

Leitura do capítulo 3 de

Pretrained Transformers for Text Ranking: BERT and Beyond

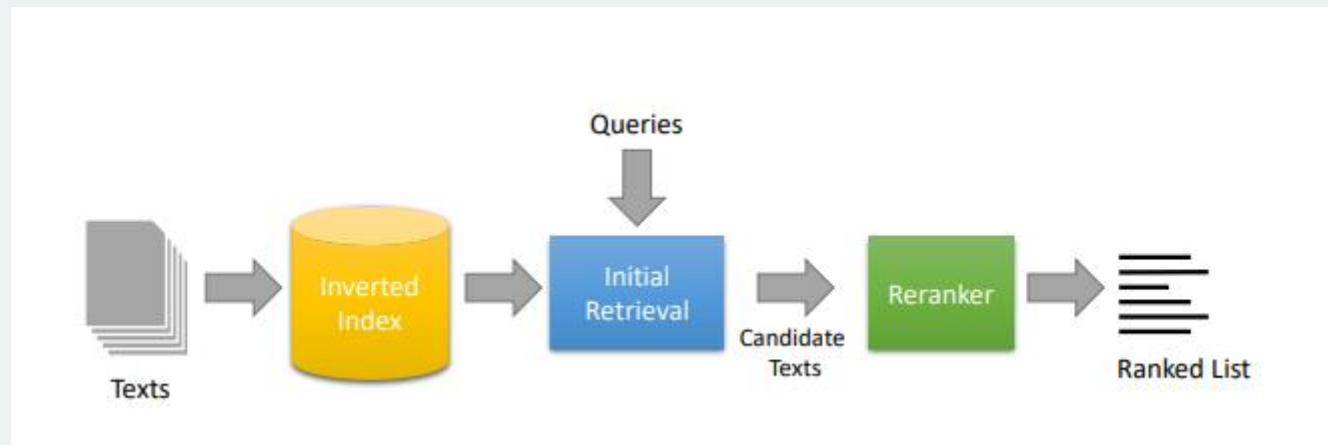
(Jimmy Lin, Rodrigo Nogueira, Andrew Yates)

Contribuições

1. Visão geral do que é o BERT, como é usado para diferentes tarefas de NLP e para reranking:
 1. Single-input classification (ex.: análise de sentimentos)
 2. Two-input classification (ex.: detecção se duas sentenças são parágrafes, **reranking**)
 3. Single-input token labeling (ex.: reconhecimento de entidades nomeadas)
 4. Two-input token labeling (ex.: question answering – que parte do texto responde a pergunta?)

