MACHINE LEARNING

1) In my view Residual Sum of Squares (RSS) is s a better measure of goodness of fit model in regression, because takes account into all variables.
2) total sum of squares (TSS) = explained sum of squares (ESS) + residual sum of squares (RSS). TSS measures variation in values of observed variable with respect to sample mean, ESS calculates deviation between estimated value and mean value of observed variable, whereas RSS is used to measure the variance in data
3) Regularization in machine learning is calibrate machine learning models in order to minimize adjusted loss function and prevent overfitting or underfitting.
4) Gini impurity is used to calculate amount of probability of a specific feature that is classified incorrectly when we select randomly.
5) Yes, decision-trees prone to overfitting because complexity and dimensionality of the decision tree model.
6) Ensemble techniques will help in improving accuracy of results in models by combining multiple models.
7) bagging gets ensemble model with less variance whereas boosting try to produce strong models less biased.
8) out-of-bag error is average error for each calculated using predictions from trees,I.e number of wrongly classifying.
9) K-fold Cross-Validation is when dataset is split into K number of folds and is used to evaluate the models ability
10) Hyperparameter tuning consists of finding a set of best hyperparameter values for learning algorithm while applying this algorithm to any data.
11) Because of large learning rate will make learning jump over minima.
12) No, we don't use because target label has no linear correlation with features.
13) Both are ensemble techniques, Ada Boost starts with building short tree whereas Gradient Boost starts making a single leaf.
14) the bias–variance trade off in machine learning is the variance of the parameter estimated across samples can be reduced by increasing bias in estimated parameters.
15) RBF is to find non-linear classifier or regression line, Linear Kernel is used when the data is Linearly separable, and polynomial kernel to map the data into higher-dimensional space