(Q7)

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST–2 EXAMINATION – 2026

B.Tech–IV Semester (ECE)

COURSE CODE (CREDITS): 25B1WEC431 (3) **MAX. MARKS:** 25

COURSENAME: Foundation of Machine Learning and Python

COURSE INSTRUCTORS: Dr. Vikas Baghel **MAX. TIME:** 1 Hour 30 Min

Note: (a) All questions are compulsory.

(b) The candidate is allowed to make suitable numeric assumptions wherever required for solving problems.

(c) Use of calculators is allowed.

Q.No	Question	CO	Marks
Q1	Logistic Regression and Regularization: For a binary classifier, $$p(y=1 \mid x)=\sigma(z), \quad z=\beta_0+\beta_1 x_1+\beta_2 x_2, \quad \sigma(z)=\frac{1}{1+e^{-z}},$$ with parameters $\beta_0=-3$, $\beta_1=0.8$, $\beta_2=1.2$ and sample $(x_1,x_2)=(2,1)$: (i) compute z and $p(y=1 \mid x)$, (ii) classify using threshold 0.5, and (iii) write Ridge and Lasso objectives and explain in one line how ℓ_2 and ℓ_1 penalties affect coefficients.	CO1	[5]
Q2	(a) KNN classification: Given training points A$(1,1) \rightarrow 0$, B$(2,1) \rightarrow 0$, C$(2,3) \rightarrow 1$, D$(3,3) \rightarrow 1$, E$(3,1) \rightarrow 1$, and query point $q=(2,2)$, compute Euclidean distances from q to all training points and predict the class using $k=3$.	CO2	[3]
	(b) Linear SVM: For decision function $f(x)=w^T x+b$ with $w=[2,-1]^T$ and $b=-1$, (i) find class labels for points $(2,1)$, $(1,3)$, and $(3,4)$ using $\text{sign}(f(x))$, and (ii) compute the margin width $m=\frac{2}{\|w\|}$.		[3]

Q.No	Question	CO	Marks
Q3	(a) Decision Tree split selection: At a parent node, class counts are [5 Yes, 5 No]. Attribute A creates two children: $A = 0$: [4 Yes, 2 No], $A = 1$: [1 Yes, 3 No]. Using entropy, compute (i) parent entropy, (ii) weighted child entropy, and (iii) information gain of split on A.	CO3	[4]
	(b) Naive Bayes classification: Assume $P(\text{Spam}) = 0.4$, $P(\text{Ham}) = 0.6$, $P(\text{free} \mid \text{Spam}) = 0.7$, $P(\text{offer} \mid \text{Spam}) = 0.6$, $P(\text{free} \mid \text{Ham}) = 0.1$, $P(\text{offer} \mid \text{Ham}) = 0.2$. For a message containing words {free, offer}, compute unnormalized posteriors and predict the class.		[4]
Q4	K-means clustering (numerical): Consider points $(1, 1)$, $(1.5, 2)$, $(3, 4)$, $(5, 7)$, $(3.5, 5)$, $(4.5, 5)$, $(3.5, 4.5)$ with $K = 2$ and initial centroids $C_1 = (1, 1)$, $C_2 = (5, 7)$. (i) Assign each point to the nearest centroid using Euclidean distance. (ii) Compute updated centroids after one iteration. (iii) Run one more assignment step using updated centroids and report final cluster membership. (iv) State one advantage and one limitation of K-means.	CO4	[6]