
Análise de Sentimentos

— Abordagens com Inteligência Artificial —

Felipe Reis, Luciana Hamanaka e Oscar Firme

Introdução

- O que é análise de sentimentos?
- Como Funciona a análise de sentimentos?
- Polaridade
 - Polaridade + = Amei o presente!
 - Polaridade - = Já estou com raiva desse governo.
 - Polaridade 0 = A camisa é branca.



Polaridade

Outras Abordagens e possibilidades:

- Emojis:



Like



Love



Haha



Yay



Wow



Sad



Angry

Cuidado: por detrás dos emojis, há alguém medindo a sua satisfação!

- Hashtag(#):



Desafios

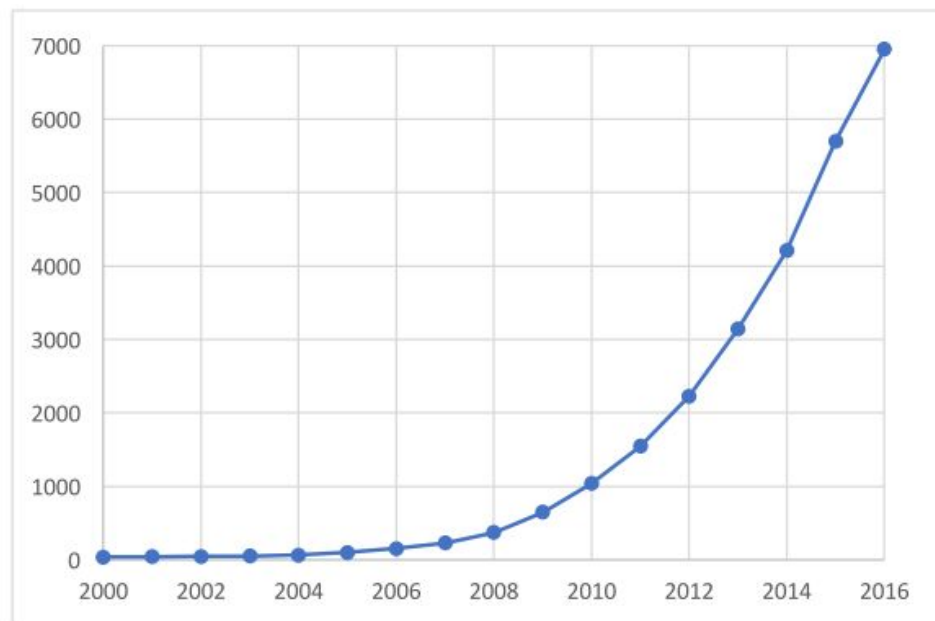
A análise de sentimentos possui diversos desafios a serem ultrapassados, como:

- Ruídos
- Ironia
- Ferramentas basicamente para o idioma inglês
- Questões morais

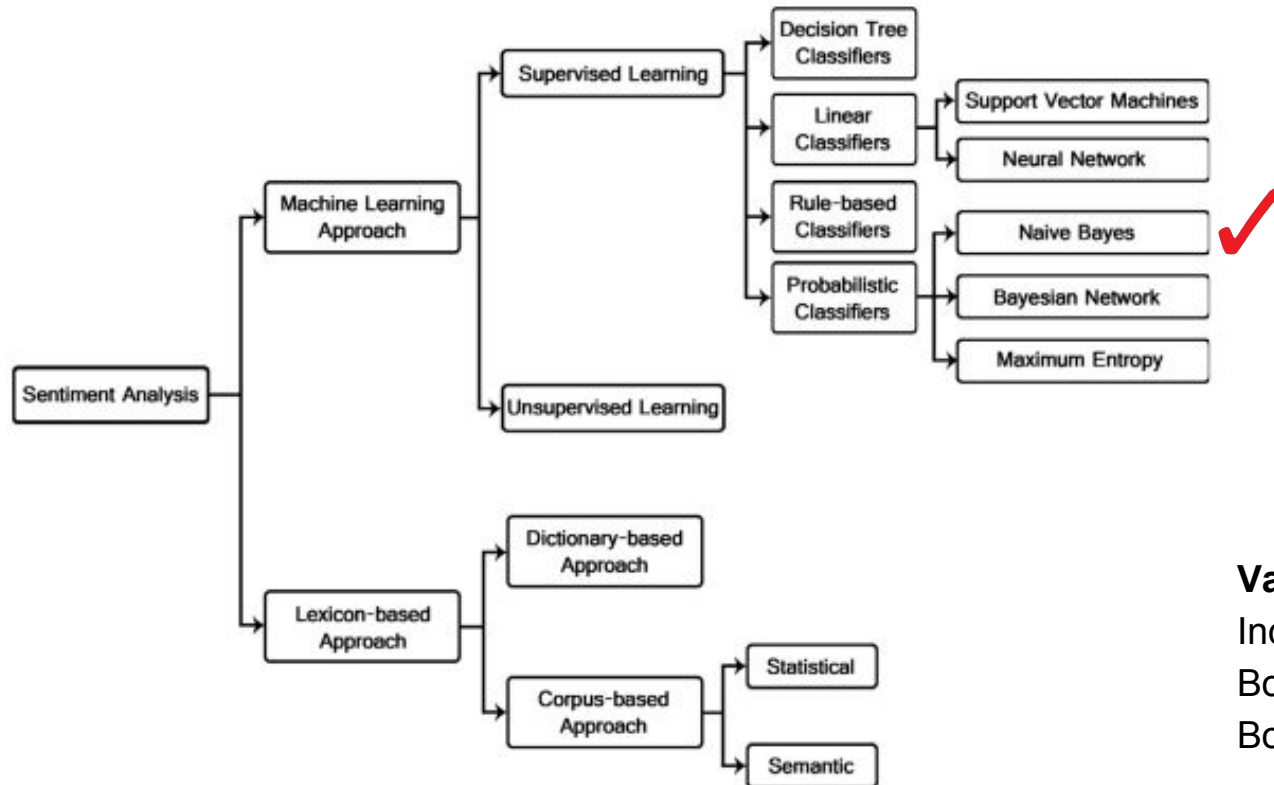
Utilizada...

