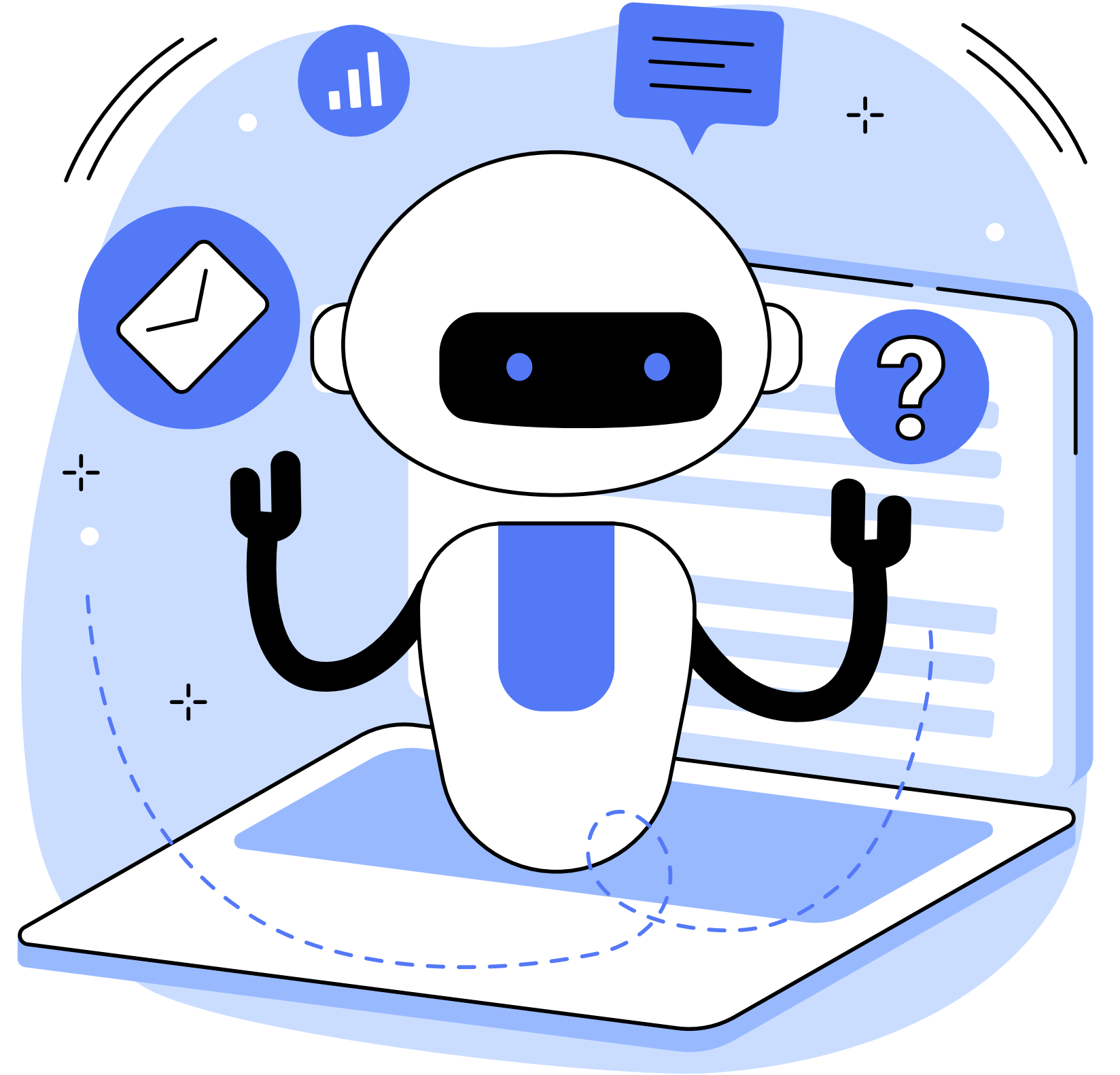




MAKİNE ÖĞRENMESİ TEMELLERİ

MAKİNE ÖĞRENMESİ NEDİR?

bilgisayarların açıkça programlanmadan, veri ve deneyimlerden öğrenmelerini sağlayan bir yapay zeka alt dalıdır. Makine öğrenmesi, bilgisayar sistemlerinin belirli görevleri yerine getirmek için verilerden örüntüleri ve ilişkileri otomatik olarak öğrenmelerini amaçlar.



TEMEL KAVRAMLAR



1

Veri: Makine öğrenmesi algoritmalarının eğitilmesi ve test edilmesi için kullanılan bilgiler.

