Bu yazının amacı, literatürde naif Bayes (Naive Bayes) olarak bilinen algoritmanın, sınıflandırma (classification) için kullanılma şeklini açıklamaktır. Herhangi bir sınıflandırma probleminde olduğu gibi, amacımız birden fazla özelliği taşıyan bir yöney (vektör) kullanarak verilen bilgilerden bir eğitim oluşturmak ve bu eğitim neticesinde gelen yeni verileri doğru bir şekilde sınıflandırmaktır.

Sınıflandırma işlemi sırasında, nasıl bir yol izlendiğini bir örnek üzerinden açıklamaya çalışalım.

Örneğimiz, aşağıda verilen tablodaki verilerden bir sınıflandırıcı eğitimi olsun (train).

Yukarıdaki ilk iki satır ortalama ve ikinci iki satır ise varyans değerleridir. Bu değerleri basitçe Excel ile hesapladık.

Şimdi beklenen değeri hesaplayacağız. Yani gelen test verimizin Muhasebe kısmında birisine ait olması veya Yazılım kısmından birisine ait olması için beklenen durum hesabı yapacağız.

Hesaplamaya geçmeden önce yazının konusu olan naive bayes kavramını hızlıca açıklayalım. Aslında naive bayes sınıflandırıcısı basiçte bütün koşullu olasılıkların çarpımıdır.

Aşağıdaki şekilde gösterilebilir: