

## CIÊNCIAS DE DADOS (BIG DATA PROCESSING AND ANALYTICS)

Arquitetura de *Big Data*







Universidade Presbiteriana  
**Mackenzie**

**Modalidade a distância**

## **RECUPERAÇÃO DA INFORMAÇÃO NA WEB E EM REDES SOCIAIS**

**Trilha 6 – Processamento de Linguagem Natural:  
análise de sentimentos**

**Professor: Luciano Moreira Camilo e Silva**

# Sumário

1. Introdução à Trilha .....	4
1.1. Caso prático .....	5
2. Recuperação de informação via API .....	7
2.1. Twitter .....	8
2.2. Tweepy .....	10
3. Análise de sentimentos .....	11
3.1. IBM Watson .....	11
3.2. Naive Bayes .....	14
3.2.1 Teoria .....	14
3.2.2 Código .....	15
4. Síntese.....	17
5. Referências .....	18

## 1. Introdução à Trilha

O *homo sapiens* é uma espécie de primatas bípedes, a única do gênero *homo* ainda viva (GOODMAN et al., 1990) Nos diferenciamos dos demais primatas e demais animais, de modo geral, pelo cérebro desenvolvido e por nossa capacidade de raciocínio abstrato, linguagem sofisticada e outros fatores que destacam nossa inteligência.

A espécie humana pode ser considerada um grupo de animais gregários, ou seja, somos uma espécie que vive em comunidades. Das tribos antigas às megalópoles modernas, o ser humano busca se reunir em sociedade e trocar experiências, mesmo que, algumas vezes, belicosas entre tribos diferentes.

Essa troca de experiência se faz presente em nosso dia a dia. Conversamos nossos amigos sobre a quais filmes assistiremos nos cinemas ou em nosso serviço de *streaming* favorito. Ouvimos nossos colegas sobre em quais restaurantes vale a pena gastar um pouco a mais para comer bem ou onde se economiza. Confiamos em nossos familiares para ouvirem nossos dilemas e esperamos dicas de como agir. Contamos uns com os outros para tomar decisões.

A tecnologia nos permitiu trocar mais informações entre nós, com os telégrafos e telefones. A internet, principalmente após as grandes evoluções da Web 2.0 da década passada (O'REILLY, 2009), permitiu experimentar o início de um metaverso digital, onde trocamos informações com pessoas que, às vezes, não conhecemos, mas que, muitas vezes, confiamos.

Figura 1 – Tim O'Reilly



Fonte: [SXSW Interactive 2011, Austin, TX](#), by [Kris Krug CC BY-SA 2.0](#)

Nesta trilha, avaliaremos algumas mensagens disponíveis publicamente na web e, usando processamento de linguagem natural, identificaremos os sentimentos que essas mensagens expressam.

## 1.1. Caso prático

Como nas outras trilhas em que foi tratado o tema de Processamento de Linguagem Natural, o aprendizado desta será acompanhado de um caso prático.

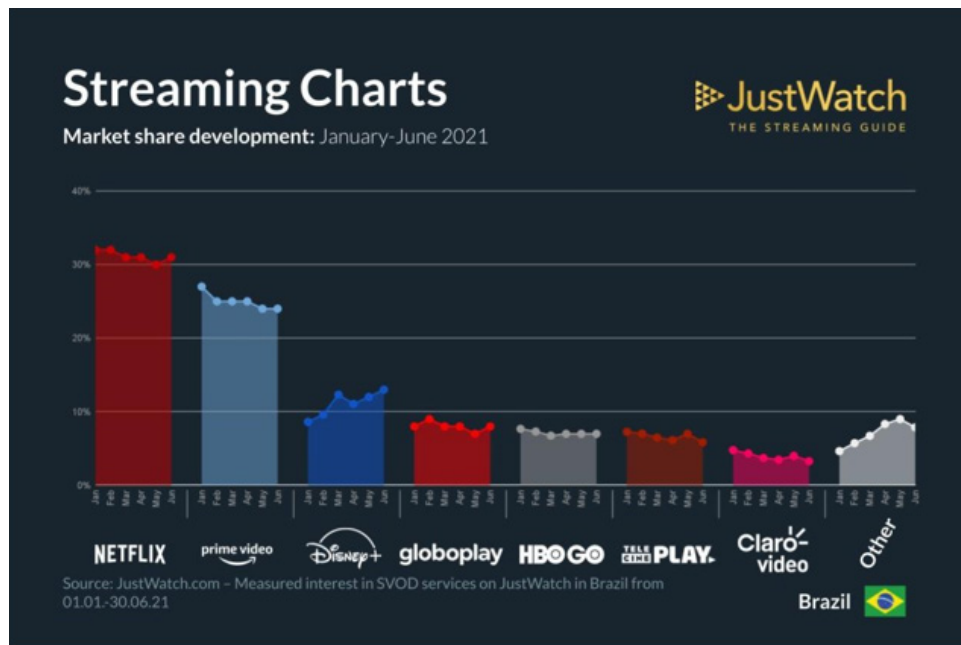
Para fugir dos clichês encontrados nos tutoriais espalhados pela web – e que são muito úteis para o aprendizado, sem dúvida – foi escolhido um exemplo não trivial: um vídeo sob demanda.

