

DETECÇÃO AUTOMÁTICA DE CONTEÚDOS PRECONCEITUOSOS UTILIZANDO TÉCNICAS DE CLASSIFICAÇÃO DE TEXTOS

Keterly Geovana Gouveia Silva¹, Patrick Pedreira Silva²

Resumo

Com o aumento exponencial da quantidade de informações textuais torna-se relevante investigar métodos e técnicas que permitam lidar com este conteúdo de forma eficiente e automática, dessa forma, o processamento automático de textos é um grande desafio. A classificação automática de texto envolve atribuir uma ou mais categorias de documentos predefinidas. Esta investigação se propõe a examinar as técnicas associadas à categorização automática, propondo uma ferramenta computacional que permita identificar conteúdos preconceituosos em redes sociais, a partir da análise estatística e linguística de textos coletados. O intuito deste trabalho foi investigar e desenvolver técnicas de classificação automática para realizar a tarefa de detectar discursos de ódio contra a comunidade LGBTQIA+. Foi desenvolvido um website responsivo, utilizando linguagem Python, banco de dados MySQL e Flask Framework, permitindo que usuários possam classificar seus comentários ou visualizar comentários classificados pelo modelo como preconceituoso ou não.

Palavras-chave: Classificação automática; algoritmos; PLN; inteligência artificial.

Abstract

With the exponential increase in the amount of textual information, it becomes relevant to investigate methods and techniques that allow dealing with this content efficiently and automatically, thus, automatic text processing is a great challenge. Automatic text classification involves assigning one or more predefined document categories. This investigation proposes to examine the techniques associated with automatic categorization, proposing a computational tool that allows identifying prejudiced content in social networks, based on statistical and linguistic analysis of collected texts. The purpose of this work was to investigate and develop automatic classification techniques to perform the task of detecting hate speech against the LGBTQIA+ community. A responsive website was developed, using Python language, MySQL database and Flask Framework, allowing users to classify their comments or view comments classified by the model as prejudiced or not.

Keywords: Automatic classification; algorithms; PLN; artificial intelligence.

1 Introdução

A internet vem aumentando exponencialmente sua quantidade de informações textuais geradas nessa última década; somente nos últimos 13 anos, o número de domínios cresceu de 15.000 para 350.000.000 (O'Dell, 2011). Esse aumento de informações também se reflete no crescimento da participação dos usuários nos conteúdos da Web. Tanta informação agrupada em um ciberespaço comum, onde são criadas centenas de milhares de novos conteúdos por dia

¹ Bacharel em Ciência da Computação pelo Centro Universitário Sagrado Coração-UNISAGRADO; E-mail: keterly.g@gmail.com.

² Doutor em Ciências da Reabilitação pelo Programa de Pós-graduação em Ciências da Reabilitação do Hospital de Reabilitação de Anomalias Craniofaciais da Universidade de São Paulo-HRAC/USP, câmpus Bauru; professor dos cursos Ciência da Computação e Jogos Digitais do Centro Universitário Sagrado Coração-UNISAGRADO, Bauru; E-mail: patrick.silva@usc.br.

torna interessante, porém difícil e extremamente demorada, a tarefa de processar informação. Entretanto, além de informações úteis, muito conteúdo impróprio, sobretudo, com carga preconceituosa é gerado e postado nas redes sociais, causando transtornos e incentivando essa prática criminosa.

Neste contexto, surgiu uma nova área de pesquisa relacionada ao Processamento de Linguagem Natural (PLN), chamada de Classificação Automática de Textos. As pesquisas nesta área visam extrair, de certo modo, dessa crescente quantidade de dados, dados úteis que permitam compreender o conteúdo divulgado pelos usuários de internet. Deste modo, é possível classificar textos considerando o conteúdo contido em determinado documento (Pang; Lee, 2008).

A classificação automática de textos tem sido bastante explorada em diversas tarefas: indexação automática para sistemas de recuperação de informação, organização de documentos e filtragem de textos, organização e filtragem de mensagens de e-mail, filtro de notícias, recomendação de documentos, etc. A utilização de especialistas humanos para realizar categorizações de forma manual é um processo custoso e lento e, diante desse enorme conteúdo gerado, motiva a investigação de técnicas de categorização automática de documentos. Os resultados obtidos nesta área de pesquisa são promissores e satisfatórios, o que incentiva e justifica o seu estudo e aplicação na tarefa de detecção de conteúdos preconceituosos (WEISS *et al.* 2005). As redes sociais têm sido palco de manifestações de caráter ofensivo, preconceituoso, discriminatório e de grave intolerância. Escondidas no anonimato que as redes sociais permitem com o distanciamento que promovem, algumas pessoas se sentem à vontade para expressar todo tipo de agressão e difusão de mentiras, ferindo a honra e dignidade das pessoas.

Detectar esses conteúdos impróprios e agir com rapidez representa o diferencial entre o sucesso ou o fracasso no combate ao preconceito, justificando os esforços envolvidos no desenvolvimento de ferramentas que possam processar esse tipo de informação. Deste modo, cada vez mais pessoas e principalmente empresas, estão interessadas em observar as postagens de um grupo de pessoas sobre temas que podem representar conteúdos impróprios e que devem ser eliminados de suas redes sociais.

