

# Lineer Sınıflandırma

Fuat Can Beylunioglu

December 23, 2017

# Giriş

- ▶ Lineer sınıflandırma, veriyi doğrusal bir çizgi ile ayırmayı amaçlar.
- ▶ Bazı durumlarda veri tamamen ayrılabilir ancak bunun mümkün olmadığı durumlar da söz konusudur.
- ▶ Bu gibi durumlarda belli bir hata payı göze alınabilir, ya da lineer olmayan yöntemler kullanılabilir.

# Lineer Sınıflandırıcılar, Lojistik Regresyonlar

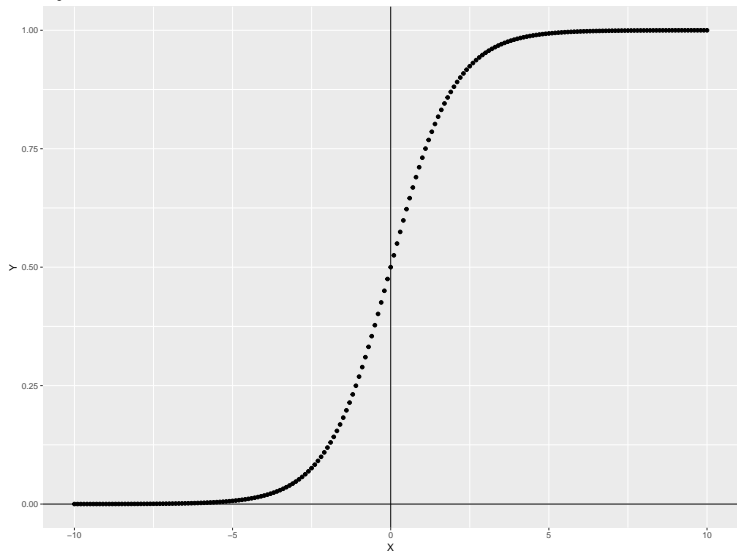
- ▶ Lojistik regresyonlar lineer regresyon modeli üzerine kuruludur ve non-lineer uygulamaları yaygın değildir.

Model denklemini aşağıdaki gibidir:

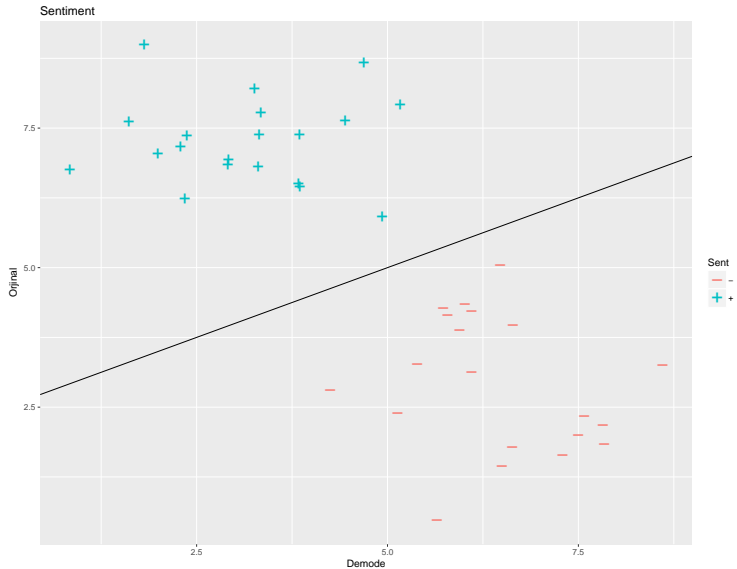
$$Y_i = \phi(\beta_0 X_{0i} + \beta_1 X_{1i} + \cdots + \beta_n X_{ni}) \quad (1)$$

- ▶ Bu denkleme göre  $X_i$ 'deki herhangi değere karşılık  $Y_i$  her zaman 0 ve 1 arası değer üretecektir.
- ▶ Bu yüzden  $Y_i$  bir olasılık olarak yorumlanır.
- ▶ Öte yandan  $X_i$  de *score* olarak tanımlanır.

