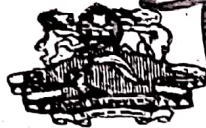SCO 409

KENYATTA UNIVERSITY

EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE

(COMPUTER SCIENCE)

SCO 409: NATURAL LANGUAGE PROCESSING

1ST SEMESTER 2022/2023   TIME: 2 HOURS

QUESTION 1 (COMPULSORY)

a) Explain the following linguistic concepts as used in NLP   (4 Marks)

   i.   Morphology

  ii.   Syntax

 iii.   Semantics

  iv.   Pragmatics

b) Briefly evaluate three potential applications of NLP   (6 Marks)

c) Explain two reasons why corpora are required in NLP projects   (4 Marks)

d) Provide the formula for a bigram model as used in NLP.   (2 Marks)

e) The idea of a context-free grammar (CFG) should be familiar from formal language theory.
Identify the four components of a CFG   (4 Marks)

f) Design / Draw a parse tree to depict how to read this sentence: (6 Marks)
        "They can fish in December"

g) With the help of suitable examples, distinguish between lemmatization and stemming as used in NLP   (4 Marks)

QUESTION TWO

a) Distinguish between the four different types of machine translation in NLP (6 Marks)

b) Distinguish between the 2 types of models in speech recognition (4 Marks)

c) With the help of a diagram, describe the process of information retrieval (10 Marks)

QUESTION THREE

a) Interpret the information extraction process depicted in the process flow below. Take note to explain the stages/concepts indicated (5 Marks)
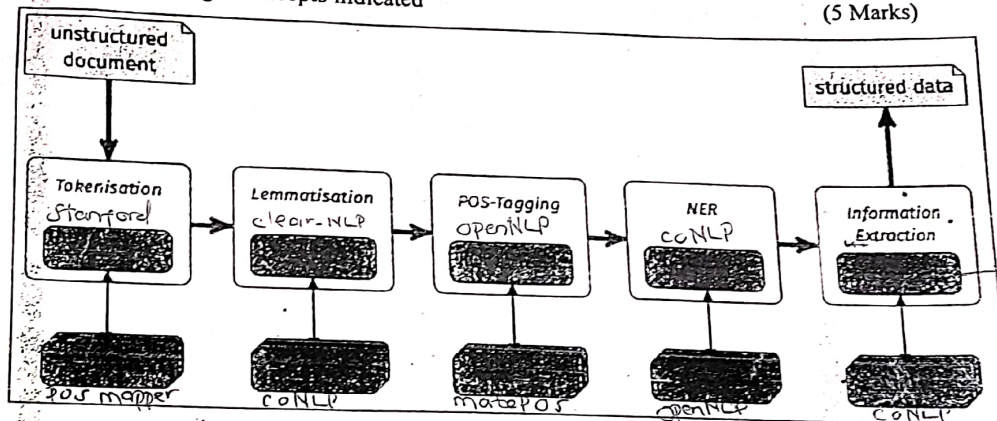
b) Briefly explain how the following most popular supervised NLP machine learning algorithms used in NLP work (9 Marks)

 i. Support Vector Machines.

 ii. Bayesian Networks.

 iii. Neural Networks/Deep Learning.

c) Evaluate the elements of a good corpus design (6 Marks)
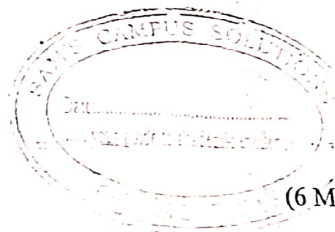
QUESTION FOUR

a) Explain the key inputs in the evaluation of Word Sense Disambiguation (8 Marks)

b) Explain the properties of Stochastic POS Tagging (5 Marks)

c) With the aid of a diagram, show the phases or logical steps in NLP (7 Marks)

