# COMP9318 Assignment 1

Name: Keshi Chen

ZID: z5142821

**Q1**.

(1)

Location Time Item Quantity

0 Sydney 2005 PS2 1400

1 Sydney 2005 ALL 1400

2 Sydney 2006 PS2 1500

3 Sydney 2006 Wii 500

4 Sydney 2006 ALL 2000

5 Sydney ALL PS2 2900

6 Sydney ALL Wii 500

7 Sydney ALL ALL 3400

8 Melbourne 2005 Xbox 360 1700

9 Melbourne 2005 ALL 1700

10 Melbourne ALL Xbox 360 1700

11 Melbourne ALL ALL 1700

12 ALL 2005 PS2 1400

13 ALL 2005 Xbox 360 1700

14 ALL 2005 ALL 3100

15 ALL 2006 PS2 1500

16 ALL 2006 Wii 500

17 ALL 2006 ALL 2000

18 ALL ALL PS2 2900

19 ALL ALL Wii 500

20 ALL ALL Xbox 360 1700

21 ALL ALL ALL 5100

(2)

select location, time, item, sum(quantity)

from test

group by location, time, item

union

select location, time, ‘ALL’, sum(quantity)

from test

group by location, time

union

select location, ‘ALL’, item, sum(quantity)

from test

group by location, item

union

select location, ‘ALL’, ‘ALL’, sum(quantity)

from test

group by location

union

select ‘ALL’, time, item, sum(quantity)

from test

group by time, item

union

select ‘ALL’, time, ‘ALL’, sum(quantity)

from test

group by time

union

select ‘ALL’, ‘ALL’, item, sum(quantity)

from test

group by item

(3)

Location Time Item Quantity

0 Sydney 2005 PS2 1400

1 Sydney 2005 ALL 1400

2 Sydney 2006 PS2 1500

3 Sydney 2006 ALL 1500

4 Sydney ALL PS2 2900

5 Sydney ALL ALL 2900

6 ALL 2005 PS2 1400

7 ALL 2005 ALL 1400

8 ALL 2006 PS2 1500

9 ALL 2006 ALL 1500

10 ALL ALL PS2 2900

11 ALL ALL ALL 2900

(4) f(Location, Time, Item) = 12\* Location + 4 \* Time + Item

The process can be represented in 3 steps as below:

Step 1:

Location Time Item Quantity

0 1 1 1 1400

1 1 1 0 1400

2 1 2 1 1500

3 1 2 3 500

4 1 2 0 2000

5 1 0 1 2900

6 1 0 3 500

7 1 0 0 3400

8 2 1 2 1700

9 2 1 0 1700

10 2 0 2 1700

11 2 0 0 1700

12 0 1 1 1400

13 0 1 2 1700

14 0 1 0 3100

15 0 2 1 1500

16 0 2 3 500

17 0 2 0 2000

18 0 0 1 2900

19 0 0 3 500

20 0 0 2 1700

21 0 0 0 5100

Step 2: f(Location, Time, Item) = 12\* Location + 4 \* Time + Item

Location Time Item Quantity ArrayIndex

0 1 1 1 1400 17

1 1 1 0 1400 16

2 1 2 1 1500 21

3 1 2 3 500 23

4 1 2 0 2000 20

5 1 0 1 2900 13

6 1 0 3 500 15

7 1 0 0 3400 12

8 2 1 2 1700 30

9 2 1 0 1700 28

10 2 0 2 1700 26

11 2 0 0 1700 24

12 0 1 1 1400 5

13 0 1 2 1700 6

14 0 1 0 3100 4

15 0 2 1 1500 9

16 0 2 3 500 11

17 0 2 0 2000 8

18 0 0 1 2900 1

19 0 0 3 500 3

20 0 0 2 1700 2

21 0 0 0 5100 0

Step 3:

ArrayIndex Quantity

1. 5100
2. 2900
3. 1700
4. 500
5. 3100
6. 1400
7. 1700
8. 2000
9. 1500

11 500

12 3400

13 2900

15 500

16 1400

17 1400

20 2000

21 1500

23 500

24 1700

26 1700

28 1700

30 1700

**Q2**

(1) The Naïve Bayes Classifier: f(x) = argmaxx∈{Cj} ∏­­i=1 P(αi|Cj) \* P(Cj)

Since Cj ∈ {0, 1}, αi∈{0, 1}, f(x) can be represented as:

0, if ∏­­i=1P(αi| C0) \* P(C0) - ∏­­i=1P(αi| C1) \* P(C1) > 0

f(x) =

1, if otherwise

Then with log∏­­i xi = ∑­i log xi, we have:

f(x) = ∑­i log +log

Let wi =log , w0 = log , i ∈ {1, 2, 3, … , d}

Then Naïve Bayes Classifier is equal to a binary linear classifier in d+1 dimensional space with input xi either be 0 or 1.

(2) We have to do partial derivatives to minimize the cost function and gradient descents to get wi in Logistic Regression Classifier, thus its calculation is complex.

In Naïve Bayes Classifier, we only need to calculate each P(xi|C) once to get the wi.

So Naïve Bayes Classifier is much easier than Logistic Regression Classifier.

**Q3**

1. Logistic regression model:

δ(x) = ①

For one training sample (xi , yi ), where yi ∈{0, 1}, we have the probability of yi:

P(yi | xi , w) = δ(x)yi (1 - δ(x))1-yi

Hence when yi = 1, P(yi | xi , w) = δ(x), otherwise P(yi | xi , w) = (1 - δ(x)).

Then we have the log-likelihood:

Log-likelihood = ln(P(yi | xi , w)) = ln(δ(xi)yi (1 - δ(xi))1-yi)

= lnδ(xi)yi + ln(1 - δ(xi))1-yi

= yi lnδ(xi) + (1 - yi) ln(1 - δ(xi))

For N training samples, we have:

Log-likelihood = ∑­i=1 yi lnδ(xi) + (1 - yi) ln(1 - δ(xi)) , i ∈[1, N] ②

And since δ(xi) = = ③

From ② ③we have:

Log-likelihood = ∑­i=1 (yi wTxi - ln(1 +exp(wTxi))) , i ∈[1, N] ④

Since loss function = - log-likelihood, we have:

l(w) = ∑­i=1 (- yi wTxi + ln(1 +exp(wTxi))) , i ∈[1, N]

Q.E.D

1. Similarly, P(yi | xi , w) = f(wTxi)yi (1 - f(wTxi))1-yi

ln P(yi | xi , w) = yi ln f(wTxi) + (1 - yi) (1 - f(wTxi))

For N samples, we have:

Log-likelihood = ∑­i=1 yi ln f(wTxi) + (1 - yi) ln(1 - f(wTxi)) , i ∈[1, N]

Hence the loss function:

l(w) = - log-likelihood = ∑­i=1 - yi ln f(wTxi) - (1 - yi) ln(1 - f(wTxi)) , i ∈[1, N]