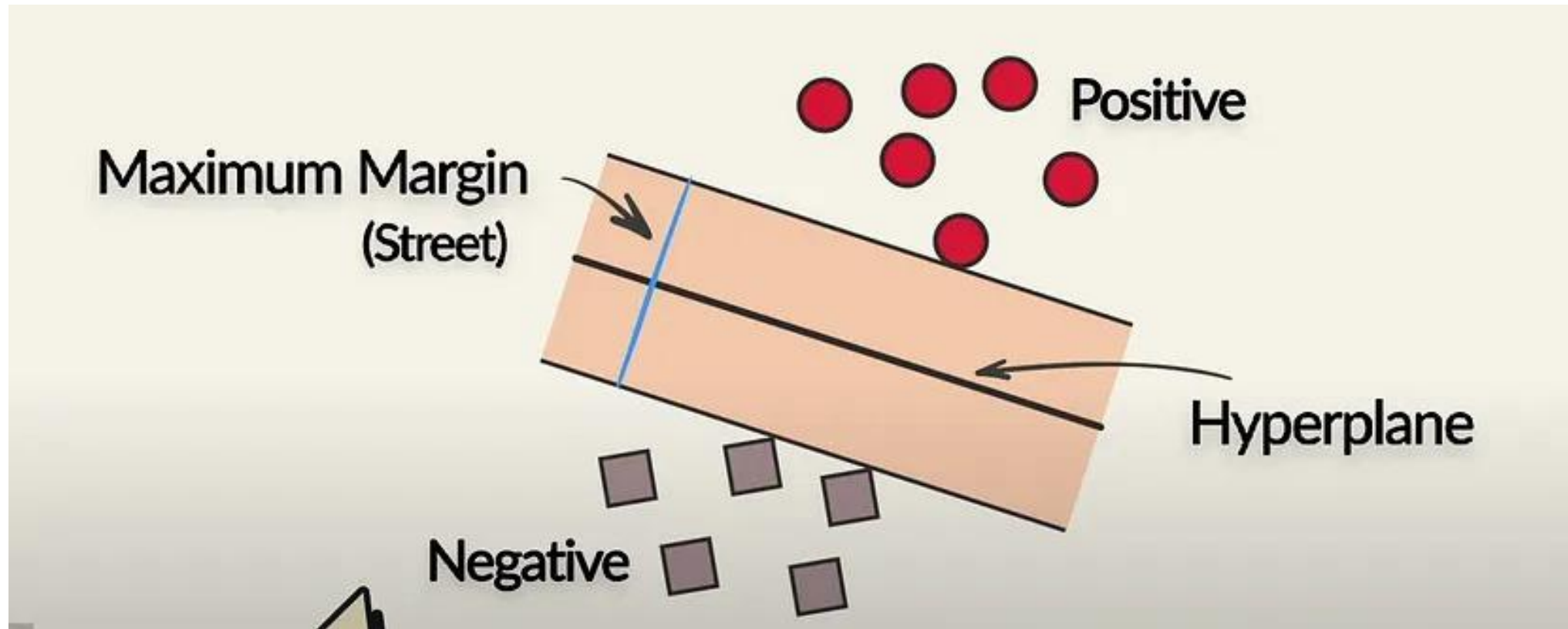