QUESTION FIVE

a) Distinguish between top-down parsing and bottom-up parsing in NLP   (4 Marks)

b) Write the Python packages for performing the following NLP tasks   ·   (4 Marks)
   i.   Divide the input text into sentences
   ii. Divide the input text into words and punctuation marks
   iii. Extract the base form of the words
   iv. Extract the base form of the word depending upon whether it is used as a noun or as a verb

c) Semantic analysis creates a representation of the meaning of a sentence. in relation to word representation or representation of the meaning, designate four building blocks that play an important role   (8 Marks)

d) Explain the following regular expressions   ·   (4 Marks)
   i. $(0 + 10^*)$
   ii. $(a + b)^*$

END

KENYATTA UNIVERSITY
EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE
(COMPUTER SCIENCE & COMPUTER AND MATHS)

SCO 411: NEURAL NETWORKS

1ˢᵀ SEMESTER, 2022/2023

TIME: 2 Hours

INSTRUCTIONS: Answer Question ONE and any other TWO questions.

Question One - Compulsory (30 marks)

a. Define the generalization as used in neural network training. (2 marks)

b) Explain four main characteristics of an Artificial Neural Networks system. (4 marks)

c) Explain how supervised learning mechanism is used to train a neural network system. (4 marks)

d) Consider the unit shown below. Suppose that the weights corresponds to the three inputs which have the following values.

$$w_1 = 2$$
$$w_2 = -4$$
$$w_3 = 1$$

and the activation of the unit is given by the step-function:

$$\varphi(v) = \begin{cases} 1 & if\ v \geq 0 \\ 0 & otherwise \end{cases}$$

Calculate what will be the output value y of the unit for each of the following input patterns:

Pattern	P_1	P_2	P_3	P_4
x_1	1	0	1	1
x_2	0	1	0	1
x_3	0	1	1	1

(4 marks)

INVOLVEMENT IN EXAMINATION IRREGULARITY SHALL LEAD TO DISCONTINUATION

e) Consider a single-layer network (perceptron) with two inputs, having weights $w_1 = 0.5$ and $w_2 = -0.25$. Determine the output of the network when given an input vector $(1,1)$ for sigmoid activation function. (4 marks)

(f) Consider the following equation for back propagation algorithm

$$\delta_k = O_k(1 - O_k)(y_k - O_k)$$

$$\delta_j = O_j(1 - O_j)\sum_k \delta_k W_{jk}$$

Explain the meaning of all terms in the equation. (4 marks)

g) Explain four main approaches to overfitting avoidance in neural network training? (8 marks)

Question Two

a) Describe how competitive learning process takes place in neutral network. (6 marks)

b) Explain how the network training time and accuracy is influenced by the size f the hidden layer (6 marks)

c) State and explain two parameters that are set in neutral networks training. (4 marks)

d) Distinguish between linearly separable and nonlinearly separable problems. Give examples. (6 marks)

Question Three

a) State and explain the purpose of the three layers within the structure of generalized radial basis function network topology. (6 marks)

b) Describe briefly the three approaches you can use to find the centre's and variances for the radial basis network training algorithm (6 marks)

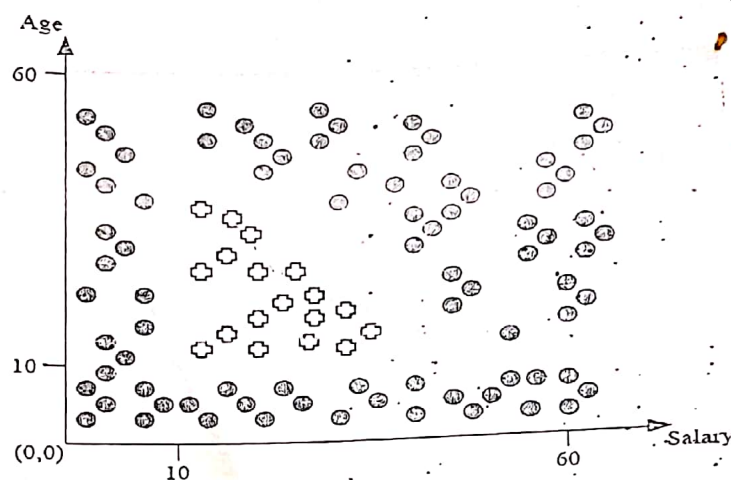c) Describe four properties of hope freld network. (8 marks)

