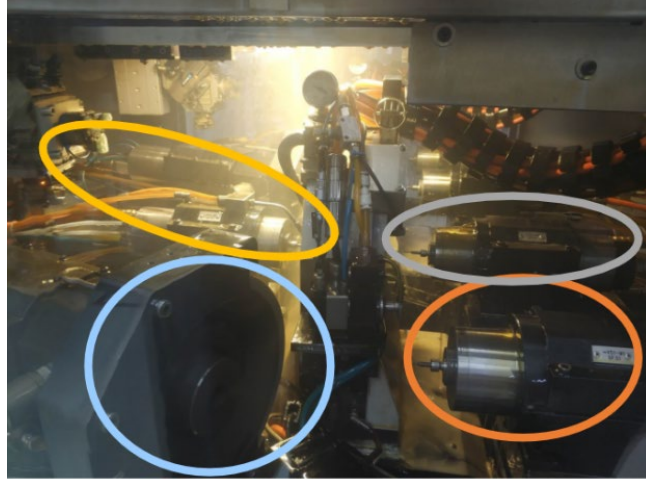
Bu çalışma, enjektör üreten bir firmadaki taşlama makinesinde ortaya çıkan ıskarta sayılarının azaltılmasını amaçlamaktadır. İncelenen çalışmalar, kalite iyileştirme çabalarında makine öğrenmesinin sıklıkla başvurulan, başarısını kanıtlamış bir yöntem olduğunu gösterirken, gerçekleştirilen çalışma için de önemli bir motivasyon kaynağı olmuştur. Bu çalışmada, ilk olarak girdi parametreleri ile ürün kalitesi arasındaki ilişki farklı makine öğrenmesi yöntemleri ile modellenmektedir. Daha sonra en yüksek performansa sahip yöntem seçilerek girdi parametrelerine ait değerler belirlenmekte ve makine operatörünün kullanımı için bir öneri sistemi geliştirilmektedir. Bu çalışmanın katkısı aşağıdaki gibi özetlenebilir:

- Çalışma, bir gerçek hayat problemini ele almakta, bu problemin çözümü için makine öğrenmesi ve öneri sistemi içeren bütünlük bir yaklaşım geliştirmektedir.
- Önerilen çözüm yaklaşımı, taşlama işlemi için spesifik olmayıp süreç parametrelerine ait değerlerin seçiminin insan faktörüne bağımlılığını azaltmaktadır. Bu yönüyle, farklı süreçlerde uygulanabilecek ve dijital dönüşüme katkı sağlayacak genel bir yaklaşımdır.
- Gerçek veri kullanılarak elde edilen test sonuçları, önerilen yaklaşımın ıskartaları %56 oranında önleyebilme potansiyelini ortaya koymaktadır.

Bu çalışma beş bölümden oluşmaktadır. Kalite iyileştirmede makine öğrenmesi çalışmalarına odaklanmış kaynak araştırmasına bu bölümde yer verilmiştir. İkinci bölümde, bu çalışmada ele alınan problem anlatılmaktadır. Üçüncü bölümde problemin çözümüne yönelik metodolojiye yer verilmektedir. Gerçek hayat problemini ele alan çalışmada elde edilen sonuçlar ise dördüncü bölümde anlatılmaktadır. Beşinci bölümde ise genel sonuçlar ve değerlendirmelerden bahsedilmektedir.

2. PROBLEM TANIMI

Bu çalışma, enjektör üreten bir firmanın taşlama hattında ortaya çıkan kalite problemini ele almaktadır. Söz konusu taşlama hattında enjektör alt parçasının taşlama işlemi gerçekleştirilmektedir. Parça ön taşlama işleminden geçerek taşlama makinelerine gelmekte ve burada çap, açı ve dış geometri taşlama işlemlerinden geçmektedir. Bu işlemler sırasında dört tip taş kullanılmaktadır (Şekil 1). Taşlama makinesi içerisinde, robotik kol tarafından getirilen parça ilk önce Taş A ve Taş B tarafından eş zamanlı olarak taşlanmaktadır. Taş A parçanın şaft çapını, Taş B ise parçanın iç çapını taşlamaktadır. Bu taşlar, parçanın şaft çapını ve iç çapını istenen tolerans değerleri sağlanana dek işlemektedir. Şaft ve iç çap taşlaması biten parça, dış çapı ve açıları taşlanmak üzere Taş D'ye iletilmeden önce Taş C tarafından taşlanmaktadır. Taş C; Taş D'nin işlem yükünü azaltmak, kullanım ömrünü arttırmak ve son ürün kalitesini arttırmak için parçaya ön taşlama işlemi uygulamaktadır. Ön taşlama işlemi bittikten sonra parça Taş D'ye iletilir. Taş D, parçanın son taşlama işlemini yapmaktadır. Bu taş parçanın açı değerlerini işledikten sonra tutucu mekanizma parçayı robotik kola geri iletmektedir. Robotik kol parçayı hassas ölçüm yapılması için hassas ölçüm masası olarak adlandırılan havalı ölçüm sistemine taşımaktadır. Parçanın kritik ölçüleri burada ölçülmekte, eğer tolerans aralıkları içinde değilse bu parça robotik kol aracılığıyla ıskarta hücrelerine bırakılmaktadır. Eğer ölçüler toleransları karşılıyorsa parça, taşlama işleminden kalan taş izleri ve yüzey pürüzlerinin giderilmesi amacıyla fırçalama işlemine gönderilmektedir.

