

**Makine Öğrenimi Araştırması**

Muhammad Rauf

Karabük Üniversitesi, TOBB Meslek Yüksekokulu, Bilgisayar Programcılığı Bölümü, Karabük

[2005113531@ogrenci.karabuk.edu.tr](mailto:2005113531@ogrenci.karabuk.edu.tr)

Özet: Makine öğrenimi, bilgisayarların daha akıllı hale gelmesi ve sahip oldukları verilerin deneyiminden öğrenmesi için insanlar tarafından tekrar tekrar programlanmaya gerek kalmadan bir şeye karar vermeyi kendi başına öğrenebilen bir sistemdir. Öğrenme tekniğine dayalı olarak, denetimli öğrenme, etiketlenmiş bir veri kümesi (eğitim verileri) kullanılarak ayırt edilebilirken, denetimsiz öğrenme, veri kümesine dayalı sonuçlar çıkarır. Veri kümesi biçimindeki girdi, doğru analizi üretmek için makine öğrenimi tarafından kullanılır. Çözüm, makine öğrenimi oluşturmak için kullanılan algoritmaları ve kitaplıkları sağlayan Python'u kullanmaktır. Yapay zeka (AI) ve makine öğrenimi, onlarca yıllık iniş ve çıkışlardan sonra şu anda yeniden patlama evresine giriyor. Yapay Zeka, uygulamasının Facebook, Twitter, Google, Amazon gibi günümüzün iş ve sosyal medya uygulamalarında ve hatta Hepsiburada, Yemek sepeti ve benzeri gibi Türkiye’den büyük uygulamalarda kitlesel olarak gerçekleştirildiği yerlerde tekrar sevilir. Bu kitaptaki tartışmanın yapısı 3 ana bölümden oluşmaktadır: (1) Makine Kavramları Öğrenme ve Yapay Zeka (2) Makine Öğrenimi için Python Programlamanın Temelleri ve (3) Denetimli ve Denetimsiz Öğrenme olmak üzere çeşitli algoritmalar uygulayarak Python Kullanarak Makine Öğrenimi Uygulama Örnekleri. Algoritmaların anlaşılmasından, veri kümelerinin işlenmesinden eğitim ve teste kadar çeşitli vaka çalışmaları ayrıntılı olarak tartışılır ve ayrıca geliştirilen makine öğrenimi modelinin sonuçlarını görselleştirir.

Anahtar Sözcükler: Makine öğrenimi · Algoritma · Yapay zeka · Veri tabanı

Abstract: Machine learning is a system that is able to learn on its own to decide something without having to be repeatedly programmed by humans so that computers become more intelligent and learn from the experience of the data they have. Based on the learning technique, supervised learning can be distinguished using a dataset (training data) that has been labeled, while unsupervised learning draws conclusions based on the dataset. The input in the form of a dataset is used by machine learning to produce the correct analysis. The solution is to use Python which provides the algorithms and libraries used to create machine learning. Artificial intelligence (AI) and machine learning are currently re-entering a booming phase after decades of ups and downs. Artificial Intelligence is again loved, where its implementation is carried out massively in today's business and social media applications such as Facebook, Twitter, Google, Amazon, and even large applications from Turkey such as Hepsiburada, Yemek Sepeti, and so on. The structure of the discussion in this book includes 3 major parts, namely (1) Machine Concepts Learning and Artificial Intelligence (2) Fundamentals of Python Programming for Machine Learning and (3) Examples of Implementing Machine Learning Using Python by applying several algorithms, both Supervised and Unsupervised Learning. Several case studies are discussed in detail from understanding algorithms, processing datasets to training and testing as well as visualizing the results of the developed machine learning model.

1. Giriş

İnsan hayatı boyunca öğrenmeyi asla bırakmaz. Bu, biz farkında olmadan ve doğal olarak gerçekleşir. Ancak derinleştirmede bir makine elbette pratik bir şey olabilir. Bir kişi günlük deneyimler yoluyla öğrenir. Bu deneyimden, kişi bilgi kazanacaktır.

