

İşten Ayrılma Durumunun Tahminlenmesinde Makine Öğrenmesi Algoritmaları Kullanımı – Demir Çelik Sektöründe Bir Uygulama

Nurselin SÜLLÜ¹, Atıl KURT^{2*}

¹Alanya Alaaddin Keykubat Üniversitesi Mühendislik Fakültesi Endüstri Mühendisliği Anabilim Dalı, Antalya

²Alanya Alaaddin Keykubat Üniversitesi Mühendislik Fakültesi Endüstri Mühendisliği Bölümü, Antalya

¹<https://orcid.org/0000-0002-8902-3959>

²<https://orcid.org/0000-0002-5438-3743>

*Sorumlu yazar: atil.kurt@alanya.edu.tr

Araştırma Makalesi

Makale Tarihiçesi:

Geliş tarihi: 07.11.2024

Kabul tarihi: 23.03.2025

Online Yayınlanma: 16.09.2025

Anahtar Kelimeler:

Makine öğrenmesi
Rastgele orman algoritması
Çalışan tutundurma
İşten ayrılma tahminlenmesi

ÖZ

Zorlu çalışma koşullarına sahip demir-çelik sektöründeki birçok firma, mevcut personeli elde tutmanın maliyetinin yeni bir çalışan işe almaktan daha düşük olması nedeniyle çalışan devir oranını düşürmeye odaklanmaktadır. Bu çalışmanın amacı, makine öğrenmesi algoritmalarının kullanılmasıyla işten ayrılma riski olan çalışanları tahmin etmek, işe alımlarda hangi kriterlere öncelik verilmesi gerektiğini belirleme ve çalışan bağlılığını artırmaya yönelik stratejik kararlar geliştirme süreçlerinde işletmelere rehberlik etmektir. Çalışma kapsamında, Knime ve Python programları kullanılarak bir demir-çelik firmasının 16 yıllık sürede toplam 2318 çalışan için 14 farklı niteliği içeren veriler analiz edilmiştir. Çalışmada beş farklı makine öğrenmesi algoritması uygulanmış ve en yüksek doğruluk oranına (%78,40) Rastgele Orman Algoritması ile ulaşılmıştır. Rastgele Orman Algoritması yardımıyla, işten ayrılmalara yol açabilecek en önemli faktörler belirlenmiştir.

Using Machine Learning Algorithms to Predict Turnover - An Application in the Iron and Steel Industry

Research Article

Article History:

Received: 07.11.2024

Accepted: 23.03.2025

Published online: 16.09.2025

Keywords:

Machine learning
Random forest algorithm
Employee retention
Turnover prediction

ABSTRACT

Since the cost of retaining the existing staff is lower than the cost of hiring a new staff for many companies, including companies in the Iron and Steel Sector, which have challenging working conditions, it aims to have a low labor turnover rate. The aim of this study is to guide businesses in the processes of predicting employees at risk of turnover by using machine learning algorithms, determining which criteria should be prioritized in recruitment, and developing strategic decisions to increase employee loyalty. The machine learning algorithms in the study were implemented through Knime and Python programming languages. In the application, the data, which took place in an Iron and Steel company in a 16-year period, consists of 14 different attributes that may be effective on the resignation of 2318 personnel. Six different machine learning algorithms were applied and the most successful of them was the Random Forest Algorithm with an accuracy rate of approximately 78.40%. Therefore, the most important criteria leading to turnover were determined using the Random Forest Algorithm. As a result of this study, it will be possible to make decisions such as whether the existing staff has the potential to leave the job, which criteria should be prioritized in recruitment and the necessary actions to be taken to ensure the continuity of the staff.

To Cite: Süllü N., Kurt A. İşten Ayrılma Durumunun Tahminlenmesinde Makine Öğrenmesi Algoritmaları Kullanımı – Demir Çelik Sektöründe Bir Uygulama. Osmaniye Korkut Ata Üniversitesi Fen Bilimleri Enstitüsü Dergisi 2025; 8(4): 1525-1543.

1. Giriş

İşletmelerdeki insan kaynakları yönetiminin, çalışanların ihtiyaçlarına karşılık verme ve ücret, izin, işe alım gibi süreçlerin yanı sıra merkezinde insan faktörü bulunduğundan, insana değer verme, çalışanları motive etme ve çalışanı tutundurma faaliyetleri de bulunmaktadır. İnsan kaynakları yönetimi makro ölçekte, çalışanların bilgi birikimine sahip olması ve kurumların pazarda rekabet edebilmeleri için oluşturdukları planda insanı baş unsur olarak gösterilmesi gerektiğine dayanmaktadır. Mikro ölçekte ise kurum içi organizasyonda uyum ve yeteneklerin yönetilmesi ile açıklanmaktadır (Karakütük, 1997).

Demir çelik sektörü inşaat, otomotiv, beyaz eşya, gemi yapımı, makine imalatı gibi pek çok endüstriye demir-çelik sağlayan ve onların üretimi, işlenmesi ve satışıyla ilgilenen sanayi koludur. Demir çelik sektöründe yer alan işletmelerin nitelikli personel ihtiyacına sahip olmasından kaynaklı işe alım süreçleri uzun soluklu olmakta ve yetenekli bilgi birikimine sahip olan mevcuttaki personellerin kaybedilmemesi konusu önem taşımaktadır. Personel kaybı yaşanması sahadaki işlerin aksamasına, çalışanların iş yükünün artmasına ve mesaili çalışma düzenine geçilmesine neden olmaktadır. Tüm bu sorunlar ile karşılaşmamak adına ve iş gören devir oranının düşük olması hedeflendiğinden mevcut personellerin tutundurulması sektör için olumlu bir olgudur. İstifa süreçleri firmayı zaman, maliyet ve operasyonel açıdan olumsuz etkileyeceğinden istifaların doğru şekilde öngörülerek ilgili tedbirlerin alınması ve çalışan memnuniyetinin sağlanması önem arz etmektedir.

Demir çelik sektöründeki karşılaşılan bu zorluklar insan kaynakları yönetiminin daha etkin ve veriye dayalı karar mekanizmalarının geliştirmesi ihtiyacını doğurmaktadır. Bu kapsamda, çalışanların işten ayrılma durumlarının tahmin edilmesi, iş gücü planlaması ve personel tutundurmayı daha etkili hale getirebilir. Bu noktada, makine öğrenmesi teknolojilerinin devreye girmesi, veri tabanlı analizler yaparak istifa eğilimlerini ve bu eğilimlerin nedenlerini belirleme imkânı sunabilmektedir. Bu tür teknoloji kullanımları mevcut personelin elde tutulmasının yanı sıra işe alım süreçlerinde daha stratejik kararlar alınmasına katkı sağlar. İnsan kaynakları alanında makine öğrenmesi uygulamaları zaman içerisinde kuruluşlar tarafından kullanılmaya başlanmıştır ancak yapay zekanın insan kaynaklarındaki uygulamadaki zorluk ve karmaşıklık nedeniyle daha fazla araştırmaya ihtiyaç duyulmaktadır (Garg ve ark., 2022; Kişi, 2023).

Araştırmalar incelendiğinde makine öğrenmesinin yapay zekanın bir alt dalı olarak tanımlandığı görülmektedir (Alkan, 2024). Makine öğrenmesi, verilerin belirli bir probleme uygun olarak modellenmesine yarayan algoritmalarıdır. Makine öğrenmesi, bilgisayar sistemleri açıkça programlanmadan veri ve deneyimlerden öğrenilmesini sağlar, büyük miktarda veriyle eğitilir ve belirli bir görevi yerine getirmek için verileri kullanarak ilişkiler çıkarır. Makine öğrenmesinin probleme özgü olması daha etkili sonuçlar vermesini sağladığı için çeşitli metotlar geliştirilmiştir. Bu metotlara denetimli, denetimsiz, pekiştirmeli ve derin öğrenme örnek olarak verilebilir (Murphy, 2012).

İstifa süreçleri firmayı zaman, maliyet ve operasyonel açıdan olumsuz etkileyeceğinden istifaların doğru şekilde öngörülerek ilgili tedbirlerin alınması ve çalışan memnuniyetinin sağlanması önem arz etmektedir. Bu kapsamda makine öğrenmesi uygulamaları kullanılarak insan kaynakları faaliyetlerinden

olan çalışan tutundurma üzerinde değerlendirmeler yapılmış olup, problem için şu katkılar elde edilmiştir: (1) İstifa durumunun tahminlenmesi, (2) istifaya neden olacak etkenlerin sınıflandırılması, (3) olası istifaları önlemek için yapılması gereken önlemler.

