THE UNIVERSITY OF HONG KONG
FACULTY OF ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE
COMP7409B Machine Learning in Trading and Finance
Final Examination

Date: May 14, 2024 Time: 6:30pm-8:30pm

INSTRUCTIONS:

a. Answer ALL questions. They are all COMPULSORY.
b. Total mark is 100. The mark value of each question (or part of a question) is indicated before the question (or part of the question)
c. Write your university number clearly at the beginning of your answer script. DO NOT write down your name.
d. Only approved calculators as announced by the Examinations Secretary can be used in the examination. It is the candidates' responsibility to ensure that their calculator operates satisfactorily, and candidates must record the name and type of the calculator used on the front page of your answer script.

1. (24%) This question may require you to either (i) determine the output of a Python code snippet, or (ii) create simple Python programs to perform specific tasks. You can assume that your programs have already imported the libraries pandas as pd, numpy as np, pytorch as torch. If you require additional libraries for your code, make sure to import them beforehand.

   (a) (4%) What is the output of the following snippet:

```
L0 = np.linspace(-10, 10, 5)
L1 = []
for _ in L0:
    if _ > 0:
        L1.append(_)
L2 = [_**2 for _ in L1]
print(L2)
```

   (b) (4%) Write a Python code snippet that generates a graph for the function f(x) = x * x, where x ranges from 0 to 10 (inclusive).

   (c) (4%) What is the output of the following snippet?

```
a = np.arange(0,24).reshape(2,-1,2)
print(a[np.newaxis,...])
```

   (d) (4%) What is the output of the following program:

```
print([0,1,2,3,4,5,6][:2:-1])
```

   (e) (4%) What is the output of the following statement:

```
t= torch.zeros(30)
t_reshape = t.reshape(5,6)
print(t_reshape)
```

   (f) (4%) Recall that torch has a method matmul() which compute the multiplication of two matrices, which is defined formally as follows:

   If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B}$ is an $n \times p$ matrix,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}$$

   the result of the matrix multiplication of A and B is the matrix

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{pmatrix}$$

such that

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj},$$

This question asks you what the output of the following snippet is.

```
m1 = torch.tensor([[1, 2],[3,4]])
m2 = torch.tensor([[6,7],[5,0]])
print(m1+m2)
print(m1-m2)
print(m1*m2)
print(torch.matmul(m1,m2))
```

2. (25%) The Kaggle website has the data set "Walmart.csv" containing the weekly sales of US retail store Walmart at different locations. Following shows the beginning and end of the data set, as well as its data fields.

A	B	C	D	E	F	G	H	I
Store	Date	Weekly_Sa	Holiday_Fl	Temperatu	Fuel_Price	CPI	Unemployment	
1	5/2/2010	1643691	0	42.31	2.572	211.0964	8.106	
1	12/2/2010	1641957	1	38.51	2.548	211.2422	8.106	
1	19/2/2010	1611968	0	39.93	2.514	211.2891	8.106	
1	26/2/2010	1409728	0	46.63	2.561	211.3196	8.106	
1	5/3/2010	1554807	0	46.5	2.625	211.3501	8.106	
1	12/3/2010	1439542	0	57.79	2.667	211.3806	8.106	
1	19/3/2010	1472516	0	54.58	2.72	211.2156	8.106	
45	31/8/2012	734297.9	0	75.09	3.867	191.4613	8.684	
45	7/9/2012	766512.7	1	75.7	3.911	191.5777	8.684	
45	14/9/2012	702238.3	0	67.87	3.948	191.6999	8.684	
45	21/9/2012	723086.2	0	65.32	4.038	191.8567	8.684	
45	28/9/2012	713174	0	64.88	3.997	192.0136	8.684	
45	5/10/2012	733455.1	0	64.89	3.985	192.1704	8.667	
45	12/10/2012	734464.4	0	54.47	4	192.3273	8.667	
45	19/10/2012	718125.5	0	56.47	3.969	192.3309	8.667	
45	26/10/2012	760281.4	0	58.85	3.882	192.3089	8.667	

Store - the store number

Date - the week of sales

Weekly_Sales - sales for the given store

Holiday_Flag - whether the week is a special holiday week 1 – Holiday week 0 – Non-holiday week

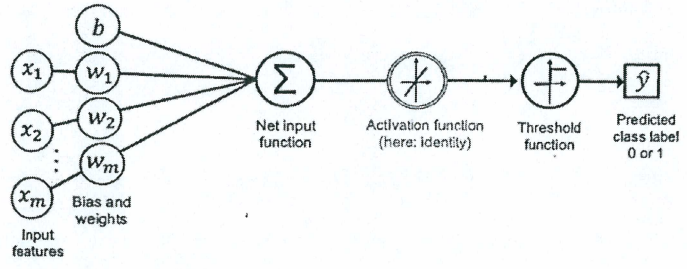
Temperature - Temperature on the day of sale

Fuel_Price - Cost of fuel in the region

CPI – Prevailing consumer price index

This question asks you to design a system that applies machine learning to predict the weekly sale of a given store of Walwart for next week. You should describe all the steps needed to implement such a system, such as data cleaning and exploratory data analysis. This is an open question, and you may include any steps that you think are appropriate. You should add some Python snippets to clarify your description.

3. (15%) Consider the following simple neural network



where

$$\text{the output of the "net input function is "} b + \sum_{1 \le i \le m} x_i \times w_i,$$

and output of the Threshold function is 1 if its input is no smaller than 0.5, and 0 otherwise. Suppose that m = 3, and

$$w_1 = 0.3, w_2 = -0.2, w_3 = 0.5 \text{ and } b = -0.01.$$

What are the sensitivity, specificity and the F1-score of the above simple neural network for the following testing data set, where the columns x1, x2 and x3 are for the three input features, and column y is for the correct class labels.

	x1	x2	x3	y
0	0.1	0.60	0.40	1
1	0.1	-0.20	0.06	0
2	0.3	0.05	0.50	1
3	0.2	0.40	0.20	0
4	1.0	0.17	0.20	1

You may provide the details of your calculations, allowing for the possibility of receiving partial credit even if your final answers are incorrect.

4. (16%) The question relates to the pricing of a put option for stock X. The option has a strike price of K and expires in 150 days from today. The closing price of X today is S. Calculate the premium of the contract for this put option in each of the following four scenarios, and provide justification for each answer.

(a) (4%) The stock price does not change over time, and it is S after 150 days. Also, the risk-free rate is 0.

(b) (4%) The stock price does not change over time, and the risk-free rate is r during this period.

(c) (4%) After 150 days, the price of the stock is S with a probability of 0.5, S/2 with a probability of 0.25 and 3S/2 with a probability of 0.25. Assuming the risk-free rate is 0.

(d) (4%) After 150 days, the price of the stock is S with a probability of 0.5, S/2 with a probability of 0.25 and 3S/2 with a probability of 0.25. Assuming the risk-free rate is r.

5.(20%) We have introduced in our leacture the OpenAI Gym's game Cart Pole, in which the action space is

Num	Action
0	Push cart to the left
1	Push cart to the right

and the observation space

Num	Observation	Min	Max
0	Cart Position	-4.8	4.8
1	Cart Velocity	-Inf	Inf
2	Pole Angle	~ -0.418 rad (-24°)	~ 0.418 rad (24°)
3	Pole Angular Velocity	-Inf	Inf

We have also given an algorithm that plays the game by random moves. Write a Python program that improves the performance by using Q-learning.

END OF PAPER