Logit Function



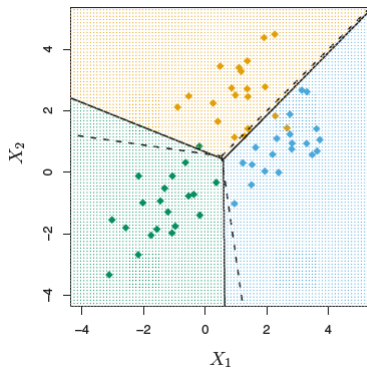
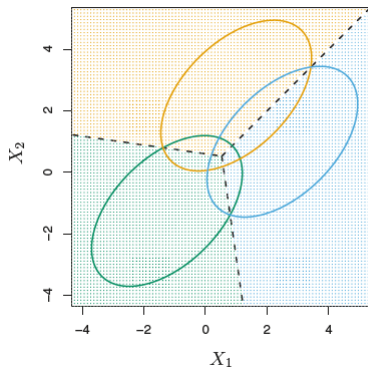
# Sentiment Örneği



# Linear Discriminant Analysis

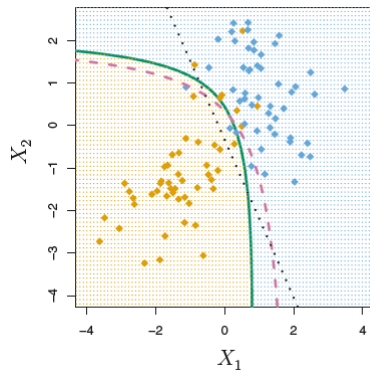
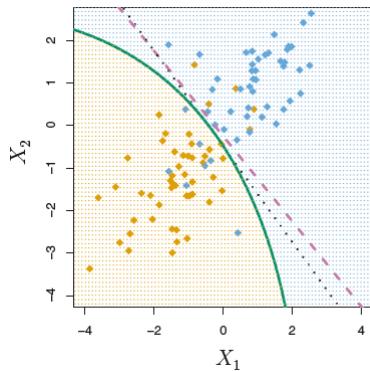
- ▶ LDA ve lojistik regresyonlar bağımsız değişkenlerin koşullu dağılımının normal olduğunu varsayar.
- ▶ Lojistik regresyon bu varsayımı hata payına yüklerken LDA bu varsayıma dayanarak veriyi sınıflara böler.
- ▶ LDA, lojistik regresyonlara benzer şekilde çalışır ve normal dağılım varsayımı sağlandığında benzer sonuçlar verir.
- ▶ Ayrıca ayırım yapılacak sınıf sayısı 2'den fazla olduğu durumlara daha uygundur.
- ▶ Ancak birçok durumda bunu elde etmek güçtür.
- ▶ Eğer sınıflar birbirinden ayrık ise lojistik regresyonların tersine LDA güvenilir (varyansı düşük) sonuçlar vermektedir.
- ▶ Ancak sınıflar lineer bir şekilde ayrımlanmamışsa o zaman iki yaklaşım da kötü sonuç verecektir. Bu durumda Quadratic Discriminant Analysis kullanılabilir

# LDA, 3 Sınıflı Veri



**Kaynak:** James vd. (2013)

## QDA, 2 Sınıflı Veri



**Kaynak:** James vd. (2013)



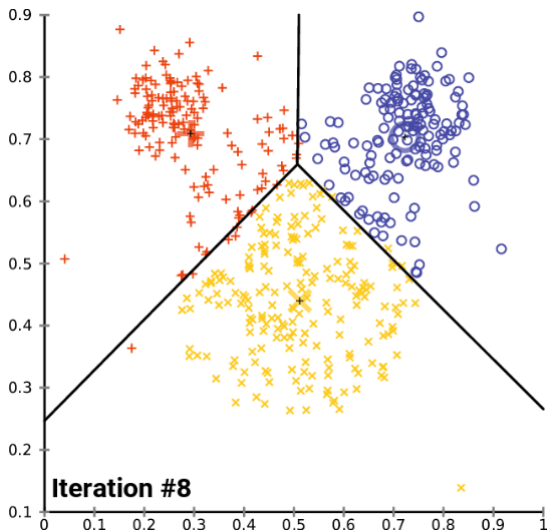
# KNN Sınıflandırma

- + En basit sınıflandırma yöntemlerinden biridir.
- + Lineer olmayan karar sınırı olduğunda başarısı daha yüksektir.
  - Ancak KNN hangi bağımlı değişkenin daha açıklayıcı olduğu hakkında bir bilgi vermez.
- ▶ Bir gözlem hakkında tahminde bulunmak istendiğinde, elde bulunan veri seti arasından gözleme en benzer (yakın) olan değeri verir.
- ▶ Uzaklık, tek bağımsız değişken olduğunda  $\text{dist} = X_i - X_{\text{train}}$  olurken birden çok bağımsız değişken olduğu durumda  $\text{dist} = f(X_i, X_{\text{train}})$  olmaktadır.
- ▶ Burada distance fonksiyonu Euclidean, Manhattan, Cosine, vs olabilir.

