

## DOĞAL DİL İŞLEME DERSİ 15.12.2022 TARİHLİ ÖDEVİ

Bu ödevde birincil hedefiniz, Laplace smoothing ve bag of words Naive Bayes sınıflandırıcısı uygulamak ve performansını birkaç veri kümesi üzerinde değerlendirmektir. Teorik arka planı 12. Hafta ders notlarında bulunmaktadır. Bu ödev için iskelet kod ekteki dosyada bulunmaktadır.

Kullanacağınız birincil veri kümesi, 2010'daki Haiti Depreminde toplanan metin mesajlarının bir külliyatı olacak ve sınıflandırma görevi, mesajın yardım isteyip istemediğini (alakalı - *relevant*) veya olmadığını (alakasız - *irrelevant*) belirlemek olacaktır. Bu veriler, [Mission 4636](#) adlı kitle kaynak kullanımı ve metin sınıflandırmasının gerçek dünyadaki bir uygulamasında toplanmıştır ve bununla ilgili daha fazla bilgiyi ilgili [bu makalede](#) okuyabilirsiniz. Ana kodunuzda `NaiveBayesClassifier` sınıfı üzerinde çalışacaksınız. `naive_bayes.py` dosyasını çalıştırarak programınızı test edebilirsiniz. Bu kodda cevabınızı girmeniz gereken yerler belirtilmiştir. **Yapmanız gerekenler:**

### `train`

`TrainNaiveBayes` kabakodunu kullanarak alakalı sınıf değişkenlerini güncelleyin: `self.vocabulary`, `self.logprior`, `self.loglikelihood`. `self.loglikelihood`, `w` kelime için string ve `c` sınıf (class) için string olmak üzere `(w, c)` şeklinde tuples formunda olmalı.

### `score`

Verilen doküman için toplanmış log-probability değerini döndürün. Return the summed log-probability for a given document; bu kısım `TestNaiveBayes` kabakodundaki for-loop kısmına benzer.

### `predict`

Verilen doküman için en muhtemel sınıfı döndürün; bu kısımda muhtemel sınıflar üzerinden dönerek (loop) yazdığınız `score` fonksiyonunu kullanın.

### `evaluate`

Modelinizin test verisi üzerindeki performansını (precision, recall, F1 score) hesaplayın.

GitHub'ta ders için oluşturduğunuz depoya ayrıntılı açıklama içeren bir video ile ekleyin. UZEM'e linkini gönderin.

Son teslim tarihinden sonra gönderilen ödevler hiçbir şartta dikkate alınmayacaktır.

**Son teslim tarihi:** 10 Ocak 2023 Salı 23:50

**Başarılar.**

Dr. Öğr. Üyesi Ayşe Saliha Sunar