Bilgi edinmenin çeşitli yolları vardır. En basit yol, hesaplanan haberleri ezberlemeyi veya kaydetmeyi öğrenmektir. Başka bir yol, uzman olan diğer insanlardan bilgi alarak zaten var. Bir kişi, yaptığı sorunları çözme deneyimi yoluyla da öğrenebilir. Bir ikilemin üstesinden başarıyla geldikten sonra, kişi yapıyı ve sorunun nasıl çözüleceğini hatırlayacaktır. İnsanlar neredeyse benzer bir ikilem yaşarlarsa, birileri sorunu daha verimli bir şekilde çözebilir.

Günlük yaşamda, makine öğreniminin günlük yaşamda faydalı olduğunu görebiliriz. Çok somut görünen bazı modeller ses tanıma, parmak izi tanıma veya el yazısı tanımadır. Günümüzde makine öğreniminin gelişimi çok hızlıdır. Daha akıllı makineler yapmak için çeşitli çalışmalar yapılıyor. Gördüğümüz en son eskrim modeli, sahibini tanıyabilen ve duygularını tanıyabilen Honda tarafından yapılmış akıllı bir robot olan ASIMO. Asimo kadar akıllı bir robot üretmek için, etkileşimde bulunabilmede ses tanıma, sahibinin yüzünü tanıyabilme ve odaları tanıyabilmede görüntü tanıma ve daha pek çok yapay zeka tekniği uygulanıyor. popüler bir bilim dalı.

Böylece Asimo kadar akıllı bir robot üretebilir, ardından etkileşim kurabilmek için ses tanıma, sahibinin yüzünü tanıyabilmek ve odaları tanıyabilmek için görüntü tanıma gibi çeşitli yapay zeka teknikleri uygulanır. artık medyada popüler olan bir bilim dalı. Yapay zekanın (yapay zeka) bir çıkış dalı olarak yaratılan, hemen hemen herkes makine öğrenimi teknikleri kullanılarak oluşturulan kişisel bilgisayar sistemleriyle etkileşime girmiş, bunları kullanmış veya duymuştur. Facebook'ta otomatik hashtag'lerin fotoğraflarını görmekten, Google'da arama önerilerini kullanmaktan, çevrimiçi alışverişte homojen ürün önerilerine tıklamaktan, spam olmadan e-posta hizmetlerinden yararlanmaya, AlphaGo'nun Go oyununda en iyi profesyonel oyuncuları yendiği haberlerini duymaya kadar.

2. Yöntemler  
  
**A. MAKİNE ÖĞRENİMİ UYGULAMA YÖNTEMLERİ**

Makine öğrenmesi (Machine Learning) uygulamasındaki bazı yöntemler ise şu şekildedir :

1. Sorununuzu Tanımlayın

Makine öğrenimi sürecindeki ilk aşama, programdaki sorunları tespit edip sıralayabilmektir. Veri toplama da dahil olmak üzere bu Makine Öğrenimi süreciyle ulaşılması gereken sorunlar veya hedefler nelerdir? Çünkü veri olmadan, Makine Öğrenmesi ile problem çözülmeyecektir. Veriler Excel, MS Access veya diğer dosyalar biçiminde olabilir. Bu adım, gelecekteki öğrenme için temel formdur. Mevcut sorunlardan, bunları aşağıdakiler gibi birkaç soruya ayırabilirsiniz :

* Amacınız ne? Neyi tahmin etmek istiyorsun?

• Hangi özellikleri (Ağırlıklar) kullanacaksınız?

• Ne tür veriler girmek istiyorsunuz? Veriler mevcut mu?

* Ne tür sorunlarla karşı karşıyayız? İkili Sınıflandırma? Yoksa Kümeleme mi?
* Ne tür iyileştirmeler bekleniyor?
* Verilerdeki Özelliği (ağırlığı) nasıl ölçersiniz?

2.Verilerin Toplanması ve Hazırlanması

Makine Öğrenimi sürecindeki bir sonraki yöntem, veri toplamak ve hazırlamaktır. Ne kadar çok ve kaliteli veriye sahip olursanız, Makine Öğreniminin performansı o kadar iyi olur. Web kazıma ve Veri Madenciliği gibi veri toplama yöntemlerine örnekler. Ancak genel olarak aşamalar aşağıdaki gibi tanımlanabilir. İlk olarak, verilerin sıralanması aşamasında "öğretilen" kullanılacak veriler, eğitim verileri (eğitilecek veriler), doğrulama verileri (doğrulama için kullanılan veriler) ve test verileri (tahmin için kullanılan veriler) olmak üzere üç kısma ayrılır. deneyler). Bu aynı zamanda Makine Öğreniminin verilerin varlığına çok bağımlı bir şey olduğunu da doğrular. Bu verilerden bir sonraki aşamada hangi algoritmayı kullanacağımıza karar verebiliriz.

