**MachineLearning(worksheet-4)**

1. (A)
2. (A)
3. (B)
4. (A)
5. (A)
6. (A)
7. (B)
8. (C)
9. (entropy- 0.971,gini index-0.4,0.6)

Random forests consist of multiple single trees each based on a random sample of the training data. They are more accurate than single decision trees. the decision boundary becomes more accurate and stable as more trees are added. While a single decision tree is often pruned, a random forest tree is fully grown and unpruned, and so, naturally, the feature space is split into more and smaller regions. Each random forest tree is learned on a random sample, and at each node, a random set of features are considered for splitting. Both mechanisms create diversity among the trees.

11. scaling is needed to bring every feature in the same precedence without any upfront importance. scaling is applied is that few algorithms like Neural network gradient descent **converge much faster** with feature scaling than without it.

Normalization and Standardization.

12. The gradient descent algorithm[3] is the core algorithm in linear regression. The gradient algorithm is used to feature the data before the iterative calculation to accelerate the convergence speed of the algorithm. There are usually two methods for feature scaling: normalization methods and interval scaling.

(1) In performing linear regression calculations, it is usually necessary to normalize the feature data to speed up the iteration. (2) When the feature data value is large, the normalization method can be used to obtain better convergence.

(3) When the feature data value is small, the normalization method can be selected according to the specific situation.

(4) Feature scaling can make machine learning algorithm work better and different machine learning algorithms can accept different range of input values

13. *The main problem of imbalanced data sets lies on the fact that they are often associated with a user preference bias towards the performance on cases that are poorly represented in the available data sample.* .imbalanced classification problems typically rate classification errors with the minority class as more important than those with the majority class. As such performance metrics may be needed that focus on the minority class, which is made challenging because it is the minority class where we lack observations required to train an effective model

14. the F-score, also called the F1-score, is a measure of a model’s accuracy on dataset. It is used to evaluate binary classification systems, which [classify](https://deepai.org/machine-learning-glossary-and-terms/classifier) examples into ‘positive’ or ‘negative’.

The F-score is a way of combining the [precision and recall](https://deepai.org/machine-learning-glossary-and-terms/precision-and-recall) of the model, and it is defined as the [harmonic mean](https://deepai.org/machine-learning-glossary-and-terms/harmonic-mean) of the model’s precision and recall.

Fscore=tp/(tp+½(fp+fn)

15. fit() just calculates the parameters (e.g. μ*μ* and σ*σ* in case of [StandardScaler](http://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html)) and saves them as an internal object's state.

transform() method to apply the transformation to any particular set of examples.

fit\_transform() joins these two steps and is used for the initial fitting of parameters on the training set x*x*, while also returning the transformed x′*x′*. Internally, the transformer object just [calls first fit() and then transform()](https://github.com/scikit-learn/scikit-learn/blob/6c3e17989a7d80c34f124365f2c436a3fdcb1497/sklearn/base.py#L659-L690) on the same data.