Yapay zekanın insan kaynaklarında kullanımı bu çalışmada teorik ve uygulamalı olarak gerçekleştirilmiştir. Teorik ve hazır veri seti kullanılan araştırmalardan farklı olarak gerçek veriler üzerinde uygulanması çalışmayı özgün kılmaktadır (Çetin, 2024, Widiyawati, 2023). Uygulama kapsamında firmadan alınan gerçek veriler üzerinde çalışılmış ve elde edilen ham veri seti, eğitim ve test seti olarak ayrılmıştır. Personellerin işten ayrılma üzerinde etkisi bulunabileceği tahminlenen nitelikler ele alınmış ve bu nitelikler girdi özelliği taşımıştır. Tahminlemedeki çıktı ise işten ayrıldı ya da ayrılmadı şeklinde değerlendirilmiştir. Dolayısıyla, çalışmada ele alınan problemin çıktıya sahip olması nedeniyle denetimli makine öğrenmesi olarak sınıflandırılmıştır. Kullanımı, karşılaştırma yapılarak değerlendirilmesi ve başarı oranları göz önüne alınarak Python ve paket program olan Knime (Konstanz Information Miner) kullanılmıştır. Her iki platformda da Rastgele Orman, Naive Bayes, Karar ağacı, Destek Vektör Makineleri ve k-en Yakın Komşu Algoritmaları kullanılarak tahminlemeler gerçekleştirilmiş ve karşılaştırmalar yapılarak en iyi tahminlemede bulunan algoritma tespiti gerçekleştirilmiştir.

Tüm bu faktörler göz önüne alındığında araştırma, ülkemizdeki insan kaynakları alanında makine öğrenmesi algoritmalarının uygulanması konusunda hem gerçek veri kullanımı hem de bu sektörde daha önce benzer bir çalışmanın yapılmamış olması sebebiyle özgün değere sahiptir. Aynı zamanda çalışmanın detaylı incelemesi yapıldığı takdirde uygulama adımları ile gelecekteki projeler için de yol gösterici niteliği taşıyabilecektir. Çalışmanın devamında konu ile ilgili yapılan çalışmaların olduğu literatür incelemesi kısmı bulunmaktadır. Demir çelik sektöründeki problem ve firmadan elde edilen veri setinin oluşturulması ve hazırlanması çalışmanın 3. kısmında değinilecek olup, 4. kısımda çözüm için kullanılan yöntemler ve elde edilen sonuçlar tartışılacaktır. Son olarak, 5. bölümde elde edilen sonuçlar değerlendirilerek öneriler sunulacak ve sonuç kısmı ile sonlandırılacaktır.

2. Literatür Taraması

Bu bölümde literatürde insan kaynakları yönetimi alanında yapılan yapay zeka uygulamaları incelenecektir. Bu alanda yapılan çalışmaları inceleyen birçok inceleme çalışması bulunmaktadır (Gelinas ve ark., 2022; Palos-Sanchez ve ark., 2022; Bujold ve ark., 2024). Bu çalışmalara ek olarak literatürde yapay zeka uygulamalarının insan kaynakları yönetimindeki birçok probleme uygulanabilirliği (Strohmeier ve Piazza, 2015), insan kaynaklarındaki yeri (Khatri ve ark., 2020) ve insan kaynakları üzerinde evrimsel süreci (Jatoba ve ark., 2019) gibi yapay zekanın belirli bir özel alandaki çalışmalar da mevcuttur. İncelemiş olduğumuz çalışmalarda daha çok çalışan kaybı tahminleme problemleri üzerinde çalışıldığı gözlemlenmiştir.

Potansiyel istifaların makine öğrenmesi algoritmaları ile tahminlenmesi problemi veri havuzu Kaggle üzerinden alınan veriler ve Knime programı kullanılarak çalışılmıştır (Widiyawati, 2023). Çalışmada

Naive Bayes algoritması kullanılmış ve %81,16 oranında doğruluk başarısı elde edilmiştir. Fallucchi ve ark. (2020) çalışmasında çalışan işten ayrılma süreci tahminleme problemi aynı veri seti sağlayıcısını kullanarak 6 farklı algoritma denenmiş ve performans değerlendirmelerine göre en başarılı sonuç Gaussian Naive Bayes ile elde edilmiştir. Arromrit ve ark. (2023) ise aynı konu ve veri seti platformunu kullanarak birçok algoritma uygulamış ve en iyi sonuç Rastgele Orman Algoritması ile elde edilmiştir. Punnoose ve Ajit (2016) çalışanların devir oranı üzerinde bir çalışma yapmış ve birçok algoritma tarafından sınanmış olup en iyi sonucun Aşırı Gradyan Güçlendirme yöntemi tarafından bulunduğu sonucuna varılmıştır. Ayrıca çalışmalarında potansiyel istifalarda ne yapılabilir sorusu üzerine de derin öğrenme algoritmaları uygulanabilirliği önerilmiştir. Atef ve ark. (2022) çalışmalarında IBM tarafından yaratılan çalışan devir verilerini kullanarak k-en yakın komşu arama ve rastgele orman algoritmalarını kullanarak tahminleme yapmış ve en iyi sonucu k-en yakın komşu ile elde etmişlerdir. Diğer bir çalışan devir tahmini çalışmasında ise tekrarlayan sinir ağları kullanılarak e-posta iletişim analizi yapılarak işten ayrılma durumunun tahminlenmesi yapılmış olup %74 oranında başarı elde edilmiştir (Korytkowski ve ark., 2023).

Endonezya'nın telekomünikasyon şirketindeki çalışan kaybı problemi için iki yıllık gerçek verileri kullanarak 3 farklı makine öğrenmesi algoritması kullanılmış ve en iyi sonuç %97.5'lik doğruluk ile Rastgele Orman Algoritması olmuştur (Alamsyah and Salma, 2018). Diğer bir telekomünikasyon şirketinin çalışan kaybı tahminlemesi problemi baz alınarak üretilen veriler ile yapılan çalışmada ise 8 farklı algoritma Python ile uygulanmış olup Rastgele Orman Algoritması (%92.2) en iyi sonucu vermiştir (Uzak, 2022). Mhatre ve ark. (2020) yine çalışan kaybı tahminlemesi çalışması yapmış olup 6 farklı algoritma denemiş ve en iyi çözüm Aşırı Gradyan Artırma (%96) algoritması ile elde etmişlerdir. Başka bir çalışan kaybı tahmini problemi için 5 farklı makine öğrenmesi algoritması kullanılmış ve %86 başarı oranı ile Lojistik Regresyon Algoritması en iyi performansı vermiştir (Mohbey, 2021). Ayrıca çalışmalarında bir ara yüz tasarlanarak erken uyarı sistemi oluşturulmuştur. Al-Darraji ve ark. (2021) Derin Sinir Ağları kullanılarak başka bir çalışan kaybı tahmini problemi çalışılmış ve makine öğrenmesi teknikleri ile insan kaynakları planlamasının verimliliğinin artırılacağı vurgulanmıştır. Guerranti ve Dimitri (2023) çalışan kaybı tahmini ile çeşitli nitelikler arasında korelasyon analizi yapmış ve birçok algoritmayı sınayarak en iyi sonucu %88'lik doğruluk performans ile Lojistik Regresyon Algoritması ile elde etmiştir.

Çalışan kaybı tahmini çalışmalarının yanı sıra insan kaynakları yönetimi için farklı yapay zeka çalışmaları da bulunmaktadır. Aday seçimi ve adayların elde tutulma problemi 12 farklı nitelik kullanılarak ele alınmış ve 8 farklı algoritma kullanılarak Marvin ve ark. (2021) tarafından çözülmüştür. En iyi çözüm Rastgele Orman Algoritması (%92) ile elde edilmiştir. Başka bir aday seçim problemi ise Tian ve ark. (2023) tarafından çalışılmış olup 5 farklı algoritma sınanmıştır ve en iyi çözüm Destek Vektör Makineleri algoritması tarafından elde edilmiştir. Zhu (2021) çalışmasında insan kaynaklarında gradyan iniş algoritması kullanan bir ücret tahmin modeli tasarlamıştır. Nawaz ve ark. (2024) çalışmalarında yapay zeka adaptasyonun insan kaynaklar yönetimi üzerine etkisini farklı çıktılara

odaklanarak incelenmiştir. Çalışma sonucunda doğruluk, bilgi işlem gücü ve kapasitesi, kişiselleştirmenin zaman tasarrufu ve maliyet azaltmayı önemli ölçüde etkilediğini, otomasyon ve gerçek zamanlı deneyimin ise etkilemediğini göstermektedir.

Ülkemizde makine öğrenmesi ile Telekomünikasyon sektöründeki çalışan kaybı tahmini üzerine bir çalışma yapılmıştır (Uzak, 2022). Yaptığımız çalışma ile bu çalışma arasında hem uygulama aşamasında farklılıklar vardır hem de uygulandığı sektörler birbirinden farklıdır. Buna ek olarak bu çalışmanın aksine çalışmamızda gerçek veriler kullanılmıştır. Bu çalışmanın daha önce yapılan araştırmalardan farkı, demir çelik sektöründe gerçekleştirilmesi, gerçek verilerin kullanılması, birden fazla makine öğrenmesi platformunda ve farklı makine öğrenmesi algoritmalarının uygulanarak karşılaştırılmasıdır.

