Naïve Bayes

Bayes Theorem:

P(A|B) = P(B|A) x P(A)

P(B)

Example : (When CSK WON)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sr. No | Toss | Venue | Outlook | Result |
| 1 | Won | Mumbai | Overcast | Won |
| 2 | Lost | Chennai | Sunny | Won |
| 3 | Won | Kolkata | Sunny | Won |
| 4 | Won | Chennai | Sunny | Won |
| 5 | Lost | Mumbai | Sunny | Lost |
| 6 | Won | Chennai | Overcast | Lost |
| 7 | Won | Kolkata | Overcast | Lost |
| 8 | Won | Mumbai | Sunny | Won |

Input = Lost, Mumbai, Sunny ( Given By user)

Predict = Yes / No ( Output Class)

Training Peroid :

In Training Phase what Naïve bayes does is create a Frequency Table for each attribute against the target. Then, molding the frequency tables to Likelihood Tables

For Toss:

|  |  |  |  |
| --- | --- | --- | --- |
|  | | Csk | |
| Won | Lost |
| Toss | Won | 4/5 | 2/3 |
| Lost | 1/5 | 1/3 |

For Outlook:

|  |  |  |  |
| --- | --- | --- | --- |
|  | | CSK | |
| Won | Lost |
| Outlook | Overcast | 1/5 | 2/3 |
| Sunny | 4/5 | 1/3 |

For Venue:

|  |  |  |  |
| --- | --- | --- | --- |
|  | | CSK | |
| Won | Lost |
| Venue | Mumbai | 2/5 | 1/3 |
| Chennai | 2/5 | 1/3 |
| Kolkata | 1/5 | 1/3 |
|  |  |  |  |

Now if we use normal Bayes theorem it will try to find out Probability of user inputs in the dataset for particular target column there is possibility that such combination isn’t present in the dataset.

So,

Probablity of Winning:

P(W|Lost , Mumbai, Sunny) = P( Lost, Mumbai, Sunny|W) x P(W) (This Combination isn’t present)

P( Lost, Mumbai, Sunny )

P(W|Lost , Mumbai, Sunny) = 0 X (5/8)

P(W|Lost , Mumbai, Sunny) = 0

Probablity of Losing:

P(L | Lost, Mumbai, Sunny) = P( Lost, Mumbai, Sunny|L) x P(L) (This Combination is Present)

P( Lost, Mumbai, Sunny)

P(L | Lost, Mumbai, Sunny) = (1/3) x(3/8) = 1/8

Assumption:

To overcome this Problem Naïve bayes made one assumption, it assumes that each and every feature made equal and independent contribution in outcome that’s why its called Naïve Bayes.

After Assumptions:

Probablity of Winning:

P(W|Lost , Mumbai, Sunny) = P(Lost|W) X P(Mumbai | W) X P(Sunny | W) X P(W)

= (1/5) x (2/5) x (4/5) x (5/8)

= 0.040

Probablity of Lost:

P(L| Lost, Mumbai, Sunny) = P(Lost|L) X P(Mumbai|L) X P(Sunny|L) X P(L)

= (1/3) x (1/3) x (1/3) x (3/8)

= 0.013

Now,

P(Won) = (0.040 / (0.040 + 0.013)) = 0.754

P(Lost) = (0.013 /(0.040 + 0.013)) = 0.246

= Prediction Class will be Won.