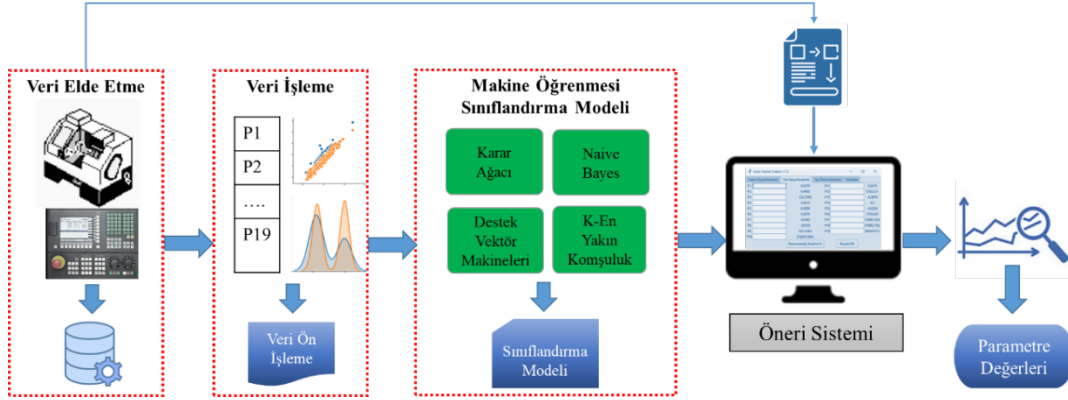


Şekil 1:
Taşlama işleminde kullanılan dört tip taş

Taşlama işlemi sonunda ortaya çıkan ıskartalar makine ve insan faktörüne bağlı olmak üzere iki başlık altında sınıflandırılmaktadır. Makine faktörüne bağlı ıskartalar, makinede meydana gelen herhangi bir aksaklık veya arıza durumunda oluşur ve makinenin durmasına neden olurlar. Taşlama işleminde kullanılan taşların, ekonomik ömrünü tamamladıklarında değiştirilmesi gerekmektedir. Makine duruşu, taş değişimi gibi nedenlerle üretime ara verildiğinde, makine tekrar üretime geçmeden önce belirli parametrelere ayar yapılmaktadır. Parametrelere atanacak değerlerin seçimi ise operatör tarafından, kendi deneyimine dayanarak deneme yanılma şeklinde gerçekleştirilmekte ve yapılan ayar işlemi kaliteli parça çıkana kadar devam etmektedir. İşte bu süreçte kaliteli parça çıkana dek oluşan ıskartalar, insan faktörüne bağlı ıskartalar olarak adlandırılmaktadır. Taşlama makinesinde oluşan toplam ıskarta miktarının %60'ını bu tür ıskartalar oluşturmaktadır. Bu çalışmada insan faktörüne bağlı ıskartalara odaklanarak, bunların azaltılması amacıyla ayar sürecinde operatörün hata payını azaltacak bir öneri sistemi geliştirilmesi amaçlanmıştır.

3. METODOLOJİ

Bu bölümde, taşlama makinesinde ortaya çıkan ıskarta sayılarının azaltılması amacıyla önerilen metodolojiden ve kullanılan makine öğrenmesi algoritmalarından bahsedilmektedir. Şekil 2'de görülmekte olan önerilen bütünleşik metodolojinin ilk aşaması, makinalarda ıskarta oluşması ile ilgili parametrelere ait tahminleme için kullanılacak gerçek verilerin elde edilmesidir. Sonrasında bu parametre değerlerinin birbirleri ile ilişkileri incelenmiş ve veriler üzerinde ön işleme gerçekleştirilmiştir. Ardından makine öğrenme algoritmaları kullanılarak bir sınıflandırma modeli belirlenmiştir. Son olarak geçmiş parametre verilerini içeren bir veri seti sayesinde çıktılarının istenilen tolerans aralıklarında kalabilmesi için en uygun girdi parametre değerlerinin ne olması gerektiği konusunda kullanıcıya yardımcı olacak bir öneri sistemi geliştirilmiştir.



Şekil 2:
Önerilen metodoloji

Yapay zekanın alt kümesi olan makine öğrenmesi, makinenin belli bir şekilde programlama yapmadan geçmiş bilgilerden otomatik olarak öğrenmesini sağlayan bir araçtır (Paturi ve Cheruku, 2020). Öğrenme şekillerine dayalı olarak, makine öğrenme yöntemleri denetimli öğrenme, denetimsiz öğrenme, yarı denetimli öğrenme ve takviyeli öğrenme olarak sınıflandırılmaktadır (Kumar ve diğ., 2019).

Denetimli öğrenme yaklaşımları, bir veri kümesi üzerinde bir algoritma eğitimi gerçekleştirir, bu sayede her eğitim noktasının kendisine eklenmiş bir "etiketi" vardır. Bu etiket, eğitim noktasının ait olduğu belirli sınıfı gösterir. Denetimli algoritmalar daha sonra bu sınıfları ayıran karar sınırlarını belirlemeye çalışır. Diğer yandan, denetimsiz yaklaşımlar, etiketlenmemiş verilerdeki anahtar örüntüleri tanımlamak için kullanılır. Yarı denetimli yaklaşımda, kullanıcı bazı etiketli verileri ve bazı etiketsiz verileri aynı anda sağlar. Model daha sonra verilen etiketli ve etiketsiz bilgilerin özelliklerine göre bir karar sınırı oluşturmaya çalışabilir ve verileri kümeler halinde sınıflandırır (Okaro ve diğ., 2019). Veri geri bildirimine dayalı makine öğrenmesine takviyeli öğrenme denir. Eğitim verileri (ödül ve cezalar şeklinde) yalnızca dinamik bir ortamda bir yapay zeka temsilcisine geri bildirim olarak verilir (Zhang, 2020).

Bu çalışmada, girdi parametre değerleri kullanılarak, bunlara karşılık oluşan çıktıların sınıflandırması gerçekleştirildiği için denetimli öğrenme uygulanmıştır. Bu amaçla denetimli öğrenme algoritmalarından karar ağacı, Naive Bayes, DVM ve k -EYK (3-EYK, 5-EYK ve 7-EYK) kullanılmıştır.

3.1. Karar Ağacı

Karar ağacı, okunabilirliği artırmak için eğer-öyleyse kurallar setine dayanarak sınıflandırma oluşturan denetimli makine öğrenmesi yaklaşımıdır. Bir karar ağacı, yaprak düğümler (nihai sonuçlar) ve karar düğümleri (alternatifler arasında seçim) olarak adlandırılan iki tür düğüm içerir. Karar ağacı, eğitim verilerinden çıkarılan karar kurallarına dayalı bir eğitim modeli oluşturarak bir sınıfı veya hedefi tahmin etmek için kullanılır (Kumar ve diğ., 2019).

Diğer makine öğrenimi teknikleri gibi, karar ağacı bir eğitim ve test aşaması gerektirir. Giriş olarak n özellikli bir X vektörü, buna karşılık gelen çıktı olarak Y etiketi olmak üzere (X, Y) m çiftlerinin S eğitim seti şu şekildedir (Lahouar ve Slama, 2015):

$$X = (x_1, \dots, x_j, \dots, x_n)^T, X \in \mathbb{R}^n \quad (1)$$

$$Y \in \mathbb{R} \quad (2)$$

$$S = \{(X_1, Y_1), \dots, (X_m, Y_m)\} \quad (3)$$

Eğitim süreci, belirli bir sonlandırma kriteri karşılanana kadar özellikleri, yinelemeli olarak farklı Y etiketlerine sahip düğümlere bölerek bir h tahmini oluşturur. Bu kriter, genellikle farklı etiketlere sahip alt düğümlere sahip olmanın mümkün olmadığı durumdur. Terminal düğümleri ağacın yaprakları olarak adlandırılır ve bunlar farklı olası Y etiketleri gösterir. Eğitim aşaması, tahmin h kullanılarak herhangi bir yeni özellik vektörünün etiketinin tahmin edilmesinden oluşur (Lahouar ve Slama, 2015):

$$Y = h(X) \quad (4)$$

3.2. Naïve Bayes

Naïve Bayes, özellikler arasında güçlü bir bağımsızlık varsayımının olduğu Bayes teoremine dayalı basit olasılıksal sınıflandırıcıdır (Siryani ve diğ., 2017).

