ANSWERS

Exercise 1 ANSWER

To estimate the probability of the game going ahead under the given conditions, we use the Naïve Bayes formula, with all the probabilities read out of the table, since all the attributes are categorical.

We denote the entire set of given weather attribute values with a single letter, E (for 'evidence'). We estimate the required conditional and a priori probabilities from the table and substitute them into the Naïve Bayes formula:

$$\begin{aligned}
P(Yes|E) &= \frac{P(E|Yes)P(Yes)}{P(E)} \\
&= \frac{P(\mathit{Rainy}|Yes)P(\mathit{Mild}|Yes)P(\mathit{Normal}|Yes)P(\mathit{True}|Yes)P(Yes)}{P(E)} \\
&= \frac{\frac{3}{9} \times \frac{4}{9} \times \frac{6}{9} \times \frac{3}{9} \times \frac{9}{14}}{P(E)} = \frac{0.0212}{P(E)}
\end{aligned}$$

We repeat this calculation for Play=No:

$$\begin{aligned}
P(No|E) &= \frac{P(E|No)P(No)}{P(E)} \\
&= \frac{P(\mathit{Rainy}|No)P(\mathit{Mild}|No)P(\mathit{Normal}|No)P(\mathit{True}|No)P(No)}{P(E)} \\
&= \frac{\frac{2}{5} \times \frac{2}{5} \times \frac{1}{5} \times \frac{3}{5} \times \frac{5}{14}}{P(E)} = \frac{0.0069}{P(E)}
\end{aligned}$$

To find the probability of Play=Yes we don't need to estimate P(E) but use:

$$P(Yes|E) + P(No|E) = 1 => \frac{0.0212}{P(E)} + \frac{0.0069}{P(E)} = 1 => P(E) = 0.0212 + 0.0069$$

We calculate the required probability estimate:

$$P(Yes|E) = \frac{0.0212}{0.0212 + 0.0069} = \mathbf{0.75}$$

It is quite probable that the game will go ahead.

Exercise 2 ANSWER

To calculate the required probability estimate, we need to use two different methods for estimating conditional probabilities: for the categorical variables we use value counts from the table, while for the numeric variables we use the formula for probability density of the Gaussian (normal) distribution.

The Gaussian probability density calculation, using the formula $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2}$, requires the mean and the standard deviation of the conditional distribution for the relevant variable, so we find those first:

The values for distribution of `Temperature` conditional on `Play=Yes`:

$$\bar{t}(Yes) = \frac{83+70+68+64+69+75+75+72+81}{9} = 73$$

$$\sigma_t(Yes) = \left\{\frac{1}{9}\left[(83-73)^2 + (70-73)^2 + (68-73)^2 + (64-73)^2 + (69-73)^2 + (75-73)^2 + (75-73)^2 + (72-73)^2 + (81-73)^2\right]\right\}^{\frac{1}{2}} = 5.81$$

Distribution of `Temperature` conditional on `Play=No`:

$$\bar{t}(No) = \frac{85+80+65+72+71}{5} = 74.6$$

$$\sigma_t(No) = \left\{\frac{1}{5}\left[(85-74.6)^2 + (80-74.6)^2 + (65-74.6)^2 + (72-74.6)^2 + (71-74.6)^2\right]\right\}^{\frac{1}{2}} = 7.06$$

Distribution of `Humidity` conditional on `Play=Yes`:

$$\bar{h}(Yes) = \frac{86+96+80+65+70+80+70+90+75}{9} = 79.11$$

$$\sigma_h(Yes) = \left\{\frac{1}{9}\left[(86-79.11)^2 + (96-79.11)^2 + (80-79.11)^2 + (65-79.11)^2 + (70-79.11)^2 + (80-79.11)^2 + (70-79.11)^2 + (90-79.11)^2 + (75-79.11)^2\right]\right\}^{\frac{1}{2}} = 9.63$$

Distribution of `Humidity` conditional on `Play=No`:

$$\bar{h}(No) = \frac{85+90+70+95+91}{5} = 86.2$$

$$\sigma_h(No) = \left\{\frac{1}{5}\left[(85-86.2)^2 + (90-86.2)^2 + (70-86.2)^2 + (95-86.2)^2 + (91-86.2)^2\right]\right\}^{\frac{1}{2}} = 8.70$$

Now that we have the mean and standard deviation values prepared, we can use a formula that combines simple probability and Gaussian probability density to calculate the required probability estimate. The letter E is used to denote the 'evidence' i.e. the set of given attribute values. Note that the first two calculations are not for an actual probability but for a probability density over a two dimensional attribute space (we use f_M to denote the multi-dimensionality of the density's nominator). However, this is of no importance as it is the ratios of these values that are eventually used to determine the probabilities.

$$
\begin{aligned}
f_M(Yes|E) &= \frac{f_M(E|Yes)P(Yes)}{P(E)} \\
&= \frac{P(Rainy|Yes)f(Temp. = 80|Yes)f(Hum. = 80|Yes)P(True|Yes)P(Yes)}{P(E)} \\
&= \frac{\frac{3}{9} \times \frac{1}{5.81\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{80-73}{5.81}\right)^2} \times \frac{1}{7.06\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{80-74.6}{7.06}\right)^2} \times \frac{3}{9} \times \frac{9}{14}}{P(E)} \\
&= \frac{0.000098}{P(E)}
\end{aligned}
$$

We repeat this calculation for Play=No:

$$
\begin{aligned}
f_M(No|E) &= \frac{f_M(E|No)P(No)}{P(E)} \\
&= \frac{P(Rainy|No)f(Temp. = 80|No)f(Hum. = 80|No)P(True|No)P(No)}{P(E)} \\
&= \frac{\frac{2}{5} \times \frac{1}{5.81\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{80-73}{5.81}\right)^2} \times \frac{1}{7.06\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{80-74.6}{7.06}\right)^2} \times \frac{3}{5} \times \frac{5}{14}}{P(E)} \\
&= \frac{0.000129}{P(E)}
\end{aligned}
$$

To find the probability of Play=Yes we don't need to estimate P(E) but use:

$$
P(Yes|E) + P(No|E) = 1 => \frac{0.000098}{P(E)} + \frac{0.000129}{P(E)} = 1 => P(E) = 0.000098 + 0.000129
$$

We calculate the required probability estimate:

$$
P(Yes|E) = \frac{0.000098}{0.000098 + 0.000129} = \mathbf{0.43}
$$

It is somewhat more probable that the game will not go ahead.