結構化機器學習模型及其應用

第一次報告

系所:應數所大數據組
學生:張格恩
學號:7106053112

一、Representation

c_{11}	c_{12}	c_{21}	c_{22}	a_1	b_1	a_2	b_2
E	2	A	C	1	0	6	2
5	B	6	3	1	0	6	2

……

a_{15}	b_{15}	a_{16}	b_{16}
4	4	E	E
4	4	E	E

Feature(x) :

$x = [a_1, b_1, \ldots, a_{10}, b_{10}, a_{13}, b_{13}, \ldots, a_{16}, b_{16}]$ (28 Dim) ， a_i , b_i =0~E $\forall i$

$x \longrightarrow [x_1, x_2, \ldots, x_{111}, x_{112}]$ (112 Dim) ， $x_i = 0 \; or \; 1 \quad \forall i$

Label($y = [y_1, y_2]$) :

$y = [y_1, y_2] = [c_{11}, c_{12}, c_{21}, c_{22}] \longrightarrow y = 16$ bit$(000\cdots 000 \sim 111\cdots 111)$

$y \longrightarrow (y_0, y_1, \ldots, y_{14}, y_{15})$ $y_i = 0$ or 1 $\forall i$=0~15

一、Representation

Feature(x) :

$x = [a_1, b_1, \ldots, a_{10}, b_{10}, a_{13}, b_{13}, \ldots, a_{16}, b_{16}]$ (28 Dim) ， a_i , b_i =0~E $\forall i$

$x \longrightarrow [x_1, x_2, \ldots, x_{111}, x_{112}]$ (112 Dim) ， $x_i = 0 \; or \; 1 \quad \forall i$

Reason:

(1)增加bit of feature的相關性

(2)降低feature distribution的variance

ex: 在Neural Network裡

x^1 = [F, 1, 6] $\longrightarrow w_1 \times 15 + w_2 \times 1 + w_3 \times 6$

$x^1 \longrightarrow x^1_{new}$ = [1, 1, 1, 1, …, 0, 1, 1, 0] $\longrightarrow w_1 \times 1 + w_2 \times 1 + \cdots + w_{11} \times 1 + w_{12} \times 0$

一、Representation

Label($y = [y_1, y_2]$) :

$y = [y_1, y_2] = [c_{11}, c_{12}, c_{21}, c_{22}] \longrightarrow y = 16\ \text{bit}(000\cdots 000 \sim 111\cdots 111)$

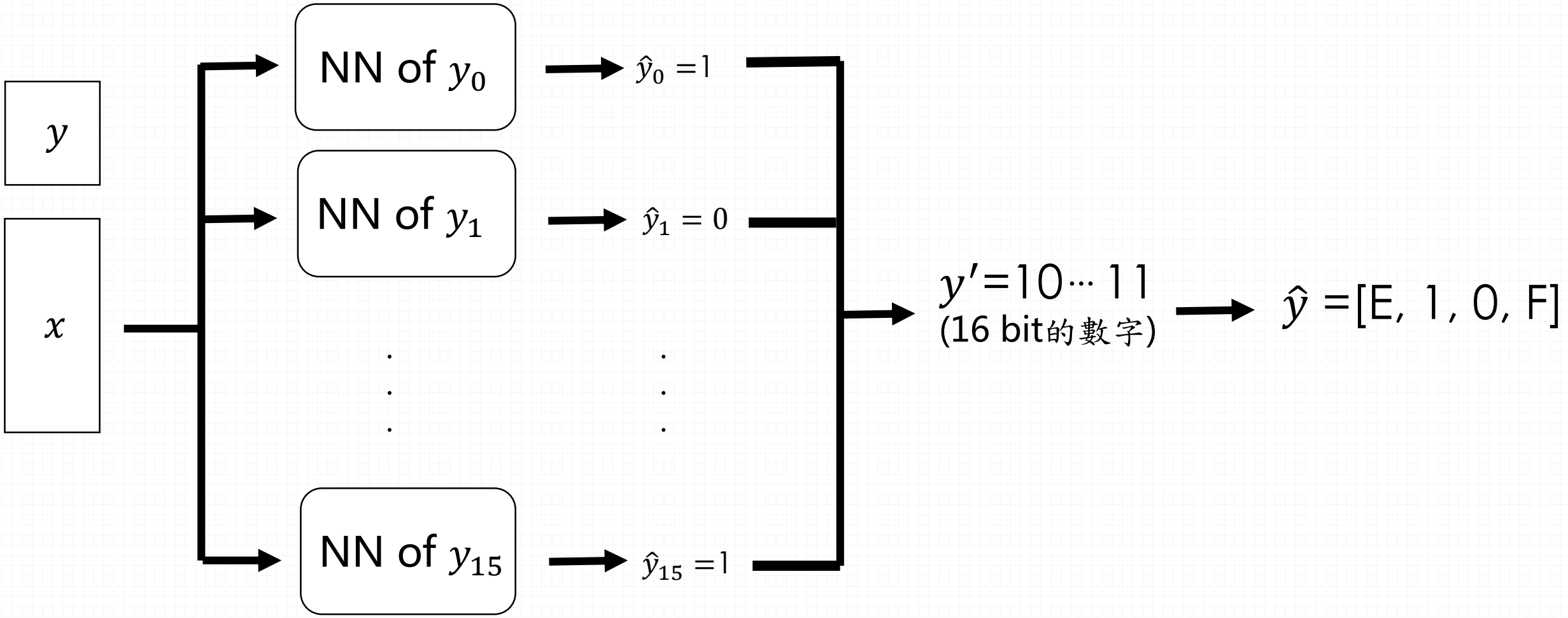
$y \longrightarrow (\ y_0\ , y_1, \ldots, y_{14}, y_{15}\ )$ $\quad y_i = 0 \text{ or } 1\ \forall i{=}0{\sim}15$

