

## ANÁLISE DA PERCEPÇÃO DAS PESSOAS SOBRE ATUAÇÃO DOS ÓRGÃOS RESPONSÁVEIS POR SEGURANÇA PÚBLICA

Carlos Henrique do Vale e Silva bolsista PIBIC/UFPI, (Glauber Dias Gonçalves Orientador, UFPI-CSHNB)

Palavras-chave: Operações Policiais; Youtube; Processamento de Linguagem Natural, Monitoramento.

### 1. INTRODUÇÃO

Redução e controle de taxas de criminalidade são desafios enfrentados nos grandes centros urbanos do mundo. Esta é uma questão ainda mais grave nos países em desenvolvimento. No Brasil, por exemplo, foram registradas 40.804 mortes por crimes violentos em 2022, contra 41.160 em 2021, 356 mortes a menos. A taxa de mortes decorrentes de crimes violentos por 100 mil habitantes varia de 1,41 a 1,69 no cenário nacional, enquanto nos estados do Nordeste essa taxa chega a mais de 3 mortes, segundo o monitor da violência G1 NEV-USP.

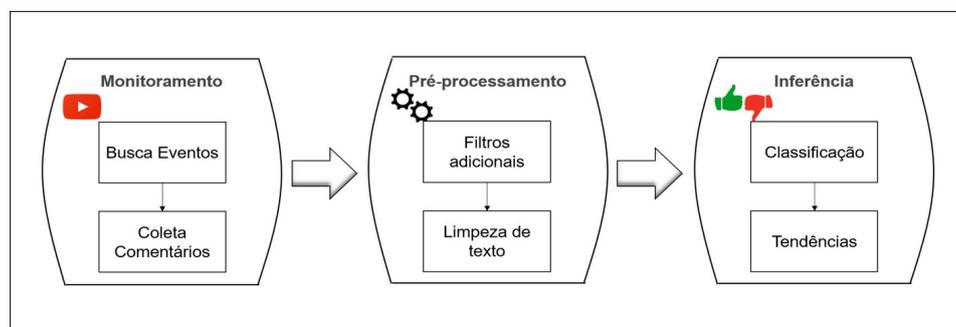
Diante desses fatos, costumeiramente a população não só brasileira, mas também mundial, vem recorrendo cada vez mais às plataformas Web para manifestar opiniões através de comentários sobre a atuação dos órgãos e autoridades responsáveis pela segurança pública. Ao passo que esses comentários são uma das principais formas de manifestação da população, eles oferecem uma oportunidade única para compreender a percepção da população sobre violência urbana, em especial, o posicionamento e sentimento das pessoas sobre operações policiais.

Neste trabalho, investigamos o potencial de comentários de usuários no Youtube para analisar a percepção das pessoas sobre operações policiais em incidentes de violência urbana com repercussão nessa plataforma. Esforços anteriores da literatura já usaram comentários de usuários em redes sociais no contexto de crimes e sensação de insegurança para avaliar sentimentos da população (CHAPARRO et. al, 2022, HAND et al., 2022). Contudo, observa-se ainda a falta de metodologias para inferir posicionamentos, i.e., aprovação, neutralidade ou desaprovação no contexto de segurança pública.

Os resultados anteriores desse projeto (FEITOSA et al., 2022a, FEITOSA et al., 2022b) mostram que modelos BERT<sup>1</sup> pré-treinados em língua local e retreinados com expressões do contexto obtiveram melhor desempenho para inferências. No projeto corrente oferecemos duas novas contribuições: (i) metodologia para extrair comentários da população sobre operações policiais baseada no serviço Youtube; e (ii) estudo de caso sobre inferência de posicionamentos dos usuários desse serviço em operações policiais no ano de 2022. Mais detalhes sobre essas contribuições podem ser obtidos em nossas bases de dados e códigos disponíveis<sup>2</sup>.

### 2. METODOLOGIA

Propomos um sistema para inferir posicionamentos de comentários das pessoas na Web sobre operações policiais no Brasil. O sistema consiste em três módulos como mostra a Figura 1.



