





# Eğitimde Makine Öğrenmesi: Araştırmalardaki Güncel Eğilimler Üzerine İnceleme

Ergin Tosunoğlu<sup>1</sup>  Ramazan Yılmaz<sup>2</sup>  Eren Özeren<sup>3</sup>  Zeynep Sağlam<sup>4</sup> 

<sup>1</sup> Öğretmen, Bartın İl Milli Eğitim Müdürlüğü, [ergintosunoglu@gmail.com](mailto:ergintosunoglu@gmail.com)  
(Sorumlu Yazar/Corresponding Author)

<sup>2</sup> Doç. Dr. Bartın Üniversitesi, [ramazanyilmaz067@gmail.com](mailto:ramazanyilmaz067@gmail.com)

<sup>3</sup> Öğretmen, Bartın İl Milli Eğitim Müdürlüğü, [erenozeren78@gmail.com](mailto:erenozeren78@gmail.com)

<sup>4</sup> Öğretmen, Çaycuma İlçe Milli Eğitim Müdürlüğü, [zeynepelicora@gmail.com](mailto:zeynepelicora@gmail.com)

## Makale Bilgileri

## ÖZ

### Makale Geçmişi

Geliş: 27.05.2021

Kabul: 20.06.2021

Yayın: 30.09.2021

### Anahtar Kelimeler:

Makine öğrenmesi,  
Eğitim araştırmaları,  
Yöntemsel eğilimler,  
Makale analizi.

Araştırmanın amacı eğitim bilimleri alanında makine öğrenmesi yöntemleri kullanılarak yazılmış makalelerdeki eğilimleri incelemektir. Web of Science veri tabanında 2015 yılından 2020 yılına kadar olan 5 yıllık süreci kapsayan, konusunda “machine learning” geçen, “Educational Research” veya “Educational Scientific Disciplines” kategorilerinde bulunan çalışmalarını kapsamaktadır. Bu kapsamda tespit edilen 201 çalışmadan 17’si kapsam dışında kalmış ve 201 makale incelemeye alınmıştır. Araştırmada makalelerin incelenmesinde içerik analiz yöntemi kullanılmıştır. Makalelerin içerik analizi araştırmacılar tarafından geliştirilen makale değerlendirme formu kullanılarak yapılmıştır. Yapılan analizler sonucunda makalelerin %42,9’unun deneysel veya uygulamalı çalışma ile gerçekleştirildiği, makalelerde kullanılan makine öğrenmesi yöntemleri incelendiğinde %22,5 karar ağaçları, %17,8 destek vektör makinesi, %14,2 navie bayes sınıflandırma yöntemlerinin kullanıldığı görülmüştür. Makalelerde kullanılan örneklem profilleri incelendiğinde lisans öğrencilerinin %35,3 oranla en fazla olduğu, örneklem sayısındaki yoğunlaşmanın %29,7 oranla 101-1000 arasında olduğu görülmüştür. Makalelerde kullanılan veriler %34 oranıyla veri tabanları ve log kayıtlarından elde edilmiştir. Bu çalışmada son 5 yıllık süreçte yapılmış çalışmalar incelenmiştir. Araştırma sonuçlarının eğitimde makine öğrenmesi yöntemlerinin kullanılmasına konusunda ulusal alanyazına katkı sağlayacağı, araştırmacılara bu konuda fikir vereceği düşünülmektedir.

## Machine Learning in Education: A Study on Current Trends in Researchs

### Article Info

### ABSTRACT

#### Article History

Received: 27.05.2021

Accepted: 20.06.2021

Published: 30.09.2021

#### Keywords:

Machine learning,  
Educational research,  
Methodological  
trends,  
Article analysis.

The aim of the research is to examine the trends in the articles written using machine learning methods in the field of educational sciences. It covers the studies in the "Educational Research" or "Educational Scientific Disciplines" categories in the Web of Science database, covering the 5-year period from 2015 to 2020, with "machine learning" in their subject. In this context, 17 of the 201 studies identified were out of the scope and 201 articles were included in the review. Content analysis method was used to analyze the articles in the research. Content analysis of the articles was made using the article evaluation form developed by the researchers. As a result of the analysis, it was seen that 42.9% of the articles were carried out with experimental or applied studies, and when the machine learning methods used in the articles were examined, 22.5% decision trees, 17.8% support vector machine, 14.2% navie bayes classification methods were used. When the sample profiles used in the articles were examined, it was seen that undergraduate students were the highest with 35.3%, and the concentration in the number of samples was between 101-1000 with a ratio of 29.7%. The data used in the articles were obtained from databases and log records with a rate of 34%. In this study, the studies carried out in the last 5 years were examined. It is thought that the results of the research will contribute to the national literature on the use of machine learning methods in education and will give an idea to the researchers.

**Atf/Citation:** Tosunoğlu, E., Yılmaz, R., Özeren, E., Sağlam, Z. (2021). Eğitimde Makine Öğrenmesi: Araştırmalardaki Güncel Eğilimler Üzerine İnceleme, *Ahmet Keleşoğlu Eğitim Fakültesi Dergisi*, 3(2), 178-199.

“This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) (CC BY-NC 4.0)”



## GİRİŞ

Teknolojinin insan hayatını her geçen gün daha fazla etkilemesi eğitimde de varlığını göstermektedir. Özellikle Covid-19 pandemisi süreci kapsamında eğitimin tamamen teknolojik cihazlar ile verilmesi teknolojiyi eğitimin olmazsa olmaz ögesi haline getirmiştir. 21.yüzyıl becerileri kavramı ile birlikte, öğrenmenin sadece sınıfta değil hayat boyu devam eden süreçte her yerde gerçekleşeceği, öğretmenin yegane bilgi kaynağı olmadığı yeni bir eğitim anlayışına ihtiyaç vardır (Dede, 2011). Öğrenciler artık kendine sunulan değil, kendi istedikleri bilgileri öğrenme eğilimindedir (Dereli, 2019) ve kendi hızlarında ilerleyerek öğrenmek daha çok ilgilerini çekmektedir (Çelen ve Seferoğlu, 2020). MOOC adı verilen kurslarda zaman ve mekan sorunu olmadan, istek ve ihtiyaçlar doğrultusunda istenilen eğitimler alınabilmektedir. Bu bağlamda ülkeler eğitim sistemlerini çağın ihtiyaçları doğrultusunda düzenlemeli (Sağlam vd., 2020), eğitim materyalleri ve ortamları düzenlenirken bireysel farklılıklar dikkate alınmalıdır (Gülcüoğlu vd., 2020).

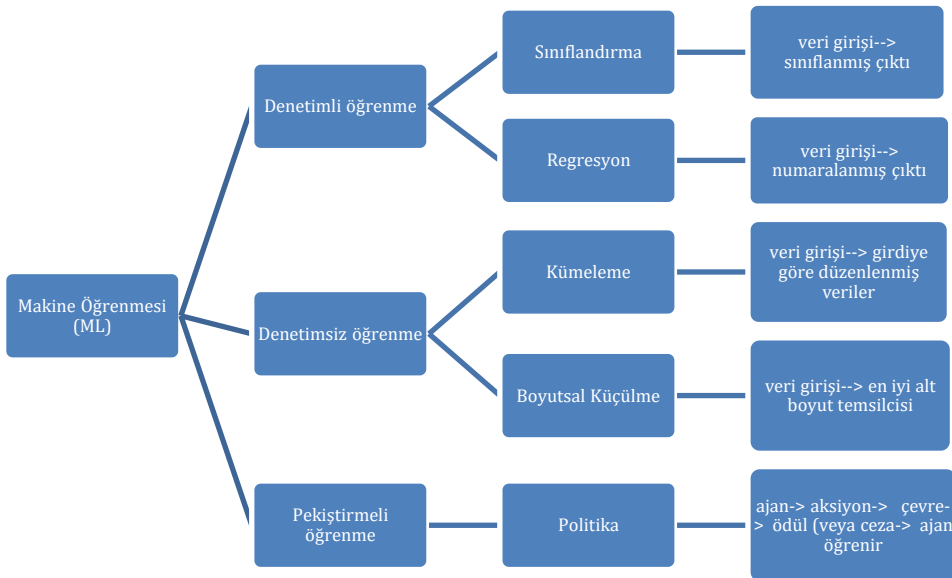
Milyarlarca insanın aktif olarak sosyal medya kullandığı bir dönemde eğitim öğretim kurumlarının dijitalleşme sürecinin nasıl olacağı merak konusudur (Parlak, 2017). Türkiye son yıllarda e-devlet uygulamaları ile birçok alanda dijital dönüşümde adımlar atmıştır. Eğitim alanında e-okul, Eğitim Bilişim Ağı (EBA), Fırsatları Artırma Teknolojiyi İyileştirme Hareketi (FATİH) eğitimin dijitalleşmesinde atılan adımlardandır. Bu adımlar olumlu yönde olmasına rağmen henüz yeterli değildir. Teknoloji bütün sektörlerde hızla dönüşüme sebep olurken eğitim alanında daha çok dönüşüm olması beklenmektedir (Akgün, 2019). Eğitimde dijital dönüşümün artması öğrencilerin teknoloji bağımlılığını azaltacak, derslere olan ilgiyi artıracaktır (Sarsıcı ve Çelik, 2019).

