UNIVERSITY OF RUHUNA

BACHELOR OF COMPUTER SCIENCE (BCS) (SPECAIL) DEGREE LEVEL IV (SEMESTER I) EXAMINATION – AUGEST 2017

COURSE UNIT: CSC4133 – Neural Networks **TIME: 2 Hours**

Answer all four Questions.

(1)

(a)

(i) What is said to be a "**pattern**" in the context of pattern recognition systems?
(ii) Explain a criterion to define what is said to be a "**good**" pattern vector.

(04 Marks)

(b)

(i) What is the difference between the **feature extraction** and **feature selection** processes in the context of pattern recognition systems?
(ii) Briefly explain using a suitable example, how feature extraction can be performed in an image recognition context using the **projection histogram** method.

(06 Marks)

(c) Explain the fundamental difference between the **Bagging** and **Boosting** ensemble learning methods.

(04 Marks)

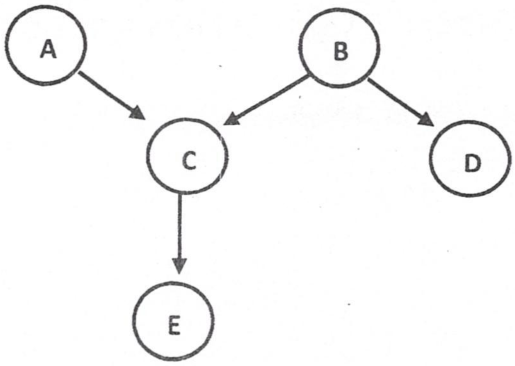
(d) Assume that three classifiers (C1, C2 and C3) are used to form multiple classifier systems as shown below to distinguish three classes (Class1, Class 2, and Class 3).

Classifier	Importance of Classifier	Posterior Probability		
		Class 1	Class 2	Class 3
C1	0.61	0.69	0.13	0.18
C2	0.57	0.21	0.66	0.13
C3	0.69	0.12	0.27	0.61

What will be the output class of the multiple classifier systems if the **linear weighted average** was used to combine the outputs of the three classifiers?

(06 Marks)

(e) Consider the Bayesian Belief Network (BBN) below.



(i) Identify and draw linear, converging and diverging type connections in the above BBN.

(ii) Write down the joint distribution P (A, B, C, D, E) of this BBN.

(05 Marks)

(2)

(a) A Neural Network based solution to a problem has an input of three colours namely blue, green and red.
Explain using a suitable sketch, how this data can be applied to the network using

(i) A single input node

(ii) Several input nodes.

(04 Marks)

(b) (i) Give two characteristics a training data set should have.

(ii) What is the adverse (negative) effect of over-fitting? Explain how over-fitting could be avoided.

(05 Marks)

(c) (i) Explain how the misclassification of data could be minimised in a Support Vector Machine.

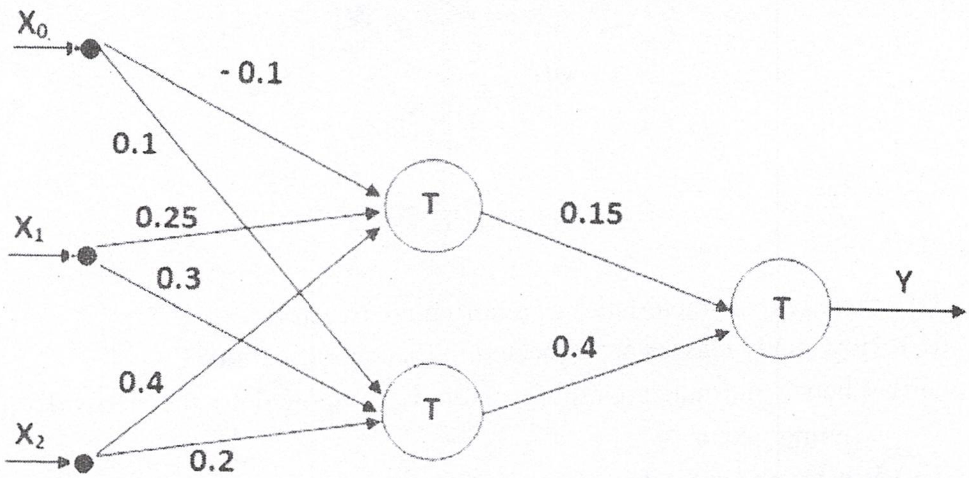
(ii) The following table shows the results of testing a neural network which has been designed for intruder detection.

No.of Data	Correctly classified	Misclassified	
		Intruder accepted	Genuine user rejected
100	84	4	12

Calculate the False Acceptance (FA) and the False Rejection (FR) rates.

