
UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE COMPUTAÇÃO

Processamento de Linguagem Natural
Professor: Thales Vieira

1a lista de exercícios

3 de junho de 2020

Instruções:

A lista deve ser respondida por grupos de até 3 pessoas (graduação) e individualmente (mestrado).

Resoluções idênticas de grupos distintos serão desconsideradas.

O código e demais dados devem ser anexados a cada questão.

Data limite para entrega: 10/06/2020.

1. Implemente uma função usando expressões regulares que teste se um string contém uma sequência de letras minúsculas seguidas por um *underscore*. Mostre exemplos.

2. Implemente uma função usando expressões regulares que testa se um string representa um CPF no formato correto. Mostre exemplos.

3. Implemente uma função usando expressões regulares que seja capaz, com uma única expressão regular, de extrair rua, número, apartamento (opcionalmente), bairro (opcionalmente), CEP, cidade e estado. Esta função deve ser capaz de funcionar em todos os exemplos abaixo:

- Rua José da Silva, 346, Farol, CEP 57002-220. Maceió, AL.
- Rua da Consolação, 9999, apt 302, CEP 11022-202. São Paulo, SP.
- Avenida Atlântica, 420, ap 1001, Copacabana, cep 22011-010. Rio de Janeiro, RJ.

4. Implemente uma função usando expressões regulares que encontre todas as URLs em um string. Mostre exemplos.

5. Considere o seguinte exemplo de referência de livro em formato de citação APA:

Manning, C. D., Manning, C. D., & Schutze, H. (1999). Foundations of

statistical natural language processing. MIT press.

Implemente uma função usando expressões regulares que extraia cada autor, ano de publicação, título e editora do livro, para qualquer referência de livro neste formato.