Test Information

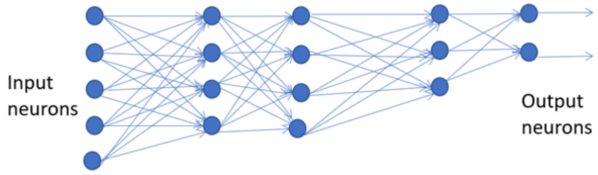
Description	After you completed the assignment, you may get a mark (probably 0), it is a dummy result and not the real result because the true answers have not been input to the systems yet. The real result will be published 2 weeks or more after the assignment deadline.
Instructions	Hard deadline: late submission is not allowed. Please click on the "save and submit" button at the end of the page after you completed the assignment. If your answer is not an integer, give it to the nearest 3 decimal places. If your answer is an integer, add decimal point has no harm. E.g. 3=3.0=3.00=3.000 etc. Please complete the assignment before the deadline. Multiple attempts are allowed.
Multiple Attempts	This test allows multiple attempts.
Force Completion	This test can be saved and resumed later.
	Your answers are saved automatically.

Question Completion Status:

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

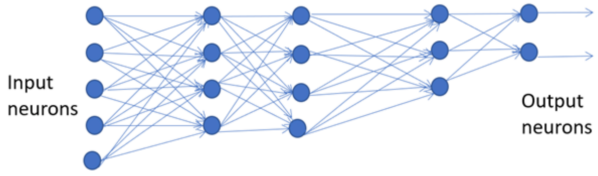
QUESTION 1 — **10 points** Save Answer

A fully connected neural network is shown below.
How many weights are in this network?



QUESTION 2 — **10 points** Save Answer

A fully connected neural network is shown below.
How many biases are in the network?



QUESTION 3 — **10 points** Save Answer

Boosting 1: An Ada-Boost algorithm based on parallel axis weak classifiers is used to build a strong classifier. There are 2029 samples, in the training data set, each sample is 2-dimension, 40% are positive (+1) examples and the rest are negative (-1) examples (round off if necessary).
At step t=1, a weak classifier hs( ) which gives the least error is found, using hs( ) 200 samples are incorrectly classified. Calculate the weight (α at time t) of this weak classifier in building the final strong classifier.

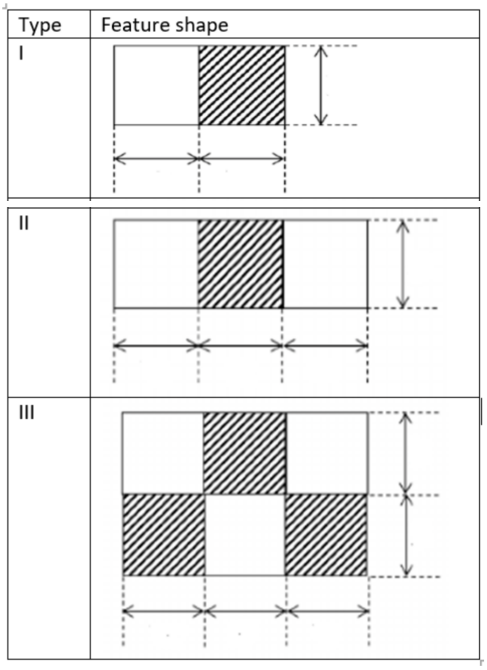
QUESTION 4 — **10 points** Save Answer

An Ada-Boost algorithm based on parallel axis weak classifiers is used to build a strong classifier. There are 2000 samples in the training data set.
At step t=1, a weak classifier hs( ) which gives the least error is selected, using hs( ) 300 samples are incorrectly classified.
Assume D=normalized incorrectly classified weight.
At time t=2, all normalized weights D (time t=2) for incorrectly classified samples should be equal. Find the value of D (time t=2).

QUESTION 5

10 points Save Answer

In an object detection system using a detection window with size width=48 pixels, and height = 24 pixels. There are 3 basic feature types as shown below. The feature value is equal to (Sum of pixels in shaded area)- (Sum of pixels in white area). The numbers of different features found in the detection window for type I (one row, each row has two rectangles) , II (one row, each row has three rectangles), III (two rows, each row has three rectangles) are N1,N2,N3 respectively. Find N1+N2+N3.



QUESTION 6

15 points Save Answer

For a neural network shown in the diagram.
Weights from input to the first hidden layer are called W.

