Q1. -**CASCADE**------feedforward network include a connection from the input and every previous layer to following layers. [1]

Q2. An unauthentic relationship develops between input-output, if number of features used to fit ANN model is relatively larger than data, resulting in **OVERFITTING**--------[1]

Q3. High bias means the model is not "fitting" well on the **TRAINING** -------- data set. [1]

Q4. -**SMOTE**---------oversampling technique in unbalanced data works by selecting examples that are close in the feature space, drawing a line between the examples in the feature space and drawing a new sample at a point along that line. [1]

Q5. AUC for the worst classifier is - **0.5**-----. [1]

Q6. The symptoms of underfitting are High bias and **LOW-**---------variance. [1]

Q7 One percent of population is infected with a rare disease. If the clinical test classifies every example as positive. [Give values in number not %]

(i)Recall = **1.0-----**TP/(TP+FN)= 1/(1+10)

(ii)Precision = **0.01** ------- TP/(TP+FP)=1/(1+99)

(iii)Specificity =**0**------ TN/(TN+FP)=0/(0+99)

(iv) F-2 measure = ----**0.048**--- $(1+2^2)(P)(R)\ /(2^2\ P+R)$ [1+1+1+2=5]

Q8. In classification problems , name the the performance measure term (NOT expression)(i)_**PRECISON**_____ answers the question- What proportion of positive identification was actually positive? (ii) __**FPR/Fallout/FAR**_____ is calculated as the ratio between the number of negative events wrongly categorized as positive and the total number of actual negative events. [2M]

Q9. In remote sensing applicatons, If the correct classified pixels in a class are divided by the total number of pixels that were classified in that class, this measure is called **USER/CONSUMER**-------- accuracy. [1]

Q10. A --**BIAS**------- value in neural network allows shifting the activation function to the left or right, which may be critical for successful learning. [1]

Q11. NOT-**XOR/XNOR**------ binary gate is linearly inseparable, [1]

Q12. The **PReLU-**--------activation function learns parameter that control the leaky-ness of the function. [1]

Q13. With a purelin node function e^{-x} is to be learnt using Delta learning. When input is 0.1,net = wx+b =-0.1. learning rate is 0.5.

Change in weight= **0.05024=** $\eta(t-o)f'(net).x = 0.5(.9048-(-0.1))(1)(0.1)$

[2]

Q14. BPA is used to predict probability of heart disease related to case A,B &C. ANN has 1 hidden layer with one hidden neuron. Activation function at output layer is Relu and at hidden layer is tanh, all the weights from hidden to output layers are unity. Bias for the hidden layer is also unity. *F*or a patient based on his tests, the expected probabilities of heart disease due to case A, B, C are 1.0,0,0 respectively. The output of the ANN network obtained for case A, B & C are 0.8,0.2, -0.1 respectively. Output of hidden node is 0.1. learning rate is unity.

i. Error vector at output layer =
ii. Error vector at hidden layer=
iii. $\Delta W(H\text{-}O_A)$=---
iv. $\Delta W(H\text{-}O_B)$=----
v. Δb=---- [3+2+1+1+1=8]

In this problem ReLu output should have been [0.8,0.2,0.1] , as ReLu output is either positive or zero, actually in my mind I was giving net of output nodes so output would have been [0.8,0.2,0], so I have given marks even if you took, last component of output as -0.1/0.1/0.

(i)$[e_{O\text{-}A}, e_{O\text{-}B}, e_{O\text{-}C}$ = **[0.2(1-.8),-0.2(0-0.2),±0.1/0]**

(ii) $e_H = (1-.1^2)[e_{O\text{-}A} + e_{O\text{-}B} + e_{O\text{-}C}]$ **=±0.099/0**

(iii) **0.02**

(iv) **-0.02**

(v) **±0.099/0**

Q15. -**AUTOENCODER**---------- is an artificial neural network that takes an input and encode it using some function and then decode it to get the output

identical to the input, with main objective to minimize the reconstruction error between the input and the output. [1]

Q16. Predicted values for class A,B,C are 0.6,0.7, 0.4 respectively, while target values for class A,B,C are 1,0,1 respectively.[use log not ln]

log loss for class B =**0.5228**= -log(1-.7)

log loss for class C=**0.397**=-log(0.4)

[1+1=2]

Q17. [1,1,1,-1] is stored in auto-associative network with no self-connection. Activation function is Bipolar with threshold of 0. What vector will be outputted if first two elements of the vector are

(i) missed [1 1 1 -1] **(1 1 1 -1)**

(ii) mistaken [1 1 -1 1] **(1 1 -1 1)** [2+2=4]

$$W = \begin{bmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

$[0\ 0\ 1\ -1]W = [2\ 2\ 1\ -1] - - > [1\ 1\ 1\ -1]$

$[-1\ -1\ 1\ -1]W = [1\ 1\ -1\ 1] - - > [1\ 1\ -1\ 1]$

Q18. Train a perceptron to learn AND gate for bipolar inputs & outputs .Initial weights and bias is zero, learning rate is unity,[activation function : f(net) =1 when net>0, f(net)=0 when net=0, f(net)=-1 when net <0]. Use the perceptron learning rule which changes weight proportional to target value, when X1=1,X2 =-1 is presented to it.

(i) Δw1=**-1**-----

(ii) Δw2=**1**--

(iii) Δb=**-1**--- [3]

net=0, y=f(net)=0, t=-1

w1=0+(1)(-1)(1)=-1

w2=0+(1)(-1)(-1)=1

b=0+(1)(-1)(1)=-1

Q19. Certain disease is relatively rare with a prevalence of 0.1% in the population of one lakh on which tests were conducted. The test has a reported sensitivity of 80%. In addition, the test comes back positive 10% of the time and negative 90% of the time. Provide number against TP,TN,FP,FN and % against Pr(Disease/test positive).

(i) TP =--**80**--
(ii) TN= **89,980**---
(iii) FP= **9920**------
(iv) FN=**20**-----
(v) Pr(Disease/Test positive)= **0.8%** [5]

Prevalence 0.1 %,i.e. 100 out of 1 lakh positive cases, 80% sensitivity means 80 are TP, 20 FN. Total positive tests are 10%, i.e. 10,000, TP is 80 so FP =10,000-80=9920, total negative 90% i.e 90,000, FN=20, TN=10,000-20=89,980

Q20. Output of network corresponding to Red, Blue,Green node is [2,1-3] is given to softmax function. What would be the output of softmax function indicating probability of color being Blue? [2]

$\text{Probability}_{\text{Blue}}$ =--e^{1} /(e^{2} + e^{1} + e^{-3})=**0.2676**--

Q21. -**SWISH**-------- activation function [f(x)=x·sigmoid(βx)] is reported to improve top-1 classification accuracy on ImageNet by 0.9% for Mobile NASNet-A and 0.6\% for Inception-ResNet-v2