

Studies on Computational Linguistics II

Transformer-based Pre-Trained Models and Prompt Tuning

Spring 2023

108.536A

Course Information

- Transformer-based pre-trained language models and application fields focused on Prompt Tuning
- Instructor: Sangah Lee (Dept. of Linguistics, Seoul National University)
(sanalee@snu.ac.kr)
- Lecture: Tue, Thu 17:00-18:15 (bldg. 2, room 320)
- Textbooks: slides and supplementary materials provided

Course Information

- Grade Policies: Relative Grading (A-F)

Item	Attendance	Assignment	Presentation	Final	Total
Rate (%)	10	30	30	30	100

- If you have a valid reason for absence, please submit:
the relevant documents and the Attendance Acknowledgment Request Form
(uploaded on the eTL page)
- Students should present some of the selected papers on relevant topics
provided in the course: expected twice a semester
- Final: project presentation
a system or a conference-level paper based on the course content

Objectives

- Reading and understanding papers

- 수업 중 논문 발제

주제 선점: https://docs.google.com/spreadsheets/d/1D0_wUdGtKQsha5TZqU58BWh5vqYF5LuEt7L-iFmOa3k/edit?usp=sharing

- 과제: 논문 핵심 내용 요약 →

(매주 1회 제출, 발제한 주 제외)

해당 주의 reading list 논문들 중 선택

0~4주차, 8~9주차 제외

- 논문 제목, 저자, url 등 기본 정보

- 논문에서 하고자 한 것

- 논문에서 사용한 방법론

- 논문에서 사용한 데이터

- 실험 결과

- 눈에 띄는 한계점이나 내 연구와의 비교 등

- Organizing and improving your work (기말평가)

- 수업에서 다룬 내용을 활용하여 내 연구 발전 -> 논문으로 작성

- 수업에서 다룬 내용에 기반한 시스템/프로그램 구현


Syllabus

Week	Date	Topic
0	3/2 (Thu)	Course Introduction
1	3/7, 3/9	Paradigms of NLP
2	3/14, 3/16	Attention
3	3/21, 3/23	Transformers
4	3/28, 3/30	Transformer-based PTMs (Basics)
5	4/4, 4/6	Transformer-based PTMs (Improvements)
6	4/11, 4/13	Transformer-based PTMs (Applications)
7	4/18, 4/20	Transformer-based PTMs (Applications)

Syllabus

Week	Date	Topic
8	4/25, 4/27	Fine-Tuning and Prompt Tuning
9	5/2, 5/4	Prompt Tuning (Basics)
10	5/9, 5/11	Prompt Tuning (Basics)
11	5/16, 5/18	Prompt Tuning (Analysis)
12	5/23, 5/25	Prompt Tuning (Improvements)
13	5/30, 6/1	Prompt Tuning (Applications)
14	6/8(Thu), 6/13(Tue)	Prompt Tuning (Applications)
15	6/15 (Thu)	Final Project Presentations

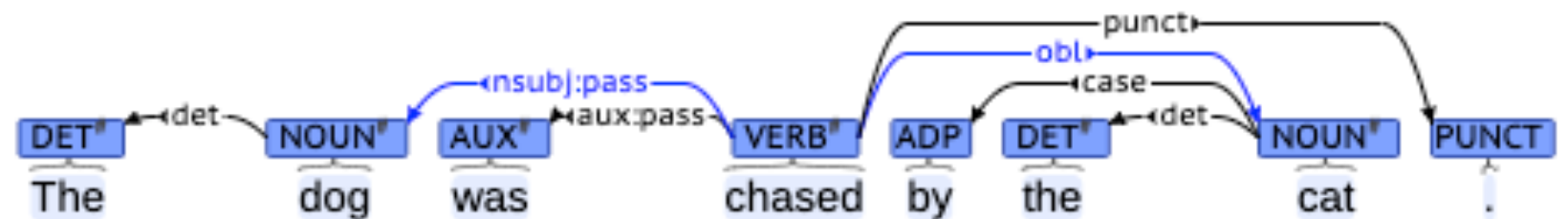
Computational Linguistics

- A subfield of linguistics and computer science
 - concerned with the interaction of human language and computers
 - Includes:
 - the analysis of written texts and spoken discourse
 - the translation of text and speech from one language into another
 - the use of human languages for communication between computers and people
 - the modeling and testing of linguistic theories
 - Statistical analysis of written texts and spoken discourse
 - analysis on corpus: relative frequencies or collocation of letters, sounds, morphemes, words, ...
- 