O próprio governo federal lançou em 2015, o chamado Pacto pelo Enfrentamento às Violações de Direitos Humanos na Internet. A iniciativa, com o nome de “Humaniza Redes” prevê uma ouvidoria online, onde denúncias podem ser feitas e, em seguida, encaminhadas a provedores de internet (Jornal do Comércio, 2015).

Segundo um levantamento realizado pelo projeto Comunica que Muda, iniciativa da agência Nova/sb, são alarmantes os números referentes à intolerância do internauta brasileiro. A pesquisa aponta que plataformas como Facebook, Twitter e Instagram têm um número elevado de textos sobre temas sensíveis, como racismo, posicionamento político e homofobia. Foram identificadas 393.284 menções, sendo 84% delas com abordagem negativa, de exposição do preconceito e da discriminação (Comunica que Muda, 2016). O que se tem visto nas redes sociais é o acirramento do discurso de ódio, de intolerância às diferenças.

Como resultado do panorama político gerado a partir das eleições de 2014, houve um intenso debate nas redes, na maioria das vezes com xingamentos e discursos rasos, que incentivam o ódio e a divisão. Do total de mensagens analisadas, com cunho político, cerca de 97,4% delas abordaram aspectos negativos. A segregação virtual foi materializada no muro erguido no gramado do Congresso Nacional para separar manifestantes contra e a favor do impeachment (Matsuura, 2016).

O segundo tema com maior número de mensagens está relacionado ao ódio às mulheres, com isso, a misoginia se alastra pelas redes sociais. Assédio, pornografia de vingança, incitação ao estupro e outras violências são, por vezes, curtidas e compartilhadas, reforçando no ambiente virtual o machismo presente na sociedade de desigualdades de gênero (Matsuura, 2016).

Pessoas com algum tipo de deficiência, também sofrem com o preconceito nas redes sociais, havendo em muitas situações uma abordagem negativa sobre o tema. Termos como “leproso” e “retardado mental” e o uso da deficiência para “justificar” direitos são usados nessas citações preconceituosas (Matsuura, 2016).

O racismo também tem forte presença nas redes sociais brasileiras, fato este evidenciado recentemente com casos de ataques a celebridades negras. O levantamento também mostra que existe intolerância pela aparência, homofobia, classes sociais, idade/geração, religião e xenofobia. Este cenário mostra a necessidade de estudos que permitam debater a tênue linha que separa o discurso de ódio do direito à liberdade de expressão. O direito à liberdade de expressão não é absoluto, legislações tratam o discurso de ódio explicitamente como um limitador da liberdade de expressão (Matsuura, 2016).

Basicamente existem dois tipos de intolerância. O visível, em que o agressor vai direto ao ponto; e o invisível, mais sutil, que se esconde em comentários que podem passar despercebidos, pois abordam discursos que já foram incorporados pela sociedade, mas não pelas vítimas (Matsuura, 2016).

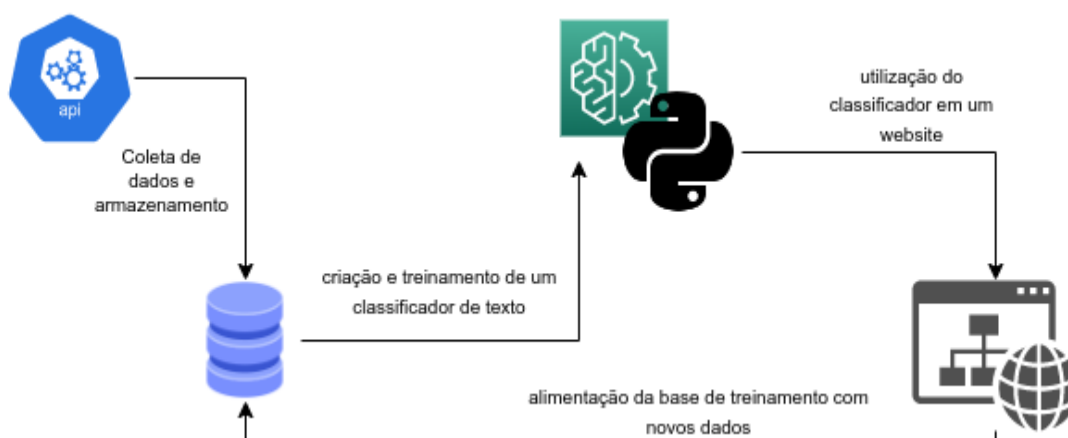
Nesta visão, ao contrário do que pode parecer, o Brasil tem se mostrado como um país intolerante. As redes sociais são um reflexo dessa realidade, amplificando o ódio e reafirmando os preconceitos que as pessoas já têm (Matsuura, 2016).

Diante desta problemática, este trabalho visa proporcionar subsídios para se estudar os métodos necessários para fazer com que postagens em redes sociais possam ser processadas e classificadas automaticamente, gerando um “entendimento” do seu conteúdo semântico, conforme a justificativa e objetivos apresentados adiante.

