**THE UNIVERSITY OF HONG KONG**

**FACULTY OF ENGINEERING**
**DEPARTMENT OF COMPUTER SCIENCE**

**FITE7410 Financial Fraud Analytics**

**Date:** **Time:**

**The exam is 120 minutes.**
**Answer ALL questions in SEPARATE ANSWER BOOK provided.**

**Download the Answer Book from Moodle**

*Candidates are permitted to refer to any electronic/printed/handwritten materials in the examination. Internet searching and crowdsourcing from group messages, online forums or social media, etc. are strictly forbidden.*

*Only approved calculators as announced by the Examinations Secretary can be used in this examination. It is 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 the examination script.*

*Students can write/type/draw answer using computer and/or on paper. Students need to submit a single WORD OR PDF file to OLEX.*

## QUESTION 1.

### QUESTION 1.1.

Given the following pairs of customer income and fraud outcome in a data set. Use the spit criteria: Annual Income < 120k, give the formulation for Information Gain (IG) based on Entropy.

Write down the steps of your formulation. **NO marks will be given if only IG value is written as the answer.**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Fraud | No | No | No | Yes | No | Yes | No | Yes | No | Yes |
| Annual Income | 90k | 100k | 130k | 200k | 300k | 110k | 105k | 250k | 150k | 140k |

### QUESTION 1.2.

Given the confusion matrix below :

| | | **Actual** | |
|---|---|---|---|
| | | **Fraud=Yes** | **Fraud=No** |
| **Prediction** | **Fraud=Yes** | 3 | 3 |
| | **Fraud=No** | 0 | 4 |

(a) Write down the steps of 3 classification metrics: Accuracy, Recall, Precision. Write down the steps of your formulation. **NO marks will be given if only classification metrics value are written as the answer.**

(b) Based on the result of the classification metrics in (a), if the cost of missing out fraudsters is higher than false alarms, will you prefer this model? Why or Why Not?

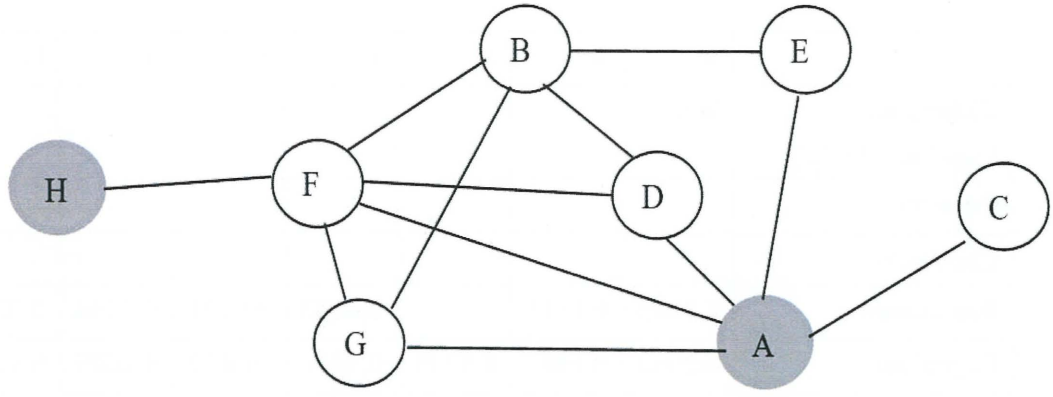
### QUESTION 1.3.

What are the 3 common stages in Money Laundering?

### QUESTION 1.4.

Please name 3 recommendations of the Financial Action Task Force (FATF).

**QUESTION 2.**




Assuming the above is an information network of a company, where

Nodes = staff of the company

Edges = email communications between 2 nodes

Grey nodes = identified staff with known fraudulent acts

The GEODESIC PATHs or SHORTEST PATHs of all nodes are as follows:

| | **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** |
|---|---|---|---|---|---|---|---|---|
| **A** | 0 | 2 | 1 | 1 | 1 | 1 | 1 | 2 |
| **B** | 2 | 0 | 3 | 1 | 1 | 1 | 1 | 2 |
| **C** | 1 | 3 | 0 | 2 | 2 | 2 | 2 | 3 |
| **D** | 1 | 1 | 2 | 0 | 2 | 1 | 2 | 2 |
| **E** | 1 | 1 | 2 | 2 | 0 | 2 | 2 | 3 |
| **F** | 1 | 1 | 2 | 1 | 2 | 0 | 1 | 1 |
| **G** | 1 | 1 | 2 | 2 | 2 | 1 | 0 | 2 |
| **H** | 2 | 2 | 3 | 2 | 3 | 1 | 2 | 0 |

## QUESTION 2.1.

Please fill in the relational neighbor, farness, closeness and betweenness centrality in the blank box below:

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| ***Relational Neighbor P(F\|?)*** | NA | | | | | | | NA |
| ***Farness*** | | | | | | | | 2.1429 |
| ***Closeness*** | | | | | | | | 0.4667 |
| ***Betweenness*** | 0.3968 | 0.1111 | | 0.0238 | 0.0238 | 0.3254 | 0.0238 | |
| ***PageRank*** | 0.2042 | 0.1587 | 0.0535 | 0.1213 | 0.0872 | 0.2009 | 0.1213 | 0.0529 |

## QUESTION 2.2.

When analysing this email communication network, please identify:

a) Who can get access to information most efficiently in this email network? Please state your reasons to support your answer.

b) Who can mostly control the information flow in this email network? Please state your reasons to support your answer.

c) Who will be most likely tend to exhibit similar behaviours as the neighbours in future? Please state your reasons to support your answer.

d) Which person(s) in the network we should monitor in the future for potential fraudulent acts? Please state your reasons based on the scores in the table listed in QUESTION 2.1 to support your answer.

**QUESTION 3.**

Case study

Case Background

A whistle blower complaint was received by the Headquarters of a European chemical company targeting at their Country Managing Director in China.

There were allegations on diversion of the company's business to his private business as well as paying himself excessive and fraudulent expense claims during the last 3 years.

Please provide an Action Plan setting out how you would investigate the potential irregularities/ issues highlighted above based on what you have learned in fraud investigation.

Please also explain what technology you would use to assist with the investigation, e.g. data analytics and/or any other tools.

=== END OF PAPER ===