ASSIGNMENT - 1

Section - A:

1) T
2) T
3) T
4) T
5) T
6) F
7) F
8) T
9) T
10)

SEC B :

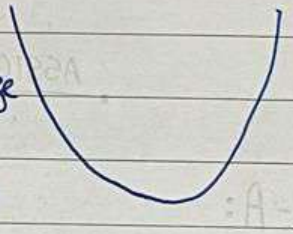
1

Model	Loss func^n	Regularizer
SVM		
LASSO	$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \lambda \sum_{i=1}^{m} \lvert w_i \rvert$	L1
RIDGE	$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \lambda \sum_{i=1}^{m} w_i^2$	L2

2)

3) a) M.S.E.

$w_{new} = w_{old} - \alpha \frac{\partial J}{\partial w}$; —— (1)



Let's say $\frac{\partial J}{\partial w} > 0$. Then to minimize loss, we need to decrease weight in order to reach the absolute minima of loss. (1) decreases w.

If $\frac{\partial J}{\partial w} < 0$, then we need to increase w. (1) increases w.

b) Absolute loss func^n.

SECTION C :

1) Reason for underfitting - Choosing inapproprite model. For eg - if we choose linear regression model for binary classification, the dist. b/w predicted output & label will be large, leading to high bias & underfitting.

2) If both training & test error remain high even after using ~~boost~~ suitable model, suitable loss func^n & applying other techniques like regularization to improve accuracy, it could mean that either data is insufficient in qty. or data is inconsistent.

3) Bagging reduces variance by ~~adding~~ giving that output which the highest no. of models vote for, thereby smoothening the effect of outliers & preventing instability.

4) Boosting has a reducing effect on both bias & variance. However, its major effect is reducing bias. It does so by using previous models to train new models, ultimately developing the optimum model. The instances misclassified by earlier model receive higher wt, & the instances correctly classified receives lower wt.

SEC E:

4) Two methods to avoid overfitting in decision trees-

① Bagging - By using the average result of a large no. of trees, overfitting can be prevented

② Regularization - By adding penalty for complexity, regularization prevents variance.

SEC - F

1) No, random forests can't use the same data for testing & training. This is because if the same data which had been used for training is used for ~~tra~~ testing then accuracy will be high whereas in reality our model's predictions may be off by large amts. when used on unseen data. ∴ the true purpose of testing is not met & we are only deceived.

2)

Bagging	Boosting
• Main aim is to decrease variance.	• Aim is to decrease bias
• Each model is given = weight	• Models are given weights acc to performance.

• Models are trained parallely i.e. independent of each other & classification is done by majority vote	• Previous models are used to train successive models, ultimately leading to optimum ~~modle~~ model.