Question Four

a) Describe how **Gradient Descent** learning process takes place in neural network. (8 marks)

b) Explain the function of the bias θ in a feed forward neural network with a single hidden layer used for classification problems? (2 marks)

c) The diagram below classifies a group of people according to whether they enjoy playing arcade games (+) or not (o). Each data point is a person surveyed, and the axes represent the age (in years) and the salary (in hourly wages) of each respondent.

i. Can this data be classified using a neural network with no hidden layer? Explain why or why not.

(3 marks)

ii. Based on the diagram above, draw the simplest neural network that can classify this data, given inputs of age and salary, and outputs of YES (people who like arcade games) and NO (people who don't like arcade games). Indicate the weights and activation functions for each node in the network.

(7 marks)

Question Five

a) Contrast Backpropagation Neural Network with Kohnen Self Organizing Neural Network algorithm.

(5 marks)

b) Consider the neural network below showing the initial weights, the inputs (0.9 and 0.5) and the output
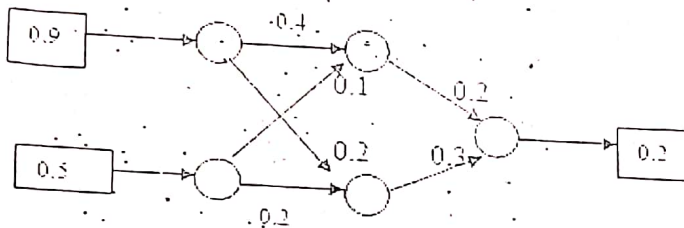
Using back-propagation algorithm, determine the new adjusted weights V_1 and W_1 after one epoch. Assume that the INPUT layer uses a LINEAR FUNCTION, the Hidden and OUTPUT layers use the SIGMOID function $y = 1/1 + e^x$ and the learning rate is 0.1

(15 marks)

KENYATTA UNIVERSITY

UNIVERSITY EXAMINATIONS 2017/2018

SECOND SEMESTER EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE (COMPUTER SCIENCE & MATHEMATICS)

SCO 411: NEURAL NETWORKS

TIME: 8.00 A.M. - 10.00 A.M.

DATE: FRIDAY 24TH AUGUST 2018

INSTRUCTIONS:

Answer Question ONE and any other TWO questions.

Question One - Compulsory (30 marks)

a. Briefly discuss three applications of artificial neural networks system in the financial industry. (6 marks)

b) Explain four main characteristics of an Artificial Neural Networks system. (4 marks)

c) Explain how supervised learning mechanism is used to train a neural network system. (4 marks)

d) Consider the unit shown below. Suppose that the weights corresponds to the three inputs which have the following values.

$$w_1 = 2$$
$$w_2 = -4$$
$$w_3 = 1$$

and the activation of the unit is given by the step-function:

$$\varphi(v) \begin{cases} 1 \text{ if } v \geq 0 \\ 0 \text{ otherwise} \end{cases}$$

Calculate what will be the output value y of the unit for each of the following input patterns:

Pattern	P_1	P_2	P_3	P_4
x_1	1	0	1	1
x_2	0	1	0	1
x_3	0	1	1	1

(4 marks)

INVOLVEMENT IN ANY EXAMINATION IRREGULARITY SHALL LEAD TO DISCONTINUATION

Consider a single-layer network (perceptron) with two inputs, having weights $w_1 = 0.5$ and $w_2 = -0.25$. Determine the output of the network when given an input vector $(1,1)$ for sigmoid activation function. (4 marks)

(g) Consider the following equation for back propagation algorithm

$$\delta_k = O_k(1 - O_k)(y_k = O_k)$$

$$\delta_j = O_j(1 - O_j)\sum_k \delta_k W_{jk}$$

Explain the meaning of all terms in the equation. (4 marks)

f) Explain four main approaches to overfitting avoidance in neural network training? (4 marks)

