



# Quiz review

Started on	Wednesday, 13 March 2024, 6:54 PM
State	Finished
Completed on	Wednesday, 13 March 2024, 6:57 PM
Time taken	2 mins 38 secs
Marks	10.00/10.00
Grade	100.00 out of 100.00

## Question 1

Correct  
Mark 1.00 out of 1.00

Which of the following models is designed for image generation?

- ☐ GPT
- ☒ DALL-E ✓
- ☐ T5
- ☐ BERT
- ☐ None of the options given

The correct answer is: DALL-E

## Question 2

Correct  
Mark 1.00 out of 1.00

What is the primary task BERT is designed for?

- ☐ Image generation
- ☐ Text generation
- ☒ Bidirectional understanding of text ✓
- ☐ None of the options given
- ☐ Language translation

The correct answer is: Bidirectional understanding of text



**Question 3**

Correct

Mark 1.00 out of 1.00

What is the main difference between pre-training and fine-tuning in Transformers?

- ☐ Fine-tuning is done without labeled data
- ☐ Pre-training uses smaller models
- ☐ None of the options given
- ☐ Both are done simultaneously
- ☒ Pre-training is on a large corpus and fine-tuning is task-specific ✓

The correct answer is: Pre-training is on a large corpus and fine-tuning is task-specific

**Question 4**

Correct

Mark 1.00 out of 1.00

47562

In the context of Transformers, what does "seq to seq" stand for?

- ☐ Sequential training
- ☐ None of the options given
- ☒ Sequence to Sequence ✓
- ☐ Sequential to Sequential
- ☐ Sequence training

The correct answer is: Sequence to Sequence

47562

**Question 5**

Correct

Mark 1.00 out of 1.00

Which Transformer model is known for generating coherent paragraphs of text?

- ☐ Image GPT
- ☐ T5
- ☒ GPT ✓
- ☐ DALL·E
- ☐ BERT

47562

The correct answer is: GPT

**Question 6**

Correct

Mark 1.00 out of 1.00

How does Multi-head attention differ from standard attention?

- ☒ It allows the model to focus on multiple parts of the input simultaneously ✓
- ☐ It uses fewer parameters
- ☐ It is only used in GPT
- ☐ None of the options given
- ☐ It is faster

The correct answer is: It allows the model to focus on multiple parts of the input simultaneously

**Question 7**

Correct

Mark 1.00 out of 1.00

47562

For which task might you use a Transformer to generate a concise summary of a long article?

- ☒ Summarization ✓
- ☐ Image Classification
- ☐ Translation
- ☐ None of the options given
- ☐ Question Answering

The correct answer is: Summarization

47562

**Question 8**

Correct

Mark 1.00 out of 1.00

Which model can be used for both image and text tasks?

- ☒ None of the options given ✓
- ☐ GPT
- ☐ BERT
- ☐ T5
- ☐ DALL-E

47562

The correct answer is: None of the options given

**Question 9**

Correct

Mark 1.00 out of 1.00

In sequence-to-sequence tasks, why is attention important?

- ☐ It simplifies the model
- ☒ It helps the model focus on relevant parts of the input ✓
- ☐ It reduces overfitting
- ☐ All of the options given
- ☐ It speeds up computation

The correct answer is: It helps the model focus on relevant parts of the input

**Question 10**

Correct

Mark 1.00 out of 1.00

47562

Which mechanism allows Transformers to weigh the importance of different words in a sequence?

- ☐ CNN layers
- ☐ None of the options given
- ☐ LSTM cells
- ☒ Self Attention Mechanism ✓
- ☐ RNN cells

The correct answer is: Self Attention Mechanism

47562

[◀ Transformers and Attention Mechanisms](#)[Case Study - Transformers in Machine Translation ▶](#)

47562