

1. Öğrenme Felsefesi, Günümüzde Makine Öğrenmesi ve Temel Kavramları, Değerlendirme Metrikleri (Ölçütleri)

Öğrenme Faaliyeti ve Makine Öğrenmesi:

İnsanda öğrenme aktivitesi, psikiyatrik olarak şöyle tanımlanmıştır:

İnsanın bilgilerinin, davranışlarının, yeteneklerinin, değerlerinin ve önceliklerinin mevcut olanlarının modifiye(değişmesi) edilmesi veya yenilerinin elde edilmesi işlemi öğrenme faaliyeti ifade edilmiştir. (Richard Gross, Psychology: The science of mind and Behaviour, UK).

Öğrenme yeteneğinin yalnız insan ile sınırlı olmadığı, hayvanlar, makineler ve bazı bitki türlerinde bile görülebildiği bilinmektedir(Wikipedia- Learning). Öğrenme canlıda sadece psikolojik bir olay değildir, fizyolojik temelleri vardır. Öğrenme insan beyinin de mikro ölçeklerde fizyolojik değişikliğe neden olan bir olaydır. Biyolojik sinir hücrelerinin bağlantısının değişimi bir öğrenme faaliyeti sonucudur. Biyolojik sinir sisteminde öğrenilen bilginin, sinir hücreleri arasındaki bağlantısalılık(bağlantı kurulumu ve bağlantının gücü) ile kodlandığı düşünülmektedir.

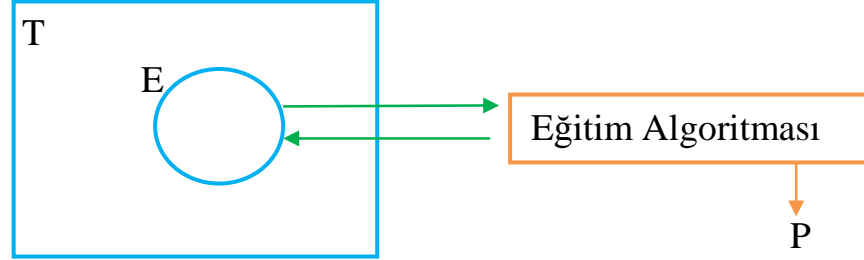
Konumuz makinelerin(bilgisayarların) öğrenmesi olduğu için bilgisayar bilimleri kapsamında inceleyelim. Bu dersimizde yapay öğrenme yerine literatürde yaygın kullanımı nedeni ile "makine öğrenmesi" ismi tercih edilecektir.

Makine öğrenmesi (yapay öğrenme) dersleri veren Tom Mitchell makine öğrenmesinin, "Tecrübeleri ile performansını iyileştirebilen bir program nasıl yapabiliriz?" cevabını arama çabasının bir sonucu olarak ortaya çıktığını belirtmiş ve "Makine Öğrenmesi" isimli ders kitabında makinenin öğrenmesini şöyle formülize etmiştir: (Tom Mitchell. Machine Learning 1997.)

"Bir makine, bir T görev kümesinden elde ettiği E tecrübe kümesi sonucunda bir P performans ölçütüne göre kendi performansını iyileştirebiliyorsa bu makine öğreniyor."

Makine öğrenmenin, 90'larda teorik olarak (T,E,P) yapısı ile formülize edildiği görülebilir. Günümüzde hala bu formülasyon geçerliliğini korumaktadır. Konuya veri analizi açısından baktığımızda T görev kümesi, bir sistemden elde edilen bütün sistem verilerini, E tecrübe kümesi ise makine öğrenmesi algoritmasının eğitiminde kullanılacak olan verilerin kümesi olarak düşünülebilir. Dolayısı ile eğitim kümesi, sistemin üretebileceği bütün verilerin alt kümesi olmak durumundadır. Eğitim kümesi $E \subseteq T$ koşulunu sağlamalıdır. Ancak, çoğu zaman uygulamada sistemin üretebileceği bütün verilerin kümesi çok büyük olur. Bu nedenle eğitim kümesi büyüklüğü veri kümesi büyüklüğünden küçük kalır. ($|E| < |T|$ durumu genelde geçerlidir.) P performans ölçütü makine öğrenmesinde eğitim performansını ölçmek için kullanılan karesel hata yada doğruluk gibi performans ölçütlerini ifade eder. Dolayısı ile

günümüz makine öğrenmesi aslında Tom Mitchell'in (T,E,P) yapısı tanımına göre gelişmiştir. aşağıdaki şekilde (T,E,P) yapısı temsili olarak gösterilmiştir.



