
**UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE COMPUTAÇÃO**

**Processamento de Linguagem Natural
Professor: Thales Vieira**

6a lista de exercícios

09 de agosto 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: 23/08/2020.

Usando sua base de textos após os pré-processamentos realizados na lista 2, realize as seguintes tarefas:

1. Resolva novamente a segunda questão da 3a lista usando pelo menos duas arquiteturas de redes neurais que utilizem camadas Embedding, convolucionais e LSTM. Compare com os resultados obtidos anteriormente nas lista 3 e 5.

2. Usando sua base de textos:

- a) Treine uma rede LSTM para gerar texto, que receba uma ou mais palavras de uma frase como entrada. O treinamento deve ser realizado considerando um conjunto supervisionado que gera a próxima palavra de uma sequência de tamanho 4, usando subsequências de sua base.
- b) Após o treinamento, exiba pelo menos 5 exemplos de textos dados de entrada, e do texto gerado em seguida pela rede treinada. Para cada exemplo, gere pelo menos 10 palavras consecutivamente.
- c) Faça o mesmo usando Cadeias de Markov com bi-grams (usando apenas 1 palavra para tentar prever a seguinte). Compare os resultados com os da LSTM.