Crescimento:

- Grande atenção nos anos 2000



Técnicas de mineração de opinião



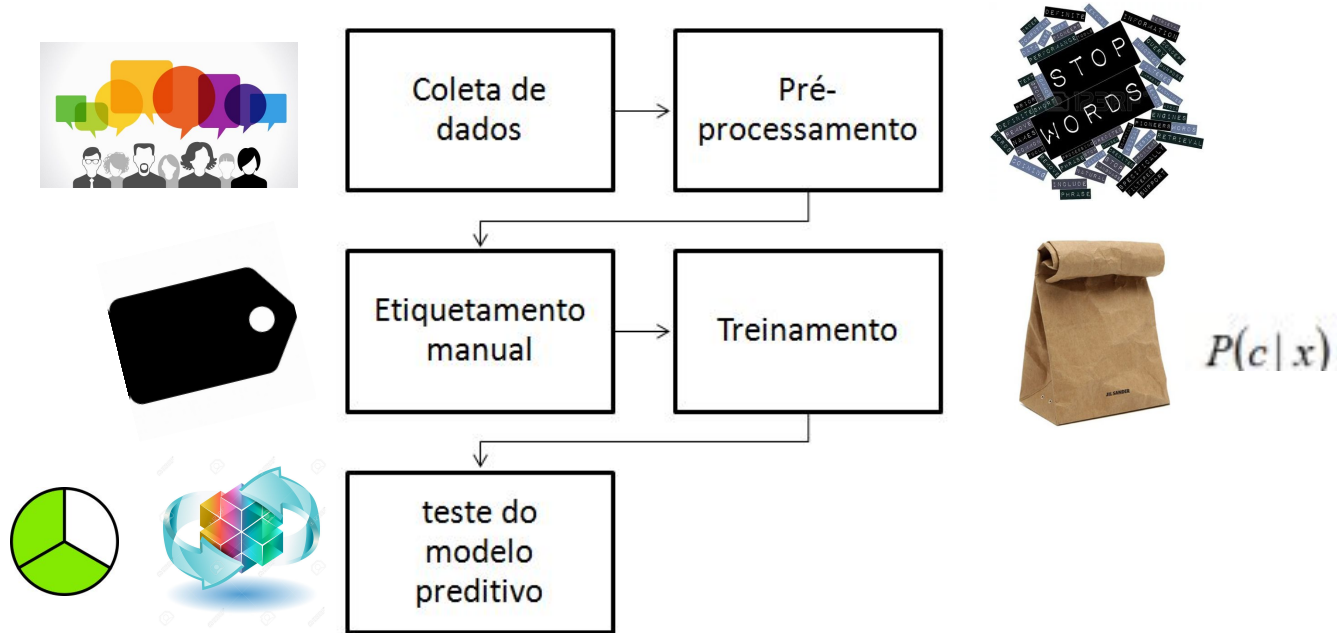
Vantagens:

Independente do idioma

Bom para textos pequenos

Bons resultados na literatura

Etapas genéricas do método



Bag of words

Recorte do algoritmo implementado

#Conjunto de importações

```
from sklearn.feature_extraction.text import CountVectorizer
```

#Cria uma instância para a bag-of-words

```
bag = CountVectorizer()
```

#Método fit_transform:

#fit = cria e aprende a bag

#transform = cria a matriz termo-documento

```
bag_treino = bag.fit_transform(dados_treino)
```

#Printa o vocabulário criado para a bag-of-words

```
print(bag.vocabulary_)
```

#A função sorted() ordena o vocabulário da bag-of-words

```
print(sorted(bag.vocabulary_))
```

#Printa a matriz termo-documento da bag-of-words

Conjunto de treinamento:

['Dois mil e dezesseis foi um ano filho da puta.', 'Tô aqui pedindo que a gente se junte nessa vibração de amor e arte.']