Dijital platformlardan elde edilen veriler istatistiksel olarak bilgi çıkarımında bulunmak amacıyla kullanılmaktadır (Avcı, 2015). Eğitim platformlarında tutulan veriler de eğitsel veri madenciliği uygulamalarında kullanılmaktadır. Çok fazla verinin olduğu bu ortamlarda verilerden anlamlı ve doğru bilgiler çıkarabilmek çok önemlidir. Eğitsel veri madenciliği yöntemleri ile öğrenene ilişkin veriler analiz edilerek öğretmenlere ve yöneticilere tavsiyeler sunulmaktadır (Türel, 2016; Pena, 2013). Makine öğrenmesi algoritmaları ile geliştirilmiş sistemlerde öğrenciler bireysel özellikleri doğrultusunda ders içeriklerini tercih ettikleri yöntemlerle alabilecek, öğrenme hızlarına göre ilerleyecek, eksik oldukları noktaları verilen geri dönütler sayesinde görebilecek ve tamamlayabilecektir. Öğretmenler hazırladığı dokümanları sistemde bir kere oluşturduğunda tekrar kullanabilecek, öğrencinin verilen görevleri yapıp yapmadığını, hangi görevleri ne kadar yaptığını, öğrencinin bireysel özelliklerine uygun olarak geri dönütleri takip edebilecek, gereken düzenlemeleri yapabilecektir. Bu yöntemler sayesinde öğretmen öğretenden çok öğrenciyi takip eden, ona yol gösteren, ilerlemesine katkı sağlayan bir lider konumuna yerleşecektir. Eğitim sistemi kapsamında öğrencinin devam durumu, hangi çalışmaları yaptığı, ilerlemesinin nasıl olduğu, bireysel özellikleri, ölçme değerlendirme sonuçları gibi öğrenciye yönelik özellikler takip edilebilecektir. Öğretmenlerin sistem içinde yaptıkları çalışmalar, sistemi nasıl kullandıkları, öğrencilere verdikleri dönütler görülebilecektir.

## Makine Öğrenmesi

Dijitalleşen dünya ile birlikte nesnelerin interneti (IoT), bulut bilişim, büyük veri (Big Data), siber güvenlik, yapay zeka, makine öğrenmesi gibi kavramlar sık kullanılmaya başlanmıştır. Büyük veri kavramından bahsedilecek olunursa, bilgisayarlar kullanılmaya başlanmadan önce defterlerde saklanabilen kayıtların günümüz dünyasında aynı şekilde muhafaza edilmesi mümkün görünmemektedir. Bilginin her geçen gün katlanarak büyümesi bu kadar çok veri yığını içinden gerekli olan bilginin bulunması karmaşasını doğurmuştur. Veri madenciliği bu noktada karşımıza çıkmaktadır, büyük miktarda verinin içinden anlamlı ilişkiler çıkararak, bu ilişkilerin analiz edilmesidir (Savaş vd., 2012).

Verinin içinden işe yarayan bilginin çıkarılmasının yanında bu bilgiyle ileride olacaklar öngörülebilir, bu öngörme işi çok önemlidir. İnsanlar yaşadığı tecrübeler, aldığı eğitimler doğrultusunda ileriye yönelik tahminlerde bulunabilmektedir. Ancak insan karar verirken duygularının da etkisiyle rasyonel olarak karar vermekte zorlanabilmekte, bunun yanında çok fazla veri olduğu durumlarda insanların gözünden kaçan ayrıntılar yanlış kararlara sebep olabilmektedir. Ancak eğitilmiş makineler bütün durumları değerlendirip, milyonlarca veriyi çok hızlı bir şekilde inceleyerek çok kısa sürede rasyonel kararlar verebilmektedir. Makine öğrenmesinde insan düşünce yapısını ve karar verme yetisini modelleyebilmek çok önemlidir (Ülker ve İnik, 2017). Bu nedenle makine öğrenmesi için, insan zekasını taklit etmek üzere tasarlanmış hesaplamalı algoritmaların gelişen bir dalı (El Naqa ve Murphy, 2015), bir problemin veriye uygun algoritma ile modellenmesidir denilebilir (Atalay ve Çelik, 2017). Büyük veri çeşitlilik, hız ve hacim nedeniyle büyük sorunları beraberinde getirmektedir (Suthaharan, 2014). Makine öğrenmesi algoritmaları büyük verileri uygun şekilde sınıflandırarak bu sorunların önüne geçmektedir. Şekil 1’de makine öğrenmesi algoritmalarının hiyerarşik yapısı gösterilmektedir.



Şekil 1. Makine öğrenmesi algoritmaları hiyerarşisi (Rashidi vd., 2019)

Verinin türüne göre kullanılması gereken farklı makine öğrenmesi algoritmaları mevcuttur. Şekil 1 makine öğrenmesi stiline göre algoritmalara ait hiyerarşiyi göstermektedir. Algoritmalar özelliğine göre kümeleme,

sınıflandırma, tahmin ve kestirim yapabilme özelliklerine sahiptirler. Bu algoritmalar denetimli, denetimsiz ve pekiştirmeli olmak üzere 3 strateji ile makinelere öğretilmektedir. Denetimli öğrenmede makineye hem girdi hem çıktı değeri verilir, denetimsiz öğrenmede makineye sadece girdi değeri verilir, pekiştirmeli öğrenme ise girdi değerlerine karşılık gelecek en uygun çıktı edilmesi sırasında katsayıların en uygun değerlerinin bulunması ile sağlanır (Ülker ve Civalak, 2002). Hangi makine öğrenmesi algoritmasının kullanılması gerektiğini uygulamayı yazan programcı belirlemektedir. Makine öğrenmesinde kullanılacak çok fazla algoritma vardır, bu algoritmalar arasından veriyi işleyebilecek en uygun algoritmayı seçmek önemlidir. Algoritmaları işlevleri açısından gruplandırmak, uygun algoritmayı seçmede yol gösterici olacaktır. Tablo 1’de işlevlerine göre makine öğrenmesi alanında kullanılan algoritmalar verilmiştir.

**Tablo 1. İşlevlerine Göre Algoritmalar (machinelearningmastery, 2020)**

Algoritma Kategorisi	Alt algoritmalar	İşlevleri
Regresyon algoritmaları	<ul style="list-style-type: none"> <li>• Sıradan en küçük kareler regresyonu (OLSR)</li> <li>• Doğrusal regresyon</li> <li>• Lojistik regresyon</li> <li>• Adımsal regresyon</li> <li>• Çok değişkenli uyarlanabilir regresyon (MARS)</li> <li>• Yerel olarak tahmini dağılım grafiği düzenleme (LOESS)</li> </ul>	Bağımsız değişkenden yola çıkarak bağımlı değişkene ilişkin kestirimde bulunmak amacıyla kullanılırlar.
Örnek tabanlı algoritmalar	<ul style="list-style-type: none"> <li>• k-En Yakın Komşu (kNN)</li> <li>• Vektör Niceleme (LVQ) Öğrenme</li> <li>• Kendi Kendini Düzenleyen Harita (SOM)</li> <li>• Yerel Ağırlıklı Öğrenme (LWL)</li> <li>• Destek Vektör Makineleri (SVM)</li> </ul>	Örnek bir veri tabanı oluşturarak, en iyi eşleşmeyi bulmak ve bir tahminde bulunmak için verileri veri tabanıyla karşılaştırır.
Düzenleştirme algoritmaları	<ul style="list-style-type: none"> <li>• Sırt Regresyonu</li> <li>• En Küçük Mutlak Büzülme ve Seçim Operatörü (LASSO)</li> <li>• Elastik ağ</li> <li>• En Küçük Açılı Regresyon (LARS)</li> </ul>	Karmaşık modeller üzerinde daha iyi olan regresyon algoritmaları.
Karar ağacı algoritmaları	<ul style="list-style-type: none"> <li>• Sınıflandırma ve Regresyon Ağacı (CART)</li> <li>• Yinelemeli dikotomatör 3 (ID3)</li> <li>• C4.5 ve C5.0 (güçlü bir yaklaşımın farklı sürümleri)</li> <li>• Ki-kare Otomatik Etkileşim Algılama (CHAID)</li> <li>• Güdük karar</li> </ul>	Verilerdeki özniteliklerin gerçek değerlerine dayalı olarak bir karar modeli oluşturur. Sınıflandırma ve regresyon problemlerine yönelik veriler

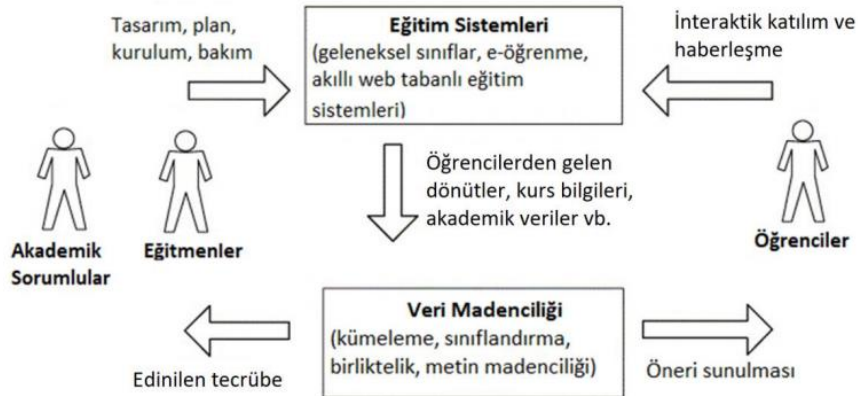
	<ul style="list-style-type: none"> <li>• M5</li> <li>• Koşullu Karar Ağaçları</li> </ul>	üzerinde eğitilir. Hızlı ve doğru karar veren bir makine öğrenimi modelidir.
Bayes algoritmaları	<ul style="list-style-type: none"> <li>• Naive Bayes</li> <li>• Gauss Saf Bayes</li> <li>• Çok terimli Naif Bayes</li> <li>• Ortalamalı Tek Bağımlılık Tahmincileri (AODE)</li> <li>• Bayes İnanç Ağı (BBN)</li> <li>• Bayes Ağı (BN)</li> </ul>	Sınıflandırma ve regresyon problemleri için Bayes teoremini uygulayan algoritmalarıdır.
Kümeleme algoritmaları	<ul style="list-style-type: none"> <li>• k-Means</li> <li>• k-Medians</li> <li>• Beklenti Maksimizasyonu (EM)</li> <li>• Hiyerarşik kümeleme</li> </ul>	Grupların ortak özellikleri maksimum olacak şekilde düzenlenmesini sağlar.
Birliktelik kuralı öğrenme algoritmaları	<ul style="list-style-type: none"> <li>• Apriori algoritması</li> <li>• Eclat algoritması</li> </ul>	Verilerdeki değişkenler arasında gözlemlenen ilişkileri en iyi açıklayan kuralları çıkarır. Büyük veri kümeleri üzerindeki önemli ve yararlı ilişkilerin keşfedilmesini sağlar.
Yapay sinir ağı algoritmaları	<ul style="list-style-type: none"> <li>• Algılayıcı</li> <li>• Çok Katmanlı Algılayıcılar (MLP)</li> <li>• Geri Yayılım</li> <li>• Stokastik Gradyan İniş</li> <li>• Hopfield Ağı</li> <li>• Radyal Temel Fonksiyon Ağı (RBFN)</li> </ul>	Biyolojik sinir ağı yapısından ilham alan modeldir. Regresyon ve sınıflandırma problemleri başta olmak üzere her türlü problemde kullanılabilen algoritmalarıdır.
Derin öğrenme algoritmaları	<ul style="list-style-type: none"> <li>• Evrişimli Sinir Ağı (CNN)</li> <li>• Tekrarlayan Sinir Ağları (RNN'ler)</li> <li>• Uzun Kısa Vadeli Bellek Ağları (LSTM'ler)</li> <li>• Yığılmış Otomatik Kodlayıcılar</li> <li>• Derin Boltzmann Makinesi (DBM)</li> <li>• Derin İnanç Ağları (DBN)</li> </ul>	Çok büyük ve karmaşık sinir ağları oluştururlar. Ses, görüntü, metin gibi büyük veri kümeleriyle çalışırlar.
Boyut azaltma algoritmaları	<ul style="list-style-type: none"> <li>• Temel Bileşen Regresyonu (PCR)</li> <li>• Kısmi En Küçük Kareler Regresyonu (PLSR)</li> <li>• Sammon Haritalama</li> <li>• Çok Boyutlu Ölçekleme (MDS)</li> </ul>	Denetimsiz öğrenme ile verilerdeki doğal yapıları arar ve kullanır, daha az bilgi kullanarak

