Exercises for Chapter 8, Part 1

8.1 Consider the data sets for two classes $X_1 = \{(0,0)\}$ and $X_2 = \{(1,0),(0,1)\}$.

a) Which classification probabilities will a naive Bayes classifier produce for the feature vector $(0,0)$?

The class probabilities are $p(1) = 1/3$ and $p(2) = 2/3$. For class 1, the probabilities of 0 for the first or second features are $p((0,?) \mid 1) = 1/1 = 1$ and $p((?,0) \mid 1) = 1/1 = 1$, so $p((0,0) \mid 1) = 1 \cdot 1 = 1$. For class 2, the probabilities of 0 for the first or second features are $p((0,?) \mid 2) = 1/2$ and $p((?,0) \mid 2) = 1/2$, so $p((0,0) \mid 2) = 1/2 \cdot 1/2 = 1/4$. With the Bayes rule we obtain

$$p(1 \mid (0,0)) = \frac{p(1) \cdot p((0,0) \mid 1)}{p(1) \cdot p((0,0) \mid 1) + p(2) \cdot p((0,0) \mid 2)} = \frac{\frac{1}{3} \cdot 1}{\frac{1}{3} \cdot 1 + \frac{2}{3} \cdot \frac{1}{4}} = \frac{2}{3}$$

$$p(2 \mid (0,0)) = \frac{p(2) \cdot p((0,0) \mid 1)}{p(1) \cdot p((0,0) \mid 1) + p(2) \cdot p((0,0) \mid 2)} = \frac{\frac{2}{3} \cdot \frac{1}{4}}{\frac{1}{3} \cdot 1 + \frac{2}{3} \cdot \frac{1}{4}} = \frac{1}{3}$$

8.2 Consider a classifier with one-dimensional input $x \in \mathbb{R}$, where the data for the positive and the negative class follow the Cauchy distributions

$$p(x \mid \oplus) = \frac{1}{\pi} \cdot \frac{1}{1 + (x+1)^2} \quad \text{and} \quad p(x \mid \ominus) = \frac{1}{\pi} \cdot \frac{1}{1 + (x-1)^2}$$

a) For 50% positive and 50% negative training data, for which x will the naive Bayes classifier yield "positive"?

At the border between the positive and negative classes, both probabilities are equal, so $p(\oplus) \cdot p(x \mid \oplus) = p(\ominus) \cdot p(x \mid \ominus) \quad \Rightarrow \quad \frac{1}{2\pi} \cdot \frac{1}{1+(x+1)^2} = \frac{1}{2\pi} \cdot \frac{1}{1+(x-1)^2} \quad \Rightarrow \quad 1 + (x+1)^2 = 1 + (x-1)^2 \quad \Rightarrow \quad x^2 + 2x + 2 = x^2 - 2x + 2 \quad \Rightarrow \quad x = 0$. Now the question is if we get "positive" for $x < 0$ or for $x > 0$. This can be easily found by e.g. looking at the maximum of the probability $p(x \mid \oplus)$ of the positive distribution, which is at $x = -1$, so we get the positive class for all $x < 0$.

b) For which percentage p of positive training data will the naive Bayes classifier *always* yield "positive"? If needed, assume $\sqrt{2} \approx 1.4$.

For $p(\oplus) = p$ and $p(\ominus) = 1 - p$, at the border between the classes we have $(1-p)(x^2 + 2x + 2) = p(x^2 - 2x + 2) \quad \Rightarrow \quad (2p-1)x^2 - 2x + 4p - 2 = 0 \quad \Rightarrow \quad x^2 - \frac{2}{2p-1}x + 2 = 0 \quad \Rightarrow \quad x = \frac{1}{2p-1} \pm \sqrt{\frac{1}{(2p-1)^2} - 2}$

The limit to obtain a unique solution (always the same class) is obtained when $\sqrt{\ldots} = 0 \quad \Rightarrow \quad \frac{1}{2p-1} = \pm\sqrt{2} \quad \Rightarrow \quad p = \frac{1}{2} \pm \frac{\sqrt{2}}{4} \approx 15\%, 85\%$
Since p corresponds to positive training data, the classifier will always yield "positive" for $p > 85\%$.

c) For an arbitrary percentage p of positive training data, for which x will the naive Bayes classifier yield "positive"?

$p < 15\%$: never "positive"
$15\% \leq p < 50\%$: "positive" for

$$x > \frac{1}{2p-1} - \sqrt{\frac{1}{(2p-1)^2} - 2} \quad \text{and} \quad x < \frac{1}{2p-1} + \sqrt{\frac{1}{(2p-1)^2} - 2}$$

$p = 50\%$: "positive" for $x < 0$
$50\% < p \leq 85\%$: "positive" for

$$x < \frac{1}{2p-1} - \sqrt{\frac{1}{(2p-1)^2} - 2} \quad \text{or} \quad x > \frac{1}{2p-1} + \sqrt{\frac{1}{(2p-1)^2} - 2}$$

$p > 85\%$: always "positive"