BERT vs Traditional NLP Approaches

Traditional: Sequential Processing

The → movie → was → not → bad

Processes left-to-right only

Result: "not bad" = negative

BERT: Bidirectional Processing

The movie was not bad

Sees full context simultaneously

Result: "not bad" = positive

Context Understanding Comparison

Legend: Traditional, BERT

Categories: Negation, Sarcasm, Idioms, Metaphors, Ambiguity

Y-axis: Understanding (%)
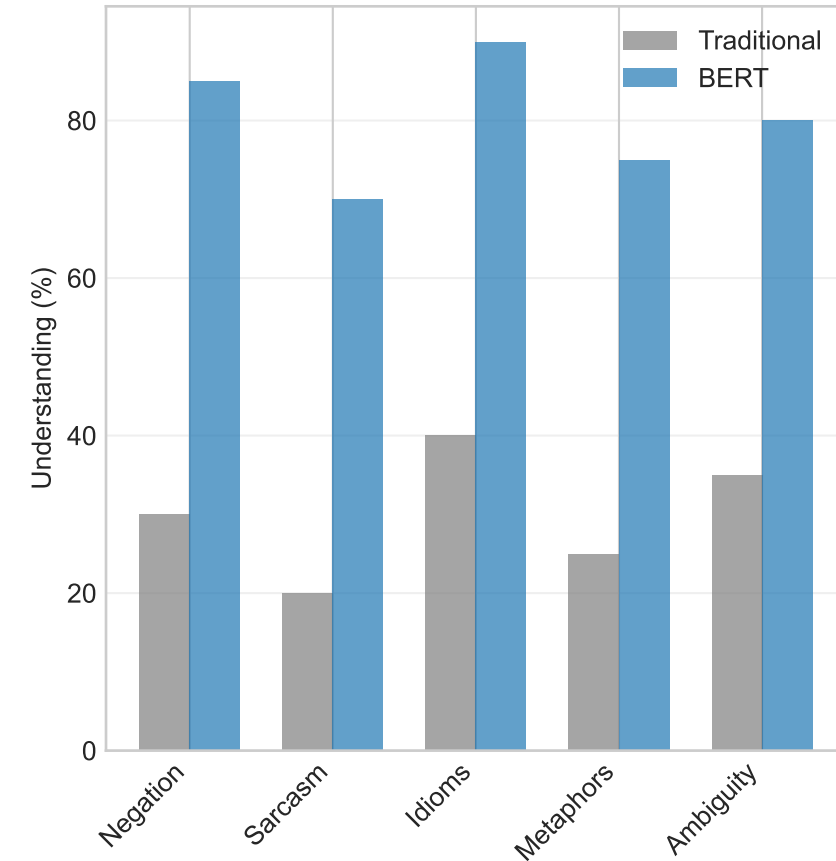
Processing Speed Comparison

X-axis: Number of Documents
Y-axis: Processing Time (seconds)

