

COMO CONTRIBUIR COM O PROJETO PEGABOT?

Pesquisa e Estratégia Grupo 2A

Ana Leticia

Pâmela

Abigail

Edital Pegabot

Categoria 3 | Pesquisa & Estratégia: A categoria 3 engloba questões de pesquisa e análise de dados, com o objetivo de pensar os próximos passos da ferramenta Pegabot e como ela pode se tornar mais versátil. Podemos citar como exemplos de desafios a criação de um sistema de monitoramento em tempo real de ataques à ONGs e Jornalistas e o estabelecimento de características para uma melhor análise de perfis pelo Pegabot.

Desafio 2

Hackathon Pegabot Match GP | Pesquisa & Estratégia - Planilhas Google

Com o objetivo de monitorar em tempo real ataques à OSCs e Jornalistas, dois atores importantes no processo democrático, o mecanismo desenvolvido nesse desafio deverá possuir explorar formas de detecção de padrões linguísticos de discurso de ódio e poderá ter, em paralelo, o desenvolvimento de um motor de inteligência artificial para monitorar de forma automatizada esses ataques. Assim, será possível observar, classificar e agir na defesa desses usuários atingidos por essa prática maliciosa nas redes.

Brainstorming:

Análise de dados

Base bibliográfica sobre Discurso de ódio

[Veja como monitorar em tempo real ataques coordenados no Twitter - Escola de Dados](#)

<https://www.kaggle.com/datasets/gpreda/presidential-debate-video-comments>

https://www12.senado.leg.br/ril/edicoes/52/207/ril_v52_n207_p143.pdf

<https://www1.folha.uol.com.br/poder/2022/08/big-techs-silenciam-sobre-equipes-de-moderacao-em-portugues-a-2-meses-das-eleicoes.shtml>

<http://preventviolentextremism.info/sites/default/files/A%20Lexicon-Based%20Approach%20for%20Hate%20Speech%20Detection.pdf>

[Twitter API Documentation | Docs | Twitter Developer Platform](#)

[Getting access to the Twitter API | Docs | Twitter Developer Platform](#)

<https://internetlab.org.br/pt/imprensa/discurso-odio-elle-fernanda-martins/>

O que significa odiar uma pessoa?

Trata-se de um sentimento negativo em que se deseja mal ao sujeito ou objeto **odiado**. O ódio está relacionado com a inimizade e a repulsão. As **pessoas** tentam

evitar ou destruir aquilo que odeiam. No caso do ódio relativamente a outro ser humano, o sentimento pode refletir-se através de insultos ou de agressões físicas.

Foi entendido que o twitter estabelece um código de conduta, e depende de utilizarem as denúncias, para que providências sejam tomadas. *(Analisaremos e tomaremos medidas contra denúncias de contas direcionando os comportamentos seguir a um indivíduo ou grupo de pessoas em Tweets ou Mensagens Diretas.)*

[Política do do Twitter contra conduta de propagação de ódio | Ajuda do Twitter](#)

Por exemplo, membros de uma categoria protegida podem se referir uns aos outros com termos que são normalmente considerados xingamentos. Quando usados de modo consensual, a intenção por trás desses termos não é o abuso

Definição de discurso de ódio: utilização de palavras “que tendem a insultar, intimidar ou assediar pessoas em virtude de sua raça, cor, etnicidade, nacionalidade, sexo ou religião” ou ainda à sua potencialidade ou “capacidade de instigar violência, ódio ou discriminação contra tais pessoas” (Winfried Brugger 2007, p. 151),

Como reconhecer o discurso de ódio? O **discurso de ódio** possui duas características básicas: (I) Insulto e/ou ofensa a uma pessoa, incluindo um grupo socialmente vulnerável ao qual ela pertence; (II) Fala, gesto, expressão que instiga a violência, seja ela explícita ou implícita na fala do agressor.

É crime discurso de ódio na internet?

No país, cumprir **discurso de ódio** não é de maneira nenhuma penalizado desde março de 2014.

[Eu vejo robôs e eles estão por toda parte: \(quase\) tudo que você precisa saber sobre bots e metodologias para detecção de automação | by ITS Rio | ITS FEED](#)

[Bots para todos os lados: os desafios da detecção de bots em redes sociais | by ITS Rio | ITS FEED](#)

[Integridade de dados de mensagens - Documentação da IBM](#)

[Protocolos de segurança criptográficos: TLS - Documentação da IBM](#)

<https://arxiv.org/abs/1911.09179>

[Aprofundamentos sobre o PEGABOT](#)

<https://pegabot.com.br/>

[GitHub: Where the world builds software · GitHub - 4 August 2022 \(loom.com\)](#)

[https://drive.google.com/drive/u/0/folders/1I1qhnP6plAbseMMr8BqsmDWOCIp7WzQ](https://drive.google.com/drive/u/0/folders/1I1qhnP6plAbseMMr8BqsmDWOCIp7WzQ4)

4

[Lab de dados - INCT.DD - Google Drive](#)

[Relatório da consultoria para Projeto PegaBot.pdf - Google Drive](#)

[FAQ · Pegabot](#)

