

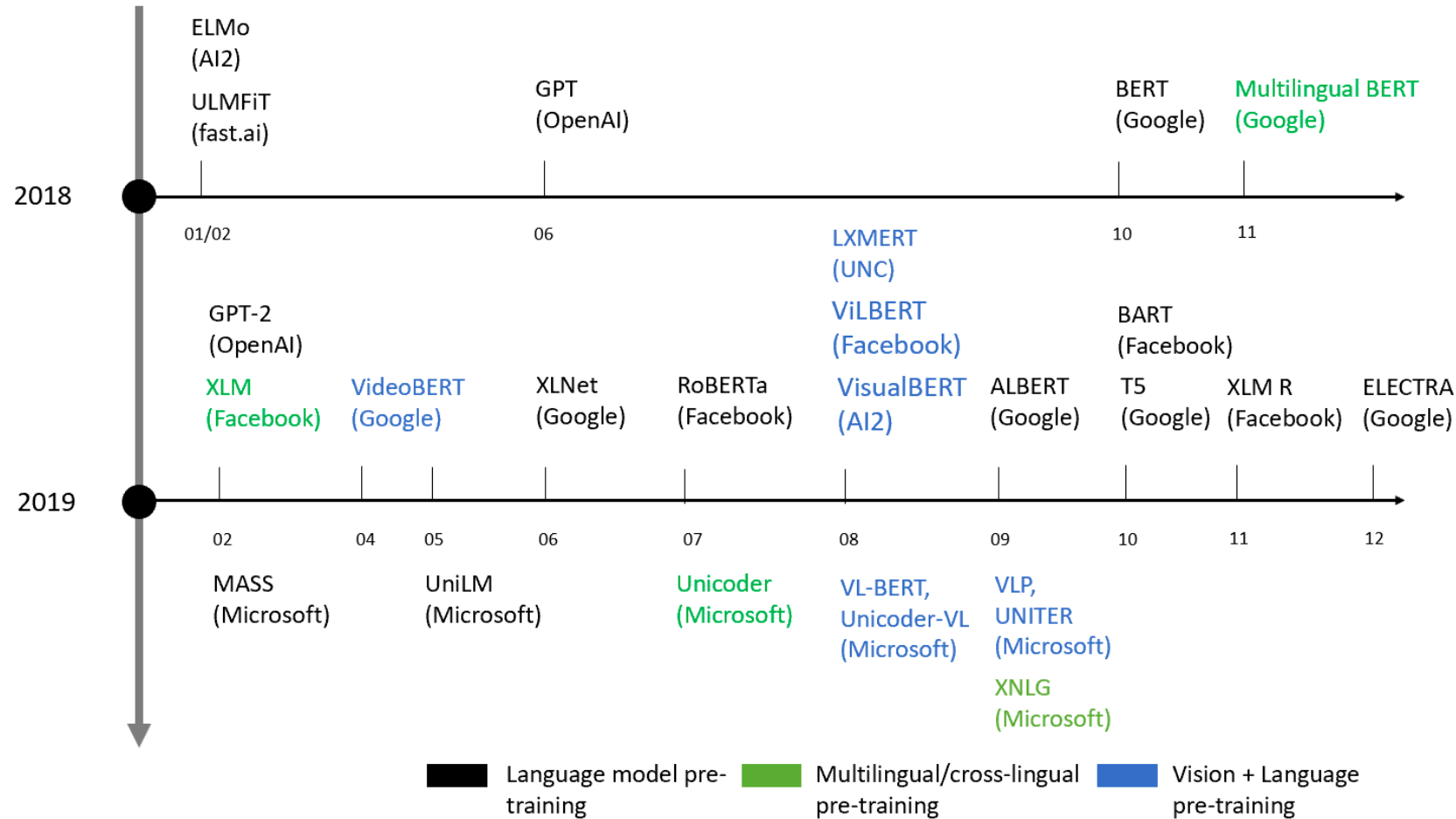
Appendix: Text Classification with BERT

Ki Hyun Kim

nlp.with.deep.learning@gmail.com

Trends in these days

- Pretraining and finetuning (Transfer learning) with Big-LM.