2 Desenvolvimento

Este artigo é derivado da monografia redigida a partir do desenvolvimento de um modelo de classificação automática de conteúdos preconceituosos, mais especificamente discursos de ódio contra a comunidade LGBTQIA+, realizado em um projeto de iniciação científica de mesmo nome. O conteúdo exposto é um exemplo de como a Inteligência Artificial, Aprendizado de Máquina e Técnicas de Processamento de Linguagem Natural podem contribuir para a filtragem dos conteúdos expostos em mídias sociais, muitas vezes de cunho racista, xenofóbico, machista, homofóbico, entre outros tipos de preconceitos que podem ser observados na sociedade atual. A figura 1 é um retrato resumido do processo de desenvolvimento deste projeto.

Figura 1 - Fluxo de desenvolvimento do projeto



Fonte: Elaborada pelos autores

2.1 Coleta de Conteúdo

Na etapa de coleta de conteúdo é feita uma busca em fontes diversas, tais como artigos em sites, comentários em mídias sociais, anúncios, documentos dentre outras. É importante a utilização de técnicas avançadas de busca, visando identificar se o conteúdo encontrado trata-

se de uma opinião ou um fato. Fatos por si só devem ser descartados, porém opiniões expressas em fatos devem ser mantidas (Becker; Tumitan, 2014).

Neste projeto foram selecionados textos digitais para compor a base de dados de treinamento do algoritmo de classificação, posteriormente exposto. Ao todo 500 comentários foram coletados e classificados manualmente, permitindo a programação das funções de classificação.

Os dados foram obtidos de algumas formas: Na plataforma do Youtube, foi feita uma coleta manual por meio da leitura de comentários em vídeos e prints exibidos pelos próprios Youtubers, em um tipo de conteúdo audiovisual com o objetivo de reagir aos comentários de seus próprios canais. Neste caso, o aplicativo Google Leans permitiu que, ao apontar a câmera para a imagem ou vídeo, o texto fosse extraído. Além desse tipo de meio para extração de palavras, foi possível copiar o texto para um dispositivo em específico, agilizando ainda mais o processo.

Essa mesma ferramenta foi utilizada quando a busca por textos foi feita no Google Imagens, assim como em sites que noticiavam cenas de preconceito e no próprio Twitter, onde muitas pessoas publicam prints em um *exposed*. Atualmente, ainda mais com a onda do que alguns chamam de “Cultura do Cancelamento”, usuários em redes sociais expõem outros usuários com prints ou informações a respeito de algum ato que gera indignação e repúdio.

A rede social Twitter teve a sua coleta de conteúdo focada na utilização de uma API, mas também de forma manual como citado anteriormente pela ferramenta Google Leans, ou pela procura manual de tweets escritos quando os resultados trazidos pela API não eram satisfatórios. A extração de conteúdo neste caso foi realizada o desenvolvimento de um script em linguagem Python e sua biblioteca tweepy, que permite a inserção de termos chave para a varredura na plataforma e retorno dos conteúdos que satisfazem as condições impostas.

Além de selecionar, foi necessário armazenar os dados. Para isso, o banco de dados MySQL foi escolhido e uma base criada, contendo inicialmente apenas uma tabela com 3 campos principais: id, comentário e emoção, sendo esta utilizada como treinamento para o modelo de classificação. Além de comentários preconceituosos, foram armazenados comentários que não possuíam discurso de ódio, objetivando ensinar o algoritmo a diferenciar textos com e sem preconceito evidente.

Por se tratar de uma base de treinamento, os dados foram classificados manualmente no campo “emoção” com os valores “preconceituoso” e “não preconceituoso”. A figura 2 demonstra uma consulta no banco e exemplo dos textos de treinamento.

Figura 2 - Leitura da base de treinamento

```
try:
    with conn.cursor() as cursor:
        query = "SELECT `comentario`, `emocao` FROM `tb_comentarios`"
        cursor.execute(query)
        result = cursor.fetchall()
        for linha in result:
            print(linha)
            base.append(linha)
except:
    print('erro')
finally:
    conn.close()
```

```
('Mano eu racho de rir com esses gays ou melhor endemoniados', 'preconceituoso')
('Um homem que se acha mulher... Um Hospicio esta cheio de gente que se acha que e o Dom Pedro, ou Tesla, etc... E nem por isso os achamos normais', 'preconceituoso')
('Um gay recebendo camisinhas Vai usar aonde? na orelha?', 'preconceituoso')
('Nao consigo ver o video, quando apertado o play aparece uma mensagem escrita Nao reproduz', 'preconceituoso')
('Quando eu vejo uma coisa dessas eu queria que voltasse o regime nazista pra acabar com essas merda de travestis ', 'preconceituoso')
('Orgao excretor nao reproduz, a igreja recrimina isso, sou a favor do bolsonaro e odeio os gays!', 'preconceituoso')
('Que nojo de voces dois ', 'preconceituoso')
('Fala devagar viado, nao da pra entender nada. Vai procurar uma fonoaudiologa bicha ridicula.', 'preconceituoso')
('Adoro essas gay que se diz bi porque ainda nao tem coragem de virar viado .', 'preconceituoso')
('Sou homofobico mesmo desgraça gayzinho', 'preconceituoso')
```

Fonte: Elaborada pelos autores.