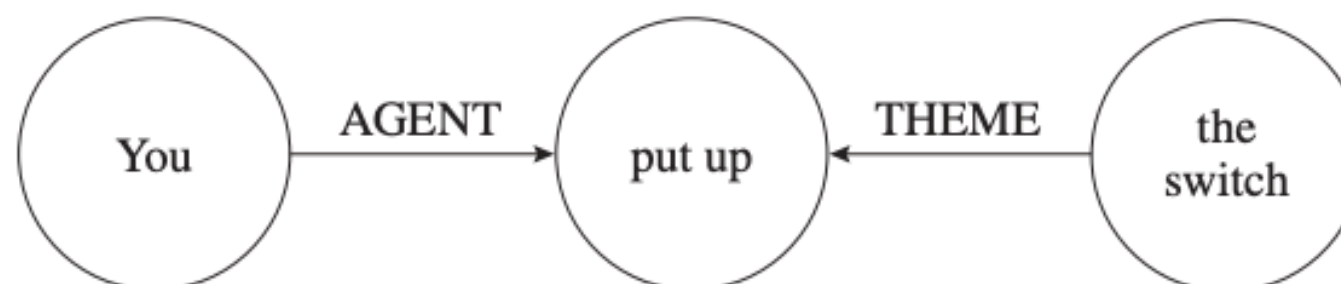
Computational Linguistics

- The interaction between language and computers in all dimensions
 - Computational Phonetics and Phonology
 - Speech Recognition, Speech Synthesis (Text-to-Speech)
 - Computational Morphology: processing of word structures
 - Computational Syntax



<https://universaldependencies.org/>

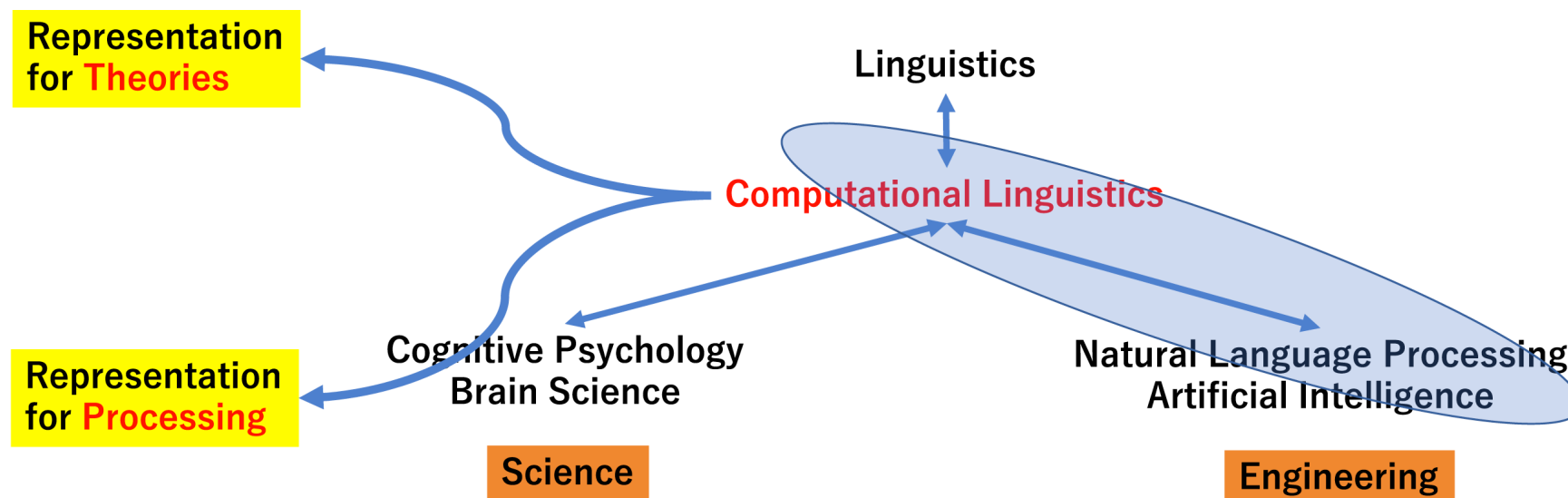
- Computational Semantics: speech understanding and generation



