How Context Length Affects Prediction Quality

- RNN
- Transformer

Transformer maintains quality

RNN degrades with distance

Next Word Prediction Accuracy (%)

Context Length (words)