What is BERT?

- Take Encoder from Transformer

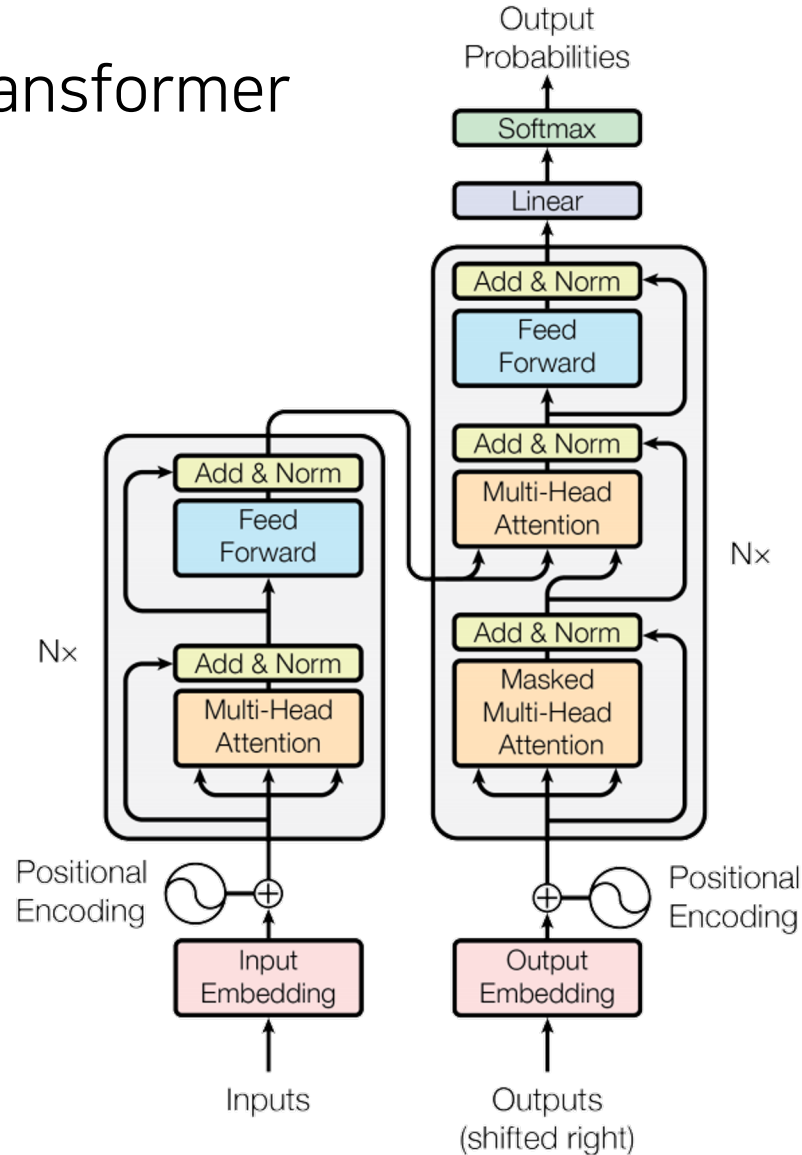
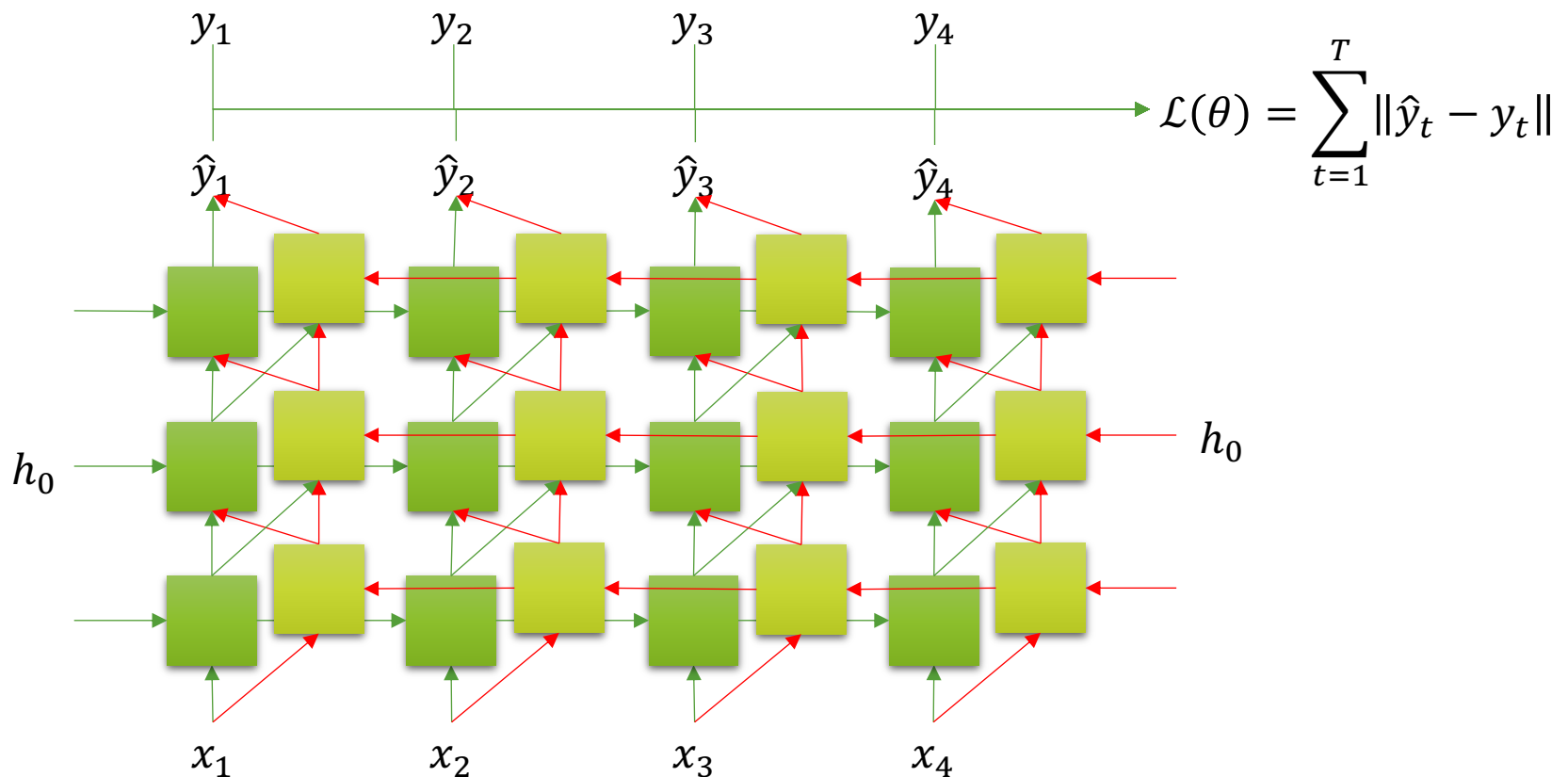


