

INDEX

Sr. No.	Details	Page No	Date	Sign
1	<ul style="list-style-type: none"> A. Install NLTK Python 3.12.4 Installation on Windows. B. Convert the given text to speech. C. Convert audio file Speech to Text 	1-5		
2	<ul style="list-style-type: none"> A. Study of various Corpus – Brown, Inaugural, Reuters, udhr with various methods like filelds, raw, words, sents, categories. B. Create and use your own corpora (plaintext, categorical). C. Study Conditional frequency distributions. D. Study of tagged corpora with methods like tagged_sents, tagged_words. E. Write a program to find the most frequent noun tags. F. Map Words to Properties Using Python Dictionaries. G. Study DefaultTagger, Regular expression tagger, UnigramTagger. H. 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. 	6-17		
3	<ul style="list-style-type: none"> A. Study of Wordnet Dictionary with methods as synsets, definitions, examples, antonyms. B. Study lemmas, hyponyms, hypernyms. C. Write a program using python to find synonym and antonym of word "active" using Wordnet. D. Compare two nouns. E. Handling stopword. 	18-25		
4	<ul style="list-style-type: none"> 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 	26-28		
5	<ul style="list-style-type: none"> A. Word tokenization in Hindi. B. Generate similar sentences from a given Hindi text input. C. Identify the Indian language of a text. 	29-30		

6	<ul style="list-style-type: none"> A. Sentence tokenization, word tokenization, Part of speech Tagging and chunking of user defined text. B. Named Entity recognition using user defined text. C. Named Entity recognition with diagram using NLTK corpus – treebank. 	31-34		
7	<ul style="list-style-type: none"> A. Define grammar using nltk. Analyze a sentence using the same. B. Accept the input string with Regular expression of Finite Automaton: 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. 	35-38		
8	Study PorterStemmer, LancasterStemmer, RegexpStemmer, SnowballStemmer Study WordNetLemmatizer.	39-39		
9	Implement Naive Bayes classifier.	40-41		
10	<ul style="list-style-type: none"> A. Speech Tagging. B. Statistical parsing. C. Malt parsing. 	42-47		
11	<ul style="list-style-type: none"> A. Multiword Expressions in NLP. B. Normalized Web Distance and Word Similarity. C. Word Sense Disambiguation. 	48-50		