3. Veri Setlerinin Oluşturulması ve Hazırlanması

Yapılan araştırmada yapay zeka alt dallarından olan makine öğrenmesi algoritmalarından yararlanarak, demir-çelik sektöründeki bulunan bir firmanın geçmişte istifa etmiş çalışanların özellikleri incelenerek, mevcuttaki veya istihdam edilmesi planlanan adayların gelecekte istifa etme olasılığı tahminlenecektir. Çalışanların görevinden istifa etme kararını almasında etkili olabilecek iç ve dış olmak üzere pek çok farklı etmen bulunmaktadır. Öncelikle personellerin işten istifa etmesinde önemli rol oynayabilecek 14 adet nitelik belirlenmiştir. Değerlendirmeye alınan nitelikler ile işten istifa edip etmediği kontrol edilerek gelecekteki istifa durumunun tahmin edilmesi amaçlanmıştır. Nitelikler, kişisel deneyim ve insan kaynakları fonksiyonlarında deneyimli karar vericilerin ortak görüşleri, literatür araştırmalarında sırasında incelenen çalışmalar, analiz ve yöntemin somut ve ölçülebilir veriler gerektirdiği dikkate alınarak seçilmiştir. İstifa üzerinde etkili olabilecek demografik özellikler ve gizlilik kuralları çerçevesinde değerlendirilebilecek diğer kriterler de göz önünde bulundurulmuştur. Ancak, istifaya etki edebilecek ast-üst ilişkileri, kurumsal aidiyet, iletişim, mobbing gibi kişisel veriler, genellikle soyut, ölçülmesi zor, eksik, yanlış veya yanlış olma riskine sahip olduğu için değerlendirmeye dahil edilmemiştir. Ayrıca, anket ve envanter sonuçlarının gizlilik açısından uygun olmaması, bu tür verilerin analiz dışında bırakılmasının bir diğer önemli sebebidir. Veriler işlenirken veri setinde nümerik olmayan nitelikler sayısal değişkenlere dönüştürülmüştür. Sayısal değerlere dönüştürülme adımı niteliğe bağlı olarak verilerin birbirilerine göre üstünlük ilişkisi olup olmadığı göz önüne alınmıştır. Bahsedilen 14 nitelik, istifa edip edilmemesi ile ilgili veriler, açıklamalar ve verilerin sayısal işleme şekilleri aşağıda listelenmiştir.

- Pozisyon: Çalışanın unvanını temsil etmekte olup veri setinde, Formen, Genel müdür, İşçi, Kıdemli mühendis, Kıdemli uzman, Müdür, Müdür yardımcısı, Mühendis, Şef, Uzman, Uzman Yardımcısı ve Vardiya amiri olmak üzere on iki farklı pozisyon türü bulunmaktadır. Unvanların birbirine karşı hiyerarşik açıdan üstünlüğü bulunmaktadır. Firmanın organizasyon şeması göz önüne alındığında bazı unvanların birbirine karşı üstünlüğü bulunmadığı bilinmektedir, dolayısıyla bu pozisyona aynı puan verilmiştir. Numaralandırma yapılırken, işçi 1, formen 2, uzman yardımcısı 3, mühendis ve uzman 4, kıdemli mühendis ve kıdemli uzman 5, vardiya

amiri ve şef 6, müdür yardımcısı 7, müdür 8 ve son olarak genel müdür ise 9 şeklinde sayısallaştırılmıştır.

- **Kıdem:** Kıdem çalışanın işyerinde fiilen çalıştığı süreyi belirtmektedir. İşten ayrılan personeller için işten ayrılış tarihinden işe giriş tarihinin çıkarılması ile elde edilmiştir. İşe devam eden personeller için ise kıdem, temsili olarak cari yılın son günü 31.12.2023 tarihinden işe giriş tarihinin çıkarılması ile elde edilmiştir.
- **Yaş:** Çalışanların yaşını göstermekte olup işten ayrılan personeller için işten ayrılış tarihinden doğum tarihinin çıkarılması ile elde edilmiştir. İşe devam eden personeller için ise yaş, temsili olarak 31.12.2023 tarihinden doğum tarihinin çıkarılması ile elde edilmiştir.
- **Doğum Yeri:** Çalışanların doğum yerini belirtmektedir ve veri setinde 68 farklı şehir bulunmaktadır. Doğum yeri niteliğinde 68 benzersiz değer bulunmaktadır ve her bir şehir birbirinden farksızdır yani herhangi üstünlük bulunmamaktadır.
- **Medeni Durum:** Evli ya da bekar olmak üzere çalışanların medeni durumunu belirtmektedir. Bu veride “Evli” olan veriler 1, “Bekar” olan veriler 0 olarak gösterilmiştir.
- **Çocuk Sahibi:** Çalışanların çocuk sahibi olup olmadığını göstermektedir. Çocuk sahibi olan çalışanla “Evet (1)”, olmayanlar ise “Hayır (0)” olarak gösterilmiştir.
- **Engellilik:** Çalışanların engelli olup olmadığını ve engelli olması durumunda kaçınıcı dereceden engelli olduğunu belirtmektedir. Engelli olmayan çalışanlar “Yok” olarak girilmiştir, engelli olan çalışanların ise birinci derece, ikinci derece, üçüncü derece engelli şeklinde açıklanmıştır. Engellilik kriterinde personelin engellilik durumundan etkilenme durumuna göre sıralama yapılmıştır, dolayısıyla bu kriterde üstünlük mevcuttur. Engellilik durumu bulunmayan personel 0, üçüncü derece yakını engelli olan personel 1, birinci derece yakını engelli olan 2, personelin kendisi üçüncü derece engelli olması 3 ve personelin kendisinin ikinci derece engelli olması ise 4 olmak üzere nümerik hale getirilmiştir.
- **Emeklilik:** Çalışanların emeklilik durumunu göstermektedir. Emekli olan personeller “Emekli”, emekli olmayan personeller ise “Normal” şeklinde ifade edilmiştir. Bu veride “Normal” olan veriler 1, “Emekli” olan veriler 0 olarak gösterilmiştir.
- **Tehlike Durumu:** Demir çelik sektörü iş sağlığı ve güvenliği bakımından yüksek risk barındırdığından tehlike durumu çalışma için önemlidir. İşin yükü ve riski değerlendirilerek yüksek risk taşıyan pozisyonlar “Çok Tehlikeli”, diğer pozisyonlar ise “Tehlikeli” olarak değerlendirilmiştir. Burada “Çok tehlikeli” olan veriler 1, “Tehlikeli” olan veriler ise 0 olarak gösterilmiştir. Bazı beyaz yaka çalışanların da sahada doğrudan tehlikeli alanlarda görev almasından dolayı bu sınıfta değerlendirilmiştir.
- **Eğitim Düzeyi:** Çalışanların eğitim seviyesini belirtmektedir. Personellerin eğitim düzeyleri; ilköğretim, lise (düz lise, açık öğretim lisesi, Anadolu lisesi vb.), meslek lisesi, ön lisans, lisans ve yüksek lisans olmak üzere veri setinde altı farklı öğrenim düzeyi bulunmaktadır. Eğitim

düzeyi niteliğinde ise verilerin birbirine üstünlüğü bulunmaktadır. Bu nedenle, eğitim seviyesi düşükten yükseğe doğru puanı artacak şekilde sıralanmıştır. Eğitim düzeyi, ilköğretim 1, meslek lisesi ve lise denk olduğundan eşit puan 2, ön lisans 3, lisans 4 ve yüksek lisans 5 şeklinde nümerik hale getirilmiştir.