$\vec{x} = (x_1, x_2, \dots, x_n)$ bir özellik vektörü ve $C = (c_1, c_2, \dots, c_m)$ bu vektörün ait olabileceği tüm olası kategorileri gösterebilir. Naïve Bayes sınıflandırıcı, her \vec{x} için p_1, p_2, \dots, p_m olasılıklarını hesaplamaktır. Burada p_j , \vec{x} özellik vektörünün c_j kategorisine ait olma olasılığıdır. (p_1, p_2, \dots, p_m) değerini belirleyerek, \vec{x} özellik vektörünün hangi kategoriye ait olduğu belirlenir. Sınıflandırma problemi için Eş. 5 ile elde edilen değer maksimumunu bulunur (Peres ve diğ., 2019):

$$P(c_j | x_1, x_2, \dots, x_n) = \frac{P(x_1, x_2, \dots, x_n | c_j) P(c_j)}{P(c_1, c_2, \dots, c_m)} \quad (5)$$

Burada $P(c_j)$, rastgele bir örneğin c_j kategorisine ait olma olasılığını belirtir. $P(x_1, x_2, \dots, x_n | c_j)$, eğitim örneğinin c_j kategorisine ait olduğunu biliyorsak, c_j kategorisinin $\vec{x} = (x_1, x_2, \dots, x_n)$ özellik vektörünü içermesinin olasılığıdır. $P(c_1, c_2, \dots, c_m)$ olası tüm kategorilerin birleşik olasılığıdır. Verilen tüm kategoriler için payda $P(c_1, c_2, \dots, c_m)$ bir sabittir, bu nedenle Eş. 5 şu şekilde basitleştirilebilir:

$$c_{NB} = \underset{c_j \in C}{\operatorname{argmax}} P(x_1, x_2, \dots, x_n | c_j) P(c_j) \quad (6)$$

Birleşik olasılık dağılımı, her terimin olasılığının çarpımına eşittir. Bu nedenle:

$$P(x_1, x_2, \dots, x_n | c_j) = \prod_{i=1}^n P(x_i | c_j) \quad (7)$$

Gerçekte sınıflandırma problemi için sadece aşağıdaki değer bulunması ile ilgilenilir:

$$c_{NB} = \underset{c_j \in C}{\operatorname{argmax}} P(c_j) \prod_{i=1}^n P(x_i | c_j) \quad (8)$$

3.3. Destek Vektör Makinaları

DVM algoritması, verilerin sınıflarını belirlemek için sonlu boyutlu uzayda hiper-düzlemler oluşturur. Daha büyük sınırlar (marjlar) daha düşük bir hataya karşılık geldiğinden, iyi bir sınıflandırma, herhangi bir sınıfın en yakın eğitim-veri noktasına olan maksimum mesafeye sahip hiper-düzlem tarafından elde edilir (Peres ve diğ., 26).

Veriler, pozitif ve negatif olmak üzere iki sınıfa sahip olduğunda, M örnek sayısı olmak üzere x_i ($i = 1, 2, \dots, M$) verisi için, pozitif sınıf $y_i = 1$ ve negatif sınıf $y_i = -1$ etiketleriyle ilişkilendirilir. Doğrusal veri durumunda, verilen verileri ayıran $f(\mathbf{x}) = 0$ hiper-düzlemini belirlemek mümkündür (Widodo ve Yang, 2007):

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b = \sum_{j=1}^M w_j x_j + b = 0 \quad (9)$$

Burada, \mathbf{w} , M -boyutlu vektör ve b , bir sabittir.

\mathbf{w} vektörü ve b sabiti, ayırıcı hiper-düzlemin konumunu tanımlamak için kullanılır. Karar fonksiyonu, giriş verilerini pozitif ve negatif olarak sınıflandıran ayırıcı hiper-düzlem oluşturmak için $f(\mathbf{x})$ işareti kullanılarak oluşturulur. Bir ayırıcı hiper-düzlem, aşağıdaki kısıtları sağlamalıdır:

$$\begin{aligned} f(\mathbf{x}) &= 1 \text{ eğer } y_i = 1 \\ f(\mathbf{x}) &= -1 \text{ eğer } y_i = -1 \end{aligned} \quad (10)$$

Bu kısıtlar aşağıdaki gibi kısaltılabilir:

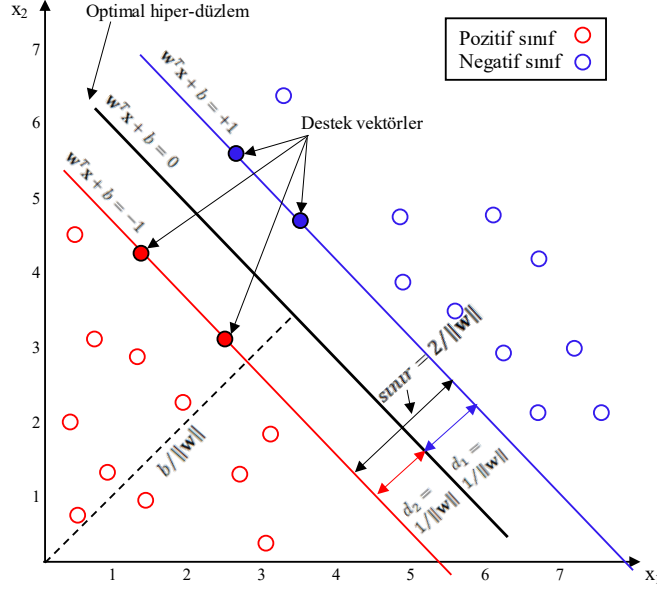
$$y_i f(x_i) = y_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \quad i = 1, 2, \dots, M \quad (11)$$

