Question 1

Which of the following tasks can be classified as a typical machine learning task (MC):

- ❏ A. Clustering gene expression profiles
- ❏ B. Detecting spam email
- ❏ C. Storing and transferring data to a secure place
- ❏ D. Recognising handwritten digit numbers.

Question 2

Consider a general setting for a machine learning problem, which of the following statements is best at defining the concept of 'label' (SC)

- ❑ A. Items or instances of data used for training and evaluating ML models.
- ❑ B. A set of attributes, each is usually in form of a vector, associated with data points
- ❑ C. Values, categories or classes, assigned to data points.
- ❑ D. None of above

Question 3

John wants to build a machine learning model which helps to recognize students' face and automatically records their attendance in his class. Which of the following machine learning approaches would you recommend for this problem (SC)

- ❑ A. A regression model
- ❑ B. A classification model
- ❑ C. A generative model
- ❑ D. None of above

Question 4

John wants to build a machine learning model which helps to recognize students' face and automatically records their attendance in a class. He decides to train the model with 50 faces of his 50 students. Which of the statement are true (MC)

- ❑ A. John will be experiencing an overfitting problem in his approach.
- ❑ B. John will be experiencing an underfitting problem in his approach.
- ❑ C. John's model cannot recognize his students very well
- ❑ D. John's model cannot recognize new students very well

Question 5

Consider an email spam detection problem where the task is to predict spam emails. The following table summarizes the confusion matrix on the test dataset. (MC)

TRUE LABELS

Predicted Class		SPAM (1)	NON-SPAM (0)
	SPAM (1)	10	15
	NON-SPAM (0)	10	15

- ❑ A. The total number of emails contained in this dataset is 50.
- ❑ B. The total number of emails labelled as SPAM in this dataset is 20.
- ❑ C. The total number of emails labelled as NON-SPAM in this dataset is 25.
- ❑ D. The total number of emails labelled as SPAM and indeed predicted as SPAM is 15.

Question 6

Consider an email spam detection problem where the task is to predict spam emails. The following table summarizes the confusion matrix on the test dataset. (MC)

TRUE LABELS

	SPAM (1)	NON-SPAM (0)
SPAM (1)	10	15
NON-SPAM (0)	10	15

Predicted Class

- A. The precision is 10/(10+15)=2/5.
- B. The precision is 10/(10+10)=1/2.
- C. The recall is 10/(10+15)=2/5.
- D. The recall is 10/(10+10)=1/2.

Question 7

Consider an email spam detection problem where the task is to predict spam emails. The following table summarizes the confusion matrix on the test dataset. (MC)

TRUE LABELS

		SPAM (1)	NON-SPAM (0)
Predicted Class	SPAM (1)	10	15
	NON-SPAM (0)	10	15

- A. The prediction accuracy on this test dataset is (10+15)/(10 + 10 + 15 + 15) = 50%.

- B. The true positive rate (TPR) is 10/(10 + 15) = 66.66%.

- C. The true negative rate (TNR) is 15/(15 + 15) = 50%.

- D. It is not possible to calculate the sensitivity in this case.

Question 8

Given an implementation and its computational graph. Choose the correct statement (MC)
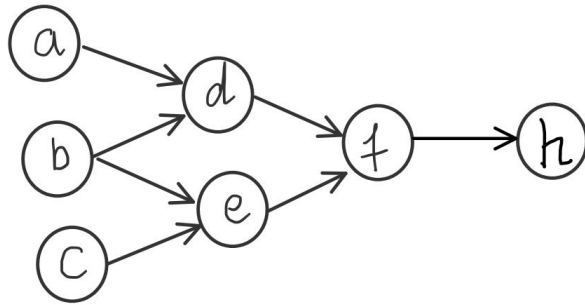

```
tf.reset_default_graph()
a = tf.constant(5)
b = tf.constant(2)
c = tf.Variable(initial_value=0., name='c')
d = tf.multiply(a,b)
e = tf.add(b,c)
f = tf.subtract(d,e)
h = tf.constant(7)
```

- ❏ A. It's a correct graph. We can run without error
- ❏ B. It is possible to get the value of tensor f without run global initialize variables
- ❏ C. It is possible to get the value of tensor h without run global initialize variables
- ❏ D. The value of d is 10

Question 9

Given an implementation as below. Choose the correct statement (MC)

```
x1 = tf.constant(name="x1", value=[1,2,3], dtype= tf.int32)
x2 = tf.constant(name="x2", value=[1], dtype= tf.int32)
x3 = tf.constant(name="x3", value=[1], dtype= tf.int64)
x4 = tf.cast(x1, tf.float32)
```

❑ A. cannot add x1 to x2 because they have different shapes

❑ B. cannot add x1 to x3 because they have different shapes

❑ C. cannot add x1 to x3 because they have different types

❑ D. cannot add x1 to x4 because they have different types

Question 10

Given an implementation as below. Choose the correct outputs (in order a, b, c) (SC)

```python
a = tf.constant([1,2,3])
b = tf.expand_dims(a, axis=1)
c = tf.expand_dims(b, axis=2)

print('a shape: {}'.format(a.get_shape()))
print('b shape: {}'.format(b.get_shape()))
print('c shape: {}'.format(c.get_shape()))
```

❑ A. (3,1), (3,1), (3,1)

❑ B. (3,1), (3,1,1), (3,1,1,1)

❑ C. (3,), (3,1), (3,2)

❑ D. (3,), (3,1), (3,1,1)