3. Çözüm İçin Seçim Algoritması

Bir sonraki yöntem, bir veri modeli oluşturmaktır. Bu aşama, uygun algoritmanın seçilmesini ve bir model şeklinde veri gösterimini içerir. Hazırlanan veriler eğitim ve test verileri olmak üzere iki kısma ayrılmıştır. Eğitim verileri, model geliştirme için kullanılır ve test verileri referans olarak kullanılır. Bundan sonra, model doğrulama, yani veri modelinin veri doğrulama ile test edilmesi gerçekleştirilir. Bu, kullanılan girdi, süreç ve çıktıdan geri bildirim almak için kullanışlıdır.

İdeal olarak, en iyi modeli elde etmek için bu araçlarda bulunan tüm algoritmalar en iyi modeli elde etmeye çalışılmalıdır. Ancak bu, oldukça deneyimli bir Veri Bilimcisi için bile çoğu zaman mümkün değildir. Bunun nedeni, yetersiz zaman veya yetersiz algoritma bilgisi olabilir. Yeni başlayanlar için, birkaç model seçmek daha iyidir veya genellikle sadece bir tane, ancak basit ve anlaşılması kolay bir model.

4. Modelin Değerlendirilmesi ve İyileştirilmesi

Bir Makine Öğrenimi modeline sahip olduktan ve model test verilerini tahmin etmek için kullanıldıktan sonra, son adım modelin performansını analiz etmek ve modeli değerlendirmektir. Modelin doğruluğunu ve performansını test etmek için veri testi (test verileri) kullanılır. Bu, önceki aşamada seçilen bir algoritmanın doğruluğunu belirleyecektir.

3. Tartışma  
  
**A. MAKİNE ÖĞRENİMİNİN TANIMI**

Makine Öğrenimi veya Makine Öğrenimi, problem çözmeye yönelik faaliyetlerin yürütülmesinde insanların rolünün yerini almak için taklit etmek için kullanılan Yapay Zekadan (AI) bir yaklaşım tekniğidir. Kısacası makine öğrenimi, kullanıcılarından yönlendirme almadan öğrenip iş yapabilmesi için yapılmış bir makinedir. Bilgisayar oyunları ve yapay zeka alanında Amerikalı bir öncü olan Arthur Samuel'e göre, makine öğrenimi, bilgisayarların yeteneklerinin nasıl tanımlanacağını inceleyen bir bilim dalıdır.

açıkça programlanmadan öğrenmek. Bir makine öğrenimi, önceki verilerden elde edilen fikirlere dayanıyorsa ve kalıpları tanımlarsa ve minimum insan veya kullanıcı müdahalesi kullanarak kararlar alırsa bunu yapabilir. Dolayısıyla Makine Öğrenmesi yeteneği öğrendiği sürece sınırlı değildir.

İlk Makine Öğrenimi terimi bilindiğinden beri, birçok kişi bunu geliştirdi, en ünlülerinden biri, bilgisayar donanımı ve yazılımı üreten ve satan bir Amerikan şirketi olan IBM tarafından 1996 yılında oluşturulan Deep Blue'dur. Deep Blue'nun en ünlü kullanımı satranç oynama yeteneğidir. Deep Blue, profesyonel satranç şampiyonlarını yenerek çok popüler.

Bu arada, Tom M. Mitchell, "Ölçülebilir bir performans (performans) P ile atanan bir görevin (Görev) T deneyiminden (deneyimi) E öğrenen bir bilgisayar programı. Örneğin, satrancı kullanmayı öğrenen bir bilgisayar programı, kendisine karşı satranç oynamaktan edindiği deneyime dayalı olarak T satranç oynarken kazanma yeteneği olarak ölçülen P yeteneğini geliştirebilir.