2.2 Classificação

A classificação é a etapa mais importante do processo e é nela que a polaridade ou orientação da opinião é definida. Esta etapa determina se uma opinião é positiva, negativa ou neutra. Na classificação ou análise de sentimentos, são as palavras opinativas que têm a maior importância, pois, através delas, é possível determinar o sentimento expresso pelo autor. Exemplos de palavras opinativas: bom, legal, ótimo, ruim, péssimo etc. (Becker; Tumitan, 2014).

Ao montar a base de dados, foi possível utilizá-la para identificar a melhor biblioteca para classificar automaticamente os textos digitais inseridos como teste. Inicialmente, com base nos estudos de cursos, o código em Python foi desenvolvido com a biblioteca NLTK e o algoritmo foi treinado. Posteriormente, acompanhando um live no Youtube, o mesmo teste foi feito utilizando a biblioteca NLTK junto a Sklearn. O objetivo desta etapa era a avaliação das bibliotecas para a escolha, e não o desenvolvimento do programa efetivamente. Após as comparações, a biblioteca NLTK foi escolhida, tendo seu algoritmo re-programado, focado em testes de classificação e na análise dos resultados.

2.3 Sumarização dos Resultados

Esta etapa é focada na apresentação dos resultados, que podem ser de forma textual, ou gráfica. A melhor forma de representar os resultados é a gráfica, pois facilita a visualização e

entendimento dos resultados sumarizados em totais e dados estatísticos (Becker; Tunitan, 2014).

No presente trabalho, após análise da melhor forma de representação dos resultados, foi escolhido que, a execução e resultados do algoritmo poderiam ser demonstrados por meio de uma interface web simples e de fácil uso e acesso para qualquer usuário.

Depois da programação do algoritmo de classificação, foi iniciada a fase de criação do produto web, dividida em duas etapas de desenvolvimento: front end e back end.

O desenvolvimento front end englobou a criação do protótipo da interface através da ferramenta Figma, sendo uma base para a programação, mas possibilitando alterações futuras. Neste protótipo decidiu-se pela criação da página principal, página de classificações corretas e incorretas, página de classificar e a página de estatísticas do algoritmo.

Para a utilização da linguagem Python na web, o framework Flask foi escolhido e com isso deu-se início a programação com o apoio de outras tecnologias: HTML, CSS, Bootstrap, Chart.js e MySQL.

Utilizando essas tecnologias, algumas rotas foram criadas e os templates desenvolvidos, inserindo informações sobre o projeto, tabela de resultados, formulário para a interação com o site e gráficos com estatísticas de erros e acertos do algoritmo, além de uma visão da quantidade de comentários obtidos com este site, e que permitiu o enriquecimento da base de dados inicial. Foi desenvolvida também uma página não visível no menu, onde o usuário valida se a classificação está ou não correta.

No desenvolvimento back-end com a linguagem Python, foi necessário analisar a adaptação do algoritmo de classificação. Por ter sido programado anteriormente utilizando Jupyter Notebook, houve uma mudança com adição de novas funções, principalmente para a interação com o banco de dados. Foi necessário também criar três novas tabelas, contendo os campos id, comentário e emoção. Uma tabela foi responsável pelo armazenamento dos comentários corretos, e as outras, dos comentários incorretos. Programaram-se funções para selecionar todos os comentários das tabelas no banco de dados além da inserção de novos comentários.

A programação backend controlou toda a lógica do site, interferindo nas tabelas de comentário, na classificação do comentário do usuário e na alimentação dos gráficos. Os resultados (telas) do protótipo desenvolvido serão descritos na seção seguinte.

2.4 Resultados e discussões

O algoritmo implementado neste projeto utilizou a biblioteca NLTK para a classificação dos comentários. Esta classificação trabalhou com diversas funções, sendo utilizadas para a limpeza dos dados, obtenção da frequência das palavras, extração das palavras únicas e a classificação propriamente dita. Por trabalhar com uma tabela de probabilidade, o algoritmo indicou a porcentagem em relação às classes “preconceituoso” e “não preconceituoso”. Na programação web essas porcentagens foram utilizadas em uma lógica de comparação, onde, ao invés de exibir em porcentagem para o usuário, o site faz uma exibição utilizando palavras, melhorando o entendimento do leitor.

A figura 3 demonstra de que forma o classificador interpreta o dado de entrada e escolhe a classe onde ele melhor se encaixa. Para este exemplo, não existe discurso de ódio e sim uma informação sobre algo ocorrido, sendo assim, classificado como não preconceituoso.

