2022 Semester Two (October 2022) Examination Period

Faculty of Information Technology

EXAM CODES:	FIT5215
TITLE OF PAPER:	Deep Learning
EXAM DURATION:	2 hours 10 mins

Rules

During your eExam, you must not have in your possession any item/material that has not been authorised for your exam. This includes books, notes, paper, electronic device/s, smart watch/device, or writing on any part of your body. Authorised items are listed above. Items/materials on your device, desk, chair, in your clothing or otherwise on your person will be deemed to be in your possession. Mobile phones must be switched off and placed face-down on your desk during your exam attempt.

You must not retain, copy, memorise or note down any exam content for personal use or to share with any other person by any means during or following your exam. You are not allowed to copy/paste text to or from external sources unless this has been authorised by your Chief Examiner.

You must comply with any instructions given to you by Monash exam staff.

As a student, and under Monash University's Student Academic Integrity procedure, you must undertake all your assessments with honesty and integrity. You must not allow anyone else to do work for you and you must not do any work for others. You must not contact, or attempt to contact, another person in an attempt to gain unfair advantage during your assessment. Assessors may take reasonable steps to check that your work displays the expected standards of academic integrity.

Failure to comply with the above instructions, or attempting to cheat or cheating in an assessment may constitute a breach of instructions under regulation 23 of the Monash University (Academic Board) Regulations or may constitute an act of academic misconduct under Part 7 of the Monash University (Council) Regulations.

Authorised Materials

CALCULATORS	☑ YES	☐ NO	Physical / Virtual Calculator
DICTIONARIES	☐ YES	☑ NO	
NOTES	☑ YES	☐ NO	Notes - Double sided A4 x 1 - Physical Only
WORKING SHEETS	☑ YES	☐ NO	

PERMITTED ITEM	☑ YES ☐ NO
if yes, items permitted are:	i) https://www.online-calculator.com/ ii) https://www.desmos.com/scientific iii) https://www.calculator.net/scientific-calculator.html

Instructions

Once your exam finishes, you will be given time to scan a QR code and upload your answers using your smartphone and laptop.

Here's how to do it.

This examination is designed for the FIT5215 Deep Learning unit, S2 2022.

- This is a **closed book** exam with specifically permitted items.
- Please answer **ALL** questions.

It contains THREE (3) parts with a total of **100 marks**:

- **Part A** contains **12** multiple-choice questions, together they are worth a total of **25 marks**. For questions with more than one answer, a correct choice will receive a partial mark and incorrect choices will reduce the mark. To receive full marks, only correct choices must be selected.
- **Part B** contains **6** short workout and knowledge questions, worth **40 marks**. These questions typically require short knowledge answers and calculations based on the knowledge you have learned from the unit. Having the calculator handy for these questions is recommended.
- **Part C** contains **7** mixed and handwritten-answer questions, worth **35 marks**. These questions typically assess the knowledge and understanding of lecture contents.
- Once the exam duration is finished, your exam will automatically submit. Please ensure you finalize your answers before the end of the allocated exam time.
- **When your exam duration has ended, you are not permitted to write on or edit any of your answer sheets.** Please note you must remain under exam conditions and are not permitted to adjust your answer sheets during the upload process once you have submitted your exam.

Good luck with your exam!

Instructions

Information

You can review your exam instructions by clicking the 'Show Instructions' button above.

Part A - Multiple Choice Questions (12 questions, 25 marks)

Information

Part A consists of **12 multiple-choice questions**, worth 25 points.