Düzlem ile en yakın veri arasındaki maksimum mesafeyi yani maksimum sınırı oluşturan ayırıcı hiper-düzlem, optimum ayırıcı hiper-düzlem olarak adlandırılır (Widodo ve Yang, 2007). Şekil 3'de görüldüğü gibi, doğrusal olarak iki sınıfa (binary class) ayrılabilen sınıflandırma problemi için DVM sınırı, aşağıdaki gibi d_1 ve d_2 'nin toplamı ile ifade edilir (Tharwat, 2019):

$$\text{sınır} = d_1 + d_2 = \frac{2}{\|\mathbf{w}\|} \quad (12)$$

d_1 ve d_2 sırasıyla birinci ve ikinci düzlemden hiper-düzleme olan mesafeyi gösterir. DVM sınıflandırıcısında, sınır genişliği Eş. 13 ile maksimize edilir:

$$\begin{aligned} &\min \frac{1}{2} \|\mathbf{w}\|^2 \\ \text{Öyle ki;} & \quad y_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \quad \forall i = 1, 2, \dots, M \end{aligned} \quad (13)$$

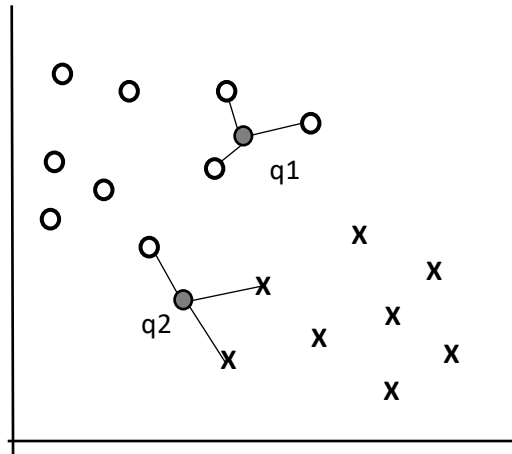


Şekil 3:
İkili sınıflandırma problemine bir örnek (Tharwat, 2019)

3.4.k-En Yakın Komşuluk

k -EYK, yeni örnek için sınıflandırmanın, eğitim setindeki en benzer örneklerle karşılaştırılarak belirlendiği örnek tabanlı öğrenme yöntemidir (Larose, 2015). Bu sınıflandırmada örnekler en yakın komşularının sınıfına göre sınıflandırılır. Burada k sınıfın belirlenmesinde kullanılan en yakın komşu örnek sayısıdır.

Şekil 4’de iki boyutlu bir özellik uzayında iki sınıflı bir problem üzerinde 3-EYK sınıflandırıcı örneği görülmektedir. Burada, q_1 örneği için verilen karar basittir, en yakın üç komşusu da “O” sınıfındadır, bu nedenle bu örnek “O” olarak sınıflandırılır. q_2 örneği için durum, “X” sınıfından iki komşusu ve bir “O” sınıfına sahip komşusu olduğundan biraz daha karmaşıktır. Hangi sınıfa ait olacağı ile ilgili karar değerlendirilmelidir. Yani k -EYK sınıflandırmasının iki aşaması vardır; birincisi en yakın komşuların belirlenmesi ve ikincisi bu komşuları kullanarak sınıfın belirlenmesidir (Cunningham ve Delany, 2020).



Şekil 4:
2-boyutlu uzayda 3-EYK sınıflandırıcı (3-NN classifier in 2-dimensional space) (Cunningham ve Delany, 2020)

Yakınlık, Öklid mesafesi gibi bir uzaklık ölçüsü cinsinden tanımlanır. n -boyutlu uzayda iki nokta arasındaki Öklid mesafesi, $X_1 = (x_{11}, x_{11}, \dots, x_{1n})$ ve $X_2 = (x_{21}, x_{22}, \dots, x_{2n})$ için (Han ve diğ., 2012):

$$\text{uzaklık}(X_1, X_2) = \sqrt{\sum_{i=1}^n (x_{1i} - x_{2i})^2} \quad (14)$$

Komşulukların belirlenmesinin ardından basit çoğunluk oylama (simple majority voting) veya mesafe ağırlıklı oylama (distance weighted voting) yöntemleri gibi yöntemlerle sınıflar belirlenebilir (Cunningham ve Delany, 2020).

4. UYGULAMA

Taşıma hattından çıkan ürünün kalitesini gösteren hedef değişken “Sağlam” ve “Iskarta” olmak üzere iki kategorik değer alabilmektedir. Köksal ve diğ. (2011), kalite tahminleme üzerine yapılan çalışmaları kalite tahmini ve kalite sınıflandırma olmak üzere iki grupta incelemiş, kalitenin nümerik değerlerle karakterize edilmesini kalite tahmini, kategorik değerlerle karakterize edilmesini ise kalite sınıflandırma olarak tanımlamıştır. Bu tanımlamaya göre, gerçekleştirilen bu çalışmadaki kalite tahminleme, kalite sınıflandırma grubuna girmektedir. Ürün kalitesinin sağlam ya da ıskarta olarak etiketlenmesine neden olan girdi parametrelerinin belirlenmesi için firmadaki uzman görüşleri dikkate alınmış ve toplamda 19 parametrenin kalitedeki değişimlere neden olabileceği sonucuna ulaşılmıştır. Bunlardan üçü taşıma aşamasına gelmeden ön işlemlerden geçen ürünün iç çapı ile ilgili parametreler olup diğerleri kesme hızı, x-y-z açılı, talaş derinliği gibi makineye bağlı parametrelerdir.

Çalışmanın gerçekleştirildiği firmada 14 adet taşıma makinesi bulunmaktadır. İlk olarak, bu makineler ıskartalar açısından karşılaştırılmış ve ıskarta oranı en yüksek makine pilot makine olarak seçilmiştir. Seçilen taşıma makinesinden manuel olarak 441 adet veri toplanmıştır. Toplanan verinin 400 adedi sağlam sınıfına, 41 adedi ise ıskarta sınıfına aittir. Veri kümesindeki her bir satır, 19 girdi parametresi ve bir hedef değerinden oluşmaktadır. Girdi parametreleri nümerik, hedef değer ise kategorik değişkenlerdir.

Veri analizi ile ilgili çalışmalarda sıklıkla verinin temizlenmesi veya dönüştürülmesi gibi işlemlere ihtiyaç duyulur. Firmanın taşıma hattında tutmuş olduğu veri, kalite sınıflandırma için ihtiyaç duyulan parametrelerden farklı olduğundan, bu çalışma için manuel olarak tutulmuştur. Veriler manuel tutulduğundan dolayı eksik veri gibi temizleme işlemi gerektirecek durumlar söz konusu değildir. Veri ön işleme aşamasında, normalizasyon işlemi gerçekleştirilerek parametrelerin birbirlerine sayısal üstünlüklerini önlemek amacıyla 0 ve 1 değerleri arasında olması sağlanmıştır. Bu amaçla, min-maks normalleştirilmesi uygulanmıştır.

