



Mariana Gonzaga dos Santos

Aplicação do algoritmo de aprendizado de máquina Naive Bayes para classificação de comentários racistas do YouTube

Formosa

2022

Mariana Gonzaga dos Santos

Aplicação do algoritmo de aprendizado de máquina Naive Bayes para classificação de comentários racistas do YouTube

Monografia submetida ao curso de graduação em Tecnologia em Análise e Desenvolvimento de Sistemas do Instituto Federal de Educação, Ciência e Tecnologia de Goiás, câmpus Formosa, como requisito parcial para obtenção do Título de Tecnólogo em Análise e Desenvolvimento de Sistemas.

Instituto Federal de Educação, Ciência e Tecnologia de Goiás – IFG
Graduação em Tecnologia em Análise e Desenvolvimento de Sistemas

Orientador: M.e. Gleidson Caetano da Silveira Pinto

Formosa

2022

Mariana Gonzaga dos Santos

Aplicação do algoritmo de aprendizado de máquina Naive Bayes para classificação de comentários racistas do YouTube/ Mariana Gonzaga dos Santos. – Formosa, 2022-41p. : il. (algumas color.) ; 30 cm.

Orientador: M.e. Gleidson Caetano da Silveira Pinto

T.C.C. (Graduação) – Instituto Federal de Educação, Ciência e Tecnologia de Goiás – IFG Graduação em Tecnologia em Análise e Desenvolvimento de Sistemas, 2022.

1. Racismo. 2. Naive Bayes. 3. Classificação de texto. 4. Aprendizado de máquina. I. Gleidson Caetano da Silveira Pinto. II. Instituto Federal de Educação, Ciência e Tecnologia de Goiás. III. Graduação em Tecnologia em Análise e Desenvolvimento de Sistemas. IV. Aplicação do algoritmo de aprendizado de máquina Naive Bayes para identificação e classificação de comentários racistas do YouTube.

Mariana Gonzaga dos Santos

Aplicação do algoritmo de aprendizado de máquina Naive Bayes para classificação de comentários racistas do YouTube

Monografia submetida ao curso de graduação em Tecnologia em Análise e Desenvolvimento de Sistemas do Instituto Federal de Educação, Ciência e Tecnologia de Goiás, câmpus Formosa, como requisito parcial para obtenção do Título de Tecnólogo em Análise e Desenvolvimento de Sistemas.

Trabalho aprovado. Formosa, 03 de fevereiro de 2022:

**M.e. Gleidson Caetano da Silveira
Pinto**
Orientador

Professora
Uyara Ferreira da Silva

Professor
Rangel Rigo

Formosa
2022

*“Nós só podemos ver um pouco do futuro, mas o
suficiente para perceber que há muito a fazer.”
(Alan Mathison Turing)*

Resumo

A utilização da mineração de dados, classificação textual e aprendizagem de máquina está em ascensão em várias áreas da tecnologia, assim, torna-se relevante os estudos acerca do tema. Nesta pesquisa está presente a aplicação e análise de desempenho do algoritmo Naive Bayes na identificação e classificação de comentários racistas retirados de vídeos do YouTube. O trabalho foi realizado em quatro etapas principais, coleta de dados (comentários do YouTube), tratamento dos dados, classificação e análise dos resultados gerados. Observou-se a partir dos resultados que o algoritmo Naive Bayes alcançou uma acurácia superior à 90%. Portanto, a partir dos resultados gerados foi possível identificar que o algoritmo apresentou uma capacidade satisfatória na classificação de 3.101 comentários em duas classes, racista ou normal.

Palavras-chave: Racismo. Naive Bayes. Classificação de texto.

Abstract

The use of data mining, textual classification and machine learning is on the rise in several areas of technology, thus, studies on the subject become relevant. This research presents the application and performance analysis of the Naive Bayes algorithm in the identification and classification of racist comments taken from YouTube videos. The work was carried out in four main stages, data collection (YouTube comments), data processing, classification and analysis of the results generated. It was observed from the results that the Naive Bayes algorithm reached an accuracy higher than 90%. Therefore, from the results generated, it was possible to identify that the algorithm presented a satisfactory ability to classify 3,101 comments into two classes, racist or normal.

Keywords: Racism. Naive Bayes. Text Classification.

Lista de ilustrações

Figura 1 – Diagrama de Árvore - Fonte: Elaborada pela Autora	20
Figura 2 – Algoritmo Naive Bayes (modelo multinomial): Treinamento e Teste [Manning, Raghavan e Schütze (2008)]	22
Figura 3 – Etapas realizadas para a classificação dos comentários. Fonte: Elaborada pela Autora	26
Figura 4 – Script utilizado para extrair os comentários. Fonte: Elaborada pela Autora	30
Figura 5 – Comentários devidamente armazenados. Fonte: Elaborada pela Autora .	30
Figura 6 – Códigos utilizados para a mineração dos dados. Fonte: Elaborada pela Autora	31
Figura 7 – Comentários antes e depois das etapas de tratamento. Fonte: Elaborada pela Autora	31
Figura 8 – Implementação: treinando o modelo e avaliando a classificação. Fonte: Elaborada pela Autora	31
Figura 9 – Nuvem de Palavras Geral Fonte: Elaborada pela Autora	32
Figura 10 – Nuvem de palavras dos comentários considerados normais Fonte: Elabo- rada pela Autora	32
Figura 11 – Nuvem de Palavras dos comentários considerados racistas. Fonte: Ela- borada pela Autora	33

Sumário

1	INTRODUÇÃO	10
1.1	Objetivos Gerais	13
1.2	Objetivos Específicos	13
2	RACISMO E CLASSIFICAÇÃO DE TEXTO	15
2.1	Racismo	15
2.1.1	Racismo e Redes Sociais	16
2.1.2	Crimes de Racismo na Internet	16
2.2	Youtube	17
2.3	Classificação de Texto	17
2.4	Naive Bayes	18
2.4.1	Teorema de Bayes	18
2.4.2	Algoritmo Naive Bayes	21
3	COLETA, MINERAÇÃO E CLASSIFICAÇÃO TEXTUAL	24
3.1	Coleta de dados	24
3.2	Tratamento de dados	24
3.2.1	Tokenização	25
3.2.2	Filtragem	25
3.2.3	Lematização e Stemização	25
3.3	Classificação dos Comentários	26
3.4	Forma de Representar Resultados	26
4	DESEMPENHO DO ALGORITMO	29
4.1	Implementação	29
4.2	Nuvem de palavras	31
4.3	Matriz de confusão	33
4.3.1	Acurácia	34
4.3.2	Precisão	34
4.3.3	Sensibilidade (recall)	34
5	CONSIDERAÇÕES FINAIS	36
5.1	Trabalhos Futuros	36
	REFERÊNCIAS	38

Introdução

CAPÍTULO 1

Introdução

As questões de desigualdade racial tem sido fonte de pesquisa em diversos trabalhos isso devido a grande necessidade de combater as várias formas de discriminações presentes na sociedade. Segundo dados do [IBGE \(2019\)](#), os brancos formavam em 2018, 43,1% da população, os pretos, 9,3% e os pardos, 46,5% ou seja, a maior parte da população brasileira. Além disso, os dados também apontam que as raças preta e parda ocupam na força de trabalho a maior parte dos postos, sendo 25,2% a mais que as pessoas de raça branca, porém também sofrem mais com a subutilização relacionando ao nível educacional e de instrução. Os dados apontam ainda desigualdades relacionadas a salário, violência sofrida, desemprego, analfabetismo, falta de representação política, dentre outros. Portanto, torna-se importante que pesquisas relacionadas a este tema estejam cada vez mais emergentes na sociedade.

