UNIT IV

VGG Nets are built on \_\_\_\_\_\_\_\_\_\_\_\_\_\_

* **Convolutional neural networks**
* Artificial Neural Network
* Forecasting
* KNN

What does VGG stands for \_\_\_\_\_\_\_\_\_?

* **Visual Geometry Group**
* Virtual Geometry Group
* Visual Geography Group
* Visual Geometry Game

\_\_\_\_\_\_\_\_\_\_is based on the VGG network?

* Alex net
* **Resnet**
* Zfnet
* Lenet

VGG’s hidden layers use \_\_\_\_\_\_\_ activation function?

* Linear Function.
* Sigmoid
* Tanh
* **ReLU**

vgg 16 supports how many layers

* 19
* 20
* **16**
* 13

## Which is an Advantage of VGG?

* **Non-linearity increased as the number of layers with smaller kernels increased, which is always a good thing in deep learning.**
* The older VGG design is slower than the newer ResNet architecture, which introduced the idea of residual learning, another key accomplishment.
* The ResNet architecture was used to tackle the vanishing gradient problem

What is kernel size in VGG16?

* **3x3**
* 6x6
* 4x4
* 2x2

What dataset is VGG16 trained on?

* **The ImageNet dataset**
* MNIST dataset
* CIFAR-10 dataset
* Stanford Cars dataset

Applications of VGG in real life?

* **image and video recognition**
* regression
* gesture control
* distance sensing

Which model is better than VGG16?

* **ResNet50**
* VGG19
* VGG16
* VGG13

1)How many GB is ImageNet?For Large Learning Capacity ,Convolutional Neural Networks consists of classes of models and their capacity can be controlled by

* Depth and Height
* Height and Width
* Breadth and Width
* Depth and Breadth

**Ans: Depth and Breadth**

2)How many GB is ImageNet?

* 100gb
* 200gb
* 150gb
* 250gb

**Ans: 150gb**

3)ImageNet contain enough example to train such models will cause without

* Severe Underfitting
* Severe Overfitting
* Lenient Overfitting
* Lenient Underfitting

**Ans:Severe Overfitting**

4)The formula for non-linearity which works particularly well with their type of normalization is

* f(x)= |(1- e^-x)^-2)|
* f(x)= |(1- e^x)^-1)|
* f(x)=|sinh x cosh x|
* f(x)= 1/(|sinh x cosh x|)

**Ans: f(x)=|sinh x cosh x|**

5)The test directory in an imagenet dataset contains data points and columns

* 100 and 1000
* 50 and 200
* 70 and 300
* 90 and 400

**ANS: 1) 100 and 1000**

6)What is the abbreviation of IISVRC?

* ImageNet Large Scale Visual Recognition Challenge
* ImageNet Least Scale Visual Recognition Challenge
* ImageNet Least Scale Virtual Recognition Challenge
* ImageNet Large Scale Virtual Recognition Challenge

**ANS :ImageNet Large Scale Visual Recognition Challenge**

7)How much accuracy is good for CNN?

* 75%
* 65%
* 95%
* 80%

**ANS: 95%**

8)ImageNet contains **\_\_\_\_\_** labeled High-Resolution Images in over 22,000 categories ?

* Over 16 Billion
* Over 15 Million
* Over 20 Billion
* Over 14 Million

**Ans: Over 15 Million**

9)Which network has the highest accuracy on ImageNet dataset?

* GoogLeNet/Inception
* alex net
* le net
* wordNet

**Ans: googlenet/inception**

**10)**What is ImageNet dataset used for?

* computer vision research.
* optical recognition
* filter oriented approach
* Natural language processing

**Ans : computer vision research**

11)What is the size of images in ImageNet?

* 224x224
* 260x260
* 234x234
* 520x520

**ANS : 224x224**

12)Who created ImageNet?

* Dr.Andrej Karpathy
* Dr.Jeremy Howard
* Dr. Fei-Fei Li
* Dr.Demis Hassabis

**Ans:Dr. Fei-Fei Li**

13 ) Choose the incorrect application for ImageNet?

