Conditions to use Naive Bayes as supervised algorithm are as follows:

1. Data should have collection of indicator variable and class variable that is data should be supervised.
2. Class variable should always be categorical in nature.
3. Naive Bayes is used to solve only probabilistic query.

NAIVE BAYES FOR CLASSIFICATION I

1. Consider we are given supervised data set.
2. The supervised data set contains set of indicator variables and a class variable. We represent set of indicator variables by X and, class variable by C.
3. It is mandatory that C is categorical in nature but X can be mixture of numerical and categorical variables.
4. The goal of Naive Bayes classification is to estimate/compute: $P(C|X)$.

   More formally, the objective is find :

   $P(C \mid x_1, x_2, x_3, \ldots, x_n)$, where $X = (x_1, x_2, x_3, \ldots, x_n)$
5. Where $P(C \mid x_1, x_2, x_3, \ldots, x_n)$ is computed using the data set given and Bayes' rule.

AMITY UNIVERSITY ONLINE
CAREERS OF TOMORROW

- The number probabilistic query to be solved depends on the number of class variable.
- If there are 2 class variables, then we solve 2 probabilistic queries. If there are 3 class variables, then we solve 3 probabilistic queries.

NAIVE BAYES FOR CLASSIFICATION II

Example

Consider data set below.

Smoke	X-ray	Cancer
Yes	Normal	No
Yes	Abnormal	Yes
Yes	Normal	No
No	Abnormal	No
No	Abnormal	Yes

Figure 63: Hypothetical data set

1. Given this data set, the objective is to predict P (Cancer|Smoke, X-ray).
2. We can estimate P (Cancer|Smoke, X-ray) from the data set given.

NAIVE BAYES FOR CLASSIFICATION III

Example

Smoke	X-ray	Cancer
Yes	Normal	No
Yes	Abnormal	Yes
Yes	Normal	No
No	Abnormal	No
No	Abnormal	Yes

New patient comes with symptoms :
(Smoke = Yes and X-ray = Abnormal) and we need to predict possibility of Cancer.

In order to declare a person with the possibility of cancer, we need to find two probabilistic queries as below:

1. P (Cancer = No | Smoke = Yes, X-ray = Abnormal)
2. P (Cancer = Yes | Smoke = Yes, X-ray = Abnormal)

We diagnosed patient with possibility of Cancer depending upon which probabilistic query out of 2 listed above is higher.
For example, let probabilistic queries above results in following probabilities after computation.

1. P (Cancer = No | Smoke = Yes, X-ray = Abnormal) = 89.3%
2. P (Cancer = Yes | Smoke = Yes, X-ray = Abnormal) = 10.7%

Since query 1 resulted in higher probability than query 2, so we declare that the person is not having Cancer with 89.3%. confidence

AMITY UNIVERSITY ONLINE
CAREERS OF TOMORROW

NAIVE BAYES FOR CLASSIFICATION IV

Using Bayes' rule for Prediction

Smoke	X-ray	Cancer
Yes	Normal	No
Yes	Abnormal	Yes
Yes	Normal	No
No	Abnormal	No
No	Abnormal	Yes

In order to solve probabilistic query as below using data set above. We use Bayes' rule:

1. P (Cancer = No | Smoke = Yes, X-ray = Abnormal)

$$P\,(\text{Cancer} = \text{No} \mid \text{Smoke} = \text{Yes, X-ray} = \text{Abnormal}) = \frac{\overbrace{P\,(\text{Smoke} = \text{Yes, X-ray} = \text{Abnormal} \mid \text{Cancer} = \text{No})}^{P\,(\beta|\alpha)} \times \overbrace{P\,(\text{Cancer} = \text{No})}^{P\,(\alpha)}}{\underbrace{P\,(\text{Smoke} = \text{Yes, X-ray} = \text{Abnormal})}_{P\,(\beta)}} \quad (24)$$

In order to solve probabilistic query as below using data set above. We use Bayes' rule:

2. P (Cancer = Yes | Smoke = Yes, X-ray = Abnormal)

$$P\,(\text{Cancer} = \text{Yes} \mid \text{Smoke} = \text{Yes, X-ray} = \text{Abnormal}) = \frac{P\,(\text{Smoke} = \text{Yes, X-ray} = \text{Abnormal} \mid \text{Cancer} = \text{Yes}) \times P\,(\text{Cancer} = \text{Yes})}{P\,(\text{Smoke} = \text{Yes, X-ray} = \text{Abnormal})} \quad (25)$$