A classificação de texto é uma área extremamente relevante da aprendizagem de máquina, através dela é possível realizar diversos estudos e análises, como análise de recomendação, auxílio na tomada de decisão, estudo de previsões, dentre outras. Existem vários algoritmos para este fim, porém nesta pesquisa foi escolhido o algoritmo de Naive Bayes, devido a capacidade de classificação apresentada pelo algoritmo em trabalhos semelhantes, as características de implementação e a capacidade de treinamento do modelo através da aprendizagem supervisionada. Realizou-se a classificação de 3.101 comentários extraídos de vídeos relacionados as questões raciais do Youtube, em duas classes, racista e normal, tendo em vista a problemática do racismo, os crimes de racismo em redes sociais e a ideia de invisibilidade proporcionada pela internet.

Para que a classificação seja realizada em um algoritmo de aprendizagem supervisionada como o Naive Bayes, é necessário um tratamento prévio dos dados textuais. Esse pré-processamento possui quatro etapas importantes sendo tokenização, filtragem, lematização e stemização. Uma vez realizadas estas etapas é possível separar os comentários em dados de treinamento e de teste. Os dados de treinamento devem estar devidamente rotulados em classes para criar o modelo classificatório. Este modelo treinado recebe os

dados de teste e realiza a classificação, gerando resultados que podem ser demonstrados em uma matriz de confusão.

A pesquisa se justifica, portanto, com base no cenário atual de desigualdades raciais e discriminação presente em todos os lugares inclusive nas redes sociais. Na grande capacidade que o aprendizado de máquina e a classificação textual possuem para auxiliar nas questões sociais.

Este trabalho está dividido em quatro partes, sendo a primeira a introdução, contendo os objetivos gerais e específicos, a segunda o referencial teórico, que demonstra bibliograficamente a importância de discutir o racismo, a classificação de texto e exemplificando como o teorema de Bayes, base matemática do algoritmo utiliza a probabilidade condicional para a predição de resultados, a terceira parte contém a metodologia é demonstrado os passos realizados na etapa de coleta de dados, tratamento textual e a classificação, na quarta e última parte estão expostos os resultados, e considerações finais da pesquisa.

Objetivos

1.1 Objetivos Gerais

- Aplicar e analisar o desempenho do algoritmo Naive Bayes na identificação e classificação de comentários racistas retirados de vídeos do YouTube.

1.2 Objetivos Específicos

- Explorar conceitos de aprendizagem de máquina, probabilidade condicional e teorema de Naive Bayes.
- Usar o YouTube Data API para realizar a coleta de dados e formar o Corpus textual.
- Utilizar o algoritmo de Naive Bayes para a classificação dos comentários.
- Demonstrar os resultados gerados pelo algoritmo.

Referencial Teórico

CAPÍTULO 2

Racismo e Classificação de Texto

Esse capítulo apresenta o referencial teórico da pesquisa está sub-dividido em seis seções sendo Racismo, Racismo e Redes Sociais, Crimes de Racismo na Internet, YouTube, Classificação de Texto e Objetivos.

2.1 Racismo

O racismo segundo [Almeida \(2019\)](#) é uma forma de discriminação baseada na raça, tem como motivação principal a sobreposição de uma raça sobre outra considerada inferior. Um dos motivos observados para a perpetuação do racismo até a atualidade se baseia no interesse de manter privilégios sobre a classe considerada minoritária. Já para [Kon \(2020\)](#) o racismo vai muito além de ato discriminatório, é um processo de hierarquização e exclusão contra uma pessoa ou uma categoria social que é considerada como diferente com base em uma característica física externa.

(..)o racismo é uma forma sistemática de discriminação que tem a raça como fundamento, e que se manifesta por meio de práticas conscientes ou inconscientes que culminam em desvantagens ou privilégios para indivíduos, a depender do grupo racial ao qual pertençam.([ALMEIDA, 2019](#), p. 22)

Muito embora casos de discriminação de raça sejam expostos periodicamente, a miscigenação presente no Brasil, causa na sociedade uma ideia de democracia racial que pode ser considerada uma grande ilusão.

Hoje, diferentemente de tempos mais distantes, a maioria dos estudiosos já não trata a democracia racial como um fato. Muitos entendem a democracia racial como uma ideia, um ideário tipicamente brasileiro, ou ainda como um mito social que, de acordo com alguns, não deve ser descartado porque carregaria a potencialidade de unir as pessoas e de mobilizá-las em torno de um ideal nobre. ([HOFBAUER, 2007](#), p. 2)

Para [Bernardino \(2002\)](#) a democracia racial ser tida como uma realidade e exibida ao mundo, causou uma sensação de inveja aos outros países que acreditaram que o paraíso

racial era real e se encontrava no Brasil. Porém, o paraíso era seletivo, não o sendo aplicável ao negro que tinha que enfrentar as dificuldades da exclusão social diariamente, a aceitação desse mito gerou algumas consequências, como a negação da existência de raças no país, a classificação baseada apenas na cor da pele e a negação acerca do tema porque conforme a democracia racial o conceito de raça é inexistente, então falar sobre discriminações raciais e políticas sociais para negros não faz sentido, e pode inclusive ser considerada uma atitude racista.

2.1.1 Racismo e Redes Sociais

Segundo [Garton, Haythornthwaite e Wellman \(1997\)](#), uma rede social é uma conexão de organizações e pessoas em uma rede de computadores. As redes sociais funcionam como reflexo do imaginário social, considerando que existe através delas a comunicação e expressão de pensamentos individuais e coletivos. Portanto analisar padrões comportamentais presentes em redes sociais, é de certa forma analisar a própria sociedade.

Os conflitos oriundos das relações étnico-raciais são travados de diversas maneiras, em diferentes contextos e espaços. E as redes sociais não estão fora desta lógica, pois é visível a postagem de comentários, imagens e vídeos que constataam a eminência do conflito racial nas redes sociais. ([MOURA, 2015](#), p. 8),

Sendo a ideia de superioridade de raça enraizada e o racismo presente na sociedade, é evidente que as conexões sociais através da internet tornam-se também palco para disseminação do preconceito racial. De acordo com [Farrington et al. \(2017\)](#), as novas formas de usar a internet, em plataformas, como o Facebook, Twitter sites de conteúdo gerado por usuários, e o YouTube. levaram a um aumento significativo de intimidação e ira online. As redes sociais são usadas para divulgar conteúdo extremista e de repulsa, aumentando a profundidade e o alcance deste discurso de ódio que confronta usuários.

2.1.2 Crimes de Racismo na Internet

A socialização por intermédio da internet gera uma falsa ideia de invisibilidade e impunidade, portanto encontrar comentários de pessoas disseminando o preconceito racial na internet e infringindo as legislações antirracistas é uma situação frequente. Em âmbito nacional, segundo [Calixto \(2010\)](#) “A Lei nº 7.716 de 1989, que define os crimes resultantes de preconceito, inicia-se relacionando dois gêneros de conduta discriminação e preconceito e cinco objetos sobre os quais recaem essas condutas raça, cor, etnia, religião e procedência nacional”. Portanto, qualquer ação discriminatória ou de preconceito, independente do ambiente em que ocorra, virtual ou não, é considerado crime. Outro conceito importante

é o de cibercrime que é, segundo [Melo \(2010\)](#), considerado qualquer ação prevista como crime ou ilegal que ocorra em meios virtuais, tecnológicos.

