

Ödev2

Ayşe Serra Şimşek
asimsek2019@gtu.edu.tr

²Elektronik Mühendisliği Bölümü, GTÜ, Kocaeli, Türkiye

I. GİRİŞ

Bu çalışmanın konusu, Spam e-posta olarak işaretlenmiş veri setiyle çalışarak Naif Bayes sınıflandırıcının kodlanması ve algoritmanın başarımının gösterilmesidir. Başarımın değerlendirilmesindeki metriklerin kullanımı açıklanmıştır.

II. TEORİK BİLGİ

Naive Bayes sınıflandırıcı, makine öğrenimi alanında sınıflandırma problemlerini çözmek için kullanılan olasılığa dayalı bir algoritmadır. Bu algoritma, Bayes Teoremi'ne dayanır ve özellikle metin sınıflandırması gibi problemlerde etkilidir.

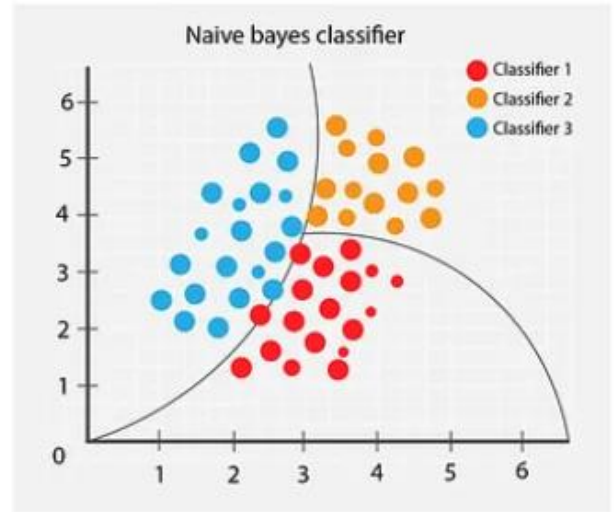
Naive Bayes, sınıflandırma yaparken her bir özelliğin (veya özneliliğin) sınıf etiketi üzerindeki etkisini bağımsız olduğunu varsayar. Bu da "naive" (saf veya basit) adının algoritmanın bu varsayıma dayanmasından gelir. Gerçekte birçok durumda özellikler arasında bağımlılıklar olabilir, ancak Naive Bayes algoritması basitliği ve hesaplama kolaylığı nedeniyle tercih edilir.

Özellikle metin sınıflandırması gibi problemlerde kullanılır. Bu ödevde de olduğu gibi bir e-postanın spam veya spam olmayan olarak sınıflandırılması gibi metin verilerinin sınıflandırılması için tercih edilir.

Naive Bayes sınıflandırıcı, bir veri setindeki özelliklerin belirli bir sınıf etiketiyle ilişkisini öğrenir ve bu bilgiyi kullanarak yeni, daha önce görmediği örneklerin sınıflandırılmasını gerçekleştirir. Öğrenme sürecinde, sınıflandırıcı her bir sınıf için özelliklerin olasılığını hesaplar. Ardından, bu olasılıkları kullanarak yeni bir veri noktasının hangi sınıfa ait olduğunu tahmin eder.

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad (1)$$

Makine öğrenmesinde sınıflandırıcılarda kullanılan "Confusion Matrix", "Accuracy", "Precision", "Recall", "F1 Score" metrikleri sınıflandırma modelinin performansını değerlendirmek için kullanılır. Confusion Matrix, sınıflar arasındaki doğru ve yanlış tahminleri gösterirken, Accuracy, Precision, Recall ve F1 Score, modelin başarısını daha ayrıntılı bir şekilde analiz etmek için kullanılır. Bu metrikler, modelin hangi durumlarda ne kadar doğru çalıştığını anlamak için birbirlerini tamamlar.



Şekil 1. Naive Bayes Sınıflandırıcı

III. ÇALIŞMA

- Öncelikle, `pd.read_csv()` kullanılarak 'spam_ham_dataset.csv' dosyası okunmuş ve 'data' adlı bir DataFrame'e yüklenmiştir.
- Veri seti, verinin %80'ini eğitim setine ve %20'sini test setine ayıracak şekilde bölünmüştür.
- Eğitim seti içinde 'ham' ve 'spam' etiketlerine sahip mailleri ayırmak için `train_data` DataFrame'inde bu filtreleme işlemleri yapılmıştır.
- Eğitim setindeki 'ham' ve 'spam' maillerin olasılıkları hesaplanır. Bunun için Bayes Teoremi kullanılmıştır. 'ham'

ELM472 Makine Öğrenmesinin Temelleri

ve 'spam' maillerinde geçen kelimelerin olasılıkları hesaplanmıştır.