* Object Recognition
* Image recognition
* Object Localisation
* Image Localisation

**Ans: Image Localisation**

Identify the reinforcement learning technique which is based on the Q-learning algorithm

**Deep Q-Network**

Linear discriminant analysis

Probabilistic clustering

Deep Deterministic Policy Gradient

GAN stands for

Generative Advertising Network

**Generative Adversarial Network**

Generate Adversarial Network

Generation Adversarial Network

Identify the most effective and well-liked GAN implementation among the following

Conditional GAN

**Deep Convolutional GAN**

Laplacian Pyramid GAN

Vanilla GAN

Name the following that is utilized to become familiar with generative models and also explain how data is produced using probabilistic models

Adversarial

**Generative**

Networks

Discriminator

Identify the need for DCGANs

**DCGANs overcomes the problem of mode collapse**

The model parameters oscillate, destabilize and never converge

The generator collapses which produces limited varieties of samples, and highly sensitive to the hyper parameter selections.

Conversion from text to speech is very easy.

Spot the style that is applied to the content image

Gram matrix

**Van gogh**

Content cost

Style cost

Identify the two networks that DCGAN uses for generating real images

Discriminator, networks

Generative, Adversarial

Adversarial, Networks

**Generator, discriminator**

How many number of parts can Generative Adversarial Networks be separated

4

1

2

**3**

Identify the AI technology which was the first computer program to defeat a professional human Go player and was developed by DeepMind Technologies

AlphaGo

DeepBlue

BKG 9.8

Libratus

DQN is trained with

**RMSProp**

Loss function

Normalisation method

CNN model

Which layer spends more time in both inference phase and training phase?

pooling layer

fully connected layer

**convolution layer**

normalization layer

Which of the following statements is true?

**GPUs are good at such massive parallel operations**

        GPUs are good at continuous and complicated calculations.

        CPU is better to accelerate deep learning operations than GPU

        CPU and GPU are not advisable to accelerate deep learning operations

How can a training be accelerated ?

        Using CPU and less datasets

**Using GPU and cuDNN library**

 Using multiple CPUs

        Using less libraries

\_\_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_  are bottlenecks in accelerating deep learning.

         GPUs , Frameworks

**Memory space , Bus bandwidth**

    Public Broadbands, Private Networks

         CPUs, memory

  How can the processing time in convolutional layer be reduced?

        Not allocating the GPU to convolution layer

**Quickly and efficiently calculating the multiply-accumulative operations**

        Allocating GPU to convolution layer at last

        Using a faster processing unit

To further accelerate the calculations required for deep learning, you may want to

**Distribute them among multiple GPUs or machines**

       Use one GPU/machine

       Use less datasets

      Using frameworks

Which of the following framework is developed to focus on distributed training?

      Google cloud architecture framework

      Google heart framework

**Microsoft's Computational Network Toolkit (CNTK)**

      Microsoft's .NET framework

Which of the following are advantages of tenserflow?

**Solves communication and data synchronization between machines.**

     Missing symbolic loops

     Benchmark tests

     No support for OpenCL

 Which of the following is sufficient in deep learning?

   32-bit single-precision floating-point representations

   64-bit double-precision floating-point representations

**16-bit half-precision floating-point representations**

 128-bit triple-precision floating-point representations

What is the technique used in reducing the bit number in deep learning ?

**binarized neural network**

     distributed training

     Im2col

     multiply-accumulate operations

By how much bit does  binarized neural network  represent the weights and intermediate data?

**1\_ bit**

 2\_bit

      5\_bit

      8\_bit

What is distinctive about GoogLeNet?

Depth in the horizontal direction

Depth in the vertical direction

**Depth in both vertical and horizontal direction**

Depth in neither vertical nor horizontal direction

What is also known as inception architecture?

ResNet

**GoogLeNet**

VGG

ImageNet

What is the use of 1x1 filter in GoogLeNet?

Reduce the number of parameters

Accelerate processing

**Both reduce the parameters and accelerate processing**

To maintain the input size

How many layers does GoogLeNet consists of ?

11

**22**

20

42

Which activation function is used in googlenet?

TanH

Sigmoid

**ReLu**

Leaky ReLu

What is the disadvantage of GoogLeNet

Input size limit

**Over fitting**

Load balancing

Complexity

Which activation function is used in the final layer of GoogLeNet architecture

Leaky ReLu

Parameterized ReLu

ReLu

**Softmax**

Which max pooling is used in GoogLeNet

1x1

2x2

**3x3**

5x5

What is the default image size for GoogLeNet

128x128

512x512

**224x224**

256x256

**Making the network deeper**

Multiple choice questions :

Which initializer is used to initialize the weight in deeper networks ?