Figura 3 - Classificação automática de novo dado de entrada

```
teste = "eu fui espancado porque sou travesti"

palavraTeste = []
for palavrasTeste in teste.split():
    testeStm = [p for p in palavrasTeste.split()]
    palavraTeste.append(str(stemmer.stem(testeStm[0])))

novaTeste = extra_palavras(palavraTeste)
print(classificador.classify(novaTeste))

nao preconceituoso

#classificando com probabilidade
dist = classificador.prob_classify(novaTeste)
for classe in dist.samples():
    print("%s %f" % (classe, dist.prob(classe)*100))

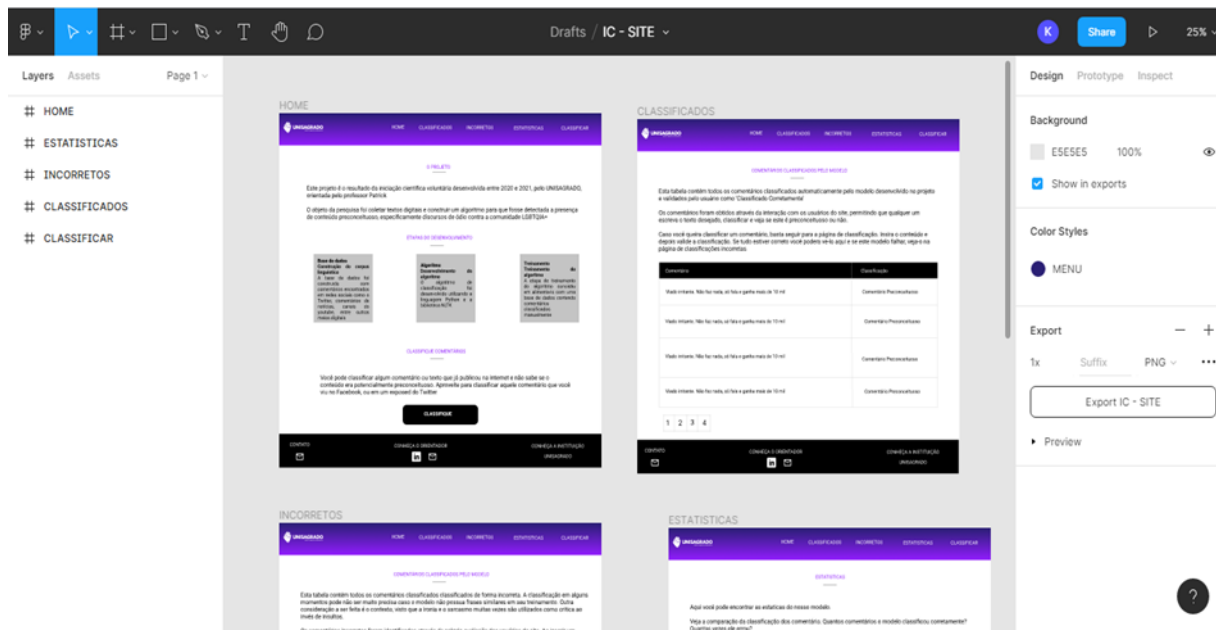
preconceituoso 38.000960
nao preconceituoso 61.999040
```

Fonte: Elaborada pelos autores.

Em uma análise dos resultados do algoritmo, a métrica acurácia atingiu o valor de 90%. No entanto, por se tratar de uma aplicação com interação com usuários, armazenando os textos classificados por eles e aumentando o número de dados na base de treino, o valor obtido alternava à medida em que essa interação acontecia.

Além da programação do modelo de classificação, utilizou-se da ferramenta Figma (Figura 4) para a prototipagem da interface web. Com ela foi possível testar cores, espaçamento, margem e a disposição dos elementos. O protótipo foi a base para o desenvolvimento front-end, mas não impossibilitou mudanças visuais conforme a necessidade.

Figura 4 – Design da interface utilizando Figma



Fonte: Elaborada pelos autores.

As imagens seguintes apresentam as páginas do site desenvolvido utilizando Flask com a linguagem Python, HTML, CSS, Bootstrap, Chart.js e o banco de dados MySQL.

A tela “home” (Figura 5) exibe todas as informações gerais do projeto, descrevendo brevemente o objetivo, o orientador e as etapas da pesquisa, além de fazer uma chamada para que o usuário classifique um conteúdo.

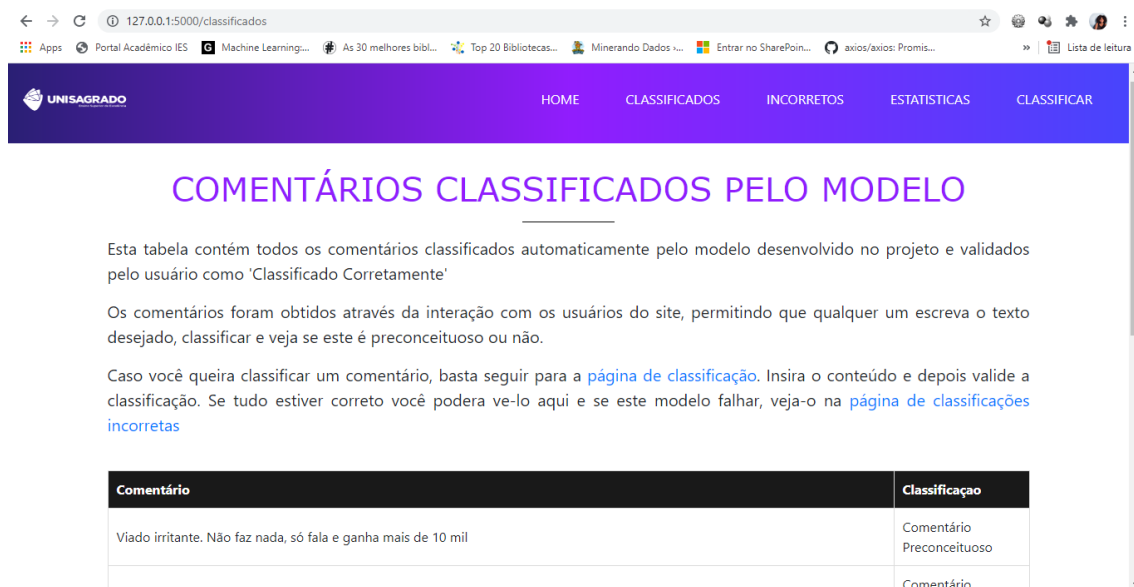
Figura 5 – Página HOME com navegação para a página classificados