- Daha sonra, verilen bir metni sınıflandırmak için Naive Bayes algoritmasını kullanan “predict()” fonksiyonu oluşturulmuştur. Oluşturulan predict() fonksiyonu, test setindeki metinleri sınıflandırmak için kullanılmıştır ve tahminler test_data DataFrame'ine 'predicted' sütununda saklanmıştır.

- Son olarak, gerçek ve tahmin edilen değerler arasındaki karşılaştırmayı gösteren bir Confusion Matrix oluşturulmuştur ve Confusion Matrix'ten elde edilen değerler kullanılarak Accuracy, Precision, Recall ve F1 Score hesaplanır.

IV. SONUÇLAR

Python kodu ile oluşturulan modelin çıktısı Şekil.2'deki gibidir.

```
<bound method NDFrame.describe of      Unnamed: 0 label
0          605   ham  Subject: enron methanol ; meter # : 988291\r\n...
1          2349   ham  Subject: hpl nom for january 9 , 2001\r\n( see...
2          3624   ham  Subject: neon retreat\r\nho ho ho , we ' re ar...
3          4685   spam Subject: photoshop , windows , office . cheap ...
4          2030   ham  Subject: re : indian springs\r\nthis deal is t...
...          ...
5166         1518   ham  Subject: put the 10 on the ft\r\nthe transport...
5167         404   ham  Subject: 3 / 4 / 2000 and following noms\r\nhp...
5168         2933   ham  Subject: calpine daily gas nomination\r\n>\r\n...
5169         1409   ham  Subject: industrial worksheets for august 2000...
5170         4807   spam Subject: important online banking alert\r\ndea...

label_num
0          0
1          0
2          0
3          1
4          0
...
5166         0
5167         0
5168         0
5169         0
5170         1

[5171 rows x 4 columns]>
```

```
Confusion Matrix:
Predicted ham spam
Actual
ham          702    18
spam          6    309
```

```
Accuracy: 0.9768115942028985
Precision: 0.9915254237288136
Recall: 0.975
F1 Score: 0.9831932773109243
```

- Gerçek ham maillerin 720'si doğru bir şekilde tahmin edilmiştir.

- Gerçek spam maillerin 315'i doğru bir şekilde tahmin edilmiştir.

- 24 mail yanlış sınıflandırılmıştır (18 ham mail spam olarak, 6 spam mail ham olarak).

- Modelin doğruluk oranı (accuracy) yaklaşık olarak **%97.68**'dir. Genel doğruluğu oldukça yüksektir.

- 'ham' sınıfı için kesinlik, 'ham' olarak tahmin edilenlerin gerçekten 'ham' olma oranıdır.'ham' sınıfı için kesinlik (precision) yaklaşık olarak **%99.15**'tir. Bu, 'ham' olarak tahmin edilen maillerin büyük çoğunluğunun gerçekten 'ham' olduğunu gösterir.

- 'ham' sınıfı için duyarlılık (recall), gerçek 'ham' maillerin ne kadarının 'ham' olarak doğru bir şekilde tahmin edildiğini gösterir. oran yaklaşık **%97.5**'tir. Bu, 'ham' olan maillerin büyük çoğunluğunun doğru bir şekilde tahmin edildiğini gösterir.

- F1 skoru, kesinlik ve duyarlılığın harmonik ortalamasıdır. Dengeli bir performans ölçüsü olarak kullanılır. F1 skoru yaklaşık olarak **%98.32**'dir. Bu, modelin 'ham' ve 'spam' sınıflarını dengeli bir şekilde tahmin ettiğini gösterir.

KAYNAKÇA

- [1] <https://towardsdatascience.com/introduction-to-na>
- [2] https://tr.wikipedia.org/wiki/Naive_Bayes
- [3] <https://mlarchive.com/machine-learning/the-ultimate-guide-to-naive-bayes/>
- [4] <https://medium.com/@ekrem.hatipoglu/machine-learning-classification-naive-bayes-part-11-4a10cd3452b4>
- [5] E. Alpaydin, Introduction to Machine Learning, 3. bs. Cambridge, MA, USA: MIT Press, 2014.

Şekil 2. Metrik Değerleri