Firma verileri kullanılarak yapılan uygulamalar, 2,40 GHz çift çekirdekli 8 GB RAM ile desteklenen Intel Core i5-6300 işlemcili bilgisayar ile gerçekleştirilmiştir. Tahminleme algoritmalarının kodlanmasında Python v3.6.9 yazılımı ve Scikit-learn kütüphanesi kullanılmıştır. Toplam verinin %67’si eğitim, kalan %33 ise test verisi olacak şekilde bölünmüştür.

4.1. Makine Öğrenmesi Performansını Değerlendirme

Karmaşıklık matrisi, ikili sınıflandırma problemlerinde sınıflandırıcının performansını değerlendirme amacıyla kullanılan 2×2 ’lik bir matristir. Gerçek değerleri bilinen bir veri seti için

gerçek ve tahmini sınıfların yer aldığı matriste (Tablo 1), Doğru sağlam (TP-True Positive), doğru sınıflandırılan sağlam sınıfa ait veri sayısını; yanlış sağlam (FP-False Positive), gerçek sınıfı ıskarta olup sağlam olarak tahmin edilen veri sayısını; doğru ıskarta (TN-True Negative), doğru sınıflandırılan ıskarta veri sayısını ve yanlış ıskarta (FN-False Negative), gerçek sınıfı sağlam olup ıskarta tahmin edilen veri sayısını göstermektedir (Power, 2011).

Tablo 1. Karmaşıklık matrisi

		Gerçek sınıf	
		Sağlam (1)	Iskarta (0)
Tahmini sınıf	Sağlam (1)	Doğru sağlam (TP)	Yanlış sağlam (FP)
	Iskarta (0)	Yanlış ıskarta (FN)	Doğru ıskarta (TN)

Karmaşıklık matrisindeki doğru ve yanlış tahmin sayıları kullanılarak sınıflandırma performansını değerleyen iki kriter, doğruluk ve F1 puanıdır. Doğruluk, sınıflandırma sonucunda elde edilen doğru tahmin sayısının, yapılan tüm tahmin sayısına oranı olarak ifade edilmektedir (Eş. 15). F1 puanı ise kesinlik ve duyarlılık değerlerinin harmonik ortalaması ile ifade edilen bir değerdir (Eş. 16). Kesinlik, tahmini sınıftaki doğru tahminleme oranını göstermektedir (Eş. 17). Duyarlılık ise gerçek sınıftaki doğru tahminleme oranı olarak ifade edilmektedir (Eş. 18). Bu kriterler aşağıdaki gibi hesaplanmaktadır (Power, 2011):

$$\text{Doğruluk} = \frac{TP + TN}{TP + FP + TN + FN} \quad (15)$$

$$F1 = \frac{2 * \text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (16)$$

$$\text{Kesinlik} = \frac{\text{Sınıfa ait doğru tahmin sayısı}}{\text{Tahmini sınıfa ait veri sayısı}} \quad (17)$$

$$\text{Duyarlılık} = \frac{\text{Sınıfa ait doğru tahmin sayısı}}{\text{Gerçek sınıfa ait veri sayısı}} \quad (18)$$

Üçüncü bölümde tanıtılan denetimli öğrenme algoritmaları, mevcut veri seti kullanılarak eğitilmiş ve sınıflandırma başarısına göre karşılaştırılmıştır. Tablo 2’de her bir algoritma için karmaşıklık matrisleri ile doğruluk ve F1 değerlerine yer verilmiştir. Gerek doğruluk, gerekse sağlam ve ıskarta sınıfları için F1 değerleri karşılaştırıldığında DVM ve 3-EYK algoritmalarının eşit ve en yüksek tahmin başarısını gösterdikleri açıktır. Bu iki algoritmanın doğruluk değeri 0,952, F1 puanı sağlam sınıfı için 0,974 ve ıskarta sınıfı için 0,720’dir. *k*-EYK’de komşuluk sayısı ile tahminleme başarısı düşerek en küçük doğruluk değeri 0,932’e, en küçük F1 puanı ise 0,963’e inmiştir. Karar ağacı ve Naïve Bayes’in sınıflandırma performanslarının ise DVM ve *k*-EYK’nin gerisinde kaldığı görülmektedir. Iskarta ve sağlam parçaların tahmin başarısı ayrı ayrı incelendiğinde sağlam verinin Naïve Bayes dışındaki diğer algoritmaların tamamı tarafından 0,929’un üzerinde kesinlik ve duyarlılık değerlerine sahip olduğu görülmektedir. Iskarta örnekleme, 1,000 ve 0,563 olan en yüksek kesinlik ve duyarlılık değerlerinin DVM ve 3-EYK

algoritmaları ile elde edildiği görülmektedir. Bu sonuç, DVM ve 3-EYK algoritmalarının ıskarta olarak tahmin ettiği parçaların tamamının ıskarta olduğunu, yani sağlam parçaların hiçbirini ıskarta olarak tahmin etmediğini göstermektedir. Fakat aynı zamanda 16 kusurlu parçadan 9 tanesinin ıskarta, 7 tanesinin ise sağlam olarak tahmin edildiği anlaşılmaktadır. EYK algoritmalarının tamamı ıskarta örnekleme 1,00'lık kesinlik değerleri verirken komşu sayısı ile duyarlılık başarısının düştüğü görülmektedir. İskartaları tahmin başarısı en düşük algoritmanın ise Naïve Bayes olduğu görülmektedir. 146 parçanın 16 tanesi ıskarta iken Naïve Bayes'in ıskarta olarak tahmin ettiği parça sayısı 52 olmuştur. Tabloda göze çarpan bir diğer sonuç, sağlam sınıfına ait verinin tahmin başarısının F1 puanı bakımından diğer sınıfa göre çok daha yüksek olduğudur. Bölüm 4'te bahsedildiği üzere, sağlam sınıfındaki veri miktarının ıskarta verisine oranı 10:1'dir. Bu oran, verideki sınıflar arası dengesizliğe işaret etmekte olup, sınıflandırma modelinin eğitim sırasında ıskarta sınıfından daha çok sağlam sınıfına odaklandığını göstermektedir. F1 puanlarında görülen sınıflar arası bu farkın, doğruluk ve F1 kriterleri açısından en iyi performansı gösteren DVM ve 3-EYK algoritmalarında diğer algoritmalara göre daha düşük olduğu görülmektedir.

