The second winter of ANN is triggered by the failing of perceptron model in solving XOR problem.

하나를 선택하세요.
- ○ 참
- ◉ 거짓 ✔

정답은 '거짓'입니다.

ADAM is an optimizer which can determine the learning rate adaptively and automatically.

하나를 선택하세요.
- ◉ 참 ✔
- ○ 거짓

정답은 '참' 입니다.

ReLU dying problem can be solved by using low learning rate and proper initialization.

하나를 선택하세요.
- ○ 참
- ◉ 거짓 ✘

정답은 '참' 입니다.

Which of the following are the factors that contribute to the success of deep learning (answers can be multiple)?

하나 혹은 하나 이상을 선택하세요.
- ☑ A. Advancement of algorithms ✔
- ☑ B. Massive training data ✔
- ☑ C. Powerful computational machines ✔
- ☑ D. Media hype ✘
- ☐ E. None of the above

Your answer is partially correct.
너무 많은 옵션을 선택하였습니다.
The correct answers are: Advancement of algorithms, Massive training data, Powerful computational machines

Which of the following **do not satisfy** the conditions (non-constant, bounded and monotonic increasing) that imposed by the universe approximation theorem on activation functions in the network (answers can be multiple)? Assume the domain of the function is range from -inf to +inf.

하나 혹은 하나 이상을 선택하세요.
- ☑ A. Hyperbolic Tangent ✘
- ☐ B. Gaussian
- ☐ C. Sigmoid
- ☐ D. Inverse Tangent
- ☐ E. ReLU (Rectifier Linear)

Your answer is incorrect.
The correct answers are: Gaussian, ReLU (Rectifier Linear)

For a network with 10 hidden layers, gradient vanishing problem is more serious in hidden layer 10 than that of hidden layer 1.

하나를 선택하세요.
- ◉ 참 ✘
- ○ 거짓

정답은 '거짓'입니다.

SGD or Mini-batch GD is good to use when training data size is huge

하나를 선택하세요.
- ◉ 참 ✔
- ○ 거짓

A MLP has 2500 input neurons, 62 output neurons and 10 hidden layers where each hidden layer contains 100 neurons. How many total number of weights in the network **including biases**?

하나를 선택하세요.

○ A. 357,000

◉ B. 347,262 ✓

○ C. 347,000

○ D. 357,262

Your answer is correct.

정답은 347,262.입니다.

Which of the following is false?

하나를 선택하세요.

○ A. If any one of the hidden neuron in MLP is non-linear (apply non-linear activation function), the whole network remains as a non-linear model.

○ B. If linear activation function is used in the hidden layers, the entire network will reduce to a linear model like perceptron.

◉ C. The activation function in the output neuron will determine whether MLP is a linear or non-linear model. ✓

○ D. For two-class classification problem, we can either use sigmoid function (one neuron) or softmax function (two neurons) at the output layer.

Your answer is correct.

정답은 The activation function in the output neuron will determine whether MLP is a linear or non-linear model..입니다.

The network with softmax function in the output layer approximates categorical probability distribution.

하나를 선택하세요.

◉ 참 ✓

○ 거짓

정답은 '참' 입니다.

How many weights update equations are there in a 10 hidden layers MLP?

하나를 선택하세요.

○ A. 8

○ B. 9

◉ C. 10 ✗

○ D. 11

Your answer is incorrect.

정답은 11.입니다.

Backpropagation with momentum is a method to speed up the convergence of the network learning.

하나를 선택하세요.

◉ 참 ✓

○ 거짓

정답은 '참' 입니다.

Network does not learn when gradient loss (d loss/d w) is zero.

하나를 선택하세요.

◉ 참 ✓

○ 거짓

정답은 '참' 입니다.

Using small size filter is useful to reduce storage complexity of CNNs.

하나를 선택하세요.
○ 참
● 거짓 ✗

정답은 '참' 입니다.

A ConvNet with following specifications is given:

Layer	# of filters/filter size (Window)/Stride/Zero Padding	Feature Map Size	# of weights
Input	-	227x227x3	NA
Conv1	64/11x11/4/0	55x55x64	
Pool1	-/3x3/2/-	27x27x64	
Conv2	128/5x5/1/2	27x27x128	(i)
Pool2	-/3x3/2/-	13x13x128	(ii)
Conv3	256/3x3/1/1	13x13x256	
Pool3	-/3x3/2/-	6x6x512	
FC1	-	4096	(iii)
Output layer	-	1000	

Determine the values of (i), (ii) and (iii).

하나를 선택하세요.
○ A. i = 64x25x128, ii = 64x25x128, iii = 6x6x512x4096
○ B. i = 64x25x128, ii = 0, iii = 1000x4096
○ C. i = 64x9x128, ii = 64x25x128, iii = 6x6x512x4096
● D. i = 64x25x128, ii = 0, iii = 6x6x512x4096 ✓

Your answer is correct.
정답은 i = 64x25x128, ii = 0, iii = 6x6x512x4096.입니다.

Local connection and weight sharing are two major characteristics of CNNs

하나를 선택하세요.
● 참 ✓
○ 거짓

정답은 '참' 입니다.

In order to freeze the weights in certain layer of the network, the corresponding learning rate should be set to zero.

하나를 선택하세요.
○ 참
● 거짓 ✗

정답은 '참' 입니다.

Which of the following is NOT a type of transfer learning?

하나를 선택하세요.
● A. None of the above ✓
○ B. Transductive Transfer Learning
○ C. Unsupervised Transfer Learning
○ D. Inductive Transfer Learning

Your answer is correct.
정답은 None of the above.입니다.

Suppose you are to design a face recognition system for Visual AI class using a Convnet with [[Conv-ReLU-Conv-ReLU-pool]x5, [FC-ReLU]x3, [FC-Softmax]]. The filter size is set to 3x3 for all convolution layers. Unfortunately, you have no massive (very large) facial training data to train this deep network from scratch. Therefore, you decide to use a ConvNet pre-built model that trained from MegaFace database (an open million-scale face database collected from web) and develope a face recognition system for class.
If you have reasonably large number of face training data per person, what should you do to the pre-built model?

하나를 선택하세요.
○ A. Remove three FC layers and re-train softmax layer.
○ B. Finetune a few late conv layers
○ C. Finetune a few early conv layers
● D. None of the above ✗

Your answer is incorrect.
정답은 Finetune a few late conv layers.입니다.

VGG Net uses 3x3 filters only throughout the network.

하나를 선택하세요.
- ⦿ 참 ✓
- ○ 거짓

정답은 '참' 입니다.

Which of the following is not the advantage of skip connection?

하나를 선택하세요.
- ○ A. Reduce optimization difficulty of deeper networks
- ○ B. Alleviate gradient vanishing problem
- ○ C. Reuse lower level features
- ⦿ D. None of the above ✓

Your answer is correct.
정답은 None of the above.입니다.

Bottleneck architecture in inception module is useful for increasing non-linearity only.

하나를 선택하세요.
- ○ 참
- ⦿ 거짓 ✓

정답은 '거짓'입니다.

Localization is a regression problem

하나를 선택하세요.
- ⦿ 참 ✓
- ○ 거짓

정답은 '참' 입니다.

Instance segmentation is the combination of panoptic and semantic segmentation.

하나를 선택하세요.
- ○ 참
- ⦿ 거짓 ✓

정답은 '거짓'입니다.