Fonte: Elaborada pelo autor

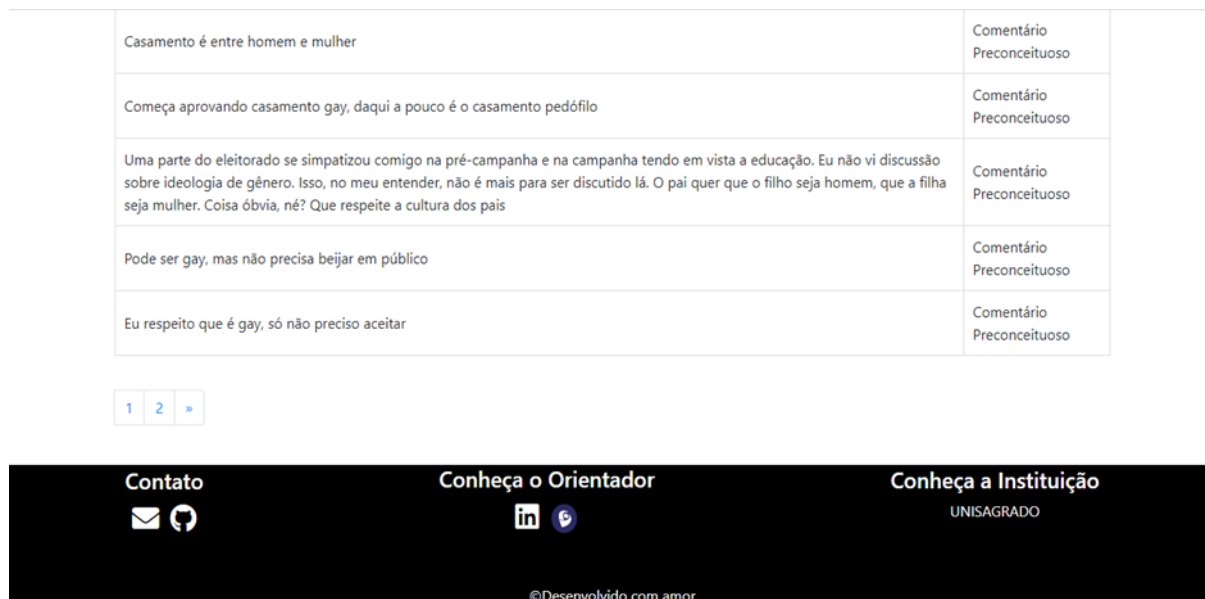
A tela “classificados” (Figura 6 e Figura 7) exibe em uma tabela todos os comentários classificados corretamente pelo modelo de acordo com a validação do próprio usuário. Além disso, ela traz uma parte informativa.

Figura 6 – Página CLASSIFICADOS



Fonte: Elaborada pelos autores.

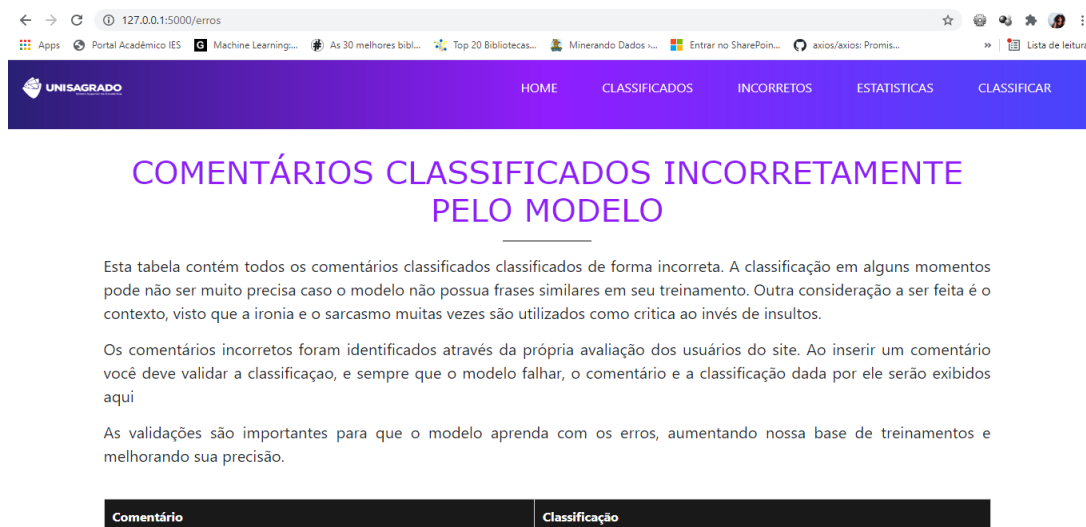
Figura 7 – Página CLASSIFICADOS com paginação de resultados



Fonte: Elaborada pelos autores.