2.2 Youtube

O Youtube é uma rede social de compartilhamento de vídeos, “Fundado por Chad Hurley, Steve Chen e Jawed Karim, ex-funcionários do site de comércio on-line PayPal, o site YouTube foi lançado oficialmente sem muito alarde em junho de 2005.” [Burgess e Green \(2009\)](#). Atualmente, segundo o site oficial da plataforma, milhares de pessoas em todo o mundo acessam o YouTube diariamente. A infinidade de conteúdo presente nesta plataforma garante que pessoas totalmente diferentes se identifiquem e assistam os vídeos.

Os autores contextualizam o YouTube na política de cultura popular participativa, apontando questões importantes de como e por quê esse site é considerado o maior aglutinador de mídia de massa da internet no início do século 21. Ressaltam ainda que o site é uma cocriação de diferentes atores que, pela própria natureza da internet e da ferramenta, se confundem e entram em choque de interesses. ([BURGESS; GREEN, 2009](#), p. 9)

De acordo com [Silva et al. \(2019\)](#), os conflitos sociais presentes no Youtube e a diversidade de pensamento geraram a necessidade da criação de políticas com foco em discursos de ódio, onde não é permitido disseminar a violência e/ou ódio contra grupos raciais, classes sociais, religião, entre outros, essa política traz uma distinção entre liberdade de expressão de pensamento e opiniões permitidos pelo Youtube, e os discursos de ódio que possam ofender outrem.

Segundo [Kavitha et al. \(2020\)](#), existem vários trabalhos e pesquisas, que utilizam os comentários do YouTube como fonte de dados para mineração textual, análise de sentimentos, aplicação de algoritmos para classificação de spam, entre outros. Para [Malik e Tian \(2017\)](#), a enorme quantidade de vídeos, comentários e opiniões de usuários presentes no YouTube tornou a plataforma uma grande geradora de dados importantes que podem ser usados em diversos campos de pesquisa.

2.3 Classificação de Texto

Um conceito importante para compreender a mineração e classificação de texto é o conceito de Machine Learning ou aprendizado de máquina, para [Monard e Baranauskas \(2003\)](#) o aprendizado de Máquina “é uma área de IA cujo objetivo é o desenvolvimento de técnicas computacionais sobre o aprendizado bem como a construção de sistemas capazes de adquirir conhecimento de forma automática.”. Portanto o aprendizado de máquina é a

capacidade de uma tecnologia computacional de “aprender”, adquirindo conhecimento de forma semelhante a um ser humano.

A classificação de texto, segundo [Lucca et al. \(2013\)](#), tem como objetivo a separação de dados textuais em classes pré-definidas, também chamadas de rótulos. Para a classificação utilizando um algoritmo de aprendizagem supervisionada, existe um processo de treinamento prévio onde dados já rotulados “ensinam” ao algoritmo o padrão a ser seguido.

As categorizações de texto, ou as atribuições de textos de linguagem natural para categorias predefinidas com base no seu conteúdo, é de crescente importância à medida que o volume de informações na Internet continua a nos sobrecarregar. O uso de classes predefinidas implica em uma abordagem de “aprendizagem supervisionada”, onde o artigo já é classificado que efetivamente define as categorias é usado como “dados de treinamento” para construir um modelo que pode ser usado para classificar novos artigos. ([SANTOS et al., 2019](#), p. 12)

Já na classificação de texto utilizando um algoritmo de aprendizagem não supervisionada, não existe o conhecimento prévio dos dados textuais que serão classificados, o algoritmo possui apenas a informação das classes em que os dados poderão ser separados, ou seja, a partir da análise de todos os dados o algoritmo reconhece padrões de cada classe e divide os dados com base em semelhanças ou diferenças encontradas.

Neste caso de aprendizagem não existe qualquer entidade ou professor que supervisione o processo de aprendizagem. Neste caso somente os padrões de entrada estão disponíveis para a rede. A partir do momento em que a rede estabelece uma harmonia com as regularidades estatísticas da entrada de dados, desenvolve-se nela uma habilidade de formar representações internas para codificar características da entrada e criar novas classes ou grupos automaticamente. ([QUINTAS, 2016](#), p. 34)

Há também a aprendizagem semi-supervisionada onde, de acordo com [de Vries e Thierens \(2021\)](#), difere dos modelos supervisionados, que exigem a rotulação previa dos dados gerando problemas quando a quantidade de dados rotulados disponíveis é limitada e dificultando o desempenho preditivo desses modelos de classificação. Rotular um grande volume de dados pode ser um processo demorado e de alto custo. Portanto, Uma solução para este problema é o uso de métodos de aprendizagem semi-supervisionada, esta forma de aprendizagem usa os dados rotulados e não rotulados para aprenderem.

2.4 Naive Bayes

2.4.1 Teorema de Bayes

O teorema de Bayes recebe esse nome devido ao seu criador, Thomas Bayes. Segundo, [Rossatto et al. \(2016\)](#), “embora não existam muitas informações sobre a vida

do cientista sabe-se que ele foi um pastor que viveu entre os anos de 1702 a 1761, na Inglaterra.” Além de ser reverendo, Bayes se interessava pelo estudo de probabilidade, escrevia suas anotações científicas, porém, nunca chegou a publicá-las, algum tempo após a sua morte seu amigo Richard Price(1732-1791) encontrou em sua casa um artigo que contia a demonstração do Teorema de Bayes, levando-o á público.

A regra ou teorema de Bayes é baseado no estudo da probabilidade, ele define a probabilidade de um evento B ocorrer, sabendo que outro evento A já ocorreu. Portanto, para realizar a aplicação do teorema é preciso ter a informação do evento anterior. De acordo com [Mitchell et al. \(1997\)](#), o teorema pode ser explicado pela seguinte afirmação, “Um evento A, dado um evento B, não depende apenas da relação entre A e B, mas também da probabilidade de observar-se A independente de observar-se B”. A probabilidade condicional é a chance de o evento A acontecer, dado que B já aconteceu. O evento B é conhecido como condicionante, então, calculamos a chance de o evento A acontecer na condição de o evento B ter acontecido..

Para uma compreensão melhor do teorema é preciso entender alguns conceitos básicos da teoria das probabilidades mais especificamente da probabilidade condicional.

A probabilidade condicional é um dos conceitos mais importantes da Teoria da Probabilidade e está relacionado com o facto de em muitas situações em que se pretende calcular a probabilidade de um acontecimento, já se dispor de alguma informação sobre o resultado da experiência que conduz à realização do acontecimento, a qual permite atualizar a atribuição de probabilidade a esse acontecimento. ([MARTINS, 2017](#), p. 1)

Em um exemplo de aplicação prática, supondo que em uma empresa de fabricação de carros exista uma divisão entre duas equipes, onde:

- A equipe 1 é responsável por 60% da fabricação total. Destes, 30% são azuis e 70% são brancos.
- A equipe 2 é responsável pelos 40% restantes do total. Destes, 60% são azuis e 40% são brancos.

A Figura 1 representa um diagrama em árvore, para facilitar a compreensão.

