

## Destek Vektör Makineleri(SVM)

Destek Vektör Makineleri (SVM), sınıflandırma ve regresyon problemleri için kullanılan güçlü bir makine öğrenimi algoritmasıdır. SVM'nin çalışma prensibi, veri noktalarını en iyi şekilde ayıran bir karar sınırı (veya hiperdüzlem) bulmaya dayanır.

**Özellik Uzayı ve Karar Sınırı:** SVM, veri noktalarını sınıflandırmak veya regresyon yapmak için özellik uzayında bir karar sınırı belirler. Özellik uzayı, veri noktalarının özelliklerini temsil eden bir uzaydır. Karar sınırı, veri noktalarını iki sınıfa bölen bir hiperdüzlem veya düzlem olarak düşünülebilir.

**En İyi Karar Sınırını Bulma:** SVM, veri noktalarını en iyi şekilde ayıran bir karar sınırı bulmaya çalışır. Bu, marjı (karar sınırı ile en yakın eğitim örnekleri arasındaki mesafe) maksimize ederek yapılır. Marj, karar sınırı ile en yakın eğitim örnekleri arasındaki mesafeyi ifade eder. SVM, marjı maksimize etmek için en uygun karar sınırını bulur.

**Destek Vektörlerinin Belirlenmesi:** SVM'nin marjı belirlemede kritik öneme sahip olan örnekler, karar sınırına en yakın olan ve marjı belirleyen noktalardır. Bu örnekler, destek vektörleri olarak adlandırılır.

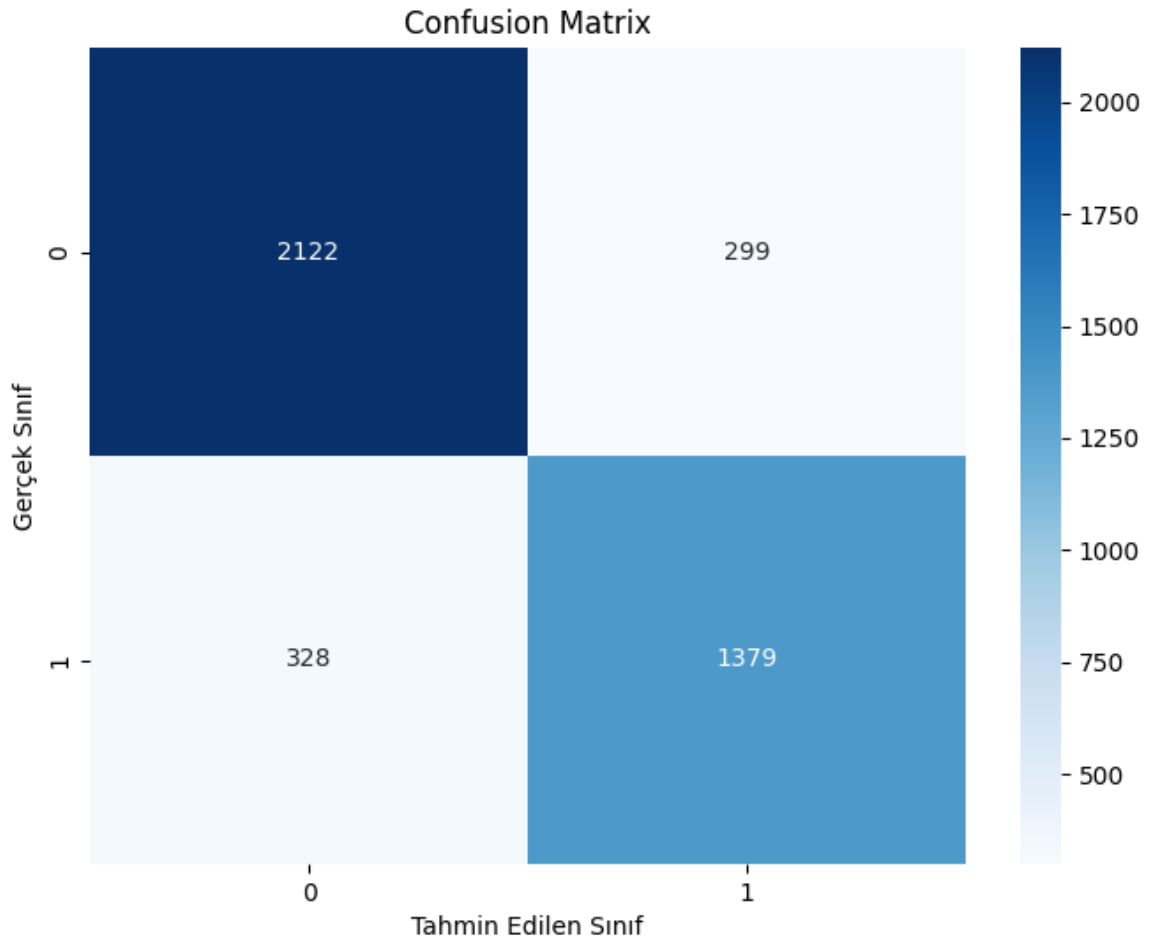
**Çekirdek Fonksiyonları ve Kernel Trick:** SVM, veri noktalarını doğrusal olarak ayıramayan durumlar için kernel trick olarak bilinen bir teknik kullanır. Bu teknik, özellik uzayını daha yüksek boyutlu bir uzaya dönüştürerek verileri doğrusal olarak ayrılabilir hale getirir. Örneğin, RBF (Radial Basis Function) çekirdeği bu amaçla sıklıkla kullanılır.

**Optimizasyon:** SVM'nin karar sınırını belirlemek için bir optimizasyon problemi olarak formüle edilir. Bu problem, veri noktalarını sınıflara en iyi şekilde ayıran bir karar sınırını bulmayı hedefler. Bu genellikle bir konveks optimizasyon problemidir ve çeşitli optimizasyon teknikleri kullanılarak çözülür.

Sonuç olarak, SVM, veri noktalarını en iyi şekilde ayıran bir karar sınırı bulmaya çalışır ve bu sınırı bulurken marjı maksimize etmeye ve destek vektörlerini kullanmaya odaklanır. Bu sayede karmaşık veri setlerinde yüksek performans sağlar ve genelleme yeteneği yüksektir.

## Uygulama

California veri kümesi kullanılarak yapılan örneğin sonuçları



Doğruluk: 0.8481104651162791