

```
In [194]: import os,re
import pickle
import numpy as np
import os
```

```
In [195]: lines= [line.lower() for line in open('evliya4.txt')]
text= [line.split() for line in lines]
text2=[list(map(lambda x: x.replace("'", ''), t )) for t in text]
```

```
In [197]: text[0]
```

```
Out[197]: ['dibâce-i', 'seyâhatnâme-i', 'seyyâh-ı', 'âlem', 'evliyâ-yı', 'bî-riyâ']
```

```
In [78]:
```

```
In [204]: evliya=[]
for e in text:
    evliya=evliya+e
evliya2=nlk.Text(evliya)
```

```
In [205]: evliya2.concordance("eyledi")

Displaying 10 of 10 matches:
öfe< emriyle evkât-ı hamseye me'mûr eyledi ve şerâ'it-i İslâm'dan salât [u] s
mazmûn-ı dîlküşâyı bu tarikke inşâ eyledi kim müverrihân-ı rûm ve nûkte-perdâ
e-i mezbûrda saltık sultân vasıyyet eyledi kim "beni gasl edüp yedi tâbût âmâd
lım top-ı pür-renk [ve] alay-ı ceng eyledi kim zemîn ü âsumân dir dir ditreyüp
z kasrın âhir-i masra'-ı târihidir: eyledi ferhâd evvân-ı şîrîn, sene 983. mes
-i lâlä-yâ mustafâ paşa şâh-ı şâhân eyledi tahmâs'a azm sene 956). andan dârü'
eyledim" deydü çâşnigirbaşına benbih eyledi ve akibince çukâdar behlül ağa ile
erzurüm'da gümrük kâtibliği İhsan eyledi ve üç kerre acem diyârına gönderdi
rdır: târih-i cülûs: hân-ı mehemed eyledi [(-)] âsûde dünyâyıl diğêr târih: o
ca (-) "cülûs-ı hân mehemed [373b] eyledi âsûde dünyâyı" [1058] fehvâsınca ci
```

```
In [206]: evliya2.concordance("Bursa")

Displaying 25 of 97 matches:
p âni gördüm, bir tedârik-i azîm ile bursa şehrinin teferrüçü ve ziyâretlerine
kadîm-i evvel, sevâd-ı mu'azzam olan bursa şehrin seyr [u] temâşâ edüp kalbün â
ki bin gurusâ hâsıl olur. ba'zı zamân bursa müvellâsına paşmaklık-bahâ İhsân olu
mânîsi üzre ber-hevâ tayarân ederken bursa hâyilinde sâbite olan ebel-i ruhân'
di. emin nebi'nin dü'âsı berekâtiyle bursa tûfândan sonra imâristân-ı emînullah
e müjde ile orhân seğirdüp vardıkda, bursa feth olduđu sâ'at osmân gâzi rûh tes
ümme'l-fütühindir." sene [626], ammâ bursa lafzı noksandır. ba'dehû ilgar ile o
rsun fakih tilâvet etdê, ez-in-cânib bursa feth-i cedit olmağıla kayseriyye ve
lûp tavattun etdiler. İlä hâze'l-ân bursa me'mûr [u] âbâdân olmadadır. ammâ ib
muz olduđu mahalde. evsâf-ı imârat-ı bursa İç kal'ada cümle iki bin hâne-i zib
e edirne sarâyı'na rağbet kal- mayup bursa sarâyı hâlâ mu'attaldır, ammâ ser-âb
ürüstdür, evsâf-ı hâkimân-ı bender-i bursa evvelâ fâtih sultân mehemed'den ber
ile hüküm eder ve bir oda yeniçeri-i bursa ağası ve cebeci çorbacısı ve yigirmi
hasan ve cennetoğlu nâm celâllilerin bursa üzre hücumların a'yân-ı bursa istimâ
flerin bursa üzre hücumların a'yân-ı bursa istimâ' edüp fermân-ı şehriyâriyle ş