Yukarıdaki şekil incelendiğinde E eğitim kümesi ile üstünde P performansı indisini iyileştiren bir algoritmanın aynı zamanda, T kümesinde bulunan ancak E kümesinde bulunmayan veriler (eğitime de kullanılmayan verilerde) içinde P performansının iyi olması gerekir. Bu durum makine öğrenmesinde ilerleyen konularda "genelleme" olarak ele alınacaktır. Eğer, $|E| < |T|$ durumunda makine öğrenmesi algoritmasının P performansı hem E kümesi hem de $T - E$ kümesinde iyileşiyor ise makine genelleme yapabiliyor denebilir. Diğer durumda, sadece E kümesinin yüksek performans ile öğrenilmesi iyi bir öğrenme faaliyeti için yeterli görülmez. Bu duruma veri kümesinin "ezberlemesi" diyebiliriz. İyi bir genelleme için,

- E eğitim kümesinin, T görev kümesini yeterince iyi temsil etmesi gerekir. Bu durum eğitim kümesinin, sistem davranışını iyi ifade eden verilerden oluşturulması anlamına gelir. Bu nedenle "eğitim kümesi tasarımı" makine öğrenmesi performansı için önemlidir. Bu durum, sahada karşılaşılan sorunları (T görev kümesi) çözebilmesi için bilgisayar mühendisleri eğitim programının (E eğitim kümesi) iyi hazırlanması durumun benzetilebilir.

- Makine öğrenmesi algoritmasının ve öğrenen modelinin, T görev kümesi verilerinin öğürebilmesi için yeterli düzeyde etkinlik ve karmaşıklıkğa sahip olması gerekir. Bu duruma şöyle örnek verilebilir. Aslında algoritma etkinliği ve model karmaşıklığı, canlılarda öğrenmenin fizyolojik altyapısına (sinir sistemi karmaşıklığına) karşılık gelir. Nispeten ilkel bir sinir sistemine ve öğrenme kabiliyetine sahip bir canlı, çok karmaşık bir süreci hiç öğrenebilir. Belki de bir sincaba diferansiyel denklem çözdürmeyi hiç bir zaman öğretemeyeceğiz.

Aşağıdaki tahmin problemi eğitim kümesi üstünde çalıştıktan sonra bir sonraki veriyi etmeye çalışın.

Öğrenme deneyi 1: $E=\{(1 \rightarrow 0), (2 \rightarrow 2)\}$ kümesinde çalışalım \Rightarrow T-E kümesinden bir soru ($6 \rightarrow ?$) cevaplayalım

Eğitim kümesi, insanın doğru matematiksel ilişkiyi görülebilmesi için yeterli olmayabilir. Bu nedenle doğru tahmin edilemeyebilir. Doğru tahmin ($6 \rightarrow 10$) olmalı idi. Şimdi doğru tahmin şansını artırmak için eğitim kümesini artıralım. Böylece, veriden daha kolay ve iyi bir genelleme yapma şansını artırmaya çalışalım.

