Practice session 2
MTAT.03.227 Machine learning
21.09.2018

Mikk
Agenda

1. KNN
2. Naive Bayes
3. Homework review (deadline on September 24th!)
4. Questions
Tennis Dataset
## Tennis Dataset

<table>
<thead>
<tr>
<th>Day</th>
<th>Outlook</th>
<th>Temp</th>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>Sunny</td>
<td>Hot</td>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>D2</td>
<td>Sunny</td>
<td>Hot</td>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>D3</td>
<td>Overcast</td>
<td>Hot</td>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>D4</td>
<td>Rain</td>
<td>Mild</td>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>D5</td>
<td>Rain</td>
<td>Cool</td>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>D6</td>
<td>Rain</td>
<td>Cool</td>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>D7</td>
<td>Overcast</td>
<td>Cool</td>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>D8</td>
<td>Sunny</td>
<td>Mild</td>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>D9</td>
<td>Sunny</td>
<td>Cool</td>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>D10</td>
<td>Rain</td>
<td>Mild</td>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>D11</td>
<td>Sunny</td>
<td>Mild</td>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>D12</td>
<td>Overcast</td>
<td>Mild</td>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>D13</td>
<td>Overcast</td>
<td>Hot</td>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>D14</td>
<td>Rain</td>
<td>Mild</td>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>
Shall we play tennis today?

<table>
<thead>
<tr>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
</tbody>
</table>
Shall we play tennis today?

\[ P(C_1) = P(\text{Yes}) = ??? \]
Shall we play tennis today?

<table>
<thead>
<tr>
<th>PlayTennis</th>
<th>P(C₁) = P(Yes) = 9/14 = 0.64</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>P(C₂) = P(No) = ???</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>
Shall we play tennis today?

<table>
<thead>
<tr>
<th>PlayTennis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

$P(C_1) = P(Yes) = \frac{9}{14} = 0.64$

$P(C_2) = P(No) = \frac{5}{14} = 0.36$
Shall we play tennis today?

\[
P(C_1) = P(\text{Yes}) = \frac{9}{14} = 0.64 \\
P(C_2) = P(\text{No}) = \frac{5}{14} = 0.36
\]

\[P(\text{Yes}) > P(\text{No}) \]
\[0.64 > 0.36\]
Shall we play tennis today?

P(Yes) = 0.64 and P(No) = 0.36 are called prior probabilities, which in the case of classification tells us how likely each of the classes is a prior, i.e., before we have observed the data (features).

\[ P(C_1) = P(Yes) = \frac{9}{14} = 0.64 \]
\[ P(C_2) = P(No) = \frac{5}{14} = 0.36 \]

P(Yes) > P(No)
0.64 > 0.36

Yes
It’s windy today. Tennis, anyone?

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>
It’s windy today. Tennis, anyone?

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>
It’s windy today. Tennis, anyone?

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[
P(\text{Weak}) = \frac{8}{14} = 0.57 \]

\[
P(\text{Strong}) = ??
\]
It’s windy today. Tennis, anyone?

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

P(Weak) = \(\frac{8}{14} = 0.57\)

P(Strong) = \(\frac{6}{14} = 0.43\)
It’s windy today. Tennis, anyone?

This is a conditional probability. What is the probability of PlayTennis=Yes given that the wind is strong?

\[ P(Yes \mid Strong) = ??? \]

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>
It’s windy today. Tennis, anyone?

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

P(Weak) = \(\frac{8}{14} = 0.57\)

P(Strong) = \(\frac{6}{14} = 0.43\)

P(Yes | Strong) = \(\frac{3}{6} = 0.5\)
It’s windy today. Tennis, anyone?

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[
P(\text{Weak}) = \frac{8}{14} = 0.57 \\
P(\text{Strong}) = \frac{6}{14} = 0.43 \\
P(\text{Yes} | \text{Strong}) = \frac{3}{6} = 0.5 \\
P(\text{No} | \text{Strong}) = ???
\]
It’s windy today. Tennis, anyone?