- **Maaş Düzeyi:** Çalışanların almış olduğu ücret seviyesini göstermektedir. Çok düşük, düşük, orta, yüksek ve çok yüksek olmak üzere beşli ölçek tercih edilmiştir. Düzeyler, işten ayrılan personeller için çalışanların işten ayrıldığı tarihteki maaşının yine aynı dönemdeki asgari ücrete oranlanması ile belirlenmiştir. İşten ayrılmayan personeller için ise mevcut ücretlerinin 2023 güncel asgari ücrete oranlanması ile elde edilmiştir. Oranlar küçükten büyüğe doğru sıralanarak çok düşükten çok yükseğe doğru sıralanmıştır. Beyaz yaka personellerin maaşları aynı şekilde alınmış, mavi yaka personellerin ise ücretleri saatlik şeklinde ödendiğinden, mavi yaka çalışanların ücretli otuz günlük bir ay şeklinde 225 saat ile saatlik ücretinin çarpılması ile aylık ücrete çevrilmiştir. Şirket politikası gereği ücret gizliliği saklı tutulduğundan gerçek veriler kullanılmış ancak oran kullanılması tercih edilmiştir. Maaş düzeyi ise veri seti oluşturulması esnasında beşli ölçeklendirilmiştir. Buna göre maaş düzeyi için sayısallaştırma işlemi çok düşükten çok yükseğe artacak şekilde numaralandırılmıştır. Çok düşük 1, düşük 2, orta 3, yüksek 4 ve çok yüksek 5 sayısallaştırılmıştır.
- **Yaka Türü:** Şirket bünyesinde çalışanların pozisyonlara göre beyaz ya da mavi yaka olmak üzere iki farklı şekilde sınıflandırılmıştır. Beyaz yaka personeller, uzman, mühendis, şef, formen, müdür vb. mavi yaka personeller ise işçiler şeklinde örneklendirilebilir. Beyaz yaka çalışanlar aylık ücretli, mavi yaka personeller ise saatlik ücretli şeklinde çalışmaktadır. Bu veride “Beyaz” olan veriler 1, “Mavi” olan veriler 0 olarak gösterilmiştir.
- **Avans:** Çalışanların şirketten nakit avans talep edip etmediğini ifade etmektedir. Sağlık sorunu vb. nedenler ile çalışanların talep etmiş olduğu avansın var olup olmadığı belirtilmiştir. Hiç avans talep etmeyen personeller 0 şeklinde gösterilmiştir. 0 ile 9 arasında avans talebi bulunmaktadır.
- **İcra:** Çalışanın icra dava dosyasının bulunup bulunmadığını göstermektedir. İcrası bulunmayan personeller “Yok”, olanlar ise “Var” şeklinde ifade edilmiştir. İcra “Yok” ise 1, “Var” ise 0 olarak gösterilmiştir.
- **İstifa Mı:** Çalışanın işten ayrılışının istifa olup olmadığını belirtmektedir. İşine devam eden ve zorunlu nedenlerle (askerlik, emeklilik, evlilik vb.) işten ayrılan personeller “Hayır”, belirsiz süreli iş sözleşmesinin işçi tarafından feshi ve deneme süresinde belirsiz süreli iş sözleşmesinin işçi tarafından feshi şeklinde çıkış kodu bulunan personeller ise “Evet” şeklinde gösterilmiştir. İstifa mı sütunu veri setinin çıktısı şeklinde belirtilmektedir, veri setinin etiketi niteliğini taşımaktadır. Bu veride istifa eden durumu “Evet” olan veriler 1, istifa etmeyenlerin durumu ise “Hayır” olan veriler 0 olarak gösterilmiştir.

Veriler firmanın SAP (System Analysis Program) sisteminden alınmış olup 2318 adet satırdan oluşmaktadır. Her satır için yukarıda bahsedilen 14 nitelik ve bir çıktı bulunmaktadır, dolayısıyla toplamda 34.770 adet veri bulunmaktadır.

4. Yöntemler ve Elde Edilen Sonuçlar

Problem için çözüm aracı olarak veri setinde “istifa” çıktısı tahminleneceğinden denetimli makine öğrenmesi algoritmalarından literatürde sıklıkla tercih edilen, küçük-orta ölçekteki ve ham veri setinde kullanımı kolay, hızlı ve yüksek doğruluk oranını hedefleyen beş tanesinin kullanımı önerilmiştir: Rastgele Orman, Naive Bayes, Karar Ağacı, Destek Vektör Makineleri ve *k*-en Yakın Komşu algoritmaları. Makine öğrenmesi algoritmalarının performans göstergelerini iki farklı platformda karşılaştırarak analiz etmek, farklı alanlarda avantajlar sunması ve algoritmaların tutarlılığının test edilmesi nedenleriyle hem Knime paket programında hem de Python programlama dilinde uygulanmıştır. Bu sayede başarı kriterlerine göre doğruluk oranları farklı algoritmalar ve programlar üzerinden karşılaştırma yapılarak değerlendirildiğinden literatürdeki diğer araştırmalara katkı sağlanması amaçlanmıştır. Knime, makine öğrenmesi, büyük veri işleme, veri bilimi ve analizi için kullanılabilen açık kaynak kodlu, sürükle – bırak yöntemi ile görselleştirilebilir ve konfigüre edilebilir nodlar ile kullanım kolaylığı sunabilmektedir. Kullanıcı dostu bir arayüz sunmasından dolayı ileri seviye bir kodlama bilgisi gerektirmemektedir. Çalışmada, verilerin %80’i eğitim, %20’si ise test için kullanılmıştır. Çalışmada kullanılan veri setine ait alternatifin kriterlerde aldığı örnek değerler Tablo 1’de gösterilmiştir. Tabloda veri gizliliği nedeniyle veriler temsili olarak gösterilmiş ancak araştırmada gerçek veriler ele alınmıştır. Çalışmanın bu kısmında bahsedilen algoritmalar kısaca açıklanacaktır.

Tablo 1. Çalışmada kullanılan veri seti örneği

Kriterler\Sıra No	1	2	3	4	5
Pozisyon	İşçi	Formen	İşçi	Vardiya Amiri	Mühendis
Kıdem	4,0	5,70	3,7	14,46	4,12
Yaş	27,9	36,78	28,4	49,11	32,55
Doğum Yeri	Osmaniye	Adana	Osmaniye	Hatay	Kahramanmaraş
Medeni Durum	Evli	Evli	Bekar	Evli	Evli
Çocuk Sahibi	Evet	Hayır	Hayır	Evet	Evet
Engellilik	Yok	Yok	3. Derece Engelli	Yok	Yok
Emeklilik	Normal	Normal	Normal	Emekli	Normal
Tehlike Durumu	Tehlikeli	Tehlikeli	Çok Tehlikeli	Tehlikeli	Tehlikeli
Eğitim Düzeyi	İlköğretim	Meslek Lisesi	İlköğretim	Ön lisans	Lisans
Maaş Düzeyi	Çok Düşük	Orta	Düşük	Yüksek	Çok Yüksek
Yaka Türü	Mavi Yaka	Beyaz Yaka	Mavi Yaka	Beyaz Yaka	Beyaz Yaka
Avans	0	0	0	1	0
İcra	Var	Yok	Yok	Yok	Yok
İşten Ayrıldı	İşten Ayrıldı	Etkin	İşten Ayrıldı	İşten Ayrıldı	İşten Ayrıldı
İstifa Mı	Evet	Hayır	Evet	Hayır	Hayır

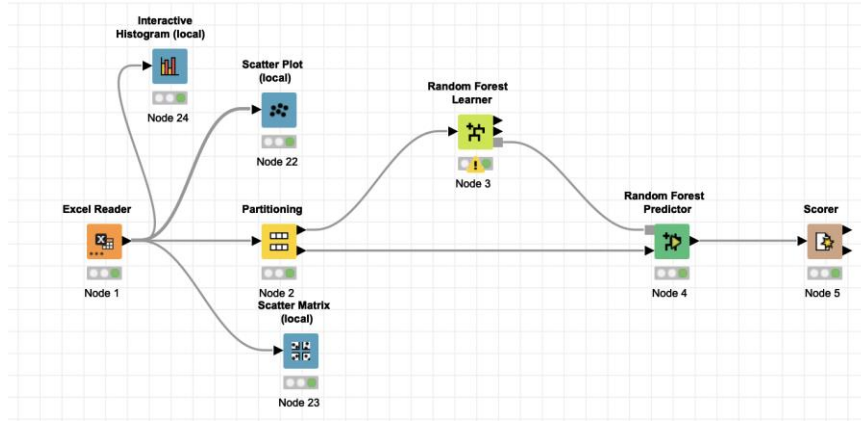
4.1. Rastgele Orman

Rastgele Orman Algoritması’nda veriler öğrenme seti için rastgele seçilir, dallanmalar ile farklı alt karar ağaçları elde edilir. Rastgele örneklenmiş veri alt kümesi ve özellikler kullanılarak her karar ağacı

eğitilir. Rastgele Orman, elde edilen karar ağaçlarını birleştirerek çalışır. Bu süreçte her ağaç bağımsız olarak eğitilir ve ağaçların tahminleri birleştirilir. Test sürecinde de elde edilen Rastgele Orman yapısına göre sınıflandırılır.

Literatürdeki araştırmalarda yüksek başarı tahmin oranı, yüksek sayıda veri ile çalışabilme, eksik veri bulunması halinde başarılı tahmin ile tamamlama gibi avantajlardan kaynaklı olarak sıklıkla tercih edilmektedir. Tüm bu avantajların yanında, aşırı öğrenme durumu oluşabilme ihtimali de dezavantaj olarak değerlendirilebilir. Bu konuda detaylı anlatım Biau and Scornet (2016) tarafından yapılmıştır.