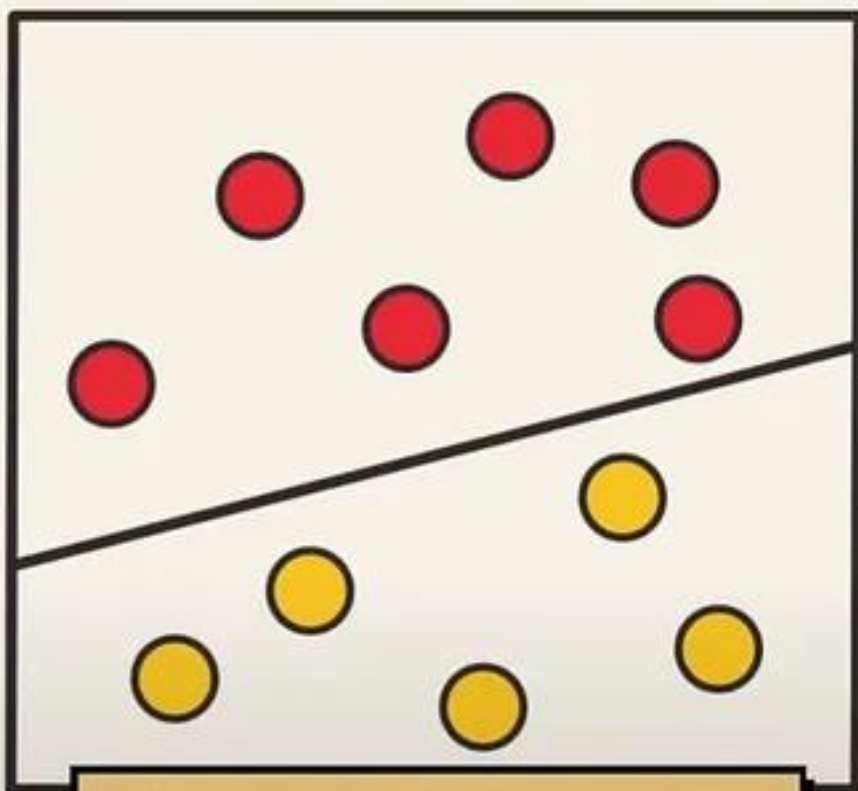


