

MACHINE LEARNING QUESTIONS

[Home](#) [Competition](#) [Past Questions](#) [Sponsorships](#) [Login](#)

IT'S MARCH 26 SOMEWHERE AND HERE IS TODAY'S QUESTION

TODAY'S QUESTION

Consulting firm

Amina works at an HR consulting firm where she spends most days processing job descriptions.

She has always been interested in using machine learning to streamline her work and make it more efficient. She has heard that transformers are a powerful type of deep learning model that can be used for natural language processing tasks.

However, Amina is not very familiar with transformers and wants to know more about their fundamental characteristics before she decides whether to use them for her job.

Which of the following is not a key feature of transformers?

- Transformers use self-attention mechanisms to weigh the input features differently.
- Transformers can process sequences of variable length in a parallel manner.
- Transformers can be trained to perform multiple tasks at once.

- Transformers use convolutional layers to process input data.
[CORRECT]

Good going!

Transformers are a type of deep learning model widely used in Natural Language processing tasks, such as machine translation and text summarization. They were introduced in the 2017 paper [Attention is All You Need](#) by Vaswani et al.

One of the key characteristics of transformers is their use of attention mechanisms. Attention allows the model to focus on specific input parts, which is useful when processing long data sequences. Another characteristic of transformers is their ability to handle variable-length sequences in parallel. This is useful when dealing with natural language data, which often has varying lengths.

We can train transformers to perform multiple tasks at once using a process known as multitask learning. In multitask learning, a single model is trained to perform multiple tasks simultaneously using a shared set of parameters. This allows the model to learn common features relevant to all tasks, which can improve the model's performance on each task.

Transformers do not use convolutional layers to process input data. Convolutional layers work by sliding a small window, or kernel, across the input data and performing an element-wise multiplication between the values in the kernel and the values in the input data. The resulting values are then summed and passed through a nonlinear activation function before being output as the output of the convolutional layer.

Recommended reading

- Check the "[Attention is All You Need](#)" paper by Vaswani et al. for an introduction to Transformers.
- The [Illustrated Transformer](#) is a fantastic visual guide that explains the inner workings of the Transformer architecture in an easy-to-understand manner.

Share this!

153 people have answered this question with 49% accuracy



SEE YOU TOMORROW FOR A NEW QUESTION

LOGIN WITH A MAGIC CODE AND GET ACCESS TO TIME TRAVEL!

Keep track of your progress, access tomorrow's question, and participate in the leaderboard.

☒ Send new questions to my inbox

Continue

We'll send you a magic code to your email

Join the other 120+ students that joined the cohort to learn how to **train, tune, deploy**, and **monitor** Machine Learning models in a production environment. Classes start April 17. [Join here!](#)

Next question in 23:43 hours. You are not logged in, [but you should](#).
For comments or feedback, contact us at hey@bnomial.com or [@0xbnomial](#).