Legend: Traditional, BERT

Performance by NLP Task

Tasks: Summary, QA, NER, Sentiment

X-axis: Accuracy (%)
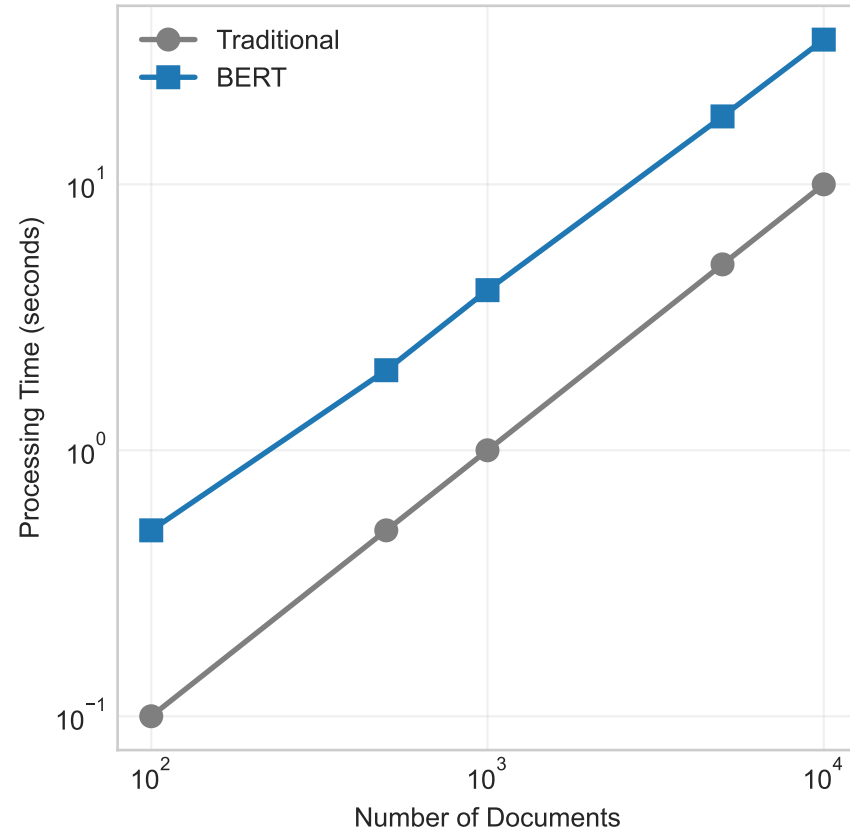
Legend: Traditional, BERT
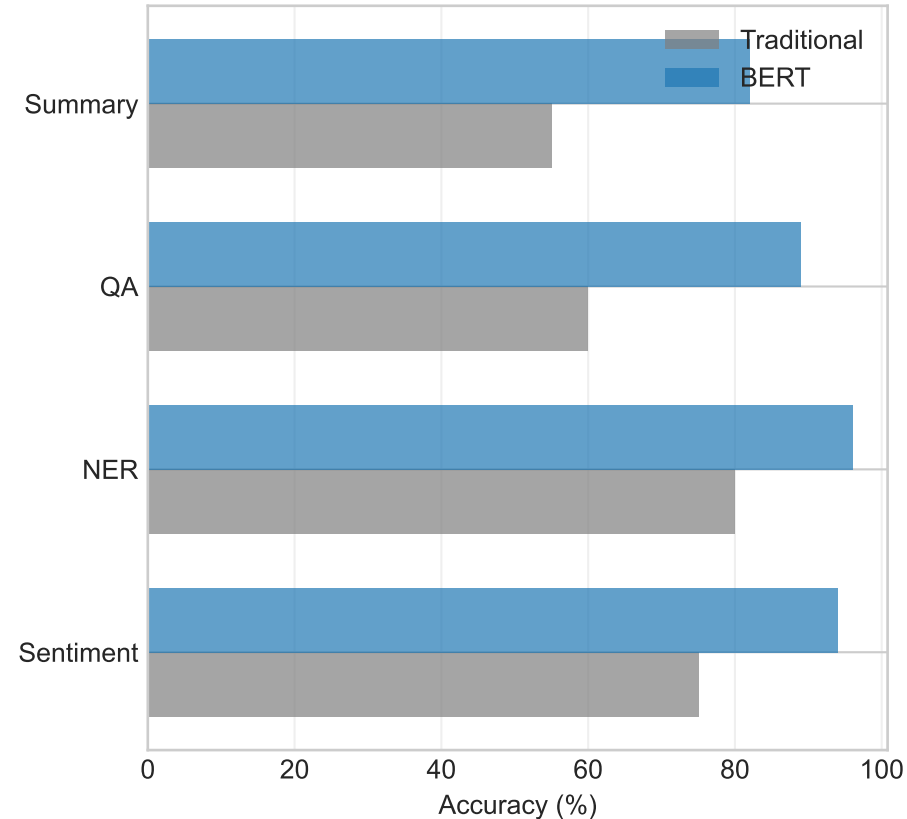
Resource Requirements

Metric	Traditional	BERT
Model Size	10 MB	440 MB
Memory	< 1 GB	4-8 GB
Training Data	Small OK	Large required
Fine-tuning	Not needed	Recommended
Languages	One	Multiple