

# SPLADE: Sparse Lexical and Expansion Model for First Stage Ranking

...

27 de abril de 2023

# Conceitos Importantes

- Ranqueamento de primeiro estágio.
- Com as representações esparsas aprendidas, é possível o uso de índice invertido.
- Com essa abordagem é possível a redução do problema de vocabulary mismatch.
- Com o uso do índice invertido o pipeline de inferência se torna bem mais rápido em comparação a busca densa.

# Contribuições do Artigo

- Ajustes de hiperparâmetros do SparTerm traz melhorias que superam amplamente os resultados relatados no artigo original
- Modelo SParse Lexical And Expansion (SPLADE), baseado em uma ativação logarítmica e regularização esparsa.
- Mostraram como a regularização da esparsidade pode ser controlada para influenciar o trade-off entre eficiência (em termos do número de operações de ponto flutuante) e eficácia.

# Resultados

model	MS MARCO dev		TREC DL 2019		FLOPS
	MRR@10	R@1000	NDCG@10	R@1000	
Dense retrieval					
Siamese (ours)	0.312	0.941	0.637	0.711	-
ANCE [25]	0.330	0.959	0.648	-	-
TCT-ColBERT [15]	0.335	0.964	0.670	0.720	-
Sparse retrieval					
BM25	0.184	0.853	0.506	0.745	0.13
DeepCT [4]	0.243	0.913	0.551	0.756	-
doc2query-T5 [18]	0.277	0.947	0.642	0.827	0.81
ST lexical-only [1]	0.275	0.912	-	-	-
ST expansion [1]	0.279	0.925	-	-	-
Our methods					
ST lexical-only	0.290	0.923	0.595	0.774	1.84
ST exp- $\ell_1$	0.314	0.959	0.668	0.800	4.62
ST exp- $\ell_{\text{FLOPS}}$	0.312	0.954	0.671	0.813	2.83
SPLADE- $\ell_1$	0.322	0.954	0.667	0.792	0.88
SPLADE- $\ell_{\text{FLOPS}}$	0.322	0.955	0.665	0.813	0.73

# Obrigado

Manoel Veríssimo

---