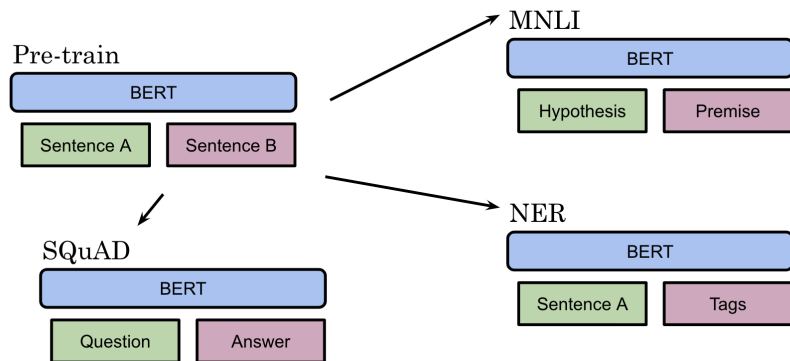


## Question Answering

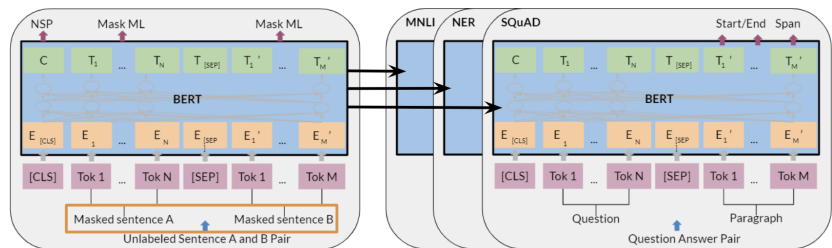
- ✓ **Video:** Week 3 Overview  
6 min
- ✓ **Reading:** Week 3 Overview  
10 min
- ✓ **Video:** Transfer Learning in NLP  
7 min
- ✓ **Reading:** Transfer Learning in NLP  
10 min
- ✓ **Video:** ELMo, GPT, BERT, T5  
7 min
- ✓ **Reading:** ELMo, GPT, BERT, T5  
10 min
- ✓ **Video:** Bidirectional Encoder Representations from Transformers (BERT)  
4 min
- ✓ **Reading:** Bidirectional Encoder Representations from Transformers (BERT)  
10 min
- ✓ **Video:** BERT Objective  
2 min
- ✓ **Reading:** BERT Objective  
10 min
- ✓ **Video:** Fine tuning BERT  
2 min
- 📖 **Reading:** Fine tuning BERT  
10 min
- 🎥 **Video:** Transformer: T5  
3 min
- 📖 **Reading:** Transformer T5  
10 min
- 🎥 **Video:** Multi-Task Training Strategy  
5 min
- 📖 **Reading:** Multi-Task Training Strategy  
10 min
- 🎥 **Video:** GLUE Benchmark  
2 min
- 📖 **Reading:** GLUE Benchmark  
10 min
- 🎥 **Video:** Question Answering  
2 min
- 📖 **Reading:** Question Answering  
10 min
- 📄 **Lab:** SentencePiece and BPE  
2h
- 📖 **Reading:** Content Resource  
10 min
- Assignment**
- 📄 **Programming Assignment:** Question Answering  
3h

# Fine tuning BERT

Once you have a pre-trained model, you can fine tune it on different tasks.



For example, given a hypothesis, you can identify the premise. Given a question, you can find the answer. You can also use it for named entity recognition. Here is a summary of the inputs.



- You can replace sentences A/B
- Paraphrase from sentence A
- Question/passage
- Hypothesis premise pairs in entailment
- Text and a  $\emptyset$  for classification/sequence tagging
- Output tokens are fed into a layer for token level tasks otherwise use [CLS] embedding as input.

Mark as completed