**B. MAKİNE ÖĞRENİMİ TÜRLERİ**

Genel olarak, makine öğrenimi algoritmaları, denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenme olmak üzere üç türe ayrılır. Makine öğrenimini kullanarak sorunları çözmeden önce, bu algoritmaların kendi işlevleri ve hedefleri olduğu için üç tür makine öğrenimini anlamamız gerekir.

1. Denetimli Öğrenme

Denetimli öğrenme, temel olarak algoritmanın, belirli bir X'in y'nin en iyi tahminini yaptığı girişi en iyi tanımlayan işlevleri seçebilmek için eğitildiği bir işlev yaklaşımı kavramını benimser. Bununla birlikte, gerçekte, en uygun işlevi bulmakta güçlük çeken pek az insan yoktur. Bunun nedeni, gerçek algoritmanın kullanılan varsayımlara bağlı olmasıdır. Karşılanmayan varsayımlar varsa, veri işleme sonuçlarının yanlılığa neden olması nadir değildir. Bu nedenle, bu algoritma, sistemin örüntüyü ve regresyonunu, sınıflandırmasını, KNN, Naive Bayes, Karar Ağaçlarını, Lineer Regresyon, Destek Vektör Makinesini ve sinir ağını öğrenebilmesi için doğru eğitim verilerini gerektirir. Denetimli öğrenme algoritmaları, en yaygın makine öğrenimi türüdür. Türkçe’de denetimli öğrenme, denetimli öğrenme olarak tanımlanır. "Denetlenen" terimi, bu algoritmaların örnek olarak öğrenmek için tasarlanması nedeniyle ortaya çıktı. Denetimli öğrenmenin çeşitli avantajları ve dezavantajları vardır. İlk avantajı, denetimli öğrenme algoritmasının basit ve anlaşılması kolay bir süreç olmasıdır. Ayrıca bu algoritma sınıflandırma için de güçlüdür. Kullanılan veriler gerçek zamanlı veriler değildir, bu nedenle sonuçları tahmin etmek için yeni veriler gerektirir. Bu algoritmanın zayıf yönü, eğitim için uzun bir hesaplama süresi gerektirmesi ve her girişi etiketlemesi gerektiğinden denetimsiz öğrenme algoritmasından daha karmaşık bir algoritma kullanmasıdır.

2. Denetimsiz Öğrenme

Denetimsiz öğrenme algoritması, etiketli veri gerektirmeyen bir algoritmadır. Denetimsiz öğrenmede, algoritma eğitim verisine ihtiyaç duymaz. Bu algoritma, doğru modeli bulmak için algoritmanın temel aldığı kategoriler veya etiketli çıktılar gerektirmeyen örüntü tespiti ve tanımlayıcı modellemede kullanılır. Bu algoritma kümeleme ve birliktelik kuralları için kullanılır. Denetimsiz öğrenmenin avantajı, etiket gerektirmediği için algoritmanın daha önce bilinmeyen kalıpları aramak için daha esnek olmasıdır. Bu algoritmanın dezavantajı, etiket olmaması ve çıktı ile girdiyi karşılaştırmanın daha zor olması nedeniyle verilerde bilgi bulmanın zorluğudur. Denetimli ve denetimsiz öğrenmeyi anladıktan sonra, tüm algoritmaların denetimli veya denetimsiz öğrenme olarak kategorize edilemeyeceği ortaya çıkıyor. Bu algoritma yarı denetimli öğrenme olarak adlandırılabilir.