Öğrenme deneyi 2: $E=\{(1 \rightarrow 0), (2 \rightarrow 2), (3 \rightarrow 4), (4 \rightarrow 6)\}$ \Rightarrow T-E kümesinden ($7 \rightarrow ?$) sorusu daha kolay cevaplanabiliyor. Çünkü, matematiksel ilişkinin ($x \rightarrow 2*x-2$) olduğunu eğitim kümesi üzerinden görebildiniz. Eğitim kümesi üstünde ($x \rightarrow 2*x-2$) ilişkisini keşfederek iyi bir genelleme yapabildiniz. Veriyi üreten sistemi artık ($x \rightarrow 2*x-2$) modeli ile ifade ettiniz ve verilerden öğrenilen bu model yardımı ile T-E (eğitim kümesinde olmayan) verileri doğru bir şekilde tahmin edebilirsiniz.

Aslında bu uygulama çözülen problem olasılık ve istatistik dersinde gördüğünüz lineer regresyon problemi çözümü. Burada $y=2x-2$ lineer regresyon modelini eğitim kümesini inceleyerek keşfettiniz.

Öğrenmenin Temel Mimarisi:

Günümüzde, konuyu pratik çalışmaları daha yakından ele alan bir Makine Öğrenmesi kitabının yazarı Phil Kim kısaca şöyle ifade etmiştir.

“Makine öğrenmesi verilerden bir model üretebilen bir tekniktir.” (Phil Kim, MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence, Apress, 2017)

Daha pratik bir tanım, (Guru99, Machine Learning Tutorial for Beginners) sitesinden şöyle verilmiştir.

“Makine öğrenmesi bir programcı tarafından özellikle kodlanmaya ihtiyaç duymadan örnekler üzerinden kendini iyileştirebilen bir sistemdir. Burada önemli nokta, makinenin daha doğru çalışmak için toplanan verilerden faydalanabilmesidir.” (Guru99)

Makine öğrenmesi, günümüzde bilgisayar bilimlerinin yapay zeka başlığı altında incelenen bir alt konusudur.

Makine öğrenmesinin geçmişine bakıldığında, makine öğrenmesi kavramı ilk olarak Arthur Samuel’in 1956’da ortaya koyduğu bilinmektedir. Şöyle kavramsallaştırılmıştır:

"Özellikle bir programlanmaya ihtiyaç duymadan tahminde bulunan, tecrübeleri ile performansını kendiliğinden iyileştiren, verilerden otomatik olarak öğrenen bir makine" (Arthur Samuel,1956) olarak konuyu ele almıştır.

Makine öğrenmesi algoritmalarından, günümüzde yaygın anlamı ile toplanan verileri kullanarak, verileri oluşturan süreçleri veya sistemleri temsil edebilen bir matematiksel model (öğrenen modeli) oluşturmaya beklenir. Bu öğrenen model yardımı ile bir görevi yürütmek için veriler ile ilgili bir programlamaya ihtiyaç duymadan otomatik olarak verilerden süreçleri öğrenmesi ve beklenen işlevleri yapabilmesi beklenir (Wikipedia- Machine Learning).

