

- 1.** Write a program to implement Text Preprocessing using given text
- 2.** Write a program to demonstrate Lemmatization and Stemming of given text
- 3.** Write a program to create an inverted index for a given text file
- 4.** Write a program to implement token Normalization of a given Text
- 5.** Write a program to search for words and patterns in a given text file using inverted Index?
- 6.** Write a program to Count Word Frequency in a given text File
- 7.** Write a program to find the term frequency and document frequency of words that contained in text file stored at location path.
- 8.** Write a program to find the frequency of the word in each document in the corpus using Term Frequency.
- 9.** Write a program to measures the rank of the specific word for its relevancy within the text document using IDF.
- 10.** Write a program to implement text vectorization using TF-IDF Model
- 11.** Write a program to implement document ranking using victor space model
- 12.** Write a Program to classify whether the given SMS is Spam or Ham using Naïve Bayes Classifier.
- 13.** Write a Program to classify whether the given review of an item is positive or negative using Naïve Bayes Classifier.
- 14.** Write a program to create multiple documents that contain normal string and convert a collection of raw documents to a matrix of TF-IDF features.
- 15.** Write a program to convert multiple text documents into dusters based on type using K-Means Clustering (use TF-IDF approach for feature Extraction)
- 16.** Write a program to convert multiple text documents into cluster documents by topics using K-Means Clustering (use bag-of-words approach for feature Extraction)
- 17.** Write a program to convert multiple text documents into clusters based on type using K-Nearest Neighbor Clustering (use TF-IDF approach for feature Extraction)
- 18.** Write a program to convert multiple text documents into cluster documents by topics using K- Nearest Neighbor Clustering (use bag-of-words approach for feature Extraction)

- 19.** Write a program to convert long list of strings (words) in a document into clusters using K-Means Clustering.
- 20.** Write a program to convert long list of strings (words) in a document into clusters using K- Nearest Neighbor Clustering.
- 21.** Write a program to create multiple documents that contain normal string and convert a collection of raw documents to a Document term matrix
- 22.** Write a program to implement web crawling using Breadth First Search?
- 23.** Write a program to implement web crawling using Depth First Search?
- 24.** Write a program to create a web crawler to crawl multiple web pages of a given URL.
- 25.** Write a program to crawl a given web page and get most frequent words.
- 26.** Write a program to crawl a given web page and scrape the complete content of given URL.
- 27.** Write a program to implement The Power Method for computing Page Rank.
- 28.** Write a program to Implement Page Rank using Random Walk method.
- 29.** Write a program to display top five ranking websites of a given keyword using Page Rank Algorithm
- 30.** Write a program to retrieve pages satisfying the (Boolean) query in the web search.