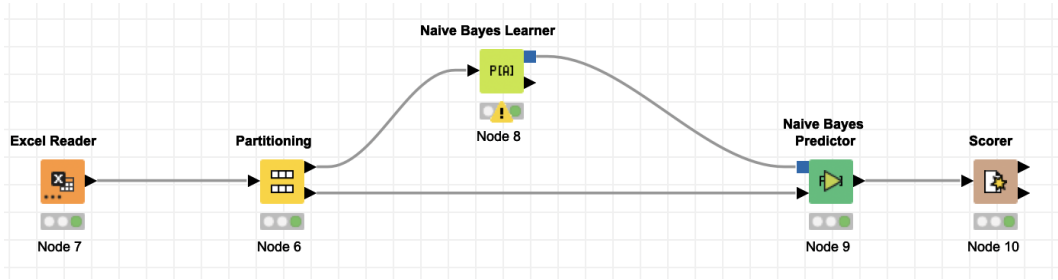
Algoritmanın Knime programındaki görsel kodu Şekil 1’de verilmiş olup algoritmaya verilerin bulunduğu Microsoft Excel dosyası Excel okuyucusu ile Knime programında içeri aktarılmasıyla başlanmaktadır. Excel okuyucusuna bağlantılar sayesinde yeni modüller eklenerek histogram grafiği, serpilme diyagramı, serpilme diyagram matrisi ya da pasta grafiği gibi görselleştirmeler yapılabilmektedir. İkinci aşamada sınıflandırıcı ile veri seti eğitim seti ve test seti olmak üzere ikiye ayrılmaktadır. Sonrasında “Rastgele Orman Eğitici” modülü ile veri eğitilip, “Rastgele Orman Tahminleyici” ile test verileri tahmini yapılır. Son olarak da “Skor” modülü ile sınıflandırma ya da tahmin sonuçlarını değerlendirip, modelin doğruluğunu ölçülür.



Şekil 1. Rastgele Orman Algoritması’nın Knime üzerinden uygulama görseli

4.2. Naive Bayes

Naive Bayes algoritması, kategorik değişkenli veri modellerinde başarılı sonuçlar vermektedir. Karmaşık problemlerde, sınıflandırma işlemleri gerçekleştirme olanağı sağlayabilmektedir. Naive Bayes algoritması Bayes teoremine dayanmaktadır ve her bir niteliğin birbirinden bağımsız olduğunu varsaymaktadır. Öncelikle eğitim verisi kullanılarak her sınıf için öncül olasılıklar ve her bir özelliğin sınıfa bağlı olasılıkları hesaplanır. Sonrasında, yeni veriler için, Bayes teoremi kullanılarak her sınıfın ardıl olasılıkları hesaplanır ve en yüksek olasılığa sahip sınıf seçilerek sınıflandırma yapılır. Naive Bayes algoritması ile ilgili detaylar ve uygulamalar için literatürde çalışmalar mevcuttur (Reddy ve ark., 2022). Bu algoritmanın görseli ise Şekil 2’de verilmiş olup, yapısı Rastgele Orman Algoritması ile aynıdır. Naive Bayes algoritmasının modülleri olan “Naive Bayes Eğitici” ve “Naive Bayes Tahminleyici” Kullanılmıştır.

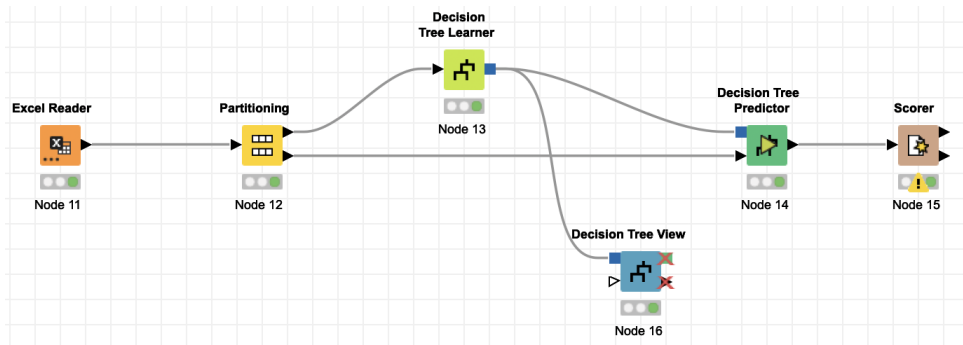


Şekil 2. Naive Bayes’in Knime üzerinden uygulama görseli

4.3. Karar Ağacı

Karar Ağacı, bir tür sınıflandırma algoritmasıdır ve sıralı, basit mantıksal testleri birleştiren bir model olarak çalışır. İlk adımda, hangi niteliğin ve hangi değerin veriyi en iyi şekilde ayırdığı tespit edilir. Her iç düğümde, algoritma en iyi bölünmeyi bulur ve veri bu bölünmeye göre alt gruplara ayrılır. Daha sonra, her alt grup için yeni düğümler oluşturulur ve bu işlem tekrarlanarak yaprak düğümlere ulaşılır. Yeni bir veri, bu bölünmeleri takip ederek bir yaprak düğüme ulaşır ve yaprak düğümdeki sınıfın değerini alır. Algoritmanın detaylarıyla ilgilenen okuyucular, Patel ve Prajapati'nin (2018) çalışmasına başvurabilir.

Karar ağacı algoritmanın KNIME programı üzerinden gösterimi Şekil 3’te gösterilmiştir. Bu algorithmada da Karar Ağacı’na ait olan eğitici ve tahmin edici modüller kullanılmıştır. Bunlara ek olarak karar ağacını görüntüleyebilmek için de “Karar Ağacı Görünümü” adlı modül kullanılmıştır.



Şekil 3. Karar Ağacı Algoritması’nın Knime uygulaması

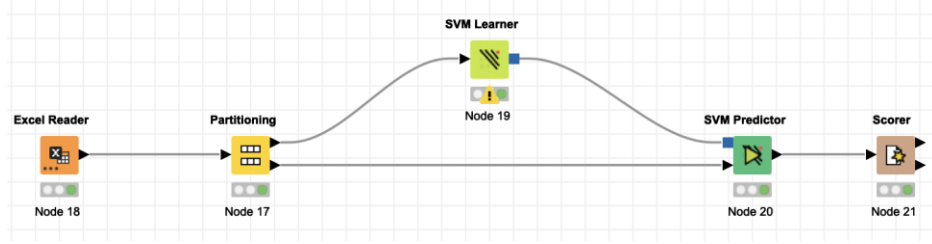
4.4. Destek Vektör Makineleri

Destek Vektör Makineleri, farklı sınıfları ayıran en uygun hiper-düzlemi bulmayı amaçlayan bir makine öğrenmesi algoritmasıdır. Algoritmanın temel prensibi, sınıflar arasındaki farkı maksimize etmek için hiper-düzlemin konumunu ayarlamaktır. Hiper-düzlem, iki sınıfı ayırt edebilmek için veriler arasında bir sınır oluşturur. İki boyutlu bir düzlemde bu sınır bir çizgi olarak tanımlanırken, daha yüksek boyutlu veri uzaylarında bir düzlem olarak ifade edilir. Algoritma, sınıflar arasındaki mesafeyi (marjı) en geniş yapacak şekilde ayarlanarak, yeni verilerin sınıflandırılmasını sağlar.

Destek Vektör Makinelerinin özellikle doğrusal olmayan sınıflandırma problemlerinde yüksek doğruluk oranına ulaşmasının arkasındaki temel neden, kernel fonksiyonları kullanarak verilerin daha yüksek boyutlara taşınabilmesidir. Bu sayede, doğrusal olarak ayrılamayan veriler, daha karmaşık sınırlarla ayrılabilir hale gelir. Bu özellikleri sayesinde, çeşitli sınıflandırma problemlerinde tercih edilen bir

yöntem olmuştur. Algoritmanın detaylarına ilgi duyan okuyucular, bu konuda yapılan kapsamlı incelemelere başvurabilirler (Cervantes ve ark., 2020).

Destek Vektör Makineleri algoritmasının Knime üzerindeki uygulaması da benzer olup Şekil 4’te görselleştirilmiştir. Burada eğitici ve tahmin ediciler için sırasıyla “SVM Eğitici” ve SVM Tahminleyici” modülleri kullanılmıştır.



Şekil 4. Destek Vektör Makineleri Algoritması’nın Knime programı üzerinden görseli

4.5. k-En Yakın Komşu

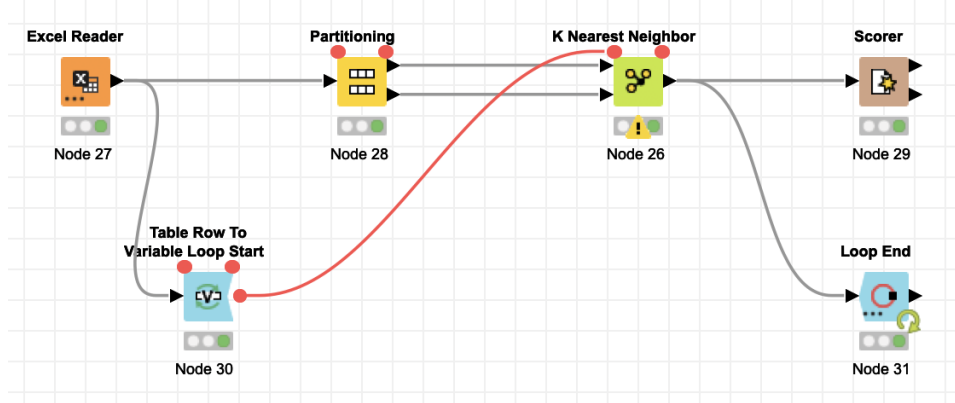
Bir makine öğrenmesi algoritması olan k-en Yakın Komşu Algoritmasında amaç, örnek uzayda bir noktaya mesafesi en az olan k tane noktanın belirlenmesidir. Veri ve k tane nokta arasında uzaklık incelenerek, yakınlıklara göre çoğunluk kararı dikkate alınarak sınıflandırma yapılır. Kullanılan algoritmanın aşamaları aşağıda listelenmiştir.