Reason:

Assume y_i are less relevant to each i，let problem(classification:65536) simplify 16 problem(classification:0/1)

二、Model

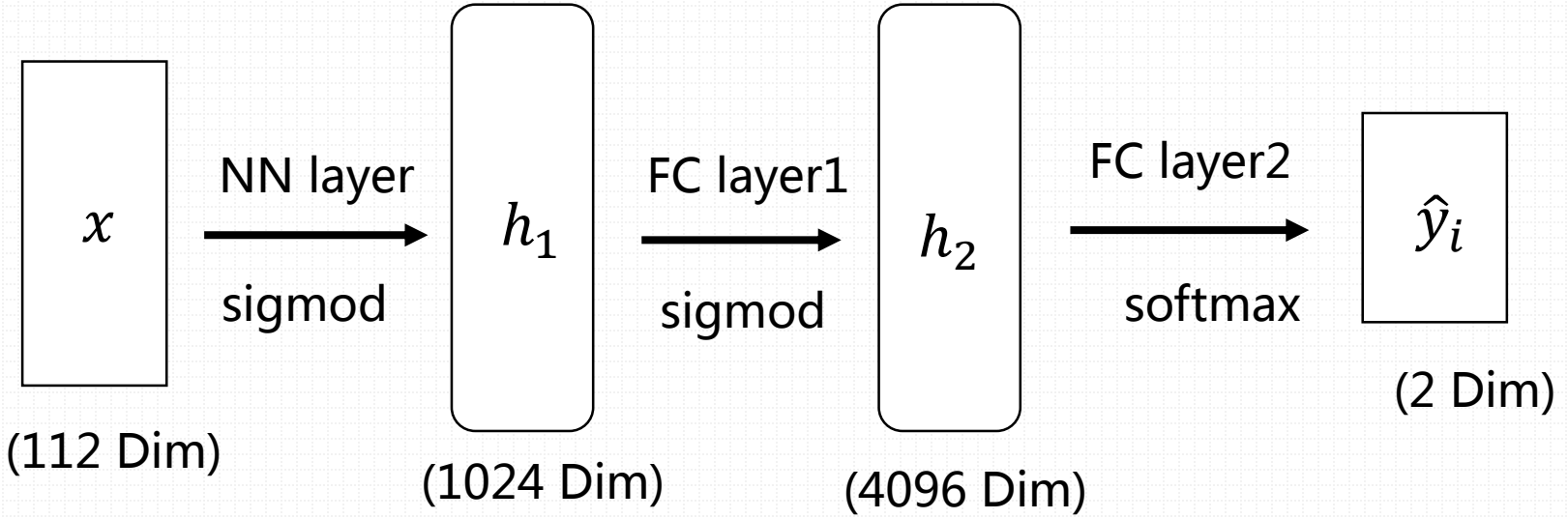
1. Neural Network(NN) for y(classification: $y = [c_{11}, c_{12}, c_{21}, c_{22}]$)



二、Model

2. Neural Network(NN) for each y_i(classification:y_i=0/1)

feature(x) : $x = [x_1, x_2, ..., x_{111}, x_{112}]$ (112 Dim) ， $x_i = 0\ or\ 1 \quad \forall i = 1 \sim 112$
label(y) : $y_i = 0 \text{ or } 1 \quad for\ each\ i = 0 \sim 15$



NN layer:
Neural work

FC layer:
Fully connected layer

二、Model

2. Neural Network(NN) for each y_i(classification:y_i=0/1)

Y_i : the true label for y_i (i bit)
$\hat{Y}_i$: the output of NN model for y_i (i bit)
C : the matrix, values are 1

Loss function: $L(W) = -Tr(Y_i^T \log(\hat{Y}_i)) - Tr((C - Y_i)^T \log(C - \hat{Y}_i))$

三、Result

Accuracy of Total test($y = [y_1, y_2] = [c_{11}, c_{12}, c_{21}, c_{22}]$) : 0.8587

Label	Acc(NN)	Label	Acc(NN)
0 bit(y_0)	0.9999	8 bit(y_8)	0.9636
1 bit(y_1)	0.9999	9 bit(y_9)	0.9613
2 bit(y_2)	0.9997	10 bit(y_{10})	0.9744
3 bit(y_3)	0.9999	11 bit(y_{11})	0.9747
4 bit(y_4)	0.9999	12 bit(y_{12})	0.9805
5 bit(y_5)	0.9987	13 bit(y_{13})	0.9967
6 bit(y_6)	0.9998	14 bit(y_{14})	0.9908
7 bit(y_7)	0.9995	15 bit(y_{15})	0.9997

THE END

感謝聆聽