Übung 05 Yi Cui 2758172

Task 1.1

Feature-based approach use takes-specific architecture that include the press-trained representations as additional features.

Fine-tuning approach introduces minimal task-specific parameters, and is trained on the downstream tasks by simply fine-tuning all pre-trained parameters.
(from mandatory paper)

Task 1.2

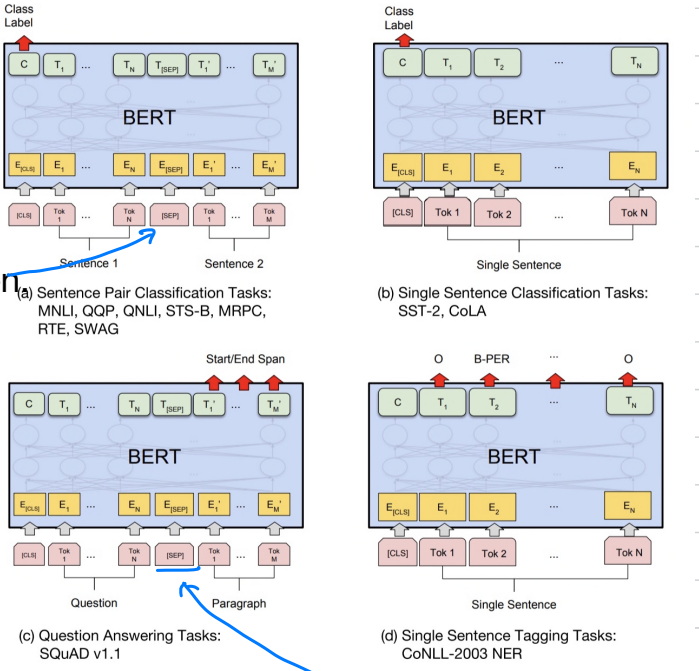
BERT uses bidirectional self-attention, which can attend to context

Task 1.3

Masked LM: simply mask some percentage of the input tokens at random and predict those masked tokens (replace or unchanged), which based on context.

Next Sentence Prediction: base on understanding the relationship between two sentences, where the second sentences could be labeled as IsNext or NotNext, instead of direct capturing by language modeling.

Task 1.4:



(a) Sentence Pair Classification Tasks: MNLI, QQP, QNLI, STS-B, MRPC, RTE, SWAG

(b) Single Sentence Classification Tasks: SST-2, CoLA

(c) Question Answering Tasks: SQuAD v1.1

(d) Single Sentence Tagging Tasks: CoNLL-2003 NER

Sentence Pair Classification:
Input: two sentences with separated token
Output: first Transformer has class label

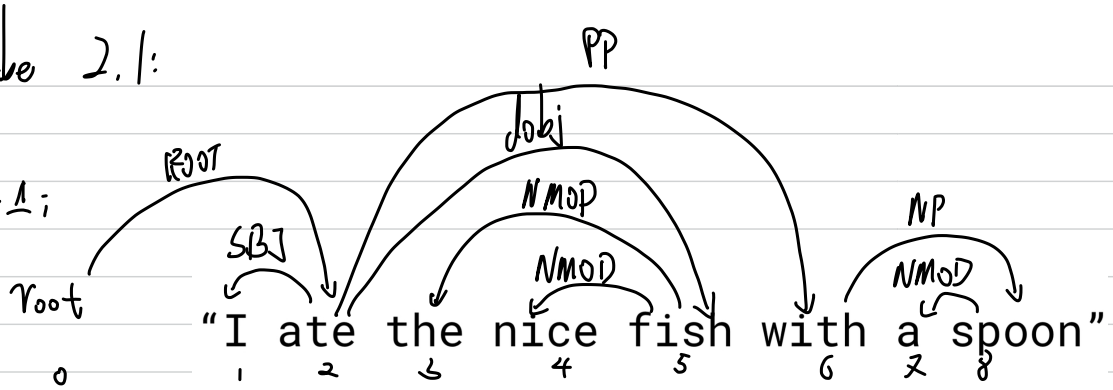
Single Sentence Classification:
Input: whole single sentence
Output: first Transformer has class label

Question Answering:
Input: Question and relevant paragraph as whole sentence with specific token.
Output: the start and end vector-Transformers

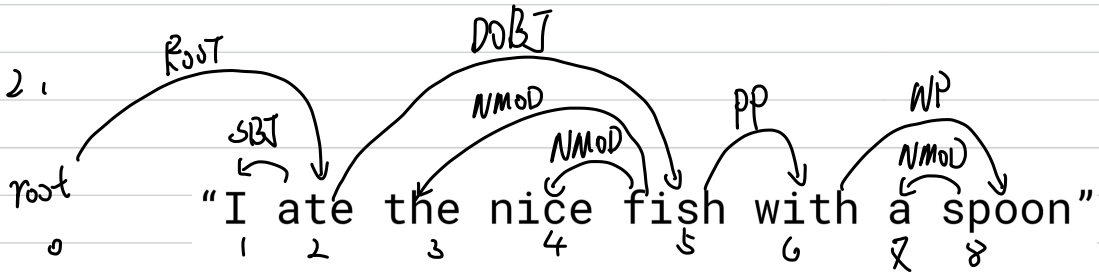
Single Sentence Tagging Tasks:
Input: whole sentence
Output: each Transformer with Tagging

Aufgabe 2.1:

Case 1:



Case 2:



Aufgabe 2.2

Case 1

Index	Word	Head
1	I	2
2	ate	0
3	the	5
4	nice	5
5	fish	2
6	with	2
7	a	8
8	spoon	6

Case 2:

Index	Word	Head
1	I	2
2	ate	0
3	the	5
4	nice	5
5	fish	2
6	with	5
7	a	8
8	spoon	6