

## Naïve Bayes Example

| Day | Outlook  | Temperature | Humidity | Wind   | Play Tennis |
|-----|----------|-------------|----------|--------|-------------|
| 1   | Sunny    | Hot         | High     | Weak   | No          |
| 2   | Sunny    | Hot         | High     | Strong | No          |
| 3   | Overcast | Hot         | High     | Weak   | Yes         |
| 4   | Rain     | Mild        | High     | Weak   | Yes         |
| 5   | Rain     | Cool        | Normal   | Weak   | Yes         |
| 6   | Rain     | Cool        | Normal   | Strong | No          |
| 7   | Overcast | Cool        | Normal   | Strong | Yes         |
| 8   | Sunny    | Mild        | High     | Weak   | No          |
| 9   | Sunny    | Cool        | Normal   | Weak   | Yes         |
| 10  | Rain     | Mild        | Normal   | Weak   | Yes         |
| 11  | Sunny    | Mild        | Normal   | Strong | Yes         |
| 12  | Overcast | Mild        | High     | Strong | Yes         |
| 13  | Overcast | Hot         | Normal   | Weak   | Yes         |
| 14  | Rain     | Mild        | High     | Strong | No          |

$P(\text{play tennis} / \text{outlook, temperature, humidity, wind}) = P(\text{outlook} / \text{play tennis}) * P(\text{temperature} / \text{play tennis}) * P(\text{humidity} / \text{play tennis}) * P(\text{wind} / \text{play tennis})$

### Training Phase

$$P(\text{play tennis} = \text{yes}) = \frac{9}{14}, P(\text{play tennis} = \text{No}) = \frac{5}{14}$$

### Outlook Table

| Feature  | Frequency  |           | Probability in Class |           |
|----------|------------|-----------|----------------------|-----------|
| Outlook  | Play = yes | Play = No | Play = yes           | Play = No |
| Sunny    | 2          | 3         | 2/9                  | 3/5       |
| Overcast | 4          | 0         | 4/9                  | 0/5       |
| Rain     | 3          | 2         | 3/9                  | 2/5       |

### Temperature Table

| Feature     | Frequency  |           | Probability in Class |           |
|-------------|------------|-----------|----------------------|-----------|
| Temperature | Play = Yes | Play = No | Play = Yes           | Play = No |
| Hot         | 2          | 2         | 2/9                  | 2/5       |
| Mild        | 4          | 2         | 4/9                  | 2/5       |
| Cool        | 3          | 1         | 3/9                  | 1/5       |

### Humidity Table

| Feature | Frequency  |           | Probability in Class |           |
|---------|------------|-----------|----------------------|-----------|
|         | Play = Yes | Play = No | Play = Yes           | Play = No |
| High    | 3          | 4         | 3/9                  | 4/5       |
| Normal  | 6          | 1         | 6/9                  | 1/5       |

### Wind Table

| Feature | Frequency  |           | Probability in Class |           |
|---------|------------|-----------|----------------------|-----------|
|         | Play = Yes | Play = No | Play = Yes           | Play = No |
| Strong  | 3          | 3         | 3/9                  | 3/5       |
| Weak    | 6          | 2         | 6/9                  | 2/5       |

### Classification Phase

$x_q = (\text{outlook} = \text{sunny}, \text{temperature} = \text{cool}, \text{Humidity} = \text{high}, \text{Wind} = \text{strong})$

$P(\text{Play Tennis} = \text{Yes}/x_q) = P(\text{Play Tennis} = \text{yes} / \text{outlook} = \text{sunny}) * P(\text{Play Tennis} = \text{yes} / \text{temperature} = \text{cool}) * P(\text{Play Tennis} = \text{yes} / \text{Humidity} = \text{high}) * P(\text{Play Tennis} = \text{yes} / \text{Wind} = \text{strong}) * P(\text{Play Tennis} = \text{Yes})$

$$= \frac{2}{9} * \frac{3}{9} * \frac{3}{9} * \frac{3}{9} = \frac{54}{6561} * \frac{9}{14} = 0.0053$$

$P(\text{Play Tennis} = \text{No}/x_q) = P(\text{Play Tennis} = \text{No} / \text{outlook} = \text{sunny}) * P(\text{Play Tennis} = \text{No} / \text{temperature} = \text{cool}) * P(\text{Play Tennis} = \text{No} / \text{Humidity} = \text{high}) * P(\text{Play Tennis} = \text{No} / \text{Wind} = \text{strong}) * P(\text{Play Tennis} = \text{No})$

$$= \frac{3}{5} * \frac{1}{5} * \frac{4}{5} * \frac{3}{5} * \frac{5}{14} = \frac{36}{625} * \frac{5}{14} = 0.0205$$

Since,  $P(\text{Play Tennis} = \text{No}) > P(\text{Play Tennis} = \text{Yes})$ .

Hence,  $x_q$  will be classified as **negative**.