



Processamento de Linguagem Natural com Transformers

Processamento de Linguagem Natural com Transformers

Lab 7

Ajuste Fino de LLM Open-Source Com Dados da
Área de Direito Extraídos via Web Scraping



Processamento de Linguagem Natural com Transformers

Lab 7

Ajuste Fino de LLM Open-Source Com Dados da Área de Direito Extraídos via Web Scraping



Formação Engenheiro de Inteligência Artificial

Neste Lab vamos fazer o ajuste fino de um LLM Open-Source com dados da área de Direito extraídos via Web Scraping. Também usaremos métricas específicas para avaliar a performance do modelo após o ajuste fino.

Usaremos como fonte de dados os dados que foram extraídos via Web Scraping com base no procedimento disponível no link abaixo (que será explicado nas aulas a seguir):

<https://huggingface.co/datasets/ymoslem/Law-StackExchange>

O dataset gerado para o Lab será fornecido a você junto com os demais arquivos do Lab disponíveis ao final do capítulo.

Recomendamos ainda visitar o link abaixo com diversas opções de papers de pesquisa e datasets da área de Direito:

<https://github.com/abdoelsayed2016/Legal-Question-Answering-Review>

Os arquivos estão ao final do capítulo.

Bons estudos.