



**Department of Computer Science and Engineering (Data Science)**  
**Department of Computer Science and Engineering (Data Science)**

COURSE CODE: **DJ19DSL602**

ACAD: **2023-2024**

DATE: **06 / 02 / 2024**

COURSE NAME: **Computational Linguistics Laboratory**

CLASS: **T.Y.B.Tech (D2)**

**EXPERIMENT NO 2**

**AIM: To implement Language model using N –Gram Model.**

**THEORY:**

Language modeling is the way of determining the probability of any sequence of words. Language modeling is used in a wide variety of applications such as Speech Recognition, Spam filtering, etc. In fact, language modelling is the key aim behind the implementation of many state-of-the-art Natural Language Processing models.

**Two types of Language Modellings:**

**Statistical Language Modellings:** Statistical Language Modeling, or Language Modelling, is the development of probabilistic models that are able to predict the next word in the sequence given the words that precede. Examples such as N-gram language modelling.

**Neural Language Modellings:** Neural network methods are achieving better results than classical methods both on standalone language models and when models are incorporated into larger models on challenging tasks like speech recognition and machine translation. A way of performing a neural language model is through word embedding's.

**Statistical language models**, in its essence, are the type of models that assign probabilities to the sequences of words. In this article, we'll understand the simplest model that assigns probabilities to sentences and sequences of words, the n-gram

**Ngram Language Model**

N-gram Language model is the sequence of N words, by that notion, a 2-gram (or bigram) is a two- word sequence of words like "please turn", "turn your", or " your homework", and a 3-gram (or trigram) is a three- word sequence of words like "please turn your", or "turn your homework"

An N-gram language model predicts the probability of a given N-gram within any sequence of words in the language. A good N-gram model can predict the next word in the sentence i.e the value of  $p(w|h)$

Example of N-gram such as unigram ("This", "article", "is", "on", "NLP") or bi-gram ('This article', 'article is', 'is on', 'on NLP').

Now, we will establish a relation on how to find the next word in the sentence using

. We need to calculate  $p(w|h)$ , where  $w$  is the candidate for the next word. For example in the above example, let's consider, we want to calculate what is the probability of the last word being "NLP" given the previous words:

$$p(NLP|this\ article\ is\ on)$$

After generalizing the above equation can be calculated as:



Shri Vile Parle Kelavani Mandal's

**DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING**

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)



### Department of Computer Science and Engineering (Data Science)

$$p(w_5|w_1, w_2, w_3, w_4) \text{ or } P(W)$$

$$= p(w_n|w_1, w_2 \dots w_n)$$

But how do we calculate it? The answer lies in the chain rule of probability:

$$P(A|B) = \frac{P(A,B)}{P(B)}$$

$$P(A, B) = P(A|B)P(B)$$

Now generalize the above equation:

$$P(X_1, X_2, \dots, X_n) = P(X_1)P(X_2|X_1)P(X_3|X_1, X_2) \dots P(X_n|X_1, X_2, \dots, X_{n-1})$$

$$P(w_1 w_2 w_3 \dots w_n) = \prod_i P(w_i|w_1 w_2 \dots w_{i-1})$$

Simplifying the above formula using Markov assumptions:

$$P(w_i|w_1, w_2, \dots, w_{i-1}) \approx P(w_i|w_{i-k}, \dots, w_{i-1})$$

**For unigram:**

$$P(w_1 w_2, \dots, w_n) \approx \prod_i P(w_i)$$

**For Bigram:**

$$P(w_i|w_1 w_2, \dots, w_{i-1}) \approx P(w_i|w_{i-1})$$

Text	N-gram
Data	1-gram
Great information	2-gram
I am fine	3-gram
Nice to meet you	4-gram

### The Ngram Language Model

$$P(w_n | w_1 \dots w_{n-1}) \approx P(w_n) \text{ unigram}$$

$$P(w_n | w_1 \dots w_{n-1}) \approx P(w_n | w_{n-1}) \text{ bigram}$$

$$P(w_n | w_1 \dots w_{n-1}) \approx P(w_n | w_{n-1} w_{n-2}) \text{ trigram}$$

$$P(w_n | w_1 \dots w_{n-1}) \approx P(w_n | w_{n-1} w_{n-2} w_{n-3}) \text{ 4-gram}$$

$$P(w_n | w_1 \dots w_{n-1}) \approx P(w_n | w_{n-1} w_{n-2} w_{n-3} w_{n-4}) \text{ 5-gram}$$



Shri Vile Parle Kelavani Mandal's

**DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING**

(Autonomous College Affiliated to the University of Mumbai)

NAAC Accredited with "A" Grade (CGPA : 3.18)



## Department of Computer Science and Engineering (Data Science)

### Illustration for N-gram probabilities

#### The Bigram Model

As the name suggests, the bigram model approximates the probability of a word given all the previous words by using only the conditional probability of one preceding word. In other words, you approximate it with the probability:  $P(\text{the} \mid \text{that})$

And so, when you use a bigram model to predict the conditional probability of the next word, you are thus making the following approximation:

This assumption that the probability of a word depends only on the previous word is also known as

#### Markov assumption.

Markov models are the class of probabilistic models that assume that we can predict the probability of some future unit without looking too far in the past. You can further generalize the bigram model to the trigram model which looks two words into the past and can thus be further generalized to the N-gram model.

#### Data Set to Be Used

Import Shakespeare Corpus from NLTK to perform following lab Experiments.

#### Lab Experiments to be performed in this Session:

##### 1. Step-by-step implementation of Ngram language model

1. Basic pre-processing
2. Code to generate N-grams
3. Creating unigrams
4. Creating bigrams
5. Creating trigrams
6. Finding Frequency Distribution
7. Finding Probabilities for Bigram
8. Finding Next word for the given word using MLE (Maximum Likelihood Estimate)

##### 2. Explore NLTK library for understanding and building Ngram Language model

# From scratch

```
import string

text = "National Science Day is celebrated on 28th February every year. The day is to commemorate the famous Indian Physicist Sir C.V. Raman's discovery of the 'Raman Effect.' Moreover, Science Day is a day where people can learn about science and its impact on our lives. It can also be a way to get children excited about science and what they might want to do in the future. On Science Day, different events worldwide promote science and give people the opportunity to participate in hands-on experiments. BYJU'S essay on Science Day helps us know some notable achievements and the objectives of celebrating Science Day. The wonder of science is immeasurable, and we must respect the significant innovations. Science Day is an annual celebration of science and technology in India. National Science Day was first observed in 1987 in India. Moreover, Science Day is celebrated across the country in different sectors like schools, colleges, government offices, universities, etc. National Science Day is also celebrated in research institutes, medical colleges, and science institutes. Now, let us learn some objectives of celebrating Science Day by reading an essay on Science Day in English. Science Day is celebrated to spread awareness and a message about the importance of scientific theories and applications in everyday life. Science day essay in English explains to us, in brief, the objectives of celebrating Science Day. The objective of celebrating Science Day is to display the efforts, accomplishments and activities in the scientific field for human welfare. Moreover, science enthusiasts and scientists discuss theories of science in everyday life and methods to improve technologies."

def preprocess_text(text):
    text = text.lower()
    text = text.translate(str.maketrans('', '', string.punctuation))
    tokens = text.split()
    return tokens

tokens = preprocess_text(text)

tokens
```

```
['national',
 'science',
 'day',
 'is',
 'celebrated',
 'on',
 '28th',
```

'february',  
'every',  
'year',  
'the',  
'day',  
'is',  
'to',  
'commemorate',  
'the',  
'famous',  
'indian',  
'physicist',  
'sir',  
'cv',  
'raman's',  
'discovery',  
'of',  
'the',  
'raman',  
'effect',  
'moreover',  
'science',  
'day',  
'is',  
'a',  
'day',  
'where',  
'people',  
'can',  
'learn',  
'about',  
'science',  
'and',  
'its',  
'impact',  
'on',  
'our',  
'lives',  
'it',  
'can',  
'also',  
'be',  
'a',  
'way',  
'to',  
'get',  
'children',  
'excited',  
'about',

'science',  
'and',  
'what',  
'they',  
'might',  
'want',  
'to',  
'do',  
'in',  
'the',  
'future',  
'on',  
'science',  
'day',  
'different',  
'events',  
'worldwide',  
'promote',  
'science',  
'and',  
'give',  
'people',  
'the',  
'opportunity',  
'to',  
'participate',  
'in',  
'hands-on',  
'experiments',  
'byju's',  
'essay',  
'on',  
'science',  
'day',  
'helps',  
'us',  
'know',  
'some',  
'notable',  
'achievements',  
'and',  
'the',  
'objectives',  
'of',  
'celebrating',  
'science',  
'day',  
'the',  
'wonder',

'of',  
'science',  
'is',  
'immeasurable',  
'and',  
'we',  
'must',  
'respect',  
'the',  
'significant',  
'innovations',  
'science',  
'day',  
'is',  
'an',  
'annual',  
'celebration',  
'of',  
'science',  
'and',  
'technology',  
'in',  
'india',  
'national',  
'science',  
'day',  
'was',  
'first',  
'observed',  
'in',  
'1987',  
'in',  
'india',  
'moreover',  
'science',  
'day',  
'is',  
'celebrated',  
'across',  
'the',  
'country',  
'in',  
'different',  
'sectors',  
'like',  
'schools',  
'colleges',  
'government',  
'offices',

'universities',  
'etc',  
'national',  
'science',  
'day',  
'is',  
'also',  
'celebrated',  
'in',  
'research',  
'institutes',  
'medical',  
'colleges',  
'and',  
'science',  
'institutes',  
'now',  
'let',  
'us',  
'learn',  
'some',  
'objectives',  
'of',  
'celebrating',  
'science',  
'day',  
'by',  
'reading',  
'an',  
'essay',  
'on',  
'science',  
'day',  
'in',  
'english',  
'science',  
'day',  
'is',  
'celebrated',  
'to',  
'spread',  
'awareness',  
'and',  
'a',  
'message',  
'about',  
'the',  
'importance',  
'of',



'scientific',  
'theories',  
'and',  
'applications',  
'in',  
'everyday',  
'life',  
'science',  
'day',  
'essay',  
'in',  
'english',  
'explains',  
'to',  
'us',  
'in',  
'brief',  
'the',  
'objectives',  
'of',  
'celebrating',  
'science',  
'day',  
'the',  
'objective',  
'of',  
'celebrating',  
'science',  
'day',  
'is',  
'to',  
'display',  
'the',  
'efforts',  
'accomplishments',  
'and',  
'activities',  
'in',  
'the',  
'scientific',  
'field',  
'for',  
'human',  
'welfare',  
'moreover',  
'science',  
'enthusiasts',  
'and',  
'scientists',

```

'discuss',
'theories',
'of',
'science',
'in',
'everyday',
'life',
'and',
'methods',
'to',
'improve',
'technologies']

def generate_ngrams(tokens, n):
    ngrams_list = []
    for i in range(len(tokens) - n + 1):
        ngrams_list.append(tuple(tokens[i:i+n]))
    return ngrams_list

def create_unigrams(tokens):
    return generate_ngrams(tokens, 1)

unigrams = create_unigrams(tokens)

unigrams

[('national',),
 ('science',),
 ('day',),
 ('is',),
 ('celebrated',),
 ('on',),
 ('28th',),
 ('february',),
 ('every',),
 ('year',),
 ('the',),
 ('day',),
 ('is',),
 ('to',),
 ('commemorate',),
 ('the',),
 ('famous',),
 ('indian',),
 ('physicist',),
 ('sir',),
 ('cv',),
 ('raman's',),
 ('discovery',),
 ('of',),

```

('the',),  
('raman',),  
('effect',),  
('moreover',),  
('science',),  
('day',),  
('is',),  
('a',),  
('day',),  
('where',),  
('people',),  
('can',),  
('learn',),  
('about',),  
('science',),  
('and',),  
('its',),  
('impact',),  
('on',),  
('our',),  
('lives',),  
('it',),  
('can',),  
('also',),  
('be',),  
('a',),  
('way',),  
('to',),  
('get',),  
('children',),  
('excited',),  
('about',),  
('science',),  
('and',),  
('what',),  
('they',),  
('might',),  
('want',),  
('to',),  
('do',),  
('in',),  
('the',),  
('future',),  
('on',),  
('science',),  
('day',),  
('different',),  
('events',),  
('worldwide',),

('promote',),  
('science',),  
('and',),  
('give',),  
('people',),  
('the',),  
('opportunity',),  
('to',),  
('participate',),  
('in',),  
('handson',),  
('experiments',),  
('byju's',),  
('essay',),  
('on',),  
('science',),  
('day',),  
('helps',),  
('us',),  
('know',),  
('some',),  
('notable',),  
('achievements',),  
('and',),  
('the',),  
('objectives',),  
('of',),  
('celebrating',),  
('science',),  
('day',),  
('the',),  
('wonder',),  
('of',),  
('science',),  
('is',),  
('immeasurable',),  
('and',),  
('we',),  
('must',),  
('respect',),  
('the',),  
('significant',),  
('innovations',),  
('science',),  
('day',),  
('is',),  
('an',),  
('annual',),  
('celebration',),

('of',),  
('science',),  
('and',),  
('technology',),  
('in',),  
('india',),  
('national',),  
('science',),  
('day',),  
('was',),  
('first',),  
('observed',),  
('in',),  
('1987',),  
('in',),  
('india',),  
('moreover',),  
('science',),  
('day',),  
('is',),  
('celebrated',),  
('across',),  
('the',),  
('country',),  
('in',),  
('different',),  
('sectors',),  
('like',),  
('schools',),  
('colleges',),  
('government',),  
('offices',),  
('universities',),  
('etc',),  
('national',),  
('science',),  
('day',),  
('is',),  
('also',),  
('celebrated',),  
('in',),  
('research',),  
('institutes',),  
('medical',),  
('colleges',),  
('and',),  
('science',),  
('institutes',),  
('now',),

('let',),  
('us',),  
('learn',),  
('some',),  
('objectives',),  
('of',),  
('celebrating',),  
('science',),  
('day',),  
('by',),  
('reading',),  
('an',),  
('essay',),  
('on',),  
('science',),  
('day',),  
('in',),  
('english',),  
('science',),  
('day',),  
('is',),  
('celebrated',),  
('to',),  
('spread',),  
('awareness',),  
('and',),  
('a',),  
('message',),  
('about',),  
('the',),  
('importance',),  
('of',),  
('scientific',),  
('theories',),  
('and',),  
('applications',),  
('in',),  
('everyday',),  
('life',),  
('science',),  
('day',),  
('essay',),  
('in',),  
('english',),  
('explains',),  
('to',),  
('us',),  
('in',),  
('brief',),

```
('the',),
('objectives',),
('of',),
('celebrating',),
('science',),
('day',),
('the',),
('objective',),
('of',),
('celebrating',),
('science',),
('day',),
('is',),
('to',),
('display',),
('the',),
('efforts',),
('accomplishments',),
('and',),
('activities',),
('in',),
('the',),
('scientific',),
('field',),
('for',),
('human',),
('welfare',),
('moreover',),
('science',),
('enthusiasts',),
('and',),
('scientists',),
('discuss',),
('theories',),
('of',),
('science',),
('in',),
('everyday',),
('life',),
('and',),
('methods',),
('to',),
('improve',),
('technologies',)]
```

```
def create_bigrams(tokens):
    return generate_ngrams(tokens, 2)
```

```
bigrams = create_bigrams(tokens)
```

bigrams

```
[('national', 'science'),
 ('science', 'day'),
 ('day', 'is'),
 ('is', 'celebrated'),
 ('celebrated', 'on'),
 ('on', '28th'),
 ('28th', 'february'),
 ('february', 'every'),
 ('every', 'year'),
 ('year', 'the'),
 ('the', 'day'),
 ('day', 'is'),
 ('is', 'to'),
 ('to', 'commemorate'),
 ('commemorate', 'the'),
 ('the', 'famous'),
 ('famous', 'indian'),
 ('indian', 'physicist'),
 ('physicist', 'sir'),
 ('sir', 'cv'),
 ('cv', 'raman's'),
 ('raman's', 'discovery'),
 ('discovery', 'of'),
 ('of', 'the'),
 ('the', 'raman'),
 ('raman', 'effect'),
 ('effect', 'moreover'),
 ('moreover', 'science'),
 ('science', 'day'),
 ('day', 'is'),
 ('is', 'a'),
 ('a', 'day'),
 ('day', 'where'),
 ('where', 'people'),
 ('people', 'can'),
 ('can', 'learn'),
 ('learn', 'about'),
 ('about', 'science'),
 ('science', 'and'),
 ('and', 'its'),
 ('its', 'impact'),
 ('impact', 'on'),
 ('on', 'our'),
 ('our', 'lives'),
 ('lives', 'it'),
 ('it', 'can'),
 ('can', 'also'),
 ('also', 'be'),
```