- Computational Pragmatics: sentence disambiguation, coreference resolution, ...

Natural Language Processing

- A schematic view of research disciplines: CL and NLP



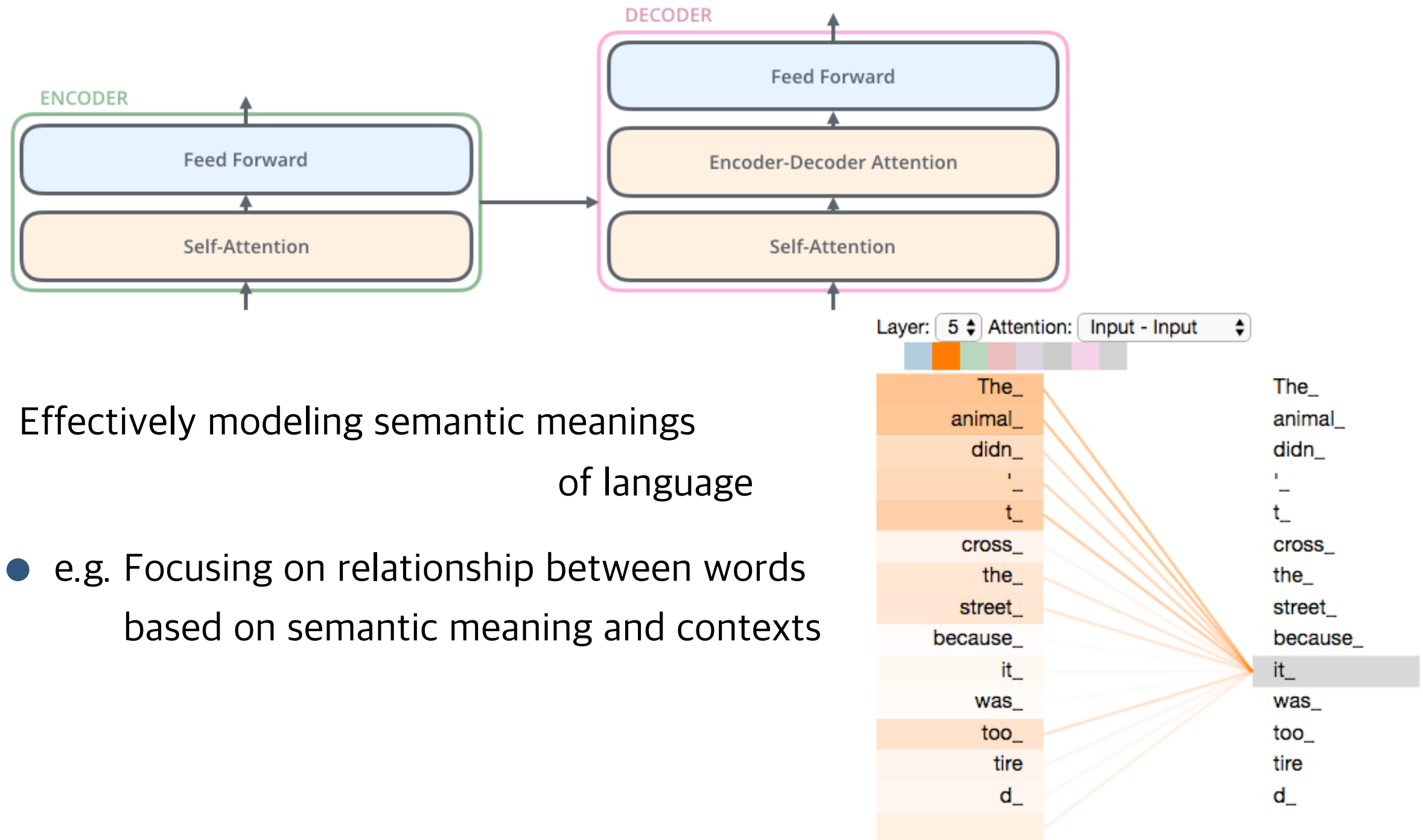
how language is processed
in our minds or our brains

