Random Forest

THEORY

1) It is a machine learning algorithm under supervised learning which can be used for both regression and classification tasks. It takes the prediction from each decision tree and outputs the prediction based on the majority votes by the decision trees.
2) Ensemble means combining multiple machine learning models. It is of two types:
   a) Bagging or Bootstrap Aggregation - Random Forest
   b) Boosting - ADA BOOST, XG BOOST
3) The step of row sampling with replacement is known as bootstrap & the step which involves combining all the results and generating output based on majority is known as aggregation.
4) Hyperparameters which increase the accuracy of the model are: n_estimators, max_features, mini_sample_leaf
5) Hyperparameters which increase the speed of the model are: n_jobs, random_state, oob_score

QUIZ

1) Which hyperparameter control the randomness of the sample?
2) Which hyperparameter decides how many processors are allowed to use?
3) Which hyperparameter stands for the number of decision trees the algorithm builds to output the predicition?

ANSWER

1) random state
2) n jobs
3) n estimators