2

Model: Veriden öğrenilen matematiksel temsil.

3

Öğrenme Algoritması: Verileri analiz edip model oluşturmak için kullanılan yöntem.

4

Özellikler : Modelin öğrenmesi için kullanılan giriş değişkenleri.

5

Etiketler : Denetimli öğrenmede modelin tahmin etmeyi öğrendiği hedef değişkenler.

MAKİNE ÖĞRENMESİ TÜRLERİ

Denetimli Öğrenme

Modelin, giriş ve çıkış veri çiftlerinden (özellikler ve etiketler) öğrenmesini içerir. Amaç, yeni verilere dayalı olarak doğru tahminler yapmaktır.

Denetimsiz Öğrenme

Etiketlenmemiş veri üzerinde çalışır ve verinin iç yapısını keşfetmeye çalışır.

Pekiştirmeli Öğrenme

bir ortam içinde belirli bir hedefe ulaşmak için nasıl hareket etmesi gerektiğini öğrenir. Bu süreç, deneme yanılma ve ödüllendirme mekanizmalarıyla gerçekleşir.

MAKİNE ÖĞRENMESİ SÜRECİ

Veri Toplama

Modelin eğitilmesi için gerekli verilerin toplanması.

Veri Hazırlama

Verilerin temizlenmesi, normalleştirilmesi ve uygun formata getirilmesi.

Özellik Mühendisliği

Modelin daha iyi performans göstermesi için verilerden anlamlı özelliklerin çıkarılması.

Model Seçimi

Göreve uygun makine öğrenmesi algoritmasının seçilmesi.

Model Eğitimi

Verileri kullanarak modelin eğitilmesi.

Model Değerlendirme

Modelin performansının test verileri üzerinde değerlendirilmesi.

Model İyileştirme

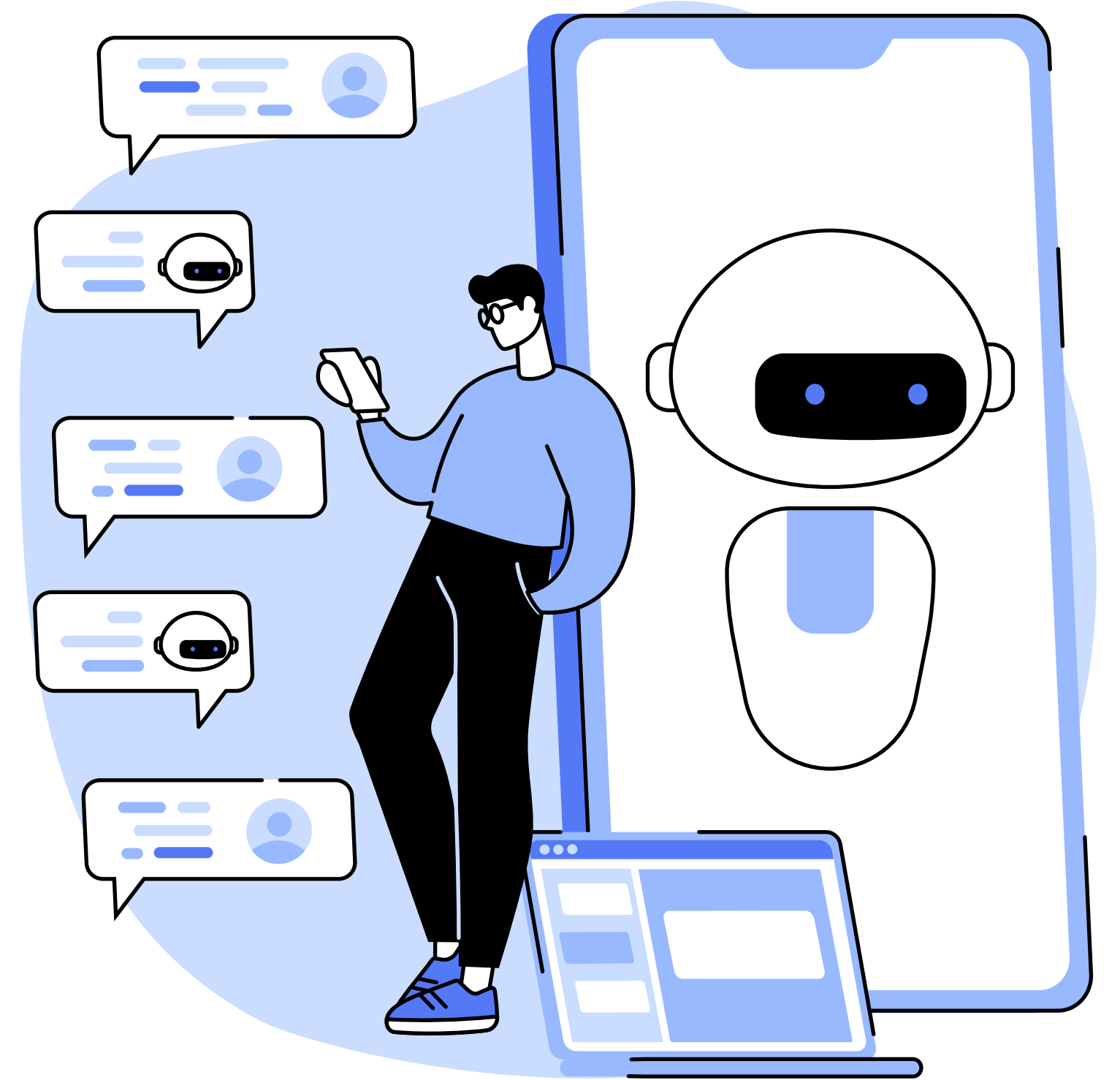
Modelin performansını artırmak yöntemlerle iyileştirilmesi.