how computer systems should be designed
to process language efficiently and effectively

- NLP may be included either in CL and other fields as their subfield.
(Any other fields can deal with and utilize language data.)
- And this kind of research can be performed through...

Transformers

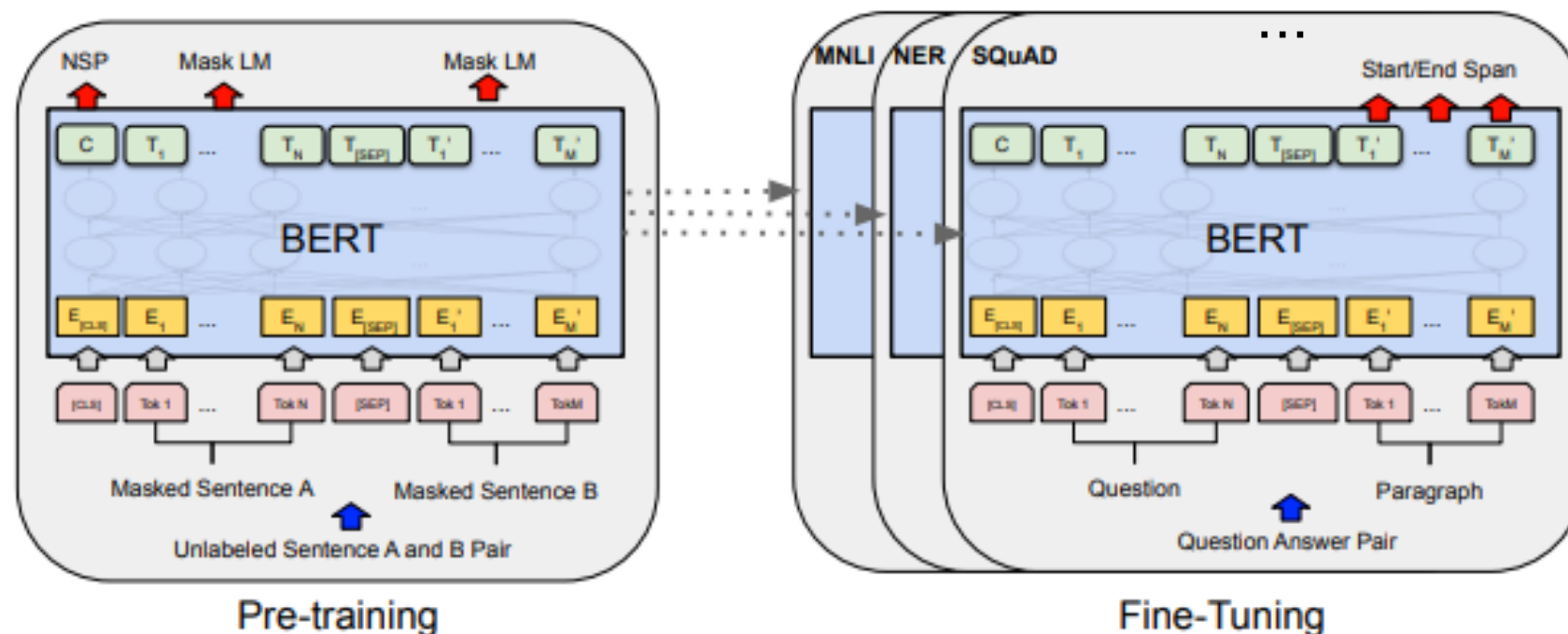
- Using very deep neural models and attention