arzında bilâd-ı seb'anm biri dahi bu bursa şehridir. evsâf-ı câmi'hâ-yı selâtin
güyâ bir kal'a-i sedd-i kahkahâ'dır. bursa şehrinin hevâdâr mürtefa' yerine vâk
öre böyle rûhâniyyetli câmi' yokdur. bursa şehrinin güyâ ayasofyası'dır. ve şeh
bâdehâne tahrir olundu. ammâ ibtidâ bursa fethinde binâ olunan yukaru ve iç ka
ya'nî şehid murâd hân-ı evvel câmi'i bursa şehrinin cânib-i garbinda nısf sâ'at
n : sene (-) târihinde binâ olunmuş, bursa şehrinin cânib-i şarkisinde şehrin r
i'dir. manzûrumuz olduđu üzre şehri-i bursa içre böyle münakkas câmi' yokdur. se
r-gir divârlardır. evsâf-ı mesâcid-i bursa bu cevâmî'lerden mâ'adâ cümle (-) ad
b] evsâf-ı medâris-i tahsil-i ulûm-ı bursa evvelâ medrese-ı (manastır'da) orhân
âğ [u] bostânları saky eder. zirâ bu bursa şehri cebel-i ruhân'ın dâmenine düğ
```

```
In [225]: evliya2.concordance("şehirnin")

Displaying 20 of 20 matches:
rdüm, bir tedârik-i azîm ile bursa şehrinin teferrüçü ve ziyâretlerine âmâde o
kal'a-i sedd-i kahkahâ'dır. bursa şehrinin hevâdâr mürtefa' yerine vâki' olmu
e rûhâniyyetli câmi' yokdur. bursa şehrinin güyâ ayasofyası'dır. ve şehrin güz
hid murâd hân-ı evvel câmi'i bursa şehrinin cânib-i garbinda nısf sâ'at ba'id
(-) târihinde binâ olunmuş, bursa şehrinin cânib-i şarkisinde şehrin ravza-i
e tahte'l-kal'ası vardır. ve bursa şehrinin cem'i tarik-ı âmları çakmak taşıy
. anıncün keşiş dağı derler. bursa şehrinin cânib-i kablesinde şehre hâ'il, ef
: "intikâl-ı emir" sene 833. bursa şehrinin hâricinde cânib-i şarkında bir mür
anı gayet latif olmağıla gekbiziye şehrinin iskelesidir. gekbiziye anın şimâli
tna'mûr ve müzeyyen değildir. bolu şehrinin iskelesidir. leb-i deryâda serâpâ
nde karadeniz'e katılır. ve samsun şehrinin suyuna kanava vü geriz derler, mâ-
n cem' olup cereyân ederek bayburd şehrinin içinden cereyân edüp iki tarafı ma
-i azîm moskov vilâyetinin (-) nâm şehrinin dağlarından gelüp üç yerden azak d
zdir, lâkin amık değildir. bu bakü şehrinin niçe gün seyr [u] temâşâsın ederke
darı ve neferâtları yokdur. tortum şehrinin cânib-i şimâline vâki' olmuşdur. c
'da sabanca somunundan sonra lâdik şehrinin memecik ekmeği rub'-ı meskûnda pem
yâtlar cereyân edüp nısfı merzifon şehrinin bâğ u bâğçelerin saky eder. ve cüm
nâ ve mü'eddebelerdir. ve bu köprü şehrinin iskeleleri karadeniz sâhilinde baf
rı birer konak yerdır. ve bu köprü şehrinin (-) tarafında beş sâ'at karib, evs
eğile germiyân hezârı derler. ammâ şehrinin ulemâ ve zurefâları şehri-i bebekbâ
```

Bigrams

```
In [211]: import nltk
from nltk.collocations import *
bigram_measures = nltk.collocations.BigramAssocMeasures()
#trigram_measures = nltk.collocations.TrigramAssocMeasures()