O mundo vive um crescimento no número de serviços de streaming de vídeo por assinatura, ou SVOD (Subscription Video On Demand) no acrônimo em inglês. A imprensa vem noticiando o duelo entre as diferentes empresas do setor como “a guerra dos streamings” como lemos nas matérias de sites estadunidenses Vulture (ADALIAN, 2021) e Observer (KATZ, 2021).

No Brasil, já temos mais de uma dezena de serviços de vídeo sob demanda disponíveis ao público (ANDRADE, 2021), partindo daqueles que atuam em nichos, como o Oldflix (MENDES, 2016), até as gigantes do streaming mundial, Netflix, HBO MAX e Amazon Prime. Dos serviços nacionais, o maior é o Globoplay da rede Globo de televisão, de acordo com pesquisa da JustWatch apresentada na Figura 2.

O caso prático dessa trilha será monitorar o sentimento dos usuários da rede social Twitter para os cinco maiores serviços de streaming no país.

Figura 2 – Serviços de streaming mais usados no Brasil



Fonte: <[justwatch.com](https://justwatch.com)>





o boom no surgimento de redes sociais. As primeiras a surgir, que inspiraram as demais, foram SixDegrees (1997), Cyworld (2001) e Friendster (2003). Aprendendo com essas, vieram o MySpace e LinkedIn e, em 2004, a rede social que passou a ser referência: o Facebook.

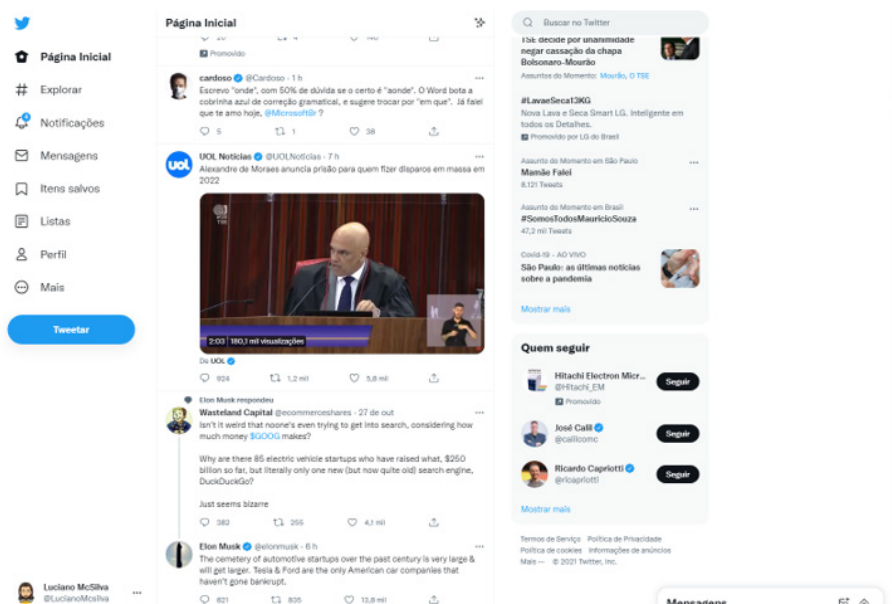
Com o Facebook dominando as redes sociais de uso mais amplo, começaram a surgir outros sistemas focados em nichos, como o Flickr para fotos (2004), YouTube para vídeos (2005) e um conceito de microblogue foi criado junto com o Twitter (2006).

