Lecture - 9: Create Input - Target pair

The last step before we create vector embedding is to create input - target pair.

* What do these input - target pair look like ?

Input → LLM | Target/Output → learn | Text sample → 1st iteration → to predict one word at a time

LLM learn | to | predict one word at a time

LLM learn to | predict | one word at a time

LLM learn to predict | one | word at a time

(Input the LLM received: LLM learn to predict)

(The LLM can't access word pass the target.)

LLM learn to predict one | word | at a time

(target to predict: word)

LLM learn to predict one word | at | a time

LLM learn to predict one word at | a | time

LLM learn to predict one word at a | time

* This is auto-regressive model, or unsupervised learning

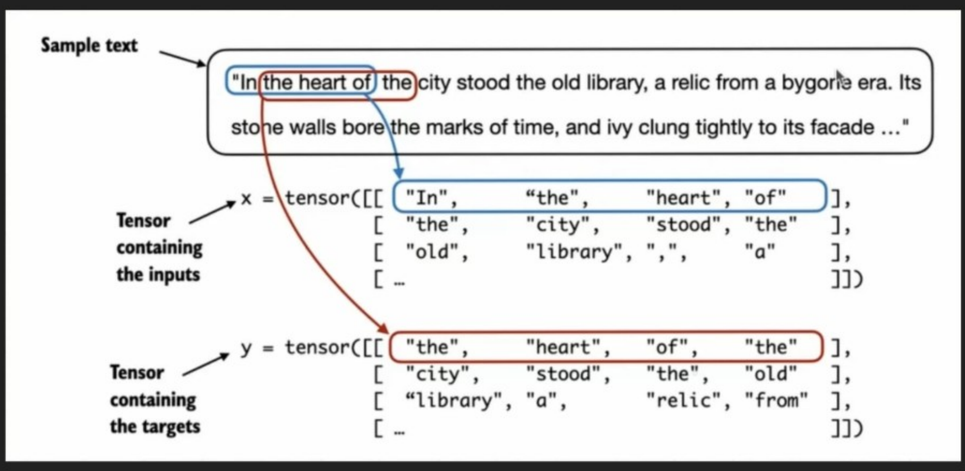
* Given a text sample

- extract input blocks as sub samples that serve as input to the LLM.
- The LLM prediction task during training to predict the next word that follows the input block
- During training, we mask out all word that are past the target

For practical - lecture - 9 code ←

* we will implement a data loader that fetches input-output target pairs using a sliding window approach

* Generating a Data Loader



To implement efficient dataloader, we collect input in a tensor x, where each row represent one input context. The second tensor y contain the corresponding prediction target (next word), which are created by shifting the input by one position.

* The meaning of Stride