3. Yarı Denetimli Öğrenme ve Pekiştirmeli Öğrenme

Daha önce açıklandığı gibi, yarı denetimli öğrenme, denetimli veya denetimsiz öğrenme olarak kategorize edilemeyen bir algoritmadır. Bu algoritma, etiketli ve etiketsiz olarak ikiye ayrılan büyük miktardaki veriler için uygundur. Yarı denetimli öğrenmenin avantajı, verilerin yalnızca yarısının etiketlenmesi ve uzmanların bunları işlemesini gerektirmemesi nedeniyle daha düşük maliyetgerektirmesidir. Takviyeli öğrenmenin amacı, çıktıyı en üst düzeye çıkaracak ve riski en aza indirecek eylemlerde bulunmak için çevre ile etkileşimlerden toplanan gözlemleri kullanmaktır. Bu algoritma tekrar tekrar öğrenmeye devam edecektir. Bu algoritmada çevreleriyle etkileşimlerinden öğrenecek ajanlar vardır. Bir model oluşturmak için, pekiştirmeli öğrenme algoritması, aracının bir karar vermek için bir eylemde bulunmasının ardından girdi verilerini gözlemlemesi dahil olmak üzere birkaç aşamadan geçer. Karar verildikten sonra, temsilci çevreden bir "ödül" veya takviye alacaktır. Daha sonra girdiyi yeniden gözlemleyin ve karar verme süreci tekrar gerçekleştirilir, ancak alınan kararların sonuçlarının daha doğru olması için çevreden ek takviye ile. Makine öğrenimi ve veri bilimi birbirini tamamlayan iki bilimdir. Bu iki bilim birbirini tamamlar ve genellikle aynı anda uygulanır. Veri bilimi, günümüzde çokça çalışılan bir bilimdir. Bunun nedeni, veri miktarının artması ve veri bilimcisi mesleğinin yüksek talep görmesidir. Bu bilimlerin her ikisi de finans, eğitim, lojistik, iletişim ve hatta moda gibi çeşitli sektörlerde uygulanabilir. İyi haber şu ki, veri bilimi herhangi bir eğitim geçmişine sahip herkes tarafından öğrenilebilir.

**C. MAKİNE MODELLERİNDE PYTHON KULLANIMI ÖĞRENME (MAKİNE ÖĞRENİMİ)**  
Makine Öğreniminde model yapımında Python programlama dilinin kullanılmasının nedenleri şunlardır:

1. Öğrenmesi kolay

Python, kısa ve okunması kolay bir kod sunar. Makine Öğrenimi ve yapay zekanın arkasındaki gerçek güç karmaşık algoritmalar ve esnek iş akışları olsa da, Python'un sadeliği geliştiricilerin ve Veri Bilimcilerinin güvenebilecekleri sistemler yazmasına olanak tanır. Veri Bilimcisi, programlama dillerinin teknik nüanslarına odaklanmak yerine makine öğrenimi sorunlarını çözmeye odaklanabilir. Ayrıca Python, öğrenmesi kolay olduğu için birçokları için çekicidir. Python kodunun insanlar tarafından anlaşılması kolaydır, bu da makine öğrenimi için model oluşturmayı çok kolaylaştırır.

2. Yüksek ve Çeşitli Kütüphane Çeşitleri

Makine Öğrenimi algoritmalarını uygulamak genellikle karmaşık ve zaman alıcı bir süreçtir. Data Scientist'in en iyi kodlama çözümünü sunabilmesi için iyi yapılandırılmış ve test edilmiş bir dijital ortama sahip olmak çok önemlidir. Programcılar, Data Scientist'in zamanını kısaltmak için bir dizi Python çerçevesi ve kitaplığı seçer. Python'un kullanılabilecek birçok kitaplığı vardır ve bir Veri Bilimcisi veya Geliştiricisinin çok karmaşık sorunları dakikalar hatta saniyeler içinde çözmesine olanak tanır.

Python'da bulunan bazı kitaplık örnekleri şunlardır :

* Makine öğrenimi için Keras, TensorFlow ve Scikit-learn
* NLP için NLTK ve spaCy (Doğal dil işleme)
* Gelişmiş bilgi işlem için SciPy
* Yüksek performanslı bilimsel bilgi işlem ve veri analizi için NumPy
* Bilgisayarla Görme için OpenCV
* Msnipulation Veri analizi için Panda

• Veri görselleştirme için Seaborn ve Matplotloib

3. Geniş Topluluk ve Çok Sayıda Açık Kaynak

Çok sayıda Python kullanıcısı, Python'u geniş ve sağlam bir topluluk haline getirir. Birçok kıdemli Python programcısı, makaleler, videolar gibi çeşitli platformlar aracılığıyla bilgilerini paylaşmaya veya stackoverflow vb. gibi çeşitli platformlarda sıklıkla karşılaşılan teknik soruları yanıtlamaya isteklidir. Bu durum, Python öğrenmeye yeni başlayan herkes için onu çok yararlı kılar. Bu durum aynı zamanda python'a yeni başlayanların python'u bağımsız olarak öğrenmesini sağlar.