# Destek Vektör Makineleri

- Makine öğrenmesi yöntemlerinden gözetimli (supervised) öğrenme yöntemlerinden biridir. Verinin olasılıksal dağılımının bilinmediği noktalarda uygulanan bir yöntemdir.
- Destek vektör makinesinin temel amacı sınıflar arasındaki en büyük marjı bulmaktır. Destek vektör makineleri bu marjın maksimize edilmesini sağlayarak sınıflandırma problemlerini çözer. Aşağıdaki şekilde pozitif ve negatif örnekleri ayırt edebilecek birçok doğru çizilebilir. Destek vektör makineleri marjini maksimum yapan en iyi hiper düzlemi bulmaya çalıştığı için maksimum marj sınıflandırıcısı olarak da adlandırılmaktadır.

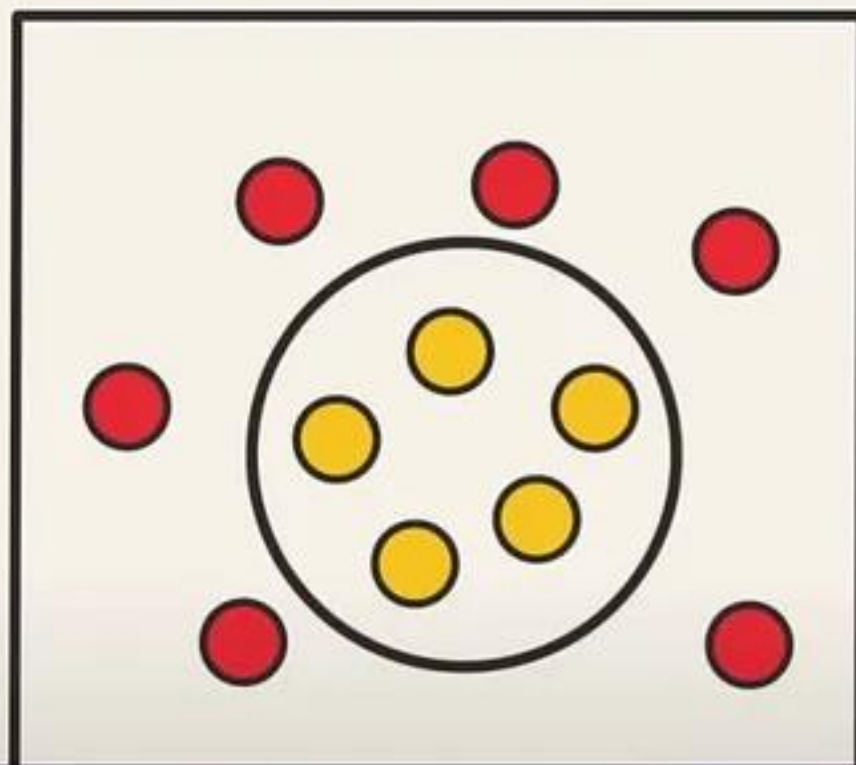


- Bu yöntemin diğer yöntemlere olan üstünlüğü; hesaplama kolaylığı, ölçeklenebilirlik, aykırı değerlere karşı dayanılabilirlik. Bu yöntem veri sayısının az olduğu durumlarda bile sınıflandırma ve regresyon problemlerinde iyi performans gösterir. Çok fazla veri olduğunda çalışabilir.
- Destek vektör makineleri doğrusal ve doğrusal olmayan şekilde 2ye ayrılmaktadır.
- Doğrusal ve doğrusal olmayan veri setlerine örnek aşağıdaki gibi verilebilir.



Linearly Separable

FIGURE 1



None-Linearly Separable

FIGURE 2

- X ve Y elemanları farklı taraflara ayırabilen bir hiperdüzlem var ise x ve y doğrusal olarak ayrılabilir. Doğrusal olarak ayrılabilen örnekler için hard marjin ve soft marjin olmak üzere 2 kısım mevcuttur.
- Gerçek hayattaki problemlerin bir çoğu doğrusal bir hiperdüzlem ile ayrılmaya uygun değildir. Destek vektör makineleri girdi uzayını daha yüksek boyutlu bir uzaya haritalar. Çekirdek fonksiyonu harita fonksiyonları yerine kullanılır. Literatürde değişik fonksiyonlar kullanılır.
- Doğrusal fonksiyon, polinomiyal fonksiyon, sigmoid fonksiyon, radyal tabanlı fonksiyonlar vb.

- Avantajları:
- Yüksek boyutlu verilerde iyi performans gösterebilir.Düşük boyutlu verilerde iyi performans gösterebilir.Aşırı uyumu engellemek için karmaşık karar sınırlarını engeller, genel yeteneklerini artırır.Doğrusal olmayan yöntemlerde uygulanabilir.Aşırı uyumu engellemek için parametre kullanır.Sayısal ve kategorik türde verilerle kullanılabilir.Probleme bağlı olarak yukarıda yazılan fonksiyonları kullanabilir.Aykırı değerlere karşı direnci yüksektir ve sonuçları etkilemez.