- Effectively modeling semantic meanings of language
- e.g. Focusing on relationship between words based on semantic meaning and contexts

Pre-Trained Models

- Such deep and complex calculations are performed and **frozen** for future use
==> Pre-Trained (Language) Models: PTM, PLM (사전학습모델)
 - Pre-trained models: BERT, GPT, ...
 - Usually provided through HuggingFace 🤗
- tasks:
Natural Language Inference
Named Entity Recognition
Question Answering

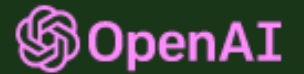


import and load a given BERT model

let the model perform specific task

Pre-Trained Models


ChatGPT: Optimizing Language Models for Dialogue



<https://openai.com/blog/chatgpt/>

- ChatGPT: one of the PTMs especially working for dialogues

 Hi ChatGPT, I am Sana. The model's decoders generate sentences (sequence of words) based on **pre-trained knowledge** and user inputs


 Hello Sana! How can I assist you today?



long contexts saved

 Say something for our students!

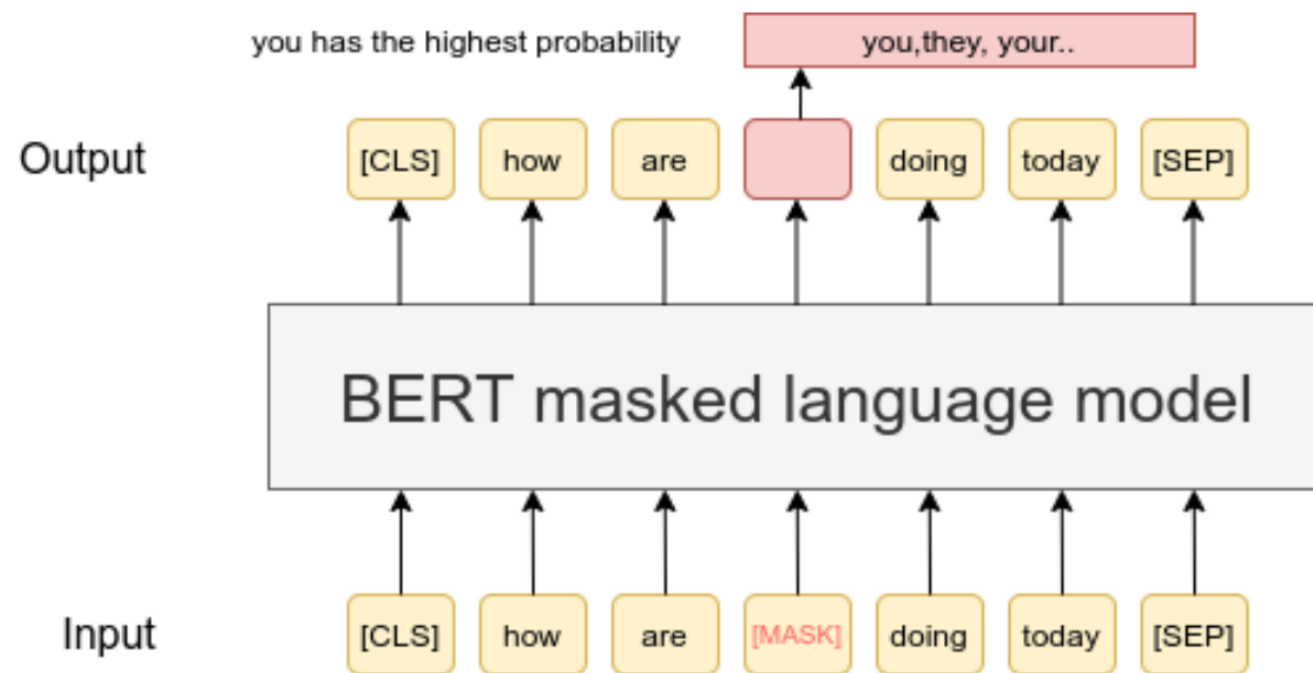
human evaluation about the generated answers

 To all the students out there: never stop learning and pursuing your passions! Education opens doors to endless possibilities and empowers you to make a positive impact in the world. Stay curious, ask questions, and never be afraid to make mistakes. Remember, they are opportunities to grow and learn. Good luck on your academic journey!