Desde então, a maioria dos serviços da web passaram a oferecer algum tipo de relação social, seja por integração com as grandes redes estabelecidas, seja procurando um novo nicho para se estabelecer.

## 2.1. Twitter

O Twitter foi a primeira e é a mais popular rede social de microblogue<sup>2</sup> do mundo, com mais de 206 milhões de usuários ativos (TWITTER, 2021) que trocam bilhões de mensagens todos os meses.

**Figura 4 – Interface gráfica do site Twitter**



<sup>2</sup> “Microblogue (do inglês, micro-blogging) é uma forma de informação de blogue que permite aos usuários que façam atualizações breves de imagens e texto (geralmente com menos de 200 caracteres) e publicá-las, para que sejam vistas publicamente ou apenas por um grupo restrito escolhido por si. Estes textos podem ser enviados por uma diversidade de meios tais como SMS, mensageiro instantâneo, correio eletrônico, MP3 ou pela Web.” Microblogue (Wikipédia).

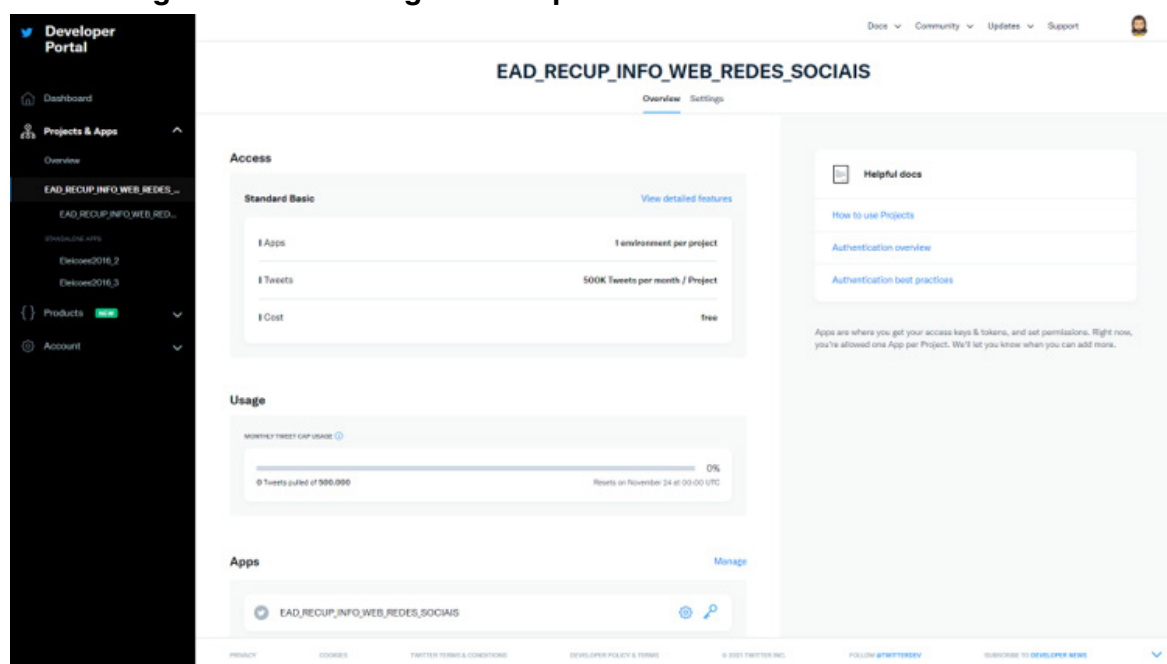


O uso da rede social no dia a dia, na troca de informações entre seus usuários tem influência maior do que apenas o entretenimento e a disseminação de dados. A rede social cresceu em presença a ponto de ter dados solicitados para averiguar possível influência nas eleições brasileiras de 2018 (GONÇALVES, 2021).

Além de eleições, é comum marcas e veículos de imprensa monitorarem as mensagens da rede social e utilizar como termômetro do que acontece. E é isso que avaliaremos nesta trilha para o assunto de assinatura por vídeos sob demanda: como os usuários dessa rede social avaliam – positiva ou negativamente – esses serviços.

A primeira etapa que precisamos construir é o sistema de recuperação de informação. Por ser um site moderno, o Twitter conta com APIs que facilitam a construção desses sistemas.

