UNIVERSITY OF NAIROBI

DEPARTMENT OF COMPUTING AND INFORMATICS

CSC3212 Machine Learning Algorithms & Programming

Due Date: 17 May 2024

**Practical Assignment Instructions**

You are provided with a dataset containing demographic information on different people and whether they earn above or below USD 50K a year. Your work is to first go through the data and understand it. Then you are to clean the data and prepare it for training the following machine learning models:

1. Naive Bayes Classifier
2. k-NN Classifier
3. Decision Trees Classifier
4. Random Forest Classifier

You are to **submit a detailed notebook** containing your code, and a summary report (max 5 pages) highlighting the following:

1. The Data Cleaning Process
2. For each of the models above, the best parameters and performance (at least accuracy and F1 score)
3. Discussion and Conclusion of Results
4. Challenges and Recommendations