

Processamento de Linguagem Natural (PLN)

Revisar todo material

2.1. - Definição de texto a ser trabalhado, explicação do seu gênero textual e objetivos de negócio

- Justifique sua escolha para este estudo. Defina os objetivos de negócio a serem atingidos.

2.2 - PLN - Tratamento do texto: Remoção de Ruídos, Homogeneização e Stopwords

- Explicar o tratamento feito ao texto se foram necessárias remoção de pontuações, stopwords dentre outros tratamentos.

2.3 - PLN - Stemming / Lemmatization

- Qual a estratégia de stemming/ou lemmatization mais adequada ao seu contexto de forma a favorecer a qualidade do retorno das buscas semânticas (similaridade de vetores).

2.4 - Chunk / Embedding

- Definir a estratégia de chunk mais adequada, com ou sem overlap ?
- Qual a técnica de embedding mais adequada ?