

Representações Pré-Treinadas

Modelo	Detalhamento das Siglas	Dados Filtrados	Linguagens	Dimensões	Tamanho do Vocabulário
xml-roberta-base	A Robustly Optimized BERT Pre Training Approach	2.5TB Common Crawl	100	768	250.002
xml-roberta-large	A Robustly Optimized BERT Pre Training Approach	2.5TB Common Crawl	100	1024	250.002
bert-base-multilingual-cased	Bidirectional Encoder Representations from Transformers	Largest Wikipedia (MLM)	104	768	119.547