



Project 3: Sentiment Analysis With Transformer

Announcement date: 10.09.2024

Due date: 11.03.2024

TA: Qian Wang



Task description

- We will do sentiment analysis on a movie review dataset.
- We want to train a transformer to understand the context of the sequence data.
- During inference, given a movie review, we want the transformer to output the probability of the sentiment of that review.



Transformers

- Transformers can effectively capture the relationships between words in a sequence.
- Unlike RNN that process inputs sequentially, transformers can directly attend to every word in the input sequence.
- Core to the transformers is the self-attention mechanism.
- Transformers add positional encodings to help the model understand the position of each word.



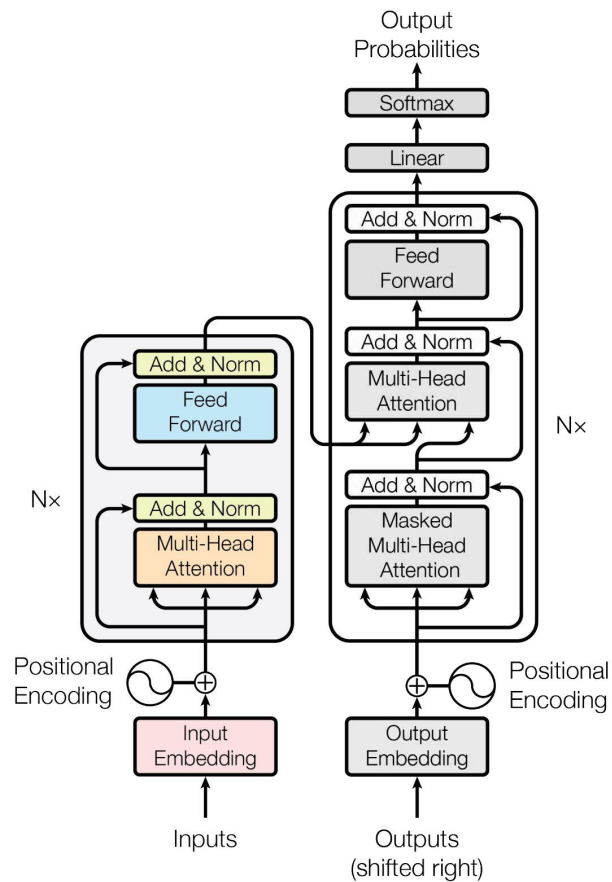
Different Transformers

- Encoder-only transformers: e.g., BERT. For tasks like text classification or Q&A.
- Decoder-only transformers: e.g., GPT. These are designed for generative tasks. No need for an encoder to process a separate input sequence.
- Encoder-Decoder transformers: e.g., T5. These are used for sequence-to-sequence tasks, such as translation or summarization.



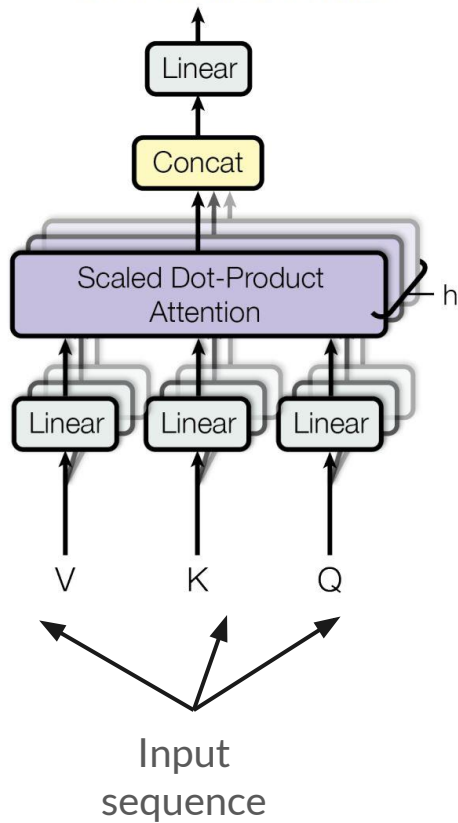
Which transformer do we use?

- Our focus is on encoding input sequence and understanding the context, not generating new sequences.
- We only use the encoder part of the transformer.

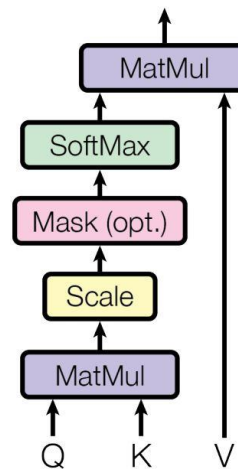




Multi-Head Attention

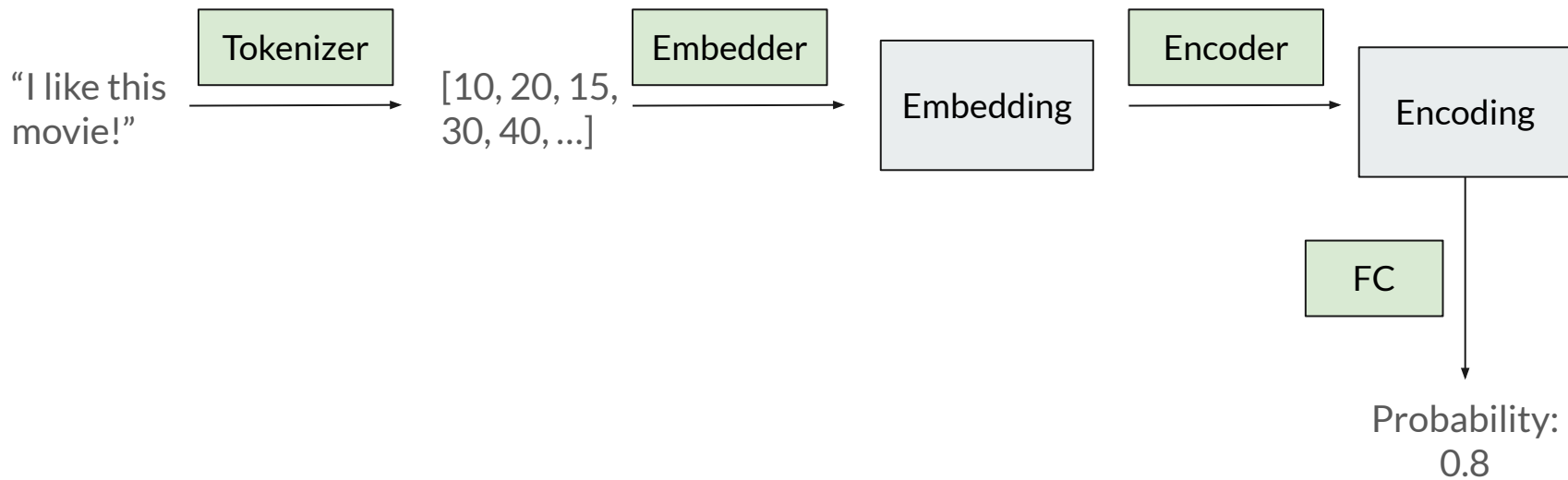


Scaled Dot-Product Attention





Pipeline





Question:

- Since the embedder can learn continuous and meaningful representations of individual words, what will happen if we remove the transformer encoder and only use the embedder for this task?