IPRC MUSANZE
Integrated Polytechnic Regional College
P.O Box 226 Musanze-Rwanda
Tel. +250 785 189 494
Email: info@iprcmusanze.rp.ac.rw
www.iprcmusanze.rp.ac.rw

MODULE CODE AND TITLE: MACHINE LEARNING

DEPARTMENT: ICT

Code: ITLML801

05/01/2024

FORMATIVE ASSESSMENT / 20

Learning Outcome II: Class L5 I8 Reg nO 23RP00815

1. What is the main goal of using a Random Forest algorithm? / 1 Mark

A. To classify data
B. To reduce variance
C. To reduce bias
D. To predict outcomes

2. What is the key difference between a classification and regression problem?/2marks

a. Classification predicts discrete values while regression predicts continuous values.
b. Classification predicts continuous values while regression predicts discrete values.
c. Classification is supervised while regression is unsupervised.
d. Classification is unsupervised while regression is supervised.

3. What is the main difference between a Decision Tree and a Random Forest? / 2 marks

A. Random Forest is more accurate
B. Decision Tree is more accurate
C. Random Forest creates multiple Decision Trees
D. Decision Tree creates multiple Random Forest

4. What is logistic regression used for in classification?/1 mark

a. To classify data points into discrete classes.
b. To identify the most important features for a given classification problem.
c. To predict the probability of a given data point belonging to a particular class.
d. To identify relationships between different classes.

5. What is the benefit of using a Random Forest algorithm over a single Decision Tree? / 1 mark

a. It reduces variance
b. It reduces bias
c. It is more accurate
d. It is faster

6. What type of data is best suited for Decision Tree algorithms? / 1 mark

(A.) Categorical data

B. Continuous data

C. Binary data

D. All of the above

7. Which of the following is an advantage of KNN? / 1 mark

A. Low bias
B. High variance
(C.) Low complexity
D. All of the Above

8. Which of the following is a key feature of KNN? / 2 marks

A. Non-parametric learning
(B.) Parametric learning
C. Both A and B
D. None of the Above

9. The decision tree can then be used to make decisions about loan applications and help the bank decide which applicants should receive the loan, list two advantages and two disadvantages of using decision tree? / **4 Marks**

10. Differentiate Pruning from Entropy/ **4 marks**

9) Advantages

- It is suitable for small datagram
- It is more accuracy
- It uses both classification and Regression
- It is faster.
- It is easy to get output.

disadvantages

- It is limited to small data set.
- It can not identify accuracy of data clearly.

10) a) Pruning is process of reducing the size of decision tree utilized by CART in dataset.

b) Entropy: is one of the measure of impurity or purity in dataset. These are information after splitting/sub