# K-means Clustering

- ▶ Unsupervised learning olarak tanımlanır.
- ▶ Sınıflandırılmamış veriyi gruplara ayırır.
- ▶ Kaç gruba ayıracağı bilgisi dışarıdan verilir.
- ▶ K grup bulunması istendiğinde algoritma her grubun merkezi olan bir gözlem bulmaya çalışır. Bu gözlemlere centroid adı verilir.
- ▶ İteratif bir yöntemdir:
  1. K tane centroid (veri noktası) rastgele seçilir.
  2. Her bir veri en yakın olduğu  $p_k$  noktasına bağlı olarak  $p_k$ 'nın içinde olduğu gruba yazılır.
  3. Her grubun ortalaması alınır. Centroid'ler bu ortalamalarla değiştirilir ve 2. ve 3. aşamalar tekrar edilir. Eğer bir önceki ortalama ile bir sonraki ortalama arasında fark belli bir seviyenin altındaysa algoritma durur.

# K-means Kümeleme

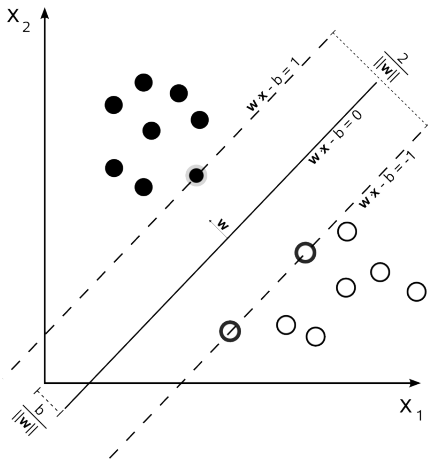


**Kaynak:** [https://www.wikipedia.org/en/K-means\\_clustering](https://www.wikipedia.org/en/K-means_clustering)

# Support Vector Machine

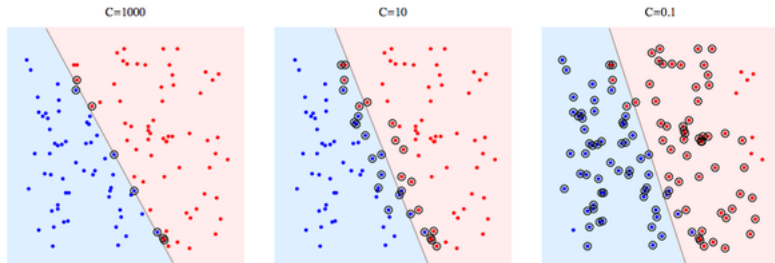
- ▶ SVM algoritması matematiksel optimizasyona dayanır.
- ▶ İki grubu birbirinden ayıran fonksiyonu bulmayı hedefler.
- ▶ Bir alt grubu olan Maximal Margin buna verilebilecek örneklerdendir.
- ▶ Bu yöntemeye göre iki grubu birbirinden ayıran sonsuz fonksiyondan aralarındaki boşluğu maksimize eden seçilir.
- ▶ Ancak iki grubu birbirinden tamamen ayıran bir fonksiyon olmadığı durumda ceza fonksiyonunu minimize eden seçilir.

# Maximal Margin Classifier



**Kaynak:** <http://francescopochetti.com/support-vector-machines/>

# SVM

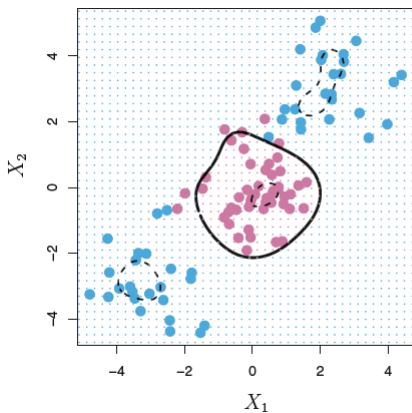
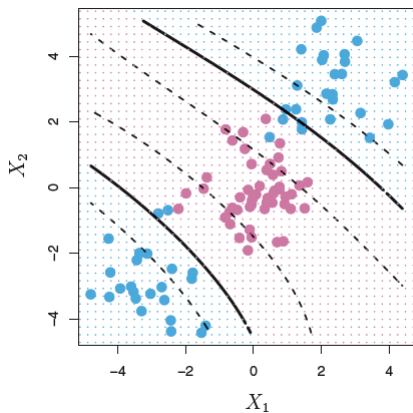


**Kaynak:** <https://www.quora.com/What-are-C-and-gamma-with-regards-to-a-support-vector-machine>

# Support Vector Machine, Kernel

- ▶ Yukarıdaki örnekler SVM'in lineer kernel kullanan çeşitleridir.
- ▶ Ancak bazı durumlarda lineer kerneller grubu ayırt etmek için yeterli değildir.
- ▶ Bu gibi durumlarda polinomial ya da radial kernel kullanılabilir.

# SVM



**Kaynak:** James vd. (2013)



# Kaynakça I



James, Gareth and Witten, Daniela and Hastie, Trevor and Tibshirani, Robert,  
*An introduction to statistical learning*,  
2013, Springer.