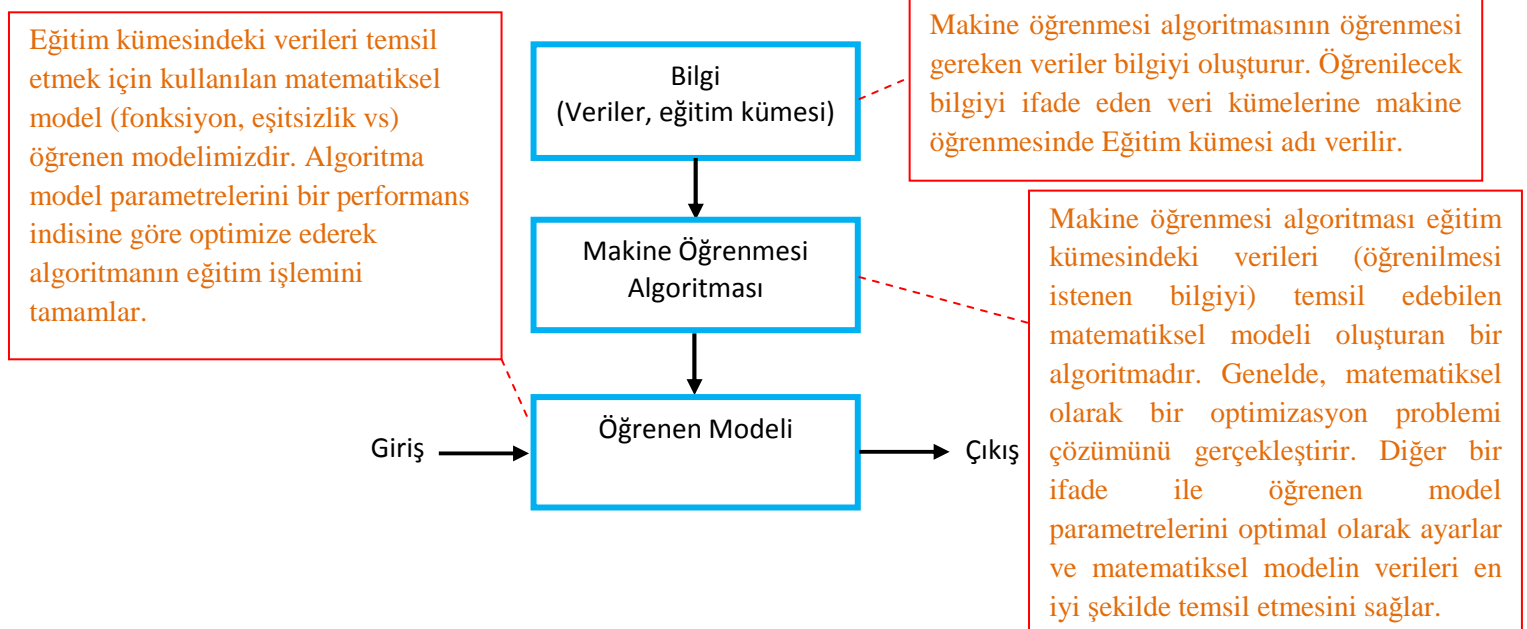
Yukarda verilen çeşitli makine öğrenmesi tanımlarda geçen temel kavramlar şunlardır:

Matematiksel Model (Öğrenen model): Makine öğrenmesi algoritması tarafından veriler (eğitim verileri, örnekler, toplanan data) yardımı ile öğrenme faaliyetinin yürütüldüğü matematiksel model bir öğrenen modeli oluşturur. Model, istenen cevabı vermek üzere eğitilir veya öğrenme faaliyeti sonucunda belirlene işlevler için performansını artırır.

Bilgi: Öğrenme işlemi için gereken eğitici bilgiler, toplanan veriler, örnekler, eğitim kümeleri, tecrübeler, etkileşimler, fizikler ölçümler ve gözlemler.. vs.

Makine öğrenme algoritması: Veriler üzerinde öğrenme modeli oluşturabilen yöntemler makine öğrenmesi algoritmalarıdır. Genelde öğrenme algoritmaları bir optimizasyon problemi çözerler.

Şekil 1’de bu bileşenleri içeren günümüz makine öğrenmesi mimarisi görülmektedir.



Şekil 1. Günümüz makine öğrenme mimarisini oluşturan temel bileşenler (Phil Kim, MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence, Apress, 2017)

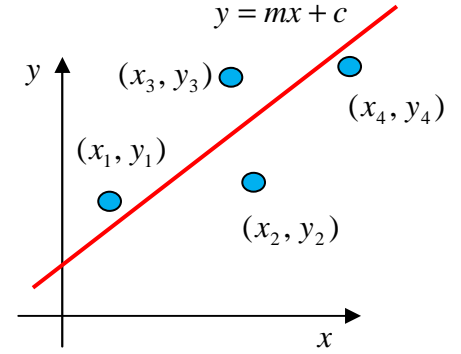
Yukarıda ifade edilen makine öğrenmesi mimarisi için basit bir örnek olması açısından olasılık ve istatistik dersinde görmüş olduğumuz Lineer Regresyon modelini hatırlayalım. Buna göre,

**Bilgi:* Veri kümesi. (Yandaki grafikte $\{ (x_1, y_1), (x_2, y_2) \}$

$(x_3, y_3), (x_4, y_4) \}$ noktaları ile gösterilmiştir.)

