

BERT Fine-tuning for Different Tasks

 Pre-trained BERT (All Parameters)

+



Sequence Classification

Uses:

[CLS]

token representation

*Sentiment Analysis,
Text Classification*



Token Classification

Uses:

T₁ T₂ T₃

each token's output

NER, POS Tagging



Question Answering

Predicts:

Start & End spans

*SQuAD, Reading
Comprehension*



End-to-End Fine-tuning

All parameters updated with task-specific data

Fast Training

Task-Specific

Transfer Learning



Only **few epochs** needed



Add layer on **top of BERT**



Few-shot & zero-shot