Weights from the hidden layer to the output layer are called OW.
Biases are called B.

The activation function for neurons is Sigmoid. The parameters are:
X1=0.3, X2=0.2, X3=0.1, X4=0.4, X5=0.5
W1=0.21, W2=0.22, W3=0.31, W4=0.42, W5=0.32,
W6=0.32, W7=0.41, W8=0.42, W9=0.42, W10=0.52
OW1=0.3, OW2=0.2, OW3=0.3, OW4=0.14
B1=0.21, B2=0.3, B3=0.25, B4=0.5
Find Y1+Y2.

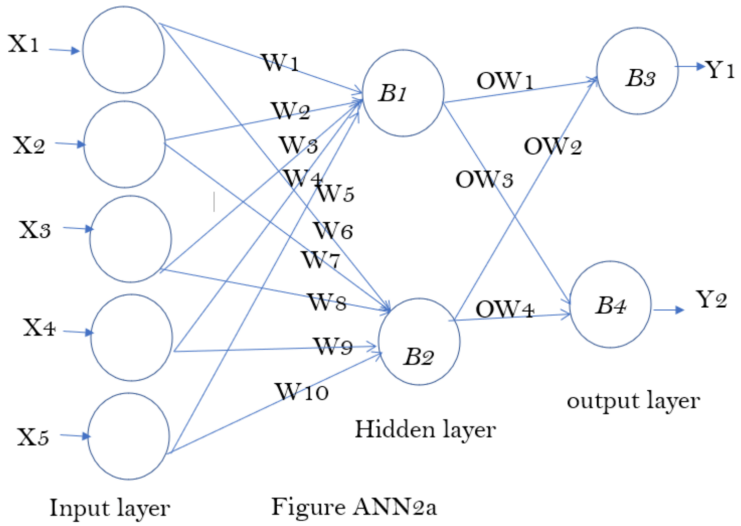


Figure ANN2a

QUESTION 7

15 points Save Answer

In the following diagram, it shows the parameters of a part of a neural network at time k. The activation function of the neurons is sigmoid. The energy to be minimized during training is E = (1/2)*(y-t)^2. such that new_w = old_w + dw, where

dw = -learning_rate*(∂E/∂w), and dw=delta_weight

Assume all the weights will be updated together only after all delta weights (dw) have been calculated for each epoch time k. So, use the current (time k) parameters for this calculation.

The current parameters at time k are:

learning_rate=0.6,
x1=0.5, x2=0.8, x3=0.7,
wh1=0.24, wh2=0.43, wh3=0.52,
ow11=0.12, ow12=0.42, ow13=0.1,
bh1=0.33, bh2=0.21, bh3 = 0.15,
s1=0.3, s2=0.2, s3=0.1 %are sensitivities

Find new wh1 at time k+1.

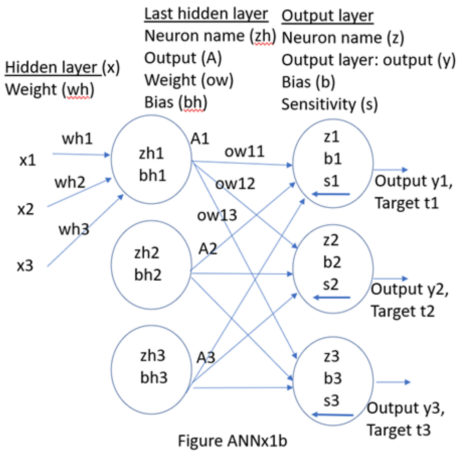


Figure ANNx1b

QUESTION 8

10 points Save Answer

A=[11, 22
23, 18];
B=[4,6
9,8];
C= conv2(A,B) %conv2 is 2D convolution

For the convolution function, all overlapped and non-overlapped cases are included.

Find the sum of all elements of C.

QUESTION 9

10 points Save Answer

The resolution of an input image is 97x97. A Convolution Neural Network CNN uses a kernel = 5x5, step size = 3, zero padding = p to generate a feature map of size DxD. Note: D must be an integer.

Find the value of D+p.

