Machine learning worksheet – 4

1ans) both gridsearchcv and randomsearchcv are used hyperparameter tunning in sklearn library

2ans) A) Random forest

3ans) B) Regularization decreases

4ans) C) A&B

5ans) C) In case of classification problem, the prediction is made by taking mode of the class labels predicted by the component trees

6ans) A) Gradient Descent algorithm can diverge from the optimal solution

7ans) D) both bias, variance decreases

8ans) B) overfitting model

10ans) randomforests are a strong modeling technique and much more robust than a single decisiontree. They aggregate many decisiontrees to limit overfitting as well as error due to bias and therefore yield useful results.

11ans) Feature Scaling or Standardization: It is a step of Data Pre-Processing which is applied to independent variables or features of data. It basically helps to normalize the data within a particular range. Sometimes, it also helps in speeding up the calculations in an algorithm, for numerical data it has done works on numbers and does not know what that number represents, types of scaling techniques are standard scaler, minimax scaler

13ans) for imbalanced dataset of classification we should prefer sensitivity-specificity and precision-recall metrics, because it doesn't give the correct prediction of the data so we shouldn’t prefer accuracy metrics

14ans) The F-score, also called the F1-score, is a measure of a model’s accuracy on a 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 it 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 **F**-**Measure** = (2 \* Precision \* Recall) / (Precision + Recall)

15ans) The fit() function calculates the values of these parameters. The transform function applies the values of the parameters on the actual data and gives the normalized value. The fit transform() function performs both in the same step