**Makine Öğrenmesi Algoritması:* En küçük kareler yöntemi ile lineer regresyon işlemi

**Öğrenen Modeli:* Lineer regresyon sonucunda elde edilen öğrenme modeli bir doğru denklemi $y = mx + b$ dir.



Yukarıda mavi renkli yuvarlak ile gösterilen veri noktalarından en az karesel hata geçen doğruyu bulma problemi lineer regresyon problemi olarak adlandırılır.

Bu veri kümesini en az hata ile elde eden $y = mx + b$ doğrusu elde edildikten sonra herhangi bir x veya y değeri için tahmin (kestirim) yapılabilir. Dolayısı ile aslında lineer regresyon problemini çözen bir program yazdığınızda bu program kendisine verilen veri kümesinden bir lineer model elde eder (öğrenme sürecini tamamlar) ve burada minimum kareler yöntemi, linear(doğrusal) tahminler yapabilen basit ve ilkel sayılabilecek bir makine öğrenmesi algoritması olarak görülebilir.

Makine Öğrenmesinin Önemi ve Diğer Disiplinlerle İlişkisi:

Günümüzde makine öğrenmesi birçok bilim dalında ve mühendislik problemlerinde uygulamaları sürekli artan güncel bir konudur:

**Veri biliminde (data science):* Verilerin modellenmesi, yorumlanması, sınıflandırılması, verilere dayalı tahmin, büyük veri (Big data) uygulamaları makine öğrenmesine ihtiyaç duyar.

İstatistik: İstatistik ile makine öğrenmesi arasında çok sıkı bağlar vardır. Bu iki dal uygulamaları ile birbirini destekler niteliktedir. Makine öğrenmesi, uygulama açısından faydalı veri analizlerine imkan sağlarken sağlar iken, istatistik alanında gelişmeler ve yeni yöntemler makine öğrenmesini yöntemlerine katkı yaparlar.

**Temel bilimler (fizik, kimya, biyoloji ve diğer bilimler):* Temel bilimlerde ortaya çıkan büyük verinin analizi, modellenmesi ve deney tasarımı gibi süreçlerde makine öğrenmesi rol oynamaya başlamıştır.

**Mühendislik alanları (makine, inşaat, elektronik, bilgisayar):* Mühendislik sistemlerinin (makine, yapı, devre..vs) tasarımı, optimizasyonu gibi süreçlerde makine öğrenmesi rol alabilmektedir.

**İşaret işleme, görüntü işlemem, medikal uygulamalar*: Ses tanıma, bilgisayarda görme, medikal tanı sistemlerinde makine öğrenmesi araçları temel bileşen haline gelmiştir.

**Ekonomi bilimi*: Ekonomik sistem modellemesi ve tahmin konularında yapay zeka aktif kullanılmaktadır.

**Robotik, Endüstriyel otomasyon, Internet of Think (IoT)*: Gelişen alanlarda makine öğrenmesi temel bir bileşen haline dönüşmüştür. Endüstri devrimi 5.0 metaforu aslında makine öğrenmesi ve yapay zeka uygulamalarının günlük hayatta ve üretim sistemlerinde rol alması durumunun yol açtığı akıllı sistemlere dönüşümünü ifade etmektedir.