Tablo 2. Karmaşıklık matrisi, doğruluk ve F1 puanı açısından sınıflandırıcıların karşılaştırılması

Alg.	Karmaşıklık matrisi			Sağlam örneklem (1)			İskarta örneklem (0)			Doğruluk
	1	0		Kesinlik	Duyarlılık	F1	Kesinlik	Duyarlılık	F1	
Karar Ağacı	1	121	9	0,931	0,931	0,931	0,438	0,438	0,438	0,877
	0	9	7							
Naïve Bayes	1	85	9	0,904	0,654	0,759	0,135	0,438	0,206	0,630
	0	45	7							
DVM	1	130	7	0,949	1,000	0,974	1,000	0,563	0,720	0,952
	0	0	9							
3-EYK	1	130	7	0,949	1,000	0,974	1,000	0,563	0,720	0,952
	0	0	9							
5-EYK	1	130	9	0,935	1,000	0,967	1,000	0,438	0,609	0,938
	0	0	7							
7-EYK	1	130	10	0,929	1,000	0,963	1,000	0,375	0,545	0,932
	0	0	6							

Küçük veya orta boyutlu veri ile çalışıldığında, performans değerlendirme için başvurulan yöntemlerden biri k -katlamalı çapraz doğrulamadır (Garcia ve diğ., 2019). k -katlamalı çapraz doğrulamada, veri k kadar küçük parçaya bölünmektedir. Bu parçalardan $k-1$ tanesi eğitim verisi, kalan bir kat ise test verisi olarak kullanılmaktadır. Bu işlem, her bir parça test verisi olarak kullanılacak şekilde k kez tekrarlanmaktadır. Bu çalışmada 5-katlamalı çapraz doğrulama

uygulanmıştır (Tablo 3). Çapraz doğrulama ile Tablo 2'deki doğruluk değerleri karşılaştırıldığında, algoritmaların sınıflandırma başarıları çok değişmeyip, DVM ve karar ağacının performanslarında düşük bir artış görülmüştür. Çapraz doğrulama sonuçlarına göre, DVM için elde edilen ortalama başarı puanı 3-EYK için elde edilen ortalama başarı puanından daha yüksektir. Bu nedenle, parametre ayarlamada kullanılmak üzere DVM sınıflandırıcısı seçilmiştir.

Tablo 3. 5-katlamalı çapraz doğrulama sonuçları

Alg.	Kat 1	Kat 2	Kat 3	Kat 4	Kat 5	Ort. Başarı Puanı
Karar Ağacı	0,917	0,917	0,932	0,925	0,955	0,929
Naïve Bayes	0,571	0,654	0,872	0,684	0,526	0,662
DVM	0,962	0,955	0,977	0,962	0,947	0,961
3-EYK	0,955	0,962	0,955	0,947	0,947	0,953
5-EYK	0,940	0,947	0,940	0,932	0,940	0,940
7-EYK	0,932	0,947	0,947	0,940	0,932	0,940

4.2. Parametre Değerleri için Öneri Sistemi

Üretim süreçlerinde kalite tahminlemesi, kalitesizlikle sonuçlanacak bir sürecin erken sonlandırılmasını, dolayısıyla zaman ve maliyet tasarrufu sağlamaktadır. Kalitesizlikle sonuçlanacak bir işlemle ilgili parametrelerin tekrardan ayarlanarak kaliteli ürün elde edilmesi ise zaman ve maliyet tasarrufunun ötesinde üretkenliği arttıracaktır. Bu çalışma, operatöre taş değişiminden sonra belirlediği parametre değerleri ile üretilecek ürünün kalitesini tahminleme imkanı sağlamaktadır. Kalite tahminlemesi, tahmini sınıfı ıskarta olan bir parametre seti için operatörün ayarlarını gözden geçirmesini ve revize etmesini sağlayacaktır. Tablo 2'de gösterilen sonuçlara göre, en iyi performansı gösteren DVM algoritması, test verisindeki sağlam sınıfa ait 130 veriden tamamını, ıskarta sınıfına ait 16 veriden ise 9'unu doğru sınıflandırmıştır. Bu sonuç, kalitesiz ürünü %56,3 oranında tahmin etme ve parametreleri kaliteli ürün ile sonuçlanacak şekilde değiştirme fırsatı anlamına gelmektedir. Bu aşamada, parametre değerlerinin seçimi operatörün deneyimine bırakıldığında, operatör sağlam çıktı elde edinceye kadar deneme-yanılma şeklinde ilerleyeceği için zaman alıcı ve maliyetli olacaktır. Dolayısıyla, operatörün deneyim durumu parametre seçim sürecinde önemli bir role sahip olmaktadır. Bu süreci operatör deneyiminden bağımsız hale getirmek için geçmiş veriden yola çıkarak sağlam çıktı verme potansiyeli yüksek parametre değerlerinin, operatöre bir sistem tarafından önerilmesi amaçlanmıştır. Operatöre önerilecek parametre değerlerinin belirlenmesi için Tablo 4'teki adımlar izlenmektedir. İlk olarak, sağlam sınıfına ait verideki her parametre için ortalama ve standart sapma hesaplanır. Daha sonra sağlam verideki parametre değerlerine standart sapma kadar ekleme ve çıkartma uygulanarak yeni parametre değerleri elde edilir. Yeni parametre değerlerinin çıktuları tahmin edilir ve sağlam sonucunu verenler belirlenir. Bu işlem, artırma ve azaltma sonucunda parametre değerleri tolerans aralığında kaldığı sürece devam eder. Her adımda standart sapmanın katsayısı artırılarak parametre değerlerinin en küçük ve en büyük değerleri arasında yayılımı sağlanır. Tolerans değerlerinin dışına çıkıldığında işlem sonlandırılır ve en çok sağlam tahmini veren belli sayıda veri seçilir.

Tablo 4. Parametre değerleri önerme aşamaları

Adım 1.	N : Veri büyüklüğü, S : Sağlam sınıfa ait veri, $n=1$.
Adım 2.	S 'deki her parametre için ortalama ve standart sapma (σ) hesapla.
Adım 3.	S 'deki her parametreyi $n\sigma$ kadar arttır ve azalt.
Adım 4.	Üretilen değerler tolerans aralığında ise
Adım 5.	Ürettiğin yeni verinin çıktısını DVM ile tahmin et.
Adım 6.	Tahmini sınıfı "sağlam" olanları belirle.
Adım 7.	Her parametre için standart sapmanın katsayısını arttır, $n=n+1$.
Adım 8.	Adım 3'e git.
Adım 9.	Aksi takdirde dur.
Adım 10.	En çok "sağlam" sonucu veren belli sayıda veriyi seç.

