

COMP 4332 – HW2 written part

Student Name: **Hu Yao-Chieh**

Student ID: **20216239**

ITSC Email: yhuag@ust.hk

Question 3.

(a)

(i)

No. of Computers	No. of Phones	Buy_NintendoLabo
2	0	0
0	2	0
4	2	1
2	4	1

(ii)

w1: 0.6141609492298914

w2: 0.7316411910756999

b: -0.49648656097013644

Training process:

```
w1: -0.5531394031500048, w2: 0.3, b: -0.12656970157500236
D: 0, Y: 0.7109495026250039, MAE Loss: 0.7109495026250039
-----
w1: -0.5531394031500048, w2: -0.43943440800315553, b: -0.4962869055765801
D: 0, Y: 0.6161953400026297, MAE Loss: 0.6161953400026297
-----
w1: 1.782259923184013, w2: 0.7282652551638533, b: 0.08756292600692434
D: 1, Y: 0.02691694736082596, MAE Loss: 0.973083052639174
-----
w1: 1.7839478911399362, w2: 0.7316411910756999, b: 0.08840690998488597
D: 1, Y: 0.9985933600367306, MAE Loss: 0.001406639963269396
-----
w1: 0.6141609492298914, w2: 0.7316411910756999, b: -0.49648656097013644
D: 0, Y: 0.9748224515917041, MAE Loss: 0.9748224515917041
```

(b)

Traditional neural network cannot treat a time series data as a stream of information, but merely segregated. This retard the processing of continuous data such as natural language, stock price and weather information. By RNN, the model can remember the past information and apply to the present prediction. Still, vanilla RNN is facing gradient vanishing problem that needs LSTM and some other techniques to solve long distance memory from the past.

Question 5. (Waive by coupon no. 87)