CS 8725: Report for assignment 2

Chanmann Lim

September 9, 2015

1. For $Y \in \{T, F\}$, one parameter is needed to describe $P(Y)$, two parameters are needed to describe $P(X_1|Y)$ and four parameters (μ_{iY}, σ_{iY}) are needed to describe $P(X_i|Y)$ for $2 \leq i \leq d$.

$P(Y)$
$P(X_1|Y), \quad Y \in \{T, F\}, \ X_1 \in \{T, F\}$
$P(X_i|Y) \sim N(\mu_{iY}, \sigma_{iY}^2), \quad Y \in \{T, F\}, \ 2 \leq i \leq d$
The total number of parameters $= 1 + 2 + 4 \times (d - 1) = 4d - 1$.

$$P(Y|X) = \frac{P(X|Y) \cdot P(Y)}{P(X)} \tag{1}$$

$$\propto P(X|Y) \cdot P(Y) \tag{2}$$

$$= P(X_1|Y) \cdot \prod_{i=2}^{d} N(x_i; \mu_{iY}, \sigma_{iY}^2) \cdot P(Y) \tag{3}$$

Where, $N(x_i; \mu_{iY}, \sigma_{iY}^2) = \dfrac{1}{\sqrt{2\pi}\sigma_{iY}} e^{-\frac{(x_i - \mu_{iY})^2}{2\sigma_{iY}}}$

2. (a) Naive Bayes decision rule, like Optimal Bayes decision rule, is the function that maximizes the posterior probability while assuming all features are independent given class. However the denominator of the posterior probability is merely the normalization term to satisfy the property of a probability and it does not depend on y thus it can be dropped.

$$f_{NB}(Sunny, Windy) = \underset{y}{\operatorname{argmax}} \, P(Sunny|y) \cdot P(Windy|y) \cdot P(y) \tag{4}$$

Where $y \in \{Hike, \neg Hike\}$ and
$$P(Hike) = P(\neg Hike) = 0.5$$

$P(y)$ can also be dropped.

$$f_{NB}(Sunny, Windy) = \underset{y}{\operatorname{argmax}} \, P(Sunny|y) \cdot P(Windy|y) \tag{5}$$

Therefore, predicting Alice and Bob go hiking if $P(Sunny|Hike) \times P(Windy|Hike) > P(Sunny|\neg Hike) \times P(Windy|\neg Hike)$ and else otherwise.

(b) The joint probability of Alice and Bob go hiking and the weather is sunny and windy can be considered as the product of the likelihood of the weather given they go hiking $P(Sunny, Windy|Hike)$ and the prior probability of going hiking $P(Hike)$.

$$P(Sunny, Windy, Hike) = P(Sunny, Windy|Hike) \cdot P(Hike) \tag{6}$$

$$= P(Sunny|Hike) \cdot P(Windy|Hike) \cdot P(Hike) \tag{7}$$

$$= 0.8 \times 0.4 \times 0.5 \tag{8}$$

$$= 0.16 \tag{9}$$

Similarly,

$$P(Sunny, Windy, \neg Hike) = P(Sunny|\neg Hike) \cdot P(Windy|\neg Hike) \cdot P(\neg Hike) \tag{10}$$
$$= 0.7 \times 0.5 \times 0.5 \tag{11}$$
$$= 0.175 \tag{12}$$

Using the above Naive Bayes decision rule, we will predict $\neg Hike$ when it's sunny and windy.

The prediction of Alice and Bob go hiking when the weather is sunny and windy $P(Hike|Sunny, Windy)$ is equal to the probability that they actually go hiking and it is sunny and windy $P(Sunny, Windy, Hike)$ divided by the probability that it's sunny and windy regardless of hiking $P(Sunny, Windy)$.

$$P(Hike|Sunny, Windy) = \frac{P(Sunny, Windy, Hike)}{P(Sunny, Windy)}$$

Therefore the probability that the weather is sunny and windy and Alice and Bob isn't go hiking $P(Sunny, Windy, \neg Hike)$ is in fact the probability of prediction error of $P(Hike|Sunny, Windy)$.

$$P_e(Hike|Sunny, Windy) = P(Sunny, Windy, \neg Hike) \tag{13}$$
$$= 0.175 \tag{14}$$

(c) From (b) we can derive the expected error rate by summing up the prediction errors in all cases.

Sunny	Windy	$P(Sunny, Windy|Hike)$	$P(Sunny, Windy|\neg Hike)$	Predict	P_e
T	T	$0.8 \times 0.4 = 0.32$	$0.7 \times 0.5 = 0.35$	$\neg Hike$	0.32×0.5
T	F	$0.8 \times 0.6 = 0.48$	$0.7 \times 0.5 = 0.35$	$Hike$	0.35×0.5
F	T	$0.2 \times 0.4 = 0.08$	$0.3 \times 0.5 = 0.15$	$\neg Hike$	0.08×0.5
F	F	$0.2 \times 0.6 = 0.12$	$0.3 \times 0.5 = 0.15$	$\neg Hike$	0.12×0.5

$$E = 0.5 \times (0.32 + 0.35 + 0.08 + 0.12) \tag{15}$$
$$= 0.435 \tag{16}$$