Question 1

Assume that we have 5 classes in **{cat = 1, dolphin = 2, monkey = 3, dog = 4, elephant = 5}** Given a data example x with ground-truth label monkey, assume that a feed-forward NN gives discriminative scores to this x as h_1=-2, h_2=1, h_3=5, h_4=2, h_5=4. What is the probability to predict x as **elephant** or **p(y=elephant | x)**?
(Single choice)

2 Marks

Select one:

- a. 0.5
- b. $\frac{e^2}{e^{-2}+e^1+e^5+e^2+e^4}$
- c. $\frac{e^{-2}}{e^{-2}+e^1+e^5+e^2+e^4}$
- d. $\log \frac{e^2}{e^{-2}+e^1+e^5+e^2+e^4}$
- e. $\frac{e^4}{e^{-2}+e^1+e^5+e^2+e^4}$

Question 2

Consider an image classification task with six classes **{cat=1, car=2, lion=3, dolphin=4, cow=5, dog = 6}**. Consider an image x. Assume that a Convolutional Neural Network gives a prediction probabilities **f(x)=[0.1, 0.2, 0.1, 0.2, 0.3, 0.1]** and categorial **ground-truth label** of x is **cow**. What is the **cross-entropy loss** suffered by this prediction?

(Single choice)

2 Marks

Select one:

- a. -log 0.2
- b. -log 0.3
- c. log 0.3
- d. -log 0.1
- e. log 0.2

Question 3

Consider the **LeakyReLU** activation function

2 Marks

$$\sigma(z) = \begin{cases} z & if\ z \geq 0 \\ 0.2z & otherwise \end{cases}$$

Assume that $h = \sigma(\bar{h})$ with $\bar{h} = [-1, 2, -3]$. What is the derivative $\frac{\partial h}{\partial \bar{h}}$?

(Single choice)

Select one:

- a. diag([-1, 2, -3])
- b. diag([0.2, 1 ,0.2])
- c. [0.2, 1 ,0.2]
- d. diag([0, 1 ,0])
- e. [0, 1, 0]

Question 4

Given the function $f(w) = \frac{1}{1000}\sum_{i=1}^{1000}(w - x_i)^2$ where $x_i=i, \forall i=1,\ldots,1000$. We need to solve $\min_w$ f(w) using stochastic gradient descent with the learning rate $\eta=0.1$. Assume we sample a batch $b_1=1, b_2=3$ of indices and at the iteration t, we have $w_t=10$. What is the value of w_{t+1} at the next iteration?

2 Marks

(Single choice)

Select one:

- a. 8.3
- b. 8.2
- c. 8.1
- d. 8.4
- e. 8.5

Question 5

Assume that the tensor before the last tensor of a CNN has shape **[128, 32, 32, 20]** and we apply **30 filters**, each of which has the shape **[3, 3, 20]** and strides= [2,2] with padding = 'valid' to obtain the last tensor. We flatten this tensor to a fully connected (FC) layer. What is the number of neurons on this FC layer?

2 Marks

(Single choice)

Select one:

- a. 15 x 15 x 30
- b. 14 x 14 x 20
- c. 128 x15 x 15 x 30
- d. 15 x 15 x 20
- e. 128 x 16 x 16 x 30

Question 6

Assume that the tensor before the last tensor of a CNN has shape **[32, 64, 64, 10]** and we apply **20 filters**, each of which has the shape **[3, 3, 10]** and strides= [2,2] with padding = 'same' to obtain the last tensor. What is the shape of the output tensor?

2 Marks

(Single choice)

Select one:

- a. [32, 29, 29, 20]
- b. [32, 32, 32, 20]
- c. [32, 31, 31, 20]
- d. [32, 30, 30, 20]
- e. [32, 33, 33, 20]

Question 7

2 Marks

According to the following code, what is the shape of h3?

```
embed_size = 64
vocab_size = 200
x = tf.keras.Input(shape=[5],dtype= 'int64')
h1 = tf.keras.layers.Embedding(vocab_size, embed_size)(x)
h2 = tf.keras.layers.GRU(16, return_sequences= True)(h1)
h3 = tf.keras.layers.GRU(8, return_sequences= True)(h2)
h4 = tf.keras.layers.GRU(16, return_sequences= True)(h3)
h5 = tf.keras.layers.Flatten()(h4)
h6 = tf.keras.layers.Dense(10, activation='softmax')(h5)
```

(Single choice)