('be', 'a'),  
('a', 'way'),  
('way', 'to'),  
('to', 'get'),  
('get', 'children'),  
('children', 'excited'),  
('excited', 'about'),  
('about', 'science'),  
('science', 'and'),  
('and', 'what'),  
('what', 'they'),  
('they', 'might'),  
('might', 'want'),  
('want', 'to'),  
('to', 'do'),  
('do', 'in'),  
('in', 'the'),  
('the', 'future'),  
('future', 'on'),  
('on', 'science'),  
('science', 'day'),  
('day', 'different'),  
('different', 'events'),  
('events', 'worldwide'),  
('worldwide', 'promote'),  
('promote', 'science'),  
('science', 'and'),  
('and', 'give'),  
('give', 'people'),  
('people', 'the'),  
('the', 'opportunity'),  
('opportunity', 'to'),  
('to', 'participate'),  
('participate', 'in'),  
('in', 'handson'),  
('handson', 'experiments'),  
('experiments', 'byju's'),  
('byju's', 'essay'),  
('essay', 'on'),  
('on', 'science'),  
('science', 'day'),  
('day', 'helps'),  
('helps', 'us'),  
('us', 'know'),  
('know', 'some'),  
('some', 'notable'),  
('notable', 'achievements'),  
('achievements', 'and'),  
('and', 'the'),

('the', 'objectives'),  
('objectives', 'of'),  
('of', 'celebrating'),  
('celebrating', 'science'),  
('science', 'day'),  
('day', 'the'),  
('the', 'wonder'),  
('wonder', 'of'),  
('of', 'science'),  
('science', 'is'),  
('is', 'immeasurable'),  
('immeasurable', 'and'),  
('and', 'we'),  
('we', 'must'),  
('must', 'respect'),  
('respect', 'the'),  
('the', 'significant'),  
('significant', 'innovations'),  
('innovations', 'science'),  
('science', 'day'),  
('day', 'is'),  
('is', 'an'),  
('an', 'annual'),  
('annual', 'celebration'),  
('celebration', 'of'),  
('of', 'science'),  
('science', 'and'),  
('and', 'technology'),  
('technology', 'in'),  
('in', 'india'),  
('india', 'national'),  
('national', 'science'),  
('science', 'day'),  
('day', 'was'),  
('was', 'first'),  
('first', 'observed'),  
('observed', 'in'),  
('in', '1987'),  
('1987', 'in'),  
('in', 'india'),  
('india', 'moreover'),  
('moreover', 'science'),  
('science', 'day'),  
('day', 'is'),  
('is', 'celebrated'),  
('celebrated', 'across'),  
('across', 'the'),  
('the', 'country'),  
('country', 'in'),

('in', 'different'),  
('different', 'sectors'),  
('sectors', 'like'),  
('like', 'schools'),  
('schools', 'colleges'),  
('colleges', 'government'),  
('government', 'offices'),  
('offices', 'universities'),  
('universities', 'etc'),  
('etc', 'national'),  
('national', 'science'),  
('science', 'day'),  
('day', 'is'),  
('is', 'also'),  
('also', 'celebrated'),  
('celebrated', 'in'),  
('in', 'research'),  
('research', 'institutes'),  
('institutes', 'medical'),  
('medical', 'colleges'),  
('colleges', 'and'),  
('and', 'science'),  
('science', 'institutes'),  
('institutes', 'now'),  
('now', 'let'),  
('let', 'us'),  
('us', 'learn'),  
('learn', 'some'),  
('some', 'objectives'),  
('objectives', 'of'),  
('of', 'celebrating'),  
('celebrating', 'science'),  
('science', 'day'),  
('day', 'by'),  
('by', 'reading'),  
('reading', 'an'),  
('an', 'essay'),  
('essay', 'on'),  
('on', 'science'),  
('science', 'day'),  
('day', 'in'),  
('in', 'english'),  
('english', 'science'),  
('science', 'day'),  
('day', 'is'),  
('is', 'celebrated'),  
('celebrated', 'to'),  
('to', 'spread'),  
('spread', 'awareness'),

('awareness', 'and'),  
('and', 'a'),  
('a', 'message'),  
('message', 'about'),  
('about', 'the'),  
('the', 'importance'),  
('importance', 'of'),  
('of', 'scientific'),  
('scientific', 'theories'),  
('theories', 'and'),  
('and', 'applications'),  
('applications', 'in'),  
('in', 'everyday'),  
('everyday', 'life'),  
('life', 'science'),  
('science', 'day'),  
('day', 'essay'),  
('essay', 'in'),  
('in', 'english'),  
('english', 'explains'),  
('explains', 'to'),  
('to', 'us'),  
('us', 'in'),  
('in', 'brief'),  
('brief', 'the'),  
('the', 'objectives'),  
('objectives', 'of'),  
('of', 'celebrating'),  
('celebrating', 'science'),  
('science', 'day'),  
('day', 'the'),  
('the', 'objective'),  
('objective', 'of'),  
('of', 'celebrating'),  
('celebrating', 'science'),  
('science', 'day'),  
('day', 'is'),  
('is', 'to'),  
('to', 'display'),  
('display', 'the'),  
('the', 'efforts'),  
('efforts', 'accomplishments'),  
('accomplishments', 'and'),  
('and', 'activities'),  
('activities', 'in'),  
('in', 'the'),  
('the', 'scientific'),  
('scientific', 'field'),  
('field', 'for'),

```
('for', 'human'),
('human', 'welfare'),
('welfare', 'moreover'),
('moreover', 'science'),
('science', 'enthusiasts'),
('enthusiasts', 'and'),
('and', 'scientists'),
('scientists', 'discuss'),
('discuss', 'theories'),
('theories', 'of'),
('of', 'science'),
('science', 'in'),
('in', 'everyday'),
('everyday', 'life'),
('life', 'and'),
('and', 'methods'),
('methods', 'to'),
('to', 'improve'),
('improve', 'technologies')]
```

```
def create_trigrams(tokens):
    return generate_ngrams(tokens, 3)
```

```
trigrams = create_trigrams(tokens)
```

```
trigrams
```

```
[('national', 'science', 'day'),
 ('science', 'day', 'is'),
 ('day', 'is', 'celebrated'),
 ('is', 'celebrated', 'on'),
 ('celebrated', 'on', '28th'),
 ('on', '28th', 'february'),
 ('28th', 'february', 'every'),
 ('february', 'every', 'year'),
 ('every', 'year', 'the'),
 ('year', 'the', 'day'),
 ('the', 'day', 'is'),
 ('day', 'is', 'to'),
 ('is', 'to', 'commemorate'),
 ('to', 'commemorate', 'the'),
 ('commemorate', 'the', 'famous'),
 ('the', 'famous', 'indian'),
 ('famous', 'indian', 'physicist'),
 ('indian', 'physicist', 'sir'),
 ('physicist', 'sir', 'cv'),
 ('sir', 'cv', 'raman's'),
 ('cv', 'raman's', 'discovery'),
 ('raman's', 'discovery', 'of'),
 ('discovery', 'of', 'the'),
```

('of', 'the', 'raman'),  
('the', 'raman', 'effect'),  
('raman', 'effect', 'moreover'),  
('effect', 'moreover', 'science'),  
('moreover', 'science', 'day'),  
('science', 'day', 'is'),  
('day', 'is', 'a'),  
('is', 'a', 'day'),  
('a', 'day', 'where'),  
('day', 'where', 'people'),  
('where', 'people', 'can'),  
('people', 'can', 'learn'),  
('can', 'learn', 'about'),  
('learn', 'about', 'science'),  
('about', 'science', 'and'),  
('science', 'and', 'its'),  
('and', 'its', 'impact'),  
('its', 'impact', 'on'),  
('impact', 'on', 'our'),  
('on', 'our', 'lives'),  
('our', 'lives', 'it'),  
('lives', 'it', 'can'),  
('it', 'can', 'also'),  
('can', 'also', 'be'),  
('also', 'be', 'a'),  
('be', 'a', 'way'),  
('a', 'way', 'to'),  
('way', 'to', 'get'),  
('to', 'get', 'children'),  
('get', 'children', 'excited'),  
('children', 'excited', 'about'),  
('excited', 'about', 'science'),  
('about', 'science', 'and'),  
('science', 'and', 'what'),  
('and', 'what', 'they'),  
('what', 'they', 'might'),  
('they', 'might', 'want'),  
('might', 'want', 'to'),  
('want', 'to', 'do'),  
('to', 'do', 'in'),  
('do', 'in', 'the'),  
('in', 'the', 'future'),  
('the', 'future', 'on'),  
('future', 'on', 'science'),  
('on', 'science', 'day'),  
('science', 'day', 'different'),  
('day', 'different', 'events'),  
('different', 'events', 'worldwide'),  
('events', 'worldwide', 'promote'),

('worldwide', 'promote', 'science'),  
('promote', 'science', 'and'),  
('science', 'and', 'give'),  
('and', 'give', 'people'),  
('give', 'people', 'the'),  
('people', 'the', 'opportunity'),  
('the', 'opportunity', 'to'),  
('opportunity', 'to', 'participate'),  
('to', 'participate', 'in'),  
('participate', 'in', 'handson'),  
('in', 'handson', 'experiments'),  
('handson', 'experiments', 'byju's'),  
('experiments', 'byju's', 'essay'),  
('byju's', 'essay', 'on'),  
('essay', 'on', 'science'),  
('on', 'science', 'day'),  
('science', 'day', 'helps'),  
('day', 'helps', 'us'),  
('helps', 'us', 'know'),  
('us', 'know', 'some'),  
('know', 'some', 'notable'),  
('some', 'notable', 'achievements'),  
('notable', 'achievements', 'and'),  
('achievements', 'and', 'the'),  
('and', 'the', 'objectives'),  
('the', 'objectives', 'of'),  
('objectives', 'of', 'celebrating'),  
('of', 'celebrating', 'science'),  
('celebrating', 'science', 'day'),  
('science', 'day', 'the'),  
('day', 'the', 'wonder'),  
('the', 'wonder', 'of'),  
('wonder', 'of', 'science'),  
('of', 'science', 'is'),  
('science', 'is', 'immeasurable'),  
('is', 'immeasurable', 'and'),  
('immeasurable', 'and', 'we'),  
('and', 'we', 'must'),  
('we', 'must', 'respect'),  
('must', 'respect', 'the'),  
('respect', 'the', 'significant'),  
('the', 'significant', 'innovations'),  
('significant', 'innovations', 'science'),  
('innovations', 'science', 'day'),  
('science', 'day', 'is'),  
('day', 'is', 'an'),  
('is', 'an', 'annual'),  
('an', 'annual', 'celebration'),  
('annual', 'celebration', 'of'),

('celebration', 'of', 'science'),  
('of', 'science', 'and'),  
('science', 'and', 'technology'),  
('and', 'technology', 'in'),  
('technology', 'in', 'india'),  
('in', 'india', 'national'),  
('india', 'national', 'science'),  
('national', 'science', 'day'),  
('science', 'day', 'was'),  
('day', 'was', 'first'),  
('was', 'first', 'observed'),  
('first', 'observed', 'in'),  
('observed', 'in', '1987'),  
('in', '1987', 'in'),  
('1987', 'in', 'india'),  
('in', 'india', 'moreover'),  
('india', 'moreover', 'science'),  
('moreover', 'science', 'day'),  
('science', 'day', 'is'),  
('day', 'is', 'celebrated'),  
('is', 'celebrated', 'across'),  
('celebrated', 'across', 'the'),  
('across', 'the', 'country'),  
('the', 'country', 'in'),  
('country', 'in', 'different'),  
('in', 'different', 'sectors'),  
('different', 'sectors', 'like'),  
('sectors', 'like', 'schools'),  
('like', 'schools', 'colleges'),  
('schools', 'colleges', 'government'),  
('colleges', 'government', 'offices'),  
('government', 'offices', 'universities'),  
('offices', 'universities', 'etc'),  
('universities', 'etc', 'national'),  
('etc', 'national', 'science'),  
('national', 'science', 'day'),  
('science', 'day', 'is'),  
('day', 'is', 'also'),  
('is', 'also', 'celebrated'),  
('also', 'celebrated', 'in'),  
('celebrated', 'in', 'research'),  
('in', 'research', 'institutes'),  
('research', 'institutes', 'medical'),  
('institutes', 'medical', 'colleges'),  
('medical', 'colleges', 'and'),  
('colleges', 'and', 'science'),  
('and', 'science', 'institutes'),  
('science', 'institutes', 'now'),  
('institutes', 'now', 'let'),



