.

Printed Page:-4 **Subject Code:- ACSBS0715**

Roll. No:

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA
(An Autonomous Institute Affiliated to AKTU, Lucknow)
B.Tech
SEM: VII - THEORY EXAMINATION (2025 - 2026)
Subject: Advance Social,Text and Media Analytics

Time: 3 Hours **Max. Marks: 100**

General Instructions:

IMP: *Verify that you have received the question paper with the correct course, code, branch etc.*
1. *This Question paper comprises of* ***three Sections -A, B, & C.*** *It consists of Multiple Choice Questions (MCQ's) & Subjective type questions.*
2. *Maximum marks for each question are indicated on right -hand side of each question.*
3. *Illustrate your answers with neat sketches wherever necessary.*
4. *Assume suitable data if necessary.*
5. *Preferably, write the answers in sequential order.*
6. *No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked.*

SECTION-A 20

1. Attempt all parts:-

1-a. State, _____of the following is not a core text mining operation? [CO1,K1] 1

(a) Information Retrieval
(b) Text Categorization
(c) Sentiment Analysis
(d) Data Encryption

1-b. State, _____ is text mining? [CO1,K1] 1

(a) Extracting minerals from text documents
(b) Extracting valuable information from unstructured text data
(c) Mining text messages for hidden meanings
(d) Mining textual patterns in multimedia files

1-c. State which algorithm is commonly used for clustering tasks? [CO2,K1] 1

(a) Decision Trees
(b) K-means
(c) Linear Regression
(d) Naive Bayes

1-d. State, what is the objective of simple predictive modeling? [CO2,K1] 1

(a) Classifying text documents
(b) Predicting numerical values based on input features
(c) Summarizing large datasets
(d) Generating random text

1-e. State, what is a "filter bubble" in social media analytics? [CO3,K1] 1

(a) The tendency of users to only consume information that confirms their existing beliefs or opinions

(b) The tendency of users to ignore information from their friends and family

(c) The tendency of users to only follow users who are similar to themselves

(d) The tendency of users to only post positive content on social media

1-f. State, what is a homophilous group in social network analysis? [CO3,K1] 1

(a) A group of people with similar characteristics who are more likely to be connected to each other

(b) A group of people with different characteristics who are more likely to be connected to each other

(c) A group of people who have no connections to each other

(d) A group of people who only connect with celebrities on social media

1-g. State what is a challenge in social media analytics related to data volume? [CO4,K1] 1

(a) Too little data to analyze

(b) Data volume is not a challenge in social media analytics

(c) Too much data to process and analyze effectively

(d) Data volume does not affect social media analytics

1-h. State what is a limitation of social media analytics in understanding user intent? [CO4,K1] 1

(a) Social media posts do not reflect user intent

(b) Users do not express their intent on social media

(c) Social media posts are often ambiguous, making it difficult to accurately determine user intent

(d) User intent is always clear in social media posts

1-i. State, what is the fundamental focus of social network analysis (SNA) in advanced analytics? [CO5,K1] 1

(a) 4

(b) Image recognition

(c) Analyzing emotional tone

(d) Predicting future connections

1-j. State, what are centrality measures used for in social network analysis? [CO5,K1] 1

(a) 2

(b) Analyzing emotional tone

(c) Identifying influential nodes in a network

(d) Predicting future connections

2. Attempt all parts:-

2.a. Describe the use of simple predictive models like linear regression in text analysis? [CO1,K2] 2

2.b. Explain the concept of document similarity in text mining? [CO2,K2] 2

2.c. Discuss and give a brief idea of the concept of web search? [CO3,K2] 2

2.d.	Describe the concept of link prediction in social network analysis? [CO4,K2]	2
2.e.	Discuss, How can link analysis be utilized to uncover relationships between different types of media content, such as images, videos, and articles? [CO5,K2]	2

SECTION-B 30

3. Attempt all parts:-

3.a. Answer any one of the following:-

3.a.(i)	Explain keyword "search over graph data"? [CO1,K2]	6
3.a.(ii)	Compare different clustering algorithms used in text mining, such as K-means and hierarchical clustering? [CO1,K4]	6

3.b. Answer any one of the following:-

3.b.(i)	Discuss what is sentiment prediction, with a relevant example? [CO2,K2]	6
3.b.(ii)	Explain with example as how are machine learning algorithms used in sentiment prediction? [CO2,K2]	6

3.c. Answer any one of the following:-

3.c.(i)	Discuss the indexing process that enables search engines to organize online content? [CO3,K2]	6
3.c.(ii)	Discuss and present a detailed explanation of ranking algorithms used for ordering search results? [CO3,K2]	6

3.d. Answer any one of the following:-

3.d.(i)	Explain how degree centrality, closeness centrality, and betweenness centrality are calculated and their respective interpretations in the context of social media networks? [CO4,K2]	6
3.d.(ii)	Discuss and provide examples of how homophily affects social connections and interactions on social media platforms. Discuss its implications for targeted marketing strategies? [CO4,K2]	6

3.e. Answer any one of the following:-

3.e.(i)	Discuss the challenges and best practices in visualizing dynamic social media data, such as real-time trends and user interactions? [CO5,K2]	6
3.e.(ii)	Explain Random Graph with the help of example? [CO5,K2]	6

SECTION-C 50

4. Answer any one of the following:-

4-a.	Explain the concept of clustering and its importance in organizing large datasets. Discuss common clustering algorithms such as k-means and hierarchical clustering. Provide a detailed explanation of topic modeling techniques, focusing on Latent Dirichlet Allocation (LDA). Discuss how LDA is used to detect topics in a collection of documents. Provide examples and discuss the challenges associated with topic detection in real-world applications? [CO1,K2]	10
4-b.	Describe the process of simple predictive modeling in text analysis. Explain how techniques like linear regression and logistic regression are applied to predict outcomes based on textual data. Discuss the importance of feature selection and feature engineering in building predictive models. Provide real-world examples of how simple predictive models are used in text analysis tasks such as predicting customer behavior or product reviews? [CO1,K2]	10

.

5. Answer any one of the following:-

5-a.	Explain Latent Dirichlet Allocation (LDA) as a probabilistic model for topic detection. How does it work, and what are its limitations? [CO2,K2]	10
5-b.	Discuss the ethical implications of sentiment analysis applications in domains like customer feedback analysis and political sentiment tracking. How can organizations ensure responsible and unbiased use of sentiment analysis technologies? [CO2,K2]	10

6. Answer any one of the following:-

6-a.	Discuss and analyze the role of user segmentation in web analytics and its impact on targeted marketing strategies and website personalization? [CO3,K2]	10
6-b.	Discuss the challenges faced in accurately measuring the success and effectiveness of SEO strategies in improving website visibility and ranking? [CO3,K2]	10

7. Answer any one of the following:-

7-a.	Discuss the concept of feature engineering in predictive modeling. How do domain-specific features enhance the accuracy and robustness of predictive models? Provide examples of feature engineering techniques tailored to text-based datasets? [CO4,K2]	10
7-b.	Explore the concept of overfitting in predictive modeling. What are the causes of overfitting, and how do researchers and data scientists employ regularization techniques to mitigate this issue? Provide case studies where regularization played a crucial role? [CO4,K2]	10

8. Answer any one of the following:-

8-a.	Discuss how does cross-cultural analysis contribute to the understanding of social network structures? [CO5,K2]	10
8-b.	In social media, Explain the role of natural language processing (NLP) in extracting information about identity from posts? [CO5,K2]	10