Assignment 1

Bratadeep Sarkar - 240285 - EEG SPELLER

SECTION-A

1) True
2) True
3) True
4) True
5) True
6) False
7) False
8) True
9) True
10) True

SECTION B

1)

Model	Loss fn	Regularization
SVM	$\max(0, 1-y(x\cdot w+b))$	loss fn + $\lambda \sum_{i=1}^{n} w_i^2$
LASSO	$(y-\hat{y})^2$	loss fn + $\lambda \sum_{i}^{n} \lvert w_i \rvert$
RIDGE	$(y-\hat{y})^2$	Loss fn + $\lambda \sum_{i}^{n} w_i^2$

3)

a) Loss fn that are differentiable & (convex). eg: Logistic loss

(b) Loss fn that are twice differentiable eg: Squared loss

1) Underfitting - model is too simple to capture str of data (high bias)
occurs when

2) If both training error & test error remain high $\Longleftrightarrow$ Model is underfitting

3) Bagging reduce variance by training multiple models independently on different subsets of data & averaging their predictions.
Averaging helps smooth out $\to$ remove noise by individual models.

4) Boosting primarily reduces bias by training models sequentially, where each new model focus on correcting errors made by previous combination of models. $\to$ ~~Reduce variance to~~.

~~Section D~~

Section D

~~1) Steps to reduce KNN time~~: Not done till now
~~-> Approx nearest~~.

Section E

1) To minimise sum of squared errors

$$\therefore \text{loss} = \sum_1^n (y_i - c)^2$$

$$\therefore \frac{\partial \text{loss}}{\partial c} = \sum_1^n 2(y_i - c)\cdot(-1) = -2\sum_1^n (y_i - c) = 0$$

$$\Rightarrow \sum_1^n y_i = \sum_1^n c \Rightarrow \sum_1^n y_i = nc$$

$$\therefore c = \frac{1}{n}\sum_1^n y_i$$

$\to$ optimal prediction is mean of targets in the leaf

2) Gini impurity $\Rightarrow$ $1 - \sum_{1}^{3} p_i^2$ $\rightarrow$ Min = 0

$\rightarrow$ This occurs when node is pure (all samples belong to one class (0,0,1))

$\therefore$ Min gini = $1 - 1 = 0$

Max $\Rightarrow$ perfect impure node $\rightarrow$ $\therefore$ $p = (1,1,1)$

$$G_{max} = 1 - \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 = \frac{2}{3}$$

3) bcoz at each step, then choose to split that best maximises info gain locally at that node, without considering how this split will affect future splits deeper in the tree.

4) Two methods to avoide overfitting
- $\rightarrow$ Limiting maximum depth
- $\rightarrow$ Remove features $\rightarrow$ simplify model $\rightarrow$ improve generalization

---

Section F

2) Done in section C 3 4 4

1) -