('now', 'let', 'us'),  
('let', 'us', 'learn'),  
('us', 'learn', 'some'),  
('learn', 'some', 'objectives'),  
('some', 'objectives', 'of'),  
('objectives', 'of', 'celebrating'),  
('of', 'celebrating', 'science'),  
('celebrating', 'science', 'day'),  
('science', 'day', 'by'),  
('day', 'by', 'reading'),  
('by', 'reading', 'an'),  
('reading', 'an', 'essay'),  
('an', 'essay', 'on'),  
('essay', 'on', 'science'),  
('on', 'science', 'day'),  
('science', 'day', 'in'),  
('day', 'in', 'english'),  
('in', 'english', 'science'),  
('english', 'science', 'day'),  
('science', 'day', 'is'),  
('day', 'is', 'celebrated'),  
('is', 'celebrated', 'to'),  
('celebrated', 'to', 'spread'),  
('to', 'spread', 'awareness'),  
('spread', 'awareness', 'and'),  
('awareness', 'and', 'a'),  
('and', 'a', 'message'),  
('a', 'message', 'about'),  
('message', 'about', 'the'),  
('about', 'the', 'importance'),  
('the', 'importance', 'of'),  
('importance', 'of', 'scientific'),  
('of', 'scientific', 'theories'),  
('scientific', 'theories', 'and'),  
('theories', 'and', 'applications'),  
('and', 'applications', 'in'),  
('applications', 'in', 'everyday'),  
('in', 'everyday', 'life'),  
('everyday', 'life', 'science'),  
('life', 'science', 'day'),  
('science', 'day', 'essay'),  
('day', 'essay', 'in'),  
('essay', 'in', 'english'),  
('in', 'english', 'explains'),  
('english', 'explains', 'to'),  
('explains', 'to', 'us'),  
('to', 'us', 'in'),  
('us', 'in', 'brief'),  
('in', 'brief', 'the'),

```
(
    ('brief', 'the', 'objectives'),
    ('the', 'objectives', 'of'),
    ('objectives', 'of', 'celebrating'),
    ('of', 'celebrating', 'science'),
    ('celebrating', 'science', 'day'),
    ('science', 'day', 'the'),
    ('day', 'the', 'objective'),
    ('the', 'objective', 'of'),
    ('objective', 'of', 'celebrating'),
    ('of', 'celebrating', 'science'),
    ('celebrating', 'science', 'day'),
    ('science', 'day', 'is'),
    ('day', 'is', 'to'),
    ('is', 'to', 'display'),
    ('to', 'display', 'the'),
    ('display', 'the', 'efforts'),
    ('the', 'efforts', 'accomplishments'),
    ('efforts', 'accomplishments', 'and'),
    ('accomplishments', 'and', 'activities'),
    ('and', 'activities', 'in'),
    ('activities', 'in', 'the'),
    ('in', 'the', 'scientific'),
    ('the', 'scientific', 'field'),
    ('scientific', 'field', 'for'),
    ('field', 'for', 'human'),
    ('for', 'human', 'welfare'),
    ('human', 'welfare', 'moreover'),
    ('welfare', 'moreover', 'science'),
    ('moreover', 'science', 'enthusiasts'),
    ('science', 'enthusiasts', 'and'),
    ('enthusiasts', 'and', 'scientists'),
    ('and', 'scientists', 'discuss'),
    ('scientists', 'discuss', 'theories'),
    ('discuss', 'theories', 'of'),
    ('theories', 'of', 'science'),
    ('of', 'science', 'in'),
    ('science', 'in', 'everyday'),
    ('in', 'everyday', 'life'),
    ('everyday', 'life', 'and'),
    ('life', 'and', 'methods'),
    ('and', 'methods', 'to'),
    ('methods', 'to', 'improve'),
    ('to', 'improve', 'technologies'])
```

```
from collections import Counter
```

```
def calculate_frequency_distribution(ngrams):
    return Counter(ngrams)
```

```
unigrams_freq_dist = calculate_frequency_distribution(unigrams)
```

unigrams\_freq\_dist

```
Counter({'national',): 3,  
        ('science',): 23,  
        ('day',): 17,  
        ('is',): 9,  
        ('celebrated',): 4,  
        ('on',): 5,  
        ('28th',): 1,  
        ('february',): 1,  
        ('every',): 1,  
        ('year',): 1,  
        ('the',): 14,  
        ('to',): 8,  
        ('commemorate',): 1,  
        ('famous',): 1,  
        ('indian',): 1,  
        ('physicist',): 1,  
        ('sir',): 1,  
        ('cv',): 1,  
        ('raman's',): 1,  
        ('discovery',): 1,  
        ('of',): 9,  
        ('raman',): 1,  
        ('effect',): 1,  
        ('moreover',): 3,  
        ('a',): 3,  
        ('where',): 1,  
        ('people',): 2,  
        ('can',): 2,  
        ('learn',): 2,  
        ('about',): 3,  
        ('and',): 12,  
        ('its',): 1,  
        ('impact',): 1,  
        ('our',): 1,  
        ('lives',): 1,  
        ('it',): 1,  
        ('also',): 2,  
        ('be',): 1,  
        ('way',): 1,  
        ('get',): 1,  
        ('children',): 1,  
        ('excited',): 1,  
        ('what',): 1,  
        ('they',): 1,  
        ('might',): 1,  
        ('want',): 1,  
        ('do',): 1,  
        ('in',): 13,
```

```
('future',): 1,
('different',): 2,
('events',): 1,
('worldwide',): 1,
('promote',): 1,
('give',): 1,
('opportunity',): 1,
('participate',): 1,
('handson',): 1,
('experiments',): 1,
('byju's',): 1,
('essay',): 3,
('helps',): 1,
('us',): 3,
('know',): 1,
('some',): 2,
('notable',): 1,
('achievements',): 1,
('objectives',): 3,
('celebrating',): 4,
('wonder',): 1,
('immeasurable',): 1,
('we',): 1,
('must',): 1,
('respect',): 1,
('significant',): 1,
('innovations',): 1,
('an',): 2,
('annual',): 1,
('celebration',): 1,
('technology',): 1,
('india',): 2,
('was',): 1,
('first',): 1,
('observed',): 1,
('1987',): 1,
('across',): 1,
('country',): 1,
('sectors',): 1,
('like',): 1,
('schools',): 1,
('colleges',): 2,
('government',): 1,
('offices',): 1,
('universities',): 1,
('etc',): 1,
('research',): 1,
('institutes',): 2,
('medical',): 1,
```

```

        ('now',): 1,
        ('let',): 1,
        ('by',): 1,
        ('reading',): 1,
        ('english',): 2,
        ('spread',): 1,
        ('awareness',): 1,
        ('message',): 1,
        ('importance',): 1,
        ('scientific',): 2,
        ('theories',): 2,
        ('applications',): 1,
        ('everyday',): 2,
        ('life',): 2,
        ('explains',): 1,
        ('brief',): 1,
        ('objective',): 1,
        ('display',): 1,
        ('efforts',): 1,
        ('accomplishments',): 1,
        ('activities',): 1,
        ('field',): 1,
        ('for',): 1,
        ('human',): 1,
        ('welfare',): 1,
        ('enthusiasts',): 1,
        ('scientists',): 1,
        ('discuss',): 1,
        ('methods',): 1,
        ('improve',): 1,
        ('technologies',): 1})

def calculate_bigram_probabilities(bigrams, unigrams_freq_dist):
    bigram_probabilities = {}
    for bigram in bigrams:
        word, next_word = bigram
        bigram_probabilities[bigram] = unigrams_freq_dist[word] /
sum(unigrams_freq_dist.values())
    return bigram_probabilities

bigram_probabilities = calculate_bigram_probabilities(bigrams,
unigrams_freq_dist)

def predict_next_word(word, bigram_probabilities):
    next_word_probabilities = {}
    for bigram, probability in bigram_probabilities.items():
        if bigram[0] == word:
            next_word_probabilities[bigram[1]] = probability
    if not next_word_probabilities:
        return None # or any other default value

```

```

    return max(next_word_probabilities,
key=next_word_probabilities.get)

given_word = "science"
next_word = predict_next_word(given_word, bigram_probabilities)
print("Given word:", given_word)
print("Next word predicted using Bigram MLE:", next_word)

Given word: science
Next word predicted using Bigram MLE: day

```

#Using NLTK

```

import nltk
from nltk.corpus import gutenber
from nltk.probability import FreqDist
from nltk.util import ngrams
from collections import defaultdict

nltk.download('punkt')

[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Package punkt is already up-to-date!

True

shakespeare_corpus = gutenber.sents('shakespeare-hamlet.txt')

```

#Basic pre-processing [lowercase and punctuations]

```

def preprocess_corpus(corpus):
    processed_corpus = []
    for sentence in corpus:
        processed_sentence = [word.lower() for word in sentence if
word.isalnum()]
        processed_corpus.append(processed_sentence)
    return processed_corpus

processed_shakespeare_corpus = preprocess_corpus(shakespeare_corpus)

```

#Function to generate N-grams

```

def generate_ngrams(tokens, n):
    return list(ngrams(tokens, n))

```

#Creating Unigrams

```
unigrams = [gram for sentence in processed_shakespeare_corpus for gram
in generate_ngrams(sentence, 1)]
```

```
unigrams
```

```
[('the',),
 ('tragedie',),
 ('of',),
 ('hamlet',),
 ('by',),
 ('william',),
 ('shakespeare',),
 ('1599',),
 ('actus',),
 ('primus',),
 ('scoena',),
 ('prima',),
 ('enter',),
 ('barnardo',),
 ('and',),
 ('francisco',),
 ('two',),
 ('centinels',),
 ('barnardo',),
 ('who',),
 ('s',),
 ('there',),
 ('fran',),
 ('nay',),
 ('answer',),
 ('me',),
 ('stand',),
 ('vnfold',),
 ('your',),
 ('selfe',),
 ('bar',),
 ('long',),
 ('liue',),
 ('the',),
 ('king',),
 ('fran',),
 ('barnardo',),
 ('bar',),
 ('he',),
 ('fran',),
 ('you',),
 ('come',),
 ('most',),
 ('carefully',),
 ('vpon',),
```

('your',),  
('houre',),  
('bar',),  
('tis',),  
('now',),  
('strook',),  
('twelue',),  
('get',),  
('thee',),  
('to',),  
('bed',),  
('francisco',),  
('fran',),  
('for',),  
('this',),  
('releefe',),  
('much',),  
('thankes',),  
('tis',),  
('bitter',),  
('cold',),  
('and',),  
('i',),  
('am',),  
('sicke',),  
('at',),  
('heart',),  
('barn',),  
('haue',),  
('you',),  
('had',),  
('quiet',),  
('guard',),  
('fran',),  
('not',),  
('a',),  
('mouse',),  
('stirring',),  
('barn',),  
('well',),  
('goodnight',),  
('if',),  
('you',),  
('do',),  
('meet',),  
('horatio',),  
('and',),  
('marcellus',),  
('the',),



('riuals',),  
('of',),  
('my',),  
('watch',),  
('bid',),  
('them',),  
('make',),  
('hast',),  
('enter',),  
('horatio',),  
('and',),  
('marcellus',),  
('fran',),  
('i',),  
('thinke',),  
('i',),  
('heare',),  
('them',),  
('stand',),  
('who',),  
('s',),  
('there',),  
('hor',),  
('friends',),  
('to',),  
('this',),  
('ground',),  
('mar',),  
('and',),  
('leige',),  
('men',),  
('to',),  
('the',),  
('dane',),  
('fran',),  
('giue',),  
('you',),  
('good',),  
('night',),  
('mar',),  
('o',),  
('farwel',),  
('honest',),  
('soldier',),  
('who',),  
('hath',),  
('relieu',),  
('d',),  
('you',),

('fra',),  
('barnardo',),  
('ha',),  
('s',),  
('my',),  
('place',),  
('giue',),  
('you',),  
('goodnight',),  
('exit',),  
('fran',),  
('mar',),  
('holla',),  
('barnardo',),  
('bar',),  
('say',),  
('what',),  
('is',),  
('horatio',),  
('there',),  
('hor',),  
('a',),  
('peece',),  
('of',),  
('him',),  
('bar',),  
('welcome',),  
('horatio',),  
('welcome',),  
('good',),  
('marcellus',),  
('mar',),  
('what',),  
('ha',),  
('s',),  
('this',),  
('thing',),  
('appear',),  
('d',),  
('again',),  
('to',),  
('night',),  
('bar',),  
('i',),  
('haue',),  
('seene',),  
('nothing',),  
('mar',),  
('horatio',),

('saies',),  
('tis',),  
('but',),  
('our',),  
('fantasie',),  
('and',),  
('will',),  
('not',),  
('let',),  
('beleefe',),  
('take',),  
('hold',),  
('of',),  
('him',),  
('touching',),  
('this',),  
('dreaded',),  
('sight',),  
('twice',),  
('seene',),  
('of',),  
('vs',),  
('therefore',),  
('i',),  
('haue',),  
('intreated',),  
('him',),  
('along',),  
('with',),  
('vs',),  
('to',),  
('watch',),  
('the',),  
('minutes',),  
('of',),  
('this',),  
('night',),  
('that',),  
('if',),  
('again',),  
('this',),  
('apparition',),  
('come',),  
('he',),  
('may',),  
('approue',),  
('our',),  
('eyes',),  
('and',),

('speake',),  
('to',),  
('it',),  
('hor',),  
('tush',),  
('tush',),  
('twill',),  
('not',),  
('appeare',),  
('bar',),  
('sit',),  
('downe',),  
('a',),  
('while',),  
('and',),  
('let',),  
('vs',),  
('once',),  
('again',),  
('assaile',),  
('your',),  
('eares',),  
('that',),  
('are',),  
('so',),  
('fortified',),  
('against',),  
('our',),  
('story',),  
('what',),  
('we',),  
('two',),  
('nights',),  
('haue',),  
('seene',),  
('hor',),  
('well',),  
('sit',),  
('we',),  
('downe',),  
('and',),  
('let',),  
('vs',),  
('heare',),  
('barnardo',),  
('speake',),  
('of',),  
('this',),  
('barn',),

('last',),  
('night',),  
('of',),  
('all',),  
('when',),  
('yond',),  
('same',),  
('starre',),  
('that',),  
('s',),  
('westward',),  
('from',),  
('the',),  
('pole',),  
('had',),  
('made',),  
('his',),  
('course',),  
('t',),  
('illume',),  
('that',),  
('part',),  
('of',),  
('heauen',),  
('where',),  
('now',),  
('it',),  
('burnes',),  
('marcellus',),  
('and',),  
('my',),  
('selfe',),  
('the',),  
('bell',),  
('then',),  
('beating',),  
('one',),  
('mar',),  
('peace',),  
('breake',),  
('thee',),  
('of',),  
('enter',),  
('the',),  
('ghost',),  
('looke',),  
('where',),  
('it',),  
('comes',),

('again',),  
('barn',),  
('in',),  
('the',),  
('same',),  
('figure',),  
('like',),  
('the',),  
('king',),  
('that',),  
('s',),  
('dead',),  
('mar',),  
('thou',),  
('art',),  
('a',),  
('scholler',),  
('speake',),  
('to',),  
('it',),  
('horatio',),  
('barn',),  
('lookes',),  
('it',),  
('not',),  
('like',),  
('the',),  
('king',),  
('marke',),  
('it',),  
('horatio',),  
('hora',),  
('most',),  
('like',),  
('it',),  
('harrowes',),  
('me',),  
('with',),  
('fear',),  
('wonder',),  
('barn',),  
('it',),  
('would',),  
('be',),  
('spoke',),  
('too',),  
('mar',),  
('question',),  
('it',),

('horatio',),  
('hor',),  
('what',),  
('art',),  
('thou',),  
('that',),  
('vsurp',),  
('st',),  
('this',),  
('time',),  
('of',),  
('night',),  
('together',),  
('with',),  
('that',),  
('faire',),  
('and',),  
('warlike',),  
('forme',),  
('in',),  
('which',),  
('the',),  
('maiesty',),  
('of',),  
('buried',),  
('denmarke',),  
('did',),  
('sometimes',),  
('march',),  
('by',),  
('heauen',),  
('i',),  
('charge',),  
('thee',),  
('speake',),  
('mar',),  
('it',),  
('is',),  
('offended',),  
('barn',),  
('see',),  
('it',),  
('stalkes',),  
('away',),  
('hor',),  
('stay',),  
('speake',),  
('speake',),  
('i',),

('charge',),  
('thee',),  
('speake',),  
('exit',),  
('the',),  
('ghost',),  
('mar',),  
('tis',),  
('gone',),  
('and',),  
('will',),  
('not',),  
('answer',),  
('barn',),  
('how',),  
('now',),  
('horatio',),  
('you',),  
('tremble',),  
('look',),  
('pale',),  
('is',),  
('not',),  
('this',),  
('something',),  
('more',),  
('then',),  
('fantasie',),  
('what',),  
('thinke',),  
('you',),  
('on',),  
('t',),  
('hor',),  
('before',),  
('my',),  
('god',),  
('i',),  
('might',),  
('not',),  
('this',),  
('beleeue',),  
('without',),  
('the',),  
('sensible',),  
('and',),  
('true',),  
('auouch',),  
('of',),