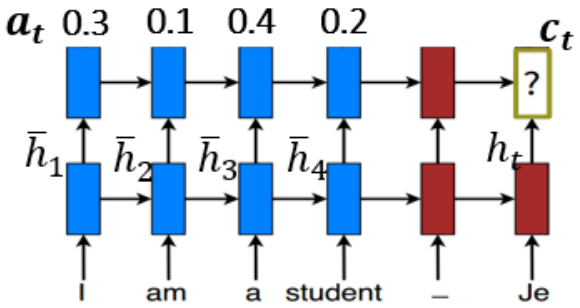
Select one:

- a. [None, 5, 10]
- b. [None, 5, 200]
- c. [None, 5, 16]
- d. [None, 5, 64]
- e. [None, 5, 8]

Question 8

Consider the below seq2seq model. We apply the global attention to compute the **context vector c_t**.
What are correct?

2
Marks



(Multiple choice)

Select one or more:

a.

$c_t = 0.1\bar{h}_2 + 0.4\bar{h}_3 + 0.2\bar{h}_4$

b.

The first word is more important than other words to the generation of the current output word.

c.

$c_t = 0.2\bar{h}_1 + 0.4\bar{h}_2 + 0.1\bar{h}_3 + 0.3\bar{h}_4$

d.

The third word is more important than other words to the generation of the current output word.

e.

$c_t = 0.3\bar{h}_1 + 0.1\bar{h}_2 + 0.4\bar{h}_3 + 0.2\bar{h}_4$

Question 9

Given a CBOW model with vocabulary size 1,000 and embedding size 250, we consider a target word with index 10 and context words with indices 15, 25, 35, 45 respectively. Let U and V be two weight matrices connecting input to hidden layers and hidden to output layers. What statements are correct?

3 Marks

(Multiple choice)

Select one or more:

a.
Shape of U is [1000,250] and shape of V is [250,1000]

b.
Input to the network is $\frac{1_{15}+1_{25}+1_{35}+1_{45}}{4}$

c.
The hidden value h is the row 10 of the matrix U

d.
Shape of U is [1000,1000] and shape of V is [250,250]

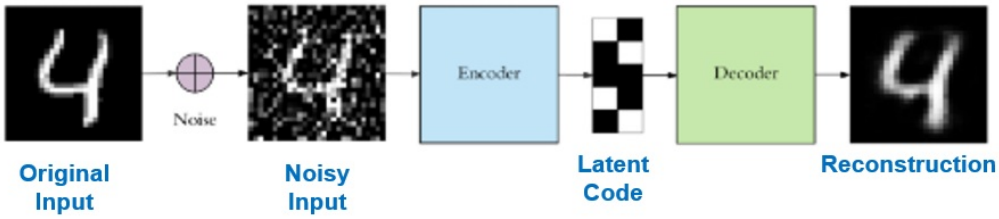
e.
Input to the network is one-hot vector 1_{10}

f.
The hidden value h is the average of rows 15, 25, 35, 45 of the matrix U

Question 10

How to train an denoising auto-encoder with encoder f_θ and decoder g_Φ?

2 Marks



(Multiple choice)

Select one or more:

a.

$$\min_{\theta,\Phi} \mathbb{E}_{\boldsymbol{x}\sim\mathbb{P}}\left[\mathbb{E}_{\boldsymbol{x}\prime\sim N(\boldsymbol{x},\eta I)}[d(\boldsymbol{x}, g_\Phi(f_\theta(\boldsymbol{x}')))]\right]$$

b.

$$\min_{\theta,\Phi} \mathbb{E}_{\boldsymbol{x}\sim\mathbb{P}}\left[d(\tilde{\boldsymbol{x}}, g_\Phi(f_\theta(\boldsymbol{x})))\right]$$

c.

$$\min_{\theta,\Phi} \mathbb{E}_{\boldsymbol{x}\sim\mathbb{P}}\left[\mathbb{E}_{\boldsymbol{\epsilon}\sim N(0,\eta I)}[d(\boldsymbol{x}, g_\Phi(f_\theta(\boldsymbol{x}+\boldsymbol{\epsilon})))]\right]$$

d.

