

Context-Based Quotation Recommendation

Ansel MacLaughlin¹, Tao Chen², Burcu Karagol Ayan², Dan Roth³

¹ Northeastern University, ² Google Research, ³ University of Pennsylvania

Introduction

Research Questions

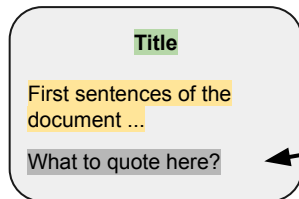
- How to recommend quotable content to an author while they are writing a document?

Our Contributions

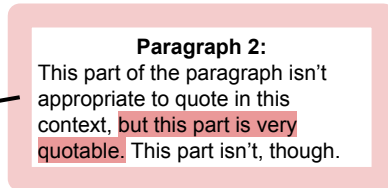
- We propose a novel approach to quote recommendation as both paragraph recommendation and span prediction.
- We are the first to apply BERT to the task of quote recommendation.

Quote Recommendation as Paragraph Ranking & Span Prediction

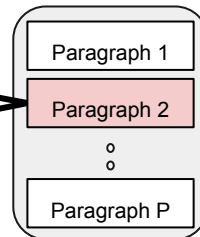
Author Writing a New Document



Relevant Source Paragraph



Source Document



Methods: Two Fine-tuned BERT Models - Paragraph & Span

Model Input (same for both BERT models):

- [CLS] New Doc Title [body_start] New Doc Left Context [SEP] Source Paragraph [SEP]
- Feed in the gold (new document, source paragraph) pair with N negative pairs (source paragraphs that weren't quoted)
- BERT processes each pair independently
- Fine-tune models with listwise (paragraph) or shared-normalization (span) loss

Results

Paragraph Ranking	TFIDF	BM25	KNRM	Conv-KNRM	Combined BERT
Mean Average Precision	39.0	39.9	39.2	38.9	53.2

Span Prediction	Full Paragraph	Paragraph (1st sentence)	Paragraph (last sentence)	Combined BERT
Exact Match	2.9	4.8	5.5	22.8