Questions 1-8 carry 1 point each, 9-11 2 points each, and 12 carries 6 points.

1. T/F For a PLA, the target function is always in the range of \{-1,1\}, or equivalently \{0,1\} or some other Boolean.

2. T/F For a machine learning problem, the target function I/O specification depends on the algorithm selected.

3. T/F A learning step in a PLA can change the values in a training data input.

4. A hypothesis in a PLA looks like
   a. some_function1(temp,humidity) = sign(w0*x0+w1*temp+w2*humidity)
   b. <temp=60,humidity=40,day='thurs',num_players=10>
   c. some_function2(temp,humidity) returns num_players
   d. all of the above

5. T/F In practical machine learning projects, some of the input training data can sometimes be set aside for running a test.

6. T/F In practical machine learning projects, the input training data may be preprocessed to correct for statistical noise, for removing outliers, etc.

7. In about 2 or so lines, describe the purpose of a test data set in machine learning.

8. The larger a training data set (check all that apply)
   a. the more likely it is that we can get a hypothesis that is a good approximation of the target function in reality.
   b. the more likely there is a risk of overfitting
   c. the more likely it is that the error in the target function will be high
   d. the longer it takes the machine learning algorithm to run in the field, or “recall” phase.

9. In a MLP, the individual neurons each implements (choose one)
   a. an activation function that generates a Boolean value from a set of inputs
   b. an activation function of the form \( y = w_0x_0 + w_1x_1 + w_2x_2 \)
   c. an activation function with input as \( y \) above, and an output that is a sigmoid function on that input.
   d. none of the above

10. The significance of using a neuron with a sigmoid activation function in a MLP is (choose 1)
a. the sigmoid can be differentiated and retains the same form after
differentiation
b. The sigmoid function is discontinuous and this is very important.
c. there is no significance, it is a random choice.

11. Back propagation in a MLP (check all that apply)
a. is a way to compute the overall error in the MLP
b. is a technique to generate changes to weights in the hidden layer
c. is a technique to minimize the error of each layer separately, starting from
the output layer and proceeding inwards.

12. Use Bayes’ rule to solve the following problem: At a party you meet a person who
claims to have been to the same school as you. You vaguely recognize them, but
cannot remember properly, so decide to work out how likely it is given that:
* 1 in 2 of the people you vaguely recognize went to school with you
* 1 in 10 of the people at the party went to school with you
* 1 in 5 people at the party you vaguely recognize.