	<ul style="list-style-type: none"> <li>• Lineer Diskriminant Analizi (LDA)</li> <li>• Karışım Diskriminant Analizi (MDA)</li> <li>• Kuadratik Diskriminant Analizi (QDA)</li> <li>• Esnek Diskriminant Analizi (FDA)</li> </ul>	verileri özetleyebilir.
Topluluk algoritmaları	<ul style="list-style-type: none"> <li>• Artırma</li> <li>• Önyüklemeli toplama (Bagging)</li> <li>• AdaBoost</li> <li>• Ağırlıklı Ortalama (Karıştırma)</li> <li>• Yığılmış Genelme (Yığınlama)</li> <li>• Gradient Güçlendirme Makineleri (GBM)</li> <li>• Gradyan Artırılmış Regresyon Ağaçları (GBRT)</li> <li>• Rastgele Orman</li> </ul>	Bağımsız olarak eğitilmiş birden fazla zayıf modelin genel tahmini yapması için bir araya getirilmiş modelidir.
Diğer algoritmalar	<ul style="list-style-type: none"> <li>• Özellik seçim algoritmaları</li> <li>• Algoritma doğruluk değerlendirmesi</li> <li>• Performans ölçüleri</li> <li>• Optimizasyon algoritmaları</li> <li>• Hesaplamalı zeka (evrimsel algoritmalar, vb.)</li> <li>• Bilgisayarla Görme (CV)</li> <li>• Doğal Dil İşleme (NLP)</li> <li>• Öneri Sistemleri</li> <li>• Pekıştirmeli Öğrenme</li> <li>• Grafik Modeller</li> </ul>	Özel görevler için geliştirilmiş algoritmalar, sayısı artırılabilir.

Tablo 1 işlevlerine göre algoritmaları göstermektedir, en yaygın makine öğrenmesi algoritmaları k-en yakın komşu, Bayes sınıflandırıcısı, karar ağaçları, lojistik regresyon, destek vektör makineleri ve yapay sinir ağlarıdır. Büyük veri eğitim, sağlık, ticaret, bankacılık, alışveriş, lojistik gibi bilgisayar ve internet kullanılan birçok alanda karşımıza çıkmaktadır. Her sektör kendi veri yapılarından veri madenciliği yöntemleri ile çıkardığı bilgileri kullanarak kendini geliştirmeye çalışmaktadır. Sağlık sektöründe doktorlar tarafından yapılan teşhis ve tedavilerde daha doğru sonuçlar elde etmek, insan kaynaklı hataları engellemek ve doktorun kararına yardımcı olmak amacıyla makine öğrenmesi tabanlı karar destek sistemleri kullanılmaktadır (Karakoyun ve Hacıbeyoğlu, 2014). Ekonometri alanındaki verilerin hacimsel olarak çok büyük olması bu verilerin analizini de zorlaştırmaktadır, verilerin analizi için makine öğrenmesi algoritmaları kullanılmaktadır (Akay, 2018). Twitter gibi sosyal medya ortamlarında bireylerin paylaşımları makine öğrenmesi algoritması ile sınıflandırılıp profil çıkarma çalışmaları yapılmaktadır (Tataroğlu, 2019). Hissedilen duyguların insan beyninde oluşturduğu elektrik sinyalleri incelenerek duygu analizlerinde makine öğrenmesi kullanılmaktadır (Er, 2019). Psikolojiden sosyolojiye, ekonomiden siyasete büyük verinin var olduğu her alanda makine öğrenmesi kullanılırken eğitim alanında kullanılmaması kaçınılmazdır.

## Eğitim Alanında Makine Öğrenmesi

Büyük verinin var olduğu alanlardan biri de eğitimidir. Dijital dönüşüm süreçlerinde okulların ve eğitim içeriklerinin de dönüşümü söz konusudur. Özellikle Covid-19 pandemisi süreçleri kapsamında geçilen uzaktan eğitim döneminde dijitalleşme isteğe bağlı değil zorunlu hale gelmiştir. Eğitim alanında tutulan veriler çok çeşitlidir, ancak eğitimde en önemli hedef öğrenenin eğitimini gerçekleştirmesidir. Eğitim alanında ortaya çıkan büyük verinin analiz edilmesi eğitimin hedeflerini tahmin etme, eğitimde alınması gereken önlemlerin erken alınması, öğrenenin öğrenme yöntemini ve öğrenme biçimini belirleyerek uygun içerikleri önerme, öğrenenin ölçme değerlendirme sonuçlarına göre eksik kaldığı konuları belirleme gibi alanlarda kullanılmaktadır. Geleneksel eğitim ortamlarında öğrenciler ve öğretmenler yüz yüze olduğu için öğrencinin durumu öğretmen tarafından izlenebilmekte ve dönüt yine aynı ortamda verilebilmektedir, ancak bu sistemde bile bazı sorunlar mevcuttur. Öncelikle aynı sınıf ortamında çok farklı bireysel özelliklere sahip öğrencilerin aynı yöntemle, aynı hızda öğrenmesi ve anlaması beklenmektedir. Öğretmen kalabalık sınıflarda öğrencilerin hepsini aynı oranda takip etmekte zorlanmaktadır. Eğitimin diğer bir boyutu olan uzaktan eğitimde ise daha az etkileşim olduğundan öğrenciyi tanımak daha da zorlaşmaktadır. Öğrencinin kendi hızında, anlayabileceği şekilde ilerlerken kendisini geliştirebileceği, anlayamadığı ve eksik olduğu konularda sistem tarafından yönlendirildiği bir yapı öğrenmeyi kolaylaştırmaktadır. Bunun yanında öğretmen kendisine yardım eden, öğrenci özelliklerini belirleyen, öğrenciye dönüt veren bir sistemde öğrenciyi daha iyi gözlemleyebilmekte ve ona çalışmalarını konusunda liderlik edebilmektedir. Eğitimsel veri madenciliği yöntemleri ile bütün bunlar sistem üzerinde gerçekleşmektedir. Şekil 2’de eğitim sistemlerinde veri madenciliği döngüsünün nasıl olduğunu göstermektedir.



**Şekil 2.** Eğitim sistemlerinde veri madenciliği döngüsü (Romero & Ventura, 2007; Akt: Çiftçi vd., 2018)

Şekil 2 incelendiğinde akademik sorumlular ve öğretmenler tarafından hazırlanan içeriklerin sistem tarafından öğrenciye sunulduğu ve öğrenciden alınan verilerin veri madenciliği yöntemleri ile analiz edilerek öğrenci ve sorumlulara dönütler sağlandığı görülebilmektedir.

Veri madenciliği yöntemleri uzaktan eğitimde elde edilen verilere uygulanabileceği gibi geleneksel sınıf ortamlarındaki verilere de uygulanabilir. Geleneksel model, sınıflarda en yaygın kullanılan eğitim modelidir. Tüm dünyada çocuklar temel eğitimi alabilmek için okullara devam etmektedir, bunun yanında halk eğitim, yetişkin eğitimi, yüksek öğretim gibi farklı kademeleri bulunur. Öğretmen öğrencilerine yüz yüze eğitim vermektedir. Sınıf ortamında öğrencilerden alınan her türlü veri veri madenciliğinde kullanılabilir, ancak bu verileri kullanabilmek için kayıt altına almak gerekir. Sınıf ortamında tek başına olan öğretmenin çok çeşitli veriyi toplayıp kayıt altına alması mümkün değildir. Genellikle öğrencilerin yaptığı çalışmalar, verilen ödevler, yapılan sınavlar, öğretmenin gözlemleri verileri oluşturmaktadır.