- k tane merkez noktası tespit edilir.
- k tane noktanın merkez noktasına olan uzaklıkları Öklid mesafesine göre hesaplanır ve kümeleme işlemi gerçekleştirilir.
- Elde edilen kümeler arasında karşılaştırmalar yapılır ve küme elemanları içerisinde yer değişiminin gerekliliği olup olmadığı kontrol edilir.
- Değişim gerekliliği devam ettiği sürece algoritma adımları sürdürülür, olmadığında ise durdurulur.
- Test edilmesi istenen yeni verinin noktası belirlenerek, öğrenme veri kümelerine en yakın olduğu küme tespit edilir.

Problemin amacına göre sınıflandırma ya da regresyon tahminlerinde uygulanabilir olduğundan kullanımı yaygındır. Oldukça basit bir mantığa dayanması, veri kümesiyle ilgili herhangi bir dağılım varsayımı içermemesi ve küçük boyutlu verilerde iyi çalışması nedeniyle yaygın olarak kullanılmaktadır. k-En Yakın Komşu Algoritması ile ilgili detayları Halder ve ark.’nın (2024) ilgili inceleme çalışmasında bulabilirsiniz.

Bu algoritmanın Knime Programı üzerinden gösterimi ise Şekil 5’te gösterilmiştir. k- En yakın komşu algoritmasında tahmin edici ve test için “K En Yakın Komşu” modülü “Tablo satırından değişkenlere döngü başlangıcı” modülü ile bir döngü başlatılmış ve bu döngü satırların okunarak değişkene çevrilmesini sağlanmıştır ve döngü “Döngü sonu” modülü ile tamamlanmıştır. Bu algorithma kullanılan komşu sayısının performansı etkilemesinden dolayı çeşitli komşu sayıları denenmiş olup en

iyi performansı sağlayan komşu sayısının 33 olduğu gözlemlenmiştir. Dolayısıyla, karşılaştırma için de bu komşu sayısı kullanılmıştır.



Şekil 5. Knime üzerinden k-En Yakın Komşu Algoritması'nın gösterimi

5. Değerlendirme ve Bulgular

Kullanılan makine öğrenmesi algoritmalarının uygunluğu birbirleri ile başarılı tahminleme performans ölçütlerine göre kıyaslanır. Her bir algoritma, doğru sınıflandırılan veri sayısı, yanlış sınıflandırılan veri sayısı, doğruluk oranı ve Cohen'in Kappa katsayısı olmak üzere dört farklı performans ölçütlerine göre değerlendirilmiştir. Başarı performans ölçütlerinden doğru sınıflandırılan veri sayısının yüksek olması, yanlış sınıflandırılan veri sayısının düşük olması, doğruluk oranının yüksek olması, Cohen'in Kappa katsayısının ise +1 ya da -1'e yakın olması beklenmektedir. Sınıflandırma işleminin doğruluğunun kararı Şekil 6'da özetlenen karışıklık matrisinde gösterilmiştir.

		Tahmin edilen sınıf		
		Pozitif	Negatif	
Gerçek Sınıf	Pozitif	Doğru Pozitif (DP)	Yanlış Negatif (YN)	Duyarlılık
	Negatif	Yanlış Pozitif (YP)	Doğru Negatif (DN)	Özgüllük
		Hassaslık	Negatif Tahmin Değeri	Doğruluk

Şekil 6. Karışıklık matrisi

Karşılaştırmalarda kullanılan doğru sayısı DP ve DN'lerin toplamı ile, yanlış sayısı YP ve YN'lerin toplamı ile bulunur. Doğruluk oranı ise toplam doğru sayısının toplam gözlem sayısına oranlanması ile bulunmaktadır. Buna ek olarak algoritmaların performansını daha kapsamlı bir şekilde değerlendirebilmek için doğruluk oranını destekleyen Cohen Kappa katsayısı da dahil edilmiştir. Cohen Kappa katsayısı, iki değerlendirici arasında gözlenen uyumu rastgele uyumdan ayırt ederek ölçen ve sınıflandırma problemlerinde sınıflandırıcı performansını değerlendirmek için kullanılan bir istatistiksel ölçüttür (Cohen, 1960). Eşitlik 1'de verilen Cohen Kappa katsayısı, $-\infty$ ile 1 arasında değer alır ve model performansının iyi olarak değerlendirilmesi için κ değerinin 1'e yakın olması beklenir. Bu, modelin

yaptığı tahminlerin rastgele yapılan tahminlere kıyasla üstün olduğunu gösterir. Performans karşılaştırmalarında, κ değerleri yüksekten düşüğe doğru sıralanarak değerlendirilir.

$$\kappa = \frac{p_0 - p_c}{1 - p_c} \quad (1)$$

Yukarıdaki formülde kullanılan p_0 değeri gerçek anlaşma oranını, p_c ise rastgele anlaşma beklenen oranı göstermektedir ve hesaplama şekilleri Denklem 2 ve 3’de verilmiştir.

$$p_0 = \frac{DP+DN}{DP+DN+YP+YN} \quad (2)$$

$$p_c = \left(\frac{DP+YN}{DP+DN+YP+YN} \times \frac{DP+YP}{DP+DN+YP+YN} \right) + \left(\frac{DN+YN}{DP+DN+YP+YN} \times \frac{DN+YP}{DP+DN+YP+YN} \right) \quad (3)$$

Şekil 6’da gösterilen gösterilen karışıklık matrisine göre algoritmaların elde ettiği sonuçlardan bir tanesi Tablo 2’de gösterilmiştir. Tablo 2’ye göre $DP+DN$ sayısı 362 ile en iyi sonuç Rastgele Orman Algoritması tarafından elde edilmiştir.

Tablo 2. Algoritmalar tarafından yapılan tahmin performansları

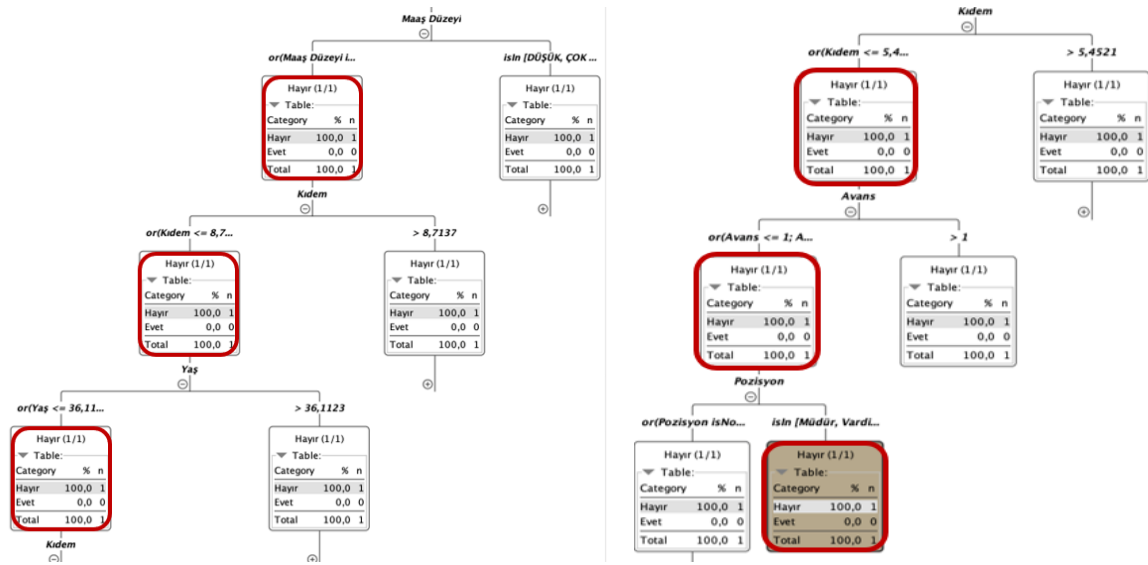
		İstifa Durumu İle İlgili Tahmin									
		Rastgele Orman		Naive Bayes		Karar Ağacı		Destek Vektör Makineleri		K-en Yakın Komşu	
		Hayır	Evet	Hayır	Evet	Hayır	Evet	Hayır	Evet	Hayır	Evet
İstifa Durumu	Hayır	289	40	247	80	272	56	331	0	294	44
	Evet	60	75	63	74	62	74	133	0	83	43

Knime ve Python programlama dillerinde uygulanmış olan algoritmaların eğitim ve test setleri rastgele olarak belirlendiğinden karşılaştırmaların anlamlı olabilmesi için iki programda da beş kere deneme gerçekleştirilmiştir. Sonuç doğrultusunda alınacak olan karar ya da yapılan tahminlerin güvenilirliğini sağlamak adına her denemeye ait sonuçlar kaydedilmiş ve ilgili değerlerin ortalaması alınmıştır. Tablo 3’teki sonuçlara baktığımızda da her iki program tarafından en iyi çözümün Rastgele Orman Algoritması ile elde edildiğini gözlemliyoruz, bununla birlikte en kötü sonuç ise Naive Bayes algoritmasında gerçekleşmiştir.