$$\min_{\theta,\Phi} \mathbb{E}_{\boldsymbol{x}\sim\mathbb{P}}\left[\mathbb{E}_{\boldsymbol{\epsilon}\sim N(0,\eta I)}[d(\boldsymbol{x}, f_\theta(g_\Phi\ (\boldsymbol{x}+\boldsymbol{\epsilon})))]\right]$$

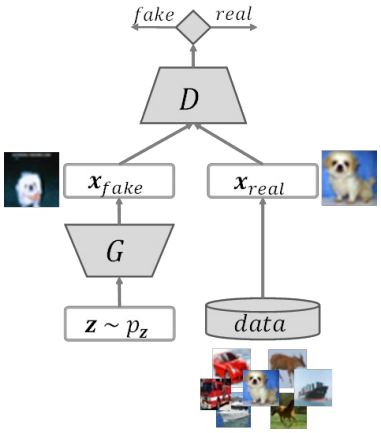
e.

$$\min_{\theta,\Phi} \mathbb{E}_{\boldsymbol{x}\sim\mathbb{P}}\left[\mathbb{E}_{\boldsymbol{x}\prime\sim N(\boldsymbol{x},\eta I)}[d(\boldsymbol{x}, f_\theta(g_\Phi(\boldsymbol{x}')))]\right]$$

Question 11

How to train GANs?

2 Marks



(Single choice)

Select one or more:

a.

$$\max_G \max_D J(G,D) = \mathbb{E}_{x \sim p_d(x)}[\log(1 - D(x))] + \mathbb{E}_{z \sim p(z)}[\log D(G(z))]$$

b.

$$\min_{\theta,\Phi} \mathbb{E}_{x \sim \mathbb{P}}[d(\tilde{x}, g_\Phi(f_\theta(x))]$$

c.

$$\max_G \min_D J(G,D) = \mathbb{E}_{x \sim p_d(x)}[\log D(x)] + \mathbb{E}_{z \sim p(z)}[\log(1 - D(G(z))]$$

d.

$$\min_G \max_D J(G,D) = \mathbb{E}_{x \sim p_d(x)}[\log D(x)] + \mathbb{E}_{z \sim p(z)}[\log(1 - D(G(z))]$$

Question 12

Given a DL model f(x;θ) parameterized by θ where f(x;θ) represents the prediction probabilities of x associated with a ground-truth label y∈{1,...,M}, we find an adversarial example by $x_{adv} = argmax_{x' \in B_\epsilon(x)} l(f(x'; \theta), y)$. Which statements are correct?

2 Marks

(Multiple choice)

Select one or more:

a.
We maximally decrease the chance to predict x with label y.

b.
It is a targeted attack.

c.
We maximally increase the chance to predict x with label y.

d.
We maximally increase the chance to predict x with any else label y'≠y.

e.
It is an untargeted attack.

Part B - Short Workout & Knowledge Questions (6 questions, 40 marks)

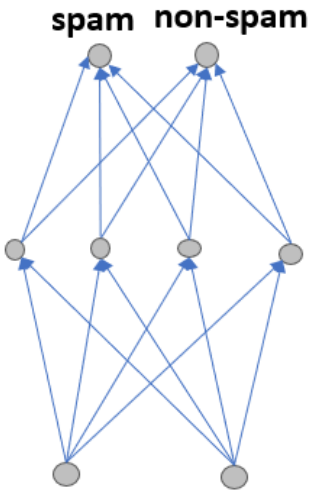
Information

Part B contains **6 short workout and knowledge questions**, worth 40 marks. These questions typically require short knowledge answers and calculations based on the knowledge you have learned from the unit. Having the calculator handy for these questions is recommended.

Question 13

Consider a feed-forward neural network as shown in the figure for spam email detection with two labels (spam=1 and non-spam=2). Assume that we feed a feature vector **x=[1,1]^T** with true label **y=2** to the network.

10 Marks



$$h^2, p = softmax(h^2)$$

$$W^2 = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix}, b^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\bar{h}^1, h^1 = ReLU(\bar{h}^1)$$

$$W^1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 1 & -1 \\ -2 & 1 \end{bmatrix}, b^1 = \begin{bmatrix} -2 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

$$x = [1,1]^T, y = 2$$

13a)
What are the formulas and the values of $\bar{h}^1, h^1$ (**hbar1** and **h1**)?