Polaridades do Conjunto de Treinamento

[-1, 1]

Bag:

['amor', 'ano', 'aqui', 'arte', 'da', 'de', 'dezesseis', 'dois', 'filho', 'foi', 'gente', 'junte', 'mil', 'nessa', 'pedindo', 'puta', 'que', 'se', 'tô', 'um', 'vibração']

Matriz termo-documento

(Frase,Palavra) TF (frequência do termo)

(0, 7) 1

(0, 12) 1

Teorema de Bayes: pensemos de forma probabilística!

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

Likelihood
Class Prior Probability
Posterior Probability
Predictor Prior Probability

$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$$

Palavra	Polaridade
Bom	1
Ruim	-1
Droga	-1
Bem	1
Droga	1
Mal	-1
Bem	1
Bom	1
Ruim	-1
Bem	-1
Droga	-1
Bom	1
Mal	-1
Droga	-1
Mal	1
Bom	1
Bem	1

Tabela de frequência		
Palavra	Polaridade 1	Polaridade -1
Bom	4	0
Ruim	0	2
Droga	1	3
Bem	3	1
Mal	1	2
Total	9	8

Tabela de probabilidade		
Palavra	Polaridade 1	Polaridade -1
Bom	4	0
Ruim	0	2
Droga	1	3
Bem	3	1
Mal	1	2
Todas	9	8
	9/17 = 0,529	8/17 = 0,471

Problema: A polaridade é positiva (1) se a palavra é "Bom"?

$$P(1|Bom) = P(Bom|1) * P(1) / P(Bom)$$

$$P(Bom|1) = 4/9 = 0,444$$

$$P(1) = 9/17 = 0,529$$

$$P(Bom) = 4/17 = 0,235$$

$$P(1|Bom) = 0,444 * 0,529 / 0,235 = 0,999 \quad (\text{Alta probabilidade})$$

Ferramenta

- Senticnet

É uma ferramenta para a mineração de opinião e análise de sentimentos que explora técnicas de Inteligência Artificial e Web Semântica. O método utiliza técnicas de Processamento de Linguagem Natural (PLN) para criar significados semânticos ou polaridade para aproximadamente 50.000 conceitos.

SenticNet

Helping machines to
learn, leverage, love.

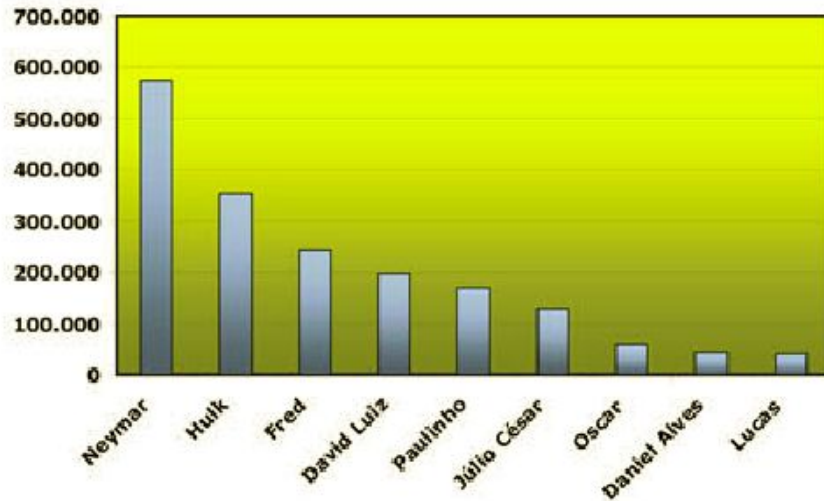
A screenshot of the SenticNet web interface. It features a light blue rounded rectangular container. Inside, there is a white text input field containing the sentence "the car is expensive but i like the color". To the right of the input field is a green button with the text "Get sentics" in white.