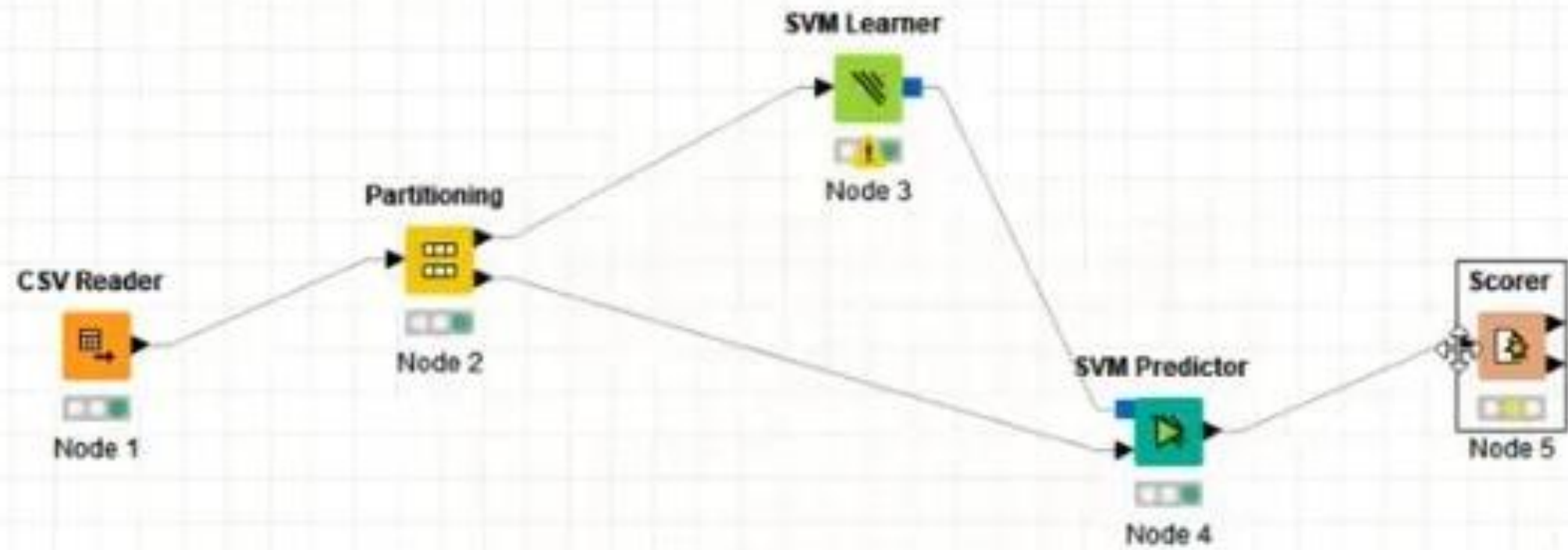
- Dezavantajları:
- Hesaplama yükü kimi durumlarda fazla olabilir ve büyük verilerde uygun olmadığı durumlar gözlenmiştir. Eğitim süreleri uzun olabilir bunun nedeni modelin karmaşık olduğu durumlardır. Matematiksel olarak algılanması zor sınırlar üretebilir. Yanlış hiperparametre seçimi modelin performansını olumsuz anlamda etkileyebilir.
- Kullanım alanları:
- Spam filtreleme, el yazısı tanıma, görüntü sınıflandırma, duygu analizi, hisse senedi fiyat tahminleri, talep tahminlemesi, fiyat tahminlemesi, nesne tanıma, ilaç keşfi, protein fonksiyonu tahmini, gen ifadesi analizi, metin kategorizasyonu, nlp, finansal sahtekarlık tespiti gibi birçok farklı alanda kullanım örnekleri bulunmaktadır.



- Python ile örnek destek vektör makinesi uygulaması
- ilgili kütüphaneler yüklenir.
- `from sklearn import datasets`
- `from sklearn.model_selection import train_test_split`
- `from sklearn.svm import SVC`
- `from sklearn.metrics import classification_report, accuracy_score`

- #pythonda var olan iris data seti kullanılabilir.
- iris = datasets.load\_iris()
- X = iris.data
- y = iris.target
- #veriler eğitim ve test verisi şeklinde ayrılır.
- X\_train, X\_test, y\_train, y\_test =
- train\_test\_split(X, y, test\_size=0.3, random\_state=42)
- #model oluşturulur
- svm\_model = SVC(kernel='linear')
- # Kernel türünü seçebilirsiniz (linear, rbf, poly, vb.)
- svm\_model.fit(X\_train, y\_train)

- #model performansı değerlendirilir.
- `y_pred = svm_model.predict(X_test)`
- #farklı performans metrikleri kullanılarak modelin başarısı ölçülür.
- `accuracy = accuracy_score(y_test, y_pred)` `print("Accuracy:", accuracy)`
- `classification_rep = classification_report(y_test, y_pred)`  
`print("Classification Report:\n", classification_rep)`



- Knime programı ile destek vektör makineleri yukarıdaki gibi uygulanabilir.
- Csv reader ile veri okunur. Partitioning node ile eğitim ve test verisi şeklinde veri ayrılır. SVM Learner node ile model oluşturulur. SVM Learner sağ tık yapılarak fonksiyon seçimi ve parametre seçimleri yapılır. SVM Predictor ile sınıflandırma yapılır ve son olarak scorer node ile modelin başarısı ölçülür.