Figure 1: The Transformer - model architecture.

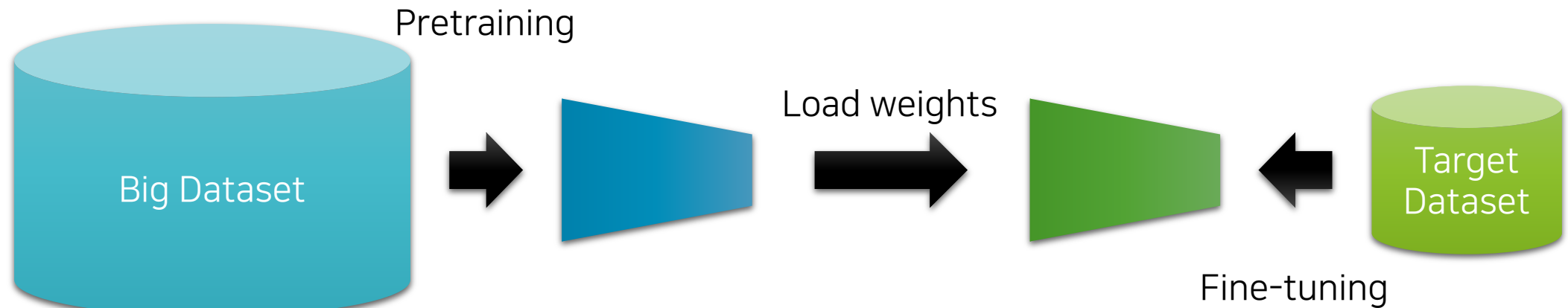
What is BERT?

