




Análise de corpus em comentários de produtos comprados online

Guilherme L. Ventre, 11201811663
Sandy M. Nascimento, 11201811732



Análise das características de um produto através dos comentários de seus consumidores

Nos últimos anos houve um forte crescimento do setor e-commerce e cada vez mais pessoas recorrem à internet para adquirir produtos de diferentes tipos.

Quais informações podem ser obtidas através dos comentários?

- Uma forma de auxiliar na decisão de comprar ou não um produto é através da leitura de comentários feitos por pessoas que já compraram aquele produto e deram sua opinião sobre o mesmo.
- Pensando nisso, utilizamos dos comentários feitos online para extrair conhecimentos sobre um produto, obtendo características citadas nas experiências de diversos consumidores.
- Nosso objetivo foi facilitar a visualização dessas características, com o intuito de demonstrar aquilo que há de mais relevante no produto na visão de outros consumidores, e assim, auxiliar na decisão de compra do produto.

Produto selecionado:

Fire TV Stick

- Possibilita usar sua TV tradicional (antiga) como uma SmartTV (acesso a aplicativos e serviços de streaming).
- Diversas pessoas não tem condição de adquirir uma nova TV ou desejam dar utilidade a uma TV antiga sem ter que assinar TV a cabo.
- Portanto é um produto que fez muito sucesso nos últimos anos e com isso acumula diversos comentários.



Base de dados

- Páginas específicas para visualizar comentários (10 comentários por página), coletamos informações de 30 páginas diferentes (300 comentários no total).
- Webscrapping para coletar os comentários e seus metadados.
 1. ID da review
 2. Nome do consumidor
 3. Nota para o produto (formato string)
 4. Nota para o produto (formato float)
 5. Título do comentário
 6. Corpo do comentário
 7. Data do comentário
 8. Data de acesso
 9. Link de acesso
 10. Região do comentário
 11. Veracidade da compra

| id | 0 | 1 | 2 |
|----------------------|---|---|---|
| Nome | Mario Camara | Ítalo Gregório | Samuel Souza |
| Rating | 5,0 de 5 estrelas | 4,0 de 5 estrelas | 5,0 de 5 estrelas |
| Num Rating | 5 | 4 | 5 |
| Título | Cumpre o que promete... | Melhor TV Stick que há! Só falta liberar HBO M... | Sensacional! |
| Comentário | Nao eh caro (saiu R\$ 250,00 com um cupom que d... | Comprei o Fire TV Stick HD c/ comandos de TV p... | Mano, não tem o que dizer! Minha tv é uma AOC ... |
| Data do comentário | Avaliado no Brasil em 23 de outubro de 2022 | Avaliado no Brasil em 31 de julho de 2022 | Avaliado no Brasil em 29 de novembro de 2022 |
| Data de acesso | 03 de dezembro de 2022 | 03 de dezembro de 2022 | 03 de dezembro de 2022 |
| Link de acesso | https://www.amazon.com.br/Fire-TV-Stick-Stream... | https://www.amazon.com.br/Fire-TV-Stick-Stream... | https://www.amazon.com.br/Fire-TV-Stick-Stream... |
| Região do comentário | Brasil | Brasil | Brasil |
| Veracidade da compra | Compra verificada | Compra verificada | Compra verificada |

Corpus

I. Corpus com o título dos comentários.

- I. Curtos
- II. Direto
- III. Elogio ou crítica

II. Corpus com o corpo dos comentários.

- I. Longos
- II. Características
- III. Informações específicas

Resultados - Títulos

Composto por:

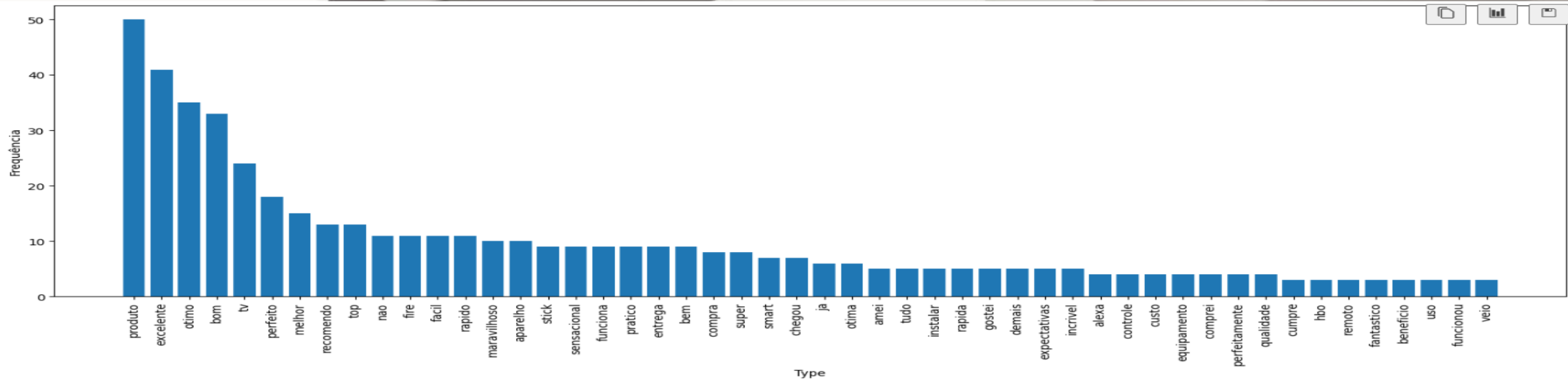
- 1.002 tokens
- 300 types

Densidade lexical:

- 29,94%

Palavras mais frequentes:

- Produto
- Excelente
- Ótimo
- Bom
- TV



Resultados - Comentários

Bigramas

Utilizando todas os types presentes no corpus:

| frequency | ngram |
|-----------|------------------|
| 27 | fire tv |
| 21 | tv smart |
| 20 | fire stick |
| 18 | tv nao |
| 17 | smart tv |
| 17 | facil instalar |
| 16 | facil instalacao |
| 14 | controle remoto |
| 13 | produto bom |
| 12 | facil usar |

Removendo types relacionados ao produto:
(TV, fire, stick, smart, controle)

| frequency | ngram |
|-----------|------------------------|
| 17 | facil instalar |
| 16 | facil instalacao |
| 13 | produto bom |
| 12 | facil usar |
| 11 | hbo max |
| 9 | funciona bem |
| 8 | vale pena |
| 8 | produto excelente |
| 8 | funciona perfeitamente |
| 7 | super recomendo |

Conclusões

- I. Os comentários são eficazes para destacar características dos produtos (usabilidade, qualidade, velocidade de entrega, cumprimento da proposta, etc)
- II. Os títulos dos comentários possuem relevância para adjetivos qualitativos do produto (bom ou ruim, funciona ou não, o quanto gostou ou não gostou, etc)

Trabalhos Futuros

- I. Criar uma ferramenta que proporciona ao usuário escolher o produto que deseja a análise.
- II. Aumentar a quantidade de informações usando outros sites de venda online ou mais páginas do próprio site para obter um resultado mais completo.

Referências

- Webscrapping: <https://www.geeksforgeeks.org/web-scraping-amazon-customer-reviews/>
- Webscrapping: <https://python.plainenglish.io/web-scraping-amazon-reviews-with-python-and-beautiful-soup-2ca41ebd2876>
- API para Webscrapping: <https://developers.whatismybrowser.com/>
- Analises básicas, WordCloud, Frequências: Atividades I, II e III de laboratório
- Bigramas: <https://towardsdatascience.com/text-analysis-basics-in-python-443282942ec5>
- Bigramas: <https://www.analyticsvidhya.com/blog/2021/09/what-are-n-grams-and-how-to-implement-them-in-python/>