**CIS 3715 Spring 2020: Quiz 6**  STUDENT NAME: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. (1 point) What is “overfitting”?
2. (2 points) Name two approaches that help us build decision trees that do not overfit?
3. (1 points) How is “Mean Squared Error” defined and what is it used for?
4. (1 points) What is the “decision boundary” for a linear predictor f(x) = 3 – x1 + 2x2?
5. (1 points) What is the purpose of “cross-validation”?
6. (1 points) What does “bagging of decision trees” refer to? Why is it expected to be better than just a “decision tree”?
7. (1 points) Give an example of a nonlinear predictor.
8. (2 points) Explain what does each line of the code below do:

k = 7

predictor = neighbors.KNeighborsClassifier(n\_neighbors = k)

predictor.fit(X\_train, y\_train);

result = predictor.score(X\_test,y\_test)