

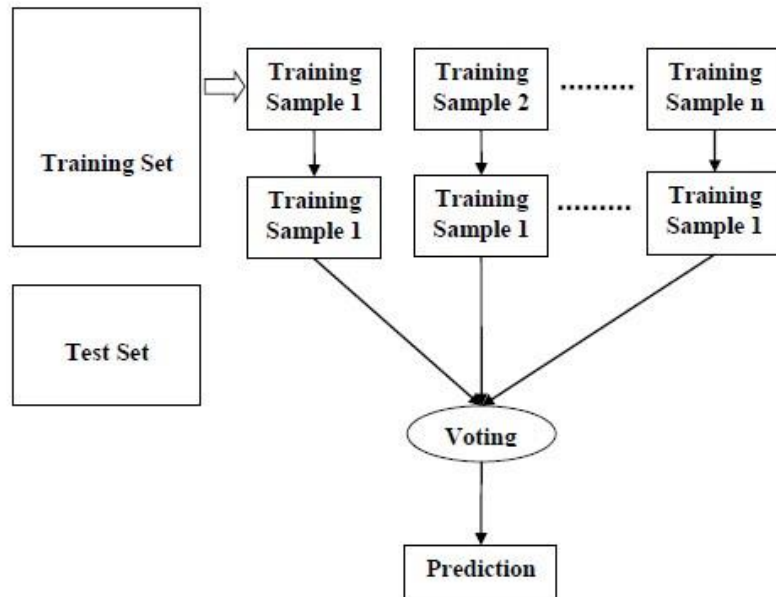
Random Forest

Random Forest, Ensemble Learning temelinde çalışır; bu, model doğruluğunu iyileştirmek ve aşırı uyumu önlemeye yardımcı olmak için farklı özellik alt kümeleriyle birden çok karar ağacı kullanmak anlamına gelir. Ensemble Learning'de burada ilgili olan iki önemli kavram vardır:

- * Torbalama (önyükleme toplama): veri kümesinden tekrar tekrar rastgele örnekler alarak varyansı azaltın. Tüm önyüklemeli veri kümeleri üzerinde eğitim alın, her küme için bir tahmin alın ve tahminlerin ortalamasını alın.
- * Güçlendirme: Ağaçlar, önceki ağaçlardan gelen bilgiler kullanılarak sırayla büyütülür. Zayıf erken öğrenicilerle (basit modellere uyan) başlayın ve hataları analiz edin: Bir girdi yanlış sınıflandırıldığında, sınıflandırma performansını iyileştirmek için ağırlığını artırın.

Aşağıdaki adımlarla bir Random Forest inşa edilebilir:

1. Önyüklenmiş bir veri kümesi oluşturun (değiştirme ile rastgele örnekleme ile örnekleme dağılımlarını tahmin edin).
2. Her adımda (düğüm) bir karar ağacı oluşturmak için önyüklemeli veri kümesinde m ($m < \text{toplam özellik sayısı}$) rastgele özellikleri seçin.
3. Çok çeşitli karar ağaçlarıyla sonuçlanan 1. ve 2. adımları tekrarlayın.
4. Her karar ağacının girdiyi sınıflandırmak için bir oyu olduğundan, her ağaçtan oyları al ve en fazla oyu olan sınıfla girdi değişkenini atarız.



Rastgele Ormanın Artıları ve Eksileri

Artıları

Aşağıdakiler Random Forest algoritmasının avantajlarıdır -

- Farklı karar ağaçlarının sonuçlarının ortalamasını alarak veya birleştirerek aşırı uyum sorununun üstesinden gelir.
- Random Forest çok çeşitli veri ögeleri için tek bir karar ağacından daha iyi çalışır.
- Random Forest, tek karar ağacına göre daha az varyansa sahiptir.
- Random Forest çok esnektir ve çok yüksek doğruluğa sahiptir.
- Random Forest algoritmasında verilerin ölçeklendirilmesi gerekmez. Ölçeklendirmeden veri sağladıktan sonra bile iyi doğruluğu korur.
- Random Forest algoritmaları, verilerin büyük bir kısmı eksik olsa bile iyi doğruluğu korur.

Eksileri

Aşağıdakiler, Random Forest algoritmasının dezavantajlarıdır - •

Karmaşıklık, Random Forest algoritmalarının ana dezavantajıdır.

- Random Forest inşası, karar ağaçlarından çok daha zor ve zaman alıcıdır.
- Random Forest algoritmasını uygulamak için daha fazla hesaplama kaynağı gereklidir.
- Karar ağaçları koleksiyonumuzun geniş olması durumunda daha az sezgiseldir.
- Random Forest kullanıldığı tahmin süreci, diğer algoritmalarla kıyasla çok zaman alıcıdır.