Model Dağıtımı

Eğitilen modelin gerçek dünya uygulamalarına entegre edilmesi.

MAKİNE ÖĞRENMESİ ALGORİTMALARI

Makine öğrenmesi algoritmaları, belirli görevleri yerine getirmek için verilerden öğrenme ve tahmin yapma sürecinde kullanılır.



MAKİNE ÖĞRENMESİ ALGORİTMALARI

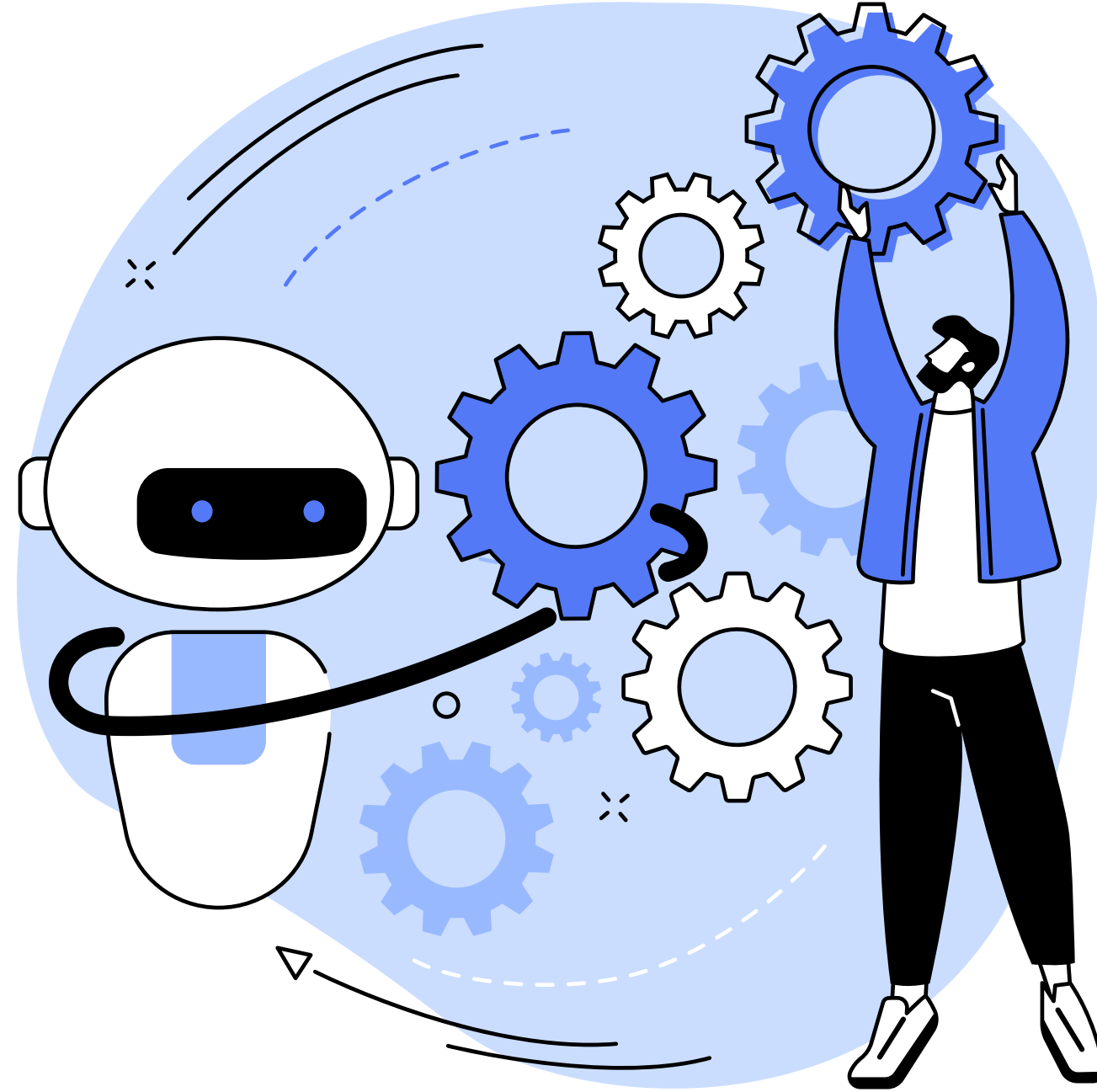
Lineer
Regresyon

Lojistik Regresyon

Destek Vektör
Makineleri

Karar Ağaçları

Rastgele Ormanlar



K-En Yakın Komşu

K-Ortalamalar

Hiyerarşik
Kümeleme

Ana Bileşen
Analizi

Q-Öğrenme

KÜMELEME (CLUSTERING)

Kümeleme, denetimsiz öğrenmenin bir yöntemidir ve birçok alanda kullanılan istatistiksel veri analizi için yaygın bir tekniktir. Veri kümesindeki verileri yorumlayarak ortak noktaları bulmak ve bunları kümeleştirme işlemi yapılarak anlamlı bir veri elde edebilmektir.



KÜMELEME ÇEŞİTLERİ

Hiyerarşik Kümeleme

Hiyerarşik kümeleme, veriyi hiyerarşik bir yapıda gruplayan bir kümeleme yöntemidir. Bu yöntem, veriyi, her biri daha küçük kümelere bölünen veya daha büyük kümelere birleştirilen küme alt kümelerine ayırır.

K-Means Algoritması:

K-Means algoritması, denetimsiz öğrenme yöntemlerinden biridir ve veriyi önceden belirlenmiş sayıda k kümeye ayırmak için kullanılır. Her küme, veri noktalarına en yakın bir merkez (centroid) ile temsil edilir.

