IN-SEMESTER EXAMINATION-I

BE (Semester-VII)

SUBJECT – Machine Learning for IT Application Development

Branch: IT
Div.: ALL
Duration: 60 Minutes
Instructions –

Date: 13/07/2026
Timing: 12.00 to 01:00PM
Maximum Marks: 20

1. All questions are compulsory.
2. Assume suitable data wherever necessary and state the assumptions made.
3. Diagrams / sketches should be given wherever necessary.
4. Use of logarithmic table, drawing instruments and non-programmable calculators is permitted.
5. Figures to the right indicate full marks.

Q.1		Answer any 5 of the following questions	Marks	Course Outcomes	Learning Levels
	a.	Relate supervised learning to labelled datasets with a suitable example.	2	CO1	AN
	b.	Differentiate between Artificial Intelligence, Machine Learning and Deep Learning with example.	2	CO1	AN
	c.	Illustrate the use of training data and test data in building a Machine Learning model with a suitable example.	2	CO1	A
	d.	Interpret the roles of data validation and data imputation in data preprocessing.	2	CO2	AN
	e.	Apply Principal Component Analysis(PCA) to illustrate its purpose in Machine Learning.	2	CO2	A
	f.	Contrast regression and classification techniques used in Machine Learning.	2	CO3	AN
	g.	Summarize dependent and independent variables in regression.	2	CO3	AN
Q.2	a.	Examine One-Hot Encoding and Label Encoding, and determine their suitability for different datasets.	5	CO2	AN
		OR			
	b.	Evaluate the importance of data preprocessing in Machine Learning model design and justify its impact on model performance.	5	CO2	E

<table>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Q.3</td><td>a.</td><td>Assess the BIRCH clustering algorithm with respect to scalability, efficiency, and clustering quality.</td><td>5</td><td>CO3</td><td>E</td></tr>
<tr><td></td><td></td><td colspan="4">OR</td></tr>
<tr><td></td><td>b.</td><td>The below table shows the marks obtain by student in midterm and final year exam.
1) Construct equation of predication (linear Regression formula)
2) Predict final year marks if the midterm marks is 40?
<table><tr><th>Midterm marks (X)</th><th>Final exam marks (Y)</th></tr><tr><td>45</td><td>60</td></tr><tr><td>70</td><td>70</td></tr><tr><td>60</td><td>54</td></tr><tr><td>84</td><td>82</td></tr><tr><td>75</td><td>68</td></tr><tr><td>84</td><td>76</td></tr></table></td><td>5</td><td>CO3</td><td>A</td></tr>
</table>

IN-SEMESTER EXAMINATION-II

BE (Semester-VII)

SUBJECT – Machine Learning for IT Application Development

Branch: IT
Div.: ALL
Duration: 60 Minutes
Date: 01/09/2026
Timing: 2:00 pm to 3:00 pm
Maximum Marks: 20

Instructions –

1. All questions are compulsory.
2. Assume suitable data wherever necessary and state the assumptions made.
3. Diagrams / sketches should be given wherever necessary.
4. Use of logarithmic table, drawing instruments and non-programmable calculators is permitted.
5. Figures to the right indicate full marks.

Q.1		Answer any 5 of the following questions	Marks	Course Outcomes	Learning Levels	PI
	a.	Interpret the fundamental principle of the k-Nearest Neighbours (KNN) algorithm.	2	CO4	AN	1.6.1
	b.	Relate Bagging to ensemble methods in Machine Learning.	2	CO4	AN	1.6.1
	c.	Interpret the role of a hyperplane in Support Vector Machine (SVM).	2	CO4	AN	1.6.1
	d.	Investigate the challenges associated with applying optimization techniques to large-scale machine learning models.	2	CO5	AN	4.4.2
	e.	Analyze the role of the validation set in hyperparameter tuning.	2	CO5	AN	3.5.1
	f.	Examine how semi-supervised learning utilizes both labeled and unlabeled data during training.	2	CO6	AN	4.4.2
	g.	Outline the key elements of reinforcement learning.	2	CO6	AN	1.6.2
Q.2	a.	Differentiate between local and global optimization strategies and examine their impact on the performance of machine learning algorithms.	5	CO5	AN	3.5.1
		OR				
	b.	Investigate how cross-validation and regularization can be employed to address overfitting and underfitting.	5	CO5	AN	2.5.2
Q.3	a.	Analyze the benefits of using semi-supervised learning techniques, especially in scenarios with limited labeled data. How does leveraging unlabeled data enhance the accuracy and efficiency of machine learning models?	5	CO6	AN	2.6.4
		OR				
	b.	Investigate the challenges associated with semi-supervised learning in Machine Learning.	5	CO6	AN	2.5.2

