**9318----Assignment**

**Jia SUN z5145482**

**Question 1**

1)

|  |  |  |  |
| --- | --- | --- | --- |
| Location | Time | Item | SUM(Quantity) |
| Sydney | 2005 | PS2 | 1400 |
| Sydney | 2005 | ALL | 1400 |
| Sydney | 2006 | PS2 | 1500 |
| Sydney | 2006 | Wii | 500 |
| Sydney | 2006 | ALL | 2000 |
| Sydney | ALL | PS2 | 2900 |
| Sydney | ALL | Wii | 500 |
| Sydney | ALL | ALL | 3400 |
| Melbourne | 2005 | XBox360 | 1700 |
| Melbourne | 2005 | ALL | 1700 |
| Melbourne | ALL | XBox360 | 1700 |
| Melbourne | ALL | ALL | 1700 |
| ALL | 2005 | PS2 | 1400 |
| ALL | 2005 | XBox360 | 1700 |
| ALL | 2005 | ALL | 3100 |
| ALL | 2006 | PS2 | 1500 |
| ALL | 2006 | Wii | 500 |
| ALL | 2006 | ALL | 2000 |
| ALL | ALL | PS2 | 2900 |
| ALL | ALL | Wii | 500 |
| ALL | ALL | XBox360 | 1700 |
| ALL | ALL | ALL | 5100 |

2.)

(SELECT Location, Year, Item, Quantity

FROM Sales)

UNION

(SELECT Location, Year, 'ALL', SUM(Quantity)

FROM Sales

GROUP BY(Location, Year))

UNION

(SELECT Location, 'ALL', Item, SUM(Quantity)

FROM Sales

GROUP BY(Location, Item))

UNION

(SELECT 'ALL', Year, Item, SUM(Quantity)

FROM Sales

GROUP BY(Year, Year))

UNION

(SELECT Location, 'ALL', 'ALL', SUM(Quantity)

FROM Sales

GROUP BY(Location))

UNION

(SELECT 'ALL', Year, 'ALL', SUM(Quantity)

FROM Sales

GROUP BY(Year))

UNION

(SELECT 'ALL', 'ALL', Item, SUM(Quantity)

FROM Sales

GROUP BY(Item))

UNION

(SELECT 'ALL', 'ALL', 'ALL', SUM(Quantity)

FROM Sales)

3)

|  |  |  |  |
| --- | --- | --- | --- |
| Location | Time | Item | SUM(Quantity) |
| Sydney | 2006 | ALL | 2000 |
| Sydney | ALL | PS2 | 2900 |
| ALL | 2005 | ALL | 3100 |
| ALL | 2006 | ALL | 2000 |
| ALL | ALL | PS2 | 2900 |
| ALL | ALL | ALL | 5100 |

4)

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

|  |  |
| --- | --- |
| SUM(Quantity) | offset |
| 5100 | 0 |
| 2900 | 1 |
| 1700 | 2 |
| 500 | 3 |
| 3100 | 4 |
| 1400 | 5 |
| 1700 | 6 |
| 2000 | 8 |
| 1500 | 9 |
| 500 | 11 |
| 3400 | 12 |
| 2900 | 13 |
| 500 | 15 |
| 1400 | 16 |
| 1400 | 17 |
| 2000 | 20 |
| 1500 | 21 |
| 500 | 23 |
| 1700 | 24 |
| 1700 | 26 |
| 1700 | 28 |
| 1700 | 30 |

**Question 2**

1)

By multinomial Naïve Bayes functions:  
①P( = , = , = ,……, = ) = y{0, 1}

Log likelihood and we can get the decision boundary:

②=0

By combining ① and ②

Therefore, we can get:

We can get easily

Then, we can get a linear form() for multinomial Naïve Bayes Classifier.

And can be got by the equation directly.

2)

For Logistic Regression

y =

P() =

P() = 1 - P()

Therefore,

P() =

Likelihood : L(w) =

The log – Likelihood is

(w) = log()l

=

=

=

=

So, we can get

Obviously, we cannot get directly.

To get maximum likelihood, we can use gradient descent and iterate many times to re-weight.

We also need a appropriate learning rate to make = - is easy to learn.

**Question 3**

1)

By using logistic regression:

Likelihood : L(w) =

The log – Likelihood is

(w) =

Because the loss function,

L(w) = -]

= -

= -

As we know , So

(w) = -

= -

= -

=-

=

2)

By using logistic regression:

Likelihood : L(w) =

The log – Likelihood is

(w) =

Due to the loss function,

L(w) = -]

= -

= -

Because

So the deduced loss function is -