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

P(Weak) = 8/14 = 0.57

P(Strong) = 6/14 = 0.43

P(Yes | Strong) = 3/6 = 0.5

P(No | Strong) = 3/6 = 0.5
It’s windy today. Tennis, anyone?

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[
P(\text{Weak}) = \frac{8}{14} = 0.57
\]

\[
P(\text{Strong}) = \frac{6}{14} = 0.43
\]

\[
P(\text{Yes} \mid \text{Strong}) = \frac{3}{6} = 0.5
\]

\[
P(\text{No} \mid \text{Strong}) = \frac{3}{6} = 0.5
\]

\[
P(\text{Yes} \mid \text{Weak}) = ???
\]

\[
P(\text{No} \mid \text{Weak}) = ???
\]
It’s windy today. Tennis, anyone?

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[
P(\text{Weak}) = \frac{8}{14} = 0.57
\]
\[
P(\text{Strong}) = \frac{6}{14} = 0.43
\]
\[
P(\text{Yes} \mid \text{Strong}) = \frac{3}{6} = 0.5
\]
\[
P(\text{No} \mid \text{Strong}) = \frac{3}{6} = 0.5
\]
\[
P(\text{Yes} \mid \text{Weak}) = \frac{6}{8} = 0.75
\]
\[
P(\text{No} \mid \text{Weak}) = \frac{2}{8} = 0.25
\]
It’s windy today. Tennis, anyone?

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[
P(\text{Weak}) = \frac{8}{14} = 0.57
\]
\[
P(\text{Strong}) = \frac{6}{14} = 0.43
\]
\[
P(\text{Yes} \mid \text{Strong}) = \frac{3}{6} = 0.5
\]
\[
P(\text{No} \mid \text{Strong}) = \frac{3}{6} = 0.5
\]
\[
P(\text{Yes} \mid \text{Weak}) = \frac{6}{8} = 0.75
\]
\[
P(\text{No} \mid \text{Weak}) = \frac{2}{8} = 0.25
\]
Conditional probability

\[ P(Y|X) = \frac{P(Y, X)}{P(X)} \]

Here, \( P(Y, X) \) is the joint probability (or likelihood) of two events occurring together. In our case, the class Y (Yes or No) and features X.
It’s windy today. Tennis, anyone?

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

P(Weak) = 8/14 = 0.57

P(Strong) = 6/14 = 0.43

P(Yes | Strong) = 3/6 = 0.5

P(No | Strong) = 3/6 = 0.5

P(Yes | Weak) = 6/8 = 0.75

P(No | Weak) = 2/8 = 0.25

\[
P(Yes | Strong) = \frac{P(Yes, Strong)}{P(Strong)}
\]

Definition of conditional probability
## Conditional probability

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[
P(Yes|Strong) = \frac{P(Yes, Strong)}{P(Strong)} = \frac{\frac{3}{14}}{\frac{6}{14}} = \frac{3 \times 14}{6 \times 14} = \frac{3}{6} = \frac{1}{2} = 0.5
\]
More attributes

<table>
<thead>
<tr>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>
More attributes

P(High, Weak) = ???

<table>
<thead>
<tr>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>
More attributes

<table>
<thead>
<tr>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[ P(\text{High, Weak}) = \frac{4}{14} = 0.29 \]
### More attributes

<table>
<thead>
<tr>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[ P(\text{High, Weak}) = \frac{4}{14} = 0.29 \]

\[ P(\text{Yes} \mid \text{High, Weak}) = \text{?} \]
More attributes

<table>
<thead>
<tr>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[
P(\text{High, Weak}) = \frac{4}{14} = 0.29
\]

\[
P(\text{Yes} \mid \text{High, Weak}) = ???
\]
More attributes

<table>
<thead>
<tr>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[ P(\text{High, Weak}) = \frac{4}{14} = 0.29 \]
\[ P(\text{Yes | High, Weak}) = \frac{2}{4} = 0.5 \]
More attributes