3 Marks

Note: you can use hbar1 and h1 for answering.

13b)
What are the formulas and the values for the logit **h2** and the prediction probability **p**?

3 Marks

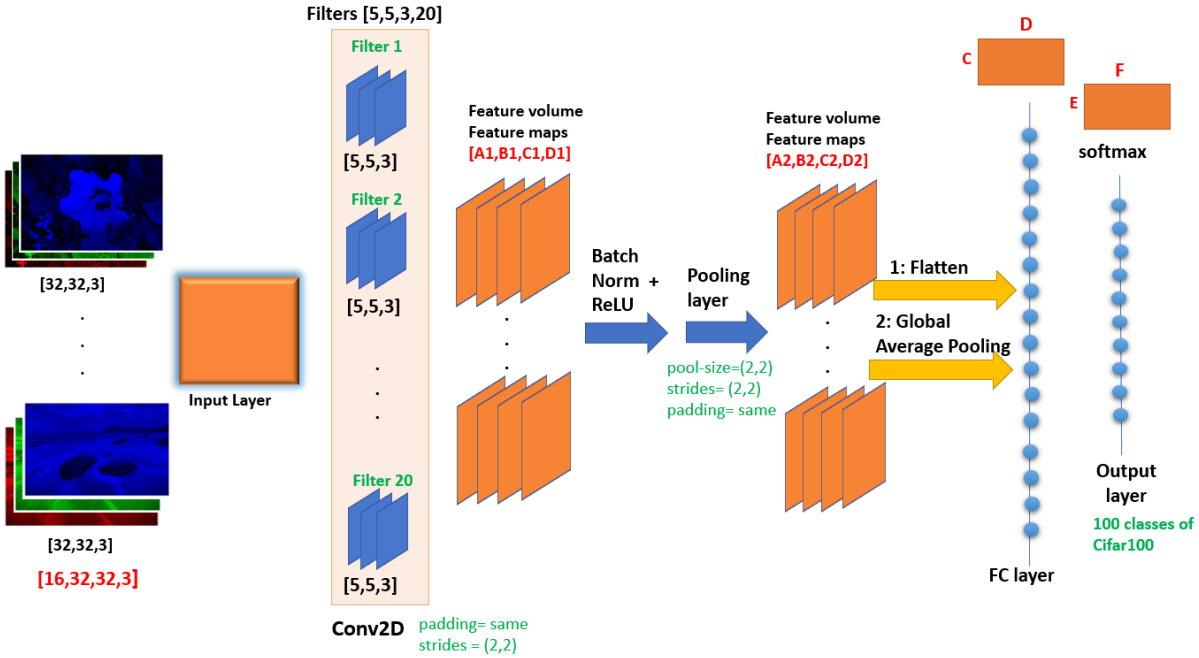
13c)
What is the predicted label $\hat{y}$ (yhat) and the cross-entropy loss for this prediction? Is it a correct or incorrect prediction?

4 Marks

Question 14

Assume that we conduct a Convolution Neural Network (CNN) with the configuration as shown in the below figure to predict the image dataset Cifar100 with 100 classes. We feed a batch of images with the shape [16, 32, 32, 3] our CNN. Answer the following questions.

10 Marks



14a)

What is the shape of the feature maps [A1, B1, C1, D1]? Show the steps of your answer.

2 Marks

14b)

What is the shape of the feature maps **[A2, B2, C2, D2]**? Show the steps of your answer.

2 Marks

14c)

Assume that we flatten the feature maps [A2, B2, C2, D2] to obtain the next output. What is the shape of the 2D tensor [C,D]? Explain your answer.

2 Marks

14d)

Assume that we apply global average pooling to the feature maps [A2, B2, C2, D2] to obtain the next output. What is the shape of the 2D tensor [C,D]? Explain your answer.