Uzaktan eğitimde ise öğretmenler ve öğrenciler aynı sınıf ortamında değildir. Uzaktan eğitim çok çeşitli yöntemlerle yapılabilmektedir, kağıt üzerinde, video ile, web üzerinden yapılabilmektedir. Uzaktan eğitim kapsamına web tabanlı kurslar, öğrenme yönetim sistemleri (learning management system) girmektedir. Web tabanlı kurslarda kişiler istedikleri kurslara kayıt olmakta, ilgi ve ihtiyaçları doğrultusunda kendi hızlarında öğrenmelerini takip etmektedirler. Bu kurslarda öğrenenlerin dersleri nasıl takip ettiği, ne sıklıkla derslere katıldığı, hangi içerikleri izlediği, hangi derste ne kadar zaman geçirdiği, izlemeden atladığı dersler takip edilmekte veriler tutulmaktadır. Öğrenme yönetim sistemleri sistemlerinde öğretmenlerin kendi sınıflarını oluşturmasına, içeriklerini üretmesine, öğrencilerle senkron-asenkron dersler hazırlamasına, sınavlar hazırlamasına izin verilmektedir. Öğrencilerin derslere katılımına, sınavlarda aldığı notlar kayıt altına alınmaktadır. Sınıf ortamında ve uzaktan eğitimde elde edilen veriler veri madenciliği yöntemleri ile analiz edilerek öğrencilere, öğretmenlere ve yöneticilere öneriler sunulmaktadır. Eğitim alanında makine öğrenmesi yöntemleri ile yapılmış birçok çalışma mevcuttur. Çalışmalar alanda olan eksikleri, hataları belirleyip, öğrenenlere ve öğretenlere tavsiyelerde bulunulmuştur. Eğitim alanında kullanılan makine öğrenmesi algoritmaları eğitsel veri madenciliği süreçlerinde kullanılan algoritmalar olarak da karşımıza çıkmaktadır.

Gök (2017), yaptığı çalışmada 6., 7. ve 8. sınıf öğrencilerine anket uygulamış, anket verileri üzerinde regresyon ve sınıflandırma algoritmaları ile yaptığı makine öğrenmesi ile öğrencilerin Türkçe, Matematik ve dönem sonu genel başarılarını tahmin etmiştir. Güner ve Çomak (2011) mühendislik fakültesine başlayan 434 öğrencinin ÖSYM ve KYK bilgilerini kullanarak Destek Vektör Makinesi (DVM) algoritmaları ile makine öğrenimi modeli uygulamış ve matematik dersinde başarılı olacak öğrencileri %86,36 doğruluk oranı ile tahmin etmiştir. Aydoğan ve Karıcı (2018) 'nın yaptığı bir diğer çalışmada ise meslek yüksek okullarına devam eden öğrencilerin başarı durumlarını etkileyen faktörler analiz edilmiştir, denetimli ve denimsiz öğrenme modelleri python programlama dili ile programlanmıştır. Araştırmada ailesinin yakınında okuyanların daha başarılı olduğu, meslek yüksek okullarına sınavsız yerleşen meslek lisesi mezunu öğrencilerin sınavla yerleşenlere göre başarı durumlarının düştüğü görülmüştür. Dönem içinde yapılan sınav, ödev ve projelerin dönem sonu başarısına etkisi makine öğrenmesi ile modellenilebilmekte ve başarı tahmin oranı doğru bir şekilde tespit edilebilmektedir (Kardaş ve Güvenir, 2020). Yıldız (2014) yaptığı doktora tezinde uzaktan eğitimde kullanılan bir öğretim yönetim sisteminin verileri üzerinde makine öğrenmesi algoritmaları kullanarak öğrenci başarısını tahmin etmiştir. Filiz (2019) yaptığı doktora tezi araştırmasında Uluslararası Matematik ve Fen Eğilimleri (TIMSS) araştırma sonuçları



üzerinde makine öğrenmesi yöntemleri ile başarıları sınıflandırabilecek en iyi algoritmaları bulmanın yanında, başarıları etkileyen öznelikleri tespit etmiştir. Makine öğrenmesi yöntemleri ile öğrenciler hakkında elde edilen veriler öğretmenlere ve yöneticilere yol gösterici olmaktadır.

Eğitsel veri madenciliği yöntemleri genellikle öğrenci başarısı ve performansı üzerine odaklanmıştır ancak eğitimcilerin performanslarının izlenmesine yönelik yapılan çalışmalar da mevcuttur. Gazi Üniversitesi'nde yapılan araştırmada öğrencilerin öğretmenleri hakkında doldurdukları ölçek eğitimsel veri madenciliği algoritmaları ile analiz edilmiş ve %97,70 oranında doğru tahmine ulaşılmıştır (Çifçi vd., 2018). Eğitsel veri madenciliğinin uygulama alanları; öğrencinin öğrenme davranışına dayalı ders adaptasyonu ve öğrenme önerileri sağlayan uygulamalar, öğrenme materyalinin ve eğitim web tabanlı derslerin değerlendirilmesi ile ilgili yaklaşımlar, e-öğrenme kurslarında hem öğretmene hem öğrencilere geri bildirim sağlayan uygulamalar ve atipik öğrencilerin öğrenme davranışlarının saptanmasına yönelik uygulamalardır (Romero ve Ventura, 2010).

### **Araştırmanın Amacı**

Makine öğrenmesi birçok alanda olduğu gibi eğitim alanında da kullanılan bir yöntemdir. Eğitim alanında kullanılan makine öğrenmesi yöntemleri ile eğitsel veri madenciliği yapabilmek, zeki öğretim sistemleri geliştirebilmek, öğretmen ve öğrenciler için sistem tarafından otomatik tavsiyeler üretebilmek, uyarlanabilir ve bireyselleştirilmiş eğitim ortamları geliştirebilmek mümkün hale gelmiştir. Ülkeler eğitimin amaçlarına ulaşmak için eğitim sistemlerinde düzenlemeler yapmaktadır. Dijital dönüşümle birlikte eğitim alanında her geçen gün artan veri hacminin eğitimin geliştirilmesi için kullanılması çok önemlidir. Eğitimsel verilerin analiz edilmesi ve elde edilen verilerin doğru bir şekilde yorumlanması için doğru makine öğrenmesi algoritmaları ile eğitilmesi önemlidir. Eğitim alanında elde edilen verilerin doğru bir şekilde değerlendirilmesi yapılmış olan çalışmaların irdelenmesi ve örnek alınması önemlidir. Alanyazın tarandığında Türkiye' de eğitim alanında makine öğrenmesi yöntemleri ile yapılmış çalışmaların analiz edildiği bir çalışmaya rastlanmamıştır. Bu nedenle gerçekleştirilen bu çalışmanın alanyazındaki bu boşluğu dolduracağı düşünülmektedir. Türkiye'de henüz makine öğrenmesi ile eğitim verilerinin analiz edilmesi çalışmaları yenidir. Bu çalışma ile web of science veri tabanında yer alan eğitimde makine öğrenmesi yönteminin kullanıldığı makaleler taranarak incelenmiştir. Çalışmanın amacı Web of Science veri tabanında 2015-2019 yıllarını kapsayan, konusunda “machine learning” geçen, “Educational Research” veya “Educational Scientific Dicipines” kategorilerinde bulunan eğitim bilimleri alanında makine öğrenmesi yöntemleri kullanılarak yazılmış makalelerdeki eğilimleri incelemektir. Bu amaç doğrultusunda aşağıdaki sorulara yanıt aranmaktadır:

1. Eğitim alanında makine öğrenmesi yöntemleri kullanılarak yazılmış makalelerde hangi algoritmalar kullanılmıştır?
2. Makalelerin türü bakımından ağırlığı nasıldır?
3. Makalelerde kullanılan araştırma yöntemlerin ağırlığı nasıldır?
4. Makalelerde kullanılan örneklem düzeyleri nasıldır?

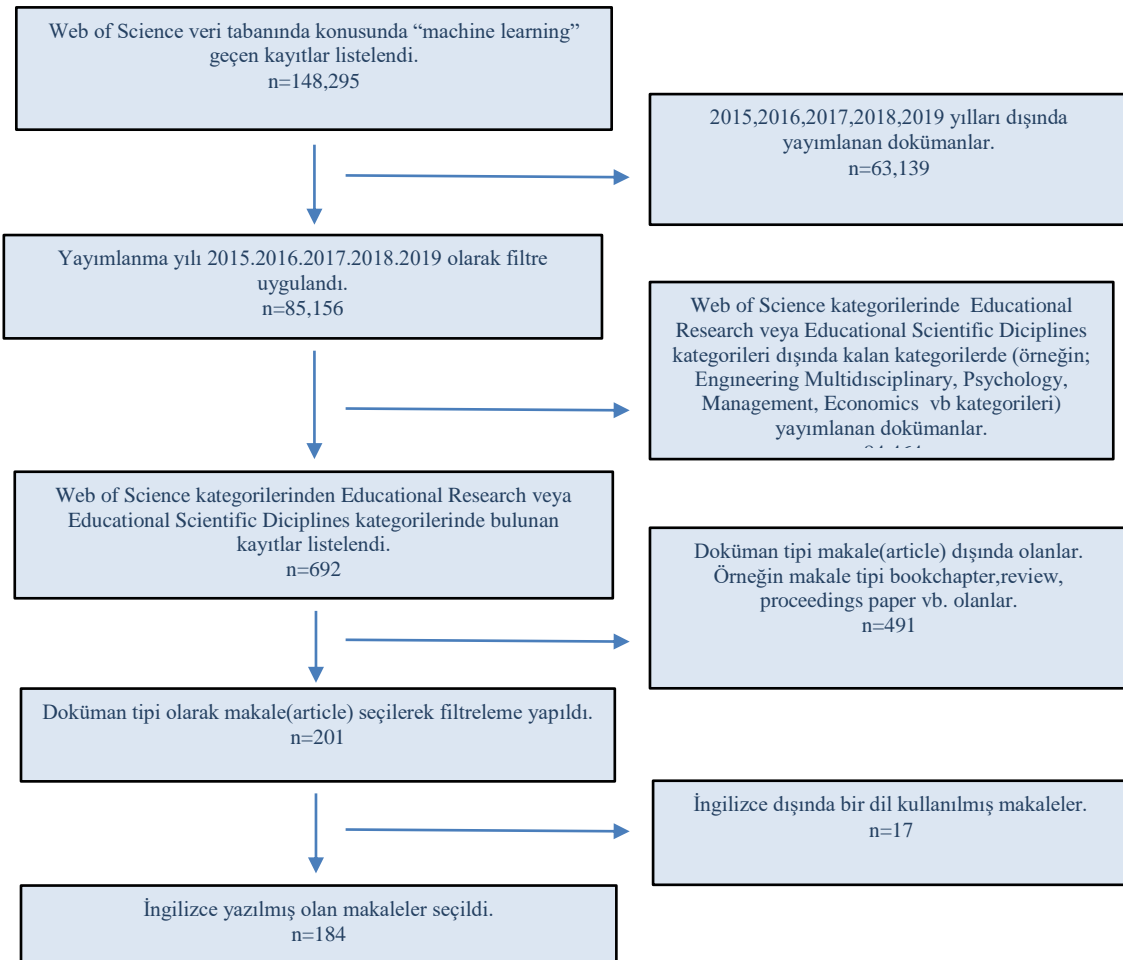
5. Makalelerde kullanılan örneklem sayılarının ağırlığı nasıldır?
6. Makaleler yoğun olarak hangi eğitim alanlarında yazılmıştır?

