

1906003132015

Doğal Dil İşleme

BAİBÜ Bilgisayar Müh.

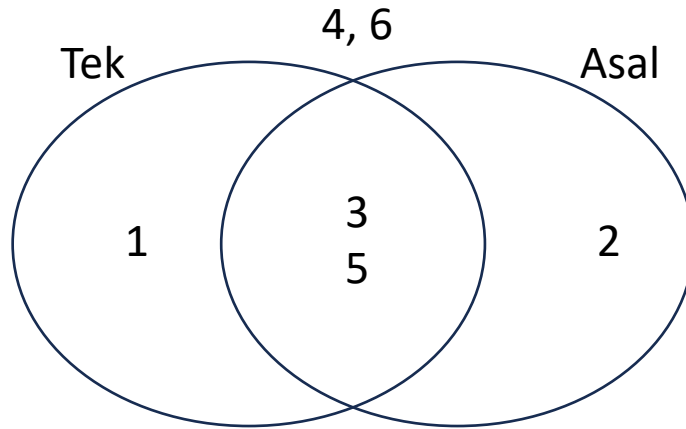
Dr. Öğr. Üyesi İsmail Hakkı Parlak

ismail.parlak@ibu.edu.tr

Oda: 335

Bayes Teoremi

- Bir zar atılıyor. Gelen sayının asal olduğu biliniyorsa bu sayının tek olma olasılığı nedir? $P(S=\text{Tek} \mid S=\text{Asal}) = ?$



$$P(S=\text{Tek} \mid S=\text{Asal}) = \frac{P(\text{Tek} \cap \text{Asal})}{P(\text{Asal})} = \frac{2/6}{3/6}$$

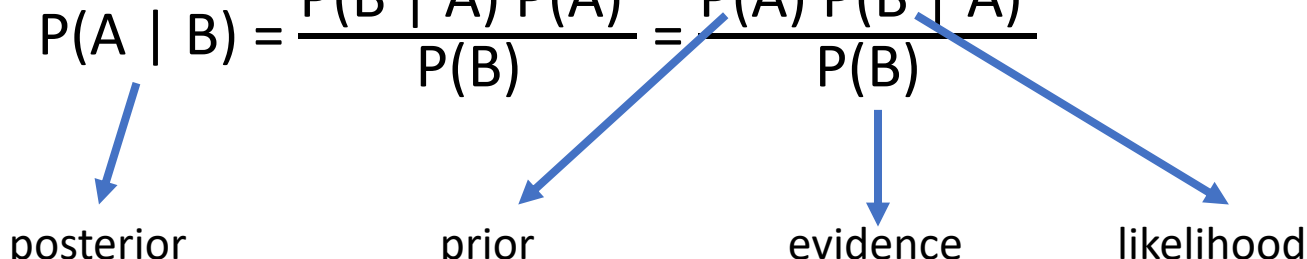
Bayes Teoremi

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{benzer şekilde} \quad P(B \mid A) = \frac{P(B \cap A)}{P(A)}$$

i. $P(A \cap B) = P(B \cap A)$

ii. $P(A \mid B) P(B) = P(A \cap B) = P(B \cap A) = P(B \mid A) P(A)$

iii. $P(A \mid B) = \frac{P(B \mid A) P(A)}{P(B)} = \frac{P(A) P(B \mid A)}{P(B)}$



posterior prior evidence likelihood

Dil Modeli

- $P(S \mid M)$ = bir dil modeli ($M = \{\pi, A1\}$) verildiğinde S dizisinin bu modelde bulunma olasılığı.
 - $\log(P(s_1)) + \sum_{t=2}^T \log(P(s_t \mid s_{t-1}))$
- $P(M \mid S) = ?$
 - S dizisi verildiğinde, M 'nin olasılığı?

