**SWS3009A Robotics and Deep Learning**

**Deep Learning Lab 2 Answer Book**

**SUBMISSION DEADLINE: SATURDAY 8 JULY 2023, 11.59 pm**

Team Member 1 Name:

Zheng Qinghe

Team Member 2 Name:

Zhao Qianyi

Team Member 3 Name (if any):

**Marks:** \_\_\_\_\_\_\_\_ / 3

**Please save as PDF before submitting to Canvas.**

**Question 1 Answer:**

1. Transformers rely on a self-attention mechanism that allows them to capture global dependencies between all positions in the input sequence. Each token in the sequence attends to all other tokens, allowing the model to learn relationships and dependencies across the entire sentence.

2. Transformers can process the entire sequence in parallel, which enables Transformers to handle longer sequences more efficiently.

3. Transformers capture contextual information by encoding each token with its surrounding context in a self-attention manner. This means that each token's representation is influenced by the entire sentence, allowing the model to understand the context and meaning of the words in relation to the entire sentence.

**Question 2 Answer:**

1. It can be more efficient if sentences are of fixed length.
2. The vectors are gathered together to form a matrix, and the dimension of the matrix needs to be fixed.
3. Using fixed-length sentences can save space consumption.

**Question 3 Answer:**

model\_name: The name or identifier of the model, which specifies the pretrained model to be loaded.

vocab\_size: The size of the vocabulary. This parameter needs to be set based on the tokenizer used and the vocabulary of the dataset.

n\_ctx: The maximum context length. It defines the maximum length of the input text that the model can handle. If the input text exceeds this length, it will be truncated or sliced to fit the model's input requirements.

bos\_token\_id: The token ID for the beginning-of-sequence marker. It indicates the starting position of the input text for the model.

eos\_token\_id: The token ID for the end-of-sequence marker. It indicates the ending position of the input text for the model.

**Question 4 Answer:**

(Fill answer here)