<sup>1</sup> Abordagem estado da arte para processamento de linguagem natural proposta pelo Google em 2019.

<sup>2</sup> [https://github.com/LABPAAD/crimes\\_stance](https://github.com/LABPAAD/crimes_stance)

**Figura 1. Módulos do sistema proposto enfatizando o fluxo contínuo de monitoramento, pré-processamento e inferência de comentários baseado em modelos de classificação pré-treinados.**

**Monitoramento:** utilizamos o YouTube como fonte de vídeos e comentários sobre notícias relacionadas a operações policiais. Escolhemos o YouTube por ser uma das redes sociais mais populares e acessadas no Brasil, e por permitir o acesso aos dados através da API do YouTube versão 3. A partir dessa, desenvolvemos um programa coletor de comentários no referido contexto de forma automatizada. Esse coletor tem dois procedimentos que são (1) busca por eventos e (2) coleta de comentários sobre esses eventos. No primeiro procedimento utilizamos comando `search().list()` da API com as seguintes palavras-chave: *assassinato ou morte ou roubo ou furto ou polícia* em períodos semanais. Assim, obtivemos uma série de 52 unidades de tempo (semanas) em 2022 e foram coletados um total de 5038 vídeos. No segundo procedimento codificamos no coletor o comando `commentThreads().list()` da API do YouTube, passando o identificador do vídeo como parâmetro para Coletar informações de todos os comentários disponíveis para cada vídeo.

**Pré-processamento:** visa refinar os dados coletados. Nesse sentido, aplicamos um novo filtro sobre os vídeos coletados para selecionar apenas aqueles que tem relação com operações policial, que é o foco do nosso trabalho. Para isso, selecionamos apenas vídeos que continham no título as palavras “policial”, “pm” e “polícia”, considerando transformação de todos os caracteres dos títulos em minúsculo e remoção de acentos e sinais ortográficos. Após a filtragem, restaram 120 vídeos, que estão associados ao total de 76373 comentários. A seguir, conduzimos outra série de pré-processamentos nos comentários via a biblioteca NLTK<sup>3</sup>, removendo links, emojis e quebras de linha, mas mantendo *stopwords* para inferências com a mesma sequência em que as palavras aparecem nos comentários.

**Inferência:** um modelo de aprendizagem de máquina previamente treinado é aplicado para classificar e identificar a tendência ao longo do tempo para posicionamento de usuários no serviço monitorado. Definimos tendências como picos de comentários em unidades de tempo que indicam claramente a percepção pública via os posicionamentos de *aprovação, desaprovação ou neutralidade*. Esses picos estão relacionados a eventos específicos relativos à incidentes de segurança polêmicos que demandam operações policiais com cobertura jornalística da imprensa. Utilizamos um modelo de processamento de linguagem natural (PLN) desenvolvido e avaliado em nosso trabalho anterior (FEITOSA et al, 2022a, FEITOSA et al., 2022b), pré-treinado na língua portuguesa com uma camada extra retreinada com expressões do contexto. O desempenho desse modelo mostrado na Tabela 1 foi o melhor dentre vários outros testados, por esse motivo ele foi escolhido para inferências nesse trabalho.

**Tabela 1. Desempenho do modelo PLN BERTimbau usando as estratégias de ajuste fino da rede neural, avaliado com as seguintes métricas: Precisão (P), Revocação (R), F1-score (F1), Acurácia (Acc) e F1-macro, extraído de Feitosa et al. (2022a).**

Negativo			Positivo			Neutro			Acc	F1-macro
P	R	F1	P	R	F1	P	R	F1		
0,58	0,63	0,61	0,69	0,68	0,69	0,77	0,74	0,76	0,70	0,68