Sentence polarity - positive

Aplicações

A análise de sentimentos pode ser utilizada para:

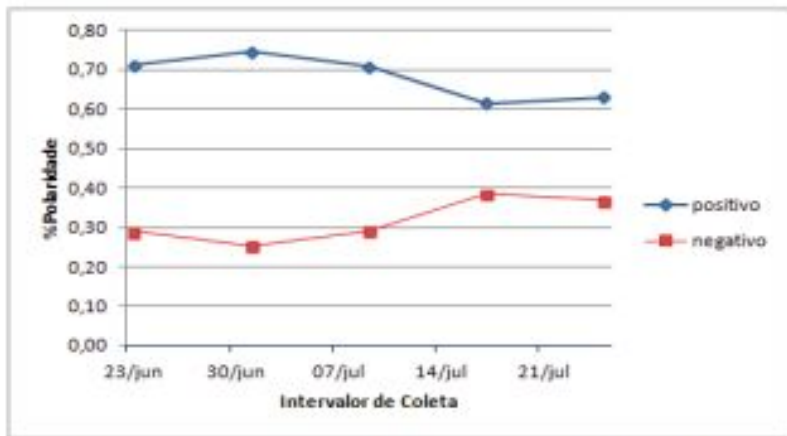
- Análise e cobertura de grandes eventos esportivos: Durante a copa das confederações a IBM analisou o sentimento da torcida, onde encontraram os seguintes dados:



Aplicações

- Análise do posicionamento da população durante os protestos em 2013.

Análise da polaridade em mensagens no Twitter. 

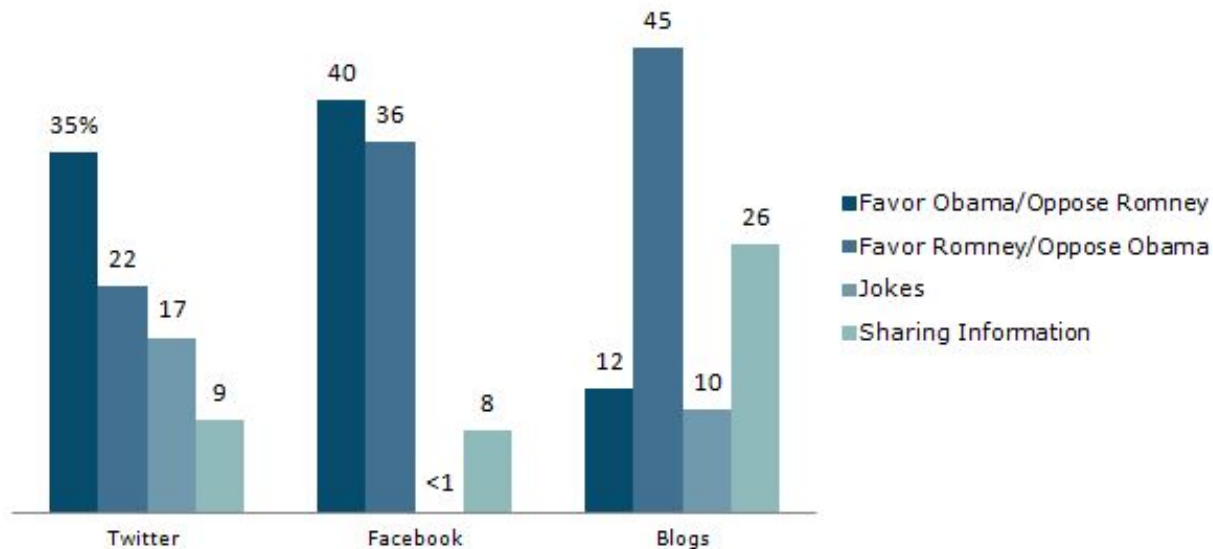


Sentimentos das Mensagens Analisadas



Aplicações

- Análise em tempo real de redes sociais sobre eleições em 2012 nos EUA. Onde é possível verificar dados interessantes.



Referências

- IBM Explica: Análise de Sentimento Social - <https://www.youtube.com/watch?v=IUHtnguLsKI>
- Análise de Sentimento de Tweets Relacionados aos Protestos que ocorreram no Brasil entre Junho e Agosto de 2013 -
<https://drive.google.com/file/d/0B5vnay4MiKMRSkVLWjVWV1Biem8/view>
- The Evolution of Sentiment Analysis - A Review of Research Topics, Venues,
and Top Cited Papers - <https://arxiv.org/ftp/arxiv/papers/1612/1612.01556.pdf>
- Política nos EUA:
[Http://www.journalism.org/2012/10/05/social-media-debate-sentiment-less-critical-obama-polls-and-press-are/](http://www.journalism.org/2012/10/05/social-media-debate-sentiment-less-critical-obama-polls-and-press-are/)

Métodos de análise

Existem diversos métodos de análise:

- Aprendizado de máquina
- Dicionários léxicos
- Processamento de linguagem natural
- Escalas psicométricas