('mine',),  
('owne',),  
('eyes',),  
('mar',),  
('is',),  
('it',),  
('not',),  
('like',),  
('the',),  
('king',),  
('hor',),  
('as',),  
('thou',),  
('art',),  
('to',),  
('thy',),  
('selfe',),  
('such',),  
('was',),  
('the',),  
('very',),  
('armour',),  
('he',),  
('had',),  
('on',),  
('when',),  
('th',),  
('ambitious',),  
('norwey',),  
('combatted',),  
('so',),  
('frown',),  
('d',),  
('he',),  
('once',),  
('when',),  
('in',),  
('an',),  
('angry',),  
('parle',),  
('he',),  
('smot',),  
('the',),  
('sledged',),  
('pollax',),  
('on',),  
('the',),  
('ice',),  
('tis',),  
('strange',),

('mar',),  
('thus',),  
('twice',),  
('before',),  
('and',),  
('iust',),  
('at',),  
('this',),  
('dead',),  
('houre',),  
('with',),  
('martiall',),  
('stalke',),  
('hath',),  
('he',),  
('gone',),  
('by',),  
('our',),  
('watch',),  
('hor',),  
('in',),  
('what',),  
('particular',),  
('thought',),  
('to',),  
('work',),  
('i',),  
('know',),  
('not',),  
('but',),  
('in',),  
('the',),  
('grosse',),  
('and',),  
('scope',),  
('of',),  
('my',),  
('opinion',),  
('this',),  
('boades',),  
('some',),  
('strange',),  
('erruption',),  
('to',),  
('our',),  
('state',),  
('mar',),  
('good',),  
('now',),

('sit',),  
('downe',),  
('tell',),  
('me',),  
('he',),  
('that',),  
('knowes',),  
('why',),  
('this',),  
('same',),  
('strict',),  
('and',),  
('most',),  
('obseruant',),  
('watch',),  
('so',),  
('nightly',),  
('toyles',),  
('the',),  
('subiect',),  
('of',),  
('the',),  
('land',),  
('and',),  
('why',),  
('such',),  
('dayly',),  
('cast',),  
('of',),  
('brazon',),  
('cannon',),  
('and',),  
('forraigne',),  
('mart',),  
('for',),  
('implements',),  
('of',),  
('warre',),  
('why',),  
('such',),  
('impreſſe',),  
('of',),  
('ſhip',),  
('wrights',),  
('whoſe',),  
('ſore',),  
('taſke',),  
('do',),  
('s',),

('not',),  
('diuide',),  
('the',),  
('sunday',),  
('from',),  
('the',),  
('weeke',),  
('what',),  
('might',),  
('be',),  
('toward',),  
('that',),  
('this',),  
('sweaty',),  
('hast',),  
('doth',),  
('make',),  
('the',),  
('night',),  
('ioynt',),  
('labourer',),  
('with',),  
('the',),  
('day',),  
('who',),  
('is',),  
('t',),  
('that',),  
('can',),  
('informe',),  
('me',),  
('hor',),  
('that',),  
('can',),  
('i',),  
('at',),  
('least',),  
('the',),  
('whisper',),  
('goes',),  
('so',),  
('our',),  
('last',),  
('king',),  
('whose',),  
('image',),  
('euen',),  
('but',),  
('now',),

('appear',),  
('d',),  
('to',),  
('vs',),  
('was',),  
('as',),  
('you',),  
('know',),  
('by',),  
('fortinbras',),  
('of',),  
('norway',),  
('thereto',),  
('prick',),  
('d',),  
('on',),  
('by',),  
('a',),  
('most',),  
('emulate',),  
('pride',),  
('dar',),  
('d',),  
('to',),  
('the',),  
('combate',),  
('in',),  
('which',),  
('our',),  
('valiant',),  
('hamlet',),  
('for',),  
('so',),  
('this',),  
('side',),  
('of',),  
('our',),  
('knowne',),  
('world',),  
('esteem',),  
('d',),  
('him',),  
('did',),  
('slay',),  
('this',),  
('fortinbras',),  
('who',),  
('by',),  
('a',),

('seal',),  
('d',),  
('compact',),  
('well',),  
('ratified',),  
('by',),  
('law',),  
('and',),  
('heraldrie',),  
('did',),  
('forfeite',),  
('with',),  
('his',),  
('life',),  
('all',),  
('those',),  
('his',),  
('lands',),  
('which',),  
('he',),  
('stood',),  
('seiz',),  
('d',),  
('on',),  
('to',),  
('the',),  
('conqueror',),  
('against',),  
('the',),  
('which',),  
('a',),  
('moity',),  
('competent',),  
('was',),  
('gaged',),  
('by',),  
('our',),  
('king',),  
('which',),  
('had',),  
('return',),  
('d',),  
('to',),  
('the',),  
('inheritance',),  
('of',),  
('fortinbras',),  
('had',),  
('he',),

('bin',),  
('vanquisher',),  
('as',),  
('by',),  
('the',),  
('same',),  
('cou',),  
('nant',),  
('and',),  
('carriage',),  
('of',),  
('the',),  
('article',),  
('designe',),  
('his',),  
('fell',),  
('to',),  
('hamlet',),  
('now',),  
('sir',),  
('young',),  
('fortinbras',),  
('of',),  
('vnimproued',),  
('mettle',),  
('hot',),  
('and',),  
('full',),  
('hath',),  
('in',),  
('the',),  
('skirts',),  
('of',),  
('norway',),  
('heere',),  
('and',),  
('there',),  
('shark',),  
('d',),  
('vp',),  
('a',),  
('list',),  
('of',),  
('landlesse',),  
('resolutes',),  
('for',),  
('foode',),  
('and',),  
('diet',),

('to',),  
('some',),  
('enterprize',),  
('that',),  
('hath',),  
('a',),  
('stomacke',),  
('in',),  
('t',),  
('which',),  
('is',),  
('no',),  
('other',),  
('and',),  
('it',),  
('doth',),  
('well',),  
('appeare',),  
('vnto',),  
('our',),  
('state',),  
('but',),  
('to',),  
('recouer',),  
('of',),  
('vs',),  
('by',),  
('strong',),  
('hand',),  
('and',),  
('termes',),  
('compulsatiue',),  
('those',),  
('foresaid',),  
('lands',),  
('so',),  
('by',),  
('his',),  
('father',),  
('lost',),  
('and',),  
('this',),  
('i',),  
('take',),  
('it',),  
('is',),  
('the',),  
('maine',),  
('motiue',),



('of',),  
('our',),  
('preparations',),  
('the',),  
('source',),  
('of',),  
('this',),  
('our',),  
('watch',),  
('and',),  
('the',),  
('cheefe',),  
('head',),  
('of',),  
('this',),  
('post',),  
('hast',),  
('and',),  
('romage',),  
('in',),  
('the',),  
('land',),  
('enter',),  
('ghost',),  
('again',),  
('but',),  
('soft',),  
('behold',),  
('loe',),  
('where',),  
('it',),  
('comes',),  
('again',),  
('ile',),  
('crosse',),  
('it',),  
('though',),  
('it',),  
('blast',),  
('me',),  
('stay',),  
('illusion',),  
('if',),  
('thou',),  
('hast',),  
('any',),  
('sound',),  
('or',),  
('vse',),

('of',),  
('voyce',),  
('speake',),  
('to',),  
('me',),  
('if',),  
('there',),  
('be',),  
('any',),  
('good',),  
('thing',),  
('to',),  
('be',),  
('done',),  
('that',),  
('may',),  
('to',),  
('thee',),  
('do',),  
('ease',),  
('and',),  
('grace',),  
('to',),  
('me',),  
('speak',),  
('to',),  
('me',),  
('if',),  
('thou',),  
('art',),  
('priuy',),  
('to',),  
('thy',),  
('countries',),  
('fate',),  
('which',),  
('happily',),  
('foreknowing',),  
('may',),  
('auoyd',),  
('oh',),  
('speake',),  
('or',),  
('if',),  
('thou',),  
('hast',),  
('vp',),  
('hoorded',),  
('in',),

```
( 'thy', ),
( 'life', ),
( 'extorted', ),
( 'treasure', ),
( 'in', ),
( 'the', ),
( 'wombe', ),
( 'of', ),
( 'earth', ),
( 'for', ),
( 'which', ),
( 'they', ),
( 'say', ),
( 'you', ),
( 'spirits', ),
( 'oft', ),
( 'walke', ),
( 'in', ),
( 'death', ),
( 'speake', ),
( 'of', ),
( 'it', ),
( 'stay', ),
...]
```

#### #Creating Bigrams

```
bigrams = [gram for sentence in processed_shakespeare_corpus for gram
in generate_ngrams(sentence, 2)]
```

bigrams

```
[('the', 'tragedie'),
('tragedie', 'of'),
('of', 'hamlet'),
('hamlet', 'by'),
('by', 'william'),
('william', 'shakespeare'),
('shakespeare', '1599'),
('actus', 'primus'),
('scoena', 'prima'),
('enter', 'barnardo'),
('barnardo', 'and'),
('and', 'francisco'),
('francisco', 'two'),
('two', 'centinels'),
('who', 's'),
('s', 'there'),
('nay', 'answer'),
```

('answer', 'me'),  
('me', 'stand'),  
('stand', 'vnfold'),  
('vnfold', 'your'),  
('your', 'selfe'),  
('long', 'liue'),  
('liue', 'the'),  
('the', 'king'),  
('you', 'come'),  
('come', 'most'),  
('most', 'carefully'),  
('carefully', 'vpon'),  
('vpon', 'your'),  
('your', 'houre'),  
('tis', 'now'),  
('now', 'strook'),  
('strook', 'twelue'),  
('twelue', 'get'),  
('get', 'thee'),  
('thee', 'to'),  
('to', 'bed'),  
('bed', 'francisco'),  
('for', 'this'),  
('this', 'releefe'),  
('releefe', 'much'),  
('much', 'thankes'),  
('thankes', 'tis'),  
('tis', 'bitter'),  
('bitter', 'cold'),  
('cold', 'and'),  
('and', 'i'),  
('i', 'am'),  
('am', 'sicke'),  
('sicke', 'at'),  
('at', 'heart'),  
('haue', 'you'),  
('you', 'had'),  
('had', 'quiet'),  
('quiet', 'guard'),  
('not', 'a'),  
('a', 'mouse'),  
('mouse', 'stirring'),  
('well', 'goodnight'),  
('if', 'you'),  
('you', 'do'),  
('do', 'meet'),  
('meet', 'horatio'),  
('horatio', 'and'),  
('and', 'marcellus'),

('marcellus', 'the'),  
('the', 'riuals'),  
('riuals', 'of'),  
('of', 'my'),  
('my', 'watch'),  
('watch', 'bid'),  
('bid', 'them'),  
('them', 'make'),  
('make', 'hast'),  
('enter', 'horatio'),  
('horatio', 'and'),  
('and', 'marcellus'),  
('i', 'thinke'),  
('thinke', 'i'),  
('i', 'heare'),  
('heare', 'them'),  
('stand', 'who'),  
('who', 's'),  
('s', 'there'),  
('friends', 'to'),  
('to', 'this'),  
('this', 'ground'),  
('and', 'leige'),  
('leige', 'men'),  
('men', 'to'),  
('to', 'the'),  
('the', 'dane'),  
('giue', 'you'),  
('you', 'good'),  
('good', 'night'),  
('o', 'farwel'),  
('farwel', 'honest'),  
('honest', 'soldier'),  
('soldier', 'who'),  
('who', 'hath'),  
('hath', 'relieu'),  
('relieu', 'd'),  
('d', 'you'),  
('barnardo', 'ha'),  
('ha', 's'),  
('s', 'my'),  
('my', 'place'),  
('place', 'giue'),  
('giue', 'you'),  
('you', 'goodnight'),  
('exit', 'fran'),  
('holla', 'barnardo'),  
('say', 'what'),  
('what', 'is'),

('is', 'horatio'),  
('horatio', 'there'),  
('a', 'peece'),  
('peece', 'of'),  
('of', 'him'),  
('welcome', 'horatio'),  
('horatio', 'welcome'),  
('welcome', 'good'),  
('good', 'marcellus'),  
('what', 'ha'),  
('ha', 's'),  
('s', 'this'),  
('this', 'thing'),  
('thing', 'appear'),  
('appear', 'd'),  
('d', 'again'),  
('again', 'to'),  
('to', 'night'),  
('i', 'haue'),  
('haue', 'seene'),  
('seene', 'nothing'),  
('horatio', 'saies'),  
('saies', 'tis'),  
('tis', 'but'),  
('but', 'our'),  
('our', 'fantasie'),  
('fantasie', 'and'),  
('and', 'will'),  
('will', 'not'),  
('not', 'let'),  
('let', 'beleefe'),  
('beleefe', 'take'),  
('take', 'hold'),  
('hold', 'of'),  
('of', 'him'),  
('him', 'touching'),  
('touching', 'this'),  
('this', 'dreaded'),  
('dreaded', 'sight'),  
('sight', 'twice'),  
('twice', 'seene'),  
('seene', 'of'),  
('of', 'vs'),  
('vs', 'therefore'),  
('therefore', 'i'),  
('i', 'haue'),  
('haue', 'intreated'),  
('intreated', 'him'),  
('him', 'along'),

('along', 'with'),  
('with', 'vs'),  
('vs', 'to'),  
('to', 'watch'),  
('watch', 'the'),  
('the', 'minutes'),  
('minutes', 'of'),  
('of', 'this'),  
('this', 'night'),  
('night', 'that'),  
('that', 'if'),  
('if', 'again'),  
('again', 'this'),  
('this', 'apparition'),  
('apparition', 'come'),  
('come', 'he'),  
('he', 'may'),  
('may', 'approve'),  
('approve', 'our'),  
('our', 'eyes'),  
('eyes', 'and'),  
('and', 'speak'),  
('speak', 'to'),  
('to', 'it'),  
('tush', 'tush'),  
('tush', 'twill'),  
('twill', 'not'),  
('not', 'appear'),  
('sit', 'down'),  
('down', 'a'),  
('a', 'while'),  
('while', 'and'),  
('and', 'let'),  
('let', 'vs'),  
('vs', 'once'),  
('once', 'again'),  
('again', 'assault'),  
('assault', 'your'),  
('your', 'ears'),  
('ears', 'that'),  
('that', 'are'),  
('are', 'so'),  
('so', 'fortified'),  
('fortified', 'against'),  
('against', 'our'),  
('our', 'story'),  
('story', 'what'),  
('what', 'we'),  
('we', 'two'),

('two', 'nights'),  
('nights', 'haue'),  
('haue', 'seene'),  
('well', 'sit'),  
('sit', 'we'),  
('we', 'downe'),  
('downe', 'and'),  
('and', 'let'),  
('let', 'vs'),  
('vs', 'heare'),  
('heare', 'barnardo'),  
('barnardo', 'speake'),  
('speake', 'of'),  
('of', 'this'),  
('last', 'night'),  
('night', 'of'),  
('of', 'all'),  
('all', 'when'),  
('when', 'yond'),  
('yond', 'same'),  
('same', 'starre'),  
('starre', 'that'),  
('that', 's'),  
('s', 'westward'),  
('westward', 'from'),  
('from', 'the'),  
('the', 'pole'),  
('pole', 'had'),  
('had', 'made'),  
('made', 'his'),  
('his', 'course'),  
('course', 't'),  
('t', 'illume'),  
('illume', 'that'),  
('that', 'part'),  
('part', 'of'),  
('of', 'heauen'),  
('heauen', 'where'),  
('where', 'now'),  
('now', 'it'),  
('it', 'burnes'),  
('burnes', 'marcellus'),  
('marcellus', 'and'),  
('and', 'my'),  
('my', 'selfe'),  
('selfe', 'the'),  
('the', 'bell'),  
('bell', 'then'),  
('then', 'beating'),