## YÖNTEM

Bu araştırmada makalelerin içeriksel analizinin gerçekleştirilmesinde içerik analizi yöntemi kullanılmıştır. İçerik analizi, sözel, yazılı ve diğer metaryallerin içerdiği mesajı nesnel ve sistematik olarak sınıflandırıp sayılara dönüştürerek çıkarımda bulunmadır(Tavşancıl ve Aslan,2001).

## Örneklem

Bu araştırmada incelenen çalışmalar Web of Science veri tabanında listelenen konusunda “machine learning” geçen Web of Science kategorilerinden “Educational Research” veya “Educational Scientific Dicipines” kategorilerinde bulunan makalelerdir. Yapılan tarama 2015 yılından 2020 yılına kadar olan 5 yıllık süreç içinde yapılan çalışmaları kapsamaktadır. Bu kriterlere göre yapılan sorgu sonucunda 201 makale tespit edilmiştir.



### Şekil 3. Örneklem Seçim Şeklinin Akış Şeması

Web of science veri tabanından sorgulama 24 Ekim 2020 tarihinde yapılmıştır. Betimsel Tarama yöntemi ile mümkün olduğu kadar genişletilmiş olan taramada her bir makale, ölçütler yardımıyla incelenmiş; araştırmacılarla yapılan ön değerlendirme sonucunda 13 çalışma kitap bölümü olduğu için, 4 çalışma ise yabancı dili İngilizce olmadığı (İspanyolca, Çince) için inceleme dışı bırakılmış olup 184 makale üzerinde inceleme işlemi yapılmıştır.

#### Verilerin Toplanması

Tarama sonucu incelenmesine karar verilen makaleler bilgisayar ortamına atılmıştır. Makalelerin sistematik bir şekilde incelenmesini sağlamak için makale inceleme formu oluşturulmuştur. Bu formun ana başlıkları; kullanılan yapay zekâ-makine öğrenmesi türü, makalenin türü, araştırma yöntemi/teknigi, veri toplama araçları, örneklem düzeyi, örneklem sayısı, öğrenme alanları olarak yazarlar tarafından belirlenmiştir. Bu ana başlıklar alt başlıklara ayrılıp gerekli kodlamalar yapılarak Microsoft Excel programında form hazırlanmıştır. Araştırmacılar tarafından bağımsız olarak işlenen formlar birleştirilip düzenlenerek veri kümesi oluşturulmuştur.

#### Verilerin Analizi

Sorgulama sonucunda kriterlere uygun olan 184 makale incelenerek oluşturulan formdaki verilere göre frekans ve yüzdeler hesaplanmıştır. Elde edilen bulgular tablolar kullanılarak sunulmuştur.

#### Geçerlilik ve Güvenilirlik

Araştırmacılar tarafından bağımsız olarak işlenen tablolar birleştirilerek uzlaşma yüzdesi hesaplanmıştır. Bu hesaplama için, Miles ve Huberman'ın (1994) “Güvenirlik = Görüş Birliği / (Görüş Birliği + Görüş Ayrılığı) x 100” denklemi kullanılmıştır. Güvenirlik katsayıları; Hangi yapay zekâ / makine öğrenmesi teknolojileri kullanılmış olduğu bölümünde %92, makalenin türü bölümünde %96, makalenin yöntemi bölümünde %98, veri toplama araçları bölümünde % 100, örneklem düzeyi bölümünde %100, örneklem sayısı bölümünde % 100 olarak hesaplanmıştır. Görüşlerin karşılaştırılması sonucunda ortaya çıkan farklılıklar üzerine tartışılarak uzlaşma sağlanmıştır. Böylece kişisel görüşlerin inceleme üzerindeki etkisi azaltılmaya çalışılarak geçerlik ve güvenilirlik hedefine ulaşılmaya çalışılmıştır.

#### BULGULAR

Bu bölümde, araştırmanın alt amaçları doğrultusunda yapılan analizler sonucunda elde edilen bulgular sunulmuştur. Makalelerin; kullanılan makine öğrenmesi algoritmalarına çalışma türleri, , araştırma yöntemi, örneklem düzeyleri, örneklem sayısı, öğrenme alanları ve veri toplama araçlarına ait dağılımları alt başlıklar halinde verilmiştir.

#### İncelenen Makalelerde Kullanılan Makine Öğrenmesi Algoritmalarına Göre Dağılımı

Araştırmanın birinci alt amacında makaleler kullanılan makine öğrenmesi algoritmalarına göre sınıflandırılmıştır. Çalışmalarda kullanılan veya üzerinde araştırma yapılan makine öğrenmesi algoritmaları ana başlıklar halinde incelenerek elde edilen sonuçlar Tablo-2' de verilmiştir.

**Tablo 2. Makalelerde Kullanılan Algoritmalar**

KULLANILAN ALGORİTMALAR	N	%
Karar Ağacı	57	22,53
Destek Vektör Makineleri	45	17,79
Naive Bayes	36	14,23
Yapay Sinir Ağları (Neural Network)	35	13,84
Doğrusal Regresyon	15	5,93
K-NN	14	5,53
Lojik Regresyon	13	5,14
NLP	10	3,95
K-Means	8	3,16
Sıralı Minimal Optimizasyon-SMO	4	1,58
Deep Learning	4	1,58
Genetik Algoritma	5	1,98
Diğer*	7	2,77

Diğer\* : Q-Learning, Bulanık Mantık, X-Means, Gause Regresyon, Expectation-Maximization (EM)

Yapılan incelemelerde çalışmaların % 39'unun birden fazla makine öğrenmesi yapısının kullanılarak gerçekleştirildiği görülmektedir. Tablo 2'deki verilere göre 57 çalışmada (%22,5) karar ağaçları kullanılmıştır. Naive bayes sınıflandırma algoritmaları 36 çalışmada (%14,2) kullanılırken, bayes algoritmaları çalışmalarda sıklıkla Karar ağacı yapıları ile beraber kullanılmıştır. Destek vektör makinesi algoritmalarının yer aldığı 45 araştırma (%17,8) bulunmakta olup bu çalışmalarda sıklıkla regresyon modellerinden (doğrusal-lojik) biride yer almaktadır. Yapay Sinir ağları 20 çalışmada tercih edilmiş olup genel olarak tek model olarak kullanılmıştır.

### İncelenen Araştırmaların Çalışma Türlerine Göre Dağılımı

Araştırmanın ikinci alt amacında makaleler türlerine göre sınıflandırılmıştır. Bu doğrultuda incelenen makalelerin türlerine göre dağılımları Tablo 3'de sunulmuştur.

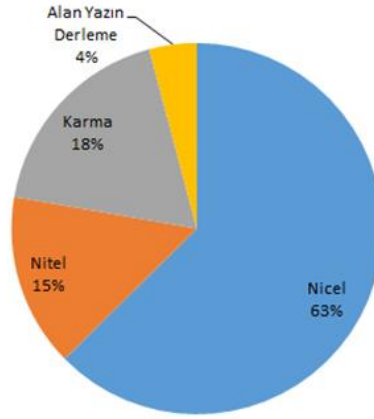
**Tablo 3. Makalelerin Türlerine Göre Dağılımı**

ARAŞTIRMANIN TÜRÜ	N	%
Alanyazın Taraması	11	5,98
Yöntem Çalışması	24	13,04
Kuramsal Çalışma	4	2,17
Deneysel/ Uygulamalı Çalışma	79	42,93
Eylem Araştırması	18	9,78
Betimsel Çalışma	19	10,33
Değerlendirme Çalışması	26	14,13
Mesleki Çalışma	3	1,63

Araştırılan alan ile ilgili incelemesi yapılan yayınlanmış 184 makaleden 79 tanesinin (%42,9) deneysel veya uygulamalı çalışma ile gerçekleştirildiği görülmektedir. Bu çalışmaların çoğunlukla uygulamaya dayalı veri toplama/veri setleri oluşturma veya mevcut verilerin uygulamalı analizi şeklinde yürütüldüğü belirlenmiştir. 26 çalışma (%14,1) değerlendirme çalışması şeklinde yapılırken 24 çalışmanın (%13) ise yöntemsel çalışma şeklinde gerçekleştirildiği görülmektedir. Değerlendirme çalışması ve yöntemsel çalışma olarak yapılan araştırmalarda genel olarak örneklem sayısının 1000'den küçük olduğu görülmektedir.

### İncelenen Makalelerin Yöntemine Göre Dağılımı

Araştırmanın üçüncü alt amacında makaleler yöntemine göre sınıflandırılmıştır. Araştırmacıların beklentisine ve konunun özelliğine uygun olarak verilerin toplanması ve yorumlanması için çalışmalarda kullanılan yöntemler farklılık göstermektedir. Taranan 184 makale yöntemsel olarak nicel, nitel, karma ve alan yazın derleme ana başlıkları altında incelenmiştir. İncelenen makalelerin yöntemlerine göre dağılımı Şekil 4' de sunulmuştur.



