

- (d) Crawling web pages
- 1-d. The term used for the process of automatically categorizing web documents into predefined categories based on their content is__(CO2) 1
- (a) Sentiment analysis
 - (b) Document clustering
 - (c) Topic modeling
 - (d) Content classification
- 1-e. The step that involves the use of algorithms to uncover patterns and insights from social media data is_(CO3) 1
- (a) Data Collection
 - (b) Data Preprocessing
 - (c) Data Analysis
 - (d) Data Visualization
- 1-f. Type of social graph that is used to represent relationships between different types of entities is__(CO3) 1
- (a) Bipartite graph
 - (b) Directed graph
 - (c) Weighted graph
 - (d) Undirected graph
- 1-g. The purpose of N-gram frequency count in text processing is(CO4) 1
- (a) To identify the most frequently occurring words in a text
 - (b) To identify the grammatical structure of a sentence
 - (c) To identify the sentiment of a text
 - (d) To identify the frequency of occurrence of sequences of N words in a text
- 1-h. To normalize keywords in NLP, which technique do we follow(CO4) 1
- (a) Lemmatization
 - (b) Parts of speech
 - (c) TF-IDF
 - (d) N-Gram
- 1-i. Trend analysis is significant for(CO5) 1
- (a) forecasting and budgeting
 - (b) profit planning
 - (c) capital rationing

(d) working capital management

- 1-j. Upward trend is also known as(CO5) 1
- (a) Bull market
 - (b) Bear market
 - (c) Range bound
 - (d) None of the above

2. Attempt all parts:-

- 2.a. Define Speech Analytics.(CO1) 2
- 2.b. Define Web Graph.(CO2) 2
- 2.c. Define Recommendation in social media.(CO3) 2
- 2.d. Describe semantic topic tagging.(CO4) 2
- 2.e. List three types of trend analysis.(CO5) 2

SECTION B

30

3. Answer any five of the following:-

- 3-a. Explain Named Entity Recognition (NER).(CO1) 6
- 3-b. Explain Sentiment Classification.(CO1) 6
- 3-c. Explain Boolean queries.(CO2) 6
- 3-d. Explain Web Scraping.(CO2) 6
- 3.e. Explain key concepts and applications of Homophily .(CO3) 6
- 3.f. Elaborate on a drawback of Extractive Summarization models.(CO4) 6
- 3.g. Differentiate between web analytics and social media analytics.(CO5) 6

SECTION C

50

4. Answer any one of the following:-

- 4-a. Elaborate Text Clustering with examples.(CO1) 10
- 4-b. Describe how N-grams represent sequences of words or characters within a text corpus and discuss the significance of different N-gram sizes in capturing linguistic patterns and dependencies.(CO1) 10

5. Answer any one of the following:-

- 5-a. Describe Web Mining , its types and Applications in detail.(CO2) 10
- 5-b. Elaborate with examples Boolean query with logical operators such as AND, OR, and NOT that specify the criteria for retrieving documents.(CO2) 10

6. Answer any one of the following:-

- 6-a. Elaborate all the steps involved in the process of Social media mining.(CO3) 10
- 6-b. Describe Key Concepts and Applications of Influence and Homophily in detail.(CO3) 10

7. Answer any one of the following:-

- 7-a. Describe the difference between Phrase Mining and N-gram frequency count.(CO4) 10
- 7-b. Explain Machine-Learned Classification, and how is it different from other machine learning tasks, such as regression and clustering.(CO4) 10

8. Answer any one of the following:-

- 8-a. Explain the uses of trend analysis in surveys. Explain few advantages of analyzing trends in survey data.(CO5) 10
- 8-b. Explain ten latest web development trends.(CO5) 10

REG. MAY 2024