Dil Modeli

- MT, ME iki farklı dil modeli olsun.
 - MT: Türkçe, ME: İngilizce.
- $S = \text{"This is a bug"}$
- $P(MT \mid S) <? P(ME \mid S)$
- $P(MT \mid S) = \frac{P(MT)P(S \mid MT)}{P(S)}$
- $P(ME \mid S) = \frac{P(ME)P(S \mid ME)}{P(S)}$

Dil Modeli

- $P(MT \mid S) <? P(ME \mid S)$
- $P(MT \mid S) = \frac{P(MT)P(S \mid MT)}{P(S)}$
- $P(ME \mid S) = \frac{P(ME)P(S \mid ME)}{P(S)}$
- $\frac{P(MT)P(S \mid MT)}{P(S)} <? \frac{P(ME)P(S \mid ME)}{P(S)}$
 - Paydalar eşit.
 - $P(S \mid Mx)$ 'yi hesaplamayı biliyoruz.
 - $P(Mx) = ?$

Dil Modeli

- $\frac{P(MT)P(S | MT)}{P(S)} <? \frac{P(ME)P(S | ME)}{P(S)}$

- Paydalar eşit.
- $P(S | Mx)$ 'yi hesaplamayı biliyoruz.
- $P(Mx)$ = Veri setindeki örneklerin x modeline ait olma

olasılığı = $\frac{\text{count}(s \in Mx)}{\text{count}(s)}$

- Ör: Veri setinde 600 Türkçe, 400 İngilizce örnek varsa, $P(MT) = 0.6$, $P(ME) = 0.4$
- Veri setindeki dağılımlar eşitse $P(Mx)$ 'ler birbirini götürür.

Üretici (Generative) vs Ayırıştırıcı (Discriminative) Modeller

Üretici Modeller	Ayırıştırıcı Modeller
Üzerinde eğitildikleri veri setine benzer veriler üretebilen ML modelleridir. Veri setindeki dağılımları öğrenip bu dağılımlara göre yeni diziler üretebilirler.	Veri setindeki sınıflar arasındaki farklılıkları öğrenip sınıfları ayırıştırabilirler.
Zeki Müren gibi şarkı sözü yazma, bir komuttan resim üretme, chat botları, ...	Bir şarkı sözünün ZM/MG'den hangisine ait olduğunu tahmin etme, resimde kedi/köpek olduğunu tahmin etme, ...
Bayesian networks, Hidden Markov Models, Latent Dirichlet Allocation	Logistic regression, Decision Trees, ANN, Naive Bayes Classifiers

Ayırıştırıcı modeller genelde denetimli öğrenme (supervised learning) yöntemi ile eğitilirler. Örn: Veri setindeki şarkı sözlerinin ZM/MG sınıflarından hangisine ait olduğu önceden bilinmektedir ve her şarkı sözü ZM veya MG etiketi ile etiketlenmiştir. Etiketlere class, label, vb. denir.

Denetimli Öğrenme

1. Veri setindeki tüm örnekler etiketlenir. (ZM:0, MG:1)

Şarkı sözü 1	0
Şarkı sözü 2	1
Şarkı sözü 3	1
...	...
Şarkı sözü 1000	0

2. Veri seti eğitim ve test seti olmak üzere rastgele parçalanır.

Eğitim seti	
Şarkı sözü 2	1
Şarkı sözü 4	0
...	...
Şarkı sözü 1000	0

Test seti	
Şarkı sözü 1	0
Şarkı sözü 3	1
...	...
Şarkı sözü 985	1

Denetimli Öğrenme

3. Eğitim setindeki örnekler ile ML modeli eğitilir.
 4. Test setindeki örnekler ile modelin başarısı test edilir.
- Eğitim setine genelde tüm veri setinin %70 - 85'lik kısmı aktarılır.
 - X: şarkı sözleri, y: şarkıcı etiketleri olsun

```
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2)
```

Denetimli Öğrenme

```
X_train, X_test, y_train, y_test = train_test_split(  
X, y, test_size=0.2)
```

- X_train ve y_train kullanılarak model (M) eğitilir.
 - M.train(X_train, y_train).
- M'in eğitimi bittikten sonra M'e X_test seti verilerek X_test setine karşılık gelen etiketler hesaplatılır.
 - y_pred = M.predict(X_test)
- y_pred ile y_test karşılaştırılarak M'in kaç sonucu doğru tahmin ettiği bulunur ve böylelikle M'in başarısı (accuracy) hesaplanabilir.