She initializer

**He initializer**

Glorat initializer

Kernel initializer

What characteristics can be seen in a network when “He initialaizer” is used to initialize weights and Adam to update weight parameters ?

Sigmoid function is used

Convolution layers which use small 2×2 filters

**A dropout layer used after a fully connected layer**

“He initializer” for initial bias values

Which of the following characteristics does not contribute to the improvement of recognition accuracy?

Learning rate decay

Ensemble learning

**Random forests**

Data augmentation

Which of the following is used for improving the recognition accuracy ?

**Learning rate decay**

Random forest

Gradient boost

Data preprocessing

Which of the following is not an advantage of making a network deeper?

Expanding the receptive field

Reducing the number of parameters

**increasing the number of parameters**

Training efficiency

Which of the following is the use of data augmentation ?

Algorithm used to expand filter size

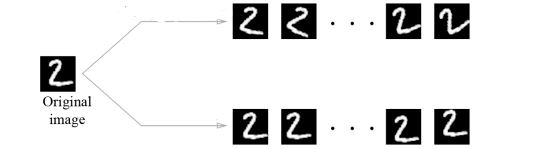
**Algorithm used to expand input image size artificially**

Algorithm for data preprocessing

Algorithm for combining weight and bias

What changes (i) and (ii) are being done in the following input image?

(i)



(ii)

Change by movement and change by rotation

**Change by rotation and change by movement**

Change by scaling and change by flipping

Change by flipping and change by scaling

list one Advantage of making a network deeper

**Training efficiency**

Reduce overfitting

It increases the number of parameter

It compress the receptive field

How can data augmentation be used to improve the recognition accuracy?

Increasing the number of testing images

**Increasing the number of training images**

Increasing the number of training and testing images

Reducing the number of training and testing images

How can we get a deeper network?

By applying a large filter several times

**By applying a small filter several times**

By applying a large filter few times

By applying a small filter few times

you can divide the problem to learn hierarchically by \_\_\_\_\_\_\_\_\_\_\_

Making the network simple

Adding the network with more hidden layers

Adding the network with less hidden layers

**Making the network deeper**

How many parameters are obtained when two 3x3 convolution operations are done?

6

9

11

**18**

Find the technique provided by a deeper network for correct training ?

**Big data**

IOT

Machine learning

Deep learning

 In a deep learning frameworks Support\_\_\_\_\_\_\_\_\_\_\_\_\_\_ To Process A Large Number Of Operations Quickly.

**GPU's**

          CPU

          NVIDIA

          CUDA

\_\_\_\_\_\_\_\_\_\_\_indicates a fully connected layer.

           conv

           pool

           norm

**fc**

convolution layers spend a lot of time in \_\_\_\_\_\_\_\_\_ Net

         GoogLeNet

**AlexNet**

         ImageNet

         ResNet

The time that AlexNet took for learning between a \_\_\_\_\_ and \_\_\_\_\_.

          SVMs and CNN

**GPU and CPU**

           NIC and RNN

          GAN and DCGAN

R-CNN converts an image into squares and uses \_\_\_\_\_\_\_\_\_\_\_ for classification.

**support vector machines**

linear regression

logistic regression

Fully Convolutional Network

To reduce such useless calculations, a technique called a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Fully Convolutional Network**

Highway network

DenseNet

Residual attention NN

A model called \_\_\_\_\_\_\_\_\_\_\_\_ is typically used to generate image captions for deep learning.

**Neural Image Caption**

Pyramidal Net

VGG

PolyNet

NIC consists of a deep CNN and a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ for handling natural language.

**Recurrent Neural Network**

Squeeze and Excitation Network

Competitive Squeeze and Excitation Network

Neural Image Caption

Handling various types of information, such as images and natural language,

is called \_\_\_\_\_\_\_\_\_\_

**multi-modal processing**

Transfer learning approach

Radial function Network

Self organizing Maps

Process flow of R-CNN

**Input image – Extract region proposal – Compute CNN features – Classify regions**

Extract region proposal -- Input image – Compute CNN features – Classify regions

Input image – Classify regions -- Extract region proposal – Compute CNN features

Input image – Extract region proposal – Classify regions -- Compute CNN features

What is Digital Image Processing?

**It’s a software that allows altering digital pictures**

It’s an application that alters digital videos

It’s a system that manipulates digital medias

It’s a machine that allows altering digital images

Which of the following image processing approaches is the fastest, most accurate, and flexible?

