✓ **Selamat! Anda lulus!**

| Nilai diterima 100% | Nilai Pengiriman Terbaru 100% | UNTUK LULUS 80% atau lebih tinggi | **Pergi ke item berikutnya** |

---

1. 1 / 1 poin

Error analysis

m_{cv} = ~~500~~ 5000 examples in cross validation set.
Algorithm misclassifies ~~100~~ 1000 of them.
Manually examine (100) examples and categorize them based on common traits.

→ Pharma: 21 → more data features
→ Deliberate misspellings (w4tches, med1cine): (3)
→ Unusual email routing: 7
→ Steal passwords (phishing): 18 → more data features
Spam message in embedded image: 5

Which of these is a way to do error analysis?

◉ Manually examine a sample of the training examples that the model misclassified in order to identify common traits and trends.

○ Collecting additional training data in order to help the algorithm do better.

○ Calculating the training error J_{train}

○ Calculating the test error J_{test}

✓ **Benar**
Correct. By identifying similar types of errors, you can collect more data that are similar to these misclassified examples in order to train the model to improve on these types of examples.

---

2. 1 / 1 poin

Data augmentation

Augmentation: modifying an existing training example to create a new training example.
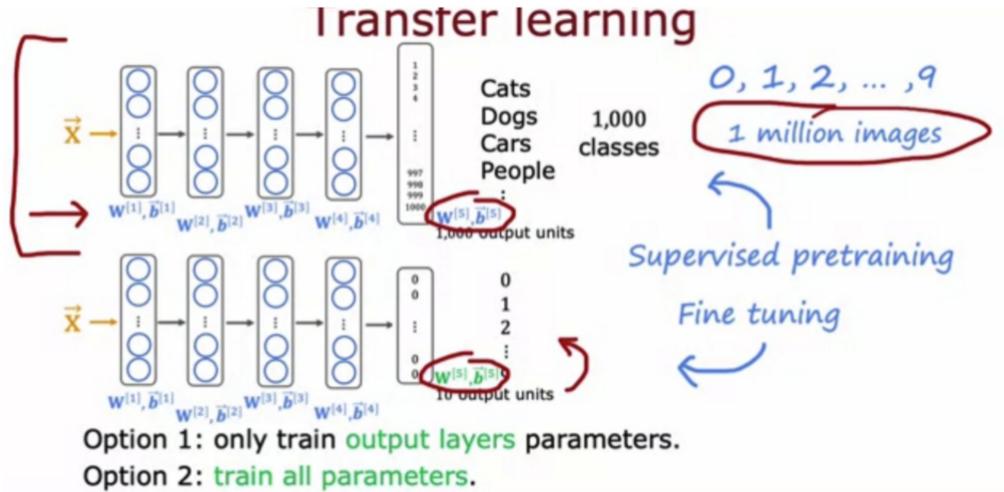
(X,Y)

We sometimes take an existing training example and modify it (for example, by rotating an image slightly) to create a new example with the same label. What is this process called?

○ Error analysis

○ Machine learning diagnostic

◉ Data augmentation

○ Bias/variance analysis

✓ **Benar**
Yes! Modifying existing data (such as images, or audio) is called data augmentation.

---

3. 1 / 1 poin

Transfer learning

$\vec{x}$ →

Cats
Dogs
Cars
People

1,000 classes

0, 1, 2, ... ,9

1 million images

Supervised pretraining

$w^{[1]}, \vec{b}^{[1]}$ $w^{[2]}, \vec{b}^{[2]}$ $w^{[3]}, \vec{b}^{[3]}$ $w^{[4]}, \vec{b}^{[4]}$ $w^{[5]}, \vec{b}^{[5]}$
1,000 output units

$\vec{x}$ →

0
1
2
⋮
9

Fine tuning

$w^{[1]}, \vec{b}^{[1]}$ $w^{[2]}, \vec{b}^{[2]}$ $w^{[3]}, \vec{b}^{[3]}$ $w^{[4]}, \vec{b}^{[4]}$ $w^{[5]}, \vec{b}^{[5]}$
10 output units

Option 1: only train **output layers** parameters.
Option 2: **train all parameters.**

What are two possible ways to perform transfer learning? Hint: two of the four choices are correct.

☐ Given a dataset, pre-train and then further fine tune a neural network on the same dataset.

☑ You can choose to train all parameters of the model, including the output layers, as well as the earlier layers.

✓ **Benar**
Correct. It may help to train all the layers of the model on your own training set. This may take more time compared to if you just trained the parameters of the output layers.

☑ You can choose to train just the output layers' parameters and leave the other parameters of the model fixed.

✓ **Benar**
Correct. The earlier layers of the model may be reusable as is, because they are identifying low level features that are relevant to your task.

☐ Download a pre-trained model and use it for prediction without modifying or re-training it.