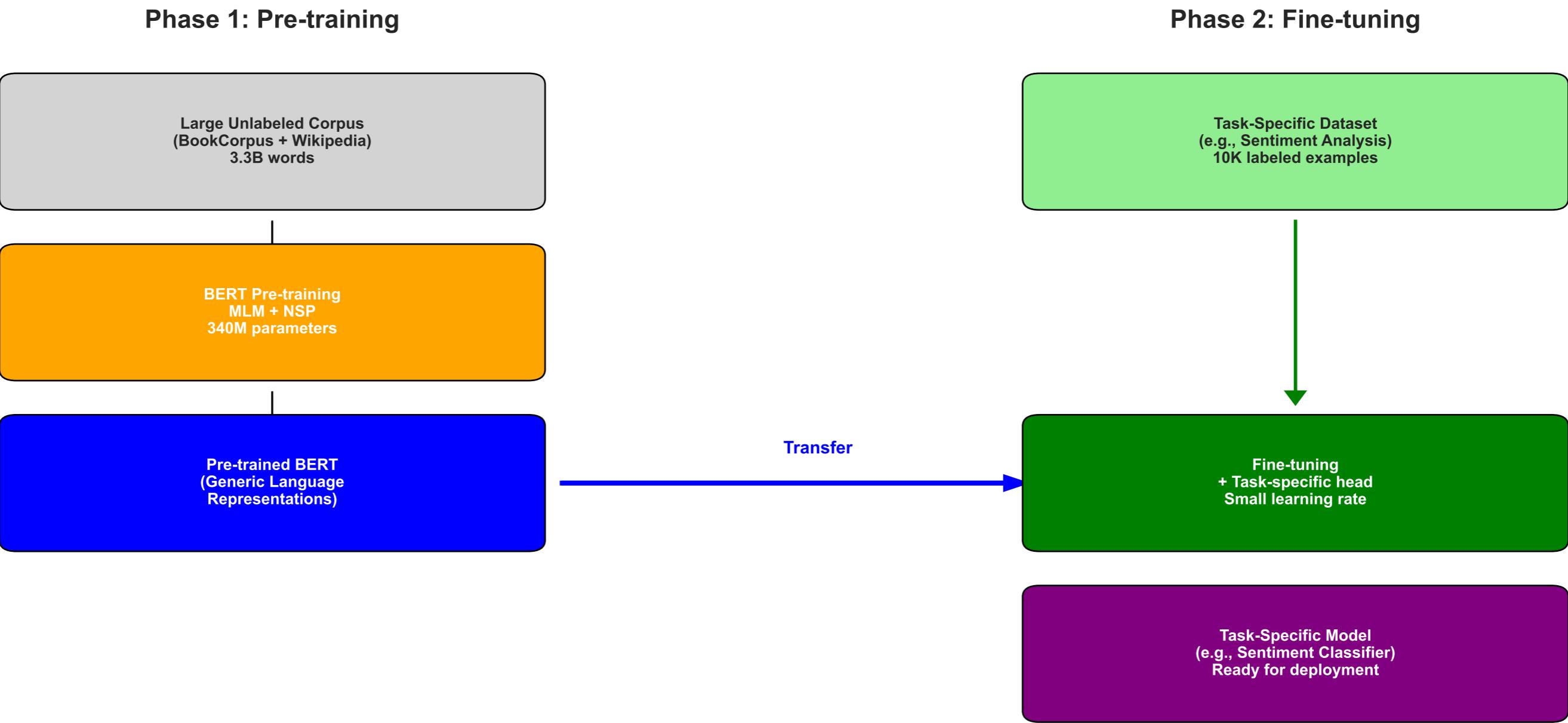


BERT Fine-tuning: Transfer Learning in Action



Comparison: From Scratch vs Fine-tuning

From Scratch

Training Data:
Training Time:
GPU Requirements:
Accuracy:

100K+ examples
2-3 days
8+ GPUs
85%

BERT Fine-tuning

1K examples
1-2 hours
1 GPU
95%