Seminar 7

Definition 1 (Naive Bayes Classifier)
Naive Bayes (NB) Classifier assumes that the effect of the value of a predictor x on a given class c is class conditional independent. Bayes theorem provides a way of calculating the posterior probability $P(c|x)$ from class prior probability $P(c)$, predictor prior probability $P(x)$ and probability of the predictor given the class $P(x|c)$

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

and for a vector of predictors $X = (x_1, \dots, x_n)$

$$P(c|X) = \frac{P(x_1|c) \dots P(x_n|c)P(c)}{P(x_1) \dots P(x_n)}.$$

The class with the highest posterior probability is the outcome of prediction.

Exercise 7/1

What is naive about Naive Bayes classifier? Briefly outline its major idea.

Answers can vary. For official definition refer to the Manning book.

Exercise 7/2

Considering the table of observations, use the Naive Bayes classifier to recommend whether to *Play Golf* given a day with *Outlook = Rainy*, *Temperature = Mild*, *Humidity = Normal* and *Windy = True*. Do not deal with the zero-frequency problem.

Outlook	Temperature	Humidity	Windy	Play Golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Overcast	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Cool	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Sunny	Mild	High	True	No

Table 1: Exercise.

First build the likelihood tables for each predictor

		Play Golf		
		Yes	No	
Outlook	Sunny	$3/9$	$2/5$	$5/14$
	Overcast	$4/9$	$0/5$	$4/14$
	Rainy	$2/9$	$3/5$	$5/14$
		$9/14$	$5/14$	

		Play Golf		
		Yes	No	
Temperature	Hot	$2/9$	$2/5$	$4/14$
	Mild	$4/9$	$2/5$	$6/14$
	Cool	$3/9$	$1/5$	$4/14$
		$9/14$	$5/14$	

		Play Golf		
		Yes	No	
Humidity	High	$3/9$	$4/5$	$7/14$
	Normal	$6/9$	$1/5$	$7/14$
		$9/14$	$5/14$	

		Play Golf		
		Yes	No	
Windy	True	$3/9$	$2/5$	$5/14$
	False	$6/9$	$3/5$	$9/14$
		$9/14$	$5/14$	

We see that probability of *Sunny* given *Yes* is $3/9 = 0.33$, probability of *Sunny* is $5/14 = 0.36$ and probability of *Yes* is $9/14 = 0.64$. Then we count the likelihoods of *Yes* and *No*

$$
\begin{aligned}
&P(Yes|Rainy, Mild, Normal, True) = \\
&\quad = P(Rainy|Yes) \cdot P(Mild|Yes) \cdot P(Normal|Yes) \cdot P(True|Yes) \cdot P(Yes) \\
&\quad = \frac{2}{9} \cdot \frac{4}{9} \cdot \frac{6}{9} \cdot \frac{3}{9} \cdot \frac{9}{14} = 0.014109347 \\
&P(No|Rainy, Mild, Normal, True) = \\
&\quad = P(Rainy|No) \cdot P(Mild|No) \cdot P(Normal|No) \cdot (True|No) \cdot P(No) \\
&\quad = \frac{3}{5} \cdot \frac{2}{5} \cdot \frac{1}{5} \cdot \frac{3}{5} \cdot \frac{5}{14} = 0.010285714
\end{aligned}
\tag{1}
$$

and suggest *Yes*. We can normalize the likelihoods to obtain the % confidence:

$$P(Yes|Rainy, Mild, Normal, True) = \frac{0.014109347}{0.014109347 + 0.010285714} = 57.84\%$$

$$P(No|Rainy, Mild, Normal, True) = \frac{0.010285714}{0.014109347 + 0.010285714} = 42.16\%$$

Definition 2 (A Linear Classifier)
Our linear classifier finds the hyperplane that bisects and is perpendicular to the connecting line of the closest points from the two classes. The separating (decision) hyperplane is defined in terms of a normal (weight) vector $\mathbf{w}$ *and a scalar intercept term* b *as*

$$f(x) = \mathbf{w} \cdot \mathbf{x} + b$$

where $\cdot$ *is the dot product of vectors. Finally, the classifier becomes*

$$class(x) = sgn(f(x)).$$

Exercise 7/3

Draw a sketch explaining the concept of our linear classifier. Include the equation of the separation hyperplane. Is our classifier equivalent to support vector machines (SVM)? What are limitations of our classifier?

Answers can vary. For official definition refer to the Manning book.

Exercise 7/4

Build a linear classifier for the training set $\{([1,1],-1),([2,0],-1),([2,3],+1)\}$.

We first take the closest two points from the respective classes: $[1,1]$ and $[2,3]$. We have $\mathbf{w} = a \cdot ([1,1]-[2,3]) = [a,2a]$. Now we calculate a and b

$$a + 2a + b = -1$$

$$2a + 6a + b = 1$$

for the points $[1,1]$ and $[2,3]$, respectively. The solution is

$$a = \frac{2}{5} \qquad b = \frac{-11}{5}$$

building the weight vector

$$\mathbf{w} = \left[\frac{2}{5}, \frac{4}{5}\right]$$

and the final classifier becomes

$$class(x) = sgn\left(\frac{2}{5}x_1 + \frac{4}{5}x_2 - \frac{11}{5}\right).$$

Exercise 7/5

Explain the concept of classification based on neural networks. Draw a sketch and comment on all components.

Answers can vary. For official definition refer to the Manning book.

Exercise 7/6

What is the difference between supervised and unsupervised learning? Give examples.

Answers can vary. For official definition refer to the Manning book.