**Digital**

Photographic

Electronic

Optical

A model called \_\_\_\_\_\_\_\_\_ is typically used to generate image captions for deep learning.

**NIC**

              RNN

              CNN

              GAN

 \_\_\_\_\_\_\_\_\_\_ Maxwell generation of GPUs supported the storage of half-accuracy floating-point numbers

**NVIDIA's**

                     CPU,s

                     CVPR

                     RNN

Experiments have shown that 16-bit \_\_\_\_\_\_\_ are sufficient in deep learning

**half float**

                 Single float

                 Double float

                precision float

**ResNet**

What is the minimum error rate recorded for the ensemble of Residual Networks as of 2015?

  7.3%

**3.57%**

  6.7%

  2.73%

Overcomes the difficulties of performance when network becomes too deep,

ResNet exclusively introduces the

Inception Architecture

**Skip Architecture**

Pooling Layers

Shallow Architecture

ResNet is a network developed by

Preferred Networks

Facebook

**Microsoft**

DeepMind

Which state-of-the-art architecture is built on ResNet?

DenseNet

**Highway Network**

EfficientNet

GoogLeNet

What is the use of implementing skip connection in RES net?

**Help skipping unwanted Convulutional layers by regularization**

Help reduce the number of parameters

Helps in image scaling and feature scaling

Helps accelerate learning

RES net is inspired from which other neural  network?

**VGG neural network**

  GoogleLeNet

  ImageNet

  All of the above

Which optimizer is best for Resnet?

  Adagrad

**Momentum**

  SGD

  Adam

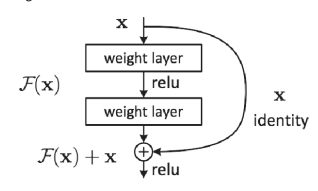
What is the problem that Resnet allows training of such deep networks by constructing the network through modules?

**Degradation problem**

  Long training Time

  Catastrophic forgetting

  Large trained Models



From the above image, from F(x)+x, what +x denotes?

  Input of previous layer

  Output of previous layer

**Skip connection**

  Bais coefficient

ResNet adopts the skip architecture. This in turn alleviates the

 complication.

  Null Value problem

  Gradient Towering problem

**Gradient Vanishing problem**

  Misspecification problem

The Skip architecture to pass the gradient as it is, it skips input data from the

**Weight layer**

Input Layer

Output Layer

Pooling Layer

UNIT V

1. The universal way to generate sequence data in deep learning is to train a model

to predict the \_\_\_\_\_\_ or \_\_\_\_\_\_\_ in a sequence, using the \_\_\_\_\_\_ as input.

a) next word, previous words, next few words

**b) next token, next few tokens, previous tokens**

c) next character, next few characters, previous characters

d) next word, next few words, previous words

2. A classical image auto encoder takes an image, maps it to a \_\_\_\_\_\_\_\_\_\_\_ via

an \_\_\_\_\_\_\_\_\_\_ module, and then decodes it back to an output with the same dimensions

as the original image, via a \_\_\_\_\_\_\_\_\_\_\_ module

a) Latent vector space, decoder, encoder

b) Vector space, decoder, encoder

**c) Latent vector space, encoder, decoder**

d) Vector space, encoder, decoder

3. \_\_\_\_\_\_\_\_\_\_\_\_\_ enable the generation of fairly realistic synthetic images by forcing the

generated images to be statistically almost indistinguishable from real ones.

**a) General Adversarial Networks**

b) Autoencoder

c) Autodecoder

d) Diffusion Model Networks

4. \_\_\_\_\_\_\_\_\_\_\_ optimization is a powerful technique that is an absolute requirement for

getting to state-of-the-art models on any task

a) Batch

**b) Hyperparameter**

c) GPU

d) CPU

5. What is a TPU ?

a) Template Production Unit

b) Tuple Processing Unit

**c) Tensor Processing Unit**

d) Tensor Production Unit

6.\_\_\_\_\_\_\_ is an artistic image modification technique that uses the representations learned by convolutional neural networks.

1. Deepdream
2. Neural style transfer
3. Autoencoders
4. Textgeneration

Ans: a) Deepdream

7.Neural style transfer is used by which network?