1. Portanto, qual é a probabilidade de ser fabricado um carro branco?

Para responder essa questão utiliza-se a probabilidade condicional, ficando:

$$P(B) = P(E_1 \cap B) \cup P(E_2 \cap B)$$

$$P(B) = P(E_1 \cap B) + P(E_2 \cap B)$$

$$P(B) = P(E_1)P(B|E_1) + P(E_2)P(B|E_2)$$

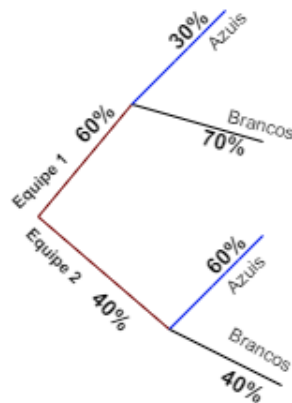


Figura 1 – Diagrama de Árvore - Fonte: Elaborada pela Autora

Onde, $P(B)$ é a probabilidade de ser um carro branco, E_1 é a equipe 1, E_2 é a equipe 2, B é um carro branco, e $P(B|E_1)$ é a probabilidade de B acontecer tendo o evento condicionante E_1 . Ou seja, sabendo que um carro branco foi produzido qual probabilidade dele ter sido feito pela equipe 1, sendo assim a probabilidade a posteriori.

É possível perceber neste caso que há dois eventos independentes entre si, ou o carro foi fabricado pela equipe 1 e é branco ou foi fabricado pela equipe 2 e é branco. Portanto, a probabilidade total de que seja fabricado um carro branco é igual à probabilidade de a equipe 1 ter fabricado o carro, multiplicada pela probabilidade de ser um carro branco sabendo que foi fabricado pela equipe 1, somada a probabilidade de a equipe 2 ter fabricado o carro, multiplicada pela probabilidade de ser um carro branco sabendo que foi fabricado pela equipe 2. Substituindo os dados e utilizando o diagrama de árvore como facilitador tem-se que:

$$P(B) = 0,60 \cdot 0,70 + 0,40 \cdot 0,40$$

$$P(B) = 0,58$$

2. Sabendo que um carro branco foi fabricado, qual a probabilidade de ele ter sido feito pela equipe 1?

Dando continuidade e entrando realmente no Teorema de Bayes, na questão 2, tem-se a informação a priori do evento, ou seja, já é apresentado que um carro branco foi produzido. Segundo [Zevarex e Santos \(2020\)](#) “O raciocínio de Thomas Bayes apresentado em seu teorema tem como ponto de partida o conhecimento a priori do evento $P(A)$ e da probabilidade condicional $P(B|A)$ para calcular a probabilidade a posteriori $P(A|B)$.” A equação do teorema apresenta a seguinte forma:

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

Adequando ao exemplo proposto tem-se:

$$P(equipe1|branco) = \frac{P(branco|equipe1)P(equipe1)}{P(branco)} = \frac{P(branco|equipe1)P(equipe1)}{P(equipe1)P(branco|equipe1)+P(equipe2)P(branco|equipe2)}$$

$$P(equipe1|branco) = \frac{0,70 \cdot 0,60}{0,60 \cdot 0,70 + 0,40 \cdot 0,40} = \frac{0,42}{0,58} = 0,72$$

2.4.2 Algoritmo Naive Bayes

Segundo, [Hubert et al. \(2021\)](#) existem vários algoritmos com o intuito de classificar dados em diferentes classes, o classificador Naive Bayes é um deles e apresenta um ótimo desempenho quando comparado a outros algoritmos, esse desempenho depende de vários fatores e um deles é a forma como os dados são pré-processados. De acordo com [Silva et al. \(2017\)](#) “o algoritmo Naive Bayes é baseado na teoria das probabilidades e supõe que os atributos vão influenciar a classe de modo independente.”

A função de classificação recebe como parâmetros, além do documento de teste, o conjunto de classes, o vocabulário e as probabilidades estimadas no treinamento. Para cada classe, a probabilidade a posteriori é calculada somando o logaritmo da probabilidade a priori com os logaritmos das probabilidades condicionais de cada termo do documento de teste. ([LUCCA et al., 2013](#), p. 2)

De acordo com [Azalia, Bijaksana e Huda \(2019\)](#), o Naive Bayes utiliza uma coleção de dados de treinamento como amostra, previamente rotulados, para estimar os parâmetros do modelo. Um dos algoritmos baseados no teorema Naive Bayes, mais usado para a classificação de texto é o Naive Bayes Multinomial, que é um dos modelos de aprendizado supervisionado de máquina que realiza o processo de classificação de textos utilizando o valor de probabilidade de uma classe em um documento.

```

TRAINMULTINOMIALNB(C,ID)
1  V ← EXTRACTVOCABULARY(ID)
2  N ← COUNTDOCS(ID)
3  for each c ∈ C
4  do Nc ← COUNTDOCSINCLASS(ID, c)
5     prior[c] ← Nc / N
6     textc ← CONCATENATETEXTOFALLDOCSINCLASS(ID, c)
7     for each t ∈ V
8     do Tct ← COUNTTOKENSOFTERM(textc, t)
9     for each t ∈ V
10    do condprob[t][c] ←  $\frac{T_{ct}+1}{\sum_{t'} (T_{ct'}+1)}$ 
11  return V, prior, condprob

APPLYMULTINOMIALNB(C, V, prior, condprob, d)
1  W ← EXTRACTTOKENSFROMDOC(V, d)
2  for each c ∈ C
3  do score[c] ← log prior[c]
4     for each t ∈ W
5     do score[c] += log condprob[t][c]
6  return arg maxc ∈ C score[c]

```

Figura 2 – Algoritmo Naive Bayes (modelo multinomial): Treinamento e Teste [Manning, Raghavan e Schütze (2008)]

Metodologia

CAPÍTULO 3

Coleta, Mineração e Classificação Textual

Este capítulo apresenta a metodologia desta pesquisa, está dividido em quatro seções, sendo na primeira seção apresentada a forma como os dados foram coletados, na segunda seção é mostrado como foi realizado o tratamento e mineração dos dados, na terceira seção é demonstrado como o algoritmo Naive Bayes foi utilizado para a classificação e na quarta é especificada as formas escolhidas para a representação dos resultados obtidos.

3.1 Coleta de dados

Para a formação da base de dados utilizada neste trabalho, a coleta dos comentários dos vídeos do YouTube, foi feita utilizando a YouTube Data API, a plataforma de scripts do Google e o Google Docs. O primeiro passo para a extração dos comentários foi criar um arquivo de extensão (.XLS ou .XLSX) no Google Docs. Após criar essa planilha foi preciso acessar o vídeo do qual os comentários seriam extraídos e copiar o ID do mesmo, segundo o manual de referência do YouTube Data API [CommentThreads](#) (2021) o ID é a propriedade única de identificação do vídeo especificado, este está localizado na parte final da URL do vídeo. Após a cópia do ID foi preciso colá-lo, na primeira coluna e primeira linha da planilha previamente criada. Para iniciar o script foi necessário acessar a aba de ferramentas da planilha e escolher a opção de “Editor de script”. O script desenvolvido teve como base a documentação do Google Scripts [Spreadsheet...](#) (2021), do YouTube Data API [CommentThreads](#) (2021), e o vídeo [Orbis](#) (2020).

3.2 Tratamento de dados

A preparação do Corpus textual é de extrema importância para a classificação, segundo [Allahyari et al. \(2017\)](#) o pré-processamento dos dados geralmente consiste em quatro etapas: Tokenization, Filtering, Lemmatization e Stemming. O processamento

dos comentários foi realizado utilizando o software livre Rstudio que é uma IDE para a linguagem de programação R.

3.2.1 Tokenização

De acordo com [Chauhan e Kumar \(2020\)](#) a tokenização é uma das etapas mais importantes do tratamento de dados, nela as cadeias de texto mais longas são divididas em pedaços menores também chamados de tokens. Depois que o texto é separado em palavras, sentenças, parágrafos, entre outros, então o algoritmo é aplicado apenas para classificar, visualizar e encontrar padrões para realizar a análise. O processo de tokenização dos comentários foi realizado separando-os e fazendo a conversão do Corpus em um vetor de caracteres.

