Assignment 2

Deep Learning KU, WS 2025/26

Team Members		
Last name	First name	Matriculation Number
Anderson	Alice	1500001
Binford	Bob	1600002

Example: Including Code and Explanation in the Report

Task: Create a character-wise encoding of the input text. Each unique character should be represented by a distinct integer.

Code Listing:

```
1  # Create character mappings
2  chars = sorted(list(set(text_data)))
3  self.string_to_int = {ch: i for i, ch in enumerate(chars)}
4  self.int_to_string = {i: ch for i, ch in enumerate(chars)}
5
6  # Encode text to integers
7  encoded_data = [self.string_to_int[c] for c in text_data]
8
9  # Convert to tensor
10 self.data = torch.tensor(encoded_data, dtype=torch.long)
```

Explanation: The `set()` function in line 2 creates a set object containing each distinct character in our input text, which is then converted to a sorted list for easier handling. Lines 3 and 4 then define dictionaries that map each character to an integer and the other way around. Line 7 then applies this encoding of characters to integers to our dataset, and line 10 transforms this dataset into a tensor. We use `dtype=torch.long` because it represents full integer values and can support a rich dictionary (although smaller integer types would also work in this specific example).