NAIVE BAYES FOR CLASSIFICATION V

Naive Bayes - Formal Definition

For a given data set with set of indicator features $X = (x_1, x_2, x_3, \ldots, x_n)$ and class variable C, Naive Bayes predicts C that maximises Equation 26.

$$P(C \mid x_1, x_2, x_3, \ldots, x_n) = \frac{P(x_1, x_2, x_3, \ldots, x_n \mid C) \times P(C)}{P(x_1, x_2, x_3, \ldots, x_n)} \tag{26}$$

In order to solve Equation 26,

1. P (C) is directly given in data.
2. $P(x_1, x_2, x_3, \ldots, x_n \mid C)$ is given in data.
3. $P(x_1, x_2, x_3, \ldots, x_n)$, we need not care for its value. For the reason it remain same for each class value in the data set.

NAIVE BAYES FOR CLASSIFICATION VI

Naive Bayes - Conditional Independence Assumption

- Naive bayes assumes that all features present in the data set are independent to each other.
- The only dependency that exist is between a class variable and features present in the data set.
- For formally, given feature set $X = (x_1, x_2, x_3, \ldots, x_n)$, Naive Bayes assumes that there is no relationship between features in X.
- In Figure 65, Naive Bayes model on data set in Table 64 is presented.

Smoke	X-ray	Cancer
Yes	Normal	No
Yes	Abnormal	Yes
Yes	Normal	No
No	Abnormal	No
No	Abnormal	Yes

Figure 64: Hypothetical data set

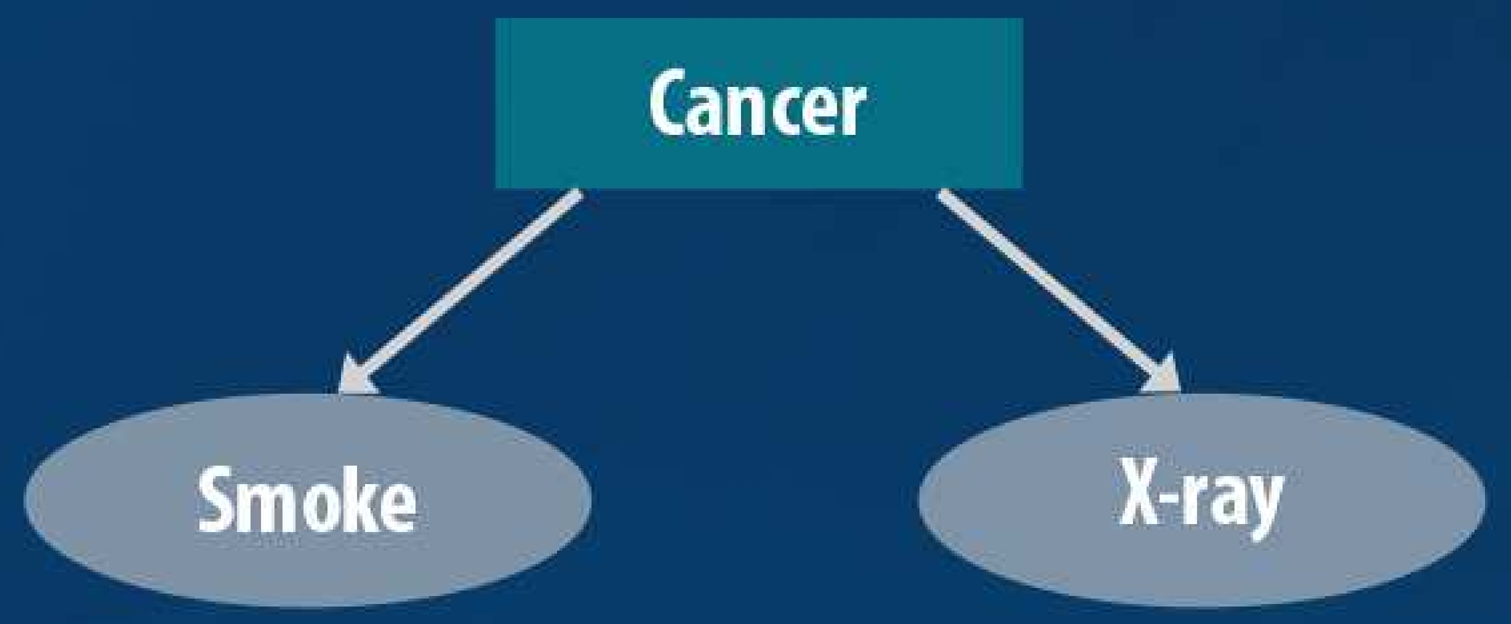

Figure 65: Naive Bayes model on data set above

NAIVE BAYES FOR CLASSIFICATION VII

Using Conditional Independence Assumption to solve Probabilistic Queries

Consider for the given data set below, we are to solve:

P (Cancer = No | Smoke = Yes, X-ray = Abnormal)

$$P(\text{Cancer} = \text{No} \mid \text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal}) = \frac{P(\text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal} \mid \text{Cancer} = \text{No}) \times P(\text{Cancer} = \text{No})}{\cancel{P(\text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal})}} \quad (27)$$

Using Conditional Independence assumption Equation above can be rewritten as below.

P (Cancer = No | Smoke = Yes, X-ray = Abnormal)

$$P(\text{Cancer} = \text{No} \mid \text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal}) = \frac{P(\text{Smoke} = \text{Yes} \mid \text{Cancer} = \text{No}) \times P(\text{X-ray} = \text{Abnormal} \mid \text{Cancer} = \text{No}) \times P(\text{Cancer} = \text{No})}{\cancel{P(\text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal})}} \quad (28)$$

All terms in right of the Equation can be computed directly using data set given.

NAIVE BAYES FOR CLASSIFICATION VIII

Naive Bayes - Estimating probabilities from Data - Example Illustration 1

Objective: Forecast chances of Cancer for symptoms (Smoke = Present, X-ray = Abnormal) using Naive Bayes.

Smoke	X-ray	Cancer
Yes	Normal	No
Yes	Abnormal	Yes
Yes	Normal	No
No	Abnormal	No
No	Abnormal	Yes

Figure 66: Hypothetical data set

1. $$P(\text{Cancer} = \text{No} \mid \text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal}) = \frac{P(\text{Smoke} = \text{Yes} \mid \text{Cancer} = \text{No}) \times P(\text{X-ray} = \text{Abnormal} \mid \text{Cancer} = \text{No}) \times P(\text{Cancer} = \text{No})}{\cancel{P(\text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal})}} \quad (29)$$

Using data set,

$P(\text{Cancer} = \text{No}) = 3/5 = 0.6$

$P(\text{Smoke} = \text{Yes} \mid \text{Cancer} = \text{No}) = 2/3 = 0.66$ $P(\text{X-ray} = \text{Abnormal} \mid \text{Cancer} = \text{No}) = 1/3 = 0.33$. Substituting all these terms in Equation above gives:

$$P(\text{Cancer} = \text{No} \mid \text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal}) = \frac{0.66 \times 0.33 \times 0.6}{\cancel{P(\text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal})}} = 13.06\% \quad (30)$$

NAIVE BAYES FOR CLASSIFICATION IX

Naive Bayes - Estimating probabilities from Data - Example Illustration 1

Smoke	X-ray	Cancer
Yes	Normal	No
Yes	Abnormal	Yes
Yes	Normal	No
No	Abnormal	No
No	Abnormal	Yes

Figure 67: Hypothetical data set

2. $$P(\text{Cancer} = \text{Yes} \mid \text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal}) = \frac{P(\text{Smoke} = \text{Yes} \mid \text{Cancer} = \text{Yes}) \times P(\text{X-ray} = \text{Abnormal} \mid \text{Cancer} = \text{Yes}) \times P(\text{Cancer} = \text{Yes})}{\cancel{P(\text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal})}} \quad (31)$$

Using data set,

$P(\text{Cancer} = \text{Yes}) = 2/5 = 0.4$

$P(\text{Smoke} = \text{Yes} \mid \text{Cancer} = \text{Yes}) = 1/2 = 0.50$

$P(\text{X-ray} = \text{Abnormal} \mid \text{Cancer} = \text{Yes}) = 2/2 = 1$. Substituting all these terms in Equation above gives:

$$P(\text{Cancer} = \text{Yes} \mid \text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal}) = \frac{0.5 \times 1 \times 0.4}{\cancel{P(\text{Smoke} = \text{Yes}, \text{X-ray} = \text{Abnormal})}} = 20.0\% \quad (32)$$

The class of new observation (Smoke = Yes, X-ray = Abnormal) is Cancer = Yes