A tela “incorretos” (Figura 8) exibe em uma tabela todos os comentários classificados incorretamente pelo modelo, de acordo com a validação do próprio usuário. Além disso, ela traz uma parte informativa.

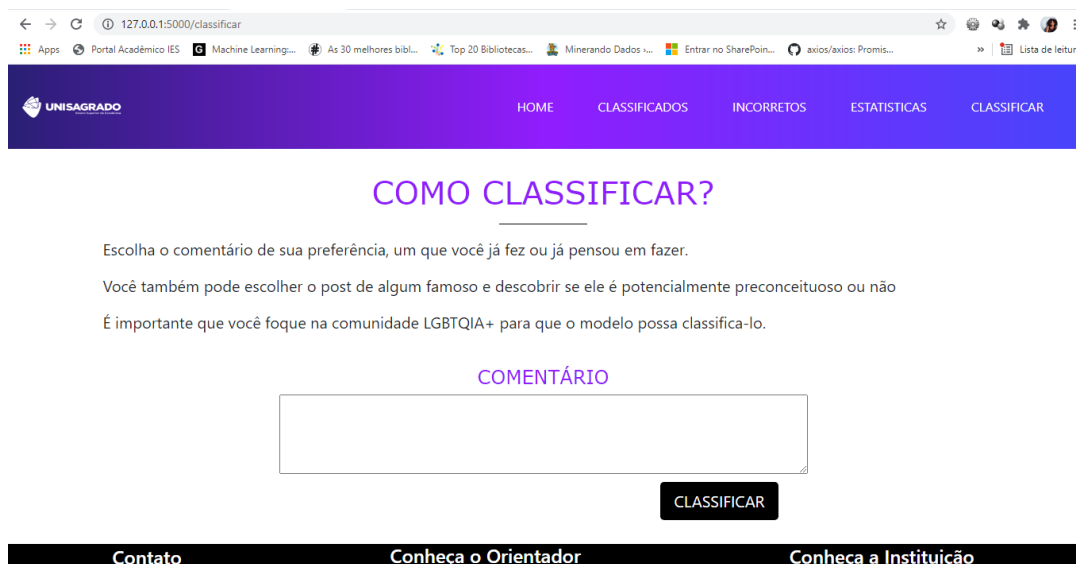
Figura 8 – Página INCORRETOS



Fonte: Elaborada pelos autores.

A tela “classifique” (Figura 9) exibe um campo para que o usuário insira um conteúdo e veja sua classificação. Ao clicar para classificar, ele é redirecionado para a página de validação.

Figura 9 – Página CLASSIFICAR



Fonte: Elaborada pelos autores.

A tela “validação” (Figura 10) é a página onde o usuário informa se a classificação dada pelo modelo está correta. Com essa validação ele é redirecionado para a página “classificados” caso o modelo tenha acertado, ou para a página “incorretos”. Se incorreto, o comentário é salvo na base de dados de comentários incorretos e na base de dados teste, onde sua classificação é invertida para que ele seja utilizado nos treinamentos. Se correto, ele é salvo somente na base de dados de comentários corretos, mas também é usado no treinamento.

Figura 10 – Página de VALIDAÇÃO

VALIDAÇÃO DA CLASSIFICAÇÃO

Antes de listarmos esse comentário em nossa tabela, precisamos saber se ele está classificado corretamente.

Nos ajude a melhorar o algoritmo informando se esta classificação está correta

Se estiver, você será redirecionado para a aba de Classificados. Se não estiver, você será redirecionado para a aba Incorretos, podendo ver outros comentários classificados de forma errada.