**Figura 5 – Interface gráfica do portal de desenvolvedores do Twitter**



Qualquer pessoa pode criar uma aplicação no site de desenvolvedores do Twitter.<sup>3</sup> A conta gratuita permite que se recupere dados de até sete dias em um limite de até 500 mil tweets por mês, mais do que suficiente para o exercício da trilha.

Ao criar a aplicação, serão fornecidas duas chaves criptográficas que não devem ser compartilhadas: API KEY e API SECRET. Para facilitar a construção do nosso robô de captura de informações do Twitter, utilizaremos a Biblioteca Tweepy.

<sup>3</sup> Acesse o site em <<https://developer.twitter.com/>>.

## 2.2. Tweepy

A biblioteca tweepy faz a abstração de boa parte das complexidades de se trabalhar com uma API, encapsulando, em classes e métodos do Python, todas as funções necessárias para requisitar uma informação.

Abaixo, um código exemplificando a busca de tweets contendo a palavra netflix, usando o tweepy.

```
1. import tweepy
2.
3. #Passa as credenciais para o tweepy fazer toda a comunicação de autenticação
4. auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
5.
6. #Com o resultado da autenticação, inicializa a API
7. api = tweepy.API(auth)
8.
9. #Executa uma busca, pela API de search
10. tweets = api.search_tweets(q='netflix')
11.
12. #Imprime o resultado
13. for tweet in tweets:
14.     #Imprime o tweet, removendo quebras de linha se houver
15.     print(''.join(tweet._json['text'].splitlines(False)), end='\n\n')
```

Utilizando as mesmas técnicas da Trilha 3, é possível construir uma aranha para puxar e armazenar todos os tweets. Para facilitar o exercício, uma amostra com nove mil tweets será disponibilizada para análise.

Para o exercício, foram recuperados os tweets entre os dias 16/10/2021 e 25/10/2021. Abaixo, uma amostra do que foi coletado.

**Figura 6 – Amostra de 15 registros dos tweets coletados**

	id_str	text	screen_name	search_term	search_term_group	source
3583	1450267147377201152	chegou Maligno na hbo max mas quem ...	taisouzax	hbo max	hbomax	Twitter for iPhone
4191	1452387178592411648	Os lançamentos da Netflix em novembro de 2021 ...	canaltech	netflix	netflix	canaltech
8009	1449871350915637248	Uma obra prima \nhttps://t.co/hU16a4q1Oy	muhkent	primevideo	primevideo	Twitter for Android
5056	1451326165713670144	Sacrifício é ter que aguentar o Luciano Hulk n...	hatakerenata	netflix	netflix	Twitter for Android
5757	1450876023554330624	@alleywaysx @SemideusGrecia @danielbfx @marcos...	sr_albertini	netflix	netflix	Twitter for Android
1051	1450589824709185536	@dudadopke Com dólar tem Disney, com bitcoin t...	brurichards	disney+	disney+	Twitter for iPhone
5708	1450897163790204928	"Rebelde": Netflix divulga clipe e data de lan...	papelpop	netflix	netflix	Twitter Web App
7095	1449921497720557568	@a_alinediniz @rvandromel Doces Magnolias na N...	patricia_gomes	netflix	netflix	Twitter for iPhone
2585	1451015167685365760	eu tô tão tentado a assinar a globoplay só pra...	ltsMarchiori	globoplay	globoplay	Twitter for iPhone
4308	1452219366603362304	Para os órfãos de Rick and Morty, achei essa o...	Tiagoarruda666	netflix	netflix	Twitter for Android
947	1450658198285860864	O mc gui salvou o rico da roça kkkkkkkkk a car...	Eduardo92375039	disney+	disney+	Twitter for iPhone
3333	1451697896487849984	hbo max me salvando c ben 10 classico liga da...	afeholipe	hbo max	hbomax	Twitter for Android
1524	1449819480117465088	@calabresadani pelo amor de Deus quero morar n...	gabriela_srs	disney+	disney+	Twitter for iPhone
2759	1450992490807537664	@ccaarrvaa Ta passando amiga. no globoplay	twitamandy	globoplay	globoplay	Twitter for iPhone
1588	1449598443538763776	@Biaaa_comenta Pior q não, está palestrando co...	porraajess	disney+	disney+	Twitter for Android

## 3. Análise de sentimentos

Análise de sentimentos é uma subárea da ciência da computação, especificamente dentro da área de processamento de linguagem natural, cujo objetivo é extrair informações a partir da interpretação do texto da mensagem.