Makine Öğrenmesi Yöntemi Türleri

Öğrenme mekanizmasına göre makine öğrenmesi 3 temel grupta sınıflanır:

1. *Denetimli Öğrenme (Supervised Learning)*: Her bir girişe karşılık doğru çıkışların tanıtıldığı öğrenme yöntemi denetimli öğrenmedir. Burada, bir denetmen (supervizer) doğru sonuçları tanımlar. Denetimli öğrenme yönteminde kullanılan veri kümeleri yaygın olarak etiketli veri (labeled veri) olarak adlandırılır. Çünkü her bir giriş verisi (özellikler) için öğrenilmesi gereken doğru çıkış(sonuç) veri kümesinde bulunur. Denetimli öğrenmede denetmenin tecrübesi ve bilgisi eğitim kümesi yardımı ile öğrenen modele aktarılabilir. Bu nedenle eğitim kümesi tasarımı denetimli öğrenme başarısı için çok önemlidir ve eğitim kümesi öğrenilmesi gereken bilgiyi ifade eder nitelikte olmalıdır. Denetimli öğrenmenin yaygın uygulandığı problemler: Regresyon problemleri ve Sınıflama problemleri

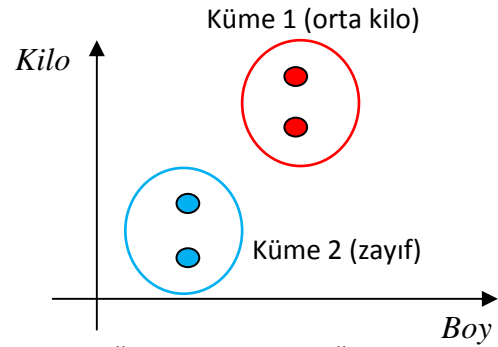
Örnek: Bir meyve sınıflama işleminin denetimli eğitimi için bir etiketli eğitim kümesi örneği şöyle ifade edilebilir.

Girişler (özellik)			Doğru Çıkış (İstenen Çıkış)
Şekil	Uzunluğu	Ağırlığı	Etiket (Denetmen Bilgisi)
Sivri	20 cm	120 gr	1 (Muz)
Sivri	16 cm	100 gr	1 (Muz)
Oval	12 cm	95 gr	2 (Elma)
Oval	3 cm	25 gr	3 (Çilek)

2. *Denetimsiz Öğrenme (Unsupervised Learning)*: Denetimsiz öğrenmede giriş verileri (özellikler) için bir doğru(istenen) çıkışın tanıtılmadığı öğrenme yöntemidir. Burada, doğru sonuçları veren bir denetmen (supervisor) bulunmaz. Bazı kaynaklarda denetimsiz öğrenme için kullanılan eğitim kümeleri etiketsiz veri (unlabelled data) olarak adlandırılmıştır. Bir denetimci olmadığı için öğrenen modelin kendi tecrübelerine dayanarak öğrenmesi beklenir. Çünkü her bir veri için doğru çıkış önceden tanımlanmamıştır. Denetimsiz öğrenmenin uygulandığı problemler: yaygın olarak kümeleme problemleridir.

Örnek: Etiketsiz bir boy-ağırlık verileri ile kümeleme işlemi yapılsın.

Etiketsiz Ölçüm verileri	
Boy	Ağırlığı
184 cm	85 kg
178 cm	80 kg
181 cm	70 kg
177 cm	68 kg



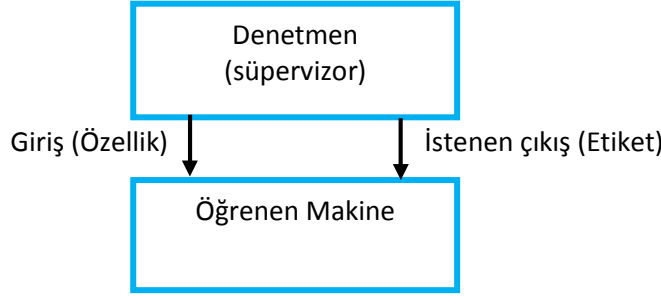
Boy-Ağırlık verilerinin dağılımı ve birbirine mesafeleri değerlendirilerek iki kümeye ayrılabilir. Küme 1 (orta kilo) ve Küme 2 (Zayıf).