2 Marks

14e)

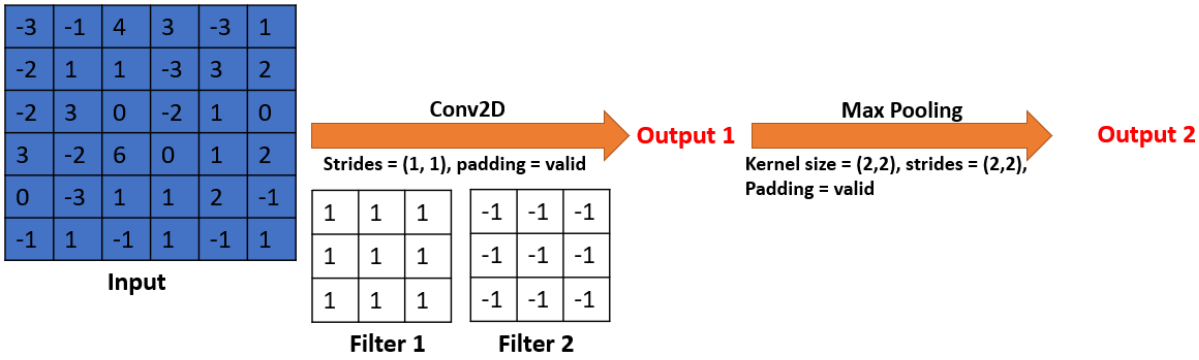
What is the shape of the 2D tensor [E,F]? Explain your answer.

2 Marks

Information

The statement below applies to Q.15-Q.16

Assume that we have [6,6,1] input tensor as shown below. We first apply a Conv2D layer with two filters (i.e., filter 1 and filter 2), strides = (1,1), and padding = valid to obtain the output 1. We then apply max pooling with kernel size = (2,2), strides = (2,2), and padding= valid to obtain the output 2.



Question 15

What are the values of the feature maps in Output 1?

5 Marks

Please answer the question on your blank piece of paper.

- After your exam finishes, you'll have **extra time** to access your phone to scan a **QR code** and **upload** your answer.
- Clearly label each page with **Student ID** and **this question number** (and **sub part** if applicable) (for example, 'Question 7a')
- **Do not** write your Name on it

No. of answer sheets: 1

Question 16

What are the values of the feature maps in Output 2?

5 Marks

Please answer the question on your blank piece of paper.

- After your exam finishes, you'll have **extra time** to access your phone to scan a **QR code** and **upload** your answer.
- Clearly label each page with **Student ID** and **this question number** (and **sub part** if applicable) (for example, 'Question 7a')
- **Do not** write your Name on it

No. of answer sheets: 1

Question 17

Read the following code and provide the shapes of the tensors x, h1, h2, h3, h4, h5, h6. Note that the shapes should contain one dimension with the Value None for batch size

5 Marks

```
embed_size = 256
vocab_size = 512
x= tf.keras.Input(shape = [20], dtype = 'int64')
h1 = tf.keras.layers.Embedding(vocab_size, embed_size)(x)
h2 = tf.keras.layers.LSTM(16, return_sequences = True)(h1)
h3 = tf.keras.layers.LSTM(32, return_sequences = True)(h2)
h4 = tf.keras.layers.LSTM(64, return_sequences = False)(h3)
h5 = tf.keras.layers.Flatten()(h4)
h6 = tf.keras.layers.Dense(15, activation = 'softmax')(h5)
```

Question 18

Read the following code and provide the shapes of the tensors x, h1, h2, z, hbar2, hbar1, xr.Note that the shapes should contain one dimension with the Value None for batch size

5 Marks

```
x = tf.keras.Input(shape = [32,32,3], dtype = 'float64')
h1 = tf.keras.layers.Conv2D(10, kernel_size=3, strides=[4,4], padding= "same", activation = 'relu')(x)
h2 = tf.keras.layers.Conv2D(20, kernel_size=3, strides=[2,2], padding= "same", activation = 'relu')(h1)
z= tf.keras.layers.Flatten()(h2)
hbar2 = tf.keras.layers.Reshape([h2.shape[1],h2.shape[2], h2.shape[3]])(h2)
hbar1 = tf.keras.layers.Conv2DTranspose(10, kernel_size = 3, strides = [2,2], padding= 'same', activation = 'relu')(hbar2)
xr = tf.keras.layers.Conv2DTranspose(3, kernel_size = 3, strides = [4,4], padding= 'same', activation = 'relu')(hbar1)
```