END SEMESTER EXAMINATION, Oct 2026

B.E. SEMESTER VII (CBCGS-HME 2023)

12 OCT 2026

Branch:	IT	**Q.P. Code:**	E7C4061-1
Subject:	Machine Learning for IT Application Development	**Duration:**	2 hours
Subject Code:	PCC-IT 701	**Max. Marks:**	60

Instructions:
1. All sections are compulsory
2. Figures to the right indicate full marks.
3. Assume suitable data if necessary and state the assumptions clearly.

Section-I Short Answer Questions (Answer any 05 questions out of 06) (10 Marks)
(Fundamental, Core Types)

Q. No.		Marks	CO	RBT Level	PI
1	Investigate the application of Unsupervised Learning using a suitable example.	2	CO1	AN	2.5.2
2	Examine the role of data validation and data imputation in improving data quality.	2	CO2	AN	2.5.2
3	Evaluate polynomial regression. How does it handle non-linear relationships?	2	CO3	AN	2.5.2
4	Interpret the role of a hyperplane in Support Vector Machine (SVM).	2	CO4	AN	1.6.1
5	Examine the role of optimization in training machine learning models.	2	CO5	AN	1.6.1
6	Analyze the potential risks or challenges of using semi-supervised learning in situations with limited labeled data.	2	CO6	AN	2.5.2

Section-II Descriptive Answer Questions (Answer any 04 out of 06) (20 Marks)
(Descriptive, Comprehension Types)

1	Examine the significance of hypothesis testing in Machine Learning.	5	CO1	AN	2.5.2
2	Analyze the process of handling outliers in a dataset and its importance in data preprocessing.	5	CO2	AN	4.5.6
3	The following table shows the BMI and Cholesterol levels of 10 patients. **BMI / Cholesterol**: 17/140, 21/189, 24/210, 28/240, 14/130, 16/100, 19/135, 22/166, 15/130, 18/170 Construct the equation of prediction using the linear regression formula. Predict the likely cholesterol level for a patient whose BMI is 27.	5	CO3	C	5.4.2
4	Differentiate Bagging and Boosting based on how their base models are combined.	5	CO4	AN	3.5.1

5	Analyze the role of the validation set in hyperparameter tuning.	5	CO5	AN	3.5.1
6	Illustrate a real-world problem where semi-supervised learning would be beneficial.	5	CO6	A	3.5.1
Section-III	**Long Answer Question (Answer any 03 out of 05) (Application, Analytical, Evaluation, Design Type)**	**(30 Marks)**			
1	Consider the following table and decide how chi-square test is useful to know if gender has anything to do with political party preference. Note: for degree of freedom 2 and alpha level of 0.05, tabulated value is 5.991 **Republican Democrat Independent Total** **Male** 100 70 30 200 **Female** 140 60 20 220 **Total** 240 130 50 420	10	CO2	C	5.4.2
2	Evaluate the factors to be considered when selecting appropriate evaluation metrics for a machine learning task.	10	CO3	E	3.5.1
3	Construct a decision tree for the given dataset by applying the Gini Index to select the appropriate splitting attributes.	10	CO4	A	3.5.1
4	Analyze the working of k-fold cross-validation and examine its significance in assessing machine learning model performance.	10	CO5	AN	3.5.1
5	Assess the benefits of semi-supervised learning when labeled data is limited and how unlabeled data contributes to model accuracy and efficiency.	10	CO6	E	4.4.2

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No