Q1-1: Which of the following is/are correct regarding benefits of ensemble model?

A. Better performance

B. Generalized models

C. Better interpretability

1. A, C
2. B, C
3. A, B
4. A, B, C

Q1-1: Which of the following is/are correct regarding benefits of ensemble model?

A. Better performance

B. Generalized models

C. Better interpretability

1. A, C
2. B, C
3. A, B 
4. A, B, C

Interpretability is lost when ensemble model is used.

Q1-2: Which of the following statement(s) is/are correct about Bagging?

A. In bagging, we choose random subsamples of the input points with replacement

B. The main purpose of bagging is to decrease the bias of learning algorithms.

C. Bagging is ineffective with logistic regression, because all of the learners learn exactly the same decision boundary

1. A
2. B, C
3. A, C
4. A, B, C

Q1-2: Which of the following statement(s) is/are correct about Bagging?

A. In bagging, we choose random subsamples of the input points with replacement

B. The main purpose of bagging is to decrease the bias of learning algorithms.

C. Bagging is ineffective with logistic regression, because all of the learners learn exactly the same decision boundary

1. A 
2. B, C
3. A, C
4. A, B, C

The main purpose of bagging is to decrease the **variance** of learning algorithms.

Q2-1: Select the correct statement about AdaBoost.

A. *In AdaBoost weights of the misclassified examples go up by the same multiplicative factor.*
B. *In AdaBoost, if the weighted training error ϵ_t of the t^{th} weak classifier is large, then this classifier has smaller weights in the weighted majority vote for prediction*

1. Both the statements are TRUE.
2. Statement A is TRUE, but statement B is FALSE.
3. Statement A is FALSE, but statement B is TRUE.
4. Both the statements are FALSE.

Q2-1: Select the correct statement about AdaBoost.

A. *In AdaBoost weights of the misclassified examples go up by the same multiplicative factor.*

B. *In AdaBoost, if the weighted training error ϵ_t of the t^{th} weak classifier is large, then this classifier has smaller weights in the weighted majority vote for prediction*

1. Both the statements are TRUE. 
2. Statement A is TRUE, but statement B is FALSE.
3. Statement A is FALSE, but statement B is TRUE.
4. Both the statements are FALSE.

Statement A = True, follows from the update equation.

Statement B = True. See the ensemble prediction formula.

Q2-2: You've just finished training a random forest for spam classification, and it is getting abnormally bad performance on your validation set, but good performance on your training set. What could be causing the problem?

A. *Your decision tree is too deep*
B. *You have too few trees in your ensemble*
C. *Your bagging implementation is randomly sampling sample points without replacement*
D. *You are randomly sampling too many features when you choose a split*

1. A, B, C
2. B, C, D
3. A, C, D
4. A, B, C, D

Q2-2: You've just finished training a random forest for spam classification, and it is getting abnormally bad performance on your validation set, but good performance on your training set. What could be causing the problem?

A. Your decision tree is too deep

B. You have too few trees in your ensemble

C. Your bagging implementation is randomly sampling sample points without replacement

D. You are randomly sampling too many features when you choose a split

1. A, B, C
2. B, C, D
3. A, C, D
4. A, B, C, D 

We observe overfitting.

A: large models can cause overfitting
B: too few trees (extreme case: one tree) will not reduce the variance, and thus doesn't prevent overfit
C: if without replacement, the random subsampled datasets = the original training dataset, so it's not doing bagging at all
D: if too many features, then the tree models have large capacity and can overfit

Q3-1: Suppose the minimum Hamming distance between any pair of codewords is 8. For which values of *x* can we still get the right classification with *x* single-bit errors? Choose the maximum possible value.

1. 1
2. 2
3. 3
4. 4

Q3-1: Suppose the minimum Hamming distance between any pair of codewords is 8. For which values of *x* can we still get the right classification with *x* single-bit errors? Choose the maximum possible value.

1. 1
2. 2
3. 3 
4. 4

d = 8, so this code can correct upto floor([8 - 1]/2) = floor(3.5) = 3 errors

Q3-2: Suppose there are 3 different classifier models with 70% accuracy. You make an ensemble model by using majority voting among these classifiers. Select the correct statement about this ensemble model.

1. *Maximum accuracy you can get = 100%.*
2. *Maximum accuracy you can get would be ≥ 70%, but cannot be 100%.*
3. *Maximum accuracy you can get would be < 70%.*
4. Can't comment on Maximum accuracy of the ensemble model.

Q3-2: Suppose there are 3 different classifier models with 70% accuracy. You make an ensemble model by using majority voting among these classifiers. Select the correct statement about this ensemble model.

1. *Maximum accuracy you can get = 100%.* 
2. *Maximum accuracy you can get would be ≥ 70%, but cannot be 100%.*
3. *Maximum accuracy you can get would be < 70%.*
4. Can't comment on Maximum accuracy of the ensemble model.

M1, M2, M3 are 3 classifiers (70% accuracy). E = output of the ensemble model. Y = Desired output. Consider the case shown, where the three classifiers make errors on different data points. As we can see, it is able to achieve 100% accuracy.

M1	M2	M3	E	Y
1	1	0	1	**1**
1	1	0	1	**1**
1	1	0	1	**1**
1	1	1	1	**1**
1	0	1	1	**1**
1	0	1	1	**1**
1	0	1	1	**1**
0	1	1	1	**1**
0	1	1	1	**1**
0	1	1	1	**1**