3.2.2 Filtragem

Na etapa de filtragem algumas palavras, pontuações e emojis devem ser removidos para que o Corpus textual não tenha uma quantidade muito alta de dados inúteis para a análise, segundo [Allahyari et al. \(2017\)](#) o processo de filtragem inclui a remoção de Stop Words que são palavras que frequentemente aparecem no texto sem ter muita informação de conteúdo, por exemplo, preposições, conjunções, etc. Também é feita a remoção de pontuação, espaços em branco e o processo para que todos os termos fiquem em caixa baixa. Para esta etapa de preparação do texto foi usado o pacote de Text Mining “tm” da linguagem R com a interface “map” que possui várias funções de transformação de dados textuais, [Feinerer, Hornik e Feuerer \(2015\)](#).

3.2.3 Lematização e Stemização

No processo de Lematização, segundo [Radygin et al. \(2021\)](#), as palavras são padronizadas por exemplo, as que estão no plural são passadas para o singular, os adjetivos são convertidos para o masculino e singular, essa normalização não ocasiona perdas pelo fato de que essas palavras da forma original, geralmente não possuem informações úteis para a análise. Já na Stemização o objetivo é encontrar a palavra raiz, então é feita a retirada de sufixos e afixos. Ainda de acordo com [Radygin et al. \(2021\)](#), esses processos são importantes pelo fato de reduzir a quantidade de dados a serem processados, economizando recursos computacionais e aumentando a eficácia no processo de classificação ou mineração de texto.

3.3 Classificação dos Comentários

A figura 3, aborda como foi realizada a separação entre os comentários que foram usados como dados de treinamento e teste, e a implementação do algoritmo Naive Bayes, utilizando a linguagem R e a IDE Rstudio.

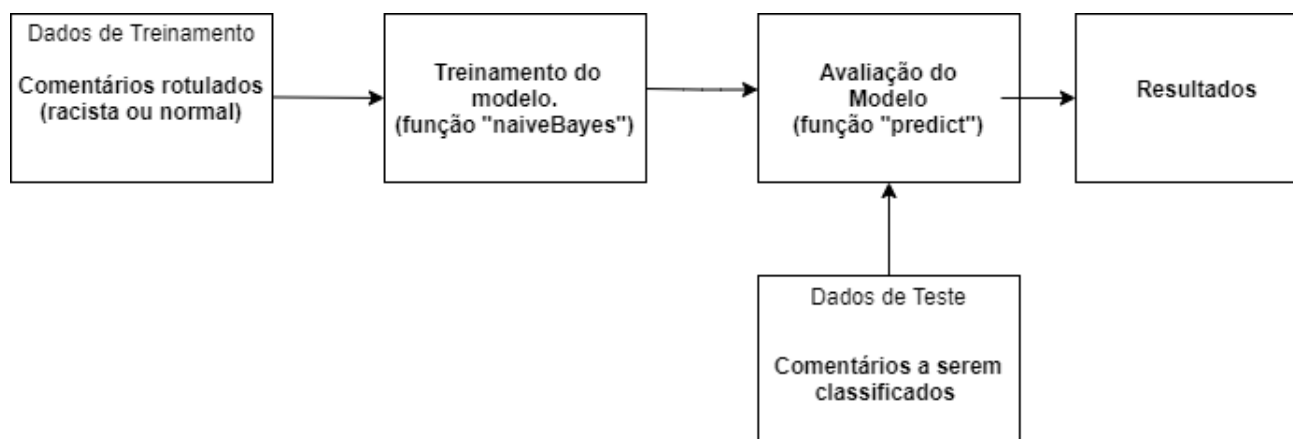


Figura 3 – Etapas realizadas para a classificação dos comentários. Fonte: Elaborada pela Autora

Segundo [Maass e Storey \(2021\)](#), o treinamento inclui a seleção, treinamento e avaliação de modelos. Treinar um modelo significa ajustar os parâmetros do modelo aos dados.

3.4 Forma de Representar Resultados

Para demonstrar os dados textuais obtidos através da classificação, o método escolhido foi a nuvem de palavras que é segundo [Ahuja \(2017\)](#), uma representação visual de dados de texto, normalmente usada para visualizar um texto em palavras únicas, a nuvem gerada fornece uma representação em imagem de um conjunto de termos com base na frequência com que a palavra aparece nos dados textuais analisados. Isso pode ser interpretado como o destaque da palavra em um domínio específico. As nuvens de palavras são, portanto, uma maneira rápida de resumir grandes quantidades de dados, é também um indicativo da filosofia associada a esse conjunto textual, essa forma de visualização pode auxiliar no processo de análise textual exploratória.

Para exibir métricas de desempenho foi usada uma matriz de confusão que de acordo com [Santra e Christy \(2012\)](#), contém informações sobre classificações reais e previsões feitas por um algoritmo de classificação. O desempenho destes algoritmos geralmente é avaliado usando os dados da matriz, avaliando a acurácia que é segundo [Balas et al. \(2019\)](#), o quão próximo o modelo chegou dos dados reais. A precisão que de acordo com

[Salmi e Rustam \(2019\)](#), mostra quão bem está o desempenho do classificador relacionado a uma classe específica, portanto, se a precisão for alta significa que a previsão é melhor. Outra métrica importante para avaliar o modelo segundo [Lever J. \(2016\)](#), é a sensibilidade, também chamada recall, esta medida é a proporção de positivos conhecidos previstos corretamente.

Resultados

CAPÍTULO 4

Desempenho do algoritmo

4.1 Implementação

Considerando que em um algoritmo de aprendizagem supervisionada, como o Naive Bayes é preciso ter dados rotulados previamente para o treinamento do modelo, a base de treinamento usada neste trabalho é composta por 1.654 comentários que foram rotulados manualmente, levando em conta a bibliografia e o conhecimento empírico, em duas classes: racista e normal. Já os dados utilizados para testar o modelo após o treinamento, consistem em 3100 comentários.

O processo de implementação é iniciado com o script desenvolvido para a coleta de dados este está demonstrado na figura 4, neste a primeira variável criada “planilha” armazena e define a informação de qual planilha está ativa. O “resultado” define que a planilha terá duas colunas “Nome” e “Comentário”, também armazena os dados que serão coletados, atribuindo estes valores a uma nova página criada posteriormente. A variável “url” refere-se a identificação do vídeo e recebe o ID que foi copiado e colado na primeira linha e coluna da página número um, na planilha. No primeiro laço de repetição a variável “dados” armazena uma lista de informações, que contém o ID do vídeo, o número máximo de comentários que serão armazenados e a informação de qual será a página seguinte. No segundo laço de repetição os dados são separados em linhas, a estrutura condicional, apresenta uma condição de parada, caso não haja mais dados o script encerra a execução. A variável “novaPlanilha” cria uma página e armazena os dados sempre que o script é executado.

Para que o script seja executado com sucesso e os comentários transferidos para a planilha, é preciso adicionar o YouTube Data API como um serviço na própria ferramenta de editor de script, e autorizar através do login em uma conta Google válida.

