



Advantages

- ✓ **Full context** - captures both past and future information
- ✓ **Better representations** - richer hidden states at each time step
- ✓ **Improved accuracy** - superior performance on many NLP tasks
- ✓ **Disambiguation** - resolves ambiguity using surrounding context



Disadvantages

- ✗ **Cannot stream** - requires full sequence before processing
- ✗ **Higher computation** - 2x parameters and memory compared to unidirectional
- ✗ **Not for generation** - unsuitable for autoregressive tasks
- ✗ **Slower training** - increased training time and inference latency



Key Applications



Named Entity Recognition



POS Tagging



Sentiment Analysis

Capture sentiment from entire text

Identify entities using full sentence context

Accurate part-of-speech labeling with context



Slot Filling

Extract structured info from utterances



Dependency Parsing

Understand syntactic relationships



Speech Recognition

Phoneme prediction with context