**Şekil 4.** Makalelerin Yöntemsel Dağılımı

Yapılan incelemede makalelerin %63'ünde nicel araştırma yöntemlerinin kullanıldığı görülmektedir. Nicel araştırma yönteminin alt başlıkları bakımından çalışmalar değerlendirildiğinde deneysel araştırma yöntemlerinin 61 araştırmada deneysel olmayan araştırma yöntemlerinin ise 105 araştırmada tercih edildiği belirlenmiştir. Nitel araştırma yöntemleri araştırmalarda %15 oranında kullanılmış olup örnek olay yöntemi bu başlık altında en fazla tercih edilen yöntemdir. 11 araştırmanın (%4) alanyazın derleme şeklinde gerçekleştirildiği belirlenmiştir. 48 araştırma nitel ve nicel araştırma yöntemlerinin birlikte kullanıldığı karma araştırma yöntemi ile gerçekleştirmiş, bu araştırma yöntemi içerisinde açıklayıcı karma araştırma yönteminin en fazla tercih edilen yöntem olduğu görülmüştür.

#### Araştırmaların Örneklem Profiline Göre Dağılımı

Araştırmanın dördüncü alt amacında makaleler örneklem profiline göre sınıflandırılmıştır. İncelenen makalelerde farklı profillerden katılımcı grubu ile çalışılmıştır. Bazı makalelerde ise birden fazla katılımcı profili seçildiği görülmüştür. İncelenen makalelerin örneklem profiline göre dağılımı Tablo 4'de verilmiştir.

**Tablo 4-** Makalelerin Örneklem Profili Dağılımı

SEÇENEKLER	N	%
İlköğretim (1-5)	2	1,09
İlköğretim (6-8)	12	6,52
Ortaöğretim (9-12)	32	17,39
Lisans (Eğitim Fak.)	11	5,98
Lisans (Diğer)*	65	35,33
Lisansüstü (Master-Doktora)	12	6,52
Öğretmenler	4	2,17

Eğitim Yöneticileri	5	2,72
Diğer**	41	22,28

\* Mühendislik Fakültesi, Tıp Fakültesi, Fizik-Kimya-Biyoloji bölümü öğrencileri.

\*\* Yetişkin Eğitimi, Şirket Çalışan Eğitimleri, Sağlık Personeli Eğitimleri, Ulaştırma Çalışanları, İşverenler, ML ve AI alan uzmanları

Çalışmalarda örneklem seçiminde en fazla (%35,3) lisans (diğer) grubunda yer alan katılımcılar kullanılmış olup bu grup içerisinde mühendislik fakültesi, tıp fakültesi ve fen bilimleri alanından katılımcıların genel itibarı ile yoğun olarak tercih edildiği söylenebilir. Eğitim bilimleri alanında lisans öğrencileri ise 11 çalışmada örneklem olarak alınmışlardır. Diğer örneklem profillerini kullanan 41 çalışma bulunup bu çalışmaların yoğun olarak yetişkin eğitimi ve çalışan eğitimi grupları ile ilgili araştırmalar yürüttükleri görülmektedir. 32 çalışmada ortaöğretim grubu öğrencileri üzerinde araştırma yapıldığı belirlenmiş olup bu çalışmaların ekseriyetle eğitim analitiği üzerine yürütüldüğü görülmektedir.

### Örneklem Sayısına Göre Dağılım

Araştırmanın beşinci alt amacında makaleler kullanılan örneklem sayısına göre sınıflandırılmıştır. Makalelere konu olan araştırmalarda kullanılan örneklem sayıları üzerine yapılan incelemelerde elde edilen sonuçlar Tablo 5' verilmiştir.

**Tablo 5** Makalelerdeki Örneklem Sayısına Göre Dağılım

ARALIK	N	%
1 - 30 arası	24	13,48
31 – 50 arası	10	5,62
51 – 100 arası	12	6,74
101 – 1000 arası	53	29,78
1001 – 10.000 arası	31	17,42
10.000 – 100.000 arası	14	7,87
100.000 – 1.000.000 arası	11	6,18
1.000.000 üzeri	9	5,06
Belirtilmemiş	14	7,87

Elde edilen bulgulara göre çalışmaların % 29,7'sinde (N=57) 101-1.000 arası örneklem ile çalışmıştır. 24 araştırma 1-30 arası örneklem düzeyinde gerçekleştirilmiş olup bu çalışmaların sıklıkla lisans düzeyindeki katılımcılarla deneysel/uygulamalı çalışma şeklinde yürütüldüğü görülmektedir. 10.000-100.000 arası veri kümeleri ile 14 araştırma yürütülürken, 100.000-1.000.000 arası veri kümeleri ile 11 çalışma yürütülmüş olup bu

çalışmaların genel olarak çevrim içi ortamlardan elde edilen (Mooc, Moodle vb.) veriler ile gerçekleştirildiği görülmektedir. 9 çalışma büyük veri (big data) kümeleri ile (1.000.000 üzeri) gerçekleştirilmiş, bu araştırmalarda genel itibari ile ulusal ve uluslar arası kuruluşların sağladığı veri setlerinin kullanıldığı belirlenmiştir

### Araştırmaların Üzerinde Çalışılan Öğretim Alanına (Disipline) Göre Dağılımları

Araştırmanın altıncı alt amacında makaleler çalışılan öğretim alanına (disipline) göre sınıflandırılmıştır. İncelenen makalelerde üzerinde çalışılan öğretim alanları (disiplini) değerlendirilerek elde edilen bulgular Tablo 6' de verilmiştir. Değerlendirme sonucunda araştırmaların büyük bir bölümünün tek bir disiplin üzerinde çalıştığı görülmekle birlikte, birden fazla disiplin ile aynı anda çalışan araştırmalarında bulunmaktadır.

**Tablo 6-** Makalelerdeki Öğretim Alanlarının Dağılımı

ÖĞRENME ALANI	N	%
Fen Bilimleri	18	9,18
Matematik	17	8,67
Yabancı Dil Öğretimi	10	5,10
Dil Eğitimi	16	8,16
Bilgisayar/Bilişim Tek.	28	14,29
Eğitim/Öğretim	38	19,39
Mühendislik	22	11,22
Sağlık	26	13,27
Diğer*	21	10,71

\* Müzik, Görselleştirme, Ulaştırma, İletişim, Sanat, Beden Eğitimi, İstatistik.

İncelenen araştırmaların %19,39'u (N=38) eğitim/öğretim alanında (disiplininde) gerçekleştirilmiş olup bu alandaki çalışmaların Lisans profil grubu üzerinde yoğunlaştığı görülmektedir. Ayrıca bu çalışmalarda genel olarak devam/devamsızlık, başarısızlık nedenleri, akademik başarıyı artırmada yöntem ve metotlar gibi başlıklarda eğitim analitiği araştırmaları yapıldığı görülmektedir. Çalışmaların %14,29'u bilgisayar ve bilişim teknolojileri üzerine gerçekleştirilmiştir. Öğretim alanı (disiplini) olarak tıp ve sağlık hizmetleri üzerine 26 araştırma yürütülmüş, 22 araştırma ise mühendislik alanları üzerine gerçekleştirildiği belirlenmiştir.

### İncelenen Makalelerde Kullanılan Veri Toplama Araçlarının Dağılımı

Araştırmanın yedinci alt amacında makaleler kullanılan veri toplama araçlarına göre sınıflandırılmıştır. Makalelerde kullanılan veri toplama araçları derlenerek Tablo 7' de verilmiştir. Veri tabanları (Databases) ve log kayıtları (Log Data) üzerinden veri toplayarak yapılan araştırmalarda veri setleri genel olarak tek veri toplama aracı olarak kullanılmıştır.



**Tablo 7- Makalelerdeki Veri Toplama Araçlarının Dağılımı**

VERİ TOPLAMA ARAÇLARI	N	%
Gözlem / Gözlem Formu	24	12,37
Görüşme (Görüşme/Odak Grup Görüşmesi)	6	3,09
Anketler	16	8,25
Sınav/Test/ödev/kurs - Sonuçları	38	19,59
Mevcut Çalışmalar/Alan yazın	12	6,19
Ölçekler (Etkileşim/Nitelik/Değerlendirme/Tutum Vb)	7	3,61
Veri Tabanları/Log Kayıtları	66	34,02
Görüş Formu	5	2,58
Sensörler/Dijital Cihazlar	20	10,31

Araştırmalarda veri toplama aracı olarak %34 oranında veri tabanları ve log kayıtları kullanılmıştır. Bu iki veri toplama aracını kullanan çalışmaların öğretim alanı olarak eğitim/öğretim grubu içerisindeki araştırmalara yöneldikleri görülmektedir. Yine bu iki veri toplama aracı ile çalışılan araştırmalarda görüntü işleme ve doğal dil işleme çalışmalarının önemli ölçüde yer tuttuğu belirlenmiştir. Sınav-test sonuçları, ödev-görev sonuçları, kurs değerlendirmeleri gibi veri toplama araçları kullanan araştırmaların genel dağılımdaki oranı ise % 19,59'dur. 16 araştırmada anketler ile veri toplama yöntemi seçildiği ve bu çalışmaların genellikle sınırlı örneklem kullanılan (1-1000) ve ağırlıklı olarak deneysel olmayan nicel araştırmalar olduğu görülmektedir.