Um estudo abrangente (MEJOVA, 2009) analisou diversas técnicas de extração de sentimento e descreveu as diferentes formas de avaliação, indo da avaliação de polarização dos textos (positivo e negativo) a avaliações de emoções (raiva, desejo, orgulho etc.).

Em comum, todas as técnicas partem de análise de textos, tratamento e extração de características de palavras e/ou frases. Parte-se de textos conhecidos, com o sentimento avaliado a priori e, então, a partir de modelos matemáticos aplicados sobre o conjunto de características extraídas, infere-se o resultado esperado. Como todo modelo matemático, há imprecisões na avaliação, porém, por ser um modelo computacional, pode ser aplicado em grande escala.

O uso em grande escala da avaliação de sentimento permite o surgimento de vários estudos, entre eles:

- Avaliação de consumidores quanto à qualidade de produtos e serviços.
- Avaliação de tendências de mercado.
- Monitoramento de marcas.
- Avaliação de discurso de ódio em redes sociais.
- Auxílio em conversas de atendimento ao consumidor.

### 3.1. IBM Watson

Atualmente, está muito fácil utilizar serviços computacionais que outrora estavam disponíveis apenas para grandes empresas e governos que podiam arcar com o salário de pesquisadores e profissionais especializados.

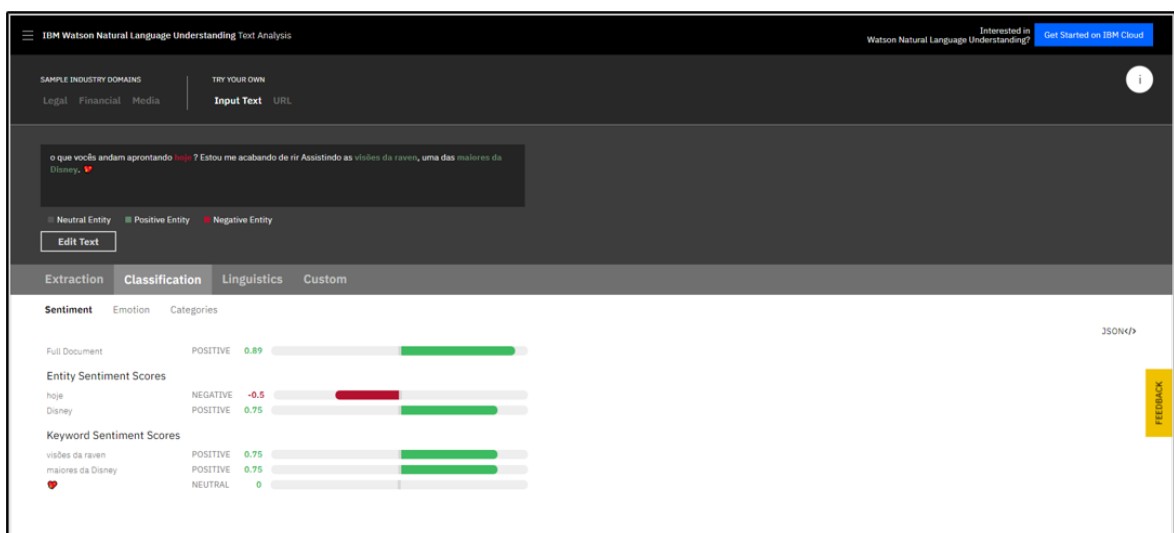
Para fazer análise de sentimento em cima dos tweets capturados, conforme apresentado, utilizaremos o serviço Watson da IBM Cloud.

O IBM Watson começou como um projeto de pesquisa da IBM e seu primeiro desafio foi disputar um concurso televisivo nos Estados Unidos, chamado Jeopardy! (MARKOFF, 2011). Um jogo de perguntas e respostas semelhante ao Show do Milhão, popular no Brasil.

O objetivo da IBM era construir um supercomputador de processamento de linguagem natural avançado, considerando:

- recuperação de informações;
- representação de conhecimento;
- raciocínio automatizado; e
- machine learning

**Figura 7 – Interface gráfica do IBM Cloud Watson**



Todos os serviços podem ser consumidos a partir de uma chamada às APIs do Watson. Um cadastro no site da empresa permite, inclusive, o uso de duas mil frases gratuitamente, todo mês.

