Unit 7 - Week 5

Assignment 5
Due on 2023-09-25, 23:59:57

1. Suppose you are growing yeast with 4 different nutrient conditions. You also want to include a control where there is no nutrient addition. Design a balanced experiment with two factors: nutrient type and nutrient concentration.

2. Consider a binary classification problem. Canada and the United States are the two classes. You have data on the number of immigrants from each country and the number of crimes committed in the US. Design a classification model to predict whether a crime was committed by a Canadian or an American based on this data.

3. Consider the Bayes classifier and the Bayes decision function. For a continuous random variable, can you express the Bayes decision function as a linear function? Explain.

4. Let's say you have a set of points in a two-dimensional space, and you want to classify them into two groups. The points are not linearly separable. Describe how you would approach this problem using a support vector machine (SVM).

5. Consider the following diagram. What is the decision boundary for the Bayes classifier?

- X-axis: number of immigrants
- Y-axis: number of crimes committed

6. Write a short essay explaining the importance of outlier detection in machine learning.

7. Suppose you are working on a project to predict customer churn in a subscription-based service. Explain how you would approach this problem using a decision tree classifier.

8. In the context of binary classification, what is the significance of the false positive rate and the false negative rate?

9. Consider a scenario where you are working on a project to predict the default status of loan applicants. Explain how you would approach this problem using logistic regression.

10. Suppose you are working on a project to predict the sentiment of movie reviews. Explain how you would approach this problem using a Naive Bayes classifier.

11. Consider the following decision tree for a classification problem. Explain how this tree is constructed.

- Root node: number of immigrants
- Left child: number of crimes committed
- Right child: number of immigrants

12. Explain the concept of feature selection in machine learning. Why is it important?