TECHNICAL UNIVERSITY OF MOMBASA

INSTITUTE OF COMPUTING AND INFORMATICS

DEPARTMENT OF COMPUTER SCIENCE & INFORMATION TECHNOLOGY

UNIVERSITY EXAMINATION FOR:

BMCS 12S

EIT4454: MACHINE LEARNING

END OF SEMESTER EXAMINATION

SERIES: APRIL 2016

TIME: 2 HOURS

DATE: 10 May 2016

Instructions to Candidates
You should have the following for this examination
-Answer Booklet, examination pass and student ID
This paper consists of **FIVE** questions. Attempt question ONE (Compulsory) and any other TWO questions.
Do not write on the question paper.

Question ONE (Compulsory)

a) State any four areas where machine learning could be applied.   (4 marks)

b) Define machine learning as earlier defined by Tom Mitchell   (2 marks)

c) Explain the meaning of the following expressions: $p(X = x)$;   (3 marks)

d) Write an expression for the following graphical representation.   (3 marks)
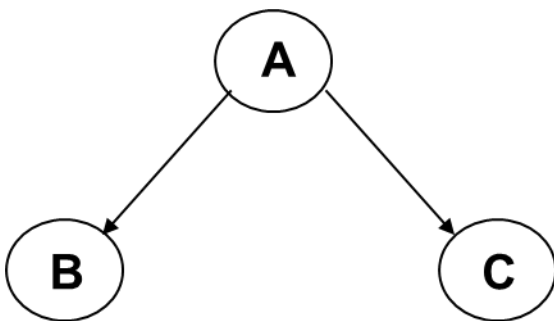
e) Explain the following terms as used in machine learning data set: training data, validation data, test data (3 marks)

f) Explain the types of tree pruning. (4 marks)

g) Explain the components of Bayesian theory. (4 marks)

h) Explain the difference between classification and regression. (4 marks)

i) Define the term outlies as used in machine learning and outline the possible reasons for their existence (3 marks)

Question TWO

a) Discuss Linear Regression by hypothesis citing the hypothesis and showing how selected parameter values determine the type of model. (7 marks)

b) Show clearly how the cost function of a logistic regression may be derived by determining deviation between the actual points and the regression line. (8 marks)

c) With suitable illustrations, show the difference between the terms: over fitting, normal fitting, and under fitting and state the implication of over fitting on the cost function. (5 marks)

Question THREE

a) Explain how the binary classifier hypothesis for logistic regression is derived. (6 marks)

b) Starting from the linear regression cost function show how the logistic regression cost function is achieved. (12 marks)

c) State any four optimization algorithms. (2 marks)

Question FOUR

a) Describe Neural Networks (5 marks)

b) With suitable a suitable Neural Network and equations, explain the concept of forward propagation

(12 marks)

c) Explain the term back propagation as used in Neural Network. (3 marks)

Question FIVE

a) Starting from logistic regression cost function, describe how the support vector model is obtained. (16 marks)

b) Define the following terms: i) unsupervised learning; ii) clustering (4 mark)