COMS 4701 Artificial Intelligence
Sec 01

HW 3

Name:	Yu Wang
Uni:	yw3025
Date:	April. 15, 2018

Question 1: Learning Methods

1. **Learning algorithm for predicting rain**

I choose to use **linear regression**

1) Since rainfall is continuous (continuous output in R), we need a repressor, so linear regression is suitable for this problem.
2) Since linear regression is a supervised learning algorithm. In the training dataset, we assume we have the rainfall given the currents and tides in the Atlantic Ocean in the past six month data.

2. **Multiple choice: boosting**

pick **c**

a: give large weights to hard example

b: slightly better than random

3. **Multiple choice: K-means**

Pick **b and c**

a: we don't assume fit normal distribution

d, e: high intra-cluster similarity and a low inter-cluster similarity.

Question 2: Naïve Bayes

1. **Build a naive Bayes classifier using simple probabilities**

1) Priors:

$$p(boy) = \frac{3}{8}$$

$$p(girl) = \frac{5}{8}$$

2) Conditional probability:

$$p(Tyler|boy) = \frac{1}{3}$$

$$p(John|boy) = \frac{1}{3}$$

$$p(Ali|boy) = \frac{1}{3}$$

$$p(salma|boy) = 0$$

$$p(Leila|boy) = 0$$

$$p(Alexanda|boy) = 0$$

$$p(Salma|girl) = \frac{1}{5}$$

$$p(Tyler|girl) = \frac{2}{5}$$

$$p(Leila|girl) = \frac{1}{5}$$

$$p(Alexandra|girl) = \frac{1}{5}$$

$$p(John|girl) = 0$$

$$p(Ali|girl) = 0$$

$$p(no|boy) = \frac{1}{3}$$

$$p(yes|boy) = \frac{2}{3}$$

$$p(no|girl) = \frac{3}{5}$$

$$p(yes|girl) = \frac{2}{5}$$

$$p(blue|boy) = \frac{2}{3}$$

$$p(brown|boy) = \frac{1}{3}$$

$$p(blue|girl) = \frac{3}{5}$$

$$p(brown|girl) = \frac{2}{5}$$

$$p(short|boy) = \frac{2}{3}$$

$$p(long|boy) = \frac{1}{3}$$

$$p(short|girl) = \frac{1}{5}$$

$$p(long|girl) = \frac{4}{5}$$

2. Prediction for a " Tall kid named Tyler with long hair and brown eyes"?

Answer: girl

By Bayes algorithm we know

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)}$$

Since the p(x) in the denominator is not important, we can do:

$\underline{\mathrm{p(boy|tall = yes, name = Tyler, hair = long, eye = brown)\ can\ be\ estimate\ by:}}$

$$\boldsymbol{p(boy) * p(tall = yes|boy) * p(name = Tyler|boy) * p(hair = long|boy) * p(eye = brown|boy) = \frac{3}{8} * \frac{2}{3} * \frac{1}{3} * \frac{1}{3} * \frac{1}{3} = \frac{1}{108}}$$

$\underline{\mathrm{p(girl|tall = yes, name = Tyler, hair = long, eye = brown)\ can\ be\ estimate\ by:}}$

$$\boldsymbol{p(girl) * p(tall = yes|girl) * p(name = Tyler|girl) * p(hair = long|girl) * p(eye = brown|girl) = \frac{5}{8} * \frac{2}{5} * \frac{2}{5} * \frac{4}{5} * \frac{2}{5} = \frac{4}{125}}$$

since $\underline{\mathbf{p(girl|tall = yes, name = Tyler, hair = long, eye = brown) > p(boy|tall = yes, name = Tyler, hair = long, eye = brown)}}$ **so it's a girl**

3. Population

In this dataset,

$$p(boy) = \frac{3}{8}$$

$$p(girl) = \frac{5}{8}$$

So we know the sample is imbalanced, it can't reflect the distribution priors in reality.

Since

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)}$$

When we have $p(girl) > p(boy)$, we would more likely to predict a girl if $p(x|girl) == p(x|boy)$(x stands for features in this question, tall, eye…etc). So this dataset has bias.

To solve such problem, we can collect more data(more boy data), or perform up/down sampling. (of course, there is another problem… the dataset is too small for application)

Question 3: Neural Networks

1. Construct a one-hidden layer neural network

First, we assume

$$g(z) = \frac{1}{1 + e^{-z}}$$

$$g(10) = 0.99995 = 1$$

$$g(-10) = 0.0004 = 0$$

And $for\ z \geq 10, g(z) \to 1, for\ z \leq -10, g(z) \to 0$

$(X\ or\ not\ Y)\ XOR\ (not\ Z\ or\ not\ T)$
$= \big((X\ or\ not\ Y) or\ (not\ Z\ or\ not\ T)\big)\ AND\ \big((X\ or\ not\ Y)\ NAND\ (not\ Z\ or\ not\ T)\big)$
$= (not\ X\ and\ Y\ and\ not\ Z)\ or\ (not\ X\ and\ Y\ and\ not\ T)\ or\ (X\ and\ Z\ and\ T)\ or\ (Z\ and\ not\ Y\ and\ T)$