QUESTION 10

10 points Save Answer

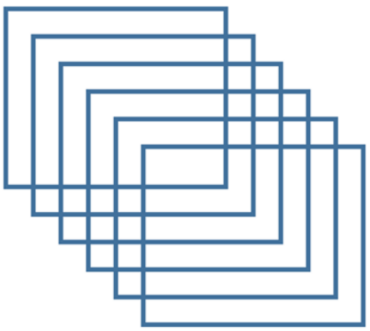
The input layer to the first convolution layer feature maps (C1) of a convolution neural network CNN is shown in the attached figure.

Input = 84 x 84 pixels
kernel size=5x5
Step size=1

For this part of the neural network, the number of weights is W and the number of biases required is B.

Find W+B.


Input is NxN


6 feature maps in C1

QUESTION 11 — 10 points — Save Answer

In a system, the performance is as follows.
True positives (TP) =100
True Negatives (TN)=60
False Positives (FP) =51
False Negatives (FN)=32
If the accuracy is A, Precision is P, Recall is R. Find A+P+R.

QUESTION 12 — 10 points — Save Answer

There are 200 apples in the picture. The system can extract 150 apples, of which 20 are incorrect. Calculate the precision of the system.

QUESTION 13 — 10 points — Save Answer

To train a neural network you are given 60000 training samples and 10000 testing samples.

Using training scheme 1 of applying the full batch method, the total iterations for 10 epochs is X1.

Using another training scheme 2 of applying the mini-batch method, the mini-batch size is 500, the total iterations for 10 epochs is X2.

Find X1+X2.

QUESTION 14 — 15 points — Save Answer

In this assignment, we use the notations as given in the lecture notes. You may calculate the result with the help of a calculator or computer.
A set of training data (X) and their classes (Y) as shown below.

X=[u v] is the coordinates of the sample.

X has 2 classes Y: "pos" and "neg"

```
neg=[16   38
     2    54
     32   4
     42   10
     30  42];
pos=[22   39
     4    33
     -22  -25
     -37  -31
     -23  -48];
```

Just after the calculation of step t =2, the highest normalized weight D is D_high, the lowest normalized weight D is D_low. find D_high - D_low.

QUESTION 15 — 20 points — Save Answer

Run the following code in COLAB

https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/tutorials/keras/overfit_and_underfit.ipynb

and answer the following questions. Partial credit will be given for correct answers.

References :

https://colab.research.google.com/ (how to use colab)

https://www.tensorflow.org/datasets/catalog/higgs

Which of the following statements are true?

- [ ] Each sample in the Higgs data set has 28 features.
- [ ] The first feature of a sample is the class label.
- [ ] BUFFER_SIZE+ BATCH_SIZE+ N_TRAIN +BUFFER_SIZE+BATCH_SIZE + STEPS_PER_EPOCH=32000.
- [ ] In the "tiny" model, the number of biases=17.
- [ ] During training the learning rate is kept stable.
- [ ] Without weight regularization, the "tiny" model can avoid overfitting compared to other models under the same testing environment.
- [ ] L2 regularization adds the absolute value of the weight coefficients, while
  L1 regularization adds the square of the value of the weight coefficients to the total loss during weight update.
- [ ] L1 regularization performs better to reduce overfitting than L2 regularization.
- [ ] The "dropout" regularization method can improve overfitting.
- [ ] In this experiment the best performance in term of reducing overfitting is Combined L2 + dropout.

QUESTION 16 — 15 points — Save Answer

Run the CNN system found in this link

https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/tutorials/images/cnn.ipynb

Based on this CNN system, after the step "Add Dense layers on top" is executed, if the total number of weights used is W and biases is B. Find W + B of this network.

Hints: Please read these documents for more information

tensorflow - what is the default activation function of dense layer in keras - Stack Overflow/

machine learning - What is the role of "Flatten" in Keras? - Stack Overflow

QUESTION 17

15 points Save Answer

In transfer learning the processing steps are:

1. Obtain the pre-trained model
2. Create a base model
3. Freeze layers
4. Add new trainable layers
5. Train the new layers on the dataset
6. Improve the model via fine-tuning

Study the process of transfer learning by running the demo code at

https://colab.research.google.com/github/keras-team/keras-io/blob/master/guides/ipynb/transfer_learning.ipynb

Just after the step “Build a model”.

Total number of parameters = x1

Number of trainable parameters. = x2

Just after the step “Do a round of fine-tuning of the entire model”

Total number of parameters = x3

Number of trainable parameters. = x4

Note: In the final tuning step there are non-trainable parameters which are used for data normalization.

Find x1+x2+x3+x4.