<table>
<thead>
<tr>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[ P(\text{High, Weak}) = \frac{4}{14} = 0.29 \]
\[ P(\text{Yes} \mid \text{High, Weak}) = \frac{2}{4} = 0.5 \]
\[ P(\text{No} \mid \text{High, Weak}) = \frac{2}{4} = 0.5 \]
### More attributes

<table>
<thead>
<tr>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

- \(P(\text{High, Weak}) = \frac{4}{14} = 0.29\)
- \(P(\text{Yes | High, Weak}) = \frac{2}{4} = 0.5\)
- \(P(\text{No | High, Weak}) = \frac{2}{4} = 0.5\)
- \(P(\text{High, Strong}) = \frac{3}{14} = 0.21\)
- \(P(\text{Yes | High, Strong}) = \frac{1}{3} = 0.33\)
- \(P(\text{No | High, Strong}) = \frac{2}{3} = 0.69\)

...
1. Estimate from data:
   \[ P(\text{Class} \mid X_1, X_2, X_3, \ldots) \]

2. For a given instance \((X_1, X_2, X_3, \ldots)\) predict class whose conditional probability is greater:
Classifier

1. Estimate from data:
   \[ P(\text{Class} \mid X_1, X_2, X_3, \ldots) \]

2. For a given instance \((X_1, X_2, X_3, \ldots)\) predict class whose conditional probability is greater:
   \[ P(C_1 \mid X_1, X_2, X_3, \ldots) > P(C_2 \mid X_1, X_2, X_3, \ldots) \rightarrow \text{predict } C_1 \]
More attributes

<table>
<thead>
<tr>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[ P(\text{High}, \text{Strong}) = \frac{3}{14} = 0.21 \]
\[ P(\text{Yes} \mid \text{High}, \text{Strong}) = \frac{1}{3} = 0.33 \]
\[ P(\text{No} \mid \text{High}, \text{Strong}) = \frac{2}{3} = 0.69 \]
More attributes

<table>
<thead>
<tr>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[
P(\text{High, Strong}) = \frac{3}{14} = 0.21
\]

\[
P(\text{Yes} \mid \text{High, Strong}) = \frac{1}{3} = 0.33
\]

\[
P(\text{No} \mid \text{High, Strong}) = \frac{2}{3} = 0.69
\]

\[P(\text{No} \mid \text{High, Strong}) > P(\text{Yes} \mid \text{High, Strong}) \rightarrow \text{predict No}\]
\[
P(\text{Yes} \mid \text{Strong}) = \frac{P(\text{Yes, Strong})}{P(\text{Strong})}
\]

\[
P(\text{Yes} \mid \text{Strong}) = \frac{3}{6} = 0.5
\]
P(Yes | Strong) = P(Yes, Strong)/P(Strong)

P(Strong | Yes) = ???
P(Yes | Strong) = P(Yes, Strong)/P(Strong)

P(Strong | Yes) = P(Strong, Yes)/P(Yes)
\[ P(\text{Yes} \mid \text{Strong}) = \frac{P(\text{Yes, Strong})}{P(\text{Strong})} \]

\[ P(\text{Strong} \mid \text{Yes}) = \frac{P(\text{Strong, Yes})}{P(\text{Yes})} \]
\[ P(\text{Yes} \mid \text{Strong}) = \frac{P(\text{Yes, Strong})}{P(\text{Strong})} \]

\[ P(\text{Strong} \mid \text{Yes}) = \frac{P(\text{Strong, Yes})}{P(\text{Yes})} \]

\[ P(\text{Yes, Strong}) = P(\text{Strong, Yes}) \quad ??? \]
\[ P(\text{Yes} \mid \text{Strong}) = \frac{P(\text{Yes, Strong})}{P(\text{Strong})} \]

\[ P(\text{Strong} \mid \text{Yes}) = \frac{P(\text{Strong, Yes})}{P(\text{Yes})} \]