Pre-Trained Models

- Pre-Training: language modeling on large-scaled datasets (e.g. BERT)



MLM (Masked Language Modeling):
modeling contextual information
in a sentence

Sentence 1

Sentence 2

Next Sentence?

I am going outside.

I will be back after 6.

YES

I am going outside.

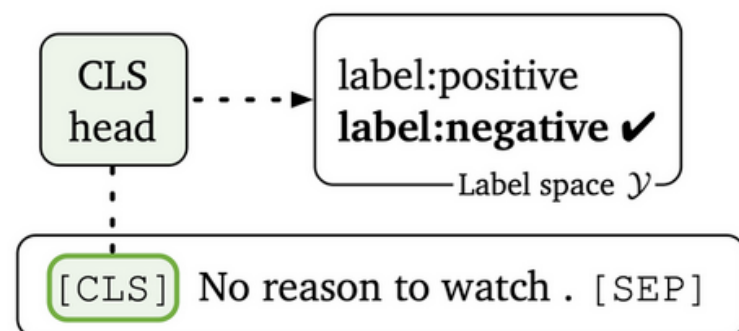
You know nothing John snow.

NO

NSP (Next Sentence Prediction):
modeling contextual information
between two sentences

Pre-Trained Models

- Transfer Learning
 - Capture important knowledge from multiple source tasks and then apply the knowledge to a target task
 - Source tasks and target tasks may have completely different data domains and task settings, yet the knowledge required to handle these tasks is consistent.
 - Using less datasets, cost and parameters
 - e.g. Fine-Tuning, Prompt Tuning, ...
- e.g. BERT fine-tuning: add a classifier layer on top of PTM

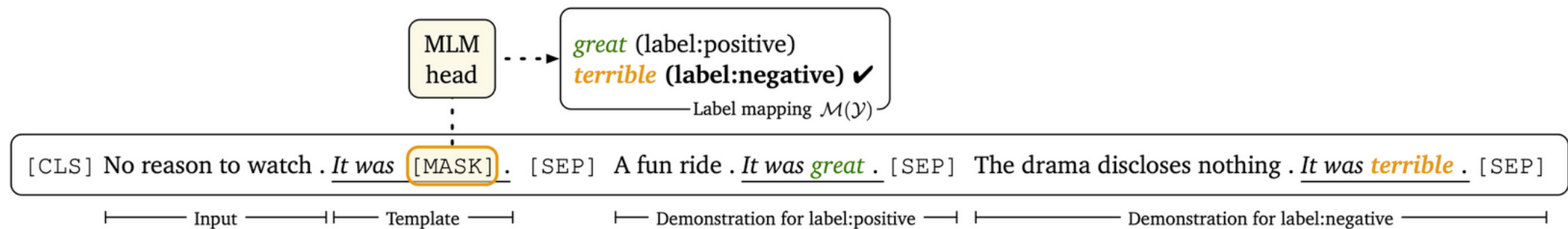


Han et al. (2021). "Pre-Trained Models: Past, Present and Future". arXiv:2106.07319.

Gao et al. (2020). "Making Pre-Trained Language Models Better Few-Shot Learners". Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). pp. 3816-3830.

Better Methods for Pre-Trained Models

- Prompt-based Learning
 - Reducing the need for large supervised datasets
 - Bridging the gap between pre-training and task-specific fine-tuning



(similar to Masked Language Modeling of pre-training phase)

- Discrete prompts, continuous (soft) prompts, ...
- Generalization for various kinds of NLP tasks