



Exercícios de Revisão da Unidade I

- 1) Tradicionalmente, em PLN, distinguem-se alguns níveis linguísticos para que seja possível estudar e tratar os vários fenômenos da língua. Considerando apenas a linguagem escrita, esses níveis incluem a morfologia, a sintaxe, a semântica, o discurso e a pragmática. Explique com suas palavras os objetivos de cada um dos níveis linguísticos, identificando a interação entre eles.

- 2) Qual o resultado da expressão regular a seguir se aplicada no texto abaixo?

"[Rr]\\$[0-9]+"

Está cada vez mais caro abastecer os veículos nos postos de combustível em Palmas. O preço em R\$/litro da gasolina, por exemplo, pode chegar a R\$5,89 na capital, segundo pesquisa realizada pelo Procon no dia 1º de março. Antes da pandemia o preço do litro do combustível não passava de R\$4,67. As constantes altas, anunciadas pela Petrobras, preocupam consumidores e afetam setores da economia.

- 3) Qual resposta melhor explica a importância do POS tagging para o PLN?
- a) Permite calcular a frequência em cada classe gramatical.
 - b) Palavras de mesma classe gramatical compartilham propriedades.
 - c) Permite identificar a raiz ou radical das palavras.
 - d) Analisa como as palavras podem ser combinadas em sentenças bem formadas.
- 4) Considerando o pipeline de PLN, quais os principais processos na fase de Pré-processamento?
- a) tokenizador, remoção stop-words, stemming
 - b) tokenizador, remoção stop-words, stemming, POS Tagger
 - c) tokenizador, remoção stop-words, stemming, Parser Sintático
 - d) tokenizador, remoção stop-words, stemming, POS Tagger, NER
- 5) Considere o texto a seguir, dita por uma mãe ao seu filho pequeno. Apresente o resultado do pré-processamento, *POS tagging* e *parser* sintático, quando aplicados no texto abaixo.

“Eu mandei você estudar, e não brincar. Você está de castigo!”

- 6) Qual o resultado da tarefa de NER no texto da questão 2, usando o SpaCy ?

- 7) Apresenta a análise de Chunks no texto abaixo, usando a abordagem por expressão regular ou por regras ChunkRule / ChinkRule?

As constantes altas no preço da gasolina, anunciadas pela Petrobras, preocupam consumidores e afetam setores da economia.

Obs: como estratégia de estudo, sugiro responder às questões práticas primeiro realizando a análise manual, e depois conferindo sua resposta com o resultado de programas Python, usando a biblioteca NLTK e SpaCy.