

Terna Engineering College
Computer Engineering Department
Program: Sem VIII

Course: Natural Language Processing


Experiment No. 1

A.1 Aim: Perform Word analysis and word generation to study morphology using Virtual Lab.

PART B
(PART B: TO BE COMPLETED BY STUDENTS)

Roll No. 50	Name: AMEY MAHENDRA THAKUR
Class: BE COMPS B 50	Batch: B3
Date of Experiment: 19-01-2022	Date of Submission: 19-01-2022
Grade:	

B.1 Virtual Lab (Input & Output):


HOMEPARTNERSCONTACT

Computer Science and Engineering

Introduction
Objective
List of experiments
Target Audience
Course Alignment
Feedback

Natural Language Processing

Natural Language is the language written or spoken by humans in their daily life. Natural Language Processing is an interdisciplinary field dealing with human-computer interaction and computer aided processing of human language. It combines major concepts from computer science, artificial intelligence, and linguistics.

HOMEPARTNERSCONTACT

Computer Science and Engineering

Introduction
Objective
List of experiments
Target Audience
Course Alignment
Feedback

Natural Language Processing

1. Word Analysis
2. Word Generation
3. Morphology
4. N-Grams
5. N-Grams Smoothing
6. POS Tagging: Hidden Markov Model
7. POS Tagging: Viterbi Decoding
8. Building POS Tagger
9. Chunking
10. Building Chunker

Word Analysis

Select a Language which you know better

English ▼

Select a word from the below dropbox and do a morphological analysis on that word

beat ▼

Select the Correct morphological analysis for the above word using dropboxes (NOTE : na = not applicable)

WORD	beat	
ROOT	beat ▼	✓
CATEGORY	verb ▼	✓
GENDER	na ▼	✓
NUMBER	singular ▼	✓
PERSON	first ▼	✓
CASE	na ▼	✓
TENSE	simple-present ▼	✓
<div style="border: 1px solid black; padding: 2px; display: inline-block; margin-right: 20px;">Check</div> Right answer!!!		

Word Analysis

Select a Language which you know better


English ▼

Select a word from the below dropbox and do a morphological analysis on that word

play ▼

Select the Correct morphological analysis for the above word using dropboxes (NOTE : na = not applicable)

WORD	play	
ROOT	play ▼	✓
CATEGORY	noun ▼	✓
GENDER	na ▼	✓
NUMBER	singular ▼	✓
PERSON	na ▼	✓
CASE	na ▼	✓
TENSE	na ▼	✓
<div style="border: 1px solid black; padding: 2px; display: inline-block; margin-right: 20px;">Check</div> Right answer!!!		



Word Analysis

Select a Language which you know better


English

Select a word from the below dropdown and do a morphological analysis on that word

watch

Select the Correct morphological analysis for the above word using dropdowns (NOTE : na = not applicable)

WORD	watch	
ROOT	watch	✓
CATEGORY	noun	✓
GENDER	female	✓
NUMBER	plural	✓
PERSON	na	✓
CASE	na	✓
TENSE	na	✓
<input type="button" value="Check"/>	Right answer!!!	



Word Analysis

Select a Language which you know better

English

Select a word from the below dropdown and do a morphological analysis on that word

sofa

Select the Correct morphological analysis for the above word using dropdowns (NOTE : na = not applicable)

WORD	sofa	
ROOT	sofa	✓
CATEGORY	noun	✓
GENDER	na	✓
NUMBER	singular	✓
PERSON	na	✓
CASE	na	✓
TENSE	na	✓
<input type="button" value="Check"/>	Right answer!!!	

Morphology

Select a Root Word

घोड़ा

Fill the add delete table here:

Delete	Add	Number	Case	Correction
आ	आ	sing	dr	✓
आ	ए	plu	dr	✓
आ	ए	sing	ob	✓
आ	ओं	plu	ob	✓

For Example for लड़का:

Delete	Add	Number	Case
आ	आ	sing	dr
आ	ए	plu	dr
आ	ए	sing	ob
आ	ओं	plu	ob

Submit

Correct Answer!

