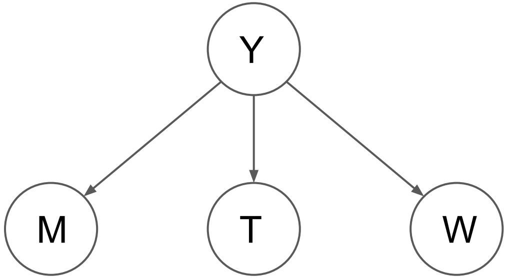
Naive Bayes

Pacman has developed a hobby of fishing. Over the years, he has learned that a day can be considered fit or unfit for fishing Y which results in three features: whether or not Ms. Pacman can show up M, the temperature of the day T, and how high the water level is W. Pacman models it as the following Naive Bayes classification problem,shown on the right:



We wish to calculate the probability a day is fit for fishing given features of the day. Consider the conditional probability tables that Pacman has estimated over the years:

Y	P(Y)
yes	0.1
no	0.9

M	Y	P(M|Y)
yes	yes	0.5
no	yes	0.5
yes	no	0.2
no	no	0.8

W	Y	P(W|Y)
high	yes	0.1
low	yes	0.9
high	no	0.5
low	no	0.5

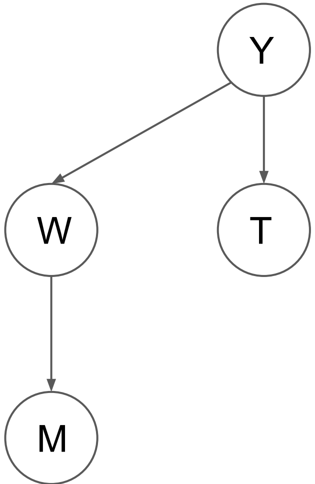
T	Y	P(T|Y)
cold	yes	0.2
warm	yes	0.3
hot	yes	0.5
cold	no	0.5
warm	no	0.2
hot	no	0.3

(a) Using the method of Naive Bayes, if Ms. Pacman is available, the weather is cold, and the water level is high, do we predict that the day is fit for fishing?

(b) Assume Pacman now has the underlying Bayes Net model, and the conditional probability tables from the previous parts do not apply. Recall that predictions are made under the Naive Bayes classification using the conditional probability, P(Y|W,T,M), and the Naive Bayes assumption that features are independent given the class. We wish to explore if the Naive Bayes model is guaranteed to be able to represent the true distribution.

(i) Does the Naives Bayes modeling assumption hold?

(ii) Can the Naive Bayes model represent the true conditional distribution P(Y|W,T,M)?



Answers

a. We will first calculate the fit / not fit probability for the given conditions, namely P(Y=yes | M=yes, W=high, T=cold), P(Y=no | M=yes, W=high, T=cold).

P(Y=yes | M=yes, W=high, T=cold)

= P(Y=yes, M=yes, W=high, T=cold) / P(M=yes, W=high, T=cold)

= P(M=yes|Y=yes) P(W=high|Y=yes) P(T=cold|Y=yes)P(Y=yes) / $\Sigma_{y \in \{yes,\ no\}}$(P(M=yes|Y=y) P(W=high|Y=y) P(T=cold|Y=y)P(Y=y))

= 0.5x0.1x0.2x0.1 / (0.5x0.1x0.2x0.1 + 0.2x0.5x0.5x0.9) = 0.022

Similarly, P(Y=no | M=yes, W=high, T=cold) = P(M=yes|Y=no) P(W=high|Y=no) P(T=cold|Y=no)P(no) / $\Sigma_{y \in \{yes,\ no\}}$(P(M=yes|Y=y) P(W=high|Y=y) P(T=cold|Y=y)P(Y=y)) = 0.978

Since P(Y=no | M=yes, W=high, T=cold) > P(Y=yes | M=yes, W=high, T=cold). So it’s not fit for fishing.

b.

(i) No. W is not independent of M given the class Y.

(ii) Yes. The model does represent the true conditional model P(Y|W, M, T). It is because M is not required to model the distribution here. Hence we can use P(Y|W, T) to represent the true conditional probability P(Y|W, M, T). To see why this is the case, based on the given model, we have:

$$
\begin{aligned}
P(Y|W,M,T) &= \frac{P(W,M,T,Y)}{\sum_y P(W,M,T|Y=y)P(Y=y)} \\
&= \frac{P(W,M|Y)P(T|Y)P(Y)}{\sum_y P(W,M,T|Y=y)P(Y=y)} \\
&= \frac{P(M|W)P(W|Y)P(T|Y)P(Y)}{\sum_y P(M|W)P(W|Y=y)P(T|Y=y)P(Y=y)} \\
&= \frac{\cancel{P(M|W)}P(W|Y)P(T|Y)P(Y)}{\cancel{P(M|W)}\sum_y P(W|Y=y)P(T|Y=y)P(Y=y)} \\
&= \frac{P(W|Y)P(T|Y)P(Y)}{\sum_y P(W|Y=y)P(T|Y=y)P(Y=y)} \\
&= P(Y|W,T)
\end{aligned}
$$