(06 Marks)

(c) The figure below depicts an artificial neural network whose neural elements are linear threshold units. T is the threshold value of each unit (node). The numerical values alongside each arrow indicate the network's weight values. If the inputs (X_1 and X_2) are restricted to binary values, the network implements a Boolean function.



(i) What is the role of X_0?

(ii) What is the role of T?

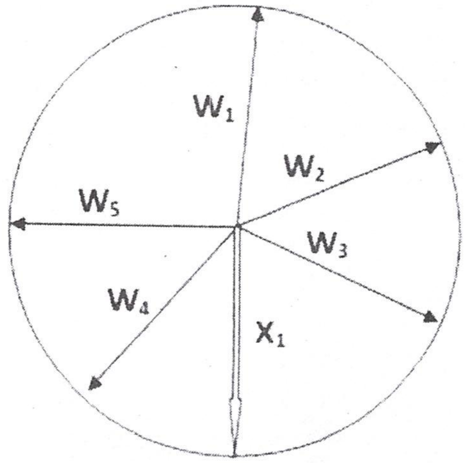
(iii) Give an alternative activation function for T.

(iv) What is the Boolean function implemented in this network when T = 0.5? Show your calculations clearly.

(10 Marks)

(3)

(a) The normalized weight vectors $W_1, W_2, \ldots\ldots, W_5$ of a Self-Organizing Map (SOM) are shown in the following diagram.



(i) State two characteristics of a normalized vector.

(ii) How many classes are expected in this classification?

(iii) When a normalized input vector $\mathbf{X_1}$ is applied to the network, identify the winning vector.

(iv) What is the most dissimilar weight vector?

(06 Marks)

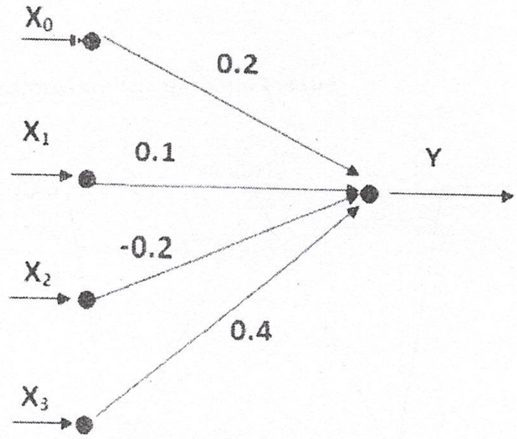
(b) One significant application of an Instar-Outstar network is, deriving a perfect output from an imperfect input. By considering the recognition of used currency notes as an example and using a suitable diagram, briefly explain the steps involved in the process of organization of an Instar-Outstar network for such an application.

(08 Marks)

(c) Using a suitable diagram of distribution of data in two classes, explain a situation in which a Radial Basis Function (RBF) network performs better than a multi-layer network.

(04 Marks)

(d) A supervised neural network (as shown in the diagram) is trained for the input vector [1, 0.3, 0.2, 0.4] and the target T = 1. Weights are as shown.



(i) Calculate the initial network output Y and the initial error δ if no activation is used.

(ii) Using the Delta rule $\Delta W = \eta\delta x$, calculate the ΔW and hence W_{new}. Symbols have their usual notations and assume that $\eta = 0.5$

(iii) Plot the Error (E) vs Weight (W) curve if the gradient descend method is used.

(07 Marks)

(4)

(a) In a Fuzzy logic based system to rank and select an all rounder (who performs well in fielding, batting and bowling) for the national cricket team using the performance of suitable candidates during their recent first class cricket tournaments, the following fuzzy functions are used.

