

Leitura do capítulo 1 de

*Pretrained Transformers for Text Ranking:
BERT and Beyond*
(Jimmy Lin, Rodrigo Nogueira, Andrew Yates)

Contribuições

1. O que é ranqueamento de texto
2. Aplicações para ranqueamento de texto
3. Histórico: desse ~1940 até BERT

Conceitos importantes

Text ranking. Aplicações:

Recuperação de informação

Serviço de busca

Question Answering (QA)

Infoboxes (identificar texto no documento)

Community Question Answering (CQA)

Quora, StackOverflow...

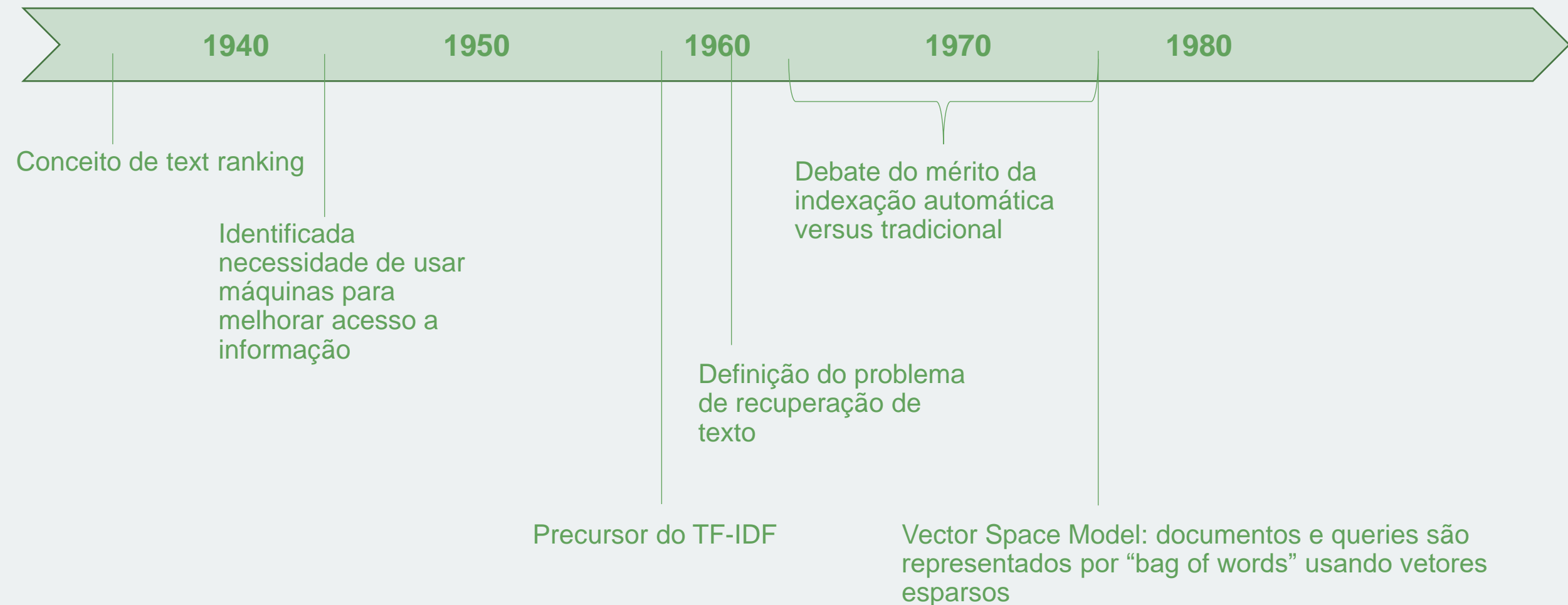
Text recommendation

Jornais, PubMed...

Suporte a outras aplicações

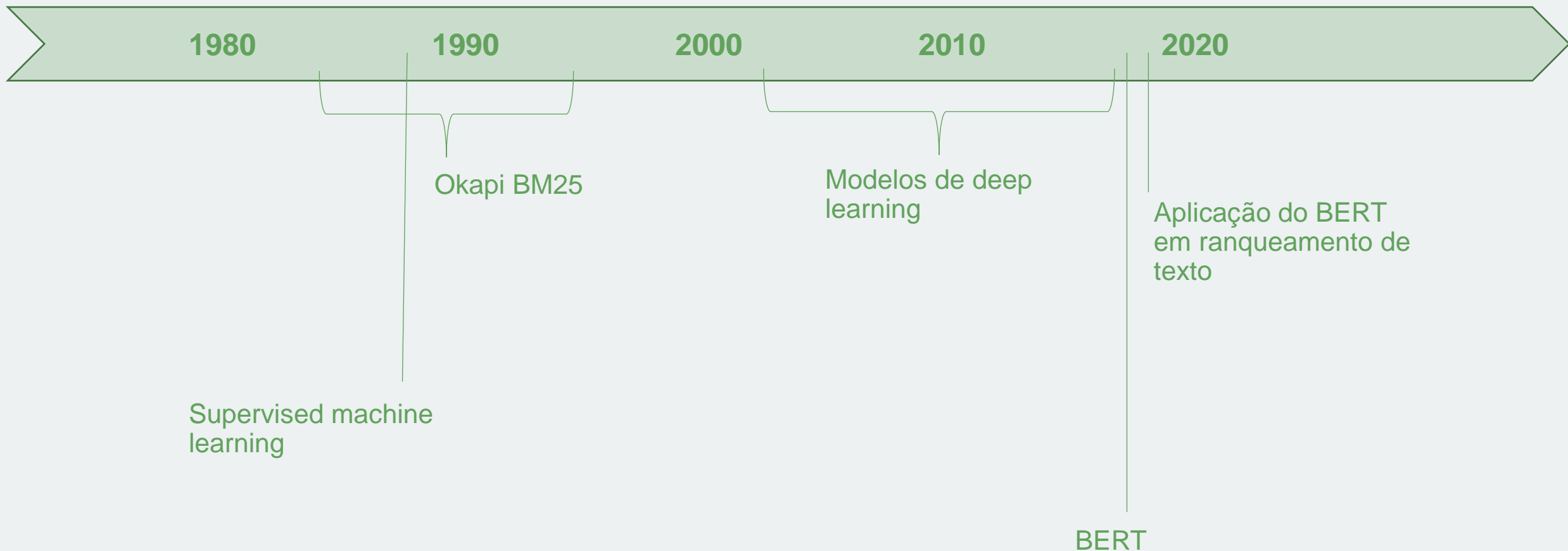
NLP: Ranquear texto para decidir entidades (por exemplo E1 vende A1 para E2)

Linha do tempo 1/2



Nessa época ainda era apenas busca exata: carro é diferente de automóvel

Linha do tempo 2/2



Aqui já se considera o significado das palavras: carro não é mais diferente de automóvel

Questão curiosa

Em 2018, Lin levantou a questão: *até que ponto modelos de ranqueamento baseados em redes neurais funcionam considerando dados que estão disponíveis publicamente?*

Em 2019, Yang et al fizeram uma análise dos trabalhos antigos e descobriram que não funcionavam.

Em 2019 isso mudou com o BERT...

Obrigado

Leandro Carísio
carisio@gmail.com