- Masked Language Models (MLM)
 - Denoising Autoencoders
- Next Sentence Prediction (NSP)



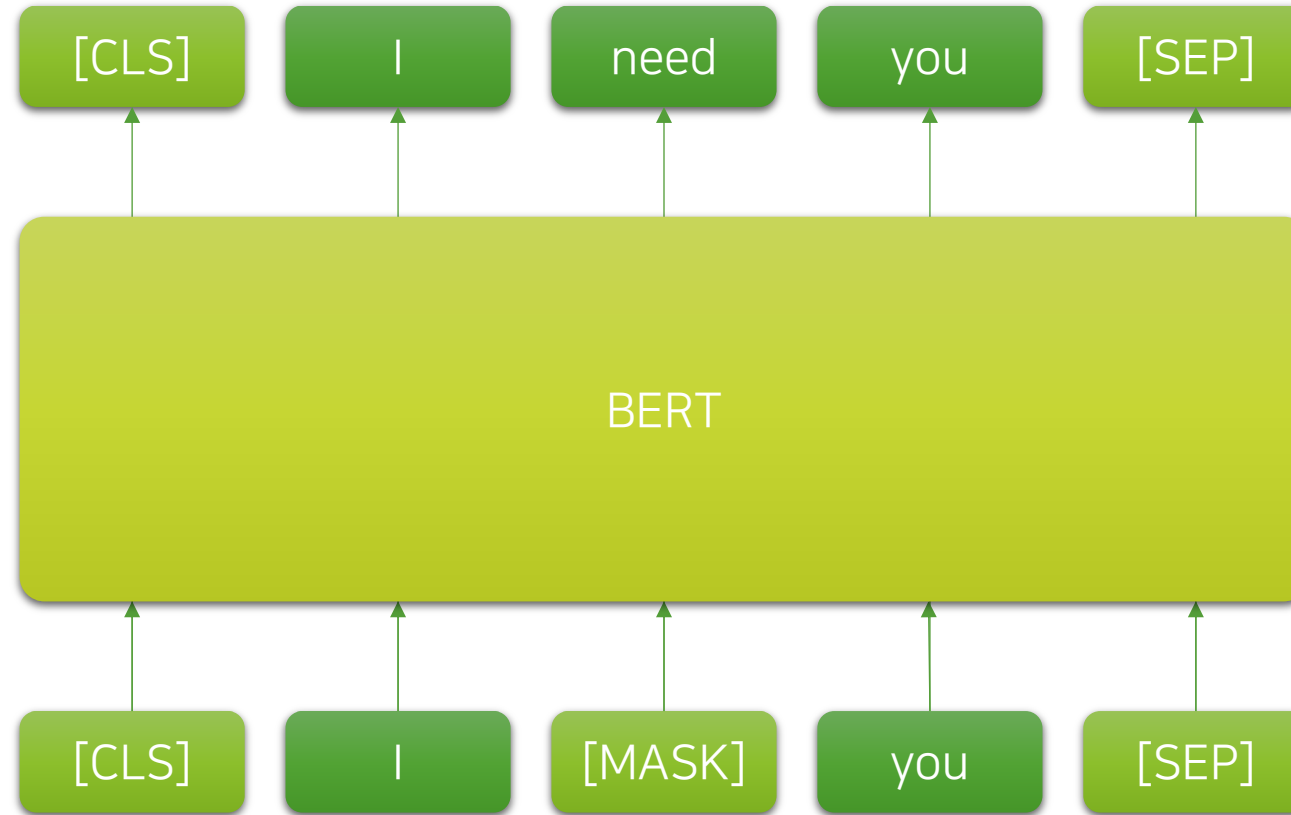
Transfer Learning via BERT

- BERT: Bi-directional Sequential Modeling (using Transformer)
- Pretraining using MLM and NSP, and fine-tuning to downstream tasks.
 - e.g. Text Classification, Machine Reading Comprehension