1. VGG19
2. VGG16
3. VGG10
4. VGG12

Ans:a) VGG19

8.In scaling up training model the double precision consists of how many bits?

1. 16bits
2. 64bits
3. 32bits
4. 8bits

Ans: c) 32bits

9.Which network is used to work by work by processing sequences of inputs one timestep at a Time, and maintaining a state throughout.

1. Recurrent neural networks (RNNs)
2. Multilayer perceptrons
3. Convolutional neural network(CNN)
4. Deep Neural Network(DNN)

Ans: a) Recurrent neural networks (RNNs)

10.Any network that Can model the probability of the next token given the previous one is

Called?

a)latent space

b)language model

c)soft max

d)entropy model

ans: b) language model

11.What is the feature that Google released in 2016, capable of automatically generating a selection of quick replies to emails or text messages?

a)Automatic email filters and categories.

b)Event details used to create calendar events

c)Smart Reply in email.

d)Writing and intelligent search suggestions.

Ans: c)

12.Any network that can model the probability of the next token given the previous ones is called

a)softmax

b)entropy model

c)language model

d)latent space

Ans:c)

13.How Neural Style transfer works?

a)Applying the style of a target image to a reference image while conserving the content of the target image.

b)Applying the style of a reference image to a target image while conserving the content of the target image.

c)Applying the style of a reference image to a target image while conserving the style of the reference image.

d)Applying the style of a target image to a reference image while conserving the style of the reference image.

Ans:b)

14.Neural Style transfer is used by which network?

a) VGG19

b)VGG16

c)VGG10

d)VGG12

Ans:a).

15.which of the following are sequence models ?

a.RNNs and Transformers

b.IMDB

c.NLP

d.TensorFlow

Answer a.RNN and transformers

16.what is the simplest way to encode a piece of text for processing ?

a.to represent an entire text as a single vector,

b.TextVectorization method

c.to discard order and treat it as a set (a “bag”) of tokens.

d.to write a reusable model-building function

Answer c.to discard order and treat it as a set (a “bag”) of tokens.

17. when you encode something via one-hot encoding, you’re making a

feature-engineering decision. You’re injecting into your model a

fundamental assumption about the structure of your feature space. What is

the assumption ?

a.the different tokens which are encoding are all independent from each

other

b.they are dependent

c.they share information with each other.

d.vectors should reflect the semantic relationship between these words.

Answer a.the different tokens which are encoding are all independent

from each other

~~18.The Embedding layer is best understood as a dictionary that maps~~

~~integer indices (which stand for specific words) to dense vectors. It takes integers as input,~~

~~looks up these integers in an internal dictionary, and returns the associated vectors.~~

~~It’s effectively a dictionary lookup. Look the flowchart .~~

~~Word index —----Embedding layer —------ X~~

~~Find x~~

~~a.word-embedding space~~

~~b.backpropagation~~

~~c.Corresponding word vector~~

~~d.RNN~~

~~Answer c.Corresponding word vector~~

19) Which artistic image-editing method that makes advantage of the representations that convolutional neural networks have learned.

a) Vision Transformers

b)DeepDream

c)DALL.E

d)Deepimg

ans:- b) DeepDream

20)The estimated second order gradient information used by which method that allows for a quicker convergence to the minimum.

a)L-BFGS

b)LBG

c)DCB

d)RFGS

ans:- a)L-BFGS

21)How many layers are there in Autoencoder?

a)2

b)3

c)4

d)5

ans:-b) 3

22) Autoencoders are trained using \_\_\_\_

a) Feed Forward

b) Reconstruction

c) Back Propagation

d) They do not require Training

ans:- c) Back propagation

23)GNN stands for \_\_\_\_\_

a) Generative Advertising Network

b) Generative Adversarial Network

c) Generate Adversarial Network

d) Generation adversarial Network

ans:- a) Generative Adversarial Network

24.Which sampling introduces randomness in the sampling process by sampling from

the probability distribution for the next character.

a)stochastic sampling

b)systematic sampling

c)stratified sampling

d) cluster sampling

ans:a)

~~25.\_\_\_\_\_\_\_\_\_\_\_\_\_\_ is an artistic image-modification technique that uses the~~

~~representations learned by convolutional neural networks.~~

~~a)Deepest Dream~~

~~b)Deep Dream~~

~~c)DeepPy~~

~~d)Clarifai~~

~~ans:b)~~

26.Which model is based on either words or characters ?

