THE CO-OPERATIVE UNIVERSITY OF KENYA

END OF SEMESTER EXAMINATION DECEMBER -2022

EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE IN COMPUTER SCIENCE
(YR IV SEM I)

UNIT CODE: BCSC 4147

UNIT TITLE: MACHINE LEARNING

DATE: WEDNESDAY, 21ST DECEMBER, 2022

TIME: 2:00 PM – 4:00 PM

INSTRUCTIONS:

- **Answer question ONE (compulsory) and any other TWO questions**

QUESTION ONE

(a) Describe the term learning as used in Machine Learning. [5 Marks]
(b) Discuss the three categories of machine learning. Be sure to mention the techniques used per category. [9 Marks]
(c) With the aid of a diagram, discuss the supervised learning process as applied in Machine Learning. [5 Marks]
(d) When building a naive Bayes classifier, we assume that the attributes are conditionally independent given the class label. What is the benefit of this assumption? [5 Marks]
(e) During preparation of machine learning models, the dataset is normally split into three sets. Discuss the three sets with a clear outline of their roles. [6 Marks]

QUESTION TWO

(a) Support Vector Machine (SVM) is a highly accurate classification technique. However, SVM classifiers suffer from slow processing when training with a large set of data tuples. Using an example of your choice, discuss the difficulty could be overcome. [8 Marks]
(b) Using appropriate examples in each case, discuss the three main approaches of using association rules for classifications. [12 Marks]

QUESTION THREE

(a) Considering accuracy, efficiency and interpretability, compare and contrast the use of Rule Induction to that of Decision Trees in the process of model training. [9 Marks]
(b) Discuss the concept of sequential covering as applied in rule induction systems. Ensure that your discussion includes a mention of the two sequential covering approaches with their differences. [11 Marks]

QUESTION FOUR

(a) Using an example of your choice, discuss the concept of Information Theory (including entropy measurement and information gain) with respect to building of decision tree classifiers. [14 Marks]
(b) Overfitting is one likely challenge expected in the use of Decision Tree classifiers. Discuss what overfitting is and the two common approaches that can be used to avoid overfitting. [6 Marks]

QUESTION FIVE

(a) Model evaluation is a critical step in the Machine Learning process. Using an example in each case, discuss any four techniques used to evaluate Machine Learning Models. [12 Marks]

(b) Discuss the four classification measure that form the set of confusion matrix measures used in Machine Learning. [8 Marks]