### TARTIŞMA VE SONUÇ

Bu çalışma kapsamında makine öğrenmesi ile ilgili Web of Science kategorilerinden “Educational Research” veya “Educational Scientific Dicipines” kategorilerinde bulunan 2015 yılından 2020 yılına kadar olan 5 yıllık süreç içinde yazılmış 184 makale incelenmiştir. Web of science veri tabanından sorgulama 24 Ekim 2020 tarihinde yapılmıştır. Makaleler; kullanılan algoritmalar, kullanılan yöntem, örneklem profiline, örneklem sayısına, üzerinde çalışılan öğretim alanına, kullanılan veri toplama araçlarına göre dağılımı özellikleri bakımından analiz edilmiştir.

Karar ağaçları, tesadüfi orman, gradyen güçlendirme (gradient boosting) gibi yöntemler, her türlü veri bilimi problemlerinde yaygın bir şekilde kullanılmaktadırlar. Tarama yapılan makaleler kullanılan algoritmalar bakımından analiz edildiğinden karar ağacı algoritmasının en çok kullanılan algoritma olduğu görülmüştür. Karar ağacı, belli kurallar çerçevesinde veriyi daha küçük kümelere bölmek için kullanılan yapıdır (Albayrak ve Yılmaz, 2009). Ağaç tabanlı öğrenme algoritmaları, en çok kullanılan gözetimli öğrenme algoritmalarındandır. Genel itibariyle ele alınan bütün problemlerin çözümüne uyarlanabildiği için tercih edilirler. Çalışmaların çoğunluğunun deneysel veya uygulamalı çalışma ile gerçekleştirildiği görülmektedir. Deneysel araştırmalar, güçlü

kanıtlar elde edilen, neden sonuç ilişkilerinin en iyi şekilde değerlendirildiği araştırmalardır (Kuzu Kurban, 2017). Bu güçlü yönleri bakımından deneysel araştırmaların bu konuda tercih edildiği söylenebilir. Araştırma yöntemleri toplanacak verinin niteliğine göre farklılık gösterir. Araştırmanın içeriğine göre yöntemin dezavantajları, avantajları, sınırlılıkları göz önünde bulundurularak yöntem seçimi yapılır. Ülkemiz eğitim bilim araştırmalarında sıklıkla nicel araştırma yöntemlerinin tercih edildiği görülmektedir (Türnüklü, 2001). Bu çalışma kapsamında incelenen makalelerde de çoğunlukla nicel araştırma yöntemlerinin kullanıldığı görülmektedir. Nitel araştırmalardan elde edilen verilerin daha derinlemesine olması nedeniyle kuram geliştirmeye yönelik araştırmaların daha az gerçekleştiği söylenebilir. Nitel araştırmaların zaman bakımından ekonomik olmaması ve yaygın kullanımı olmamasından dolayı daha az gerçekleştirildiği düşünülmektedir (Arık ve Türkmen, 2009). İncelenen makalelerde farklı profillerden katılımcı grubu ile çalışılmıştır. Bazı makalelerde ise birden fazla katılımcı profili seçildiği görülmüştür. Örneklem seçiminde en fazla lisans grubunda yer alan katılımcılar kullanılmış olup bu grup içerisinde mühendislik fakültesi, tıp fakültesi ve fen bilimleri alanından katılımcıların genel itibari ile yoğun olarak tercih edildiği söylenebilir. Eğitim teknolojisi araştırmalarında genellikle lisans düzeyinde eğitim gören katılımcıların örneklem olarak seçildiği görülmektedir (Çakmak vd., 2016). Bu bilgi ile bizim bulgularımız uyuşmaktadır. Araştırmacılar kolay ulaşılabilen, hali hazırda olan grup olduğu için lisans öğrencilerinden örneklem grubu oluşturmuş olabilirler. Örneklem seçiminde örneklem yanlılığı hatası yapılmış olabileceği söylenebilir. Genellikle, yapılan yanlılık hatası ‘orada bulunma’ ile oluşturulan örneklem grubunun kullanılmasıdır (Özen ve Gül, 2007). Bu seçimin zaman ve emek açısından araştırmacı için ekonomik olduğundan tercih edildiği söylenebilir (Ozan ve Köse, 2014). Okul öncesi, ilköğretim ve ortaöğretim öğrencilerin örneklem olarak tercih edilmeme nedenleri, gereken izinlerin prosedürlerinin zorluğu ve zaman kaybı olduğu düşünülmektedir (Arık ve Türkmen, 2009). İncelenen makalelerde örneklem sayısı genellikle 101 – 1000 aralığındadır. Seçer ve diğerleri (2014) yaptığı içerik analizinde incelenen makalelerin çoğunluğunun nicel çalışma olduğu ve örneklem büyüklüğünün de nicel araştırmalar için 301-1000 arasında olmasının uygun olduğu belirtilmiştir. İncelenen araştırmaların çoğunlukla eğitim/öğretim alanında olduğu görülmektedir. Eğitim öğretimin kalitesini arttırmak için sadece bilgisayara dayalı bir sistem değil akıllı sistemlerin tercih edilmesi yararlı olacaktır (İşler ve Kılıç,2021).Eğitim/öğretim alanında bilgisayar teknolojisi ile birlikte yapay zeka teknolojisinin kullanılması eğitim kalitesini arttırabileceği için çalışmaların bu alanda yoğunlaştığı söylenebilir.

Bu araştırmada 5 yıllık süreç içinde yayımlanan belirlediğimiz kriterlere uyan makalelerin içerikleri analiz edilmiştir. Elde edilen bulgular yorumlanmış ancak karşılaştırma yapılmamıştır. Farklı yıl aralıklarıyla yapılacak bir çalışma ile karşılaştırma yapılarak daha fazla bilgi sağlanabilir. Çalışma gruplarının çoğunluğunun lisans öğrencisi olduğu görülmektedir. Okul önesi ve ilkokul öğrencileri ile yapılacak çalışmalar alanyazına katkı sağlayacaktır. EBA(Eğitim Bilişim Ağı) gibi ortamlarda kullanıcıların kullanım süreleri ve kullanım şekilleri verilerinden kullanıcıların kaliteli ve verimli süre geçirmesi için stratejiler geliştirilebilir. Bunu sağlamak için yapay zekanın raporlama imkanından faydalanılabilir. Değerlendirmeler, öğrenciler ve öğreticiler açısından yapay zeka ile farklı bir boyut kazanabilir. Dönütler, yapay zeka teknolojisi ile ihtiyaçlar doğrultusunda otomatik olarak yapılarak bireyselleştirilmiş öğrenme daha etkin bir şekilde sağlanabilir.

## KAYNAKÇA

- Akay, E. Ç. (2018). Ekonometride Yeni Bir Ufuk: Büyük Veri ve Makine Öğrenmesi. *Sosyal Bilimler Araştırma Dergisi*, 7(2), 41-53.
- Akgün, E. (2019). *2023 Eğitim Vizyonunda Dijital Dönüşüm*. Seta - Perspektif (C. 233)
- Albayrak, A. S., & Yılmaz, S. K. (2009). Veri Madenciliği: Karar Ağacı Algoritmaları ve İMKB Verileri Üzerine Bir Uygulama. *Suleyman Demirel University Journal of Faculty of Economics & Administrative Sciences*, 14(1).
- Arık, R. S. ve Türkmen, M. (2009). Eğitim Bilimleri Alanında Yayımlanan Bilimsel Dergilerde Yer Alan Makalelerin İncelenmesi. I. Uluslararası Türkiye Eğitim Araştırmaları Kongresi, Çanakkale Onsekiz Mart Üniversitesi, Antalya.
- Atalay, M., & Çelik, E. (2017). Büyük Veri Analizinde Yapay Zekâ Ve Makine Öğrenmesi Uygulamaları-Artificial Intelligence and Machine Learning Applications in Big Data Analysis. *Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9(22), 155-172.
- Aydın, U.(2012). Sorunsaldan Sonuçlara Bilimsel Araştırma Süreci: Bir Araştırma Raporu Modeli Örneği. *ODÜ Sosyal Bilimler Araştırmaları Dergisi (ODÜSOBİAD)*, 3(5), 135-161.
- Çakmak, E. K., Özüdoğru, G., Bozkurt, Ş. B., Ülker, Ü., Ünsal, N. Ö., Boz, K., ... & Bahadır, H. (2016). 2014 Yılında eğitim teknolojileri alanındaki yayımlanan makalelerin incelenmesi. *Eğitim Teknolojisi Kuram ve Uygulama*, 6(1), 80-108.
- Özlem, A. V. C. I. (2015). Dijital Yaşamın Dijital Özne (l) leri: Herkes ya da Hiç Kimse. *Uşak Üniversitesi Sosyal Bilimler Dergisi*, 8(1), 249-266.
- Aydoğan, M., & Karcı, A. (2018). Meslek Yüksekokulu Öğrencilerinin Başarı Performanslarının Makine Öğrenmesi Yöntemleri ile Analizi. *International Symposium on Multidisciplinary Studies and Innovative Technologies*. October, Turkey.
- Çelen, F. K., & Seferoğlu, S. S., (2020). Yeni nesil öğrenme kültürü: çevrimiçi öğrenme toplulukları üzerine bir inceleme. *Gençlik ve dijital çağ*, 80. Hacettepe: Ankara.
- Çifçi, F., Kaleli, C., & Günal, S. (2018). Öznitelik seçme ve makine öğrenmesi yöntemleriyle eğitim performansının tahmin edilmesi. *Anadolu Üniversitesi Eğitim Bilimleri Enstitüsü Dergisi*, 8(2), 419-440.
- Dede, Chris. (2011). Emerging technologies, ubiquitous learning, and educational transformation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6964 LNCS, 1–8. doi:10.1007/978-3-642-23985-4\_1
- Dereli, Z. (2019). *Dijital Yerliler Yeni Nesil Eğitim Modeli*. İstanbul: Hümanist Ajans A.Ş. El Naqa, I., & Murphy, M. J. (2015). What is machine learning?. In *Machine Learning in Radiation Oncology* (pp. 3-11). Springer, Cham.
- Er, M. B. (2019). *Makine öğrenmesi ile Türk müziğinde duygu analizi* (Master's thesis, Maltepe Üniversitesi, Fen Bilimleri Enstitüsü).
- Filiz, E. (2019). Makine Öğrenmesi Yöntemleri ve Eğitim Verisi Üzerine Bir Uygulama: Uluslararası Matematik ve Fen Eğilimleri Araştırması 2015 Türkiye Örneği. *Yayınlanmamış Doktora Tezi. Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü: İstanbul*.
- Gök, M. (2017). Makine Öğrenmesi Yöntemleri İle Akademik Başarının Tahmin Edilmesi. *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, 5(3), 139-148.
- Gülcüoğlu, E., Özeren, E., Tosunoğlu, E. (2020). 2010-2019 Yılları Arasında Uyarlanabilir Öğrenme ile İlgili Yapılmış Yüksek Lisans Tezlerindeki Yöntemsel Eğilimler Üzerine Bir İnceleme. *Öğretim Teknolojisi ve Hayat Boyu Öğrenme Dergisi*, 1 (1) , 107-121.
- Güner, N., & Çomak, E. (2011). Mühendislik Öğrencilerinin Matematik I Derslerindeki Başarısının Destek Vektör Makineleri Kullanılarak Tahmin Edilmesi. *Pamukkale University Journal of Engineering Sciences*, 17(2).