Part C - Mixed & Handwritten-Answer Questions (7 questions, 35 marks)

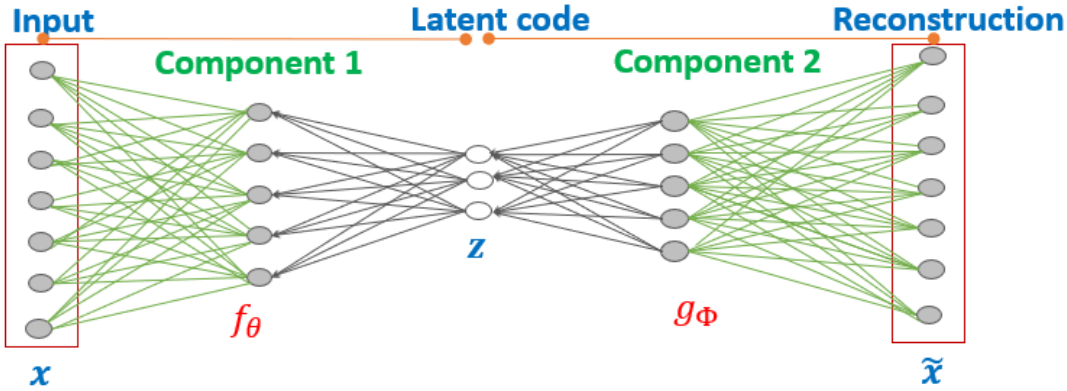
Information

Part C contains **7 mixed and handwritten-answer questions**, worth 35 marks . These questions typically assess the knowledge and understanding of lecture contents.

Information

The statement below applies to Q.19 - Q.20 - Q21 - Q22

Consider a deep learning model with the network architecture as shown below.



Question 19

What are the names of the deep learning model, component 1, and component 2? Write down the objective function to train this deep learning model with the explanations of the meaning of each term in this objective function.

5 Marks

Please answer the question on your blank piece of paper.

- After your exam finishes, you'll have **extra time** to access your phone to scan a **QR code** and **upload** your answer.
- Clearly label each page with **Student ID** and **this question number** (and **sub part** if applicable) (for example, 'Question 7a')
- **Do not** write your Name on it

No. of answer sheets: 1

Question 20

To be able to perform denoising, we add some Gaussian noises to inputs and aim to reconstruct benign inputs from noisy inputs. Write down the objective function to train this denoising variant with the explanations of the meaning of each term in this objective function.

2.5 Marks

Please answer the question on your blank piece of paper.

- After your exam finishes, you'll have **extra time** to access your phone to scan a **QR code** and **upload** your answer.
- Clearly label each page with **Student ID** and **this question number** (and **sub part** if applicable) (for example, 'Question 7a')
- **Do not** write your Name on it

No. of answer sheets: 1

Question 21

To gain sparser latent codes z, we apply a spare regularization term to latent codes z. Write down the objective function to train this sparse variant with the explanations of the meaning of each term in this objective function.

2.5 Marks

Please answer the question on your blank piece of paper.

- After your exam finishes, you'll have **extra time** to access your phone to scan a **QR code** and **upload** your answer.
- Clearly label each page with **Student ID** and **this question number** (and **sub part** if applicable) (for example, 'Question 7a')
- **Do not** write your Name on it

No. of answer sheets: 1

Question 22

To strengthen this model, we leverage with the principle of generative adversarial networks (GAN). To this end, we can view g_Φ as a generator and devise a discriminator D_γ (i.e., γ is its parameters) to discriminate $\tilde{x}$ and x. Give your further thoughts and comments about this extension. Write down the purposes of the discriminator D_γ and the component 2 (i.e., g_Φ) in this context. Write down the objective function to train D_γ. Write down the objective function to train f_θ and g_Φ when leveraging with the GAN principle.