<table>
<thead>
<tr>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[ P(\text{Strong, Yes}) = P(\text{Yes, Strong}) \]
\[ P(\text{Yes} | \text{Strong}) = \frac{P(\text{Yes}, \text{Strong})}{P(\text{Strong})} \]
\[ P(\text{Strong} | \text{Yes}) = \frac{P(\text{Strong}, \text{Yes})}{P(\text{Yes})} \]
\[ P(\text{Strong, Yes}) = P(\text{Yes, Strong}) \]
\[ P(\text{Yes, Strong}) = ??? \]
P(Yes | Strong) = P(Yes, Strong)/P(Strong)

P(Strong | Yes) = P(Strong, Yes)/P(Yes)

P(Strong, Yes) = P(Yes, Strong)

P(Yes, Strong) = P(Yes | Strong)P(Strong)
\[ P(Yes \mid Strong) = \frac{P(Yes, Strong)}{P(Strong)} \]

\[ P(Strong \mid Yes) = \frac{P(Strong, Yes)}{P(Yes)} \]

\[ P(Strong, Yes) = P(Yes, Strong) \]

\[ P(Yes, Strong) = P(Yes \mid Strong)P(Strong) = \]

\[ = ??? \]
P(Yes | Strong) = P(Yes, Strong)/P(Strong)

P(Strong | Yes) = P(Strong, Yes)/P(Yes)

P(Strong, Yes) = P(Yes, Strong)

P(Yes, Strong) = P(Yes | Strong)P(Strong) =
= P(Strong | Yes)P(Yes)
\[ P(\text{Yes} \mid \text{Strong}) = \frac{P(\text{Yes}, \text{Strong})}{P(\text{Strong})} \]

\[ P(\text{Strong} \mid \text{Yes}) = \frac{P(\text{Strong}, \text{Yes})}{P(\text{Yes})} \]

\[ P(\text{Strong}, \text{Yes}) = P(\text{Yes}, \text{Strong}) \]

\[ P(\text{Yes}, \text{Strong}) = P(\text{Yes} \mid \text{Strong})P(\text{Strong}) = \]

\[ = P(\text{Strong} \mid \text{Yes})P(\text{Yes}) \]
\[ P(\text{Yes} \mid \text{Strong}) = \frac{P(\text{Yes}, \text{Strong})}{P(\text{Strong})} \]

\[ P(\text{Strong} \mid \text{Yes}) = \frac{P(\text{Strong}, \text{Yes})}{P(\text{Yes})} \]

\[ P(\text{Strong}, \text{Yes}) = P(\text{Yes}, \text{Strong}) \]

\[ P(\text{Yes} \mid \text{Strong})P(\text{Strong}) = P(\text{Strong} \mid \text{Yes})P(\text{Yes}) \]
\[ P(\text{Yes} \mid \text{Strong}) = \frac{P(\text{Yes, Strong})}{P(\text{Strong})} \]
\[ P(\text{Strong} \mid \text{Yes}) = \frac{P(\text{Strong, Yes})}{P(\text{Yes})} \]
\[ P(\text{Strong, Yes}) = P(\text{Yes, Strong}) \]
\[ P(\text{Yes} \mid \text{Strong})P(\text{Strong}) = P(\text{Strong} \mid \text{Yes})P(\text{Yes}) \]
\[ P(\text{Yes} \mid \text{Strong}) = \text{???} \]
\[ P(\text{Yes} \mid \text{Strong}) = \frac{P(\text{Yes, Strong})}{P(\text{Strong})} \]

\[ P(\text{Strong} \mid \text{Yes}) = \frac{P(\text{Strong, Yes})}{P(\text{Yes})} \]

\[ P(\text{Strong, Yes}) = P(\text{Yes, Strong}) \]

\[ P(\text{Yes} \mid \text{Strong})P(\text{Strong}) = P(\text{Strong} \mid \text{Yes})P(\text{Yes}) \]

\[ P(\text{Yes} \mid \text{Strong}) = \frac{P(\text{Strong} \mid \text{Yes})P(\text{Yes})}{P(\text{Strong})} \]
P(Yes | Strong) = P(Yes, Strong)/P(Strong)

P(Strong | Yes) = P(Strong, Yes)/P(Yes)

P(Strong, Yes) = P(Yes, Strong)