Morphology

Select a Root Word

नदी

Fill the add delete table here:

Delete	Add	Number	Case	Correction
ई	ई	sing	dr	✓
ई	इयाँ	plu	dr	✓
ई	ई	sing	ob	✓
ई	इयों	plu	ob	✓

For Example for लड़का:

Delete	Add	Number	Case
आ	आ	sing	dr
आ	ए	plu	dr
आ	ए	sing	ob
आ	ओं	plu	ob

Submit

Correct Answer!

B.2 Observations & Learning:

- NLP (Natural Language Processing) is a branch of artificial intelligence that examines how computers interact with human languages, specifically how to design computers to process and evaluate huge volumes of natural language data. NLP is frequently used to categorise text data.
- Text classification is the challenge of categorising text data based on its content. The process of producing features for a machine learning model from raw text input is the most significant aspect of text categorization.

B.3 Conclusion:

We comprehend the ideas of Word Analysis and different sorts of Morphologies related to languages after successfully completing the task.

B.4 Questions of Curiosity:

Q1: Choose a typical masculine noun, ending in 'A', from your language. Write down its various forms along with various features and their values associated with them.

ANS:

WORD	Sofa
ROOT	Sofa
CATEGORY	Noun
GENDER	Masculine
NUMBER	Singular
PERSON	NA
CASE	NA
TENSE	Na

Q2. English has a suffix –en whose use is illustrated in the following lists:

List A	List B
red	redden
mad	madden
soft	soften
wide	widen
sharp	sharpen

Regarding these data, answer the following questions:

A. What part of speech does the suffix -en attach to? That is, what is the part of speech of the words in list A?

ANS:

- The suffix is attached to the root. In this list A, the suffix is attached to red, mad, soft, wide, sharp which are root words.

B. When the suffix -en is attached to a word, what part of speech is the resulting word? Give some specific morphological properties of one of the words in list B, in order to justify your answer.

ANS:

- When the suffix -en is attached to a word, the resulting word is a verb.

WORD	Soften
ROOT	Soft
CATEGORY	Verb
GENDER	Male
NUMBER	Singular
PERSON	Third
CASE	NA
TENSE	Past Perfect

Q3: Take one verb from your mother tongue, gloss it (i.e., give the English meaning) and conjugate it in all tenses and aspects and persons.

INDICATIVE			
Pronouns	PRESENT	PAST	FUTURE
I	read	read	will read
You	read	read	will read
He/She/It	reads	read	will read
We	read	read	will read
You	read	read	will read
They	read	read	will read

PERFECT TENSE			
Pronouns	PRESENT PERFECT	PAST PERFECT	FUTURE PERFECT
I	have read	had read	will have read
You	have read	had read	will have read
He/She/It	has read	had read	will have read
We	have read	had read	will have read
You	have read	had read	will have read
They	have read	had read	will have read

CONTINUOUS AND EMPHATIC TENSES				
Pronouns	PRESENT CONTINUOUS	PAST CONTINUOUS	PRESENT EMPHATIC	PAST EMPHATIC
I	am reading	was reading	do read	did read
You	are reading	were reading	do read	did read
He/She/It	is reading	was reading	does read	did read
We	are reading	were reading	do read	did Read
You	are reading	were reading	do read	did read
They	are reading	were reading	do read	did read

Q4: Refer to the following data and answer the question below:

List 1: taller, shorter, higher, lower, smarter
List 2: mower, teacher, sailor, caller, operator
List 3: never, cover, finger, river

Do the words ending with 'er'/'or' have some common features?

ANS:

1. List 1: The words ending with -er are comparative words that show a relationship between two entities
2. List 2: All these words are related to a particular profession or kind of work a person is doing. For example 1. Teacher- a person whose job is to teach; 2. Operator- a person whose job is to work a particular machine or piece of equipment
3. List 3: The words don't have a common feature except they are rhyming.

Q5: Identify root and suffix in the following words:

kissed
stronger
goodness
teacher

ANS:

WORD	ROOT	SUFFIX
kissed	kiss	ed
stronger	strong	er
goodness	good	ness
teacher	teach	er