5 Marks

Please answer the question on your blank piece of paper.

- After your exam finishes, you'll have **extra time** to access your phone to scan a **QR code** and **upload** your answer.
- Clearly label each page with **Student ID** and **this question number** (and **sub part** if applicable) (for example, 'Question 7a')
- **Do not** write your Name on it

No. of answer sheets: 1

Information

The statement below applies to Q.23 - Q.24

Consider a Convolution Neural Network (CNN) with the model parameter θ. Specifically, given an image x with the ground-truth label y, the CNN returns the prediction probabilities $f(x;\theta)$ over M classes and suffers the loss $l(f(x;\theta),y)$ where l is a loss function (e.g., the cross-entropy loss).

Question 23

a) Adversarial examples are a serious issue of CNNs. Give a definition of adversarial examples. Give a practical example to explain why adversarial examples circumvent the applications of CNNs in reality. (**3 points**)
b) Given a benign example x and the ϵ-ball $B_\epsilon=\{x':\|x'-x\|_\infty \leq \epsilon\}$, describe and give the formula to find out a targeted adversarial example for x. (**3 points**)
c) Given a benign example x and the ϵ-ball $B_\epsilon=\{x':\|x'-x\|_\infty \leq \epsilon\}$, describe and give the formula to find out a untargeted adversarial example for x. (**3 points**)

9 Marks

Please answer the question on your blank piece of paper.

- After your exam finishes, you'll have **extra time** to access your phone to scan a **QR code** and **upload** your answer.
- Clearly label each page with **Student ID** and **this question number** (and **sub part** if applicable) (for example, 'Question 7a')
- **Do not** write your Name on it

No. of answer sheets: 2

Question 24

a) Given a mini-batch $B=\{(x_1,y_1),\ldots,(x_b,y_b)\}$ at an iteration, describe how to perform adversarial training for this mini-batch to improve model robustness. (**3 points**)
b) How adversarial training is similar to data augmentation? How adversarial training is different to data augmentation? (**3 points**)

6 Marks

Please answer the question on your blank piece of paper.

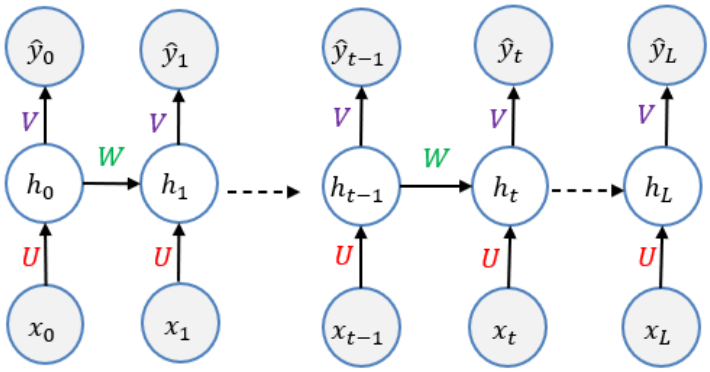
- After your exam finishes, you'll have **extra time** to access your phone to scan a **QR code** and **upload** your answer.
- Clearly label each page with **Student ID** and **this question number** (and **sub part** if applicable) (for example, 'Question 7a')
- **Do not** write your Name on it

No. of answer sheets: 1

Question 25

5 Marks

Given a standard recurrent neural network (RNN) as shown in the following figure, assume that we feed a sequence $x=x_0, x_1, x_2,\ldots,x_L$ to the RNN.



Explain why we can consider the **hidden state h_L** as a **lossy summary** of the sequence $x=x_0, x_1, x_2,\ldots,x_L$.

Please answer the question on your blank piece of paper.

- After your exam finishes, you'll have **extra time** to access your phone to scan a **QR code** and **upload** your answer.
- Clearly label each page with **Student ID** and **this question number** (and **sub part** if applicable) (for example, 'Question 7a')
- **Do not** write your Name on it

No. of answer sheets: 1