M. Sc (Information Technology)		Semester – IV	
Course Name: Natural Language Processing Practical		**Course Code: PSIT4P2a**	
Periods per week (1 Period is 60 minutes)		4	
Credits		2	
		Hours	**Marks**
Evaluation System	**Practical Examination**	2	50
	Internal	--	-

List of Practical:	
1.	a. Install NLTK b. Convert the given text to speech c. Convert audio file Speech to Text.
2.	a. Study of various Corpus – Brown, Inaugural, Reuters, udhr with various methods like fields, raw, words, sents, categories, b. Create and use your own corpora(plaintext, categorical) c. Study Conditional frequency distributions Study of tagged corpora with methods like tagged_sents, tagged_words. d. Write a program to find the most frequent noun tags. e. Map Words to Properties Using Python Dictionaries f. Study DefaultTagger, Regular expression tagger, UnigramTagger g. Find different words from a given plain text without any space by comparing this text with a given corpus of words. Also find the score of words.
3.	a. Study of Wordnet Dictionary with methods as synsets, definitions, examples, antonyms. b. Study lemmas, hyponyms, hypernyms, entailments, c. Write a program using python to find synonym and antonym of word "active" using Wordnet d. Compare two nouns e. Handling stopword. Using nltk Adding or Removing Stop Words in NLTK's Default Stop Word List Using Gensim Adding and Removing Stop Words in Default Gensim Stop Words List Using Spacy Adding and Removing Stop Words in Default Spacy Stop Words List
4.	Text Tokenization a. Tokenization using Python's split() function b. Tokenization using Regular Expressions (RegEx) c. Tokenization using NLTK d. Tokenization using the spaCy library e. Tokenization using Keras f. Tokenization using Gensim
5.	Important NLP Libraries for Indian Languages and perform: a. word tokenization in Hindi b. Generate similar sentences from a given Hindi text input c. Identify the Indian language of a text

6.	Illustrate part of speech tagging. a. Part of speech Tagging and chunking of user defined text. b. Named Entity recognition of user defined text. c. Named Entity recognition with diagram using NLTK corpus – treebank
7.	a. Define grammer using nltk. Analyze a sentence using the same. b. Accept the input string with Regular expression of FA: 101+ c. Accept the input string with Regular expression of FA: (a+b)*bba d. Implementation of Deductive Chart Parsing using context free grammar and a given sentence.
8.	Study PorterStemmer, LancasterStemmer, RegexpStemmer, SnowballStemmer Study WordNetLemmatizer
9.	Implement Naive Bayes classifier
10.	Speech Tagging: a. Speech tagging using spacy b. Speech tagging using nktl Statistical parsing: a. Usage of Give and Gave in the Penn Treebank sample b. probabilistic parser Malt parsing: Parse a sentence and draw a tree using malt parsing.
11.	a. Multiword Expressions in NLP b. Normalized Web Distance and Word Similarity c. Word Sense Disambiguation