KISII UNIVERSITY

UNIVERSITY EXAMINATIONS

FOURTH YEAR EXAMINATION FOR THE AWARD OF THE DEGREE OF BACHELOR OF INFORMATION TECHNOLOGY /BACHELOR OF APPLIED COMPUTER SCIENCE
FIRST SEMESTER, 2023/2024
(AUGUST-DECEMBER, 2023)
COMP 460/ACMP 496: MACHINE LEARNING LOGARITHIMS

STREAM: Y4 S1	TIME: 2 HOURS
DAY: THURSDAY, 9.00 -11:00 AM	DATE: 16/11/2023

INSTRUCTIONS
1. *Do not write anything on this question paper.*
2. *Answer Question ONE [Compulsory] and any other TWO Questions*

QUESTION ONE

a) Explain a False Positive rate and False Negative rate with an illustration and How they are Significant? [6mks]

b) Discuss the stages of data preprocessing architecture. Use a flow chart to illustrate this [10mks]

c) What are challenges of conventional machine learning algorithms like Random forest, K-nearest neighbor and support vector machines for cancer prediction? [9mks]

d) How do you choose an appropriate evaluation metric for a given problem? [5mks]

QUESTION TWO

a) Explain supervised, unsupervised and reinforcement learning in detail? [6mks]

b) Discuss the performance evaluation metrics suitable for predicting the future occurrence of breast cancer? [9mks]

c) How do you evaluate the performance of a recommendation system? [5mks]

QUESTION THREE

a) A data set is given to you and it has missing values which spread along 1 standard deviation from the mean. How much of the data would remain untouched? [6mks]
b) Explain the handling of missing or corrupted values in the given dataset. [6mks]
c) Explain the differences between Random Forest and Gradient Boosting machines [8mks]

QUESTION FOUR

a) Differentiate between regression and classification [4mks]
b) Differentiate between K-Means and KNN algorithms? Giving exam-le in each case [6mks]

c) Write a simple code to binarize data [10mks]

QUESTION FIVE

a) What are the steps for supervised learning [5mks]
b) Explain the distance metrics that are used for KNN [8mks]
c) Explain the main approaches that you can use to analyze your data whether image dataset or non-image dataset and also methodologies behind them [7mks]