Question Two

a) State and explain the purpose of the three layers within the structure of generalised radial basis function networks topology. (6 marks)

b) Describe briefly the three approaches you can use to find the centre's and variances for the radial basis networks training algorithm? (6 marks)

c) Describe four properties of hope field network. (8 marks)
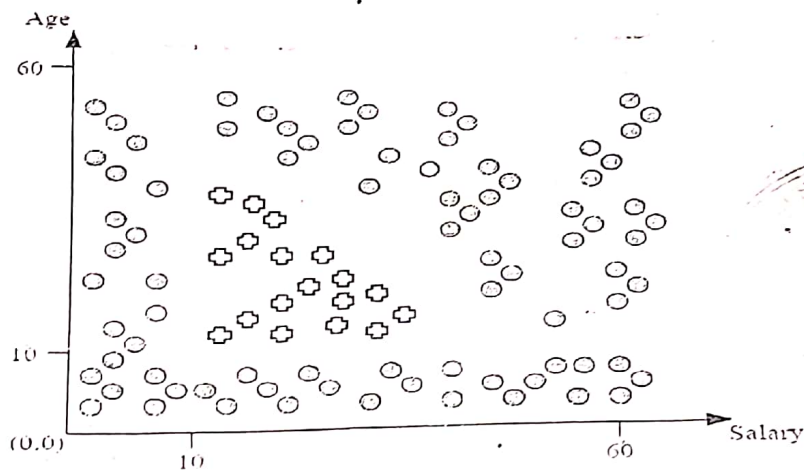
Question Three

a) Describe how competitive learning process takes place in neural network. (6 marks)

b) Explain the term Hebb's rule? (4 marks)

c) State and explain two parameters that are set in neural networks training. (4 marks)

d) Distinguish between linearly separable and nonlinearly separable problems. Give examples. (6 marks)

Question Four

a) Describe how **linear neuron learning** process takes place in neural network. (6 marks)

b) Define the following terms as used in neural network training. (4 marks)

i) Genralization

ii) Overfitting

c) Explain why XOR problem cannot be solved by a single layer perceptron and how it is solved by a Multilayer Perceptron. (6 marks)

d) Explain how the network training time and accuracy is influenced by the size of the hidden layer. (4 marks)

INVOLVEMENT IN ANY EXAMINATION IRREGULARITY SHALL LEAD TO DISCONTINUATION

Question Five

a) . Describe how **Gradient Descent** learning process takes place in neural network. (8 marks)

b) Explain the function of the bias Θ in a feedforward neural network with a single hidden layer used for classification problems? (2 marks)

c) The diagram below classifies a group of people according to whether they enjoy playing arcade games (+) or not (o). Each data point is a person surveyed, and the axes represent the age (in years) and the salary (in hourly wages) of each respondent

i. Can this data be classified using a neural network with no hidden layer? Explain why or why not. (3 marks)

ii. Based on the diagram above, draw the simplest neural network that can classify this data, given inputs of age and salary, and outputs of YES (people who like arcade games) and NO (people who don't like arcade games). Indicate the weights and activation functions for each node in the network. (7 marks)

****

INVOLVEMENT IN ANY EXAMINATION IRREGULARITY SHALL LEAD TO DISCONTINUATION

KENYATTA UNIVERSITY
EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE
(COMPUTER SCIENCE & COMPUTER AND MATHS)

SCO 411: NEURAL NETWORKS

1ST SEMESTER, 2022/2023   TIME: 2 Hours

INSTRUCTIONS: Answer Question ONE and any other TWO questions.

SCO 413

KENYATTA UNIVERSITY
EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE
(COMPUTER SCIENCE)
SCO 413: ROBOTICS

1ST SEMESTER 2022/2023   TIME: 2 Hours

INSTRUCTIONS:

Answer question One and any other Two Questions.

Question One is Compulsory an carries 30 marks. The remaining questions carry 20 marks each.