Os comentários foram extraídos, sem as respostas, de onze vídeos de canais diferentes onde a temática e/ou o criador do conteúdo, estivesse ligado as relações raciais, ou fosse



```

1 function extrairComentarios(){
2   var planilha = SpreadsheetApp.getActiveSpreadsheet();
3   var resultado=[['Nome', 'Comentário']];
4   var url = planilha.getSheets()[0].getRange(1,1).getValue();
5   var proximaPagina = undefined;
6
7   while(1){
8     var dados = YouTube.CommentThreads.list('snippet', {videoId: url, maxResults: 40000, pageToken: proximaPagina});
9     proximaPagina = dados.nextPageToken
10    for (var linha=0; linha< dados.items.length; linha++) {
11      resultado.push([dados.items[linha].snippet.topLevelComment.snippet.authorDisplayName,
12        dados.items[linha].snippet.topLevelComment.snippet.textDisplay]);
13    }
14    if(proximaPagina == "" || typeof proximaPagina === "undefined"){
15      break;
16    }
17  }
18  var novaPlanilha = planilha.insertSheet(planilha.getNumSheets())
19  novaPlanilha.getRange(1, 1, resultado.length, 2).setValues(resultado)
20 }
21

```

Figura 4 – Script utilizado para extrair os comentários. Fonte: Elaborada pela Autora

negro, ao todo foram extraídos 11.620 comentários. Após a análise visual dos dados, alguns foram retirados restando ao fim 4.755 comentários, estes foram armazenados em um único arquivo CSV, separados por vírgula, a planilha final é apresentada na figura 5.

	A	B	C	D	E	F	G	H
1	Nome	Comentário						
2		Me ensinaram na escola, mesmo que poucos tenham aprendido, sobre Chico Buarque, mas nunca vi uma referência sob						
3		Como essa mulher canta 💜💜👍👍						
4		Elza Soares Anjo Negro. Esta mulher Vencedora!!!						
5		Sempre uma cantora e intérprete fenomenais! Mereceu o título de cantora do milênio.						
6		Não é fácil provar que você tem competência como qualquer outra pessoa						
7		Sou professor e pedi para os alunos fazerem uma peça sobre o racismo no cotidiano e no final cantamos essa musica, el						
8		Uma triste realidade						
9		já imaginaram um som da Elza com a Erykah Badu?						
10		Nossa dura realidades... século 21 ... triste						

Figura 5 – Cometários devidamente armazenados. Fonte: Elaborada pela Autora

As etapas de mineração de dados foram realizadas através da linguagem R, utilizando os códigos presentes na figura 6. Após o tratamento textual os dados apresentam modificações identificáveis, na figura 7 é possível perceber a diferença entre os comentários antes do processo de tratamento e depois.

A implementação do classificador foi realizada utilizando as funções “naiveBayes” e “predic” do R. A função “naiveBayes” recebe como parâmetro os dados de treinamento que são os comentários já rotulados e as classes existentes: racista e normal, quando executada a função realiza o treinamento do modelo. A avaliação é dada utilizando a função “predict” que recebe o modelo previamente treinado e os dados de teste que serão classificados, com base no teorema de Bayes e na probabilidade condicional, calcula a probabilidade de

```

77 # Palavras em minúsculo
78 dados_corpus_filtrado <- tm_map(dados_corpus, content_transformer(tolower))
79 # Remove números
80 dados_corpus_filtrado <- tm_map(dados_corpus_filtrado, removeNumbers)
81 # Remove stop words
82 dados_corpus_filtrado <- tm_map(dados_corpus_filtrado, removeWords, stopwords("portuguese"))
83 # remove pontuação
84 dados_corpus_filtrado <- tm_map(dados_corpus_filtrado, removePunctuation)
85 # Elimina espaços em branco
86 dados_corpus_filtrado <- tm_map(dados_corpus_filtrado, stripWhitespace)
87

```

Figura 6 – Códigos utilizados para a mineração dos dados. Fonte: Elaborada pela Autora

```

> # Antes da limpeza
> lapply(dados_corpus[1:3], as.character)
$`1`
[1] "Irmão, 1000 acertos nenhum elogio, 01 erro, 1000 condenações...infelizmente é assim"

$`2`
[1] "Esperar o que de um cara que apoiou o Bozo e casou com uma loira kkkkkkkkkkkk"

$`3`
[1] "karaka Thiago virei fã"

> # Depois da limpeza
> lapply(dados_corpus_clean[1:3], as.character)
$`1`
[1] "irmão acerto nenhum elogio erro condenaçõesinfelizment é assim"

$`2`
[1] "esperar cara apoiou bozo casou loira kkkkkkkkkkkk"

$`3`
[1] "karaka thiago virei fã"

```

Figura 7 – Comentários antes e depois das etapas de tratamento. Fonte: Elaborada pela Autora

um comentário ser ou não racista. O processo de implementação presente na figura 8, é demonstrado com detalhes por Rocha (2020).

```

179 # Treinando o modelo
180 naiveB_classificador <- naiveBayes(dadosTreino, dadosTreinoRotulos)
181
182 # Avaliando o modelo
183 testePredição <- predict(naiveB_classificador, teste)
184

```

Figura 8 – Implementação: treinando o modelo e avaliando a classificação. Fonte: Elaborada pela Autora

4.2 Nuvem de palavras

Neste projeto foi utilizado o pacote gerador de nuvem de palavras “wordcloud” da linguagem R, para gerar três nuvens de palavras de modo a exibir os termos mais frequentes nos comentários classificados como racistas, normais e também geral que inclui todos os comentários.

Na figura 10, estão representadas as principais palavras encontradas nos comentários considerados normais, é possível perceber que a maioria das palavras não possui cunho

negra cabeluda, há também o xingamento que se limita apenas à menção à própria raça ou cor, pressupondo que somente seu chamamento já carregue seu conteúdo negativo.

Estão representadas, na figura 11, as principais palavras encontradas nos comentários considerados racistas, nesta é possível perceber que algumas palavras, são insultos a aparência física, intelectual ou moral, como: ridícula, feia, horrível, etc. Também existe a presença de termos como “ruim” e “bombril”, crítica e comparação racista ao cabelo das pessoas negras.

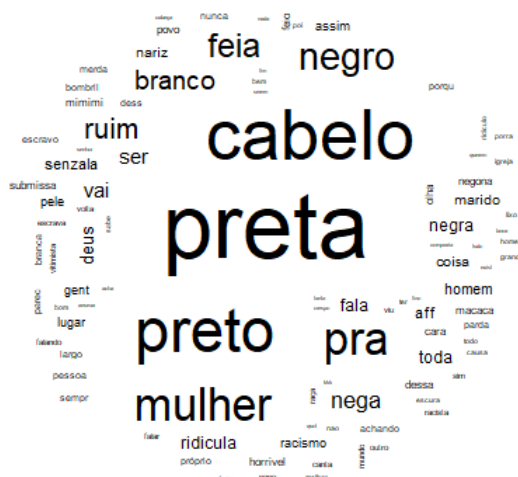


Figura 11 – Nuvem de Palavras dos comentários considerados racistas. Fonte: Elaborada pela Autora

4.3 Matriz de confusão

Neste trabalho a matriz de confusão (tabela 1), contendo os resultados do classificador Naive Bayes foi gerada utilizando a função “crossTable” do R, e editada para uma melhor representação gráfica dos dados. Sendo: TP (True Positive - Verdadeiro Positivo), TN (True Negative - Verdadeiro Negativo), FP (False Positive - Falso Positivo), FN (False Negative - Falso Negativo).

Na segunda linha da matriz estão os dados acerca da classificação dos comentários rotulados como normal, é possível perceber que o algoritmo previu 2886 comentários como normal e acertou, porém, previu que 113 eram racistas e errou. Já na terceira linha, o algoritmo previu que 63 comentários eram racistas e acertou, mas, previu que 38 eram racistas e na realidade eram comentários normais. A partir desses dados tem-se a possibilidade de analisar a acurácia, precisão e sensibilidade do modelo gerado.

