



Processamento de Linguagem Natural

UNIDADE 03

Similaridade textual

Olá, estudante! Nesta semana, vamos entender o conceito de similaridade textual, que é muito importante para algoritmos de PLN, pois, sabendo medir a similaridade entre palavras, nossos algoritmos são capazes de entender diferentes representações de um mesmo significado. Vamos lá?

| O que é similaridade léxica?

O conceito de similaridade textual é muito importante para diversas tarefas de PLN e existem diversos métodos de cálculo de similaridade. Antes, porém, precisamos entender o conceito de similaridade léxica, que é apresentado na Unidade de Aprendizagem **Similaridade léxica**, disponível na semana.

Vamos agora trabalhar de maneira prática com os métodos de similaridade léxica, tomando como base o *notebook* a seguir. Você também encontrará um vídeo tutorial em que executamos o *notebook* para entender os métodos de similaridade léxica.

Métodos de similaridade léxica



Notebook – Google Colab

Explore o *notebook* a seguir, que demonstra a utilização de métodos de similaridade léxica.

https://colab.research.google.com/drive/1-4uYqTexZ2wlwCmdjDdi4dMVnxXF9JQ_

| Como a similaridade semântica difere da léxica?

Além dos diversos métodos de cálculo de similaridade léxica, apresentados na unidade anterior, devemos estar atentos ao conceito de similaridade semântica, que está mais associado ao significado das palavras do que à sua forma. Nesta Unidade de Aprendizagem **Similaridade semântica**, disponível na semana, você será apresentado ao conceito de similaridade semântica e a métodos de cálculo.

| Conclusão

Nesta unidade, você aprendeu a:

- Definir o conceito de similaridade textual.
- Diferenciar o papel da similaridade textual em diferentes aplicações de PLN.
- Demonstrar métodos de aplicação de similaridade textual.
- Diferenciar as similaridades textuais dos tipos léxico e semântico.
- Determinar a similaridade semântica entre palavras.
- Descrever métodos baseados em ontologias de conhecimento.

| Referências

GAN, M.; DOU, X.; JIANG, R. From ontology to semantic similarity: calculation of ontology-based semantic similarity. **The Scientific World Journal**, [s. l.], v. 2013, p. 1-11, 2013.

GOMAA, W. H.; FAHMY, A. A. A survey of text similarity approaches. **International Journal of Computer Applications**, [s. l.], v. 68, n. 1, p. 13-18, 2013.

JURAFSKY, D.; MARTIN, J. H. **Speech and language processing**. Disponível em: <https://web.stanford.edu/~jurafsky/slp3/>. Acesso em: 15 dez. 2020.

