Quiz 1

Email *

Bagging is an ensemble technique that: *

- Uses a committee of experts to make predictions
- Trains multiple models on different subsets of the data ✗
- Combines predictions using a weighted average
- Constructs an ensemble by iteratively updating weights

Which of the following is/are Limitations of deep learning? *

- None of the above
- Obtain huge training datasets
- Data labeling
- Both Data labeling and Obtain huge Training datasets ✗

What happens when you increase the regularization hyperparameter lambda?

the model will perform better as it will reduce the overfitting of the model and the weights will become smaller

Describe Early Stopping *

regularization method used to avoid overfitting, it stops training when it find that there is no difference when we update the parameters and weights, train parameters with lowest validation error

Name *

Ensemble method combines multiple models to form a better model.

- True ✗
- False

In Bagging, the variance of the model can be reduced by averaging *

- True ✗
- False

is a kind of function that tells you how good the prediction of a network is. cross-entropy *

- weight decay
- Loss function ✗
- Softmax
- Softmax Loss

bagging is more sensitive to noise *

- No ✗
- Maybe
- Yes

Select the reason(s) for using a Deep Neural Network *

- Some patterns are very complex and can't be deciphered precisely by alternate means
- We finally have the technology - GPUs - to accelerate the training process by several folds of magnitude
- All of the above ✗

Deep Nets are great at recognizing patterns and using them as building blocks in deciphering inputs

What is TRUE about the functions of a Multi Layer Perceptron? *

- It predicts which group a given set of inputs falls into.
- It generates a score that determines the confidence level of the prediction.
- All of above. ✗

The first neural nets that were born out of the need to address the inaccuracy of an early classifier, the perceptron.

the weights may be reduced to zero here *

- L1 and L2
- L1 ✗
- None of the Above
- L2

Which of these techniques are useful for reducing variance (reducing overfitting)? *

- Dropout ✗
- L2 regularization ✗
- Data augmentation ✗

Which of the following would have a constant input in each epoch of training a Deep Learning model? *

- Weight between input and hidden layer ✗
- Biases of all hidden layer neurons
- Weight between hidden and output layer
- Activation function of output layer

What is weight decay? *

regularization technique adds a small penalty to the loss function to make the model generalize better

If we want to create an optimal set of weights, we choose to minimize this loss function concerning w over the entire training data set.

- True ✗
- False

Bagging reduces the variance of the final model with increasing the bias.

- True
- False ✗

In which of the following applications can we use deep learning to solve the problem? *

- All of the above ✗
- Protein structure prediction
- Detection of exotic particles
- Prediction of chemical reactions

ID *

The purpose of using ensemble learning is to: *

- Increase training time and complexity
- Eliminate the need for labeled data
- Reduce overfitting and improve generalization ✗
- Decrease the number of models required

Which of the following functions can be used as an activation function in the output layer if we wish to predict the probabilities of n classes (y^1, y^2, y^n) such that sum of p over all n equals to 1? *

- Tanh
- ReLu
- Softmax ✗
- Sigmoid

Why is the vanishing gradient a problem? *

- The gradient is calculated multiplying two numbers between 0 and 1
- With backprop, the gradient becomes smaller as it works back through the net
- All of above. ✗

Training is quick if the gradient is large and slow if its small

that balance the right amount of bias and variance that lead to the minimal error. *

- optimal model complexity ✗
- Fit
- Other