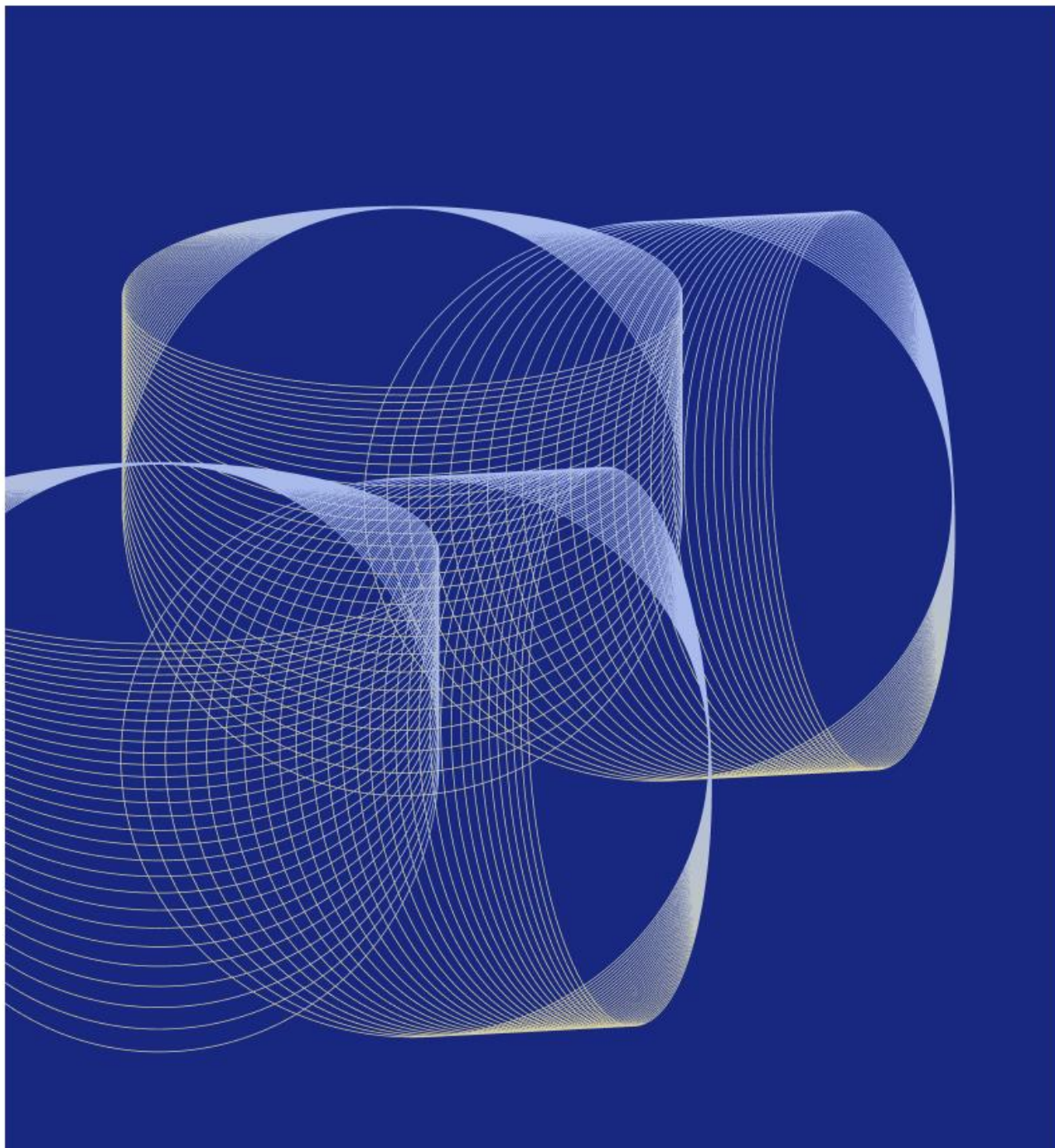


Univ.AI

AI Quizbook

2020 | Issue 3



Founded by Harvard Professors
to bring you **the world's best**
AI/ML training



Rahul Dave
Chief Scientist,
Univ.AI
Former Lecturer,
Harvard University



Pavlos Protopapas
Scientific Program
Director, IACS,
Harvard University



Achuta Kadambi
Asst. Prof. UCLA
PhD, MIT Media Lab

All Programs
Live! Online

Elite Faculty
from Top
Universities

Cutting-Edge
Curriculum

Live!
Hands-On
Mentorship

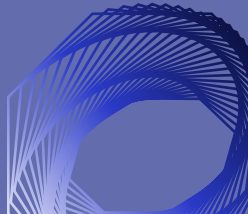
Learning by
Lots of
Doing!

Train to be among the best

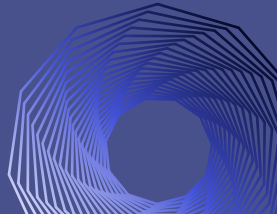
Univ.AI offers you a 3-step path to a GUARANTEED high-end job in Machine Learning & Artificial Intelligence, Data Science, or Programming (Coming Soon)

- Learn LIVE-online with Harvard faculty
- Gain expertise in 3 sequential, standalone certifications of 10-12 weeks each
- Get job-ready outcomes and certifications at each level
- The beginner programs are offered at very low costs
- Advance to the Master Certification, risk your investment with a state-of-the-art employment guarantee

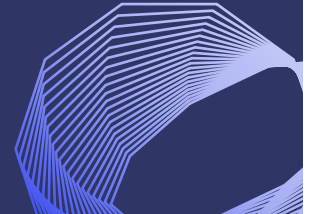
Artificial
Intelligence: X



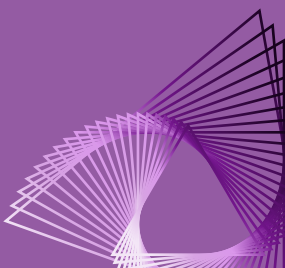
Artificial
Intelligence: Y



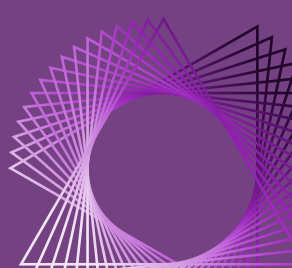
Artificial
Intelligence: Z



Data Science: X



Data Science: Y



Data Science: Z



The output of which of the following activation functions is zero-centred?

