

2017-2018 Güz Yarıyılı Makine Öğrenmesi 1. Ödevi

Konu : Naive Bayes Yöntemi ile Doküman Sınıflandırma

Amaç : Bu ödevde Naive Bayes yönteminin **doküman sınıflandırmadaki** başarısı değerlendirilecektir. Seçeceğiniz bir tweet veri setini kullanarak, Naive Bayes ile tweet'leri içeriğine göre sınıflandırınız.

Teslim Edilecekler:

1. Rapor : Giriş, uygulama ve sonuç bölümlerinden oluşan makale düzeninde bir rapor hazırlayınız.

- **Giriş** bölümünde yöntemi anlatınız.
- **Uygulama** bölümünde
 - seçtiğiniz konu, sınıf sayısı ve bilgileri, sözlükteki kelime sayısı, sınıflardaki kelime sayısı gibi bilgileri veriniz.
 - sınıflandırma başarısını eğitimden kullandığınızdan **farklı** test örnekleri ile hesaplayınız.
 - Test örneklerinin sınıflandırma sonuçlarını karışıklık (confusion) matrisi ile gösteriniz.
 - Başarılı ve başarısız sınıflandırmalara örnek 5'er tweet vererek başarısız sınıflandırma ile ilgili yorum yapınız.
- **Sonuç** bölümünde, Naive Bayes Sisteminin tweet sınıflandırma başarısını genel olarak değerlendiriniz.

2. Kaynak Kod : Yazdığınız programların **kaynak kodlarını** açıklamalı olarak hazırlayarak veriniz. Kaynak kodda Naive Bayes'e ait hazır kütüphane kullanmayınız, kodu **kendiniz** yazınız ☺. İnternette bulunan bir koda en az %60 benzeyen kodlar değerlendirilmeyecektir.

Teslim edilecekleri tek bir pdf dosya içinde birleştirip dosyanın ismini okul numaranız ve ödev numarası olarak veriniz. (17501002hw1.pdf gibi).

Teslim Tarihi: Ödevinizi 29 Ekim 2017 Pazar 23.59'a kadar e-maile gönderiniz. Geciken hergün 10 puan eksiltilerek değerlendirilir.