Name: | | SID: | |

COMP 4446 / 5046 Quiz 3 (week 8) - version B

If the choices have ◯ then select exactly one option. If the choices have □, select all correct options. Indicate your answer by filling the shape, e.g., ●. If you make a mistake, draw an X over your answer, e.g.,  .

1. (1 mark) How do self-attention and cross-attention differ? Select all true statements.
   - □ The similarity calculation.
   - ■ **The way the output is used.**
   - □ The way the weights (α) are used.
   - ■ **The source of the keys, values, and queries.**

2. (1 mark) What is the purpose of Q, K, and V matrices? Select all true statements.
   - □ To enable self-attention to account for the position of a word in the input.
   - ■ **To give flexibility in which parts of the vectors are used in each step.**
   - □ To introduce a non-linear step in the self-attention process.
   - □ To make learning easier by making it easier for gradients to propagate.

3. (1 mark) What is the purpose of teacher forcing? Select all true statements.
   - ■ **To help the model not produce output that is completely different from the true answer.**
   - □ To improve the speed of training.
   - □ To enable parallel processing in the decoder.
   - □ To make the dimensionality of inputs to the cell match at all steps.
   - □ To help the model not produce output that diverges too much from the original input.

   Solution: The second and fifth options could also be included. The first option must be selected.

4. (1 mark) Which of these parts of the transformer always contain a weight matrix that is learned in training? Select all true statements.
   - □ Positional encoding
   - ■ **Feedforward layers**
   - □ Layer normalisation
   - ■ **Self-attention**
   - □ Residual connections

   Solution: Note that layer normalisation sometimes uses trainable parameters, but not matrices.

5. (1 mark) Using the lines below, implement code that uses spaCy to count occurrences of pronouns in a provided string `text`. Provide your answer by writing the line numbers in the boxes to the right, in order, top to bottom.

```
1  nlp = spacy.load("en_core_web_sm")
2  counts[token.text] = counts.get(token.text, 0) + 1
3  counts[token.pos_] = counts.get(token.pos_, 0) + 1
4  counts = {}
5  for token in spaCy(text):
6  for pronoun in nlp(text):
7  for token in nlp(text):
8  for pronoun in spaCy(text):
9  if token.pos_ == "PRON":
10 if token.text in ["I", "we", "he", "she", "it", "you"]:
```

Solution: (1, 4), 7, 9, 2