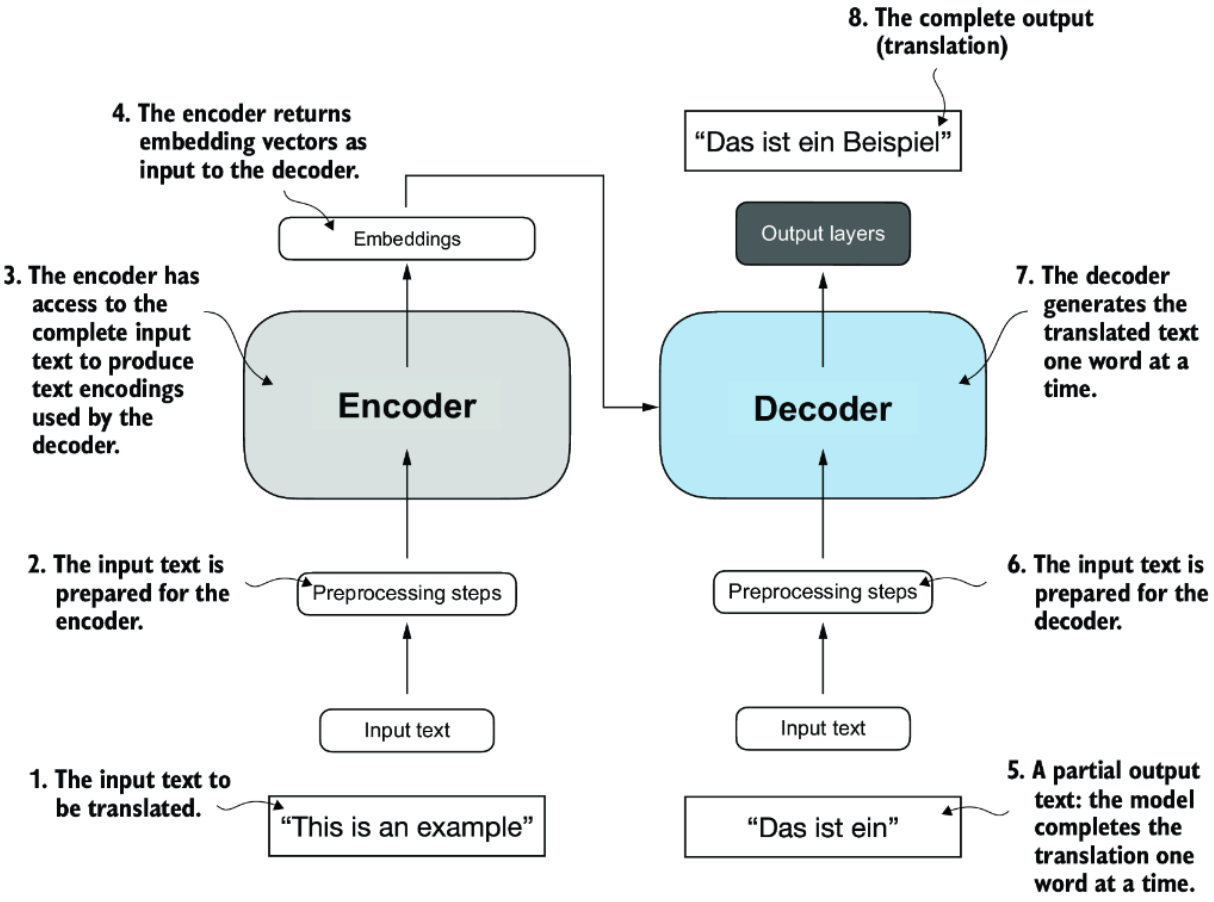
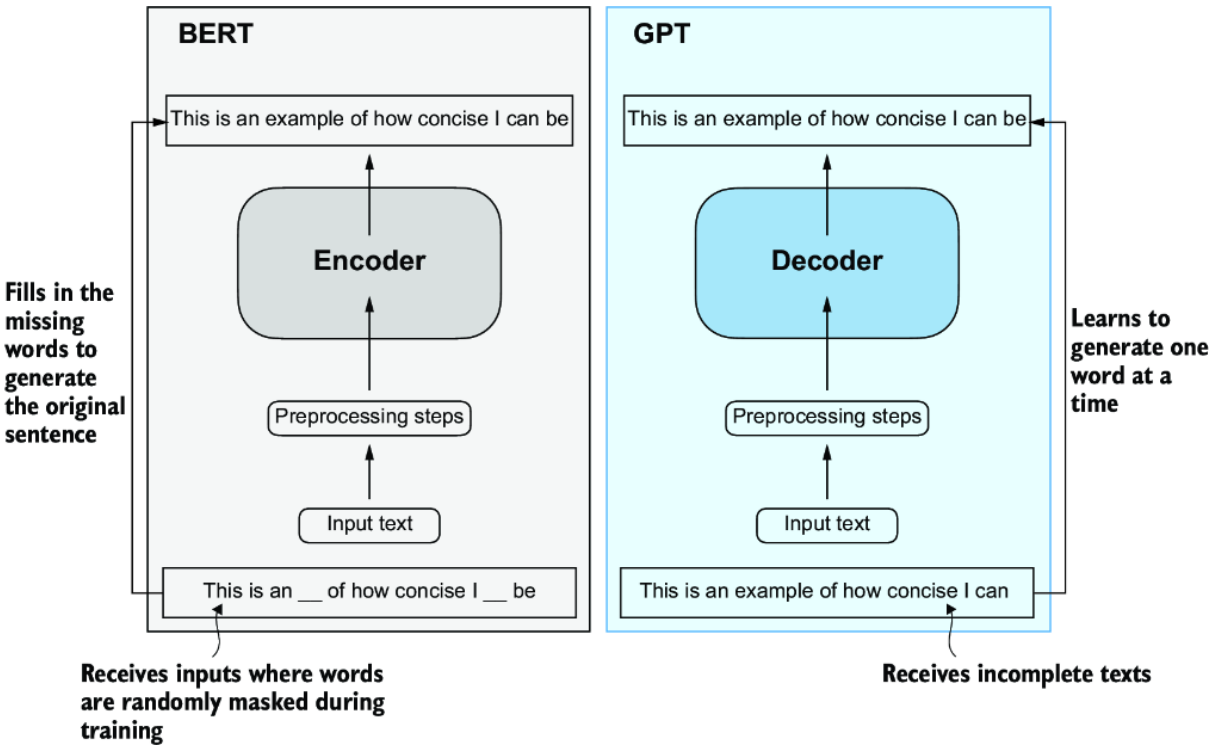
chapter 1