Tabela 1 – Matriz de confusão: Valores gerados pelo modelo. Fonte: Elaborada pela Autora

	Valores previstos		Total Linhas
	NORMAL	RACISTA	
Normal	2886 (TP)	113 (FP)	2999
Racista	38 (FN)	63 (TN)	101
Total Colunas	2924	176	3100

4.3.1 Acurácia

A acurácia do classificador é calculada utilizando a seguinte fórmula.

$$acurácia = \frac{TP+TN}{TP+FP+TN+FN}$$

Utilizando os dados gerados pela matriz de confusão tem-se, a quantidade de vezes que o modelo acertou a classificação dividido pelo total de valores:

$$acurácia = \frac{2886+63}{2886+63+38+113} = \frac{2949}{3100} = 95\%$$

É possível definir, portanto, que o modelo alcançou uma acurácia satisfatória alcançando de cerca de 95% de acerto na classificação dos dados textuais propostos.

4.3.2 Precisão

A precisão do modelo é calculada utilizando a seguinte fórmula.

$$precisão = \frac{TP}{TP+FP}$$

Utilizando os dados gerados pela matriz de confusão do modelo desenvolvido, tem-se, a quantidade de vezes que o modelo acertou a classificação de comentários rotulados como normal, dividido por esse mesmo valor, somado ao total de vezes que o classificador errou definindo os comentários que eram normais, como racista, ficando:

$$precisão = \frac{2886}{2886+113} = 96\%$$

Portanto, o modelo atingiu também uma alta precisão alcançando cerca de 96% demonstrando assim uma grande capacidade de prevê os dados corretamente.

4.3.3 Sensibilidade (recall)

A sensibilidade do modelo classificatório é calculada através da fórmula:

$$Sensibilidade = \frac{TP}{TP+FN}$$

Utilizando os dados gerados pela matriz de confusão, tem-se que o modelo alcançou uma sensibilidade de 98%.

$$Sensibilidade = \frac{2886}{2886+38} = 98\%$$

Tabela 2 – Desempenho do classificador a partir das métricas analisadas. Fonte: Elaborada pela Autora

Desempenho Geral do Classificador %

Acurácia	Precisão	Sensibilidade
95%	96%	98%

É possível observar que os resultados gerados nesta pesquisa pelo classificador assemelham-se aos resultados de trabalhos relacionados, como no projeto de [Ururahy et al. \(2018\)](#), onde o modelo Naive Bayes teve o melhor desempenho de 75% de acurácia em uma abordagem multirótulo, classificando filmes em gêneros pré definidos, ou de [Batista, Bagatini e Frozza \(2018\)](#) onde o classificador alcançou uma acurácia maior que 90%.

CAPÍTULO 5

Considerações finais

Esta pesquisa teve como objetivo principal realizar a aplicação e analisar o desempenho do algoritmo Naive Bayes para a classificação de comentários racistas retirados de vídeos do YouTube, para que esse objetivo fosse alcançado foi necessário aprofundar conhecimentos acerca da API para desenvolvedores do YouTube, da mineração e tratamento de texto, entender o funcionamento do algoritmo Naive Bayes e de como deveria ser aplicado para alcançar os melhores resultados, além de pesquisa bibliográfica relacionadas às questões raciais no Brasil e Machine learning.

Foi preciso também realizar um tratamento minucioso no corpus textual, descrito com detalhes na metodologia, prevenindo assim o modelo de classificação da influência de dados desnecessários ou incorretos. Através deste trabalho e dos resultados obtidos foi possível entender como o algoritmo de aprendizagem de máquina supervisionada Naive Bayes realiza a classificação de texto em duas classes.

Portanto, é possível concluir que este trabalho alcançou os objetivos propostos, foram explorados conceitos de aprendizagem de máquina, probabilidade condicional, teorema de Naive Bayes e mineração de texto. A YouTube Data API serviu como ferramenta para realizar a coleta dos dados e formar o corpus textual, o algoritmo Naive Bayes foi utilizado para a classificação dos comentários e os resultados gerados foram exibidos.

5.1 Trabalhos Futuros

Considerando a abrangência dos conceitos explorados, existem várias possibilidades para trabalhos futuros como:

- Realizar um estudo comparativo com outros algoritmos de aprendizagem supervisionada para classificação textual, utilizando a base de dados formada.

-
- Explorar o comportamento de algoritmos de aprendizagem não supervisionada e semi-supervisionada, para a classificação textual.
 - Realizar novos experimentos com o classificador Naive Bayes, utilizando um número maior de comentários.
 - Explorar o Processamento de Linguagem Natural (PLN) e análise de sentimentos.

Referências

- AHUJA, M. S. V. Twitter presence of jet airways-deriving customer insights using netnography and wordclouds, *procedia computer science*. v. 122, p. 17–24, 2017. ISSN 1877-0509. Acesso: 14 de ago. de 2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050917325656>>. Citado na página 26.
- ALLAHYARI, M. et al. A brief survey of text mining: Classification, clustering and extraction techniques. *arXiv preprint arXiv:1707.02919*, 2017. Citado 2 vezes nas páginas 24 e 25.
- ALMEIDA, S. *Racismo estrutural*. [S.l.]: Pólen Produção Editorial LTDA, 2019. Citado na página 15.
- AZALIA, F. Y.; BIJAKSANA, M. A.; HUDA, A. F. Name indexing in indonesian translation of hadith using named entity recognition with naïve bayes classifier. *Procedia Computer Science*, v. 157, p. 142–149, 2019. ISSN 1877-0509. The 4th International Conference on Computer Science and Computational Intelligence (ICCSCI 2019) : Enabling Collaboration to Escalate Impact of Research Results for Society. Acesso: 07 de jul. 2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050919310695>>. Citado na página 21.
- Chapter 13 - medical big data mining and processing in e-healthcare. In: BALAS, V. E. et al. (Ed.). *Internet of Things in Biomedical Engineering*. Academic Press, 2019. p. 323–339. ISBN 978-0-12-817356-5. Acesso: 05 de ago. de 2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780128173565000164>>. Citado na página 26.
- BATISTA, R. de A.; BAGATINI, D. D.; FROZZA, R. Classificação automática de códigos ncm utilizando o algoritmo naïve bayes. *iSys-Revista Brasileira de Sistemas de Informação*, v. 11, n. 2, p. 4–29, 2018. Citado na página 35.
- BERNARDINO, J. Ação afirmativa e a rediscussão do mito da democracia racial no brasil. *Estudos afro-asiáticos*, SciELO Brasil, v. 24, p. 247–273, 2002. Citado na página 15.
- BURGESS, J.; GREEN, J. Youtube e a revolução digital. *São Paulo: Aleph*, p. 24, 2009. Citado na página 17.
- CALIXTO, C. C. Breves reflexões sobre a imprescritibilidade dos crimes de racismo. *Revista Eletrônica do Curso de Direito da UFSM*, v. 5, n. 2, 2010. Citado na página 16.
- CHAUHAN, A.; KUMAR, Y. Spam filtration project using nlp. Jaypee University of Information Technology, Solan, HP, 2020. Citado na página 25.