Eldeki veriyi analiz ederek operatöre uygun parametre değerlerinin önerilmesi için Şekil 5’de görüldüğü gibi örnek bir ara yüz hazırlanmıştır. Bu çalışma, manuel olarak tutulmuş 441 adet veriyi dikkate alarak gerçekleştirilmesine rağmen, hazırlanan ara yüz sayesinde güncel verilerin dahil edilerek parametre önerisinin revize edilmesini kolaylaştıracaktır. Ayrıca, belirli aralıklarla değiştirilmesi gereken taşların kalan ömrü ile ilgili kontroller ve parametre değerleri ile ilgili grafikler, operatörün süreci takibini kolaylaştırmak amacıyla ara yüze eklenmiştir. Bu çalışma için veri toplanan taşlama makinesinde operatör ayar işlemleri için günlük ortalama 8,3 dk harcamaktadır. Bu süre taşlama hattındaki 14 makine için ortalama 116,2 dk anlamına gelmektedir. Bu süre, operatörün günlük mesaisinin yaklaşık iki saatini ayar işlemleri ile geçirdiğini göstermektedir. Firmanın, önerilen çözüm yaklaşımını mevcut sistemlerine entegre ederek kullanması durumunda taşlama makinesinde ortaya çıkan ıskartalar yaklaşık %56 oranında tahmin edilebilecektir. Ayrıca, önerilen parametre değerleri ile söz konusu %56’lık üretim için parametre ayarları operatör deneyiminden bağımsız olarak değiştirilebilecektir.

Toplu Parça Kontrolü	Tek Parça Kontrolü	Taş Ömrü Kontrolü	Grafikler
P1		P11	
P2		P12	
P3		P13	
P4		P14	
P5		P15	
P6		P16	
P7		P17	
P8		P18	
P9		P19	
P10			

Parametreleri Kontrol Et Sınıflandır

Şekil 5:

Operatöre uygun parametre değerlerinin önerilmesi için hazırlanan ara yüz

5. SONUÇ

Günümüz dijitalleşme sürecinin en önemli amaçlarından biri üretim/hizmet alanlarındaki faaliyetleri otonom hale dönüştürmektir. İnsan faktörünün hatalarından ve kararlarından bağımsız olarak çalışan süreçlerin düşük maliyet ve yüksek kalite ile sonuçlanması bu ihtiyacı doğurmaktadır. Bu çalışmada, insan faktörüne bağlı ıskartaların dijital çözümler ile azaltılması amaçlanmıştır. Bu amaçla, bir firmanın taşıma süreci incelenmiş ve gerçek veri kullanılarak makine öğrenmesi ile kalite tahminlemesi yapılmıştır. Farklı makine öğrenmesi sınıflandırma algoritmaları modellenmiş ve DVM algoritmasının en yüksek performansa sahip olduğu belirlenmiştir. Ayrıca kaliteli ürün ile sonuçlanması beklenen uygun parametre değerleri belirlenerek, bu değerlerin kullanıcılara sunulabileceği bir öneri sistemi geliştirilmiştir.

Yoğun bir veri akışının olduğu gerçek bir üretim sürecinde, toplanan verinin makine öğrenmesi algoritmalarıyla işlenmesi ve faydaya dönüştürülmesi üç önemli aşamadan oluşmaktadır. Bunlardan ilki verinin toplanmasıdır. Çalışma kapsamında, firmadaki mevcut veri kullanılmadığı için manuel tutulmuştur. Bu durum, sınırlı bir sürede sınırlı sayıda veri toplanmasına neden olmuş, bu da sağlam ve ıskarta sınıflarına ait verideki dengesiz dağılımı beraberinde getirmiştir. Bu etkiyi ortadan kaldırmak amacıyla dijitalleşmeyi başarmak isteyen firmalarda ilk aşama olan veri toplama adımının, sürece entegre edilecek sensörler yardımıyla dijitalleşmesi bir ihtiyaçtır. İkinci aşama, toplanan verideki karmaşık ilişkilerin modellenerek ürün kalitesinin tahmin edilmesi; üçüncü aşama ise kalite tahmininin imalat sürecine güncel parametre değerleri ile geri bildirimidir. Gerçekleştirdiğimiz çalışma, dijital dönüşüm sürecindeki ikinci ve üçüncü aşamaya hizmet etmekte olup, ilk aşamanın dijitalleşmesi ile birlikte, güncel verinin analizi ve güncel parametre değerlerinin belirlenmesini sağlayacak yazılımı oluşturmaktadır. Gerçek veri kullanılarak elde edilen sonuçlar, önerilen yaklaşımın ıskartaları %56 oranında azaltılabileceğini göstermektedir.

Bu çalışmanın devamı olarak, dijitalleşme sürecinin gereği olan yatırımların yapılmasıyla elde edilecek yeni veri setleriyle daha fazla testler yapılabilir. Bununla birlikte son aşamada oluşturulan öneri sistemi için meta-sezgisel yaklaşımlar kullanılabilir.

TEŞEKKÜR

Bu çalışmanın gerçekleştirilmesi için gösterdikleri ilgi ve destekleri için tüm firma çalışanlarına sonsuz teşekkürlerimizi sunarız.

ÇIKAR ÇATIŞMASI

Yazarlar, bilinen herhangi bir çıkar çatışması veya herhangi bir kurum/kuruluş ya da kişi ile ortak çıkar bulunmadığını onaylamaktadırlar.

YAZAR KATKISI

Emine Eş Yürek, çalışmanın kavramsal ve tasarım süreçlerinin belirlenmesi ve yönetimi, uygun yöntemlerle analizi ve sonuçların yorumlanması aşamalarında; Betül Yağmahan, çalışmanın kavramsal ve tasarım süreçlerinin belirlenmesi ve yönetimi, uygun yöntemlerle analizi ve sonuçların yorumlanması aşamalarında; Burak Celal Akyüz, çalışmanın kavramsal ve tasarım süreçlerinin belirlenmesi, verinin toplanması, dönüştürülmesi, uygun yöntemlerle analizi ve sonuçların yorumlanması aşamalarında; Ebubekir Sıddık Samast, çalışmanın kavramsal ve tasarım süreçlerinin belirlenmesi, verinin toplanması, dönüştürülmesi, uygun yöntemlerle analizi ve sonuçların yorumlanması aşamalarında; Nezire Dilan Çetrez, çalışmanın kavramsal ve tasarım

süreçlerinin belirlenmesi, verinin toplanması, dönüştürülmesi, uygun yöntemlerle analizi ve sonuçların yorumlanması aşamalarında katkı sağlamışlardır.