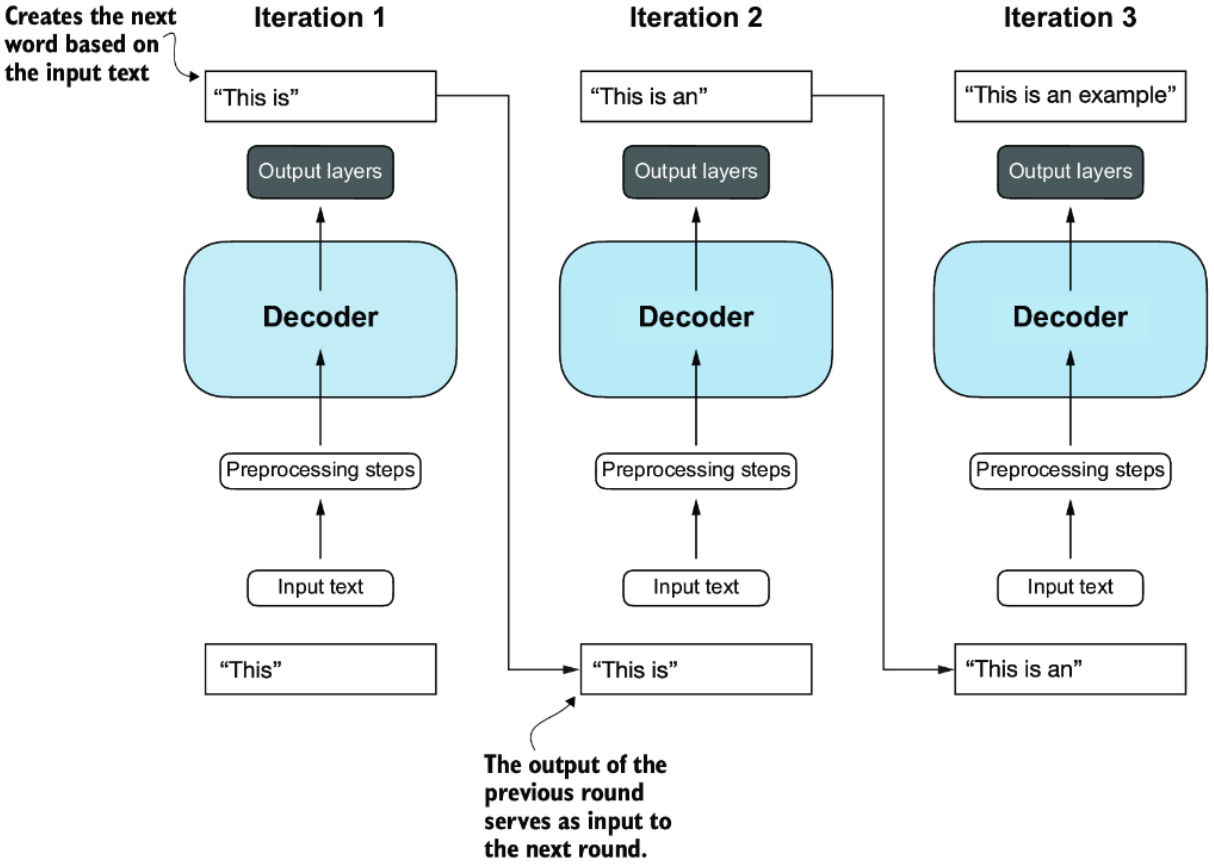


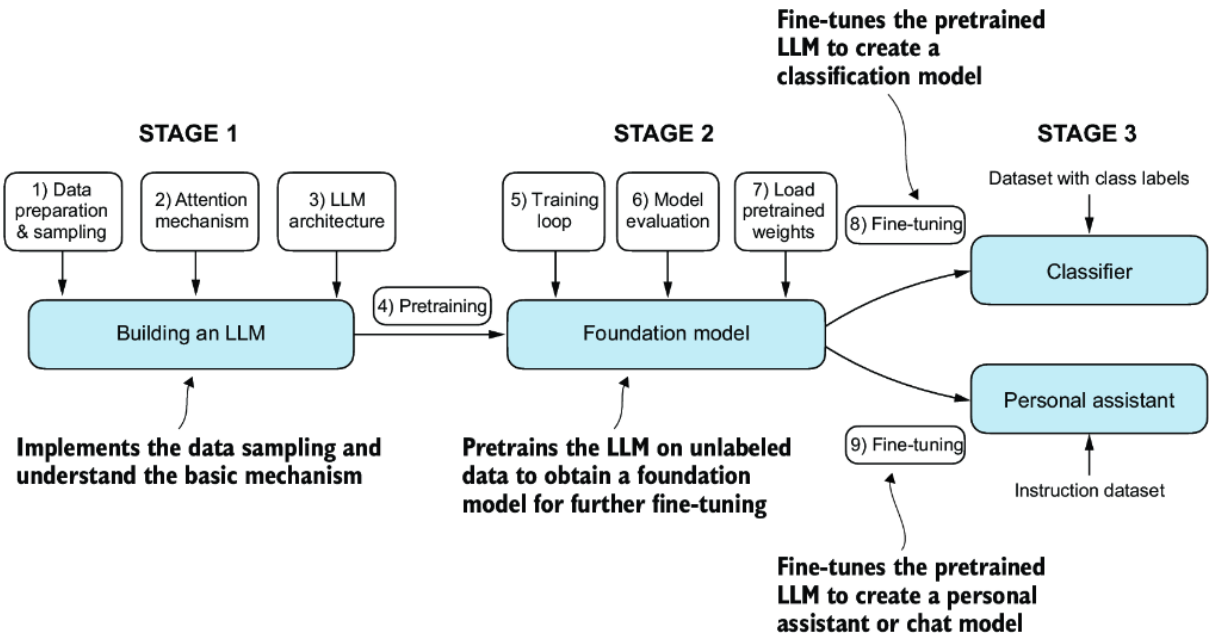
BERT : Encoder only

GPT : Decoder only


BERT
GPT
This is an example of how concise I can be
Encoder
Preprocessing steps
Input text
This is an __ of how concise I __ be
Fills in the missing words to generate the original sentence
Receives inputs where words are randomly masked during training
This is an example of how concise I can be
Decoder
Preprocessing steps
Input text
This is an example of how concise I can
Learns to generate one word at a time
Receives incomplete texts

GPT working:


Creates the next word based on the input text
Iteration 1
Iteration 2
Iteration 3
"This is"
Output layers
Decoder
Preprocessing steps
Input text
"This"
"This is an"
Output layers
Decoder
Preprocessing steps
Input text
"This is"
"This is an example"
Output layers
Decoder
Preprocessing steps
Input text
"This is an"
The output of the previous round serves as input to the next round.


Fine-tunes the pretrained LLM to create a classification model
STAGE 1
STAGE 2
STAGE 3
1) Data preparation & sampling
2) Attention mechanism
3) LLM architecture
5) Training loop
6) Model evaluation
7) Load pretrained weights
8) Fine-tuning
Dataset with class labels
Classifier
Building an LLM
4) Pretraining
Foundation model
Personal assistant
9) Fine-tuning
Instruction dataset
Implements the data sampling and understand the basic mechanism
Pretrains the LLM on unlabeled data to obtain a foundation model for further fine-tuning
Fine-tunes the pretrained LLM to create a personal assistant or chat model