

FireFairies Sonuçları

- Hocam öncelikle dictionary'leri bastırırken alt alta formatlamayı çözemedim bu konuda kusura bakmayın lütfen, sonuçlar biraz karışık duruyor.
- İkinci olarak bigram olasılıklarını hesapladığım formülü uygularken sorun yaşadım, Bigram dictionary'sindeki key'ler ile unigram dictionary'sindeki key'leri eşleştirip formülü uygulamanın bir yolunu bulamadım malesef . Ödevde ilerlemek adına bigram'ları sabit bir çarpana böldüm, bu yüzden aynı sayıda bulunan bigram'ların olasılıkları eşit duruyor.
- Son olarak son kısımda girdiğim cümlelerin güncel bigram sonuçlarını alabiliyorum fakat dictionary'den keyleri kullanarak olasılık değerlerini çekip for döngüsüne yazmayı çözemedim, bu yüzden olasılık değerlerini verip bırakmak durumunda kaldım.

Sonuçlarım aşağıdadır, saygılarımla.

```
there are 89 sentences in the text
there are 932 words in text
there are 371 Unique Words in the text
-----UNIGRAM Probabilities(Descending)-----
[[('the',), 0.0987], (('and',), 0.044), (('of',), 0.0268),
(('a',), 0.0182), (('was',), 0.0182), (('boy',), 0.0139),
(('to',), 0.0139), (('fire',), 0.0139), (('light',),
0.0139), (('in',), 0.0129)]
-----BIGRAM Probabilities(Descending)-----
[[('the', 'boy'), 0.8], (('the', 'fire'), 0.6667), (('of',
'the'), 0.5333), (('the', 'girl'), 0.5333), (('said',
'the'), 0.4), (('the', 'original'), 0.3333), (('original',
'light'), 0.3333), (('in', 'the'), 0.2667), (('and',
'the'), 0.2667), (('from', 'the'), 0.2667)]
-----BIGRAM FREQS WITH UNK WORDS and SMOOTHED k
VALUES(Descending)-----
[[('the', 'boy'), 0.0302], (('the', 'fire'), 0.0253),
(('of', 'the'), 0.0205), (('the', 'girl'), 0.0205),
(('said', 'the'), 0.0157), (('the', 'original'), 0.0133),
(('original', 'light'), 0.0133), (('in', 'the'), 0.0109),
(('and', 'the'), 0.0109), (('from', 'the'), 0.0109)]

input a sentence 1-> the boy from the fire said: i am the
fire!
ORNEK CUMLE 1: -> ['the', 'boy', 'from', 'the', 'fire',
'said', 'i', 'UNK', 'the', 'fire']
-----SMOOTHED BIGRAM PROBABILITIES OF INPUT SENTENCE 1
[[('the', 'boy'), 0.0146], (('the', 'fire'), 0.0136),
(('of', 'the'), 0.0094), (('the', 'girl'), 0.0094),
(('said', 'the'), 0.0073), (('from', 'the'), 0.0063),
(('the', 'original'), 0.0063), (('original', 'light'),
0.0063), (('in', 'the'), 0.0052), (('and', 'the'), 0.0052)]

input a sentence 2-> original light is from fire fairies
ORNEK CUMLE 2: -> ['original', 'light', 'is', 'from',
'fire', 'fairies']
-----SMOOTHED BIGRAM PROBABILITIES OF INPUT SENTENCE 2
[[('the', 'boy'), 0.0136], (('the', 'fire'), 0.0115),
(('of', 'the'), 0.0094), (('the', 'girl'), 0.0094),
(('said', 'the'), 0.0073), (('original', 'light'), 0.0073),
(('the', 'original'), 0.0063), (('in', 'the'), 0.0052),
(('and', 'the'), 0.0052), (('from', 'the'), 0.0052)]
```