P(Yes | Strong)P(Strong) = P(Strong | Yes)P(Yes)

P(Yes | Strong) = P(Strong | Yes)P(Yes)/P(Strong)
P(Yes | Strong) = P(Yes, Strong)/P(Strong)
P(Strong | Yes) = P(Strong, Yes)/P(Yes)

P(Strong, Yes) = P(Yes, Strong)
P(Yes | Strong)P(Strong) = P(Strong | Yes)P(Yes)

\[
P(Yes \mid Strong) = \frac{P(Strong \mid Yes)P(Yes)}{P(Strong)}
\]

You have just derived the Bayes’ rule!
The Bayes’ rule

\[ P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \]

\( X \) - is a feature vector
\( Y \) - is a class or a label (Yes or No)
The Bayes’ rule

\[ P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \]
The Bayes’ rule

\[ P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \]

Here, \( P(Y) \) is the **prior probability**, which in the case of classification tells us how likely each of the classes is a prior, i.e., before we have observed the data \( X \).
Shall we play tennis today?

<table>
<thead>
<tr>
<th>PlayTennis</th>
<th>P(C₁) = P(Yes) = 9/14 = 0.64</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>P(C₂) = P(No) = 5/14 = 0.36</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>
The Bayes’ rule

\[ P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \]

\( P(X) \) is the probability of the data, aka evidence, which is independent of \( Y \).
### More attributes

<table>
<thead>
<tr>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

\[
P(\text{High, Weak}) = \frac{4}{14} = 0.29 \]

\[
P(\text{High, Strong}) = \frac{3}{14} = 0.21 \]
The Bayes’ rule

\[ P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \]

\( P(X|Y) \) is a conditional probability, it represents the **likelihood** of data X occurring given class Y.
The Bayes’ rule

\[ \text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}} \]
Probability vs Likelihood
Probability vs Likelihood

In non-technical parlance, “likelihood” is usually a synonym for “probability”, but in statistical usage there is a clear distinction:

1. **probability** allows us to predict unknown outcomes based on known features

2. **likelihood** allows us to estimate the likelihood of the class - that is the probability of getting these observed feature values given the class.

\[
P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}
\]

*posterior probability*
The **Bayes** Classifier

1. Estimate from data:
   \[ P(\text{Class} | X_1, X_2, X_3, \ldots) \]

2. For a given instance \((X_1, X_2, X_3, \ldots)\) predict class whose
   **conditional probability** is **greater**:

   \[ P(\text{No} | \text{High, Strong}) > P(\text{Yes} | \text{High, Strong}) \rightarrow \text{predict No} \]

   \[ P(\text{Yes} | \text{High, Strong}) = 1/3 = 0.33 \]
   \[ P(\text{No} | \text{High, Strong}) = 2/3 = 0.69 \]
The Bayes classifier

\[ P(\text{Yes} \mid \text{Strong}) > P(\text{No} \mid \text{Strong}) \quad \rightarrow \quad \text{predict Yes} \]
The Bayes classifier

\[ P(\text{Yes} \mid \text{Strong}) > P(\text{No} \mid \text{Strong}) \implies \text{predict Yes} \]

Bayes’ rule

???
The Bayes classifier

\[ P(\text{Yes} \mid \text{Strong}) > P(\text{No} \mid \text{Strong}) \quad \rightarrow \quad \text{predict Yes} \]

Bayes’ rule

\[ \frac{P(\text{Strong} \mid \text{Yes}) P(\text{Yes})}{P(\text{Strong})} > \frac{P(\text{Strong} \mid \text{No}) P(\text{No})}{P(\text{Strong})} \quad \rightarrow \quad \text{predict Yes} \]
The Bayes classifier

$$P(Yes \mid Strong) > P(No \mid Strong) \quad \rightarrow \quad \text{predict Yes}$$

