

PCS - Departamento de Engenharia de Computação e Sistemas Digitais

Engenharia Elétrica – Ênfase Computação

Tema:

Topicalização de Nível de Serviço

Introdução

Conjuntamente com a virada do século XXI, novos comportamentos de interação sociais surgiram, sendo que grande parte ocorre no meio online. Dada esta realidade, o monitoramento da comunicação nestes meios tecnológicos, veem se tornando cada vez mais importante para empresas. Desta forma, algoritmos de processamento de linguagem natural são cada vez mais necessários e desejados.

Objetivo

Topicalizar comentários sobre nível de serviço, com índice de acerto superior a 80%.

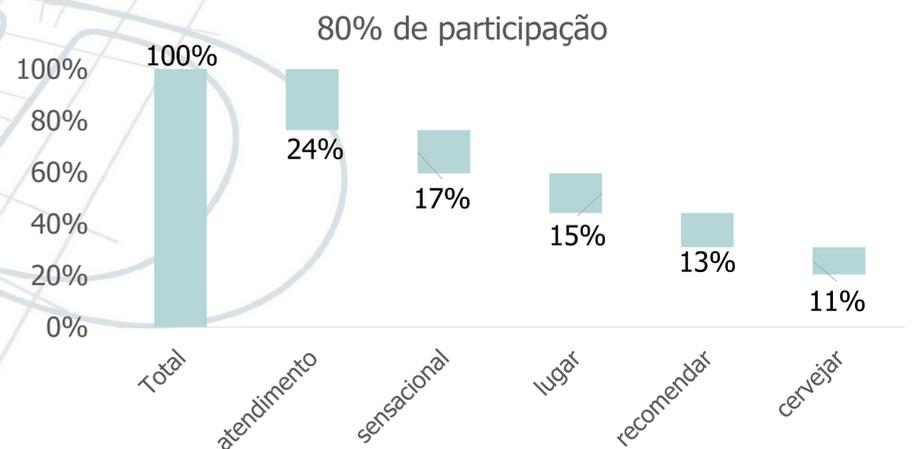
Metodologia

O algoritmo de topicalização foi organizado em módulos unitários funcionais a fim de conferir maior eficiência no desenvolvimento do produto. Cada um destes foi trabalhado em ciclos de planejamento, produção, testagem e correção e por fim integrados entre si. As unidades funcionais são listadas abaixo em ordem temporal de execução.

- Importação de dados
- Pré-Processamento
- Vetorização
- Clusterização
- Consolidação de Clusters

Resultados e Discussões

Inúmeras técnicas e modelos foram testados para se alcançar os objetivos definidos neste trabalho de conclusão de curso. Como desfecho foi perceptível que o modelo com melhores resultados foi o utilizando métodos de NLP de pré-processamento, Word embedding, Clusterização baseado em densidade e TF-IDF. Abaixo é possível observar um exemplo de resultados do algoritmo de topicalização em uma base de *feedback* de um bar.



Topicalização Completa		
Cluster	Classificados	Participação
atendimento	123	24%
sensacional	87	17%
lugar	80	15%
recomendar	69	13%
cervejar	55	11%
voltar	49	9%
comida	17	3%
ambiente	10	2%
local	6	1%
valer	5	1%
notar	4	1%
sempre	3	1%
chopp	3	1%
pena (valer a pena)	3	1%
show	3	1%
esplendor (esplendido)	3	1%

Considerações Finais

Em um contexto de recursos humanos e temporais limitados, foi cumprido o objetivo proposto com um produto funcional. Porém, existem incrementos que podem ser realizados para o refino do algoritmo, como por exemplo a utilização de modelos baseados em BERT e bi-gram e tri-gram.

Integrantes: Eric Koji Yang Imai
Henrique Minoru Hattori

Professor Orientador: Prof. Dr. Ricardo Luiz de Azevedo da Rocha