QUESTION 1

A) Define the following terms: (3 marks)

    i)   Biomimetic

    ii)   Reactive control

    iii)   Echolocation

B) Suppose you were to design a machine that could be trained, learn and speculate, Which FIVE main components and capabilities should it have   (5 marks)

C) Differentiate the following terms

    i)   Photophilic and photophobic   (2 marks)

    ii)   Excitatory connection and inhibitory connection   (2 marks)

    iii)   Observable, Partially Observable, and Hidden State   (3 marks)

D) List FIVE things a machine needs to use for it to be intelligent   (5 marks)

E) Outline the THREE fields that played a big part in Robotics field   (6 marks)

F) Explain the Main functions of Actuators   (4 marks)

QUESTION 2

A) List and explain FOUR power issues encountered in Robotics   (8 marks)

B) Explain how a Passive and Active sensor works   (4 marks)

C) Outline THREE ways in which Switches can be used as sensors   (6 marks)

D) There are two basic ways in which reflective optosensors can be arranged, based on the relative positions of the emitter and the detector, list them   (2 marks)

INVOLVEMENT IN EXAMINATION IRREGULARITY SHALL LEAD TO DISCONTINUATION

QUESTION 3

A) Light sensors can be cleverly positioned, oriented, and shielded in order to improve their accuracy and range properties. Explain these three properties   (6 marks)

B) Explain what can be done with simple idea of measuring light reflectance/reflectivity?

          (10 marks)

C) Explain why legged locomotion is a more difficult robotic problem compared with wheeled locomotion   (4 marks)

QUESTION 4

A) The ratio of the controllable DOF (let's call them CDOF) to total DOF (let's call them CDOF) on a robot tells us quite a bit about how easy it is to control its movement. Briefly explain the three   (6 marks)

B) State the TWO general rules of gearing in terms of Torque and Speed   (8 marks)

C) List FOUR things that a robot needs in order to process sensory inputs   (4 marks)

D) Describe Hybrid Control   (2 marks)

QUESTION 5

A) Describe FIVE characteristics of Desirable robot gaits   (5 marks)

B) Explain why tele-operating complicated manipulators is a great challenge in Robotics   (6 marks)

C) List and explain FIVE sources of Uncertainty in robotics   (5 marks)

D) Explain FOUR ways of detecting people in the environment   (4 marks)

KENYATTA UNIVERSITY
EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE
(COMPUTER SCIENCE)
SCO 415: DATA WAREHOUSE AND DATA MINING

1ST SEMESTER 2022/2023   TIME: 2 HOURS

INSTRUCTIONS: Answer Question ONE and any other TWO questions

Question 1

(a) The range of real life problems potentially addressable through data mining can be grouped into a few task types. List four such task types.   (4 marks)

(b) With the aid of suitable illustrative examples outline the significance of the following concepts:
- (i) The ETL Process
- (ii) Model Overfitting
- (iii) Split Sample Testing   (12 marks)

(c) The following statements were made in the context of Data Science for Business. Expound on the possible meaning of each of them.
- (i) "Getting information off internet is like taking drink from a firehose:
- (ii) "A data warehouse is a "subject oriented, non- volatile, time- variant and integrated collection of data".
- (iii) "Our biggest challenge of making the evolution from a knowing culture to a learning culture is really not the cost".   (6 marks)

(d) A Majority of data mining project managers insist on the use of a formal methodology. Explain four reasons that could support the insistence of business analytics project manager on the use of a formal methodology   (8 marks)

                              (Total: 30 marks)

Question 2

"To find signals in data, we must learn to reduce the noise-not just the noise that resides in the data but also the noise that resides in us. It is nearly impossible for noisy minds to perceive anything but noise in data"

(a) With the aid of appropriate illustrations describe four examples of situations that could be fit the description of 'noise that resides in the data' in the above statement. For each, briefly outline a way out for the data- mining practitioner   (12 marks)

(b) (i) What could the term 'noisy minds' be making reference to?   (4 marks)

     (ii) How can we address the problem of 'noisy minds'?   (4 marks)

                              (Total: 20 marks)