('beating', 'one'),  
('peace', 'breake'),  
('breake', 'thee'),  
('thee', 'of'),  
('of', 'enter'),  
('enter', 'the'),  
('the', 'ghost'),  
('looke', 'where'),  
('where', 'it'),  
('it', 'comes'),  
('comes', 'again'),  
('in', 'the'),  
('the', 'same'),  
('same', 'figure'),  
('figure', 'like'),  
('like', 'the'),  
('the', 'king'),  
('king', 'that'),  
('that', 's'),  
('s', 'dead'),  
('thou', 'art'),  
('art', 'a'),  
('a', 'scholler'),  
('scholler', 'speake'),  
('speake', 'to'),  
('to', 'it'),  
('it', 'horatio'),  
('lookes', 'it'),  
('it', 'not'),  
('not', 'like'),  
('like', 'the'),  
('the', 'king'),  
('marke', 'it'),  
('it', 'horatio'),  
('most', 'like'),  
('like', 'it'),  
('it', 'harrowes'),  
('harrowes', 'me'),  
('me', 'with'),  
('with', 'fear'),  
('fear', 'wonder'),  
('wonder', 'barn'),  
('it', 'would'),  
('would', 'be'),  
('be', 'spoke'),  
('spoke', 'too'),  
('question', 'it'),  
('it', 'horatio'),  
('what', 'art'),

('art', 'thou'),  
('thou', 'that'),  
('that', 'vsurp'),  
('vsurp', 'st'),  
('st', 'this'),  
('this', 'time'),  
('time', 'of'),  
('of', 'night'),  
('night', 'together'),  
('together', 'with'),  
('with', 'that'),  
('that', 'faire'),  
('faire', 'and'),  
('and', 'warlike'),  
('warlike', 'forme'),  
('forme', 'in'),  
('in', 'which'),  
('which', 'the'),  
('the', 'maiesty'),  
('maiesty', 'of'),  
('of', 'buried'),  
('buried', 'denmarke'),  
('denmarke', 'did'),  
('did', 'sometimes'),  
('sometimes', 'march'),  
('march', 'by'),  
('by', 'heauen'),  
('heauen', 'i'),  
('i', 'charge'),  
('charge', 'thee'),  
('thee', 'speake'),  
('it', 'is'),  
('is', 'offended'),  
('see', 'it'),  
('it', 'stalkes'),  
('stalkes', 'away'),  
('stay', 'speake'),  
('speake', 'speake'),  
('speake', 'i'),  
('i', 'charge'),  
('charge', 'thee'),  
('thee', 'speake'),  
('exit', 'the'),  
('the', 'ghost'),  
('tis', 'gone'),  
('gone', 'and'),  
('and', 'will'),  
('will', 'not'),  
('not', 'answer'),

('how', 'now'),  
('now', 'horatio'),  
('you', 'tremble'),  
('tremble', 'look'),  
('look', 'pale'),  
('pale', 'is'),  
('is', 'not'),  
('not', 'this'),  
('this', 'something'),  
('something', 'more'),  
('more', 'then'),  
('then', 'fantasie'),  
('what', 'thinke'),  
('thinke', 'you'),  
('you', 'on'),  
('on', 't'),  
('before', 'my'),  
('my', 'god'),  
('god', 'i'),  
('i', 'might'),  
('might', 'not'),  
('not', 'this'),  
('this', 'beleeue'),  
('beleeue', 'without'),  
('without', 'the'),  
('the', 'sensible'),  
('sensible', 'and'),  
('and', 'true'),  
('true', 'auouch'),  
('auouch', 'of'),  
('of', 'mine'),  
('mine', 'owne'),  
('owne', 'eyes'),  
('is', 'it'),  
('it', 'not'),  
('not', 'like'),  
('like', 'the'),  
('the', 'king'),  
('as', 'thou'),  
('thou', 'art'),  
('art', 'to'),  
('to', 'thy'),  
('thy', 'selfe'),  
('selfe', 'such'),  
('such', 'was'),  
('was', 'the'),  
('the', 'very'),  
('very', 'armour'),  
('armour', 'he'),

('he', 'had'),  
('had', 'on'),  
('on', 'when'),  
('when', 'th'),  
('th', 'ambitious'),  
('ambitious', 'norway'),  
('norway', 'combatted'),  
('combatted', 'so'),  
('so', 'frown'),  
('frown', 'd'),  
('d', 'he'),  
('he', 'once'),  
('once', 'when'),  
('when', 'in'),  
('in', 'an'),  
('an', 'angry'),  
('angry', 'parle'),  
('parle', 'he'),  
('he', 'smot'),  
('smot', 'the'),  
('the', 'sledded'),  
('sledded', 'pollax'),  
('pollax', 'on'),  
('on', 'the'),  
('the', 'ice'),  
('tis', 'strange'),  
('thus', 'twice'),  
('twice', 'before'),  
('before', 'and'),  
('and', 'iust'),  
('iust', 'at'),  
('at', 'this'),  
('this', 'dead'),  
('dead', 'houre'),  
('houre', 'with'),  
('with', 'martiall'),  
('martiall', 'stalke'),  
('stalke', 'hath'),  
('hath', 'he'),  
('he', 'gone'),  
('gone', 'by'),  
('by', 'our'),  
('our', 'watch'),  
('in', 'what'),  
('what', 'particular'),  
('particular', 'thought'),  
('thought', 'to'),  
('to', 'work'),  
('work', 'i'),

('i', 'know'),  
('know', 'not'),  
('not', 'but'),  
('but', 'in'),  
('in', 'the'),  
('the', 'grosse'),  
('grosse', 'and'),  
('and', 'scope'),  
('scope', 'of'),  
('of', 'my'),  
('my', 'opinion'),  
('opinion', 'this'),  
('this', 'boades'),  
('boades', 'some'),  
('some', 'strange'),  
('strange', 'eruption'),  
('eruption', 'to'),  
('to', 'our'),  
('our', 'state'),  
('good', 'now'),  
('now', 'sit'),  
('sit', 'downe'),  
('downe', 'tell'),  
('tell', 'me'),  
('me', 'he'),  
('he', 'that'),  
('that', 'knowes'),  
('knowes', 'why'),  
('why', 'this'),  
('this', 'same'),  
('same', 'strict'),  
('strict', 'and'),  
('and', 'most'),  
('most', 'obseruant'),  
('obseruant', 'watch'),  
('watch', 'so'),  
('so', 'nightly'),  
('nightly', 'toyles'),  
('toyles', 'the'),  
('the', 'subiect'),  
('subiect', 'of'),  
('of', 'the'),  
('the', 'land'),  
('land', 'and'),  
('and', 'why'),  
('why', 'such'),  
('such', 'dayly'),  
('dayly', 'cast'),  
('cast', 'of'),  
('of', 'brazon'),

('brazon', 'cannon'),  
('cannon', 'and'),  
('and', 'forraigne'),  
('forraigne', 'mart'),  
('mart', 'for'),  
('for', 'implements'),  
('implements', 'of'),  
('of', 'warre'),  
('warre', 'why'),  
('why', 'such'),  
('such', 'impresse'),  
('impresse', 'of'),  
('of', 'ship'),  
('ship', 'wrights'),  
('wrights', 'whose'),  
('whose', 'sore'),  
('sore', 'taske'),  
('taske', 'do'),  
('do', 's'),  
('s', 'not'),  
('not', 'diuide'),  
('diuide', 'the'),  
('the', 'sunday'),  
('sunday', 'from'),  
('from', 'the'),  
('the', 'weeke'),  
('weeke', 'what'),  
('what', 'might'),  
('might', 'be'),  
('be', 'toward'),  
('toward', 'that'),  
('that', 'this'),  
('this', 'sweaty'),  
('sweaty', 'hast'),  
('hast', 'doth'),  
('doth', 'make'),  
('make', 'the'),  
('the', 'night'),  
('night', 'ioynt'),  
('ioynt', 'labourer'),  
('labourer', 'with'),  
('with', 'the'),  
('the', 'day'),  
('day', 'who'),  
('who', 'is'),  
('is', 't'),  
('t', 'that'),  
('that', 'can'),  
('can', 'informe'),

('informe', 'me'),  
('that', 'can'),  
('can', 'i'),  
('i', 'at'),  
('at', 'least'),  
('least', 'the'),  
('the', 'whisper'),  
('whisper', 'goes'),  
('goes', 'so'),  
('so', 'our'),  
('our', 'last'),  
('last', 'king'),  
('king', 'whose'),  
('whose', 'image'),  
('image', 'euen'),  
('euen', 'but'),  
('but', 'now'),  
('now', 'appear'),  
('appear', 'd'),  
('d', 'to'),  
('to', 'vs'),  
('vs', 'was'),  
('was', 'as'),  
('as', 'you'),  
('you', 'know'),  
('know', 'by'),  
('by', 'fortinbras'),  
('fortinbras', 'of'),  
('of', 'norway'),  
('norway', 'thereto'),  
('thereto', 'prick'),  
('prick', 'd'),  
('d', 'on'),  
('on', 'by'),  
('by', 'a'),  
('a', 'most'),  
('most', 'emulate'),  
('emulate', 'pride'),  
('pride', 'dar'),  
('dar', 'd'),  
('d', 'to'),  
('to', 'the'),  
('the', 'combate'),  
('in', 'which'),  
('which', 'our'),  
('our', 'valiant'),  
('valiant', 'hamlet'),  
('hamlet', 'for'),  
('for', 'so'),

('so', 'this'),  
('this', 'side'),  
('side', 'of'),  
('of', 'our'),  
('our', 'knowne'),  
('knowne', 'world'),  
('world', 'esteem'),  
('esteem', 'd'),  
('d', 'him'),  
('him', 'did'),  
('did', 'slay'),  
('slay', 'this'),  
('this', 'fortinbras'),  
('fortinbras', 'who'),  
('who', 'by'),  
('by', 'a'),  
('a', 'seal'),  
('seal', 'd'),  
('d', 'compact'),  
('compact', 'well'),  
('well', 'ratified'),  
('ratified', 'by'),  
('by', 'law'),  
('law', 'and'),  
('and', 'heraldrie'),  
('heraldrie', 'did'),  
('did', 'forfeite'),  
('forfeite', 'with'),  
('with', 'his'),  
('his', 'life'),  
('life', 'all'),  
('all', 'those'),  
('those', 'his'),  
('his', 'lands'),  
('lands', 'which'),  
('which', 'he'),  
('he', 'stood'),  
('stood', 'seiz'),  
('seiz', 'd'),  
('d', 'on'),  
('on', 'to'),  
('to', 'the'),  
('the', 'conqueror'),  
('conqueror', 'against'),  
('against', 'the'),  
('the', 'which'),  
('which', 'a'),  
('a', 'moity'),  
('moity', 'competent'),



('competent', 'was'),  
('was', 'gaged'),  
('gaged', 'by'),  
('by', 'our'),  
('our', 'king'),  
('king', 'which'),  
('which', 'had'),  
('had', 'return'),  
('return', 'd'),  
('d', 'to'),  
('to', 'the'),  
('the', 'inheritance'),  
('inheritance', 'of'),  
('of', 'fortinbras'),  
('fortinbras', 'had'),  
('had', 'he'),  
('he', 'bin'),  
('bin', 'vanquisher'),  
('vanquisher', 'as'),  
('as', 'by'),  
('by', 'the'),  
('the', 'same'),  
('same', 'cou'),  
('cou', 'nant'),  
('nant', 'and'),  
('and', 'carriage'),  
('carriage', 'of'),  
('of', 'the'),  
('the', 'article'),  
('article', 'designe'),  
('designe', 'his'),  
('his', 'fell'),  
('fell', 'to'),  
('to', 'hamlet'),  
('now', 'sir'),  
('sir', 'young'),  
('young', 'fortinbras'),  
('fortinbras', 'of'),  
('of', 'vnimproued'),  
('vnimproued', 'mettle'),  
('mettle', 'hot'),  
('hot', 'and'),  
('and', 'full'),  
('full', 'hath'),  
('hath', 'in'),  
('in', 'the'),  
('the', 'skirts'),  
('skirts', 'of'),  
('of', 'norway'),

('norway', 'heere'),  
('heere', 'and'),  
('and', 'there'),  
('there', 'shark'),  
('shark', 'd'),  
('d', 'vp'),  
('vp', 'a'),  
('a', 'list'),  
('list', 'of'),  
('of', 'landlesse'),  
('landlesse', 'resolutes'),  
('resolutes', 'for'),  
('for', 'foode'),  
('foode', 'and'),  
('and', 'diet'),  
('diet', 'to'),  
('to', 'some'),  
('some', 'enterprize'),  
('enterprize', 'that'),  
('that', 'hath'),  
('hath', 'a'),  
('a', 'stomacke'),  
('stomacke', 'in'),  
('in', 't'),  
('t', 'which'),  
('which', 'is'),  
('is', 'no'),  
('no', 'other'),  
('other', 'and'),  
('and', 'it'),  
('it', 'doth'),  
('doth', 'well'),  
('well', 'appeare'),  
('appeare', 'vnto'),  
('vnto', 'our'),  
('our', 'state'),  
('state', 'but'),  
('but', 'to'),  
('to', 'recouer'),  
('recouer', 'of'),  
('of', 'vs'),  
('vs', 'by'),  
('by', 'strong'),  
('strong', 'hand'),  
('hand', 'and'),  
('and', 'termes'),  
('termes', 'compulsatiue'),  
('compulsatiue', 'those'),  
('those', 'foresaid'),

('foresaid', 'lands'),  
('lands', 'so'),  
('so', 'by'),  
('by', 'his'),  
('his', 'father'),  
('father', 'lost'),  
('lost', 'and'),  
('and', 'this'),  
('this', 'i'),  
('i', 'take'),  
('take', 'it'),  
('it', 'is'),  
('is', 'the'),  
('the', 'maine'),  
('maine', 'motiue'),  
('motiue', 'of'),  
('of', 'our'),  
('our', 'preparations'),  
('preparations', 'the'),  
('the', 'sourse'),  
('sourse', 'of'),  
('of', 'this'),  
('this', 'our'),  
('our', 'watch'),  
('watch', 'and'),  
('and', 'the'),  
('the', 'cheefe'),  
('cheefe', 'head'),  
('head', 'of'),  
('of', 'this'),  
('this', 'post'),  
('post', 'hast'),  
('hast', 'and'),  
('and', 'romage'),  
('romage', 'in'),  
('in', 'the'),  
('the', 'land'),  
('enter', 'ghost'),  
('ghost', 'again'),  
('but', 'soft'),  
('soft', 'behold'),  
('behold', 'loe'),  
('loe', 'where'),  
('where', 'it'),  
('it', 'comes'),  
('comes', 'again'),  
('again', 'ile'),  
('ile', 'crosse'),  
('crosse', 'it'),