KAYNAKLAR

1. Adesanya A., Abdulkareem A. ve Adesina L.M. (2020) Predicting extrusion process parameters in Nigeria cable manufacturing industry using artificial neural network, *Heliyon*, 6(7).
2. Arif F., Suryana N. ve Hussin B. (2013) A data mining approach for developing quality prediction model in multi-stage manufacturing, *International Journal of Computer Applications*, 69(22), 35-40.
3. Bai Y., Sun Z., Deng, L., Li L., Long J. ve Li C. (2018) Manufacturing quality prediction using intelligent learning approaches: A comparative study, *Sustainability*, 10(1), 85.
4. Chou P.H., Wu M.J. ve Chen K.K. (2010) Integrating support vector machine and genetic algorithm to implement dynamic wafer quality prediction system, *Expert Systems with Application*, 37(6), 4413-4424.
5. Ciurana J., Arias G. ve Ozel T. (2009) Neural network modeling and particle swarm optimization (PSO) of process parameters in pulsed laser micromachining of hardened AISI H13 steel, *Materials and Manufacturing Processes*, 24, 358-368.
6. Cunningham P. ve Delany S.J. (2020) *k*-Nearest Neighbour Classifiers: 2nd Edition (with Python examples). arXiv preprint arXiv:2004.04523.
7. Doğan A. ve Birant D. (2021) Machine learning and data mining in manufacturing, *Expert Systems with Applications*, 166, 114060.
8. Feng W., Sun J., Zhang L., Cao C. ve Yang Q. (2016) A support vector machine based naive Bayes algorithm for spam filtering, *IEEE 35th International Performance Computing and Communications Conference (IPCCC)*, 1-8.
9. Gao H., Zhang Y., Fu Y., Mao T., Zhou H ve Li D. (2018) Process parameters optimization using a novel classification model for plastic injection molding, *Intelligent Journal of Advanced Manufacturing Technology*, 94, 357-370.
10. Garcia V., Sanchez J.S., Rodriguez-Picon L.A., Mendez-Gonzalez L.C. ve Ochoa-Dominguez H.J., Using regression models for predicting the product quality in a tubing extrusion process, *Journal of Intelligent Manufacturing*, 30, 2535-2544, 2019.
11. Han J., Kamber M. ve Pei J. (2012) *Data Mining: Concepts and Techniques*, Morgan Kaufmann, Waltham, MA.
12. Jurkovic Z., Cukor G., Brezocnik M. ve Brajkovic T. (2018) A comparison of machine learning methods for cutting parameters prediction in high speed turning process, *Journal of Intelligent Manufacturing*, 29, 1683-1693.
13. Köksal G., Batmaz İ. ve Testik M.C. (2011) A review of data mining applications for quality improvement in manufacturing industry, *Expert Systems with Applications*, 38(10), 13448-13467.
14. Kumar D.P., Amgoth T. ve Annavarapu C.S.R. (2019) Machine learning algorithms for wireless sensor networks: A survey, *Information Fusion*, 49, 1-25.
15. Lahouar A. ve Slama J.B.H. (2015) Random forests model for one day ahead load forecasting, *IREC2015 The Sixth International Renewable Energy Congress*, IEEE, 1-6, March.
16. Larose D.T. (2015) *Data Mining and Predictive Analytics*, John Wiley & Sons.

17. Okaro I.A., Jayasinghe S., Sutcliffe C., Black K., Paoletti P. ve Green P.L. (2019) Automatic fault detection for laser powder-bed fusion using semi-supervised machine learning, *Additive Manufacturing*, 27, 42-53.
18. Mohammadi P. ve Wang Z. J. (2016) Machine learning for quality prediction in resistant material manufacturing process, *IEEE Canadian Conference on Electrical and Computer Engineering*, Vancouver-Canada, 15-18 May.
19. Paturi U.M.R. ve Cheruku S. (2020) *Application and performance of machine learning techniques in manufacturing sector from the past two decades: A review*, *Materials Today: Proceedings*, (in press).
20. Peres R.S., Barata J., Leitao P. ve Garcia G. (2019) Multistage quality control using machine learning in the automotive industry, *IEEE Access*, 7, 79908-79916.
21. Power D.M.W. (2011) Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation, *Journal of Machine Learning Technologies*, 2(1), 37-63.
22. Ramulu V., Ramana E.V. ve Kumar N.K. (2019) Predictive modelling for quality prediction and assurance of extrusion blow molding, *International Journal of Innovative Technology and Exploring Engineering*, 8(11), 1364-1368.
23. Sapounas I., Vosniakos G.C. ve Papazetis G. (2020) A simulation-based robust methodology for operator guidance on injection moulding machine settings, *International Journal on Interactive Design and Manufacturing*, 14, 519-533.
24. Schorr S., Möller M., Heib J., Fang S. ve Bahre D. (2020) Quality prediction of reamed bores based on process data and machine learning algorithm: A contribution to a more sustainable manufacturing, *Procedia Manufacturing*, 43, 519-526.
25. Silva J.A, Abellan-Nebot J.V., Siller H.R. ve Guedea-Elizalde F. (2014) Adaptive control optimisation system for minimising production cost in hard milling operations, *International Journal of Computer Integrated Manufacturing*, 27(4), 348-360.
26. Siryani J., Tanju B. ve Eveleigh T.J. (2017) A machine learning decision-support system improves the internet of things' smart meter operations, *IEEE Internet of Things Journal*, 4(4), 1056-1066.
27. Strasser S., Tripathi S. ve Kerschbaumer R. (2018) An approach for adaptive parameter setting in manufacturing processes, *7th International Conference on Data Science, Technology and Application*, Porto-Portugal, 24-32, 26-28 July.
28. Tharwat A. (2019) Parameter investigation of support vector machine classifier with kernel functions, *Knowledge and Information Systems*, 61(3), 1269-1302.
29. Teixidor D., Grzenda M., Bustillo A. ve Ciurana J. (2015) Modeling pulsed laser micromachining of micro geometries using machine-learning techniques, *Journal of Intelligent Manufacturing*, 26, 801-814.
30. Weichert D., Link P., Stoll A., Ihlenfeldt S. ve Wrobel S. (2019) A review of machine learning for the optimization of production processes, *The International Journal of Advanced Manufacturing Technology*, 104, 1889-1902.
31. Widodo A. ve Yang B.S. (2007) Support vector machine in machine condition monitoring and fault diagnosis, *Mechanical Systems and Signal Processing*, 21(6), 2560-2574.
32. Zhang X.D. (2020) *Machine learning. In: A Matrix Algebra Approach to Artificial Intelligence*, Springer, Singapore, 223-440.