1) For $(X\ and\ Z\ and\ T)$:

X	Z	T	$X\ and\ Z\ and\ T$
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1

$$g(w_0 + w_x + w_z + w_t) = g(10)$$

$$g(w_0) = g(-10)$$

$$g(w_0 + w_x) = g(-10)$$

$$g(w_0 + w_z) = g(-10)$$

$$g(w_0 + w_t) = g(-10)$$

$$g(w_0 + w_x + w_z) = g(-10)$$

$$g(w_0 + w_t + w_z) = g(-10)$$

$$g(w_0 + w_x + w_t) = g(-10)$$

Solving these equation, one possible solution is $w_0 = -50, w_x = 20, w_z = 20, w_t = 20$

2) For $(not\ X\ and\ Y\ and\ not\ Z)$:

X	y	Z	$not\ X\ and\ Y\ and\ not\ Z$
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	0

$$g(w_0) = g(-10)$$
$$g(w_0 + w_x) = g(-10)$$
$$g\left(w_0 + w_y\right) = g(10)$$
$$g(w_0 + w_z) = g(-10)$$
$$g\left(w_0 + w_x + w_y\right) = g(-10)$$
$$g(w_0 + w_x + w_z) = g(-10)$$
$$g\left(w_0 + w_y + w_z\right) = g(-10)$$
$$g\left(w_0 + w_x + w_y + w_z\right) = g(-10)$$

Solving these equation, one possible solution is $w_0 = -10, w_x = -20, w_y = 20, w_z = -20$

3) For $not\ X\ and\ Y\ and\ not\ T$

Same as 2) just change variable, so $w_0 = -10, w_x = -20, w_y = 20, w_t = -20$

4) For $Z\ and\ not\ Y\ and\ T$

Z	Y	T	$Z\ and\ not\ Y\ and\ T$
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0

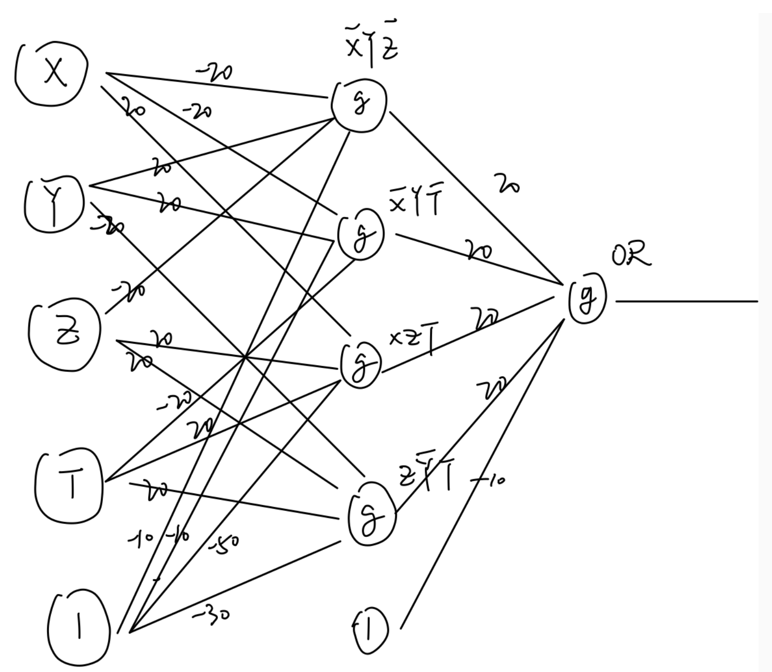
1	0	1	1
1	1	0	0
1	1	1	0

$$g(w_0) = g(-10)$$
$$g(w_0 + w_z) = g(-10)$$
$$g(w_0 + w_y) = g(-10)$$
$$g(w_0 + w_t) = g(-10)$$
$$g(w_0 + w_z + w_y) = g(-10)$$
$$g(w_0 + w_z + w_t) = g(10)$$
$$g(w_0 + w_y + w_t) = g(-10)$$
$$g(w_0 + w_z + w_y + w_t) = g(-10)$$

Solving these equation, one possible solution is $w_0 = -30, w_z = 20, w_y = -20, w_t = 20$

5) For the weights from hidden layer to output layer, we have four ORs, so just setting wights $w_0 = -10, w1 = 20, w2 = 20, w3 = 20, w4 = 20$ will suffice, since only $g(w_0) = g(-10)$

So the **final result** is as follows:



2) **Bonus**

No, it will not improve the quality of the model.

Let's define

$$y_1 = w_1^T x$$
$$y_2 = w_2^T x$$

Then

$$y = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} w_1^T \\ w_2^T \end{bmatrix} x = w'x$$
$$where\ w' = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} w_1^T \\ w_2^T \end{bmatrix}$$

Then we still have $y = w'x$

So, stacking of linear functions still give linear function. The way to solve this problem is to use an activation function such as Sigmoid or Relu.