☐

Tanh

☐

Sigmoid

☐

ReLU

☐

None of these

Answer on the next page

Why shouldn't we have stride value greater than K , the filter dimension?

☐

It will need more training

☐

It will consume more time

☐

It will leave gaps in between two parts of the image to be processed, and we might lose valuable information

☐

None of the above

The input image has been converted into a matrix of size 28 X 28 and a kernel/filter of size 7 X 7 with a stride of 1, with no padding. What will be the size of the convoluted matrix?

☐

22 x 22

☐

21 x 21

☐

28 x 28

☐

7 x 7

Previous Answer

It will leave gaps in between two parts of the image to be processed, and we might lose valuable information.

Which of the following is true with respect to a GRU (Gated Recurrent Unit)?

1. Units with short-term dependencies will have active reset gates
2. Units with long-term dependencies will have active reset gates
3. Units with short-term dependencies will have active update gates
4. Units with long-term dependencies will have active update gates

☐

1 and 4

☐

2 and 3

☐

1 and 2

☐

3 and 4

Previous Answer
22 x 22

Suppose we are generating text, one character at a time. There are C possible characters to generate. If the last hidden layer has H units, how much computation is required by a fully connected softmax output layer?

☐ $O(H)$ ☐ $O(C^2)$ ☐ $O(C*H)$ ☐ $O(C)$

Previous Answer
1 and 4

If we imagine gradient descent as a ball rolling down the hill, which parameter controls how fast the ball rolls down the hill?

Here W is a weight (or weights), α is a positive number, and $\frac{\partial L}{\partial w}$ is the change of the loss with weight W calculated at the current value of the weights.

☐ α ☐ w ☐ L ☐ $\frac{\partial L}{\partial w}$

Previous Answer
O(C*H)

Imagine a classification problem with a highly imbalanced class (i.e the majority class is observed 99% of the time in the training data). Which of the following is true in such a case?

☐

Accuracy is a good metric for imbalanced class problems

☐

Precision and Recall are good metrics for imbalanced class problems

☐

Both of the above

☐

None of the above

Previous Answer

α

Changing Sigmoid activation to ReLU will help to get over the vanishing gradient issue. True or False?

☐

True

☐

False

What is a consequence of tuning hyperparameters using a test dataset?

☐

It overfits on the test data

☐

The model over-generalizes on the train data

☐

The performance on a different test dataset will increase

☐

None of the above

Which of the following applications can we most likely program into a robot using a **single** feed-forward network?

☐

Rotating arms and legs

☐

Watering plants

☐

Following a wall

☐

Dancing to a tune

Previous Answer
It overfits on the test data

In the following sentence, fill in the blanks with “increases”, “decreases” or “makes no change to” to explain the effect of the operation on the bias and variance.

Using dropout to train a deep neural network _____ the bias and _____ the variance.

Which of the following hyper-parameters, when increased may cause overfitting of data when random forest is used?

1. Depth of tree
2. Number of trees

☐

Only 1

☐

Only 2

☐

Both 1 and 2

☐

None of the above

How can you prevent a clustering algorithm from getting stuck in bad local optima?

☐

Use multiple random initializations

☐

Increase the number of iterations

☐

Both of the above

☐

None of the above

Previous Answer
Only 1

If the number of hidden layers is increased in a multi-layer perceptron, the error on test data always decreases. True or False?

☐

True

☐

False

Previous Answer

Use multiple random initializations

In which neural net architecture does weight sharing occur?

☐

Convolutional Neural Networks

☐

Fully Connected Neural Networks

☐

Both of the above

☐

None of the above

Previous Answer
False

Which of the following are true of convolutional neural networks (CNNs) for image analysis?

Select one or more.

☐

Filters in earlier layers tend to include edge detectors

☐

Pooling layers reduce the spatial resolution of the image

☐

They have more parameters than fully connected networks with the same number of layers and the same numbers of neurons in each layer

☐

A CNN can be trained for unsupervised learning tasks, whereas an ordinary neural net cannot

Which of the following is true about k-NN?

☐

The classification accuracy is better with larger values of k

☐

The decision boundary is smoother with smaller values of k

☐

The decision boundary is linear

☐

k-NN does not require an explicit training step

Previous Answer

Filters in earlier layers tend to include edge detectors

AND

Pooling layers reduce the spatial resolution of the image

Which of the following is a deterministic algorithm?

☐

PCA

☐

K-Means

☐

Both of the above

☐

None of the above

Previous Answer

k-NN does not require an explicit training step

Which of the following filters can be used for vertical edge detection in images?

☐

$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$

☐

$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$

☐

$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

☐

$\begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$

What are the axes of an ROC curve?

☐

Vertical axis: True negative rate;
Horizontal axis: False negative rate

☐

Vertical axis: True positive rate;
Horizontal axis: False positive rate

☐

Vertical axis: False negative rate;
Horizontal axis: False positive rate

☐

Vertical axis: False positive rate;
Horizontal axis: True negative rate

Univ. AI

To learn more, drop us a note at hello@univ.ai

Previous Answer

Vertical axis: True positive rate;
Horizontal axis: False positive rate