CPE-695 Applied ML

Name : Hinal Patel

CWID : 10473912

Q:1 If the training model is simple and has less parameters then chances are that it is high bias and low variance. On the other hand, if the training model has large number of parameters then it's going to have high variance and low bias. So, the goal is to find the right/good balance without overfitting and underfitting the data. Hence, we perform this tradeoff where there is a tradeoff between bias and variance. An algorithm can't be more complex and less complex at the same time.

Techniques to reduce variance are,

- increasing training set
- decreasing features
- increasing lambda.

Techniques to reduce bias are,

- increasing features.
- Perform feature engineering
- Decreasing the alpha parameter of Regularization.

Q:2

$$\text{Precision} = \frac{TP}{TP+FP} = \frac{50}{50+40} = 0.555$$

$$\text{Recall} = \frac{TP}{TP+FN} = \frac{50}{50+30} = 0.625$$

$$F_1\text{-score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0.555 \times 0.625}{0.555 + 0.625}$$

$$= 0.587$$

Q:3 $\text{Gain}(\text{outlook}) = E(\text{play tennis}) - E(\text{outlook, Playtennis})$

$$= 0.97 - 0.646 = 0.322.$$

$$\text{Gain}(\text{Temp.}) = E(\text{PlayTennis}) - E(\text{temp, playtennis})$$

$$= 0.97 - 0.874 = 0.096$$

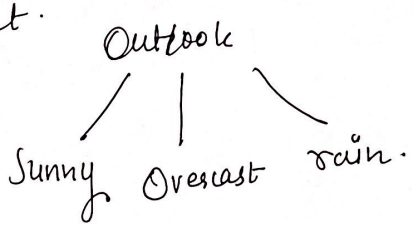
$$\text{Gain}(\text{Humidity}) = E(\text{play\_tennis}) - E(\text{humidity, Playtennis})$$

$$= 0.97 - 0.846 = 0.124$$

$$\text{Gain}(\text{Wind}) = E(\text{PlayTennis}) - E(\text{Wind, PlayTennis})$$

$$= 0.97 - 0.795 = 0.195$$

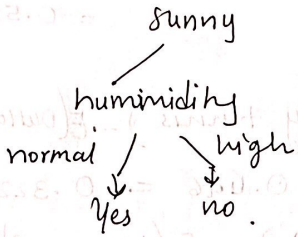
Gain for Outlook is larger so it is chosen as root.



Same for Sunny.

$$Gain(Humidity) = E(Playtennis) - E(humidity, Playten$$
$$= 0$$

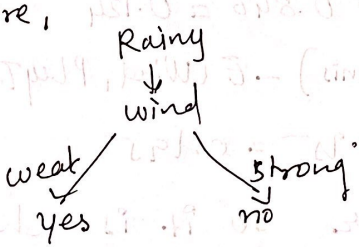
Now, gain for humidity is 0, we will get that as leaf



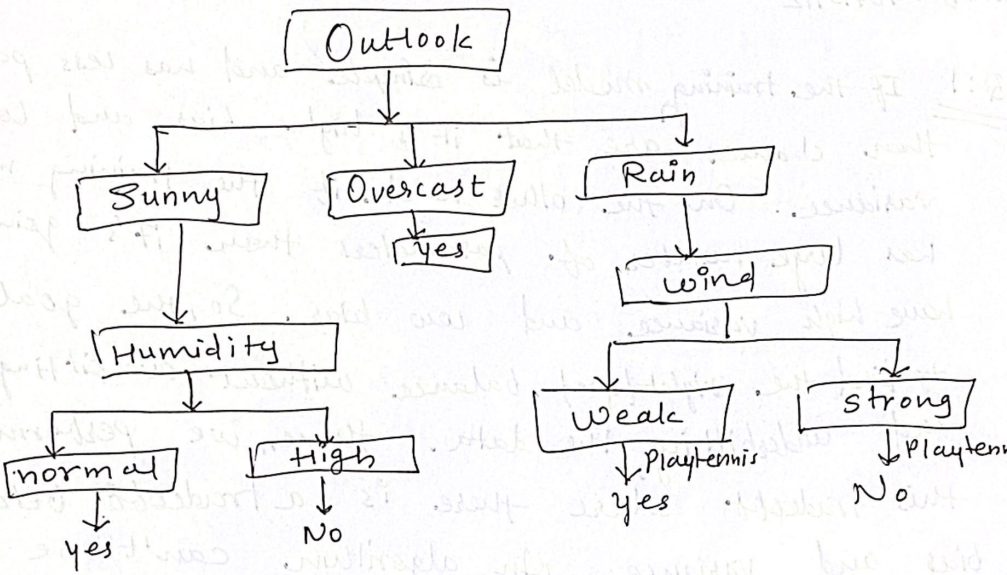

for Rainy,

$$Gain(wind) = E(Playtennis) - E(wind, Playtennis)$$
$$= 0$$

Therefore,



# final Result :-



Q:4 Using, Naive Bayes,

Predicted results are :-

classifier : 1 class 1, classifier 2 : class 1

classifier : 3 class 2

class : 1

$\hat{P}(\omega_1 / d_{1,1}(n) = 1) = \frac{40}{70}$

class : 2

$\hat{P}(\omega_2 / d_{1,1}(n) = 1) = \frac{30}{70}$

$\hat{P}(\omega_1 / d_{2,1}(n) = 1) = \frac{20}{40}$ $\hat{P}(\omega_2 / d_{2,1}(n) = 1) = \frac{20}{40}$

$\hat{P}(\omega_1 / d_{3,2}(n) = 1) = \frac{0}{10}$ $\hat{P}(\omega_2 / d_{3,2}(n) = 1) = \frac{10}{10}$

$P(\text{class } 1) = \frac{40}{70} \times \frac{20}{40} \times \frac{0}{10} = 0$.

$P(\text{class } 2) = \frac{30}{70} \times \frac{20}{40} \times \frac{10}{10} = 0.214$

$P(\text{class } 2) > P(\text{class } 1)$

Hence, we go with class 2.