Bayes’ rule

$$\frac{P(Strong \mid Yes) \cdot P(Yes)}{P(Strong)} > \frac{P(Strong \mid No) \cdot P(No)}{P(Strong)} \quad \rightarrow \quad \text{predict Yes}$$

$$P(Strong \mid Yes) \cdot P(Yes) > P(Strong \mid No) \cdot P(No) \quad \rightarrow \quad \text{predict Yes}$$
The Bayes classifier

\[ P(Yes \mid Strong) > P(No \mid Strong) \quad \rightarrow \quad \text{predict Yes} \]

Bayes’ rule

\[
\frac{P(Yes \mid Strong) \cdot P(Yes)}{P(Strong)} > \frac{P(No \mid Strong) \cdot P(No)}{P(Strong)}
\]

\rightarrow \quad \text{predict Yes}

\[
P(Strong \mid Yes) \cdot P(Yes) > P(Strong \mid No) \cdot P(No)
\]  

\rightarrow \quad \text{predict Yes}

\[
P(Strong \mid Yes)/P(Strong \mid No) > P(No)/P(Yes)
\]  

\rightarrow \quad \text{predict Yes}
What about Naïve Bayes classifier?
What about **Naïve** Bayes classifier?

Assume (naively) that features $X_1, X_2, X_3, ...$ are independent given the category or class C.
What about **Naïve** Bayes classifier?

Assume (naively) that features $X_1, X_2, X_3, ...$ are independent given the category or class $C$

Then:

$$P(X_1, X_2 \mid C) = P(X_1 \mid C) \cdot P(X_2 \mid C) \quad \text{and} \quad P(X_1 \mid X_2, C) = P(X_1 \mid C)$$
What about **Naïve** Bayes classifier?

Assume (naively) that features $X_1, X_2, X_3, \ldots$ are independent given the category or class $C$

Then:

$$P(X_1, X_2 \mid C) = P(X_1 \mid C) \cdot P(X_2 \mid C) \quad \text{and} \quad P(X_1 \mid X_2, C) = P(X_1 \mid C)$$

And this gives us tools to use:

$$P(X_i \mid X_{i+1}, X_{i+2}, \ldots, X_n, C_k) = P(X_i \mid C_k)$$

$$P(C_k \mid x_1, x_2, \ldots, x_n) = P(C_k) \prod_{i=1}^{n} P(x_i \mid C_k)$$
Naïve Bayes Classifier

\[ P(C_k|x_1, x_2, \ldots, x_n) = P(C_k) \prod_{i=1}^{n} P(x_i|C_k) \]

The naïve Bayes classifier combines this Naive Bayes probability model with a decision rule. One common rule is to pick the class that is most probable. This rule is also known as the maximum a posteriori (MAP) decision rule.

\[ \hat{y} = \arg\max_{k \in \{1, \ldots, k\}} P(C_k) \prod_{i=1}^{n} P(x_i|C_k) \]

https://en.wikipedia.org/wiki/Naive_Bayes_classifier
Naïve Bayes Classifier

If we would assume a uniform prior distribution (i.e. $P(Y)$) is the same for every value of $Y$ or we want to ignore prior distribution then this reduces to **maximum likelihood** (ML) decision rule.

$$
\hat{y} = \arg\max_{k \in (1, \ldots, k)} P(C_k) \prod_{i=1}^{n} P(x_i | C_k)
$$
Naïve Bayes Classifier

If we have **only two classes** it is convenient to work with ratios of posterior probabilities or likelihood ratios and prior odds.

\[
\frac{P(C_1 | x_1, x_2, ..., x_n)}{P(C_2 | x_1, x_2, ..., x_n)} = \frac{P(C_1) \prod_{i=1}^{n} P(x_i | C_1)}{P(C_2) \prod_{i=1}^{n} P(x_i | C_2)} > 1 \quad \rightarrow \quad \text{predict } C_1
\]

\[
\prod_{i=1}^{n} \frac{P(x_i | C_1)}{P(x_i | C_2)} > \frac{P(C_2)}{P(C_1)} \quad \rightarrow \quad \text{predict } C_1
\]
Naïve Bayes Classifier

\[ P(C_k|x_1, x_2, ..., x_n) = P(C_k) \prod_{i=1}^{n} P(x_i|C_k) \]