3. Takviyeli Öğrenme (Reinforced Learning): Bazı kaynaklarda ödüllü (rewarded) öğrenme olarak da isimlendirilir. Burada denetimci doğru sonuçları doğrudan tanımlamaz. Denetimsiz öğrenmede olduğu gibi veriler etiketsizdir yani doğru(istenen) sonuç tanımlanmamıştır. Ancak, öğrenen model doğru veya yanlış sonuç ürettiğinde bir ödül ve ceza mekanizması yardımı ile öğrenme işlemi yönlendirilir. Doğru sonuçlarda ödül verilir ve yanlış sonuçlar ceza alır. Öğrenme modelinin ödülü maksimize ve/veya cezayı minimize edecek şekilde davranması sağlanarak takviyeli öğrenme gerçekleşir. Dolayısı ile takviyeli öğrenme, ne tam anlamı ile bir denetimsiz öğrenmedir, nede tam bir denetimli öğrenmedir. İkisi arasında hibrid bir öğrenme yöntemi olarak görülebilir. Doğru(istenen) sonuçların denetimlide olduğu gibi doğrudan verilmediği ancak ürettiği çıkışlara(aksiyonlar) ödül ve ceza mekanizması ile denetimsiz öğrenmenin yönlendirildiği bir öğrenme yaklaşımıdır. Bu yöntem tecrübeye bağlı öğrenme için çok uygundur ve genellikle bir ortamında yapay zeka ajanlarının (agent) tecrübelerine dayanarak performansını iyileştirdiği uygulamalara uygundur. Yapay zekada robotik ve ortamda gerçek zamanlı öğrenen otonom sistem uygulamaları için çok daha uygun bir öğrenme türüdür. Çünkü, yapay zeka ajanı (etmen) ortam ile tecrübeleri sonucunda doğru kararlar vermiş ise ödülünü alır, yanlış kararları sonucunda maliyet(ceza) ile karşılaşır. Algoritma ödül maksimizasyonu veya ceza minimizasyonu sağlayarak, makinenin gerçek ortamda işlevsel performansını sürekli artırabilir ve deneyinden öğrenme gerçekleşebilir.

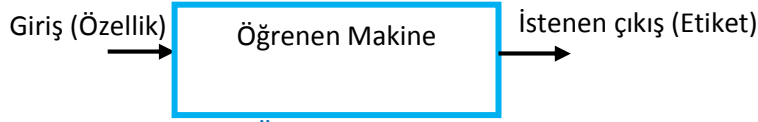
Örnek: Etiketsiz bir boy-ağırlık verileri ile kümeleme işlemi yapılsın.

Etiketsiz Ölçüm verileri		Öğrenme Modeli Çıktısı	Takviye
Boy	Ağırlığı	Tahmin	Ceza-Ödül
184 cm	85 kg	Aşırı kilo	Ceza (-5 puan)
178 cm	80 kg	Orta kilo	Ödül (+5 puan)
190 cm	70 kg	Orta kilo	Ceza (-5 puan)
177 cm	68 kg	Düşük kilo	Ödül (+5 puan)
175 cm	65 kg	Düşük kilo	Ödül (+5 puan)

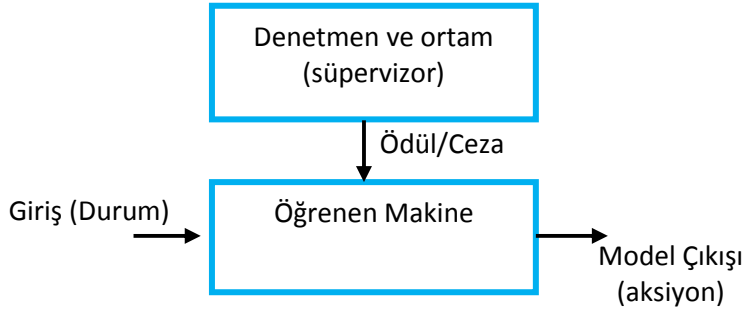
Üç farklı öğrenme yöntemi:



Denetimli Öğrenme: Sadece denetmenin tecrübesinin öğrenilmesi beklenir. Keşfe (kendi kendine) öğrenmeye açık değildir.