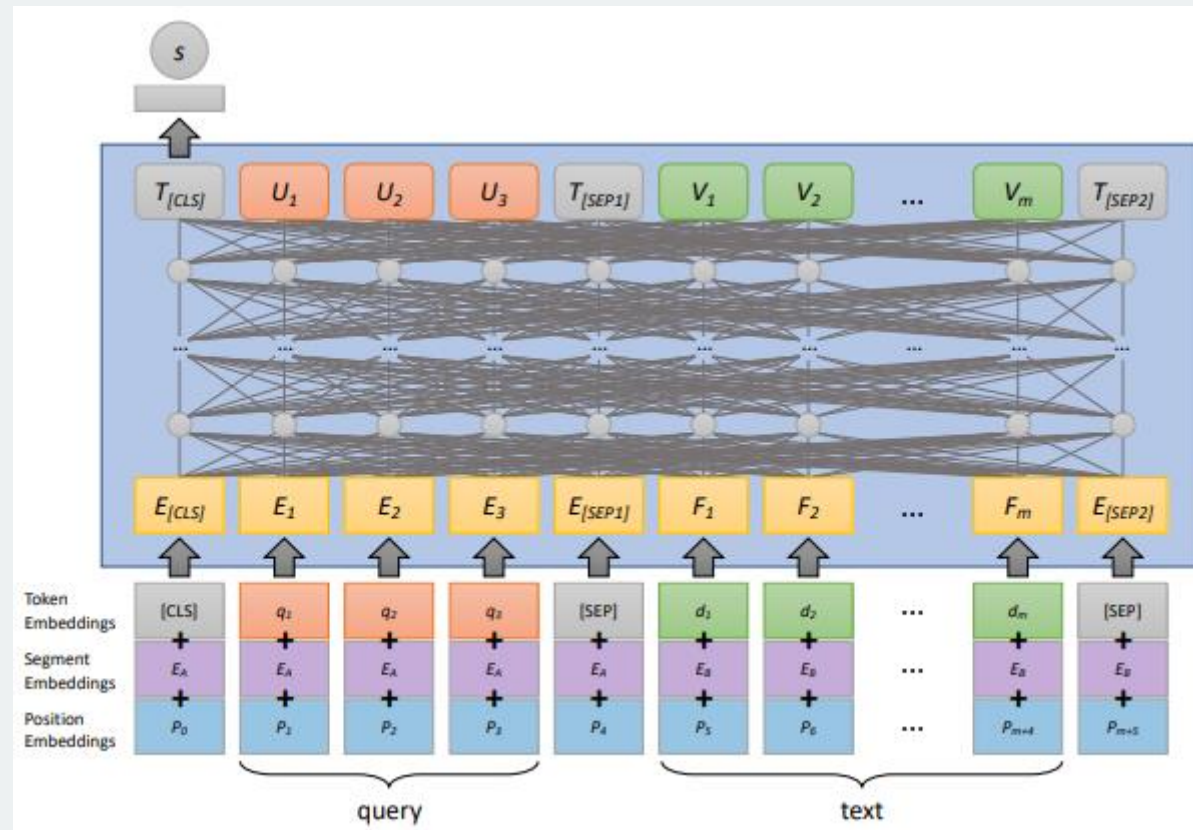
Explicação de conceitos importantes

1. O reranking é executado após uma pesquisa inicial:



Explicação de conceitos importantes

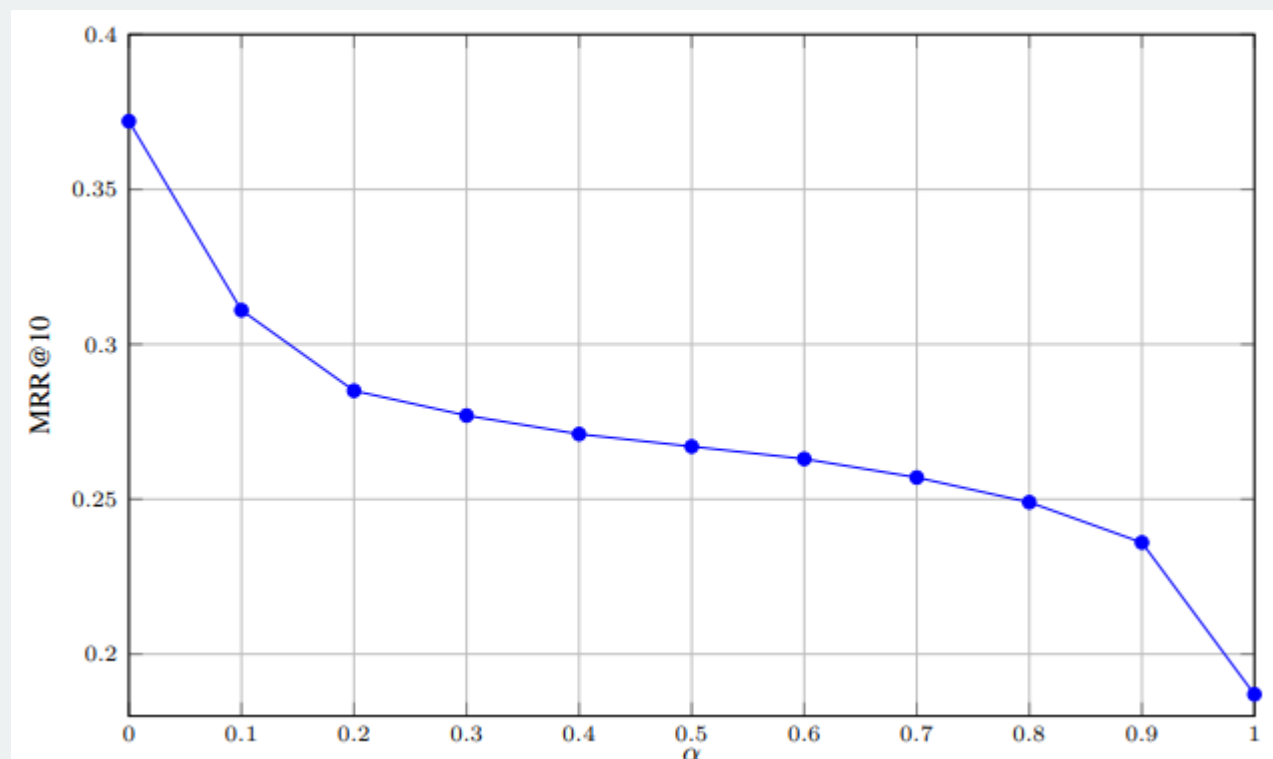
2. É usado um modelo de linguagem pré-treinado (checkpoint), que em seguida é treinado com textos classificados (ex.: relevante vs não-relevante) da tarefa que se quer fazer.



Resultado inesperado

Nos resultados apresentados, tentou-se combinar o ranking do BM25 com o do BERT e o resultado foi que usar apenas o BERT produzia um resultado melhor que qualquer combinação linear possível.

Achei inesperado pois em problemas de engenharia normalmente é possível melhorar o resultado incluindo sensores, mesmo que o novo sensor seja pior que os anteriores.



Dúvida que restou depois da leitura

Nos resultados foi mostrado um intervalo de confiança calculado com 5 novas simulações.

Não ficou claro se essas novas simulações foram com novos conjuntos de dados (entradas totalmente diferentes, representando experimentos diferentes) ou se foram com o mesmo conjunto de dados e apenas alteração na seed (nesse caso o que estaria sendo testado é mais o efeito da apresentação dos dados à rede).

Obrigado

Leandro Carísio
carisio@gmail.com