('it', 'though'),  
('though', 'it'),  
('it', 'blast'),  
('blast', 'me'),  
('stay', 'illusion'),  
('illusion', 'if'),  
('if', 'thou'),  
('thou', 'hast'),  
('hast', 'any'),  
('any', 'sound'),  
('sound', 'or'),  
('or', 'vse'),  
('vse', 'of'),  
('of', 'voyce'),  
('voyce', 'speake'),  
('speake', 'to'),  
('to', 'me'),  
('if', 'there'),  
('there', 'be'),  
('be', 'any'),  
('any', 'good'),  
('good', 'thing'),  
('thing', 'to'),  
('to', 'be'),  
('be', 'done'),  
('done', 'that'),  
('that', 'may'),  
('may', 'to'),  
('to', 'thee'),  
('thee', 'do'),  
('do', 'ease'),  
('ease', 'and'),  
('and', 'grace'),  
('grace', 'to'),  
('to', 'me'),  
('me', 'speak'),  
('speak', 'to'),  
('to', 'me'),  
('if', 'thou'),  
('thou', 'art'),  
('art', 'priuy'),  
('priuy', 'to'),  
('to', 'thy'),  
('thy', 'countries'),  
('countries', 'fate'),  
('fate', 'which'),  
('which', 'happily'),  
('happily', 'foreknowing'),  
('foreknowing', 'may'),

('may', 'auoyd'),  
('auoyd', 'oh'),  
('oh', 'speake'),  
('or', 'if'),  
('if', 'thou'),  
('thou', 'hast'),  
('hast', 'vp'),  
('vp', 'hoorded'),  
('hoorded', 'in'),  
('in', 'thy'),  
('thy', 'life'),  
('life', 'extorted'),  
('extorted', 'treasure'),  
('treasure', 'in'),  
('in', 'the'),  
('the', 'wombe'),  
('wombe', 'of'),  
('of', 'earth'),  
('earth', 'for'),  
('for', 'which'),  
('which', 'they'),  
('they', 'say'),  
('say', 'you'),  
('you', 'spirits'),  
('spirits', 'oft'),  
('oft', 'walke'),  
('walke', 'in'),  
('in', 'death'),  
('death', 'speake'),  
('speake', 'of'),  
('of', 'it'),  
('stay', 'and'),  
('and', 'speake'),  
('stop', 'it'),  
('it', 'marcellus'),  
('shall', 'i'),  
('i', 'strike'),  
('strike', 'at'),  
('at', 'it'),  
('it', 'with'),  
('with', 'my'),  
('my', 'partizan'),  
('do', 'if'),  
('if', 'it'),  
('it', 'will'),  
('will', 'not'),  
('not', 'stand'),  
('tis', 'heere'),  
('tis', 'heere'),

('tis', 'gone'),  
('exit', 'ghost'),  
('we', 'do'),  
('do', 'it'),  
('it', 'wrong'),  
('wrong', 'being'),  
('being', 'so'),  
('so', 'maiesticall'),  
('maiesticall', 'to'),  
('to', 'offer'),  
('offer', 'it'),  
('it', 'the'),  
('the', 'shew'),  
('shew', 'of'),  
('of', 'violence'),  
('violence', 'for'),  
('for', 'it'),  
('it', 'is'),  
('is', 'as'),  
('as', 'the'),  
('the', 'ayre'),  
('ayre', 'invulnerable'),  
('invulnerable', 'and'),  
('and', 'our'),  
('our', 'vaine'),  
('vaine', 'blowes'),  
('blowes', 'malicious'),  
('malicious', 'mockery'),  
('it', 'was'),  
('was', 'about'),  
('about', 'to'),  
('to', 'speake'),  
('speake', 'when'),  
('when', 'the'),  
('the', 'cocke'),  
('cocke', 'crew'),  
('and', 'then'),  
('then', 'it'),  
('it', 'started'),  
('started', 'like'),  
('like', 'a'),  
('a', 'guilty'),  
('guilty', 'thing'),  
('thing', 'vpon'),  
('vpon', 'a'),  
('a', 'fearfull'),  
('fearfull', 'summons'),  
('i', 'haue'),  
('haue', 'heard'),

('heard', 'the'),  
('the', 'cocke'),  
('cocke', 'that'),  
('that', 'is'),  
('is', 'the'),  
('the', 'trumpet'),  
('trumpet', 'to'),  
('to', 'the'),  
('the', 'day'),  
('day', 'doth'),  
('doth', 'with'),  
('with', 'his'),  
('his', 'lofty'),  
('lofty', 'and'),  
('and', 'shrill'),  
('shrill', 'sounding'),  
('sounding', 'throate'),  
('throate', 'awake'),  
('awake', 'the'),  
('the', 'god'),  
('god', 'of'),  
('of', 'day'),  
('day', 'and'),  
('and', 'at'),  
('at', 'his'),  
('his', 'warning'),  
('warning', 'whether'),  
('whether', 'in'),  
('in', 'sea'),  
('sea', 'or'),  
('or', 'fire'),  
('fire', 'in'),  
('in', 'earth'),  
('earth', 'or'),  
('or', 'ayre'),  
('ayre', 'th'),  
('th', 'extrauagant'),  
('extrauagant', 'and'),  
('and', 'erring'),  
('erring', 'spirit'),  
('spirit', 'hyes'),  
('hyes', 'to'),  
('to', 'his'),  
('his', 'confine'),  
('and', 'of'),  
('of', 'the'),  
('the', 'truth'),  
('truth', 'heerein'),  
('heerein', 'this'),

```
('this', 'present'),  
('present', 'object'),  
...]
```

### #Creating Trigrams

```
trigrams = [gram for sentence in processed_shakespeare_corpus for gram  
in generate_ngrams(sentence, 3)]
```

trigrams

```
[('the', 'tragedie', 'of'),  
 ('tragedie', 'of', 'hamlet'),  
 ('of', 'hamlet', 'by'),  
 ('hamlet', 'by', 'william'),  
 ('by', 'william', 'shakespeare'),  
 ('william', 'shakespeare', '1599'),  
 ('enter', 'barnardo', 'and'),  
 ('barnardo', 'and', 'francisco'),  
 ('and', 'francisco', 'two'),  
 ('francisco', 'two', 'centinels'),  
 ('who', 's', 'there'),  
 ('nay', 'answer', 'me'),  
 ('answer', 'me', 'stand'),  
 ('me', 'stand', 'vnfold'),  
 ('stand', 'vnfold', 'your'),  
 ('vnfold', 'your', 'selfe'),  
 ('long', 'liue', 'the'),  
 ('liue', 'the', 'king'),  
 ('you', 'come', 'most'),  
 ('come', 'most', 'carefully'),  
 ('most', 'carefully', 'vpon'),  
 ('carefully', 'vpon', 'your'),  
 ('vpon', 'your', 'houre'),  
 ('tis', 'now', 'strook'),  
 ('now', 'strook', 'twelue'),  
 ('strook', 'twelue', 'get'),  
 ('twelue', 'get', 'thee'),  
 ('get', 'thee', 'to'),  
 ('thee', 'to', 'bed'),  
 ('to', 'bed', 'francisco'),  
 ('for', 'this', 'releefe'),  
 ('this', 'releefe', 'much'),  
 ('releefe', 'much', 'thankes'),  
 ('much', 'thankes', 'tis'),  
 ('thankes', 'tis', 'bitter'),  
 ('tis', 'bitter', 'cold'),  
 ('bitter', 'cold', 'and'),  
 ('cold', 'and', 'i'),
```



('and', 'i', 'am'),  
('i', 'am', 'sicke'),  
('am', 'sicke', 'at'),  
('sicke', 'at', 'heart'),  
('haue', 'you', 'had'),  
('you', 'had', 'quiet'),  
('had', 'quiet', 'guard'),  
('not', 'a', 'mouse'),  
('a', 'mouse', 'stirring'),  
('if', 'you', 'do'),  
('you', 'do', 'meet'),  
('do', 'meet', 'horatio'),  
('meet', 'horatio', 'and'),  
('horatio', 'and', 'marcellus'),  
('and', 'marcellus', 'the'),  
('marcellus', 'the', 'riuals'),  
('the', 'riuals', 'of'),  
('riuals', 'of', 'my'),  
('of', 'my', 'watch'),  
('my', 'watch', 'bid'),  
('watch', 'bid', 'them'),  
('bid', 'them', 'make'),  
('them', 'make', 'hast'),  
('enter', 'horatio', 'and'),  
('horatio', 'and', 'marcellus'),  
('i', 'thinke', 'i'),  
('thinke', 'i', 'heare'),  
('i', 'heare', 'them'),  
('stand', 'who', 's'),  
('who', 's', 'there'),  
('friends', 'to', 'this'),  
('to', 'this', 'ground'),  
('and', 'leige', 'men'),  
('leige', 'men', 'to'),  
('men', 'to', 'the'),  
('to', 'the', 'dane'),  
('giue', 'you', 'good'),  
('you', 'good', 'night'),  
('o', 'farwel', 'honest'),  
('farwel', 'honest', 'soldier'),  
('honest', 'soldier', 'who'),  
('soldier', 'who', 'hath'),  
('who', 'hath', 'relieu'),  
('hath', 'relieu', 'd'),  
('relieu', 'd', 'you'),  
('barnardo', 'ha', 's'),  
('ha', 's', 'my'),  
('s', 'my', 'place'),  
('my', 'place', 'giue'),

('place', 'giue', 'you'),  
('giue', 'you', 'goodnight'),  
('say', 'what', 'is'),  
('what', 'is', 'horatio'),  
('is', 'horatio', 'there'),  
('a', 'peece', 'of'),  
('peece', 'of', 'him'),  
('welcome', 'horatio', 'welcome'),  
('horatio', 'welcome', 'good'),  
('welcome', 'good', 'marcellus'),  
('what', 'ha', 's'),  
('ha', 's', 'this'),  
('s', 'this', 'thing'),  
('this', 'thing', 'appear'),  
('thing', 'appear', 'd'),  
('appear', 'd', 'again'),  
('d', 'again', 'to'),  
('again', 'to', 'night'),  
('i', 'haue', 'seene'),  
('haue', 'seene', 'nothing'),  
('horatio', 'saies', 'tis'),  
('saies', 'tis', 'but'),  
('tis', 'but', 'our'),  
('but', 'our', 'fantasie'),  
('our', 'fantasie', 'and'),  
('fantasie', 'and', 'will'),  
('and', 'will', 'not'),  
('will', 'not', 'let'),  
('not', 'let', 'beleefe'),  
('let', 'beleefe', 'take'),  
('beleefe', 'take', 'hold'),  
('take', 'hold', 'of'),  
('hold', 'of', 'him'),  
('of', 'him', 'touching'),  
('him', 'touching', 'this'),  
('touching', 'this', 'dreaded'),  
('this', 'dreaded', 'sight'),  
('dreaded', 'sight', 'twice'),  
('sight', 'twice', 'seene'),  
('twice', 'seene', 'of'),  
('seene', 'of', 'vs'),  
('of', 'vs', 'therefore'),  
('vs', 'therefore', 'i'),  
('therefore', 'i', 'haue'),  
('i', 'haue', 'intreated'),  
('haue', 'intreated', 'him'),  
('intreated', 'him', 'along'),  
('him', 'along', 'with'),  
('along', 'with', 'vs'),

('with', 'vs', 'to'),  
('vs', 'to', 'watch'),  
('to', 'watch', 'the'),  
('watch', 'the', 'minutes'),  
('the', 'minutes', 'of'),  
('minutes', 'of', 'this'),  
('of', 'this', 'night'),  
('this', 'night', 'that'),  
('night', 'that', 'if'),  
('that', 'if', 'again'),  
('if', 'again', 'this'),  
('again', 'this', 'apparition'),  
('this', 'apparition', 'come'),  
('apparition', 'come', 'he'),  
('come', 'he', 'may'),  
('he', 'may', 'approve'),  
('may', 'approve', 'our'),  
('approve', 'our', 'eyes'),  
('our', 'eyes', 'and'),  
('eyes', 'and', 'speak'),  
('and', 'speak', 'to'),  
('speak', 'to', 'it'),  
('tush', 'tush', 'twill'),  
('tush', 'twill', 'not'),  
('twill', 'not', 'appear'),  
('sit', 'down', 'a'),  
('down', 'a', 'while'),  
('a', 'while', 'and'),  
('while', 'and', 'let'),  
('and', 'let', 'vs'),  
('let', 'vs', 'once'),  
('vs', 'once', 'again'),  
('once', 'again', 'assail'),  
('again', 'assail', 'your'),  
('assail', 'your', 'ears'),  
('your', 'ears', 'that'),  
('ears', 'that', 'are'),  
('that', 'are', 'so'),  
('are', 'so', 'fortified'),  
('so', 'fortified', 'against'),  
('fortified', 'against', 'our'),  
('against', 'our', 'story'),  
('our', 'story', 'what'),  
('story', 'what', 'we'),  
('what', 'we', 'two'),  
('we', 'two', 'nights'),  
('two', 'nights', 'have'),  
('nights', 'have', 'scene'),  
('well', 'sit', 'we'),

('sit', 'we', 'downe'),  
('we', 'downe', 'and'),  
('downe', 'and', 'let'),  
('and', 'let', 'vs'),  
('let', 'vs', 'heare'),  
('vs', 'heare', 'barnardo'),  
('heare', 'barnardo', 'speake'),  
('barnardo', 'speake', 'of'),  
('speake', 'of', 'this'),  
('last', 'night', 'of'),  
('night', 'of', 'all'),  
('of', 'all', 'when'),  
('all', 'when', 'yond'),  
('when', 'yond', 'same'),  
('yond', 'same', 'starre'),  
('same', 'starre', 'that'),  
('starre', 'that', 's'),  
('that', 's', 'westward'),  
('s', 'westward', 'from'),  
('westward', 'from', 'the'),  
('from', 'the', 'pole'),  
('the', 'pole', 'had'),  
('pole', 'had', 'made'),  
('had', 'made', 'his'),  
('made', 'his', 'course'),  
('his', 'course', 't'),  
('course', 't', 'illume'),  
('t', 'illume', 'that'),  
('illume', 'that', 'part'),  
('that', 'part', 'of'),  
('part', 'of', 'heauen'),  
('of', 'heauen', 'where'),  
('heauen', 'where', 'now'),  
('where', 'now', 'it'),  
('now', 'it', 'burnes'),  
('it', 'burnes', 'marcellus'),  
('burnes', 'marcellus', 'and'),  
('marcellus', 'and', 'my'),  
('and', 'my', 'selfe'),  
('my', 'selfe', 'the'),  
('selfe', 'the', 'bell'),  
('the', 'bell', 'then'),  
('bell', 'then', 'beating'),  
('then', 'beating', 'one'),  
('peace', 'breake', 'thee'),  
('breake', 'thee', 'of'),  
('thee', 'of', 'enter'),  
('of', 'enter', 'the'),  
('enter', 'the', 'ghost'),

('looke', 'where', 'it'),  
('where', 'it', 'comes'),  
('it', 'comes', 'again'),  
('in', 'the', 'same'),  
('the', 'same', 'figure'),  
('same', 'figure', 'like'),  
('figure', 'like', 'the'),  
('like', 'the', 'king'),  
('the', 'king', 'that'),  
('king', 'that', 's'),  
('that', 's', 'dead'),  
('thou', 'art', 'a'),  
('art', 'a', 'scholler'),  
('a', 'scholler', 'speake'),  
('scholler', 'speake', 'to'),  
('speake', 'to', 'it'),  
('to', 'it', 'horatio'),  
('lookes', 'it', 'not'),  
('it', 'not', 'like'),  
('not', 'like', 'the'),  
('like', 'the', 'king'),  
('marke', 'it', 'horatio'),  
('most', 'like', 'it'),  
('like', 'it', 'harrowes'),  
('it', 'harrowes', 'me'),  
('harrowes', 'me', 'with'),  
('me', 'with', 'fear'),  
('with', 'fear', 'wonder'),  
('fear', 'wonder', 'barn'),  
('it', 'would', 'be'),  
('would', 'be', 'spoke'),  
('be', 'spoke', 'too'),  
('question', 'it', 'horatio'),  
('what', 'art', 'thou'),  
('art', 'thou', 'that'),  
('thou', 'that', 'vsurp'),  
('that', 'vsurp', 'st'),  
('vsurp', 'st', 'this'),  
('st', 'this', 'time'),  
('this', 'time', 'of'),  
('time', 'of', 'night'),  
('of', 'night', 'together'),  
('night', 'together', 'with'),  
('together', 'with', 'that'),  
('with', 'that', 'faire'),  
('that', 'faire', 'and'),  
('faire', 'and', 'warlike'),  
('and', 'warlike', 'forme'),  
('warlike', 'forme', 'in'),

