4.5 Naive Bayes

Naïve Bayes Sınıflandırıcı adını İngiliz matematikçi Thomas Bayes'ten (yak. 1701 – 7 Nisan 1761) alır. Naïve Bayes Sınıflandırıcı Örüntü tanıma problemine ilk bakışta oldukça kısıtlayıcı görülen bir önerme ile kullanılabilen olasılıkcı bir yaklaşımdır. Bu önerme örüntü tanıma da kullanılacak her bir tanımlayıcı öznitelik ya da parametrenin istatistik açıdan bağımsız olması gerekliliğidir. Her ne kadar bu önerme Naive Bayes sınıflandırıcının kullanım alanını kısıtlasa da, genelde istatistik bağımsızlık koşulu esnetilerek kullanıldığında da daha karmaşık Yapay sinir ağları gibi metotlarla karşılaştırabilir sonuçlar vermektedir. Bir Naive Bayes sınıflandırıcı, her özniteliğin birbirinden koşulsal bağımsız olduğu ve öğrenilmek istenen kavramın tüm bu özniteliklere koşulsal bağlı olduğu bir Bayes ağı olarak da düşünülebilir.

Bayes Teoremi
Naive Bayes sınıflandırıcısı Bayes teoreminin bağımsızlık önermesiyle basitleştirilmiş halidir. Bayes teoremi aşağıdaki denklemle ifade edilir;

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

P(A|B) ; B olayı gerçekleştiği durumda A olayının meydana gelme olasılığıdır
P(B|A) ; A olayı gerçekleştiği durumda B olayının meydana gelme olasılığıdır
P(A) ve P(B) ; A ve B olaylarının önsel olasılıklarıdır.
Burada önsel olasılık Bayes teoreminine öznellik katar. Diğer bir ifadeyle örneğin P(A) henüz elde veri toplanmadan A olayı hakkında sahip olunan bilgidir. Diğer taraftan P(B|A) ardıl olasılıktır çünkü veri toplandıktan sonra, A olayının gerçekleşmiş olduğu durumlarda B olayının gerçekleşme ihtimali hakkında bilgi verir.

Sınıflandırma Problemi
Naive Bayes Sınıflandırması Makine öğreniminde öğreticili öğrenme alt sınıfındadır. Daha açık bir ifadeyle sınıflandırılması gereken sınıflar(kümeler) ve örnek verilerin hangi sınıflara ait olduğu bellidir. E-posta kutusuna gelen e-postaların spam olarak ayrıştırılması işlemi buna örnek verilebilir. Bu örnekte spam e-posta ve spam olmayan e-posta ayrıştırılacak iki sınıfı temsil eder. Elimizdeki spam ve spam olmayan e-postalardan yaralanarak gelecekte elimize ulaşacak e-postaların spam olup olmadığına karar verecek bir Algoritma da öğreticili makina öğrenmesine örnektir.
Sınıflandırma işleminde genel olarak elde bir örüntü (pattern) vardır. Buradaki işlem de bu örüntüyü daha önceden tanımlanmış sınıflara sınıflandırmaktır. Her örüntü nicelik (feature ya da parametre) kümesi tarafından temsil edilir.

Nicelik Kümesi
Yine yukarıda bahsedilen spam e-posta örneğinden devam edilecek olunursa; Posta kutumuzda bulunan spam e-postaları spam olmayan e-postalardan ayıran parametrelerden oluşan bir küme, mesela ikramiye, ödül gibi sözcüklerden oluşan, nicelik kümesine örnektir. Matematiksel bir ifadeyle nicelik kümesi;

$$x(i), i = 1, 2, \ldots, L$$

ise

$x = [x(1), x(2), \ldots, x(L)]^T$ $\in R^L$ L-boyutlu nicelik vektörünü oluşturur.x $\in R^L$ verildiğine göre ve S ayrıştırılacak sınıflar kümesiyse, Bayes teoremine göre aşagıdaki ifade yazılır.

$$P(S_i|x) \times p(x) = p(x|S_i) \times P(S_i)$$

ve

$$p(x) = \sum_{i=1}^{L} p(x|S_i)P(S_i)$$

- P(Si); Si'nin öncel olasılığı i = 1, 2, . . . , L,
- P(Si|x); Si'nin ardıl olasılığı
- p(x); x in Olasılık yoğunluk fonksiyonu (oyf)
- p(x|Si); i =1 = 2, . . . , L, x'in koşullu oyf'si

Bayes Karar Teoremi

Elimizde sınıfı belli olmayan bir örüntü olsun. Bu durumda

$$x = [x(1), x(2), \ldots, x(L)]^T$$

sınıfı belli olmayan örüntünün L-boyutlu nicelik vektörüdür. Spam e-posta örneğinden gidecek olursak spam olup olmadığını bilmediğimiz yeni bir e-posta sınıfı belli olmayan örüntüdür.

Yine Si x'in atanacağı sınıf ise;

Bayes karar teorisine göre x sınıf Si'ya aittir eğer

$$P(S_i|x) > P(S_j|x)$$

diğer bir ifadeyle eğer

$$P(x|S_i)P(S_i) > P(x|S_j)P(S_j)$$

Naive Bayes Sınıflandırma

Verilen bir x'in (x = [x(1), x(2), . . . , x(L)]T $\in$ RL) sınıf Si'ye ait olup olmadığına karar vermek için kullanılan yukarıda formüle edilen Bayes karar teoreminde istatistik olarak bağımsızlık önermesinden yararlanılırsa bu tip sınıflandırmaya Naive bayes sınıflandırılması denir. Matematiksel bir ifadeyle

$$P(x|S_i)P(S_i) > P(x|S_j)P(S_j)$$

ifadesindeki

$P(x|S_i)$ terimi yeniden aşağıdaki gibi yazılır

$$P(x|S_i) \approx \prod_{k=1}^{L} P(x_k|S_i)$$

böylece Bayes karar teoremi aşagıdaki şekli alır. Bayes karar teorisine göre x sınıf Si'ya aittir eğer

$$P(S_i) \prod_{k=1}^{L} P(x_k|S_i) > P(S_j) \prod_{k=1}^{L} P(x_k|S_j)$$

$P(S_i)$ ve $P(S_j)$) i ve j sınıflarının öncel olasılıklarıdır. Elde olan veri kümesinden değerleri kolayca hesaplanabilir.

Naive bayes sınıflandırıcının kullanım alanı her ne kadar kısıtlı gözükse de yüksek boyutlu uzayda ve yeterli sayıda veriyle x'in (nicelik kümesi) bileşenlerinin istatistik olarak bağımsız olması koşulu esnetilerek başarılı sonuçlar elde edilebilinir.

Uygulama Alanları
Naive Bayes sınıflandırıcısı genel olarak veri madenciliğinde, biyomedikal mühendisliği alanında, hastalıkların ya da anormalliklerin tıbbi tanımlanmasında (otomatik olarak mühendislik ürünü tıbbi cihazlar tarafından tanı konulması) ,elektrokardiyografi (EKG) grafiğinin sınıflandırılmasında, elektroensefalografi (EEG) grafiklerinin ayrıştırılmasında, genetik araştırmalarında, yığın mesaj tanımlanmasında, metin ayrıştırılmasında, ürün sınıflandırma ve diğer bazı alanlarda kullanılır.