



Processamento de Linguagem Natural com Transformers

Processamento de Linguagem Natural com Transformers

Projeto 2 Modelagem Avançada de Tópicos com Inteligência Artificial e BERTopic



Neste projeto vamos construir um modelo de IA capaz de fazer modelagem de tópicos com alto nível de precisão. E usaremos o incrível BERTopic para isso. Além da modelagem de tópicos faremos a interpretação dos resultados de forma visual com o gráfico dos c-TF-IDF scores, o Mapa de Distância Intertópico e o Hierarchical Cluster.

Os dados usados na modelagem de tópicos serão fornecidos a você e foram extraídos do link abaixo:

<https://www.kaggle.com/datasets/therohk/million-headlines>

Esse dataset contém textos de manchetes de notícias publicadas durante um período de dezenove anos. O dataset é originário da respeitável fonte de notícias australiana ABC (Australian Broadcasting Corporation): <https://www.abc.net.au>. O dataset é grande, mas usaremos uma amostra neste projeto.

Acompanhe as aulas na sequência. Bons estudos.