Ağırlık Ortalama Kaydırma Kümelemesi:

veri noktalarını, yoğunluklarının yüksek olduğu bölgelerde kümeler oluşturacak şekilde gruplamayı amaçlar. Bu yöntem, özellikle kümelerin sayısının önceden belirlenemediği durumlarda etkili olabilir.

DBSCAN

DBSCAN, ortalama kaymaya ekseninde, benzer ve yoğunluklu bölgeleri kümeleyen bir algoritmadır, ancak birkaç önemli avantajı vardır. Minimum bölge sayısında ve uzaklıkta maksimum yoğunluk bölgesi oluşturulması hedeflenir

SINIFLANDIRMA

Sınıflandırma, makine öğrenmesi ve veri madenciliğinde kullanılan, verileri belirli kategorilere veya sınıflara ayırma sürecidir. Bu süreçte, algoritmalar verileri analiz ederek, daha önce görülmemiş verilere ilişkin sınıflandırma yapmayı öğrenirler.



SINIFLANDIRMA ÇEŞİTLERİ

Karar Ağaçları

Bu algoritma, veri noktalarını ardışık if-else kararları kullanarak sınıflandırır veya tahmin eder. Karar ağaçları, ağaç yapısında bir model oluşturur; kök düğümden başlayarak dallar ve yapraklar boyunca ilerler.

K-En Yakın Komsu

hem sınıflandırma hem de regresyon problemlerinde kullanılan basit ama güçlü bir denetimsiz öğrenme yöntemidir. Bu algoritma, yeni veri noktalarının sınıfını veya değerini belirlerken, eğitim verisindeki en yakın k komşusunu dikkate alır.

