

COURSE CODE (CREDITS): 18B11BI611 (3)

MAX. MARKS: 15

COURSE NAME: Machine Learning for Bioinformatics

COURSE INSTRUCTORS: D. Gupta

MAX. TIME: 1 Hour

Note: (a) All questions are compulsory. (b) The candidate is allowed to make suitable numeric assumptions wherever required for solving problems. (c) Use of calculator is allowed. (d) Be concise.

Q. No.	Question	CO	Marks																
Q1(a)	What are the main challenges encountered in developing machine learning applications? Discuss the key difficulties and their potential solutions.	1	2																
(b)	What are the fundamental differences between supervised, unsupervised, and reinforcement learning in terms of their objectives, data requirements, and real-world applications?	1	3																
Q2 (a)	<p>A machine learning model predicts animals, Cat, Dog and Rabbit on the basis of their behavior. Using the following confusion matrix, compute Recall, Accuracy and Precision for Cat class:</p> <p style="text-align: center;">Predicted</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td></td><td>Cat</td><td>Dog</td><td>Rabbit</td></tr> <tr> <td>Actual Cat</td><td>10</td><td>2</td><td>1</td></tr> <tr> <td>Dog</td><td>1</td><td>8</td><td>1</td></tr> <tr> <td>Rabbit</td><td>2</td><td>1</td><td>9</td></tr> </table>		Cat	Dog	Rabbit	Actual Cat	10	2	1	Dog	1	8	1	Rabbit	2	1	9	3	3
	Cat	Dog	Rabbit																
Actual Cat	10	2	1																
Dog	1	8	1																
Rabbit	2	1	9																
(b)	What is ROC AUC (receiver operating characteristic area under curve) and why is it considered a useful metric for evaluating classification models, especially for imbalanced datasets?	3	2																
Q3 (a)	In linear regression, what is the total sum of residuals for a given dataset? Additionally, derive the formulas for the regression coefficients β_0 (intercept) and β_1 (slope) by minimizing the sum of squared residuals.	2	3																
(b)	What is gradient descent? Outline the step-by-step process for finding the local minimum of a differentiable function in the case of a regression line.	2	2																