Tablo 3. Kullanılan programlara göre algoritmaların performansı

Algoritma	Knime Paket Programı				Python			
	Doğru	Yanlış	DO (%)	κ	Doğru	Yanlış	DO (%)	κ
Rastgele Orman	364	100	%78,40	0,45	357	107	%76,88	0,40
Naive Bayes	320	144	%69,05	0,28	310	154	%66,82	0,13
Karar Ağacı	345	119	%74,47	0,37	333	131	%71,79	0,30
Destek Vektör Makineleri	331	133	%71,42	0,00	330	134	%71,14	0,00
K-En Yakın Komşu	337	127	%72,71	0,24	344	120	%74,07	0,30

En başarılı tahmini gerçekleştirdiğinden çalışanların istifa edip etmemesi ile ilgili öngörülerimizi Rastgele Orman Algoritmasının sonuçları ile yapılacak olup önerdiği model kullanılmıştır. Knime üzerinden Rastgele Orman algoritmasının bir karar ağacına ait örnek görüntüsü Şekil 7’de gösterilmiştir. Rastgele Orman algoritması, ayrıştırma işlemini yaparken her bir düğümde kullanılan özellikleri etkinlik için gösterge olabilecek Gini indeksi, bilgi kazancı gibi ölçütlere göre seçer ve birden fazla karar ağacından oluşur. Şekilde verilen karar ağacına göre, ilk ayrıştırma maaş düzeyi üzerinden yapılmış olup, bunu kıdem, yaş, avans ve pozisyon özellikleri takip etmektedir.

**Şekil 7.** Rastgele Orman Algoritması karar ağacı

Ayrıştırmalar sırasıyla maaş düzeyi, kıdem, yaş olduğundan bu kriterlere ait açıklama ve çalışanların istifasını önleyebilecek tutundurma önerileri Tablo 4’te açıklanmıştır. Öneriler şirket çalışanı olan bir insan kaynakları uzmanının yardımlarıyla yapılmıştır.

Tablo 4. Rastgele Orman Algoritmasından elde edilen ayrıştırma kriterleri için açıklama ve tutundurma önerileri

Ayrıştırma Kriteri	Sorunlar	Çözüm Önerileri
Maaş Düzeyi	Ücret artışı önerisinin sunulması, ücret skalasının kişiye özgü olamaması sebebiyle adaletsizlik yarattığından, istifa düşüncesi olmayanların ise ücret artışı amacı ile gerçekçi olmayan istifa başvurusu yapmasına neden olmaktadır.	Bayram harçlığı, prim, yakacak, kırtasiye yardımları, özel/tamamlayıcı sağlık sigortası, doğum, evlilik, vefat yardımı gibi yan haklar arttırılacaktır.
		Ödül yönetimi ve takdir tanıma sistemi oluşturulacaktır.
		Öneri sistemi ile kazanım sağlayan fikirler toplanarak yıl sonundaki değerlendirmelerde en iyi öneri veren çalışanlara ödül verilecektir.
Kıdem	Kıdemin artması işe hakimiyet ve bilgi birikimini beraberinde getirdiğinden yüksek kıdeme sahip çalışanların işten ayrılması yetenek ve zaman kaybı oluşturmaktadır.	İstifa etmek için başvuran çalışana, potansiyel kariyer yol haritası, eğitim ve gelişim planları konularında yönlendirme yapılması ile yetkinlik iyileştirmeleri sağlanacaktır.
		Şirketteki maaş, kıdem vb. niteliklerin sürekliliğinin sağlanmasının öneminden bahsedilerek istifanın önüne geçmeye çalışılacaktır.
Yaş	Yaş ile kıdemin orantılı olduğu göz önüne alındığında genç çalışanların yoğunlukta olması deneyim, kıdem, bağlılık ve motivasyonu olumsuz etkilemektedir.	Gönüllü faaliyetlerden olan, sportif aktiviteler, yürüyüşler, piknik organizasyonu, boya, çizim, seramik, tatlı atölyesi vb. etkinlikler planlanacaktır. Takım çalışmasını destekleyecek aktivitelerin planlanması, ekip içi iletişimi ve sosyalleşmeyi arttıracaktır.
		İş koşullarının mümkün olduğu pozisyonlarda esnek veya hibrit çalışma ile iş-yaşam arasındaki dengeyi sağlayarak kuruma bağlılığı artırması beklenebilmektedir. Aynı zamanda genç çalışanlar teknolojiye hızlı adapte olabildiğinden üretim-yönetimde verimli çalışmayı ve tatmin olmayı sağlayacaktır.
		En başarılı toplam kalite yönetimi uygulaması yapan çalışanlar “Aynı Çalışan” olarak seçilerek motivasyon sağlanır ve çalışanlar arası pozitif rekabet ortamı ile teşvik edilmiş olunacaktır.
Avans Talep Sayısı	Çalışanın sağlık sorunları için maaşlarından eşit taksitlerle kesilmek üzere almış olduğu taksitli avans talepleri çalışanların finansal dengesini bozarak ödeme gücüne neden olmaktadır.	Sağlık sorunlarının iş kazası ya da meslek hastalığı ile ilgili olması durumunda ödemelerin işyeri tarafından karşılanması önerilebilir.
		Esnek geri ödeme seçenekleri ya da rapor vb. durumların beyanı ile taksit öteleme imkânı sunulabilir.
		Finansal planlama konusunda çalışanın bilgilendirilmesi ve borç sürecinden olumsuz etkilenmemesi adına finansal ve psikolojik açıdan çalışan destek programı oluşturulabilir.
Pozisyon	İletişim eksikliği, iş-yetkinlik uyumsuzluğu, deneyim eksikliği nedenlerinden pozisyona bağlı istifalara neden olmaktadır.	Adil görev dağılımı için iş dengeleme gerçekleştirilmelidir. Norm kadro açığı nedeniyle oluşan fazla iş yükü ise norm kadro analizi ile tespit edilecektir.
		Çok tehlikeli sınıfında yer alan görevlerde, ilgili pozisyon için çalışma ortamında iyileştirme ya da görev değişikliği imkânı sunulacaktır.
		Kariyer yönetimi doğrultusunda hedefleri ile ilgili eğitimler organize edilerek gelişim imkânı sunulacak bu sayede pozisyona aidiyet duygusu oluşturulabilecektir.

6. Sonuç

Demir çelik sektörü, ağır sanayi koşullarının getirdiği zorluklar ve yüksek risk faktörleri nedeniyle çalışanlar açısından zorlu bir ortam sunmakta ve bu durum sıklıkla işten ayrılmalara yol açmaktadır. Ayrıca, kuşak değişimleri de çalışanların işten ayrılma oranlarını artıran faktörlerden biridir. İşletmeler için en önemli kaynaklardan biri olan insan kaynağının verimli ve etkili bir şekilde yönetilmesi, çalışanların bağlılıklarının artırılması açısından önemlidir. Bu doğrultuda, insan kaynakları yönetiminin fonksiyonları çerçevesinde çalışanları elde tutmaya yönelik faaliyetlerin gerçekleştirilmesi gerekmektedir.

Bu çalışmada, demir-çelik sektöründeki bir işletmeden alınan ham veriler kullanılarak çalışanların işten ayrılma durumları makine öğrenmesi algoritmaları ile tahmin edilmiştir. Knime ve Python platformları üzerinden uygulanan bu algoritmalar arasında, %78,40 başarı oranı ile en yüksek doğruluk Rastgele Orman Algoritması ile elde edilmiştir. Elde edilmiş başarı oranı literatürdeki diğer çalışmalarla karşılaştırıldığında gerçek veri seti kullanmasına rağmen kabul edilebilir bir seviyededir.

Çalışmada, işten ayrılma olasılığını etkileyen 14 farklı nitelik belirlenmiş ve bu nitelikler önem sırasına göre analiz edilmiştir. Knime programında gerçekleştirilen analiz sonucunda, Rastgele Orman Algoritması karar ağacındaki ilk üç kırılıma göre, işten ayrılmada en etkili faktörlerin maaş düzeyi, kıdem ve yaş olduğu görülmüştür. Rastgele Orman algoritması ile işten ayrılma riski taşıyan çalışanlar tespit edilerek bu çalışanların elde tutulmasına yönelik öneriler sunulmuştur.