- COMMENTTHREADS. 2021. <<https://developers.google.com/youtube/v3/docs/commentThreads/>>. Acesso: 17 de jul. de 2021. Citado na página 24.
- de Vries, S.; THIERENS, D. A reliable ensemble based approach to semi-supervised learning. *Knowledge-Based Systems*, v. 215, p. 106738, 2021. ISSN 0950-7051. Acesso em 01 de dez. 2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0950705121000010>>. Citado na página 18.
- FARRINGTON, N. et al. *Sport, racism and social media*. [S.l.]: Routledge, 2017. Citado na página 16.
- FEINERER, I.; HORNIK, K.; FEINERER, M. I. Package ‘tm’. *Corpus*, v. 10, n. 1, 2015. Citado na página 25.
- GARTON, L.; HAYTHORNTHWAITE, C.; WELLMAN, B. Studying online social networks. *Journal of computer-mediated communication*, Oxford University Press Oxford, UK, v. 3, n. 1, p. JCMC313, 1997. Citado na página 16.
- HOFBAUER, A. Branqueamento e democracia racial: sobre as entranhas do racismo no brasil. *Por que “raça*, p. 2, 2007. Citado na página 15.
- HUBERT et al. Classifying promotion images using optical character recognition and naïve bayes classifier. *Procedia Computer Science*, v. 179, p. 498–506, 2021. ISSN 1877-0509. 5th International Conference on Computer Science and Computational Intelligence 2020. Acesso em 14 de nov. de 2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050921000387>>. Citado na página 21.
- IBGE. Desigualdades sociais por cor ou raça no brasil. p. 12, 2019. ISSN 9788524045134, note = acesso em 05 de fev. 2022, url = https://biblioteca.ibge.gov.br/visualizacao/livros/liv101681_informativo.pdf, author = IBGE. Citado na página 10.
- KAVITHA, K. et al. Analysis and classification of user comments on youtube videos. *Procedia Computer Science*, v. 177, p. 593–598, 2020. ISSN 1877-0509. Acesso em 13 de jun. de 2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050920323553>>. Citado na página 17.
- KON, N. M. *O racismo e o negro no brasil*. [S.l.]: Editora Perspectiva SA, 2020. Citado na página 15.
- LEVER J., K. M. . A. N. Classification evaluation. 2016. (Acesso: 08 de fevereiro de 2021). Citado na página 27.
- LUCCA, G. et al. Uma implementação do algoritmo naïve bayes para classificação de texto. *CENTRO DE CIÊNCIAS COMPUTACIONAIS DA UNIVERSIDADE FEDERAL DO RIO GRANDE. IX Escola Regional de Banco de Dados–ERBD*, p. 1–4, 2013. Citado 2 vezes nas páginas 18 e 21.
- MAASS, W.; STOREY, V. C. Pairing conceptual modeling with machine learning. *Data Knowledge Engineering*, v. 134, p. 101909, 2021. ISSN 0169-023X. Acesso em 01 de nov. de 2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169023X21000367>>. Citado na página 26.

- MACHADO, M. R. d. A.; LIMA, M.; NERIS, N. Racismo e insulto racial na sociedade brasileira: Dinâmicas de reconhecimento e invisibilização a partir do direito 1. *Novos estudos CEBRAP*, SciELO Brasil, v. 35, p. 11–28, 2016. Citado na página 32.
- MALIK, H.; TIAN, Z. A framework for collecting youtube meta-data. *Procedia Computer Science*, v. 113, p. 194–201, 2017. ISSN 1877-0509. Acesso: 17 de nov. de 2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S187705091731757X>>. Citado na página 17.
- MANNING, C. D.; RAGHAVAN, P.; SCHÜTZ, H. *Introduction to Information Retrieval*. Cambridge, UK: Cambridge University Press, 2008. Acesso: 02 de jul. de 2021. ISBN 978-0-521-86571-5. Disponível em: <<http://nlp.stanford.edu/IR-book/information-retrieval-book.html>>. Citado 2 vezes nas páginas 7 e 22.
- MARTINS, M. E. G. Probabilidade condicional. *Revista de Ciência Elementar*, Casa das Ciências, v. 5, n. 3, 2017. Citado na página 19.
- MELO, C. E. S. d. *Racismo e violação aos direitos humanos pela internet: estudo da Lei nº 7.716/89*. Tese (Doutorado) — Universidade de São Paulo, 2010. Citado na página 17.
- MITCHELL, T. M. et al. *Machine learning*. McGraw-hill New York, 1997. Citado na página 19.
- MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. *Sistemas inteligentes-Fundamentos e aplicações*, Manole, v. 1, n. 1, p. 32, 2003. Citado na página 17.
- MOURA, T. M. de. Racismo nas redes sociais perpetuação do imaginário social de inferiorização do negro na sociedade brasileira. *Emblemas*, v. 12, n. 2, 2015. Citado na página 16.
- ORBIS, L. *Download or Scrape all YouTube Comments and Replies*. 2020. Acesso: 14 de nov. de 2021. Disponível em: <<https://www.youtube.com/watch?v=uD58-EHwael>>. Citado na página 24.
- QUINTAS, A. S. M. S. *Aprendizagem não supervisionada de padrões de interação homem-computador*. Tese (Doutorado), 2016. Citado na página 18.
- RADYGIN, V. et al. Application of text mining technologies in russian language for solving the problems of primary financial monitoring. *Procedia Computer Science*, v. 190, p. 678–683, 2021. ISSN 1877-0509. Acesso 16 de mar. de 2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050921013314>>. Citado na página 25.
- ROCHA, E. *Machine Learning Naive Bayes - Classificação de Spam - Text Mining*. 2020. Acesso: 01 de jun. de 2021. Disponível em: <https://www.youtube.com/watch?v=LnSx4D7R2q8&t=2075s&ab_channel=C3AAnciadosDados>. Citado na página 31.
- ROSSATTO, F. C. et al. Teorema de bayes: Estudo e aplicação. *Salão do Conhecimento*, 2016. Citado na página 18.

- SALMI, N.; RUSTAM, Z. Naïve bayes classifier models for predicting the colon cancer. *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, v. 546, p. 052068, jun 2019. Acesso: 15 de abril de 2021. Disponível em: <<https://doi.org/10.1088/1757-899x/546/5/052068>>. Citado na página 27.
- SANTOS, K. B. C. d. et al. Categorização de textos por aprendizagem de máquina. Universidade Federal de Alagoas, 2019. Citado na página 18.
- SANTRA, A.; CHRISTY, C. J. Genetic algorithm and confusion matrix for document clustering. *International Journal of Computer Science Issues (IJCSI)*, Citeseer, v. 9, n. 1, p. 322, 2012. Citado na página 26.
- SILVA, C. L. da et al. Usando o classificador naive bayes para geração de alertas de risco de Óbito infantil. *Revista Electronica de Sistemas de Informacao*, v. 16, n. 2, p. 1–15, May 2017. Copyright - © 2017. This work is published under <https://creativecommons.org/licenses/by/3.0/> (the “License”). Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License; Última atualização em - 2018-11-30; SubjectsTermNotLitGenreText - United States–US; Rio de Janeiro Brazil. Citado na página 21.
- SILVA, L. R. L. et al. A gestão do discurso de ódio nas plataformas de redes sociais digitais: um comparativo entre facebook, twitter e youtube. *Revista ibero-americana de ciência da informação*, v. 12, n. 2, p. 470–492, 2019. Citado na página 17.
- SPREADSHEET Service. 2021. <<https://developers.google.com/apps-script/reference/spreadsheet>>. Acesso: 17 de jul. de 2021. Citado na página 24.
- URURAHY, J. M. K. et al. Classificação multirrótulo de documentos de texto utilizando a relevância binária e o algoritmo naïve bayes. *Revista do Seminário Internacional de Estatística com R*, v. 3, n. 2, 2018. Citado na página 35.
- ZEVARER, V. A. de O.; SANTOS, C. H. da S. Estudos das implicações do teorema de bayes na computação natural. *Revista Brasileira de Iniciação Científica*, v. 7, n. 5, p. 58–79, 2020. Citado na página 20.