Machine Learning worksheet- 4 Solution

1. A
2. A
3. A
4. A
5. C
6. B
7. B
8. B
9. Entropy is the measure of impurity as given by (for binary class):

Entropy= -p log2p – q log2q therefore :

-((40/100)log2(40/100) + (60/100)log2(60/100))= 0.96

And Gini is given as Gini=1- ((P+)2 +(P-)2 ) therefore,

1-((40/100)2 + (60/100)2)=0.48

1. Random forests overcome several problems with decision trees, including:

* Reduction in overfitting: by averaging several trees, there is a significantly lower risk of overfitting.
* Less variance: By using multiple trees, you reduce the chance of stumbling across a classifier that doesn’t perform well because of the relationship between the train and test data.

1. After using various techniques to clean and process the data,we need to scale the data on the same scale for our algorithm. And, when all features are in the same scale, it also helps algorithms to understand the relative relationship better.

Some of the techniques we used for scaling is:

* Standard Scaler
* MinMax Scaler

1. **Gradient Descent** is the most common optimization algorithm in machine learninganddeep learning*.* It is a first-order optimization algorithm. This means it only takes into account the first derivative when performing the updates on the parameters.

**We have to make** sure to scale the data if it’s on a very different scales. If we don’t scale the data, the level curves (contours) would be narrower and taller which means it would take longer time to converge.

1. In highly imbalanced dataset,accuracy is not the good metrics to measure the performance. Accuracy can be misleading. Sometimes it may be desirable to select a model with a lower accuracy because it has a greater predictive power on the problem.

For example, in a problem where there is a large class imbalance, a model can predict the value of the majority class for all predictions and achieve a high classification accuracy

1. F1-score is the harmonic mean of Precision and Recall and gives a better measure of the incorrectly classified cases than the Accuracy Metric.

F1-Score= 2 \* (Precision \* Recall)/(Precision + Recall)

1. "**fit"** computes the mean and std to be used for later scaling. (just a computation), nothing is given to user. **"transform"** uses a previously computed mean and std to autoscale the data (subtract mean from all values and then divide it by std). **"fit\_transform"** does both at the same time.