Denetimsiz Öğrenme: Bir denetmen desteği yoktur. Tamamen keşfe dayalı (kendi-kendine) öğrenmesi beklenir



Takviyeli Öğrenme: Kendi tecrübelerinden denetmenin yönlendirmesi ile öğrenir. Hem keşfe açıktır hem de denetmen yönlendirmesi vardır.

Basitleştirilmiş bir örnek üzerinde üç öğrenme yöntemini inceleyelim:

Elektrik devresi laboratuvarına katılan 3 farklı öğrenci grubunu inceleyelim:

Denetimli Öğrenme: Laboratuvar görevlisi (denetimci), deney öncesinde devre elemanlarını tanıtır, kurulacak devreyi ve kuruluşunu gösterir ve bu bilginin anlaşıldığından emin olduktan sonra öğrencilerden deneyi yapmasını isterse bu denetimli öğrenme türünde bir eğitim olarak görülebilir. Öğrenciler deneye başlamadan önce bilmeleri gereken her şeyi öğrenmiştir ve görevlinin istediği devreyi en az maliyetle (zaman ve malzeme kaybı ile) kurup deneyi tamamlar. (Öğrenilen bilgi kalitesi)/(Eğitim maliyeti) oranı bakımından en verimli yöntem olacaktır. Denetmen kaliteli bilgiyi sunar öğrenci en az maliyetle (zaman, cihaz israfı) öğrenir. Burada öğrencinin kendi kendine başka bir devre kurması, denemeler yapması ve denetimcinin verdiği bilgi dışında çalışması beklenmeyebilir. Ancak, denetmen deneyleri

öğrencilerin devre elemanları ve işlevlerini anlayabilmesi (iyi bir genelleme sağlaması) için iyi tasarlaması gerekir. Burada denetmenin eğitim sürecini tasarlamada başarısı önemlidir.

Denetimsiz Öğrenme: Laboratuvar görevlisi (denetimci) yoktur. Öğrencilerin laboraturvardaki devre elemanları ve cihazları kendileri tecrübe ederek çalışmaları istenir. Ön bilgi veren denetimci olmadığı için öğrenciler kendi kendilerine önce devre elemanlarını tek tek deneyerek işlevlerini anlamaya çalışacak ve keşfetdikleri işlevlerine göre elemanları gruplara ayıracaktır. Sonra devre kurmaya çalışarak devrelerin işlevlerini tecrübe edip bu devreleri işlevlerine göre gruplamaya çalışırlar vs. Denetimsiz olan bu öğrenmenin önemli bir sorunu öğrenilen bilginin niteliği ve işlevselliğinin düşük olması olacaktır. Ayrıca, deneme sürecinde yanlış kullanımlar, bozulan devre elemanları vs. maliyeti ve eğitim için harcanan süreyi artıracaktır. Bireysel tecrübeler ve keşfe çok açık olmasına karşın (Öğrenilen bilgi kalitesi)/(Eğitim maliyeti) oranı bakımından verimsiz bir yöntem olarak görülebilir.

Takviyeli Öğrenme: Laboratuvar görevlisi (denetimci) laboraturvarda bulunur ve öğrencilerin laboraturvardaki devre elemanları ve cihazlarını kendileri tecrübe ederek çalışmaları ister. Ancak, laboratuvar süresince öğrencilerin arasında gezerek tebrik veya ikazları ile öğrencilerin çalışmalarını yönlendirir. Dolayısı ile öğrencilerin kendilerinin özel tecrübe ve keşiflerine açık olmasına rağmen belli bir amaca doğru yönlendirilen bir laboratuvar ortamı sağlanır. Hem bireysel tecrübeler ve keşfe açık hem de (Öğrenilen bilgi kalitesi)/(Eğitim maliyeti) oranı bakımından denetimsiz öğrenmeden daha verimli olur.

Bu derste çoğunlukla denetimli öğrenme yöntemleri üzerinde durulacaktır. Aşağıda bu ders kapsamında işlenmesi planlanan makine öğrenmesi problemleri ve öğrenme modelleri görülmektedir.

