Assignment 3

Illustrate Naive Bayes on the dataset to predict whether we can pet an animal or not. Find $P(x_i|y_j)$ for each x_i in X and each y_j in Y. All these calculations must be demonstrated.

Illustrate decision tree on the data set to predict whether we can pet an animal or not and all the entropy calculations must be demonstrated in the form.

	Animals	Size of Animal	Body color	Can we pet them
0	Dog	Medium	Black	Yes
1	Dog	Big	White	NO
2	Rat	Small	white	yes
3	Cow	Big	white	yes
4	Cow	Small	Brown	No
5	Cow	Big	Black	yes
6	Rat	Big	Brown	NO
7	Dog	Small	Brown	yes
8	Dog	Medium	Brown white	Yes
9	Cow	Medium	Black	No
10	Dog	Small	Black	Yes
11	Rat	medium		NO
12	Rat	Small	Brown	NO
13	Cow	Big	white	yes

1. Naive Bayes classifier -

for Naive Bayes, we need to calculate the conditional probabilities $P(x_i|y_i)$ for each feature value given the class labels (Yes or No for "can we pet them").

Let's break down the task -

Features (x) -

Animals - Dog, Rat, Cow

size of Animal - Small, medium, Big

Body color - Black, white, Brown

class label (Y) -

can we pet them - Yes, no

steps for Naive Bayes -

1. Calculate Prior Probabilities -

P(Yes) - Probability of an animal being pettable

P(No) - probability of an animals not being pettable

2. Calculate Conditional Probabilities for each feature -

P(Animal = Dog | Yes), P(Animal = Dog | No)

P(Size = Small | Yes), P(Size = small | No)

P(color = Black | yes), P(color = Black | No)

3. Combine Conditional probabilities using the Naive Bayes formula -

Given a animal, we predict P(Yes | X) & P(No|X) and classify based on which is higher

4. Example Calculation - Let's take an example of predicting if a Dog of medium size & black color can be petted.

Calculation of Prior Probabilities -

$$P(Yes) = \frac{\text{no. of Yes}}{\text{Total Samples}} = \frac{5}{14}$$

$$P(No) = \frac{6}{14}$$

Calculation of Conditional Probabilities - for "Dog, medium, Black" to predict Yes -

P(Animal = Dog | Yes)
P(Size = medium | Yes)
P(Body Color = Black | Yes)

These will be derived based on the occurences in the dataset.

2. Decision Tree.

For Decision Trees, we use entropy & information gain to build a tree that splits the data based on features in a way that best separates the Yes & No outcomes for "Can we Pet them".

Steps for Decision Tree -

1. Calculate Entropy of the Dataset - The formula for entropy is $H(S) = -\sum_{i=1}^{c} p_i \log_2(p_i)$

where p_i is the probability of class i.

2. Calculate Information Gain for each feature - for each feature (Animal, Size of Animal, Body color) Calculate the information gain

$$IG(S,A) = H(S) - \sum_{u \in Value(A)} \frac{|S_v|}{|S|} H(S_v)$$

where S_v is the subset of S for which feature A takes the value u.

3. Choose the feature with the highest Information Gain - Split the dataset based on the feature with the highest information gain & repeat the process for the child notes.

4. Example Calculation - Calculate the entropy for the initial dataset & then find the information gain for "Animals, Size of Animal & Body Color".

Example Entropy Calculation for "Can we pet them".

Total Yes = 8

Total No = 6

$$P(Yes) = \frac{8}{14}, \quad P(NO) = \frac{6}{14}$$

Entropy of the whole dataset.

$$H(S) = -\left(\frac{8}{14}\log_2\frac{8}{14} + \frac{6}{14}\log_2\frac{6}{14}\right)$$