Em Python, o código para chamar a API pode ser:



```

1. # IBM Watson
2. from ibm_watson import NaturalLanguageUnderstandingV1
3. from ibm_watson import LanguageTranslatorV3
4. from ibm_cloud_sdk_core.authenticators import IAMAuthenticator
5. from ibm_watson.natural_language_understanding_v1 import Features, CategoriesOptions, SentimentOptions
6.
7. #Puxa as credenciais de variáveis de ambiente (nunca salve credencias em seu código)
8. NATURAL_LANGUAGE_UNDERSTANDING_APIKEY = os.environ['IBM_APIKEY']
9. NATURAL_LANGUAGE_UNDERSTANDING_URL = os.environ['NLU_URL']
10.
11. #Chama o serviço da IBM Watson
12. response = natural_language_understanding.analyze(language='pt',
13.     text='inclusive muito feliz que entrou pro catálogo da amazon prime',
14.     features=Features(sentiment=SentimentOptions())).get_result()
15.
16. #Imprime o resultado
17. print(json.dumps(response, indent=2))
18.
19. """
20. Retorno do IBM Watson para a frase
21. 'inclusive muito feliz que entrou pro catálogo da amazon prime'
22. {
23.   "usage": {
24.     "text_units": 1,
25.     "text_characters": 61,
26.     "features": 1
27.   },
28.   "sentiment": {
29.     "document": {
30.       "score": 0.862283,
31.       "label": "positive" }
32.     },
33.   "language": "pt"
34. }
35.
36. """

```

Foi aplicado esse código para os quase 10 mil tweets que coletamos. Para criar uma gama maior de sentimentos, eles foram separados em sete valores, a partir do score obtido, variando sempre de 0,25 pontos entre cada escala. A primeira análise pode ser vista na Tabela 1.<sup>1</sup>

**Tabela 1 – Resultado da análise de sentimento, usando o IBM Watson**

Termo de busca	super negativo	negativo	levemente negativo	neutro	levemente positivo	positivo	super positivo	positivo – negativo
disney+	22%	12%	6%	24%	6%	13%	17%	-3%
globoplay	23%	14%	6%	18%	6%	12%	21%	-4%
hbomax	13%	10%	5%	32%	6%	15%	18%	10%
netflix	20%	13%	7%	18%	7%	13%	22%	3%
primevideo	9%	8%	5%	30%	6%	15%	27%	27%
Total Geral	20%	12%	6%	21%	7%	13%	20%	1%

<sup>1</sup> Em uma análise de sentimentos, é importante verificar possíveis erros de avaliação e se todos os tweets são realmente relacionados a serviços de *streaming*.

## 3.2. Naive Bayes

Como visto anteriormente, hoje, com o uso de uma API dos grandes players, como IBM, Google (GCP), Amazon (AWS), Microsoft (Azure), ou ainda startups especializadas, conseguimos utilizar serviços de machine learning e, com pouco esforço, integrá-los em nossos projetos.

Porém, o que está por trás dessas ferramentas de análise de sentimento? Existem diversos modelos matemáticos que podem ser utilizados para identificar o sentimento. Todos, porém, usam o mesmo princípio praticamente.

Assim como nas trilhas anteriores, prepara-se o texto para análise. A partir do texto preparado, inicia-se a extração de características, como o simples saco de palavras ou mesmo o TF-IDF.<sup>2</sup> Pode-se utilizar unigramas e n-gramas. Em avaliação de sentimento, também é muito útil manter alguns caracteres especiais, como os emojis, os quais, como visto anteriormente, são caracteres como outros quaisquer em Unicode.

A partir dessas características e conhecendo uma boa amostragem de textos positivos e negativos, um modelo matemático procura a relação com a presença (ou ausência) das palavras e o resultado positivo ou negativo.

Um dos modelos mais simples, porém com boa acurácia em classificação de textos, é o *Naive Bayes*.

### 3.2.1 Teoria

Naive Bayes é uma técnica de classificação baseada na teoria da probabilidade estudada e nomeada em homenagem ao pastor e matemático britânico Thomas Bayes.

Pelo teorema de Bayes, uma situação com probabilidades conhecidas a priori terá sua probabilidade alterada pelo conhecimento prévio de eventos ocorridos e relacionados à situação em questão, como mostrado na equação abaixo:

#### Equação 1 – Equação do teorema de Bayes

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

<sup>2</sup> Outro modelo bastante utilizado em processamento de linguagem natural é o Word2Vec, criado por cientistas do Google e patentado em 2013.