- İnik, Ö. & Ülker, E. (2017). Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri. *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, 6(3), 85-104.
- İşler, B., & Kılıç, M.(2021). Eğitimde Yapay Zekâ Kullanımı Ve Gelişimi. *Yeni Medya Elektronik Dergisi*, 5(1), 1-11.
- Karakoyun, M., & Hacıbeyoğlu, M. (2014). Biyomedikal Veri Kümeleri İle Makine Öğrenmesi Sınıflandırma Algoritmalarının İstatistiksel Olarak Karşılaştırılması. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, 16(48), 30-42.
- Kardaş, K., & GÜVENİR, H. A. (2020). Kısa Sınavların, Ödevlerin ve Projelerin Dönem Sonu Sınavına Olan Etkilerinin Farklı Makine Öğrenmesi Teknikleri ile Araştırılması. *EMO Bilimsel Dergi*, 10(1), 21-27.
- Kuzu Kurban N.(2017). I. Ulusal Hemşirelikte Araştırma Kongresi, Hemşirelikte Araştırma Geliştirme Derneği, Ankara.
- Machine learning mastery (2020):<https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/> (Erişim:02.06.2021) Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Los Angeles: Sage.
- Ozan, C., & Köse, E. (2014). Eğitim programları ve öğretim alanındaki araştırma eğilimleri. *Sakarya University Journal of Education*, 4(1), 116-136.
- Parlak, B. (2017). Dijital çağda eğitim: Olanaklar ve uygulamalar üzerine bir analiz. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 22(15), 1741-1759.
- Peña-Ayala, A. (Ed.). (2013). *Educational data mining: applications and trends* (Vol. 524). Springer.
- Rashidi, H. H., Tran, N. K., Betts, E. V., Howell, L. P., & Green, R. (2019). Artificial intelligence and machine learning in pathology: the present landscape of supervised methods. *Academic pathology*, 6, 2374289519873088.
- Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert systems with applications*, 33(1), 135-146.
- Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6), 601-618.
- Sağlam, Z., Pekiş, M., Yılmaz, R. (2020). PISA 2018 Araştırmasına Etki Eden Duygusal Faktörlerin Veri Madenciliği Yöntemleri İle İncelenmesi. *Bilgi ve İletişim Teknolojileri Dergisi*, 2 (2) , 113-148.
- SARSICI, E., & ÇELİK, A. İ. (2019). Eğitimde Dijital Dönüşüm İçin Bir Model Önerisi. *Uluslararası Eğitimde Ve Sosyal Bilimlerde Yenilikler'sanal Sempozyumu*, 339.
- Savaş, S., Topaloğlu, N., & Yılmaz, M. (2012). Veri madenciliği ve Türkiye'deki uygulama örnekleri. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 21, s. 1-23
- Seçer, İ., İsmail, A. Y., Ceyhan, O. Z. A. N., & YILMAZ, B. (2014). Rehberlik ve psikolojik danışma alanındaki araştırma eğilimleri: Bir içerik analizi. *Türk Psikolojik Danışma ve Rehberlik Dergisi*, 5(41), 49-60.
- Suthaharan, S. (2014). Big data classification: Problems and challenges in network intrusion prediction with machine learning. *ACM SIGMETRICS Performance Evaluation Review*, 41(4), 70-73.
- Tataroğlu, V. (2019). *Derin öğrenmeye dayalı sosyal medya profillemesi* (Master's thesis, Pamukkale Üniversitesi Fen Bilimleri Enstitüsü).
- Tavşancıl, E. ve Aslan, E. (2001). Sözel, Yazılı ve Diğer Materyaller için İçerik Analizi ve Uygulama Örnekleri. Epsilon Yayınevi, İstanbul. TÜREL, Y. K., & Engin, B. A. Z. (2016). Eğitsel Veri Madenciliği Üzerine Bir Araştırma. In *President Of The Symposium* (p. 54).
- Türnüklü, A. (2001). Eğitim bilim alanında aynı araştırma sorusunu yanıtlamak için farklı araştırma tekniklerinin birlikte kullanılması. *Eğitim ve Bilim*, 26(120).
- Usluel, Y. K., Avcı, Ü., Kurtoğlu, M., & Uslu, N. (2013). Yeniliklerin benimsenmesi sürecinde rol oynayan değişkenlerin betimsel tarama yöntemiyle incelenmesi. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 33(1), 53-71.

Ülker, M., & Civalek, O. (2002). Yapay Sinir Ağları ile Eksenel Yüklü Kolonların Burkulma Analizi. *Turkish J. Eng. Env. Sci*, 26, 117-125.

Yıldız, O. (2014). Makine Öğrenmesi İle Uzaktan Eğitim Öğrencilerinin Performanslarının Değerlendirilmesi. *Yayınlanmamış Doktora Tezi. İstanbul Üniversitesi Fen Bilimleri Enstitüsü.*

### EXTENDED ABSTRACT

The fact that technology affects human life more and more every day shows its presence in education. In particular, within the scope of the Covid-19 pandemic process, providing education entirely with technological devices has made technology an indispensable element of education. Along with the concept of 21st century skills, there is a need for a new understanding of education in which learning will take place not only in the classroom but everywhere throughout life, where the teacher is not the only source of information. Students no longer tend to learn the information they want, rather than the information presented to them. In this context, countries should organize their education systems in line with the needs of the age, and individual differences should be taken into account while organizing educational materials and environments. In a period when billions of people actively use social media, it is a matter of curiosity how the digitalization process of education and training institutions will be. Increasing digital transformation in education will reduce students' technology addiction and increase their interest in lessons..

Data obtained from digital platforms are used for statistical information extraction. The data held on educational platforms are also used in educational data mining applications. In systems developed with machine learning algorithms, students will be able to take the course content with the methods they prefer in line with their individual characteristics, progress according to their learning speed, and see and complete the missing points thanks to the feedback given. Within the scope of the education system, student-oriented features such as the student's attendance, what studies he/she does, how his/her progress is, individual characteristics, assessment and evaluation results can be followed. It will be possible to see the work done by the teachers in the system, how they use the system, and the feedback they give to the students.

Data mining methods used in education are described as educational data mining. Countries make arrangements in their education systems to achieve the objectives of education. With the digital transformation, it is very important to use the increasing data volume in the field of education for the development of education. It is important to analyze educational data and to be trained with the correct machine learning algorithms in order to interpret the available data correctly. With this study, case studies published in prestigious journals were examined by scanning them from the web of science database. The aim of the study is to examine the articles and make guiding inferences for future research. In this study, it is aimed to make a systematic analysis of the articles written by using machine learning method in the field of education.

For this purpose, answers to the following questions are sought:

1. Which algorithms were used in the articles written using machine learning methods in the field of education?
2. What is the weight of the articles in terms of type?
3. What is the weight of the research methods used in the articles?
4. What are the sample levels used in the articles?
5. What is the weight of the sample numbers used in the articles?
6. In which fields of education were the articles written intensively?

. In this study, content analysis method was used in the analysis of the articles. The studies examined in the research are articles in the Educational Research or Educational Scientific Dicipines categories of the Web of Science categories that have "machine learning" in their subject listed in the Web of Science database. The screening covers the studies carried out in the 5-year period from 2015 to 2020. As a result of the query made according to these criteria, 201 articles were found. As a result of the inquiry, 184 articles that comply with the

criteria were examined and the frequency and percentages were calculated according to the data in the form created. The findings obtained are presented using tables.

Methods such as decision trees, random forest, gradient boosting are widely used in all kinds of data science problems. Since the searched articles were analyzed in terms of the algorithms used, it was seen that the decision tree algorithm was the most used algorithm. It is seen that quantitative research methods are mostly used in the articles examined within the scope of this study. It can be said that, because the data obtained from qualitative researches are more in-depth, the studies aimed at developing theories are less. Participants in the undergraduate group were used the most in the selection of the sample, and it can be said that the participants from the engineering faculty, medical faculty and science fields were generally preferred in this group. In educational technology research, it is seen that the participants who are educated at the undergraduate level are generally selected as a sample. This information matches our findings. Since the researchers are an easily accessible and readily available group, they may have formed a sample group of undergraduate students. It can be said that sampling bias error may have been made in the sample selection. Generally, the bias error made is the use of the sample group created by 'being there'. It can be said that this choice is preferred because it is economical for the researcher in terms of time and effort.

In the articles examined, it was worked with a group of participants from different profiles. In some articles, it was seen that more than one participant profile was selected. Participants in the undergraduate group were used the most in the selection of the sample, and it can be said that the participants from the engineering faculty, medical faculty and science fields were generally preferred in this group.

In this research, the contents of the articles published in a 5-year period that meet the criteria we have determined were analyzed. Findings obtained were interpreted, but no comparison was made. More information can be provided by making a comparison with a study to be conducted with different year intervals.