**True/false**

1. Decreasing the depth of your decision tree (through pruning) will reduce test error. - True

 Pruning helps prevent overfitting by reducing the complexity of the tree, making it generalize better to unseen data.

1. In k-fold cross validation you leave k inputs out, train your classifier on the remaining n−k inputs and evaluate it on the leave-out inputs. You do this repeatedly and average to obtain a good estimate of your classifier’s performance.- false

 In k-fold cross-validation, the data is split into k subsets. Each subset is used as a test set once while the remaining k-1 subsets are used for training. This process is repeated k times, not leaving k inputs out.

1. Gradient Boosting is performing (stage-wise) gradient descent in function space.- True

 Gradient boosting builds models sequentially, each new model correcting errors of the previous ones by performing gradient descent in the space of functions.

1. When you split your data into train and test you have to make sure you always do the splitting uniformly at random. - True

 Uniform random splitting ensures that both the training and test sets are representative of the overall data distribution, reducing bias.

1. If a classifier obtains 0% training error it cannot have 100% testing error.- false

 A 0% training error implies perfect fit on the training data, making a 100% test error unlikely unless the test data is entirely different, which usually indicates a problem like data leakage or extreme overfitting.

1. Increasing regularization tends to reduce the bias of your classifier. - false

 Increasing regularization reduces variance but increases bias, as it simplifies the model, potentially underfitting the data.

1. If run without a depth limit, the ID3 algorithm returns the maximally compact decision tree that is consistent with a data set (if it exists). - false

 Without a depth limit, ID3 will create a very large tree that perfectly fits the training data, which may not be compact and can lead to overfitting.

1. The best classifiers make no assumptions about your data at all. - false

 Every classifier makes some assumptions about the data, known as inductive bias, which helps it generalize from training data to unseen data.

1. Random Forests learn many high variance CART trees and reduce this variance by averaging the results. That’s basically Bagging applied to (slightly modified) CART trees. - true

 Random forests use bagging to combine multiple CART trees, each trained on a random subset of the data and features, reducing overall variance.

1. As your training data set size, n, approaches infinity, the k−nearest neighbor classifier is guaranteed to have an error no worse than twice the Bayes optimal error. - true

 In theory, with an infinite amount of data, the error of k-nearest neighbor approaches a value that is at most twice the Bayes error, which is the best possible error for any classifier.

1. Squared loss regression trees require a time complexity O(n^2) per split. - false

 The time complexity for finding the best split in squared loss regression trees is typically O(n log n) per split, not O(n^2).

1. The Bayes optimal error is the best classification error you could get if there was no noise. - false

 The Bayes optimal error is the lowest possible error considering the inherent noise in the data, not assuming no noise.

Sorry for incomplete assignment. I just forgot the theoretical assignment and did this in the last hours.