Por hipótese, ao avaliar todos os tweets de nossa amostra revisada na Tabela 2, podemos afirmar, a priori e sem ler uma palavra do tweet, que há maior chance de ele ser negativo. Mais precisamente, na proporção de 1,68 vezes mais fácil de ele possuir sentimento negativo do que positivo. Simplesmente porque nosso histórico nos conta isso.

Entretanto, se, ao ler o tweet, encontrarmos a palavra **amo** e conhecendo essa informação, poderemos recalcular a probabilidade. Dentro da nossa amostra de 174 tweets, quantos tinham a palavra **amo**? Destes, quantos eram negativos, quantos eram positivos?

E se tivermos também as palavras **amo** e **muito**? Uma das formas de se evitar calcular todas as combinações de palavras existentes para a avaliação é utilizar o Teorema de Bayes e, a partir da probabilidade individual dos eventos **amo** e **muito**, reestimar a probabilidade, conhecida a priori, de ser negativo.

Ela é considerada inocente por assumir a premissa de que todos os eventos – em nossa avaliação, a aparição de palavras diferentes nos documentos – são independentes. Em outras palavras, assume-se que a chance de a palavra **muito** aparecer independe da palavra **amo** já ter aparecido no documento. Porém, ignorando-se o valor absoluto da probabilidade, mas comparando os valores resultantes da nova probabilidade  $P(\text{negativo} \mid \text{amo}, \text{muito})$  e  $P(\text{positivo} \mid \text{amo}, \text{muito})$ , podemos inferir se o tweet continua com maior probabilidade de ser negativo, ou agora, com o conhecimento dessas duas palavras, tem mais chance de ser positivo.

### 3.2.2 Código

Felizmente, criar um algoritmo utilizando a metodologia Naïve Bayes é bastante simples em Python, e boa parte do cálculo associado fica embutido nas funções das bibliotecas.

Para avaliar a eficácia do modelo, foram selecionados 200 tweets aleatoriamente para a revisão da precisão do IBM Watson. Esses 200 tweets foram avaliados individualmente pelo professor e reclassificados caso fosse necessário.

A avaliação pode ser vista na Tabela 2. Com base nos números, calcula-se que a precisão e cobertura estão em 74%.

Tabela 2 – Avaliação da precisão do IBM Watson

		Watson			
		negative	neutral	positive	
Revisão	negative	92	0	17	109
	neutral	8	0	18	26
	positive	10	0	55	65
		110	0	90	200

A partir da revisão, e para ter um modelo mais simples, descartou-se 26 tweets classificados como neutros e criou-se um modelo com os 174 restantes. Foram separados, aleatoriamente, 150 para treinamento, deixando 24 para auferir a precisão, que ficou em cerca de 70%.

```

1.  #Importa as bibliotecas
2.  ...
3.
4.  #Carrega as amostras de tweet para treinamento e remove as neutras
5.  df_nbr = pd.read_csv('amostras-revisado.csv', delimiter=',')
6.  df_nbr = df_nbr.loc[df_nbr['sentiment_revised'].ne(0)]
7.
8.  #Faz o pré-tratamento
9.  ...
10.
11. #Separa as palavras dos textos
12. df_nbr['words'] = df_nbr['text_treated'].apply(lambda x: word_tokenize(x))
13.
14. #Faz a contagem de palavras
15. all_words = nltk.FreqDist([word.lower() for words in df_nbr['words'].to_list() for word in words])
16.
17. #Transforma em lista, para usar a biblioteca NLTK
18. words = df_nbr['words'].to_list()
19. sentiments = df_nbr['sentiment_label_revised'].to_list()
20.
21. #Prepara os documentos, no formato do NLTK (tuplas)
22. documents = [(words[i], sentiments[i]) for i, sentiment in enumerate(sentiments)]
23.
24. #Extrai as características do texto para treinamento do modelo
25. def find_features(document):
26.     words = set(document)
27.     features = {}
28.     for w in word_features:
29.         features[w] = (w in words)
30.     return features
31. featuresets = [(find_features(rev), category) for (rev, category) in documents]
32.
33. #Separa 150 registros para treinamento
34. training_set = featuresets[:150]
35. classifier = nltk.NaiveBayesClassifier.train(training_set)
36.
37. #Testa o modelo nos registros que havíamos separado para testes
38. testing_set = featuresets[150:]
39. print("Acurácia:", (nltk.classify.accuracy(classifier, testing_set))*100)

```



## 4. Síntese

Esta trilha talvez seja a que mais trouxe conteúdo novo até o momento. Começamos falando de uma nova forma de recuperação de informação, utilizando APIs.