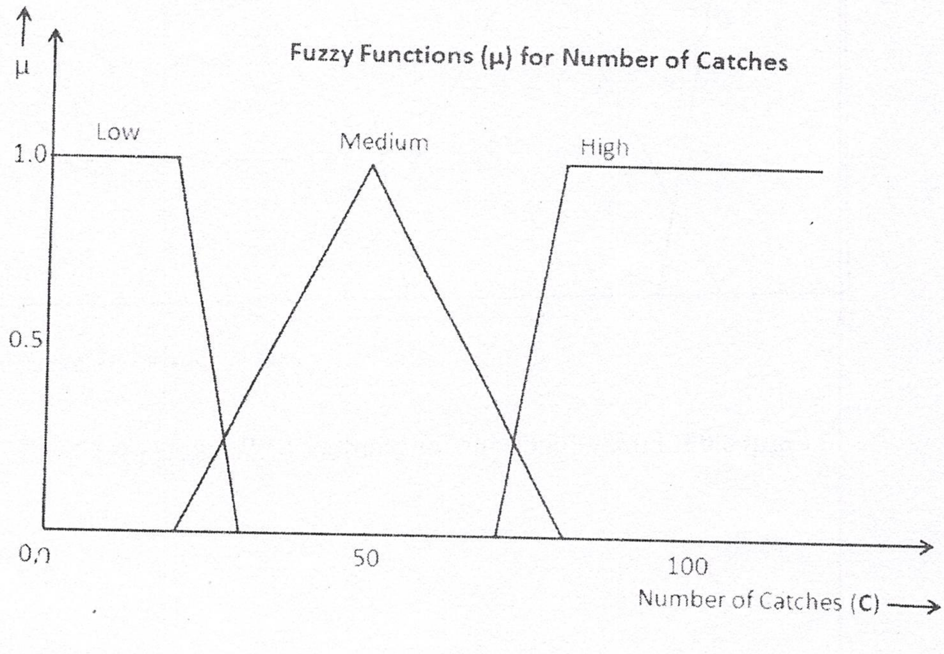


Figure 01. Fuzzy functions for Number of Catches (C)

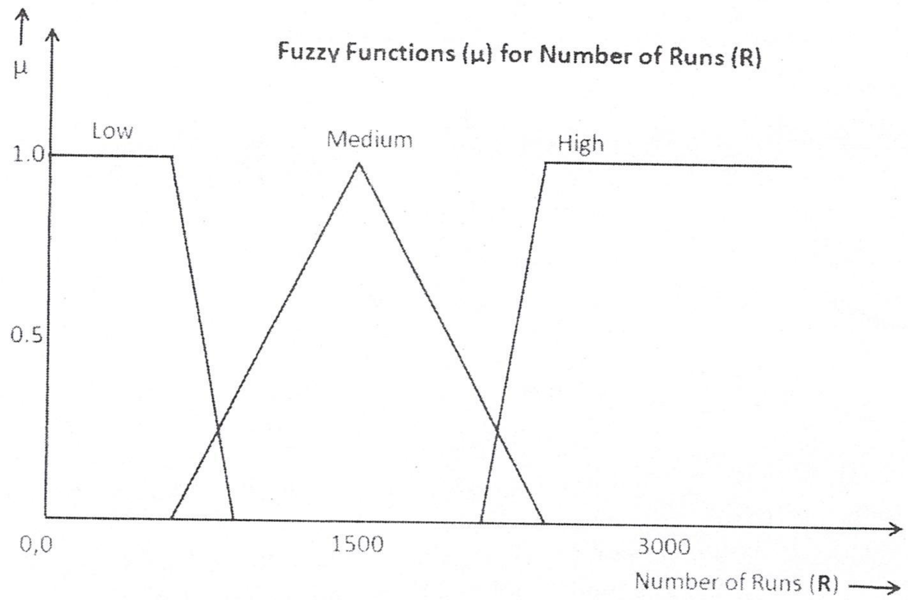


Figure 02. Fuzzy functions for Number of Runs (R)

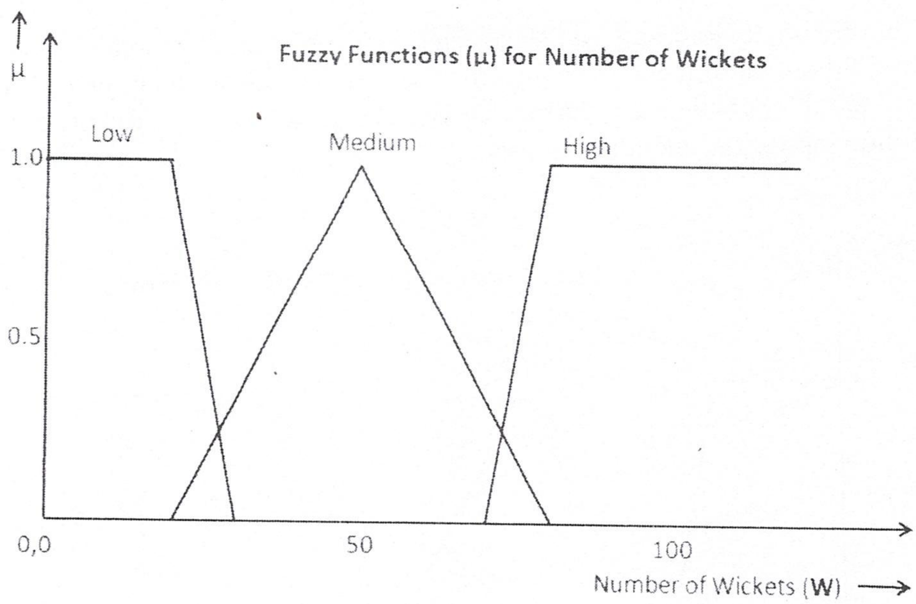


Figure 03. Fuzzy functions for Number of Wickets (W)