finder = BigramCollocationFinder.from_documents(text2)
finder.apply_freq_filter(10)
scored = finder.score_ngrams(bigram_measures.raw_freq)
```

```
In [ ]:
```

```
In [216]: sorted(bigram for bigram, score in scored) [40:60]
```

```
Out[216]: [('ahmed', 'ağa'),
('ahmed', 'hân'),
('ahmed', 'paşa'),
('ahmed', 'paşa'ya'),
('akçe', 'kazâdir.'),
('akçe', 'şerif'),
('alay-ı', 'azîm'),
('alaybeğisi', 've'),
('aleyhim', 'ecma'in.'),
('allah', 'allah'),
('altun', 've'),
('altı', 'aded'),
('altı', 'bin'),
('altı', 'pâre'),
('altı', 'yüz'),
('ali', 'ağa'),
('ali', 'paşa'),
('ali', 'paşanın'),
('ammâ', 'bir'),
('ammâ', 'bu')]
```

Significant Bigrams (PMI)

```
In [218]: finder.nbest(bigram_measures.pmi, 15)
```

```
Out[218]: [('rub'-ı', 'meskûnda'),
('neye', 'müncer'),
('aleyhim', 'ecma'in.'),
('sulehâ-yı', 'ümmedden'),
('timur-ı', 'bî-nür'),
('ân-ı', 'vâhidde'),
('ebû', 'îshâk'),
('harc-ı', 'râh'),
('şeb-i', 'muzlimde'),
('kat-ı', 'menâzil'),
('istimâd', 'taleb'),
('huzâr-ı', 'meclis'),
('mürür-ı', 'eyyâm'),
('reng-i', 'rüyaları'),
('cemâ'at-i', 'kesireye')]
```

Trigrams

```
In [219]: import nltk
from nltk.collocations import *

#bigram_measures = nltk.collocations.BigramAssocMeasures()
trigram_measures = nltk.collocations.TrigramAssocMeasures()

finder = TrigramCollocationFinder.from_documents(text2)
finder.apply_freq_filter(10)
scored = finder.score_ngrams(trigram_measures.raw_freq)
```

```
In [220]: sorted(bigram for bigram, score in scored) [40:60]
```

```
Out[220]: [('bâlâda', 'tahrir', 'olunan'),
('bâğ', 'u', 'bâğçeli'),
('bâğ', 'u', 'bâğçesi'),
('bâğlı', 've', 'bâğçeli'),
('bâğlı', 've', 'bâğçesi'),
('ceng', '[ü]', 'cidâl'),
('câmi', 've', 'hân'),
('câmi'li', 've', 'hân'),
('cümle', '(-)', 'aded'),
('cümle', 'guzât-ı', 'müslimîn'),
('cümle', 'iki', 'bin'),
('cümle', 'on', 'bir'),
('cümle', 'üç', 'bin'),
('cümle', 'üç', 'yüz'),
('defterdârzâde', 'mehemed', 'paşa'),
('deryâ-misâl', 'asker', 'ile'),
('derûn-ı', 'kal'ada', 'olan'),
('devlet-i', 'âl-i', 'osmân'da'),
('dipi', 've', 'boran'),
('dizdâri', 've', 'neferâtları')]
```

Significant Trigrams (PMI)

```
In [222]: finder.nbest(bigram_measures.pmi, 15)
```

```
Out[222]: [('rahmetullahi', 'aleyhim', 'ecma'in.'),
('sancağ', 'beği', 'tahtıdır.'),
('eflâke', 'ser', 'çekmiş'),
('âyende', 'vü', 'revendeye'),
('âyende', 'vü', 'revendeye'),
('...', '...', '...'),
('evc-i', 'âsumâna', 'ser'),
('âsumâna', 'ser', 'çekmiş'),
('hacı', 'bektâş-ı', 'veli'),
('devlet-i', 'âl-i', 'osmân'da'),
('(-)', '(-)', '(-)'),
('ceng', '[ü]', 'cidâl'),
('zâhir', 'ü', 'bâhirdir.'),
('ta'mir', 'u', 'termim'),
('(2.5', 'satır', 'boş')]
```