Estudamos também o assunto das redes sociais, destacando o crescimento delas nos últimos anos, assim como sua presença massiva na vida das pessoas; inclusive criando uma espécie de metaverso primitivo.

As redes sociais são locais de geração de informação não estruturada por milhares de pessoas, descentralizada. As mensagens e relações existentes em uma rede social são, em alguma medida, um recorte da nossa sociedade. Avaliar mensagens trocadas permite entender como esse recorte funciona.

Ainda nesta trilha, aprendemos a utilizar serviços como o IBM Watson que facilita muito na aplicação de técnicas mais avançadas de processamento de linguagem natural, com custos baixos.

Por fim, trabalhamos com as técnicas de classificação (Inocente) de Bayes, uma das mais utilizadas em classificação de textos, com o intuito de aumentar o nosso repertório e permitir que nossa compreensão do que está por de trás desses serviços cada vez mais acessíveis e populares, com forte impacto em nossas vidas.

## 5. Referências

- ADALIAN, J. The Hottest Streamer (Right Now) Who is winning the platform wars? *Vulture*, 21 jun. 2021. Disponível em: <<https://www.vulture.com/2021/06/who-is-winning-the-streaming-platform-wars.html>>. Acesso em: 3 nov. 2021.
- ANDRADE, V. HBO Max, Netflix, Globoplay e mais: Quanto custa assinar serviços de streaming? *Notícias da TV*, 29 jun. 2021. Disponível em: <<https://noticiasdatv.uol.com.br/noticia/mercado/hbo-max-netflix-globoplay-e-mais-quanto-custa-assinar-servicos-de-streaming-60383>>. Acesso em: 3 nov. 2021.
- GONÇALVES, A. L. TSE solicita dados de disseminadores de fake news ao Twitter. *Tecmundo*, 26 out. 2021. Disponível em: <<https://www.tecmundo.com.br/redes-sociais/227578-tse-solicita-dados-disseminadores-fake-news-twitter.htm>>. Acesso em: 11 nov. 2021.
- GOODMAN, M. et al. Primate evolution at the DNA level and a classification of hominoids. *Journal of Molecular Evolution*, p. 260-266, mar. 1990.
- KATZ, B. State of the Streaming Wars in 2021: It's Netflix vs HBO Max vs Everyone Else. *Observer*, 16 set. 2021. Disponível em: <<https://observer.com/2021/09/amazon-apple-netflix-disney-hbo-max-streaming-wars/>>. Acesso em: 3 nov. 2021.
- KAWAMOTO, D. TheGlobe.com's IPO one for the books. *CNET.com*, 13 nov. 1998. Disponível em: <<https://archive.md/20130119215323/http://news.com.com/2100-1023-217913.html#selection-1017.1-1017.37>>. Acesso em: 11 nov. 2021.
- MARKOFF, J. Computer Wins on 'Jeopardy!': Trivial, It's Not. *The New York Times*, 16 fev. 2011. Disponível em: <<https://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html>>. Acesso em: 11 nov. 2021.
- MEJOVA, Y. *Sentiment analysis*: an overview. University of Iowa, Computer Science Department, 2009.
- MENDES, T. Serviço de streaming Oldflix chega ao mercado para fugar os cinéfilos fãs de clássicos. *Adoro Cinema*, 31 mar. 2016. Disponível em: <<https://www.adorocinema.com/noticias/filmes/noticia-120386/>>. Acesso em: 3 nov. 2021.
- O'REILLY, T. *What is Web 2.0?* O'Reilly Media, Inc, 2009.
- TWITTER. *Q2 2021 – Letter to Shareholders*. 22 jul. 2021. Disponível em: <[https://s22.q4cdn.com/826641620/files/doc\\_financials/2021/q2/Q2'21-Shareholder-Letter.pdf](https://s22.q4cdn.com/826641620/files/doc_financials/2021/q2/Q2'21-Shareholder-Letter.pdf)>. Acesso em: 11 nov. 2021.

