Name:

Enrolment No:

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES

End Semester Examination, May 2022

Course: Neural Networks. **Semester: VI**

Program: B.Tech (CSE+AIML) **Time : 03 hrs.**

Course Code: CSAI 3001 **Max. Marks: 100**

Instructions:

SECTION A

(5Qx4M=20Marks)

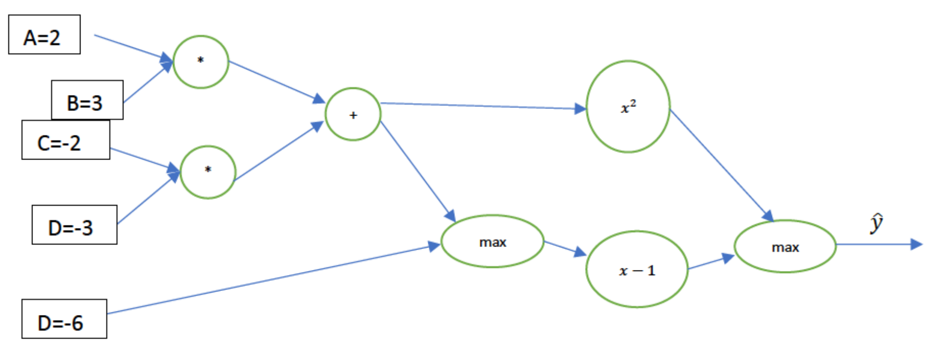
.

S. No.	Give Answer with a figure in atmost 5 lines.	Marks	CO
Q 1	Discuss the perceptron model.	**4**	**CO1**
Q 2	Discuss Adam optimization algorithms.	**4**	**CO2**
Q 3	Explain the use of the backpropagation algorithm to find the gradient in a graph with an example.	**4**	**CO3**
Q4	Discuss the use of weight initialization methods in ANN.	**4**	**CO4**
Q5	Discuss the limitations of the Autoencoder.	**4**	**CO5**

SECTION B

(4Qx10M= 40 Marks)

Q 6	$I=$¿ $K=$¿ $O=I\otimes K$ *where* $\otimes$ is convolution operation with stride 2 and 0 padding. Calculate the size and value of O.	**2+8**	**CO1**
Q 7	Suppose you have a build a classifier model with the following details. • Weights and biases initialized to $\aleph(0,1)*10^{-4}$. Where $\aleph$ denotes the normal distribution. • Relu is used as activation function • Only one neuron the final layer • Loss is binary cross entropy You observe that during the training loss is flat from the start. What is the cause of the problem and how to solve it?	**3+7**	**CO2**

Q 8	 Calculate $\frac{\partial \hat{y}}{\partial A}, \frac{\partial \hat{y}}{\partial B}, \frac{\partial \hat{y}}{\partial C}, \frac{\partial \hat{y}}{\partial D}$ using backpropagation. (Note: x represents the input given to a node and node levels denotes the function computed at the node)	**4*2.5**	**CO3**
Q 9	Illustrate the shortcomings of RNN and Discuss possible solutions (at least 3).	**5+5**	**CO5**
SECTION-C (2Qx20M=40 Marks)			
Q 10	1. Sketch the GAN architecture. 2. Write the objective of Generator 3. Write the Objective of Discriminator 4. Define the training loop to train GAN	**6+2+2+10**	**CO5**
Q 11	You are given a content image I and a style image, S. The neural style transfer method allows you to obtain an output image Y that has I 's content and S 's style. In order to perform this operation, a pretrained VGG-16 network should be used. Apply and discuss the role of the VGG-16 network in neural style transfer with figures and equations. OR In many countries routine vital statistics are of poor quality, and often incomplete or unavailable. In countries where vital registration and routine health information systems are weak, the application of verbal autopsy (VA) in demographic surveillance systems or cross-sectional surveys has been suggested for assessing cause-specific burden of mortality. The technique involves taking an interviewer-led account of the symptoms and signs that were present preceding the death of individuals from their caretakers. Traditionally the information obtained from caretakers is analysed by physicians and a cause(s) of death is reached if a majority of physicians on a panel agreed on a cause(s). The accuracy of physician reviews has been tested in several settings using causes of death assigned from hospital records as the 'gold standard'. Although physician reviews of VA gave robust estimates of cause-	**20** **OR** **10+10**	**CO4**

specific mortality fractions (CSMF) of several causes of death, the sensitivity, specificity and predictive values varied between causes of death and between populations and had poor repeatability of results.

Method

In brief, data were collected at three sites (a regional hospital in Ethiopia, and two rural hospitals in Tanzania and Ghana). Adults dying at these hospitals who lived within a 60-km radius of the institution were included in the study. A VA questionnaire was administered by interviewers with at least 12 years of formal education.

The reference diagnoses (gold standard) were obtained from a combination of hospital records and death certificates by one of the authors (DC) together with a local physician in each site. A panel of three physicians reviewed the VA data and reached a cause of death if any two agreed on a cause (physician review). The method used to derive algorithms from the data using logistic regression models has been described elsewhere.4 Each Subject was randomly assigned to the train dataset (n = 410) or test dataset (n = 386), such that the number of deaths due to each cause (gold standard) was the same in both datasets. If a cause of death had odd numbers, the extra subject was included in the train dataset. Symptoms (includes signs) with odds ratio (OR) “2 or “0.5 in univariate analyses were included in a logistic model and then those symptoms that were not significant statistically ($P > 0.1$) were dropped from the model in a backward stepwise manner. Coefficients of each symptom remaining in the model were summed to obtain a score for each subject. A cut-off score was identified for each cause of death (included 16 primary causes of adult death) that gave the estimated number of deaths closest to the true number of cause-specific deaths, such that the sensitivity was at least 50%.

We used the same train and test datasets used by Quigley et al. for training and testing an ANN. The data were ported to Microsoft Excel™ and analysed using NeuroSolutions 3.0™ (Lefebvre WC. NeuroSolution Version 3.020

Limitations of the technique

	At various points we have alluded to some of the difficulties and limitations of using neural networks for the analysis. These are summarized in following points. 1. Selecting input is not straight forward 2. Prioritizing cause specific mortality fraction over sensitivity or specificity is an manual process 3. Designing optimal network for each task is time consuming 4. Sensitivity and specificity may not high enough for the algorithms to generalizable to a variety of settings Even with sensitivity analysis, we had no way of working out which were going to be the most important inputs prior to creating a model and conducting a sensitivity analysis on it. There is some correlation with linearly predictive inputs that helps in the initial stages. Determining the weighting for the output for providing the optimum estimate of the CSMF was time-consuming. The soft- ware provides an option for prioritizing sensitivity over specificity, but no way of balancing the number of false positives and false negatives that would give an accurate CSMF estimate. Designing the optimal network topology requires building numerous networks in search of the one with the lowest least mean squared error. The number of hidden nodes, inputs and training time all affect the performance of the network. Whilst training is relatively quick compared to the many hours it took to train ANN in the early days of their development, it is still time-consuming to build and train multiple networks for each model. Cross-validation to prevent over-training required compromising the number of training examples to allow for a cross-validation dataset. Sensitivity and specificity of the ANN algorithms were not high enough to be generalizable to a variety of settings. Further- more, the accuracy of individual and summary estimates of CSMF obtained in this study could be due to the similarity in the CSMF between the training and test datasets. Thus large datasets from a variety of settings are needed to identify optimal algorithms for each site with different distributions of causes of death. Q A. Critically analyse the methods used in the above Scenario. What		

	do you think would be a better approach to the problem? Q B. Identify and explain the limitations in the methods used. How can the issues be resolved?		