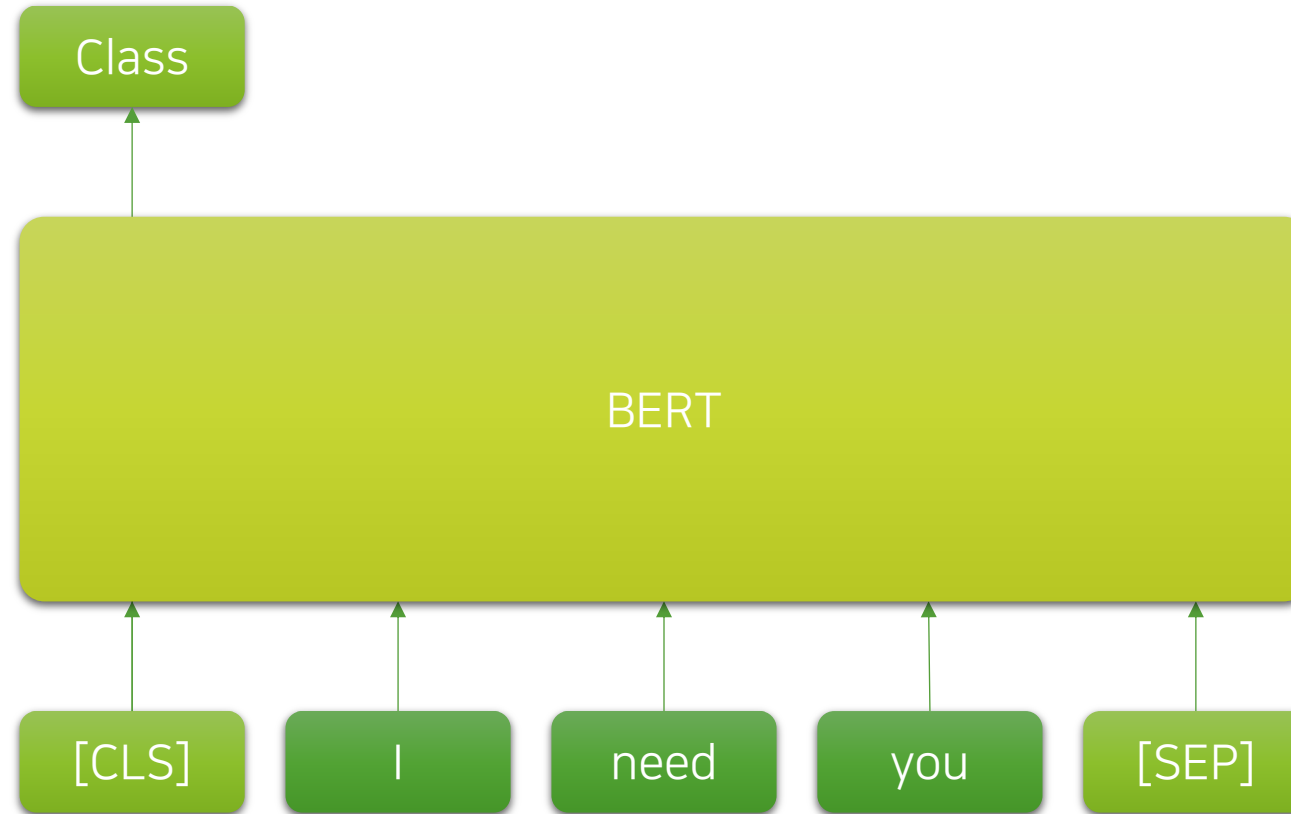
Text Classification via Transfer Learning

- Pretraining (MLM)



Text Classification via Transfer Learning

- Fine-tuning



Where to Use

- SKTBrain KoBERT
 - <https://github.com/SKTBrain/KoBERT#using-with-pytorch>
- Hugging Face
 - <https://huggingface.co/>