Araştırma kapsamında, veriler işletmenin kullandığı SAP sistemi üzerinden sağlanmıştır. Ancak geçmiş verilerin eksik ve güncel olmaması nedeniyle bazı nitelikler sayısal olmasına rağmen sınıflandırılarak sözel hale getirilmiştir. Ek olarak, mesai süresi ve işe geç gelme sıklığı gibi büyük boyutlu veriler SAP sisteminden çekilememiştir. Çalışmada kullanılan programlar, ileri düzey bilgisayar becerisi veya yapay zeka bilgisi gerektirdiğinden, insan kaynakları yetkilileri için daha kullanıcı dostu bir yapıya dönüştürülemedi.

Gelecekteki çalışmalarda farklı nitelikler, örneğin çalışanların geçmiş iş deneyimleri, izin, devamsızlık kayıtları, çalışan memnuniyet anketlerin sonuçları ile elde edilebilecek ast-üst ilişkileri, kuruma aidiyet, iş-yetkinlik düzeyi ilişkisi, iletişim gibi değişkenler dikkate alınabilir. Ayrıca, madencilik, inşaat, sağlık ve denizcilik gibi benzer koşullara sahip sektörlerde de benzer araştırmalar yapılabilir. Ek olarak, Lojistik Regresyon, Doğrusal Regresyon ve Doğrusal Diskriminant Analizi gibi farklı makine öğrenmesi algoritmalarının performansları değerlendirilebilir. Gelecek çalışmalarda, nesne yönelimli programlama dillerinde kullanıcı dostu arayüzler tasarlanarak insan kaynakları uzmanlarının tahminleri daha kolay yapabilmesi sağlanabilir. Bu sayede, İK profesyonelleri yapay zeka bilgisi gerekmeksizin ilgili nitelik değerlerini girerek çalışanların işten ayrılma olasılıklarını tahmin edebilir. Ayrıca, Knime üzerinden Karar Ağacı algoritması çıktıları dışa aktararak görselleştirilmiş sonuçlar ile işten ayrılma olasılıkları kolayca analiz edilebilir.

Çıkar Çatışması Beyanı

Makale yazarları herhangi bir çıkar çatışması olmadığını beyan eder.

Araştırmacıların Katkı Oranı Beyan Özeti

Yazarlar makaleye eşit oranında katkı sağlamış olduğunu beyan eder.

Kaynakça

- Ajit P. Prediction of employee turnover in organizations using machine learning algorithms. *International Journal of Advanced Research in Artificial Intelligence* 2016; 5(9): 22-26.
- Alamsyah A., Salma N. A comparative study of employee churn prediction model. In 2018 4th International Conference on Science and Technology (ICST), Ağustos 2018, IEEE: 1-4.
- Al-Darraj S., Honi DG., Fallucchi F., Abdulsada AI., Giuliano R., Abdulmalik HA. Employee attrition prediction using deep neural networks. *Computers* 2021; 10(11): 141.
- Alkan A. Öğrencilerin sınavlardaki performansının makine öğrenmesi teknikleriyle tahminlenmesi. *Osmaniye Korkut Ata Üniversitesi Fen Bilimleri Enstitüsü Dergisi* 2024; 7(3): 1116-1128. <https://doi.org/10.47495/okufbed.1420959>
- Arromrit T., Srisakaew K., Roswhan N., Mahikul W. A supervised machine learning method for predicting the employees turnover decisions. *IEEE 8th International Conference On Software Engineering and Computer Systems (ICSECS)*, Ağustos 2023, IEEE:122-127.
- Atef M., S Elzanfaly D., Ouf S. Early prediction of employee turnover using machine learning algorithms. *International Journal of Electrical and Computer Engineering Systems* 2022; 13(2): 135-144.
- Biau G., Scornet E. A random forest guided tour. *Test* 2016; 25: 197-227.
- Bujold A., Roberge-Maltais I., Parent-Rochelleau X., Boasen J., Sénécal S., Léger PM. Responsible artificial intelligence in human resources management: a review of the empirical literature. *AI and Ethics* 2024; 4: 1185-1200.
- Cervantes J., Garcia-Lamont F., Rodríguez-Mazahua L., Lopez A. A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing* 2020; 408: 189-215.
- Çetin G. Endüstri 4.0 bağlamında yapay zekânın insan kaynakları yönetimine etkileri-Effects of artificial intelligence on human resources management in the context of industry 4.0 (Master's thesis, Sakarya Üniversitesi). 2004.
- Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1960; 20(1): 37-46.
- Fallucchi F., Coladangelo M., Giuliano R., William DLE. Predicting employee attrition using machine learning techniques. *Computers* 2020; 9(4): 86.
- Garg S., Sinha S., Kar AK., Mani M. A review of machine learning applications in human resource management. *International Journal of Productivity and Performance Management* 2022; 71(5): 1590-1610
- Gélinas D., Sadreddin A., Vahidov R. Artificial intelligence in human resources management: A review and research agenda. *Pacific Asia Journal of the Association for Information Systems* 2022; 14(6): 1.

- Guerranti F., Dimitri GM. A comparison of machine learning approaches for predicting employee attrition. *Applied Sciences* 2022; 13(1): 267.
- Halder RK., Uddin MN., Uddin MA., Aryal S., Khraisat A. Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications. *Journal of Big Data* 2024; 11(1): 113.
- Jatobá M., Santos J., Gutierrez I., Moscon D., Fernandes PO., Teixeira JP. Evolution of artificial intelligence research in human resources. *Procedia Computer Science* 2019; 164: 137-142.
- Karakütük, K. İnsan kaynakları. *Ankara University Journal of Faculty of Educational Sciences (JFES)* 1997; 30(1).
- Khatri S., Pandey DK., Penkar D., Ramani J. Impact of artificial intelligence on human resources. In *Data Management, Analytics and Innovation: Proceedings of ICDMAI 2019 2020 Springer Singapore*; 2: 365-376.
- Kişî N., Özer MA. İnsan kaynakları yönetiminde yapay zekâ teknolojisinin benimsenmesi üzerine güç alanı analizi. *Afyon Kocatepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi* 2023; 26(Özel Sayı): 35-52.
- Korytkowski M., Nowak J., Scherer R., Zbieg A., Żak B., Relikowska G., Mader P. Employee turnover prediction from email communication analysis. In *International Conference on Artificial Intelligence and Soft Computing Cham: Springer International Publishing, Haziran 2022*: 252-263.
- Marvin G., Jackson M., Alam MGR. A machine learning approach for employee retention prediction. In *2021 IEEE Region 10 Symposium (TENSYP)*, Ağustos 2021; IEEE: 1-8.
- Mhatre A., Mahalingam A., Narayanan M., Nair A., Jaju S. Predicting employee attrition along with identifying high risk employees using big data and machine learning. *2nd International Conference on Advances in Computing, Communication Control and Networking, Aralık 2020*; IEEE: 269-276.
- Mohbey KK. Employee's attrition prediction using machine learning approaches. In *Machine Learning and Deep Learning in Real-Time Applications 2020*; IGI Global: 121-128.
- Murphy KP. *Machine learning: a probabilistic perspective*. MIT press 2012; 2-25.
- Nawaz N., Arunachalam H., Pathi BK., Gajenderan V. The adoption of artificial intelligence in human resources management practices. *International Journal of Information Management Data Insights* 2024; 4(1): 100208.
- Palos-Sánchez PR., Baena-Luna P., Badicu A., Infante-Moro J. Artificial intelligence and human resources management: A bibliometric analysis. *Applied Artificial Intelligence* 2022; 36(1): 2145631.
- Patel HH., Prajapati P. Study and analysis of decision tree based classification algorithms. *International Journal of Computer Sciences and Engineering* 2018; 6(10): 74-78.

- Reddy EMK., Gurralla A., Hasitha VB., Kumar KVR. Introduction to Naive Bayes and a review on its subtypes with applications. *Bayesian Reasoning and Gaussian Processes for Machine Learning Applications* 2022; 1: 1-14.
- Strohmeier S., Piazza F. Artificial intelligence techniques in human resource management—a conceptual exploration. *Intelligent Techniques in Engineering Management: Theory and Applications* 2015; 87: 149-172.
- Tian X., Pavur R., Han H., Zhang L. A machine learning-based human resources recruitment system for business process management: using LSA, BERT and SVM. *Business Process Management Journal* 2023; 29(1): 202-222.
- Uzak B. Telekomünikasyon sektöründe çalışan kaybı tahmini için makine öğrenmesi modeli seçimi. Bursa Uludağ Üniversitesi Fen Bilimleri Enstitüsü, Yayınlanmamış Yüksek Lisans Tezi, 2022.
- Widiyawati LM. Prediction of company employee resignation using naïve bayes algorithm. *International Journal of Scientific Engineering and Science* 2023 Jakarta, Indonesia.
- Zhu H. Research on human resource recommendation algorithm based on machine learning. *Scientific Programming* 2021; 1: 8387277.