

ÖĞRENME KAVRAMI

Yapay zekâ, insanların yapmakta bilgisayarlardan daha iyi olduğu eylemleri bilgisayarların yapabilmesini sağlamak amacıyla yürütülen çalışmalar olarak tanımlanmaktadır. Kendisine verilmiş kurallar dizini kapsamında çalışma mekanizmasını yürüten günümüz bilgisayarlarının öğrenme yeteneğinin, sadece kurallara bağlı olarak değil bu kurallar arasında bağlantı kurarak ve tecrübelерinden çıkarım yaparak çalışan insan zekasının sahip olduğu öğrenme yeteneğinden daha düşük olduğu âşikardır. Bilgisayarların insan zekasını taklit ederek öğrenme yeteneğine sahip olmasını amaçlayan bu yapay zekâ çalışmalarını öğrenme mekanizmalarına ilişkin çeşitli araştırmalar takip etmektedir.

İnsan zekasında önemli bir yeri olan “ezberlemek” ve “öğrenmek” fiillerinin, bilgisayarlar için ne anlama geldiğinin ve aralarındaki farkın bilinmesi konunun devamındaki teknik kısımların anlaşılabilmesi için önem teşkil etmektedir. İnsan beyni için kısmen zor bir iş olan ezberlemek eylemi, bilgisayarlar için bir dosyaya veri kaydetmekten ibarettir. Bu sebeple ezberlemek fiili yapay zekâ algoritmaları için ilgi çekici değildir. Buna karşılık matematiksel becerilerin kazanılması gibi işler insan beyninde ezber yoluyla gerçekleşmez. Ezber yoluyla gerçekleşmesi mümkün olmayan bu eylemler gözlem, tecrübe ve çıkarım yolu ile elde edilir. Bu şekilde elde edilmiş kazanımlar “öğrenme” olarak tanımlanırken, yapay zekâ ve özellikle makine öğrenmesi çalışmaları, bilgisayarın öğrenme işini yapabilmesini hedeflemektedir.

Gözlem ve tecrübe yolu ile eldeki sonlu verinin kullanılarak daha önce karşılaşılmamış durumlar için çıkarım yapılması *genelleştirme* olarak adlandırılmaktadır. Genelleştirilmiş kavramların çeşitli *özellikleri* referans alınarak gruplara ayrılması ise *sınıflandırma* olarak tanımlanır. Sınıflandırma görevini özellik vektörlerini sonlu sayıda sınıfa bölebilen, bir özellik vektörünü bir sınıf değerine eşleyen *sınıflandırıcılar* gerçekleştirmektedir.

Bilgisayarlar için öğrenme, toplanmış ve sınıflandırılmış veriler kullanılarak, kavram için özelliklerden sınıf değeri hesaplayan bir fonksiyon oluşturmaktır. Bu şekilde bir özellik vektörünü bir sınıf değerine eşleyen fonksiyon ise *öğrenme aracı/learning agent/perceptron/algılayıcı* olarak tanımlanır. Sınıflandırılmış verilerle beslenen algılayıcı, programlanmadan, yalnızca eğitim verilerini kullanarak öğrenmeye, kendini geliştirmeye ve bunların sonucunda özellik vektöründen sınıf değerine mümkün olduğunca iyi eşlemeler oluşturmaya çalışır. Bir algılayıcı, zaman içinde birçok eğitim örneğini gördükten sonra; yeni ve bilinmeyen veriler üzerindeki uygun bir kriterle ölçülen performansını geliştirmektedir. Algılayıcı; amacı bir eşlemeyi öğrenmek olan *görev*, hangi öğrenme algoritmasının kullanacağıının belirlendiği *ajan sınıfı*, öğrenme algoritmasının öğrenmesi gereken bilgileri içeren *eğitim verileri*, eğitilmiş aracın genelleme yapıp yapamayacağını değerlendirmesini için test *verileri* ve ajanın ne kadar doğru öğrenip öğrenmediğinin kontrolü için *performans ölçüsü* gibi kavramlarla karakterize edilir.

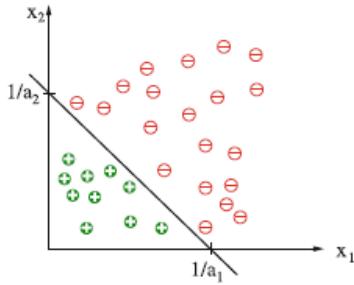


Fig.1: Lineer olarak sınıflandırılabilen iki boyutlu eğitim veri seti.

Sınıflandırma doğrusunun denklemi: $a_1x_1 + a_2x_2 = \theta$

Öğrenme ajanının, Fig.1'de olduğu gibi bazı iki boyutlu eğitim verilerini tek bir doğru ile sınıflandırması mümkün değildir ve bu tür eğitim verileri *lineer olarak ayrırlabildir*dir. Boyutu n olan bir eğitim setinin lineer olarak ayrılabilmesi için ise tek bir doğru yeterli olmamakta ve bu işlem için $n-1$ boyutlu bir hiper düzlemler kullanılmaktadır. \mathbb{R}^n (n boyutlu uzay) bulunan $n-1$ boyutlu hiper düzlemler ise, Θ eşik/sınır değeri olmak üzere

$$\sum_{i=1}^n a_i x_i = \theta$$

şeklinde tanımlanır.

M_1 ve M_2 , \mathbb{R}^n uzayının alt kümeleri olmak ve a_1, a_2, \dots, a reel sayılar olmak üzere; M_1 kümesinin tüm elemanları için,

$$\sum_{i=1}^n a_i x_i > \theta$$

ve M_2 kümesinin tüm elemanları için

$$\sum_{i=1}^n a_i x_i \leq \theta$$

şartı sağlanıyorsa M_1 ve M_2 lineer olarak ayrılabilirdir.

Bu denklemler göz önüne alındığında özellik vektörlerinin hangi sınıf değerine eşleniği bilinen verilerin kullanılarak denklemde bulunan ağırlık kat sayılarının bulunup bu kat sayılar ışığında çıkarımlar yaparak hangi sınıf değerine eşleniği bilinmeyen yeni verilerin de öğrenme aracı/algılayıcısında bulunan fonksiyonlarla gerçekleştirilmesi bilgisayarlar için “öğrenme” olarak tanımlanmaktadır.

Kaynak:

- Wolfgang ERTEL, Introduction to Artificial Intelligence, Second Edition, Springer