('forme', 'in', 'which'),  
('in', 'which', 'the'),  
('which', 'the', 'maiesty'),  
('the', 'maiesty', 'of'),  
('maiesty', 'of', 'buried'),  
('of', 'buried', 'denmarke'),  
('buried', 'denmarke', 'did'),  
('denmarke', 'did', 'sometimes'),  
('did', 'sometimes', 'march'),  
('sometimes', 'march', 'by'),  
('march', 'by', 'heauen'),  
('by', 'heauen', 'i'),  
('heauen', 'i', 'charge'),  
('i', 'charge', 'thee'),  
('charge', 'thee', 'speake'),  
('it', 'is', 'offended'),  
('see', 'it', 'stalkes'),  
('it', 'stalkes', 'away'),  
('stay', 'speake', 'speake'),  
('speake', 'speake', 'i'),  
('speake', 'i', 'charge'),  
('i', 'charge', 'thee'),  
('charge', 'thee', 'speake'),  
('exit', 'the', 'ghost'),  
('tis', 'gone', 'and'),  
('gone', 'and', 'will'),  
('and', 'will', 'not'),  
('will', 'not', 'answer'),  
('how', 'now', 'horatio'),  
('you', 'tremble', 'look'),  
('tremble', 'look', 'pale'),  
('look', 'pale', 'is'),  
('pale', 'is', 'not'),  
('is', 'not', 'this'),  
('not', 'this', 'something'),  
('this', 'something', 'more'),  
('something', 'more', 'then'),  
('more', 'then', 'fantasie'),  
('what', 'thinke', 'you'),  
('thinke', 'you', 'on'),  
('you', 'on', 't'),  
('before', 'my', 'god'),  
('my', 'god', 'i'),  
('god', 'i', 'might'),  
('i', 'might', 'not'),  
('might', 'not', 'this'),  
('not', 'this', 'beleeue'),  
('this', 'beleeue', 'without'),  
('beleeue', 'without', 'the'),

('without', 'the', 'sensible'),  
('the', 'sensible', 'and'),  
('sensible', 'and', 'true'),  
('and', 'true', 'auouch'),  
('true', 'auouch', 'of'),  
('auouch', 'of', 'mine'),  
('of', 'mine', 'owne'),  
('mine', 'owne', 'eyes'),  
('is', 'it', 'not'),  
('it', 'not', 'like'),  
('not', 'like', 'the'),  
('like', 'the', 'king'),  
('as', 'thou', 'art'),  
('thou', 'art', 'to'),  
('art', 'to', 'thy'),  
('to', 'thy', 'selfe'),  
('thy', 'selfe', 'such'),  
('selfe', 'such', 'was'),  
('such', 'was', 'the'),  
('was', 'the', 'very'),  
('the', 'very', 'armour'),  
('very', 'armour', 'he'),  
('armour', 'he', 'had'),  
('he', 'had', 'on'),  
('had', 'on', 'when'),  
('on', 'when', 'th'),  
('when', 'th', 'ambitious'),  
('th', 'ambitious', 'norwey'),  
('ambitious', 'norwey', 'combatted'),  
('norwey', 'combatted', 'so'),  
('combatted', 'so', 'frown'),  
('so', 'frown', 'd'),  
('frown', 'd', 'he'),  
('d', 'he', 'once'),  
('he', 'once', 'when'),  
('once', 'when', 'in'),  
('when', 'in', 'an'),  
('in', 'an', 'angry'),  
('an', 'angry', 'parle'),  
('angry', 'parle', 'he'),  
('parle', 'he', 'smot'),  
('he', 'smot', 'the'),  
('smot', 'the', 'sledded'),  
('the', 'sledded', 'pollax'),  
('sledded', 'pollax', 'on'),  
('pollax', 'on', 'the'),  
('on', 'the', 'ice'),  
('thus', 'twice', 'before'),  
('twice', 'before', 'and'),

('before', 'and', 'iust'),  
('and', 'iust', 'at'),  
('iust', 'at', 'this'),  
('at', 'this', 'dead'),  
('this', 'dead', 'houre'),  
('dead', 'houre', 'with'),  
('houre', 'with', 'martiall'),  
('with', 'martiall', 'stalke'),  
('martiall', 'stalke', 'hath'),  
('stalke', 'hath', 'he'),  
('hath', 'he', 'gone'),  
('he', 'gone', 'by'),  
('gone', 'by', 'our'),  
('by', 'our', 'watch'),  
('in', 'what', 'particular'),  
('what', 'particular', 'thought'),  
('particular', 'thought', 'to'),  
('thought', 'to', 'work'),  
('to', 'work', 'i'),  
('work', 'i', 'know'),  
('i', 'know', 'not'),  
('know', 'not', 'but'),  
('not', 'but', 'in'),  
('but', 'in', 'the'),  
('in', 'the', 'grosse'),  
('the', 'grosse', 'and'),  
('grosse', 'and', 'scope'),  
('and', 'scope', 'of'),  
('scope', 'of', 'my'),  
('of', 'my', 'opinion'),  
('my', 'opinion', 'this'),  
('opinion', 'this', 'boades'),  
('this', 'boades', 'some'),  
('boades', 'some', 'strange'),  
('some', 'strange', 'erruption'),  
('strange', 'erruption', 'to'),  
('erruption', 'to', 'our'),  
('to', 'our', 'state'),  
('good', 'now', 'sit'),  
('now', 'sit', 'downe'),  
('sit', 'downe', 'tell'),  
('downe', 'tell', 'me'),  
('tell', 'me', 'he'),  
('me', 'he', 'that'),  
('he', 'that', 'knowes'),  
('that', 'knowes', 'why'),  
('knowes', 'why', 'this'),  
('why', 'this', 'same'),  
('this', 'same', 'strict'),



('same', 'strict', 'and'),  
('strict', 'and', 'most'),  
('and', 'most', 'obseruant'),  
('most', 'obseruant', 'watch'),  
('obseruant', 'watch', 'so'),  
('watch', 'so', 'nightly'),  
('so', 'nightly', 'toyles'),  
('nightly', 'toyles', 'the'),  
('toyles', 'the', 'subiect'),  
('the', 'subiect', 'of'),  
('subiect', 'of', 'the'),  
('of', 'the', 'land'),  
('the', 'land', 'and'),  
('land', 'and', 'why'),  
('and', 'why', 'such'),  
('why', 'such', 'dayly'),  
('such', 'dayly', 'cast'),  
('dayly', 'cast', 'of'),  
('cast', 'of', 'brazon'),  
('of', 'brazon', 'cannon'),  
('brazon', 'cannon', 'and'),  
('cannon', 'and', 'forraigne'),  
('and', 'forraigne', 'mart'),  
('forraigne', 'mart', 'for'),  
('mart', 'for', 'implements'),  
('for', 'implements', 'of'),  
('implements', 'of', 'warre'),  
('of', 'warre', 'why'),  
('warre', 'why', 'such'),  
('why', 'such', 'impresse'),  
('such', 'impresse', 'of'),  
('impresse', 'of', 'ship'),  
('of', 'ship', 'wrights'),  
('ship', 'wrights', 'whose'),  
('wrights', 'whose', 'sore'),  
('whose', 'sore', 'taske'),  
('sore', 'taske', 'do'),  
('taske', 'do', 's'),  
('do', 's', 'not'),  
('s', 'not', 'diuide'),  
('not', 'diuide', 'the'),  
('diuide', 'the', 'sunday'),  
('the', 'sunday', 'from'),  
('sunday', 'from', 'the'),  
('from', 'the', 'weeke'),  
('the', 'weeke', 'what'),  
('weeke', 'what', 'might'),  
('what', 'might', 'be'),  
('might', 'be', 'toward'),

('be', 'toward', 'that'),  
('toward', 'that', 'this'),  
('that', 'this', 'sweaty'),  
('this', 'sweaty', 'hast'),  
('sweaty', 'hast', 'doth'),  
('hast', 'doth', 'make'),  
('doth', 'make', 'the'),  
('make', 'the', 'night'),  
('the', 'night', 'ioynt'),  
('night', 'ioynt', 'labourer'),  
('ioynt', 'labourer', 'with'),  
('labourer', 'with', 'the'),  
('with', 'the', 'day'),  
('the', 'day', 'who'),  
('day', 'who', 'is'),  
('who', 'is', 't'),  
('is', 't', 'that'),  
('t', 'that', 'can'),  
('that', 'can', 'informe'),  
('can', 'informe', 'me'),  
('that', 'can', 'i'),  
('can', 'i', 'at'),  
('i', 'at', 'least'),  
('at', 'least', 'the'),  
('least', 'the', 'whisper'),  
('the', 'whisper', 'goes'),  
('whisper', 'goes', 'so'),  
('goes', 'so', 'our'),  
('so', 'our', 'last'),  
('our', 'last', 'king'),  
('last', 'king', 'whose'),  
('king', 'whose', 'image'),  
('whose', 'image', 'euen'),  
('image', 'euen', 'but'),  
('euen', 'but', 'now'),  
('but', 'now', 'appear'),  
('now', 'appear', 'd'),  
('appear', 'd', 'to'),  
('d', 'to', 'vs'),  
('to', 'vs', 'was'),  
('vs', 'was', 'as'),  
('was', 'as', 'you'),  
('as', 'you', 'know'),  
('you', 'know', 'by'),  
('know', 'by', 'fortinbras'),  
('by', 'fortinbras', 'of'),  
('fortinbras', 'of', 'norway'),  
('of', 'norway', 'thereto'),  
('norway', 'thereto', 'prick'),  
('thereto', 'prick', 'd'),

('prick', 'd', 'on'),  
('d', 'on', 'by'),  
('on', 'by', 'a'),  
('by', 'a', 'most'),  
('a', 'most', 'emulate'),  
('most', 'emulate', 'pride'),  
('emulate', 'pride', 'dar'),  
('pride', 'dar', 'd'),  
('dar', 'd', 'to'),  
('d', 'to', 'the'),  
('to', 'the', 'combate'),  
('in', 'which', 'our'),  
('which', 'our', 'valiant'),  
('our', 'valiant', 'hamlet'),  
('valiant', 'hamlet', 'for'),  
('hamlet', 'for', 'so'),  
('for', 'so', 'this'),  
('so', 'this', 'side'),  
('this', 'side', 'of'),  
('side', 'of', 'our'),  
('of', 'our', 'knowne'),  
('our', 'knowne', 'world'),  
('knowne', 'world', 'esteem'),  
('world', 'esteem', 'd'),  
('esteem', 'd', 'him'),  
('d', 'him', 'did'),  
('him', 'did', 'slay'),  
('did', 'slay', 'this'),  
('slay', 'this', 'fortinbras'),  
('this', 'fortinbras', 'who'),  
('fortinbras', 'who', 'by'),  
('who', 'by', 'a'),  
('by', 'a', 'seal'),  
('a', 'seal', 'd'),  
('seal', 'd', 'compact'),  
('d', 'compact', 'well'),  
('compact', 'well', 'ratified'),  
('well', 'ratified', 'by'),  
('ratified', 'by', 'law'),  
('by', 'law', 'and'),  
('law', 'and', 'heraldrie'),  
('and', 'heraldrie', 'did'),  
('heraldrie', 'did', 'forfeite'),  
('did', 'forfeite', 'with'),  
('forfeite', 'with', 'his'),  
('with', 'his', 'life'),  
('his', 'life', 'all'),  
('life', 'all', 'those'),  
('all', 'those', 'his'),

('those', 'his', 'lands'),  
('his', 'lands', 'which'),  
('lands', 'which', 'he'),  
('which', 'he', 'stood'),  
('he', 'stood', 'seiz'),  
('stood', 'seiz', 'd'),  
('seiz', 'd', 'on'),  
('d', 'on', 'to'),  
('on', 'to', 'the'),  
('to', 'the', 'conqueror'),  
('the', 'conqueror', 'against'),  
('conqueror', 'against', 'the'),  
('against', 'the', 'which'),  
('the', 'which', 'a'),  
('which', 'a', 'moity'),  
('a', 'moity', 'competent'),  
('moity', 'competent', 'was'),  
('competent', 'was', 'gaged'),  
('was', 'gaged', 'by'),  
('gaged', 'by', 'our'),  
('by', 'our', 'king'),  
('our', 'king', 'which'),  
('king', 'which', 'had'),  
('which', 'had', 'return'),  
('had', 'return', 'd'),  
('return', 'd', 'to'),  
('d', 'to', 'the'),  
('to', 'the', 'inheritance'),  
('the', 'inheritance', 'of'),  
('inheritance', 'of', 'fortinbras'),  
('of', 'fortinbras', 'had'),  
('fortinbras', 'had', 'he'),  
('had', 'he', 'bin'),  
('he', 'bin', 'vanquisher'),  
('bin', 'vanquisher', 'as'),  
('vanquisher', 'as', 'by'),  
('as', 'by', 'the'),  
('by', 'the', 'same'),  
('the', 'same', 'cou'),  
('same', 'cou', 'nant'),  
('cou', 'nant', 'and'),  
('nant', 'and', 'carriage'),  
('and', 'carriage', 'of'),  
('carriage', 'of', 'the'),  
('of', 'the', 'article'),  
('the', 'article', 'designe'),  
('article', 'designe', 'his'),  
('designe', 'his', 'fell'),  
('his', 'fell', 'to'),

('fell', 'to', 'hamlet'),  
('now', 'sir', 'young'),  
('sir', 'young', 'fortinbras'),  
('young', 'fortinbras', 'of'),  
('fortinbras', 'of', 'vnimproued'),  
('of', 'vnimproued', 'mettle'),  
('vnimproued', 'mettle', 'hot'),  
('mettle', 'hot', 'and'),  
('hot', 'and', 'full'),  
('and', 'full', 'hath'),  
('full', 'hath', 'in'),  
('hath', 'in', 'the'),  
('in', 'the', 'skirts'),  
('the', 'skirts', 'of'),  
('skirts', 'of', 'norway'),  
('of', 'norway', 'heere'),  
('norway', 'heere', 'and'),  
('heere', 'and', 'there'),  
('and', 'there', 'shark'),  
('there', 'shark', 'd'),  
('shark', 'd', 'vp'),  
('d', 'vp', 'a'),  
('vp', 'a', 'list'),  
('a', 'list', 'of'),  
('list', 'of', 'landlesse'),  
('of', 'landlesse', 'resolutes'),  
('landlesse', 'resolutes', 'for'),  
('resolutes', 'for', 'foode'),  
('for', 'foode', 'and'),  
('foode', 'and', 'diet'),  
('and', 'diet', 'to'),  
('diet', 'to', 'some'),  
('to', 'some', 'enterprize'),  
('some', 'enterprize', 'that'),  
('enterprize', 'that', 'hath'),  
('that', 'hath', 'a'),  
('hath', 'a', 'stomacke'),  
('a', 'stomacke', 'in'),  
('stomacke', 'in', 't'),  
('in', 't', 'which'),  
('t', 'which', 'is'),  
('which', 'is', 'no'),  
('is', 'no', 'other'),  
('no', 'other', 'and'),  
('other', 'and', 'it'),  
('and', 'it', 'doth'),  
('it', 'doth', 'well'),  
('doth', 'well', 'appeare'),  
('well', 'appeare', 'vnto'),