Destek Vektor Makineleri

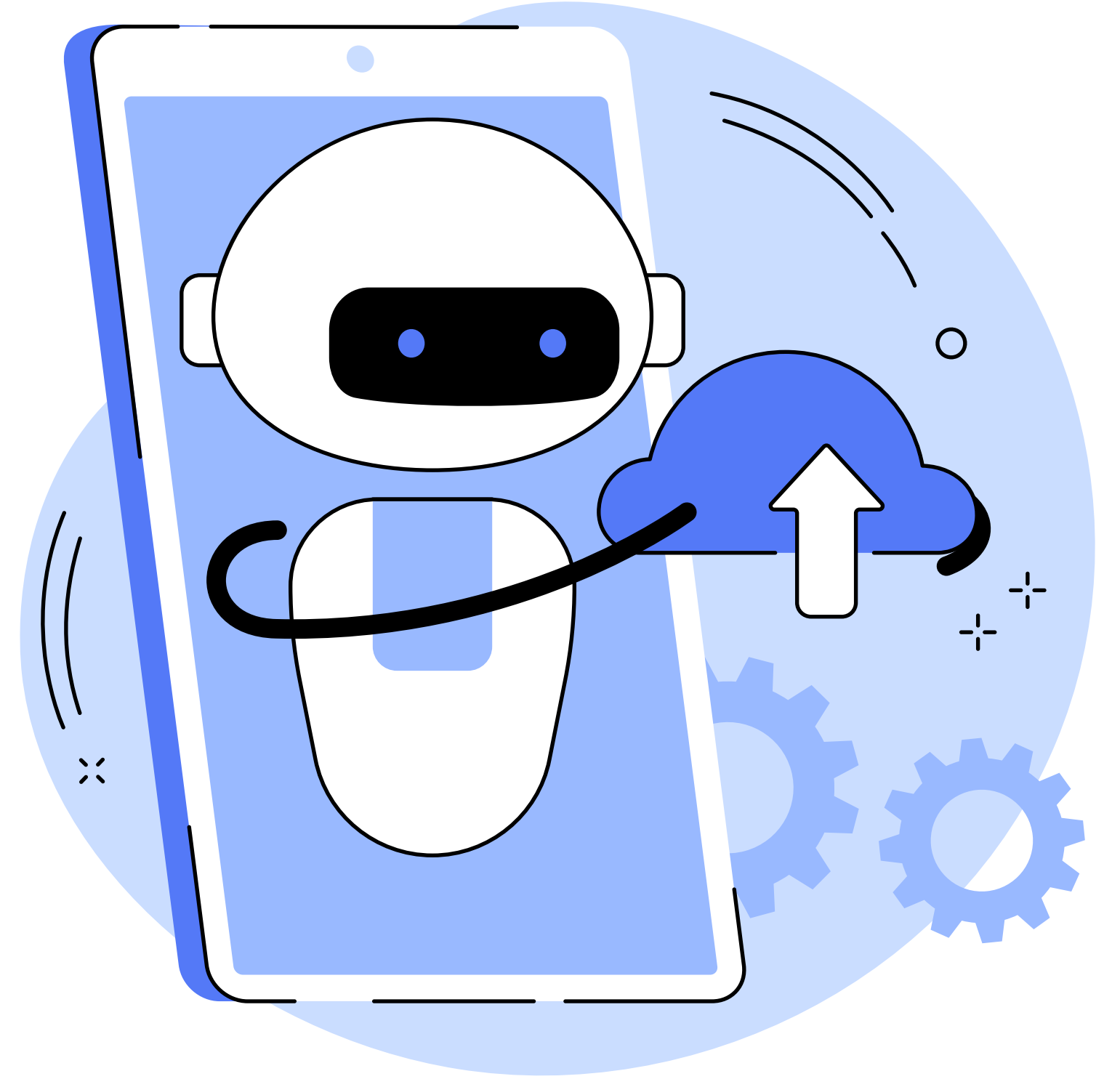
özellikle yüksek boyutlu veri setlerinde ve doğrusal olarak ayrılabilir veri setlerinde etkili bir şekilde çalışır. Temel prensibi, veri noktalarını en iyi şekilde ayıran hiper düzlemi bulmaktır.

Naive Bayes

istatistiksel bir sınıflandırma tekniği olup Bayes teoremini temel alır. genellikle gerçek dünyada doğru olmasa da, pratikte iyi performans gösterir.

SIRA DESEN MADENCİLİĐİ

Sıra desen madenciliĐi, belirli bir sırayla gelen istatiksel olarak veri örnekleri arasındaki ilgili örüntüleri bulmaya çalışır.



RASSAL ORMAN MODELİ ALGORİTMASI

Rastgele ormanlar, denetimli bir öğrenme algoritmasıdır. Hem sınıflandırma hem de regresyon için kullanılabilir. Örüntüleri bulmaya çalışır.





TEŞEKKÜRLER!