Dense Passage Retrieval for Open-Domain Question Answering

Monique Monteiro
moniquelouise@gmail.com

Conceitos importantes

- Open-domain question answering

  Tarefa que responde a perguntas usando uma grande coleção de documentos

  Requer recuperação eficiente de passagens para selecionar contextos candidatos

  Espaços vetoriais esparsos (TF-IDF, BM25) são o método "de-facto"

  Framewors de 2 estágios:

  1. *Context retriever: seleiona um subconjunto de passagens algumas das quais contêm a resposta à pergunta*
  2. *Machine reader: examina os contextos recuperados e identifica a reposta correta.*

Contribuição

Mostrar que a recuperação pode ser implementada praticamente usando apenas representações densas

Arquitetura dual-encoder (Dense Passage Retriever – DPR):

- Os embeddings são otimizados para maximizar o produto interno entre o vetor da pergunta e os vetores das passagens relevantes
- Objetivo que compara todos os pares de pergunta-passagem em um batch
- Fine-tuning em etapa única, sem pré-treinamento adicional
- Arquitetura BERT com uso do token [CLS]

Função objetivo

$$L(q_i, p_i^+, p_{i,1}^-, \cdots, p_{i,n}^-) \quad (2)$$

$$= -\log \frac{e^{\text{sim}(q_i, p_i^+)}}{e^{\text{sim}(q_i, p_i^+)} + \sum_{j=1}^{n} e^{\text{sim}(q_i, p_{i,j}^-)}}.$$

- Exemplos negativos:

  Passagens positivas relacionadas a outras questões que aparecem no conjunto de treinamento (do mesmo mini-batch) + uma passagem BM25 negative.

  Para um batch size B: treinamento em B^2 em exemplos em cada batch.

Datasets

- Wikipedia

  Dúvida: contém apenas os documentos. E as perguntas?

- QA datasets:

  Natural Questions

  Trivia QA

  WebQuestions

  CuratedTREC

  SQuAD v1.1

Resultados inesperados/interessantes

- L2 tem desempenho comparável a produto interno
- L2 e produto interno são superiores a similaridade de cossenos
- Alta capacidade de generalização
- Maior acurácia do *retriever* → melhores resultados finais da tarefa de QA
- Treinamento separado do *retriever* e do *reader* leva a melhor resultado do que treinamento em conjunto