\[ P(C_1|x_1, x_2, ..., x_n) > P(C_2|x_1, x_2, ..., x_n) \quad \rightarrow \quad \text{predict } C_1 \]

\[ P(C_1) \prod_{i=1}^{n} P(x_i|C_1) > P(C_2) \prod_{i=1}^{n} P(x_i|C_2) \quad \rightarrow \quad \text{predict } C_1 \]

\[ \prod_{i=1}^{n} \frac{P(x_i|C_1)}{P(x_i|C_2)} > \frac{P(C_2)}{P(C_1)} \quad \rightarrow \quad \text{predict } C_1 \]

\[ P(\text{Strong} | \text{Yes})/P(\text{Strong} | \text{No}) > P(\text{No})/P(\text{Yes}) \quad \rightarrow \quad \text{predict Yes} \]
Naïve Bayes Classifier

1. Estimate from data:
   \[ P(\text{Class} \mid X_1, X_2, X_3, \ldots) \]

2. For a given instance \((X_1, X_2, X_3, \ldots)\) predict class whose conditional probability is greater:

   \[ P(C_1 \mid X_1, X_2, X_3, \ldots) > P(C_2 \mid X_1, X_2, X_3, \ldots) \implies \text{predict } C_1 \]

Use Bayes’ rule to calculate

Assume \(X_1, X_2, X_3, \ldots\) are independent
We have a new day:
Outlook = Rain
Temp = Mild
Humidity = Normal
Windy = Strong

Shall we play tennis?
We have a new day:
Outlook=Rain; Temp=Mild; Humidity=Normal; Wind=Strong

\[
\frac{P(\text{Rain}|\text{Yes}) \times P(\text{Mild}|\text{Yes}) \times P(\text{Normal}|\text{Yes}) \times P(\text{Strong}|\text{Yes})}{P(\text{Rain}|\text{No}) \times P(\text{Mild}|\text{No}) \times P(\text{Normal}|\text{No}) \times P(\text{Strong}|\text{No})} > \frac{P(\text{No})}{P(\text{Yes})}
\]
<table>
<thead>
<tr>
<th>Day</th>
<th>Outlook</th>
<th>Temp</th>
<th>Humidity</th>
<th>Wind</th>
<th>PlayTennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>Sunny</td>
<td>Hot</td>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>D2</td>
<td>Sunny</td>
<td>Hot</td>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>D3</td>
<td>Overcast</td>
<td>Hot</td>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>D4</td>
<td>Rain</td>
<td>Mild</td>
<td>High</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>D5</td>
<td>Rain</td>
<td>Cool</td>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>D6</td>
<td>Rain</td>
<td>Cool</td>
<td>Normal</td>
<td>Strong</td>
<td>No</td>
</tr>
<tr>
<td>D7</td>
<td>Overcast</td>
<td>Cool</td>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>D8</td>
<td>Sunny</td>
<td>Mild</td>
<td>High</td>
<td>Weak</td>
<td>No</td>
</tr>
<tr>
<td>D9</td>
<td>Sunny</td>
<td>Cool</td>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>D10</td>
<td>Rain</td>
<td>Mild</td>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>D11</td>
<td>Sunny</td>
<td>Mild</td>
<td>Normal</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>D12</td>
<td>Overcast</td>
<td>Mild</td>
<td>High</td>
<td>Strong</td>
<td>Yes</td>
</tr>
<tr>
<td>D13</td>
<td>Overcast</td>
<td>Hot</td>
<td>Normal</td>
<td>Weak</td>
<td>Yes</td>
</tr>
<tr>
<td>D14</td>
<td>Rain</td>
<td>Mild</td>
<td>High</td>
<td>Strong</td>
<td>No</td>
</tr>
</tbody>
</table>

We have a new day: Outlook=Rain; Temp=Mild; Humidity=Normal; Wind=Strong

\[
\prod_{i=1}^{n} \frac{P(X_i|C_1)}{P(X_i|C_2)} > \frac{P(C_2)}{P(C_1)} \rightarrow \text{predict } C_1
\]