

# Sınıflandırma Algoritmaları

Sınıflandırma algoritmaları, makine öğrenimi ve veri madenciliği alanlarında yaygın olarak kullanılan ve veri analizinde önemli bir rol oynayan yöntemlerdir. Bu algoritmalar, veri setlerindeki örnekleri belirli sınıflara veya kategorilere atayan modellerdir. İşte bazı temel sınıflandırma algoritmaları ve nasıl çalıştıkları hakkında bilgi:

## 1. K-En Yakın Komşu (K-Nearest Neighbors - k-NN)

K-NN, belirli bir örneği sınıflandırmak için en yakın k komşusunun çoğunluğuna dayanan bir algoritmadır. Temel prensibi, benzerlik ölçüleri (genellikle uzaklık metrikleri) kullanarak örneğin hangi sınıfa ait olduğunu belirlemektir. K büyüklüğü algoritmada belirlenen bir parametredir.

## 2. Destek Vektör Makineleri (Support Vector Machines - SVM)

SVM, iki sınıf arasında bir karar sınırüretmeye çalışırken maksimum marjini (sınırlar arasındaki mesafe) maksimize etmeye odaklanır. Non-lineer problemlerde çekirdek fonksiyonları kullanarak verileri daha yüksek boyutlu uzaylara taşıyabilir.

## 3. Karar Ağaçları (Decision Trees)

Karar ağaçları, bir dizi karar kuralı kullanarak veriyi sınıflandıran ağaç yapısında modellerdir. Her düğüm, bir özelliğin bir kararını temsil eder ve dallanmalarla veriyi sınıflandırmak için kullanılır.

## 4. Rastgele Ormanlar (Random Forests)

Rastgele ormanlar, birçok karar ağacının (ensemble yöntemi) bir araya getirilmesiyle oluşturulur. Her ağaç rastgele bir alt örnek kümesiyle eğitilir ve sonuçlar toplanarak sınıflandırma yapılır. Bu yöntem, overfitting'i azaltırken modelin genel performansını artırır.

## 5. Naive Bayes

Naive Bayes, Bayes teoreminin temelinde çalışan bir olasılık tabanlı sınıflandırma algoritmasıdır. Her sınıf için özellikler arasındaki bağımsızlık varsayımını yapar ve bu özelliklerin olasılıklarını kullanarak sınıflandırma yapar.

## 6. Yapay Sinir Ağları (Artificial Neural Networks - ANN)

Yapay sinir ağları, biyolojik sinir ağlarını taklit eden ve çok katmanlı perceptron (sinir hücresi) yapısını kullanarak sınıflandırma problemlerini çözen modellerdir. Derin öğrenme alanında daha karmaşık yapılar kullanılarak daha derin ve karmaşık özellikleri öğrenme yeteneği bulunur.

## 7. Gradient Boosting Machines (GBM)

GBM, zayıf öğreniciler olarak bilinen modelleri (genellikle karar ağaçları) bir araya getirerek güçlü bir öğrenici oluşturur. Boosting yöntemiyle, hatalı sınıflandırmaları düzeltmek için ardışık olarak yeni modeller ekler.

## **8. Doğrusal ve Lojistik Regresyon**

Doğrusal regresyon, girdi değişkenleri arasında doğrusal ilişkileri modelleyen basit bir sınıflandırma yöntemidir. Lojistik regresyon ise doğrusal regresyonu sınıflandırma problemleri için adapte eden bir yöntemdir, sınıflandırma kararını olasılık tahminlerine dönüştürür.

Bu sınıflandırma algoritmaları, farklı veri yapıları ve problemler için avantajlara sahiptir ve uygulandıkları alanlarda çeşitli güçlü ve zayıf yönler gösterebilirler. Hangi algoritmanın kullanılacağı, veri setinin doğası, boyutu, gürültü düzeyi ve istenen çıktıya bağlı olarak belirlenir.