a) Dynamic model

b) Words model

c**) language model**

d) Neural model

27.In 2002 Douglas Eck and then at which lab in Switzerland, applied LSTM to

music generation for the first time with promising results?

a).Alex graves

b).Magenta

c).Keras

d).Schmidhuber

ANS: d).SCHMIDHUBER

28.Any network that can model the probability of the next token given the previous ones is

called ?

a).conditioning data

b).language model

c).linear model

d).regression model

ANS: b).LANGUAGE MODEL

29.Introducing randomness in the sampling process by sampling from

the probability distribution for the next character this is called ?

a).stochastic sampling

b).decision tree

c).regression tree

d).clustering

ANS: a).STOCHASTIC SAMPLING

30.In order to control the amount of stochasticity in the sampling process, we’ll intro-

duce a parameter called ?

a).softmax temperature

b).temperature increase

c).stochastic temperature

d).keras temperature

ANS: a).SOFTMAX TEMPERATURE

31. Who introduced neural style transfer in the summer of 2015?

a).john carles

b).robin megas

c).Leon Gatys et al

d).benjamin.

ANS: c).LEON GATYS ET AL

32)Applications of sequence data generation with recurrent networks

a)SLTM

b)CNN

c)RNN

d)LSTM

33)Who did pioneering work on using recurrent networks for sequence data generation?

a)Alex Graves

b)Geoffrey Hinton

c) John McCarthy

d)Arthur Samuel

34)which layer to compute the vocabulary in networks?

a)regrex\_replace

b)tenserflow

c)textvectorization

d)NLP

35)which network was introduced in 2014 by Goodfellow et al.,7

a)CNN

b)GAN

c)Generative adversarial

d)Feed forward Neural network

36)which attempts to provide premade search spaces that are

relevant to broad categories of problems such as image classification.

a)Activation Function

b)KerasTuner

c)Softmax

d)Kerasflow

37.If we were able to mathematically define content and style, then an appropriate loss function

to minimize would be the following:

Loss =

a)(distance(style(reference\_image) - style(combination\_image)) +

distance(content(original\_image) - content(combination\_image)))

b)(distance(style(reference\_image) - style(combination\_image)) -

distance(content(original\_image) - content(combination\_image)))

c)(distance(style(reference\_image) + style(combination\_image)) +

distance(content(original\_image) + content(combination\_image)))

d)(distance(style(reference\_image) + style(combination\_image)) +

distance(content(original\_image) - content(combination\_image)))

Ans:a)

38.The content loss only uses a \_\_\_\_\_\_-upper layer, but the style loss as defined by Gatys

et al.

a)Quadrables

b)Double

c)Triple

d)Single

Ans:d)

39.What is the full form of VAE?

a)Variational Auto Encoders

b)Variation Auto Encoders

c)Vectored Auto Encoders

d)Variation Automatic Encoders

Ans:a)

40.\_\_\_\_\_\_\_\_\_\_\_\_ is an unsupervised learning technique for neural networks that learns efficient

data representations (encoding) by training the network to ignore signal “noise.”

a)Rotation

b)Convolutional

c)Autoencoder

d)Linear

Ans:c)

41.\_\_\_\_\_\_\_\_\_\_\_ is used as the loss function of the reconstructed process in autoencoder.

a)Mean Square Error

b)Variational Auto Encoders

c)Generative Adversarial Networks

d)Artificial Neural Network

Ans:a)

42.Applying the style of a reference image to a target image while maintaining the target image's information is known as **neural style transfer**.

a)deep dream

**b)neural style transfer**

c)model ensembling

d)adversial networks

**43. Style transfer** entails producing a new image that captures both the reference image's style and the contents of the target image.

**a)style transfer**

b)style loss

c)content loss

d)variation loss

44.An encoder module converts the input sample input image into two parameters, **z mean and z log variance**, in a latent space of representations.

A)z mode and z log variance

**b)z mean and z log variance**

c)x z mode and x z log variance

d) xz mean and xz log variance

45.Using a random vector or random point in the latent space as input, the **generator network** decodes the data into a synthetic image.

a)discriminator network

b)adversarial network

**c)generator network**

d)neural network

46.Since discrete choices often make up the hyperparameter space, it is neither continuous nor differentiable As a result, **gradient descent** is often not possible in hyperparameter space.

**a)gradient descent**

b)bagging

c)boosting

d)stacking