NAO TENHO PRECONCEITO, SO NAO QUERO GAY PERTO DE MIM

Este comentário foi considerado preconceituoso

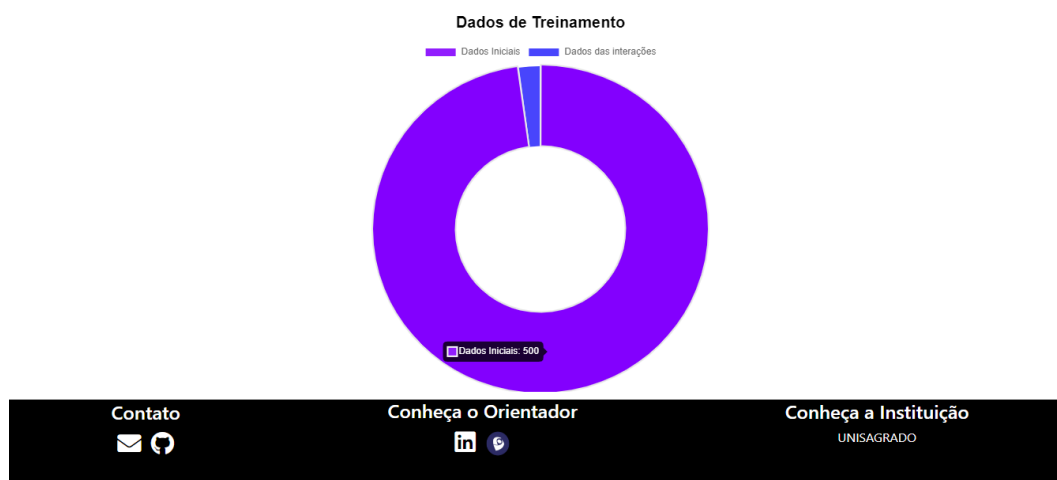
Correto

Incorreto

Fonte: Elaborada pelos autores.

A tela “estatística” (Figura 11) exibe dois gráficos alimentados dinamicamente usando o banco de dados. Um gráfico mostra a estatística de erros e acertos do modelo e o outro faz uma comparação entre a quantidade de dados iniciais e os novos dados obtidos com a interação dos próprios usuários.

Figura 11 – Página de ESTATÍSTICAS com gráfico de treinamento



Fonte: Elaborada pelos autores.

3 Conclusão

A ferramenta desenvolvida é uma maneira de demonstrar de que forma é possível aplicar classificação automática de conteúdos em sistemas web para a identificação de discursos de ódio. Hoje é possível ver filtragem em comentários identificados como impróprios, mas ainda assim se vê muitas postagens que expõem sua repulsa por tudo aquilo que é diferente de sua percepção do que pode ser visto como “normal” na nossa sociedade.

Este site classificador pode ser disponibilizado para mais usuários e assim aprimorado em trabalhos futuros, podendo ser utilizado servidores gratuitos para isso, bem como a melhoria de

sua interface e de seu algoritmo, uma vez que os avanços de estudos na área e o seu aprofundamento permite aplicações de novas técnicas e testes de diferentes modelos de aprendizado de máquina para a resolução do problema proposto neste trabalho.

O desenvolvimento deste projeto, que é resultado de uma iniciação científica trouxe, além do conhecimento nas tecnologias e ferramentas utilizadas, uma reflexão sobre o comportamento humano no mundo digital e como comentários, principalmente em redes sociais, podem ser carregados de ódio e preconceito.

Referências

BECKER, K.; TUMITAN, D. Introdução à Mineração de Opiniões: Conceitos, Aplicações e Desafios. In: Joao Eduardo Ferreira. (Org.). **Lectures of the 28th Brazilian Symposium on Databases**. 1ed. Pernambuco: CIN - UFPE, 2013, p. 27-52. Disponível em: https://www.inf.ufrgs.br/~kbecker/lib/exe/fetch.php?media=minicursosbbd_versaosubmetida.pdf. Acesso em: 04 jul. 2024.

COMUNICA QUE MUDA, 2016. **Nova/sb identifica quadro de intolerância no Brasil**. São Paulo, 08/04/2016. Disponível em: <http://www.clubedecriacao.com.br/ultimas/comunica-que-muda-2/>. Acesso em: 04 jul. 2024.

JORNAL DO COMÉRCIO. **Governo Federal Lança Ação Para Defesa de Direitos Humanos na Internet**. Porto Alegre, 07/04/2015. Disponível em: <https://www.jornaldocomercio.com/site/noticia.php?codn=192894>. Acesso em: 04 jul. 2024.

MATSUURA, S. **Brasil Cultiva Discurso de Ódio nas Redes Sociais, Mostra Pesquisa**. O Globo, Rio de Janeiro, 03 ago. 2016. Disponível em: <https://oglobo.globo.com/brasil/brasil-cultiva-discurso-de-odio-nas-redes-sociais-mostra-pesquisa-19841017#newsletterLink>. Acesso em: 04 jul. 2024.

O'DELL, J. **How Big Is the Web & How Fast Is It Growing?** Mashable, Nova York, 19 jun. 2011. Disponível em: <https://mashable.com/2011/06/19/how-many-websites/>. Acesso em: 04 jul. 2024.

PANG, B.; LEE, L. Opinion Mining and Sentiment Analysis. **Foundations and Trends in Information Retrieval**, v. 2, p. 1–135. Disponível em: <https://ieeexplore.ieee.org/document/8187070/metrics#metrics>. Acesso em: 04 jul. 2024.

PUBLICO. Comentários racistas publicados no Facebook divulgados à porta de quem os escreveu. Lisboa, 1 dez. 2015. Disponível em: <https://www.publico.pt/2015/12/01/tecnologia/noticia/comentarios-racistas-publicados-no-facebook-divulgados-a-porta-de-quem-os-escreveu-1716118>. Acesso em: 04 jul. 2024.

WEISS, S. M. *et al.* **Text Mining: Predictive Methods for Analyzing Unstructured Information**. New York, Springer Science & Business Media, 1st ed. 1, 2005.