Unified Pre-training: close-style tasks for language model (LM) pre-training
- Left-to-right unidirectional LM
- Right-to-left unidirectional LM
- Bidirectional LM
- Sequence-to-sequence LM

Unified Fine-tuning: UniLM (the same model) can be fine-tuned as a unidirectional decoder, a bidirectional encoder, or a sequence-to-sequence model to support various downstream natural language understanding and generation tasks.

Unidirectional LM: the context of the masked word to be predicted consists of the words on both the right and the left (same in BERT)

Sequence-to-sequence LM: the context consists of all the words in the source sequence and the words on its left and right (same in BERT)

Generative QA: generate answers as a sequence model

Experiments
- Abstractive Summarization
- Document -> summary: sequence-to-sequence fine-tuning
- Extractive QA: classify answer spans

Unified Language Model Pre-training for Natural Language Understanding and Generation