('appeare', 'vnto', 'our'),  
('vnto', 'our', 'state'),  
('our', 'state', 'but'),  
('state', 'but', 'to'),  
('but', 'to', 'recouer'),  
('to', 'recouer', 'of'),  
('recouer', 'of', 'vs'),  
('of', 'vs', 'by'),  
('vs', 'by', 'strong'),  
('by', 'strong', 'hand'),  
('strong', 'hand', 'and'),  
('hand', 'and', 'termes'),  
('and', 'termes', 'compulsatiue'),  
('termes', 'compulsatiue', 'those'),  
('compulsatiue', 'those', 'foresaid'),  
('those', 'foresaid', 'lands'),  
('foresaid', 'lands', 'so'),  
('lands', 'so', 'by'),  
('so', 'by', 'his'),  
('by', 'his', 'father'),  
('his', 'father', 'lost'),  
('father', 'lost', 'and'),  
('lost', 'and', 'this'),  
('and', 'this', 'i'),  
('this', 'i', 'take'),  
('i', 'take', 'it'),  
('take', 'it', 'is'),  
('it', 'is', 'the'),  
('is', 'the', 'maine'),  
('the', 'maine', 'motiue'),  
('maine', 'motiue', 'of'),  
('motiue', 'of', 'our'),  
('of', 'our', 'preparations'),  
('our', 'preparations', 'the'),  
('preparations', 'the', 'source'),  
('the', 'source', 'of'),  
('source', 'of', 'this'),  
('of', 'this', 'our'),  
('this', 'our', 'watch'),  
('our', 'watch', 'and'),  
('watch', 'and', 'the'),  
('and', 'the', 'cheefe'),  
('the', 'cheefe', 'head'),  
('cheefe', 'head', 'of'),  
('head', 'of', 'this'),  
('of', 'this', 'post'),  
('this', 'post', 'hast'),  
('post', 'hast', 'and'),  
('hast', 'and', 'romage'),

('and', 'romage', 'in'),  
('romage', 'in', 'the'),  
('in', 'the', 'land'),  
('enter', 'ghost', 'again'),  
('but', 'soft', 'behold'),  
('soft', 'behold', 'lo'),  
('behold', 'lo', 'where'),  
('lo', 'where', 'it'),  
('where', 'it', 'comes'),  
('it', 'comes', 'again'),  
('comes', 'again', 'ile'),  
('again', 'ile', 'crosse'),  
('ile', 'crosse', 'it'),  
('crosse', 'it', 'though'),  
('it', 'though', 'it'),  
('though', 'it', 'blast'),  
('it', 'blast', 'me'),  
('stay', 'illusion', 'if'),  
('illusion', 'if', 'thou'),  
('if', 'thou', 'hast'),  
('thou', 'hast', 'any'),  
('hast', 'any', 'sound'),  
('any', 'sound', 'or'),  
('sound', 'or', 'vse'),  
('or', 'vse', 'of'),  
('vse', 'of', 'voyce'),  
('of', 'voyce', 'speake'),  
('voyce', 'speake', 'to'),  
('speake', 'to', 'me'),  
('if', 'there', 'be'),  
('there', 'be', 'any'),  
('be', 'any', 'good'),  
('any', 'good', 'thing'),  
('good', 'thing', 'to'),  
('thing', 'to', 'be'),  
('to', 'be', 'done'),  
('be', 'done', 'that'),  
('done', 'that', 'may'),  
('that', 'may', 'to'),  
('may', 'to', 'thee'),  
('to', 'thee', 'do'),  
('thee', 'do', 'ease'),  
('do', 'ease', 'and'),  
('ease', 'and', 'grace'),  
('and', 'grace', 'to'),  
('grace', 'to', 'me'),  
('to', 'me', 'speak'),  
('me', 'speak', 'to'),  
('speak', 'to', 'me'),

('if', 'thou', 'art'),  
('thou', 'art', 'priuy'),  
('art', 'priuy', 'to'),  
('priuy', 'to', 'thy'),  
('to', 'thy', 'countries'),  
('thy', 'countries', 'fate'),  
('countries', 'fate', 'which'),  
('fate', 'which', 'happily'),  
('which', 'happily', 'foreknowing'),  
('happily', 'foreknowing', 'may'),  
('foreknowing', 'may', 'auoyd'),  
('may', 'auoyd', 'oh'),  
('auoyd', 'oh', 'speake'),  
('or', 'if', 'thou'),  
('if', 'thou', 'hast'),  
('thou', 'hast', 'vp'),  
('hast', 'vp', 'hoorded'),  
('vp', 'hoorded', 'in'),  
('hoorded', 'in', 'thy'),  
('in', 'thy', 'life'),  
('thy', 'life', 'extorted'),  
('life', 'extorted', 'treasure'),  
('extorted', 'treasure', 'in'),  
('treasure', 'in', 'the'),  
('in', 'the', 'wombe'),  
('the', 'wombe', 'of'),  
('wombe', 'of', 'earth'),  
('of', 'earth', 'for'),  
('earth', 'for', 'which'),  
('for', 'which', 'they'),  
('which', 'they', 'say'),  
('they', 'say', 'you'),  
('say', 'you', 'spirits'),  
('you', 'spirits', 'oft'),  
('spirits', 'oft', 'walke'),  
('oft', 'walke', 'in'),  
('walke', 'in', 'death'),  
('in', 'death', 'speake'),  
('death', 'speake', 'of'),  
('speake', 'of', 'it'),  
('stay', 'and', 'speake'),  
('stop', 'it', 'marcellus'),  
('shall', 'i', 'strike'),  
('i', 'strike', 'at'),  
('strike', 'at', 'it'),  
('at', 'it', 'with'),  
('it', 'with', 'my'),  
('with', 'my', 'partizan'),  
('do', 'if', 'it'),



('if', 'it', 'will'),  
('it', 'will', 'not'),  
('will', 'not', 'stand'),  
('we', 'do', 'it'),  
('do', 'it', 'wrong'),  
('it', 'wrong', 'being'),  
('wrong', 'being', 'so'),  
('being', 'so', 'maiesticall'),  
('so', 'maiesticall', 'to'),  
('maiesticall', 'to', 'offer'),  
('to', 'offer', 'it'),  
('offer', 'it', 'the'),  
('it', 'the', 'shew'),  
('the', 'shew', 'of'),  
('shew', 'of', 'violence'),  
('of', 'violence', 'for'),  
('violence', 'for', 'it'),  
('for', 'it', 'is'),  
('it', 'is', 'as'),  
('is', 'as', 'the'),  
('as', 'the', 'ayre'),  
('the', 'ayre', 'invulnerable'),  
('ayre', 'invulnerable', 'and'),  
('invulnerable', 'and', 'our'),  
('and', 'our', 'vaine'),  
('our', 'vaine', 'blowes'),  
('vaine', 'blowes', 'malicious'),  
('blowes', 'malicious', 'mockery'),  
('it', 'was', 'about'),  
('was', 'about', 'to'),  
('about', 'to', 'speake'),  
('to', 'speake', 'when'),  
('speake', 'when', 'the'),  
('when', 'the', 'cocke'),  
('the', 'cocke', 'crew'),  
('and', 'then', 'it'),  
('then', 'it', 'started'),  
('it', 'started', 'like'),  
('started', 'like', 'a'),  
('like', 'a', 'guilty'),  
('a', 'guilty', 'thing'),  
('guilty', 'thing', 'vpon'),  
('thing', 'vpon', 'a'),  
('vpon', 'a', 'fearfull'),  
('a', 'fearfull', 'summons'),  
('i', 'haue', 'heard'),  
('haue', 'heard', 'the'),  
('heard', 'the', 'cocke'),  
('the', 'cocke', 'that'),

('cocke', 'that', 'is'),  
('that', 'is', 'the'),  
('is', 'the', 'trumpet'),  
('the', 'trumpet', 'to'),  
('trumpet', 'to', 'the'),  
('to', 'the', 'day'),  
('the', 'day', 'doth'),  
('day', 'doth', 'with'),  
('doth', 'with', 'his'),  
('with', 'his', 'lofty'),  
('his', 'lofty', 'and'),  
('lofty', 'and', 'shrill'),  
('and', 'shrill', 'sounding'),  
('shrill', 'sounding', 'throate'),  
('sounding', 'throate', 'awake'),  
('throate', 'awake', 'the'),  
('awake', 'the', 'god'),  
('the', 'god', 'of'),  
('god', 'of', 'day'),  
('of', 'day', 'and'),  
('day', 'and', 'at'),  
('and', 'at', 'his'),  
('at', 'his', 'warning'),  
('his', 'warning', 'whether'),  
('warning', 'whether', 'in'),  
('whether', 'in', 'sea'),  
('in', 'sea', 'or'),  
('sea', 'or', 'fire'),  
('or', 'fire', 'in'),  
('fire', 'in', 'earth'),  
('in', 'earth', 'or'),  
('earth', 'or', 'ayre'),  
('or', 'ayre', 'th'),  
('ayre', 'th', 'extrauagant'),  
('th', 'extrauagant', 'and'),  
('extrauagant', 'and', 'erring'),  
('and', 'erring', 'spirit'),  
('erring', 'spirit', 'hyes'),  
('spirit', 'hyes', 'to'),  
('hyes', 'to', 'his'),  
('to', 'his', 'confine'),  
('and', 'of', 'the'),  
('of', 'the', 'truth'),  
('the', 'truth', 'heerein'),  
('truth', 'heerein', 'this'),  
('heerein', 'this', 'present'),  
('this', 'present', 'object'),  
('present', 'object', 'made'),  
('object', 'made', 'probation'),

('it', 'faded', 'on'),  
('faded', 'on', 'the'),  
('on', 'the', 'crowing'),  
('the', 'crowing', 'of'),  
('crowing', 'of', 'the'),  
('of', 'the', 'cocke'),  
('some', 'sayes', 'that'),  
('sayes', 'that', 'euer'),  
('that', 'euer', 'gainst'),  
('euer', 'gainst', 'that'),  
('gainst', 'that', 'season'),  
('that', 'season', 'comes'),  
('season', 'comes', 'wherein'),  
('comes', 'wherein', 'our'),  
('wherein', 'our', 'sauours'),  
('our', 'sauours', 'birch'),  
('sauours', 'birch', 'is'),  
('birch', 'is', 'celebrated'),  
('is', 'celebrated', 'the'),  
('celebrated', 'the', 'bird'),  
('the', 'bird', 'of'),  
('bird', 'of', 'dawning'),  
('of', 'dawning', 'singeth'),  
('dawning', 'singeth', 'all'),  
('singeth', 'all', 'night'),  
('all', 'night', 'long'),  
('night', 'long', 'and'),  
('long', 'and', 'then'),  
('and', 'then', 'they'),  
('then', 'they', 'say'),  
('they', 'say', 'no'),  
('say', 'no', 'spirit'),  
('no', 'spirit', 'can'),  
('spirit', 'can', 'walke'),  
('can', 'walke', 'abroad'),  
('walke', 'abroad', 'the'),  
('abroad', 'the', 'nights'),  
('the', 'nights', 'are'),  
('nights', 'are', 'wholsome'),  
('are', 'wholsome', 'then'),  
('wholsome', 'then', 'no'),  
('then', 'no', 'planets'),  
('no', 'planets', 'strike'),  
('planets', 'strike', 'no'),  
('strike', 'no', 'faery'),  
('no', 'faery', 'talkes'),  
('faery', 'talkes', 'nor'),  
('talkes', 'nor', 'witch'),  
('nor', 'witch', 'hath'),

```
(
    ('witch', 'hath', 'power'),
    ('hath', 'power', 'to'),
    ('power', 'to', 'charme'),
    ('to', 'charme', 'so'),
    ('charme', 'so', 'hallow'),
    ('so', 'hallow', 'd'),
    ('hallow', 'd', 'and'),
    ('d', 'and', 'so'),
    ('and', 'so', 'gracious'),
    ('so', 'gracious', 'is'),
    ('gracious', 'is', 'the'),
    ('is', 'the', 'time'),
    ('so', 'haue', 'i'),
    ('haue', 'i', 'heard'),
    ('i', 'heard', 'and'),
    ('heard', 'and', 'do'),
    ('and', 'do', 'in'),
    ('do', 'in', 'part'),
    ('in', 'part', 'beleeue'),
    ('part', 'beleeue', 'it'),
    ('but', 'looke', 'the'),
    ('looke', 'the', 'morne'),
    ('the', 'morne', 'in'),
    ('morne', 'in', 'russet'),
    ('in', 'russet', 'mantle'),
    ('russet', 'mantle', 'clad'),
    ('mantle', 'clad', 'walkes'),
    ('clad', 'walkes', 'o'),
    ('walkes', 'o', 're'),
    ('o', 're', 'the'),
    ...]

```

## #Finding Frequency Distribution

```

unigram_freq_dist = FreqDist(unigrams)
bigram_freq_dist = FreqDist(bigrams)
trigram_freq_dist = FreqDist(trigrams)

unigram_freq_dist
FreqDist({'the',): 993, ('and',): 863, ('to',): 685, ('of',): 610,
('i',): 574, ('you',): 527, ('a',): 511, ('my',): 502, ('it',): 419,
('in',): 400, ...})

bigram_freq_dist
FreqDist({'my', 'lord'): 175, ('in', 'the'): 73, ('to', 'the'): 60,
('of', 'the'): 59, ('i', 'haue'): 53, ('it', 'is'): 52, ('i', 'am'):
43, ('the', 'king'): 40, ('that', 'i'): 38, ('i', 'will'): 35, ...})

```

```
trigram_freq_dist
```

```
FreqDist({'my', 'lord', 'i'): 18, ('good', 'my', 'lord'): 14, ('i', 'my', 'lord'): 13, ('i', 'pray', 'you'): 12, ('that', 'i', 'haue'): 11, ('my', 'good', 'lord'): 8, ('it', 'is', 'a'): 8, ('lord', 'i', 'haue'): 7, ('let', 'me', 'see'): 7, ('i', 'haue', 'seene'): 6, ...})
```

```
#Finding Probabilities for Bigrams
```

```
def calculate_bigram_probability(word, next_word):
    bigram_count = bigram_freq_dist[(word, next_word)]
    unigram_count = unigram_freq_dist[(word,)]
    return bigram_count / unigram_count
```

```
#Finding Next Word for the Given Word using MLE
```

```
def predict_next_word(word):
    next_word_probabilities = defaultdict(float)
    for next_word in set([gram[1] for gram in bigrams if gram[0] == word]):
        next_word_probabilities[next_word] = calculate_bigram_probability(word, next_word)
    return max(next_word_probabilities, key=next_word_probabilities.get)
```

```
given_word = "to"
next_word = predict_next_word(given_word)
print("Given word:", given_word)
print("Next word predicted using Bigram MLE:", next_word)
```

```
Given word: to
Next word predicted using Bigram MLE: the
```

```
given_word = "the"
next_word = predict_next_word(given_word)
print("Given word:", given_word)
print("Next word predicted using Bigram MLE:", next_word)
```

```
Given word: the
Next word predicted using Bigram MLE: king
```

```
given_word = "king"
next_word = predict_next_word(given_word)
print("Given word:", given_word)
print("Next word predicted using Bigram MLE:", next_word)
```

```
Given word: king
Next word predicted using Bigram MLE: and
```