[Categoria 2 | Narrativas - Google Drive](#)

BOTPONTO

[BotPonto \(nucleo.jor.br\)](#)

[Projeto 7C0 | Home](#)

GitHub tips

[Pesquisas-e-Narrativas/CONTRIBUTING.md at main · Ana2030/Pesquisas-e-Narrativas \(github.com\)](#)

[Megabots/Pesquisas-e-Narrativas: Esse repositório objetiva agrupar diferentes frentes de desenvolvimento do Pegabot com o objetivo de construir coletivamente novas formas de funcionamento e atuação da ferramenta. \(github.com\)](#)

[Pegabot · GitHub](#)

[GitHub - AppCivico/pegabot](#)

Criação de perfil github

<https://github.com/pamelaferreiralima/Pesquisas-e-Narrativas>

Direcionamento de ideias:

[Hackathon PegaBot - Grupo 2A - Google Jamboard](#)

[Passo a Passo Github - Hackathon - Documentos Google](#)

[GitHub: Where the world builds software · GitHub - 4 August 2022 \(loom.com\)](#)

[Pegabot · GitHub repositório](#)

Compartilhamento de trabalhos

<https://sebraecanvas.com/#/dashboard/meus-canvas/967121>

[Hackathon PegaBot - Grupo 2A - Google Jamboard](#)

[Hackaton PegaBot - Documentos Google](#)

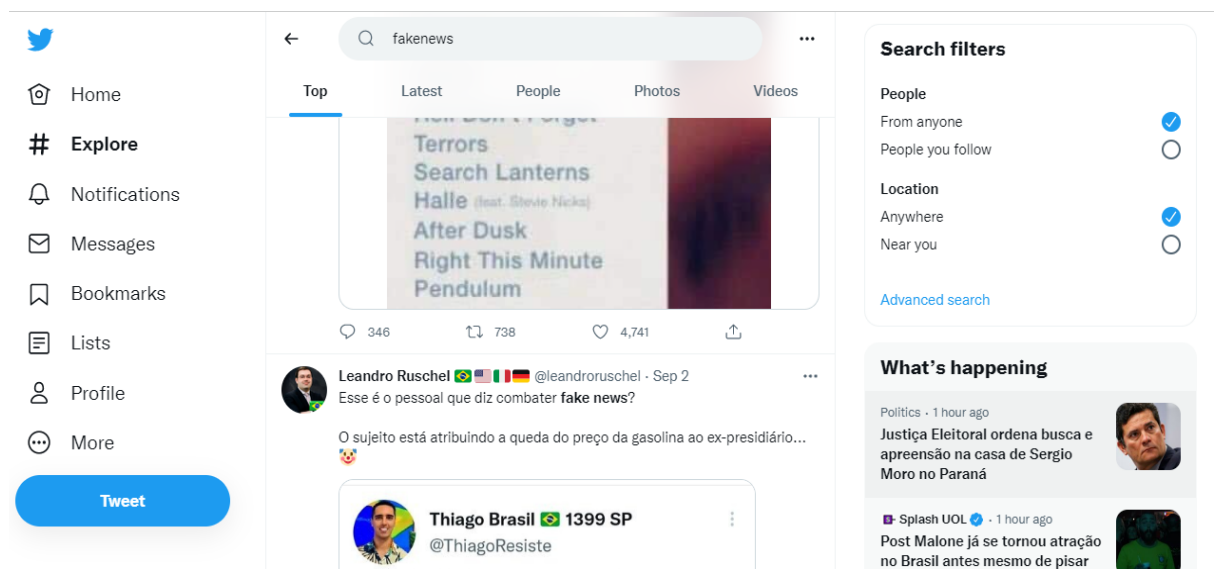
Quais as ferramentas necessárias para fazer essa análise ?

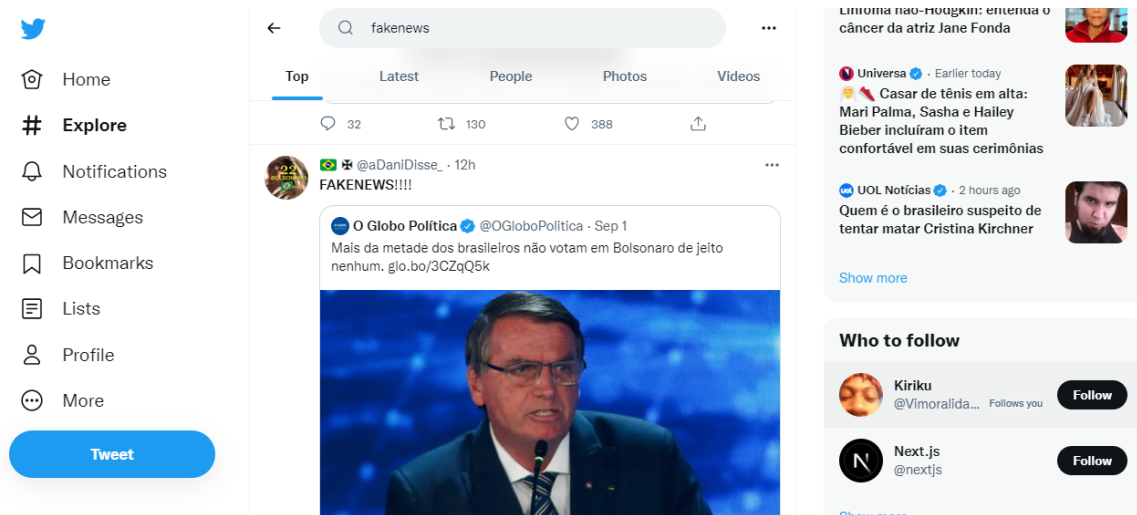
Índice de exportações que o pegador consegue capturar:

- Falso/Negativo Falso/Positivo
- Criação de dicionário de palavras violentas/ofensivas (processamento de linguagem natural)
- <http://www.iramuteq.org/>
- Análise do contexto que a palavra está inserida

- Análise de redes através do [Gephi](#) (uma das ferramentas)
- https://files.cercomp.ufg.br/webby/up/771/o/Gephi-Guia_B%C3%A1sico_de_Interface.pdf
- <https://developer.twitter.com/en/docs/twitter-api/getting-started/getting-access-to-the-twitter-api>
- Protocolo de segurança
- Analisar dados (discurso, palavras, contexto, histórico do usuário, material compartilhado)
 - identificar anomalias de ataques à uma conta específica
 - verificar se a conta se manifestou sobre alguma pauta recente e é um agente da sociedade civil com relevância
 - suspender comentários
 - analisar os discurso que estão sendo feitos e verificar se estão associados a algum tipo de discurso de ódio com referência e leituras jurídicas sobre discurso de ódio
 - verificar se é um ataque coordenado por meio de Análise de rede
 - Se for um discurso de ódio, aplicar as punições

Definir indicadores de análise? focar em análises qualitativas? criar parâmetros
 Como nosso bot vai ser encontrado? A partir de que ponto ele começa a trabalhar?
 -Analisar um evento (recente?) com discurso de ódio?
 -Validar nossa proposta (entrevista)
 o público?
 -





Edital Hackaton Pegabot:

https://www.hackathonpegabot.com.br/_files/ugd/6dff39_b99634f741d544328d60a572d8d3c889.pdf

Pela leitura do artigo sobre “RELATÓRIO DE CONSULTORIA PARA O PROJETO PEGABOT”, foi feita uma análise a respeito da identificação do que é um bot e o que não é bot, para isso foi desenvolvido alguns tipos de análises como análise temporal, análise de sentimento, análise de usuário e análise de rede.

Social bots são softwares programados para realizar tarefas automatizadas dentro das plataformas e agir como usuários comuns.

Em um dos resultados apresentados com os perfis avaliados se eram ou não bots, os que correspondiam a porcentagem igual ou acima a 70% são provavelmente bots e no relatório foram apresentados dois exemplos de bots, um que tem como funcionalidade ser replicador ou seja, foco em retweets e o outro perfil bot que foca em repetição de hashtags.

Ferramenta de software para análise dos dados :

IRAMUTEQ

Introdução:

Buscar cases voltados ao público alvo jornalista.

buscar palavras do discurso de ódio para o dicionário (Judeus pela Democracia - Oficial no Twitter: ""Deus, Pátria, Família". O candidato bolsonaro termina o debate com o lema criado pelo fascismo e usado pelo integralismo no Brasil. #DebateNaBand
<https://t.co/Si6DI22uWU> / Twitter)

Machine learn, usar aprender palavras, aumentar dicionário....

Desenvolvimento:

Análise Lexical:

<https://linguistic-datasets-pt.etica.ai/>

Detecção de discurso de ódio utilizando vetores de features aplicados a uma base nova de comentários em português

<https://revistas.unifacs.br/index.php/rsc/article/download/6463/4034>

Arquivo no Colab para extrair dados e fazer análises

https://colab.research.google.com/drive/1OtxU2sWT29qh3Q_gmVMiMP6mI7aE2rk1?usp=sharing

https://www.researchgate.net/publication/338950421_Analise_Lexical_por_meio_do_softwar_e_Iramuteq_Estudo_do_Significado_do_Trabalho_do_Juiz

<http://www.iramuteq.org/documentation/fichiers/tutoriel-en-portugais>

Tutorial iramuteq: [Tutorial_Iramuteq_2013_portugues-with-cover-page-v2.pdf](#)
(d1wqtxts1xzle7.cloudfront.net)

Perfil para ser analisado da jornalista dos direitos humanos:

<https://twitter.com/biancasantana>

<https://twitter.com/jpdoficial1/status/1564081270971326466>

Tutorial para usar a ferramenta iramuteq:

- ▶ Como usar o Iramuteq?

- ▶ Primeiros passos com o Iramuteq

ENTREGA

Submissão | Hackathon Pegabot (typeform.com)

PITCH

-tempo de 5 min máx

→ Nele, será necessário informar:

- Nome do grupo e dos participantes
- Link para o Pull Request no GitHub
- Link para o vídeo do Pitch, salvo no Youtube (Como Não Listado)

prencher forms