

Maharshi Karve Stree Shikshan Samstha's

Smt. Hiraben Nanavati Institute of Management & Research for Women

INTERNAL EXAMINATION – Sem IV – (Jan- July 2026)

SUBJECT – Predictive Analytics and Machine learning using Python

SUBJECT CODE – 410BA

Date : 14-05-2026

Pattern : 2024

Duration : 150 min

Max Marks : 50

Instructions for students :

- *Marks are indicated for each question.*
 - *Handwriting should be eligible for evaluation.*
 - *Marks will be given for quality, not quantity.*
-

Q.1] Solve any five . (10 Marks)

-) Define Predictive Analytics and mention any two business applications.
- b) Differentiate between Regression and Classification models.
- c) What is Exploratory Data Analysis (EDA)?
- d) Define Label Encoding and One-Hot Encoding.
- e) What is feature scaling? Name any two feature scaling techniques.
- f) Define Precision and Recall in classification models.
- g) What is Principal Component Analysis (PCA)?
- h) Define Responsible AI and mention one ethical concern in Predictive Analytics.

Q.2] Attempt any Two – 5 Marks Each

A] A retail company wants to predict customer purchasing behavior using Predictive Analytics.

Explain:

- Scope and applications of Predictive Analytics
 - Types of predictive models
 - Role of Python libraries such as NumPy, Pandas, Matplotlib, and Seaborn
- Illustrate your answer with a suitable business example.

B] An online learning platform stores the following student information:

- Student Name
- Course Category
- Course Completion Status

- Number of Assignments Submitted
- Student Rating

Identify the most suitable data type for each field and explain how missing values and inconsistent data can be handled during preprocessing.

C] A company wants to prepare a dataset for machine learning analysis.

Explain:

- Data cleaning techniques
- Standardization and normalization
- Correlation-based feature selection
- Importance of feature engineering in predictive analytics

Q.3] Attempt any One – 10 Marks

A] Describe the following supervised machine learning techniques with suitable examples:

- Linear Regression
- Logistic Regression
- Decision Tree
- Random Forest

Also explain the following evaluation metrics:

- RMSE
- Accuracy
- Precision
- Recall
- F1-Score

B] A bank wants to predict whether a customer will default on a loan using machine learning techniques.

The dataset contains:

- Customer Income
- Credit Score
- Loan Amount
- Employment Type
- Previous Loan Status

Answer the following:

- Explain how the dataset should be preprocessed before training the model.
- Describe the Train-Test Split method.
- Explain Cross-Validation and its importance.
- Compare Logistic Regression and SVM for this problem.
- Explain how Grid Search CV can improve model performance

Q.4] Attempt any One – 10 Marks

A] Explain the concept of Unsupervised Learning.

Describe the following clustering techniques with suitable examples:

- K-Means Clustering
- Hierarchical Clustering
- DBSCAN

Also explain:

- Elbow Method
- Principal Component Analysis (PCA)
- Applications of clustering in customer segmentation

B] An e-commerce company wants to group customers based on purchasing patterns for marketing analysis.

The management wants to:

- Identify high-value customers
- Detect unusual buying behavior
- Improve personalized marketing

Answer the following:

- i) Explain how clustering techniques can help in customer segmentation.
- ii) Describe the working of K-Means clustering.
- iii) Differentiate between Hierarchical Clustering and DBSCAN.
- iv) Explain the role of PCA in dimensionality reduction.
- v) Discuss the business applications of unsupervised learning.

Q.5] Attempt any One – 10 Marks

(10 Marks)

A] A healthcare organization wants to develop a machine learning model for predicting patient disease risk.

Answer the following:

- i) Explain the steps involved in designing a machine learning workflow.
- ii) Describe the importance of Exploratory Data Analysis (EDA).
- iii) Explain how machine learning models can be deployed using Pickle.
- iv) Discuss ethical concerns and bias in machine learning systems.
- v) Explain the importance of privacy and responsible AI in predictive analytics.

B] As a Data Analyst in a multinational company, explain how Predictive Analytics and Machine Learning can be applied in the following domains:

- Banking and Financial Forecasting
- Customer Churn Prediction

- Healthcare Analytics
- Sales Forecasting
- Recommendation Systems

Also explain:

- Importance of hyperparameter tuning
- Role of ROC-AUC in model evaluation
- Ethical implications of AI-driven decision systems.
