MACHINE LEARNING

1. Which of the following are disadvantages of using Hard Margin SVM classifier?

A) They allow misclassifications, that's why they are not optimal.
B) They cannot be used when the data is not completely linearly separable while allowing no errors.
C) They are not optimal to use in case of outliers.
D) None of the above.

Ans: B) They cannot be used when the data is not completely linearly separable while allowing no errors.

2. Which of the following statements are true regarding maximal margin classifier?

A) It is the most optimal classifier in a completely linearly separable data.
B) It's the classifier for which the margin length or the distance between the closest data-point on either side of the classifier and the classifier is maximized.
C) Any possible classifier which can linearly separate the data of two classes is called maximal margin classifier.
D) All of the above.

Ans: A) It is the most optimal classifier in a completely linearly separable data.
B) It's the classifier for which the margin length or the distance between the closest data-point on either side of the classifier and the classifier is maximized.

3. Which of the following statements are true regarding soft margin SVM classifier?

A) They are less sensitive to outliers and can be used even in their presence.
B) They make sure that there is no data point present in the margin area.
C) They allow some degree of errors or misclassification.
D) They can be used in case data is not completely linearly separable.

Ans: C) They allow some degree of errors or misclassification.
D) They can be used in case data is not completely linearly separable.

4. Which of the following statements are true regarding SVMs?

A) They take the data from lower dimensional space to some higher dimensional space in case the data is not likely to be linearly separable.
B) They use the kernel tricks to escape the complex computations required to transform the data.
C) If the data is not linearly separable SVM technique cannot be used.
D) All of the above.

Ans: B) They use the kernel tricks to escape the complex computations required to transform the data.
A) They take the data from lower dimensional space to some higher dimensional space in case the data is not likely to be linearly separable.

5. Which of the following Statements are true regarding the Kernel functions used in SVM?

A) These functions gives value of the dot product of pairs of data-points in the desired higher. dimensional space without even explicitly converting the whole data in to higher dimensional space.
B) We have to first convert the whole data in to the higher dimensional space before applying the kernel function.
C) The data product values given by the kernel functions are used to find the classifier in the higher dimensional space.
D) None of the above

Ans: A) These functions gives value of the dot product of pairs of data-points in the desired higher dimensional space without even explicitly converting the whole data in to higher dimensional space.
C) The data product values given by the kernel functions are used to find the classifier in the higher dimensional space.

6. How can SVM be classified?

A) It is a model trained using unsupervised learning. It can be used for classification and regression.
B) It is a model trained using unsupervised learning. It can be used for classification but not for regression
C) It is a model trained using supervised learning. It can be used for classification and regression.
D) It is a model trained using supervised learning. It can be used for classification not for regression.

Ans: C) It is a model trained using supervised learning. It can be used for classification and regression

7. The quality of an SVM model depends upon:

A) Selection of Kernel
B) Kernel Parameters
C) Soft Margin Parameter C
D) All of the above

Ans: D) All of the above

8. The SVM's are less effective when:

A) The data is linearly separable.
B) The data is clean and ready to use.
C) The data is noisy and contains overlapping points.
D) None of these

Ans: C) The data is noisy and contains overlapping points.

9. What would happen when you use very small C (C~0)?

A) Misclassification would happen.
B) Data will be correctly classified.
C) Can't say
D) None of these.

Ans: A) Misclassification would happen.

10. What do you mean by generalization error in terms of the SVM?

A) How far the hyperplane is from the support vectors.
B) How accurately the SVM can predict outcomes for unseen data.
C) The threshold amount of error in an SVM.
D) None of these

Ans: B) How accurately the SVM can predict outcomes for unseen data.