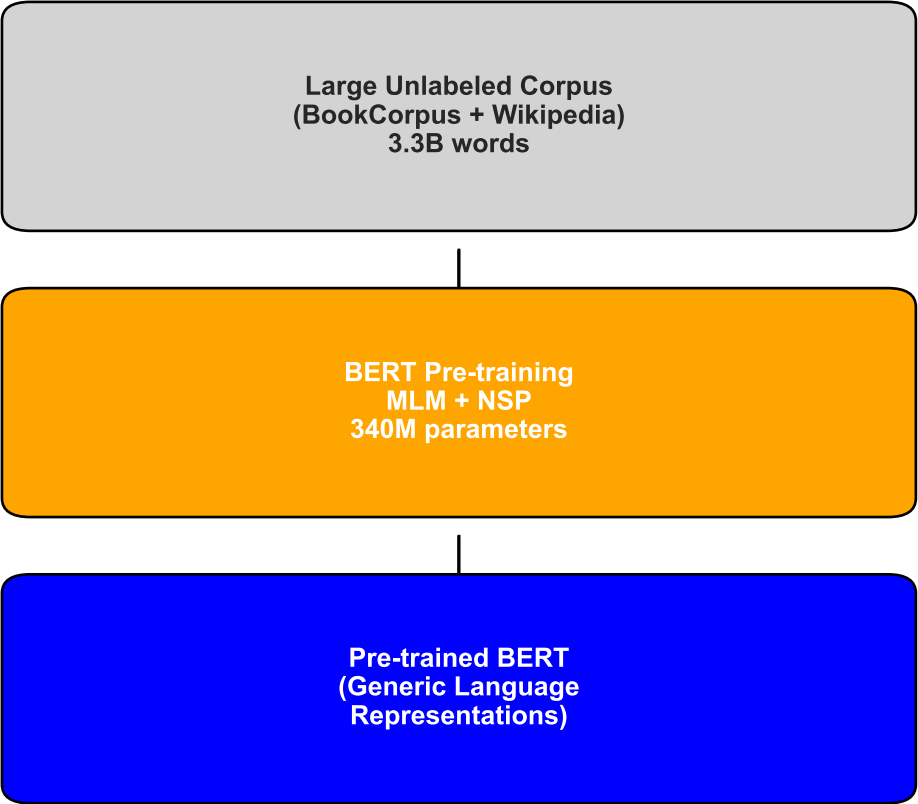
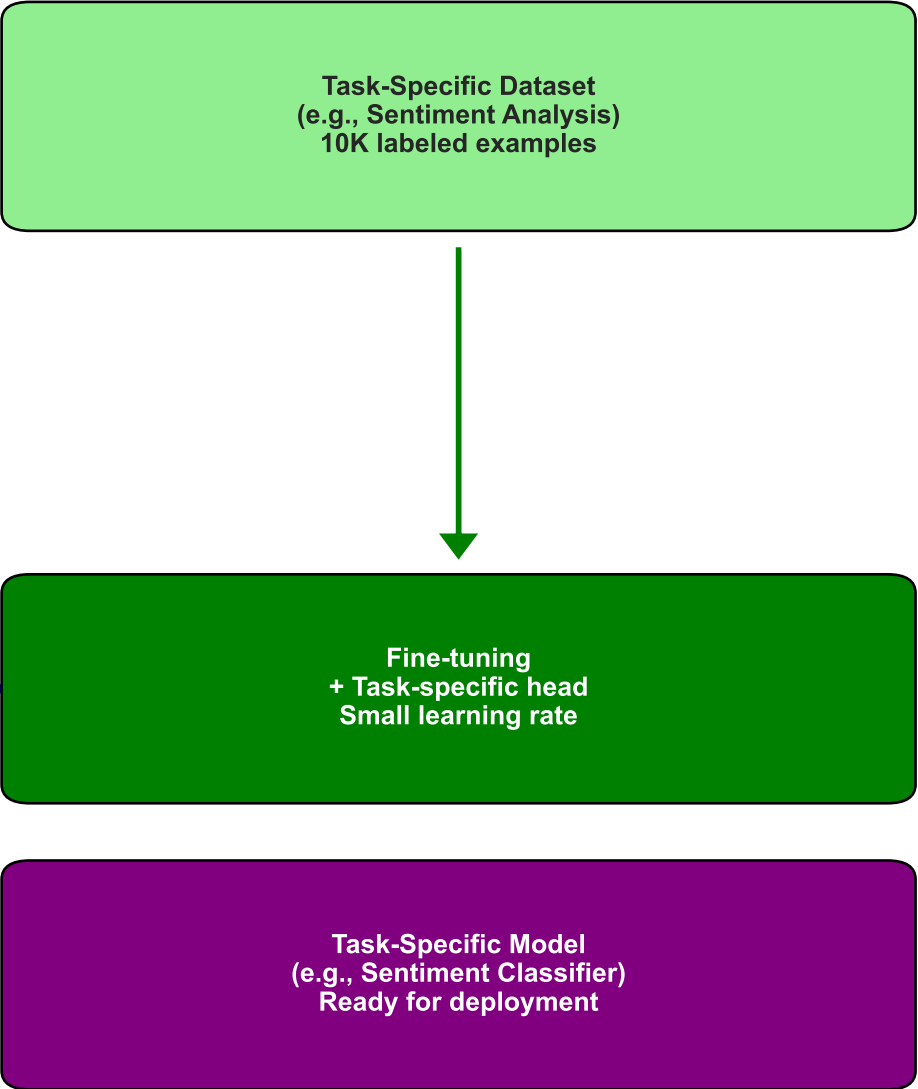


# BERT Fine-tuning: Transfer Learning in Action

## Phase 1: Pre-training



## Phase 2: Fine-tuning



Transfer

## Comparison: From Scratch vs Fine-tuning

### From Scratch

Training Data: 100K+ examples  
Training Time: 2-3 days  
GPU Requirements: 8+ GPUs  
Accuracy: 85%

### BERT Fine-tuning

1K examples  
1-2 hours  
1 GPU  
95%