

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA
(An Autonomous Institute Affiliated to AKTU, Lucknow)

MBA (Integrated)

SEM: VIII - THEORY EXAMINATION (2025 - 2026)

Subject: Predictive Analytics

Time: 2.5 Hours

Max. Marks: 60

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

15

1. Attempt all parts:-

- | | | |
|-------|--|---|
| (1.1) | Unsupervised learning helps in discovering what type of structures? (CO1,K2) | 1 |
| | (a) Labeled outputs | |
| | (b) User intentions | |
| | (c) Fixed boundaries | |
| | (d) Hidden patterns | |
| (1.2) | Ordinal data is understood as data arranged (CO2,K3) | 1 |
| | (a) Colors | |
| | (b) Labels | |
| | (c) Order | |
| | (d) Random | |
| (1.3) | Data cleaning is : (CO3,K3) | 1 |
| | (a) A one-time task | |
| | (b) A rarely needed task | |
| | (c) An optional step | |
| | (d) An ongoing process | |
| (1.4) | The validation set assists model tuning by (CO4,K6) | 1 |
| | (a) helping adjust hyperparameters | |
| | (b) storing final predictions | |
| | (c) replacing the test set | |
| | (d) reducing dataset size | |
| (1.5) | Example of time series (CO5,K4) | 1 |

- (a) Stock Prices at daily intervals
- (b) Employee names
- (c) Survey data
- (d) Customer lists

2. Attempt all parts:-

- | | | |
|-------|---|---|
| (2.1) | Explain the purpose of data preprocessing in analytics workflows. (CO1,K2) | 2 |
| (2.2) | Explain nominal data with one example. (CO2,K3) | 2 |
| (2.3) | Describe the role of data integration in the preprocessing workflow. (CO3,K3) | 2 |
| (2.4) | Describe the purpose of data partitioning in predictive analytics. (CO4,K6) | 2 |
| (2.5) | Describe the nature of time series data and highlight its chronological structure. (CO5,K4) | 2 |

SECTION-B

15

3. Answer any three of the following:-

- | | | |
|-------|---|---|
| (3.1) | Discuss how predictive analytics enhances operational efficiency in organizations. (CO1,K2) | 5 |
| (3.2) | Explain bias and variance with simple short examples. (CO2,K3) | 5 |
| (3.3) | Explain how different data integration methods improve dataset quality. (CO3,K3) | 5 |
| (3.4) | Critique the limitations of Decision Trees, focusing on overfitting and instability issues. (CO4,K6) | 5 |
| (3.5) | Evaluate the usefulness of additive and multiplicative models in different business scenarios. (CO5,K4) | 5 |

SECTION-C

30

4. Answer any one of the following:-

- | | | |
|-------|--|---|
| (4.1) | A clinic wants to understand the common reasons patients miss appointments. Prepare a data analysis method and discuss how the results can guide better scheduling. (CO1,K2) | 6 |
| (4.2) | A new online shop wants to study customer buying behavior. Build a basic predictive analytics flow and explain how each step helps the business. (CO1,K2) | 6 |

5. Answer any one of the following:-

- | | | |
|-------|--|---|
| (5.1) | A school stores student grades (A/B/C/D) and satisfaction ratings (1–10). Evaluate how these variables should be treated during data preparation. (CO2,K3) | 6 |
| (5.2) | A bank uses a variable “Loan_Flag = Yes/No”. Assess how flag and Boolean variables help in segmentation. (CO2,K3) | 6 |

6. Answer any one of the following:-

- | | | |
|-------|--|---|
| (6.1) | A logistics firm depends on live vehicle tracking. Explain how real-time integration supports its operations. (CO3,K3) | 6 |
| (6.2) | A telecom company maintains millions of customer entries. Discuss the challenges faced during cleaning of extremely large datasets. (CO3,K3) | 6 |

7. Answer any one of the following:-

- | | | |
|-------|--|---|
| (7.1) | A dataset contains inconsistent values and different date ranges. The team plans to apply custom 70:10:20 splitting before modeling. Propose a step-by-step approach | 6 |
|-------|--|---|

to execute this custom split and explain why consistency is important. (CO4,K6)

- (7.2) A Decision Tree built for employee churn prediction becomes extremely deep and fits almost every detail of the training data. Develop a strategy to handle this situation and explain how overfitting can be reduced. (CO4,K6) 6

8. Answer any one of the following:-

- (8.1) A retail chain tracks monthly footfall for eight years and notices sharp seasonal spikes. Develop a forecasting plan using additive or multiplicative models and justify your choice. (CO5,K4) 6
- (8.2) A financial analyst needs to decide whether AR or MA is best for forecasting stock volatility. Develop a comparison based on error dependence and lag structure. (CO5,K4) 6

REG_JAN_JUNE_2026