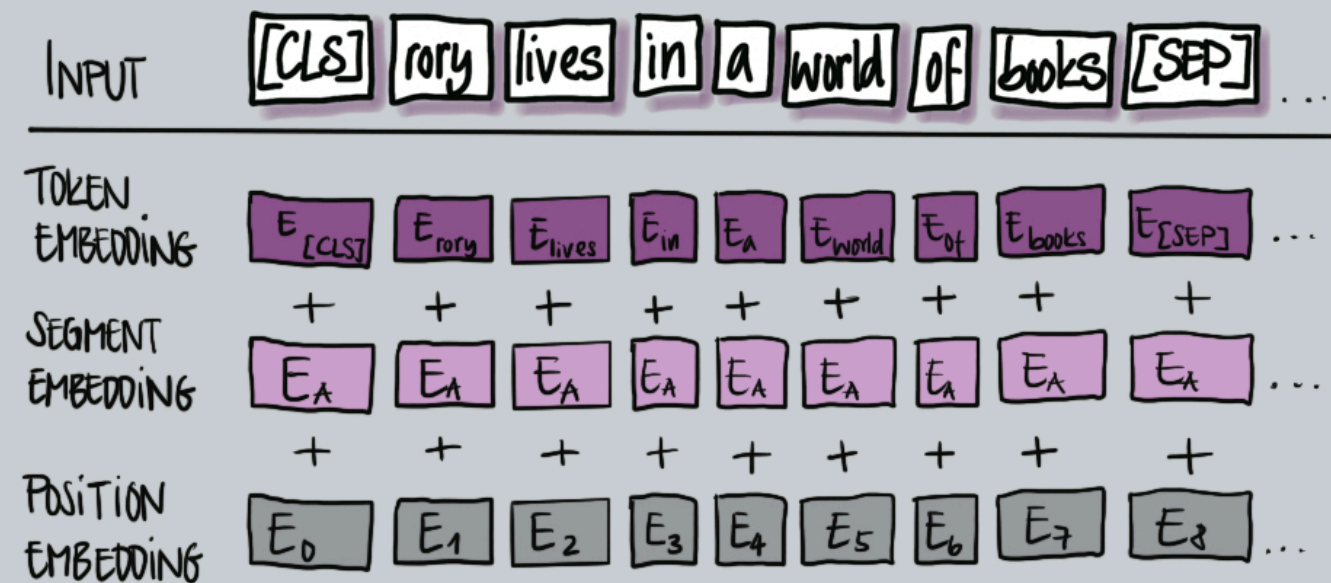
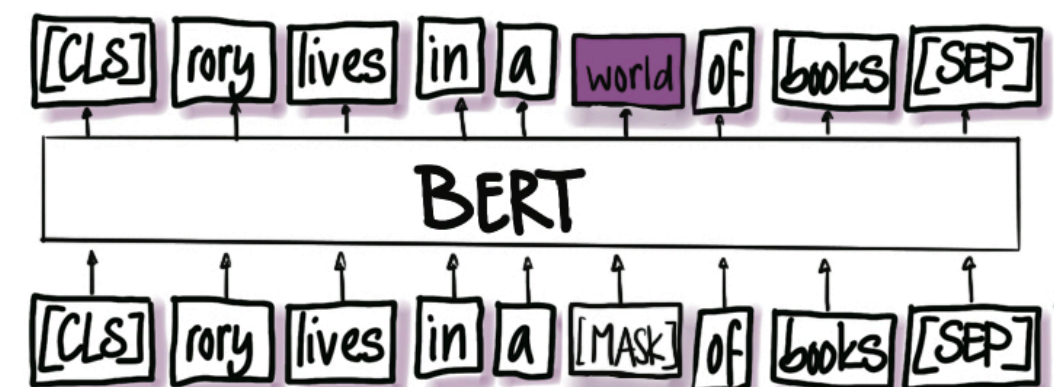


# HOW TO TRAIN YOUR BERT MODEL

Identify order of input (with embeddings)

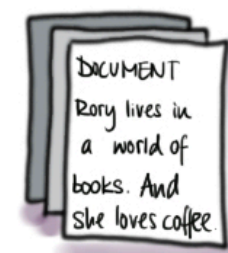


## 1. Masked language modeling (Mask ML)



With masked language modeling, you mask one word at a time of your text to let the model predict which words should be there.

## 2. Next sentence prediction



A: Rory lives in a world of books.  
B: And she loves coffee.



With next sentence prediction, the model predicts which sentence comes next.