Problem 1: Recall (2pts) (Answer in one sentence only.)
What is boosting?
Solution:
Boosting refers to effective methods of producing accurate classifier by combining the outcome from moderately accurate classifiers.

Problem 2: Work (8 pts) (Show all derivations/work and explain.)
In AdaBoost we combine \( M \) ‘weak leaners’ to get an accurate classifier. During training, we need to select the data distribution wisely so that the classifier at each round is focusing on different sub-problems. According to AdaBoost, how should we choose the data distribution at each round?
Solution:
The data weight distribution is uniform initially. In the following iterations, for data that are misclassified, the weight will be increased; and for the part of data that are correctly classified, the weight will be decreased.

How to combine the \( M \) classifier output to get the final prediction?
Solution:
The decision is a weighted combination of the classifier output, where the weight for each classifier is determined based on the weighted error. We decide the class based on the sign of the combined output.

Now consider that we have a multi-class problem with lots of classes (\( \geq 100 \) classes), is it a good idea to directly apply AdaBoost on multi-class learners?
Solution:
Directly apply AdaBoost on multi-class problem may not be a good idea, since it assumes each learner have error less than 0.5. It is very hard to get classifier that can achieve error less than 0.5 when there are a lot of classes.