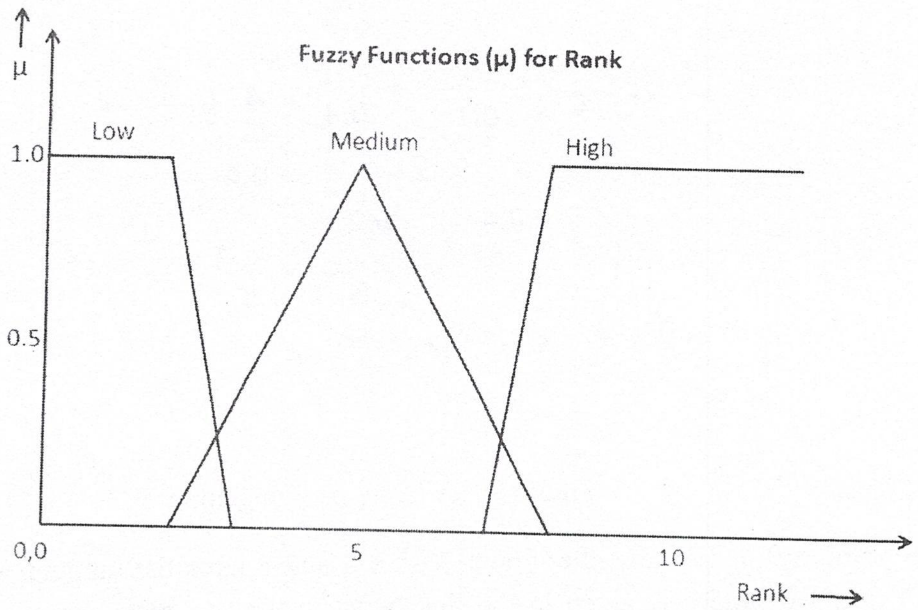


Figure 04. Fuzzy functions for Rank

Some fuzzy rules used in this system are given below.
(C – No. of Catches, R – No. of Runs, W – No. of Wickets)

I. *If **C** is **Low** and **R** is **Low** and **W** is **Low** then **Rank** is **Low***

II. *If **C** is **Medium** and **R** is **Low** and **W** is **Medium** then **Rank** is **Medium***

III. *If **C** is **Medium** and **R** is **Medium** and **W** is **Medium** then **Rank** is **Medium***

IV. *If **C** is **High** and **R** is **Medium** and **W** is **Medium** then **Rank** is **Medium***

V. *If **C** is **High** and **R** is **High** and **W** is **High** then **Rank** is **High***

(i) A candidate who is eligible to be an all-rounder has the following values for C, R and W.

- Number of Catches (C) = 75
- Number of Runs (R) = 750
- Number of Wickets (W) = 40

Derive the respective Fuzzy values for each of the above.

(ii) What are the valid Fuzzy rules for the values of C, R and W in (i) above?

(iii) Derive the Rank of the candidate in (i) above.

(12 Marks)

(b) (i) Explain briefly how the Boltzmann machine can be used to solve the problem of local minima experienced in training. Present your answer in terms of energy and probability.

(03 Marks)

(ii) The structure of a Boltzmann machine is given in figure 06.

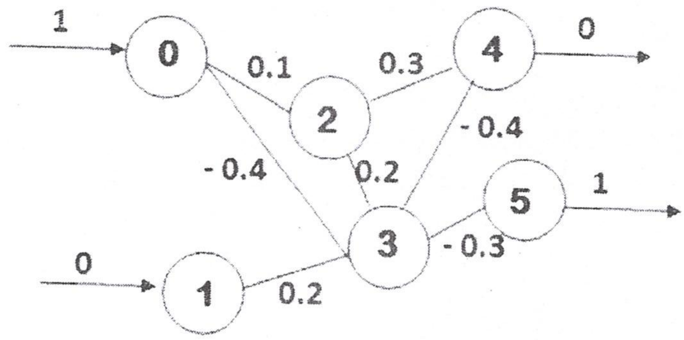


Figure 06. A Boltzmann machine

Nodes 0 and 1 have the inputs. Nodes 4 and 5 have the outputs. Weight between any two nodes is given in the link between the two respective nodes. Inputs and outputs are *clamped* with the given binary values and the nodes 2 and 3 are initially set to 1 and 0 respectively.

A node fires (i.e. its state is set to 1) if the energy of the node $E = \sum x_i w_i >= 0$. Otherwise, it does not fire (i.e. the state is set to 0).

(I) Copy the following table to your answer book and complete the columns 2, 3 and 4 for the functioning of the above Boltzmann machine. (Show the relevant calculations in Column 3).

Node Id.	Initial state	Energy (E)	State 2
0			
1			
2			
3			
4			
5			

(II) Has the Boltzmann machine achieved a stable state? Justify your answer.

(III) If the answer to (II) above is "Yes", explain how you could verify this stable state as the global minimum.

If the answer to (II) above is "No", explain how you could proceed to achieve a stable state.

(10 Marks)

****