4. Sonuç

Makine Öğrenimi veya Makine Öğrenimi, problem çözmeye yönelik faaliyetlerin yürütülmesinde insanların rolünün yerini almak için taklit etmek için kullanılan Yapay Zekadan (AI) bir yaklaşım tekniğidir. Kısacası makine öğrenimi, kullanıcılarından yönlendirme almadan öğrenip iş yapabilmesi için yapılmış bir makinedir.

Genel olarak, makine öğrenimi algoritmaları, denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenme olmak üzere üç türe ayrılır.

Makine öğrenmesi (Machine Learning) uygulamasındaki adımlar şu şekildedir :

1. Probleminizi Tanımlayın

2. Verileri Toplayın ve Hazırlayın

3. Çözüm İçin Algoritmayı Seçin

4. Modelin Değerlendirilmesi ve İyileştirilmesi

Makine Öğreniminde model yapımında Python programlama dilinin kullanılmasının nedenleri şunlardır:

1. Öğrenmesi Kolay  
2. Yüksek ve Çeşitli Kitaplık Çeşitleri  
3. Geniş Topluluk ve Çok Sayıda Açık Kaynak

5. Öneriler

Bu çağda teknolojik gelişmelerin sonuçlarının nicelik ve nitelik olarak artması. Makine Öğreniminin keşfi, başlangıçta insanların faaliyetlerini gerçekleştirmesini kolaylaştıran bir teknoloji olarak kabul edildiğinden çok fazla desteği de davet etti. Ancak pek çok insan, teknolojinin çeşitli şirketlerdeki birçok iş pozisyonunu azar azar değiştirdiğini fark etmeden, farkında olmadan yalnızca makine öğreniminden anlık sonuçların keyfini çıkarıyor. Ama öyle olsa bile, bu teknolojinin gelişimini kontrol ederken onun varlığını inkar etmemeliyiz. Makine Öğrenimi de dahil olmak üzere ileri teknoloji ile ilgili becerileri uygulayabiliriz. Böylece birçok kişinin iş pozisyonları kaydedilebilir ve hatta eskisinden daha fazla gelir elde edilebilir.

6. Kaynaklar

[1] Nurvinda Galuh, Makine Öğrenimi Türleri.

Gading, Serpong, Tangerang, Banten. 2014 https://www.dqlab.id/contact

[2] Peter Norvig, Yapay Zeka, Üçüncü Baskı, Pearson, 2010

[3] Frindy Johana, 2020 , Evde Veri Bilimini Öğrenin, dqlab.id

[4] Tom M. Mitcell, Makine Öğrenimi, The McGraw-Hill, 1997

[5] Model, IEEE Uluslararası Elektrik Mühendisliği ve Bilgisayar Bilimleri Konferansı, 2017

[6] Bachtiar, Y. 2013. Dijital Çağda Endonezya İnsan Karakteri İnşasının Direnişi. Temel Eğitim Dergisi.

[7] Sebastian Raschka. 2015. Python makine öğrenimi

[8] Darujati, C. ve Gumelar, AB (2012). Endonezyaca Metinlerin Sınıflandırılması için Denetimli Tekniklerin Kullanılması. Bandung Metin Madenciliği Dergisi, 16(1), 5-1.

[9] Aractırma Referanslar <https://www.researchgate.net/publication/359394894_Research_on_Machine_Learning_Algorithms_Real-World_Applications_and_Research_Directions>

[10] Gözetimsiz Makine Öğrenme Teknikleri ile Miktara Dayalı Negatif  
<https://pdfs.semanticscholar.org/9d70/416152327b4844ffcf3acaaae5415f5f08f8.pdf>

[11] Robot ASIMO

<https://tr.wikipedia.org/wiki/ASIMO>

[12] Python Öğrenme <https://www.w3schools.com/python/default.asp>

[13] Makine öğrenimi nedir <https://www.youtube.com/watch?v=HcqpanDadyQ>

[14] Makine Öğrenimi Türleri <https://www.javatpoint.com/types-of-machine-learning>

[15] Makine Öğrenimi Nasıl Öğrenilir https://elitedatascience.com/learn-machine-learning