### 3. RESULTADOS E DISCUSSÃO

A Figura 1-a mostra as tendências de posicionamento sobre operações policiais em 2022 classificados pelo sistema. Observamos majoritariamente neutralidade, em especial nos picos de comentários cujas datas são indicadas na figura. Isso ocorre devido a politização das operações policiais em ano de eleição presidencial. Contudo, desaprovações dominam aprovações nesses picos, acima da porcentagem de erro do modelo de 30%, indicando rejeição à polícia nesses incidentes. Na Figura 1-b a tendência de desaprovação está mais evidente. Considerando o acumulado total no período temos posicionamentos de aprovação 12,46%, neutralidade 64,11% e desaprovação 23,43% desaprovação das operações policiais. Observa-se que comentários não neutros aumentam significativamente em junho: desaprovações, inicialmente pouco superior a 1400 no primeiro pico, aumenta para 6930 a partir de 19 de

<sup>3</sup> <https://www.nltk.org/>

junho, tornando esse o mês com os maiores percentuais de opiniões de desaprovação, impactando no total acumulado notavelmente em 2022.

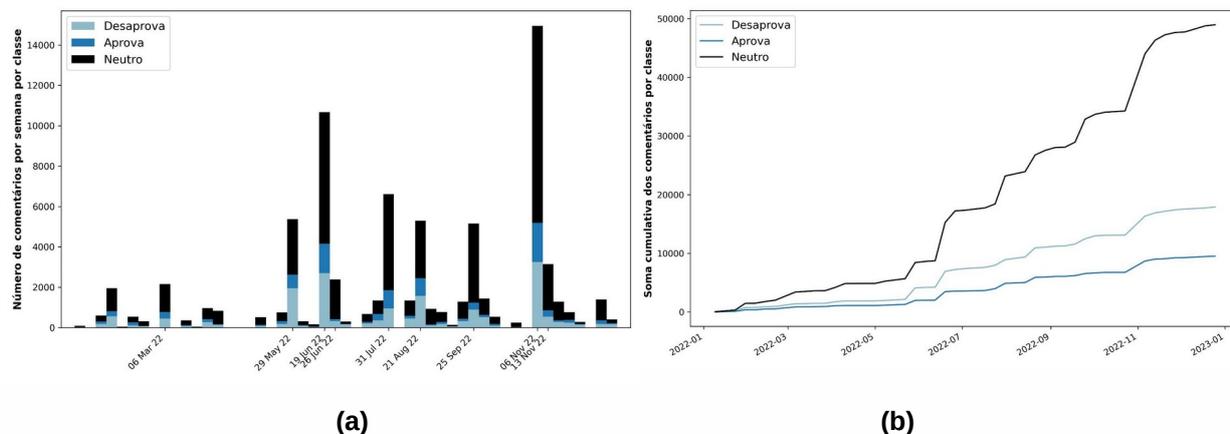


Figura 1. Número semanal (a) empilhado (b) acumulado de comentários classificados em 2022.

#### 4. CONCLUSÃO

Violência e insegurança estão entre os principais problemas dos centros urbanos. Plataformas web vêm sendo utilizadas pelas pessoas para manifestar de diferentes formas o impacto desse problema em suas vidas cotidianamente. Nesse trabalho analisamos o potencial de uso desse tipo de conteúdo para monitorar a opinião pública, inferindo posicionamentos e sentimentos de comentários de usuários no Youtube sobre a atuação policial em incidentes de segurança com repercussão nessa plataforma. Propomos um sistema com módulos de coleta, pré-processamento e inferência desses comentários, baseado em modelos PLN BERT. Nossos experimentos mostram os desafios dessa inferência: impacto de temas fora do contexto alvo, e.g., política, alta neutralidade e sarcasmo observado nos comentários. Não obstante essas questões, foi possível inferir a polaridade dominante dos comentários (excluindo neutralidade) tendendo à desaprovação das operações policiais na maioria dos eventos monitorados mesmo considerando a porcentagem de erro do modelo utilizado. Trabalhos futuros incluem a melhoria do módulo de coleta do sistema visando diminuir impacto de fatores não alvo ao contexto analisado e estimativa mais acurada dos erros do modelo que permitam maior confiança no uso de comentários em plataformas web como Youtube para inferir opinião da população em temas de importância para sociedade como segurança pública.

#### 5. REFERÊNCIAS

- CHAPARRO, Luisa Fernanda et al. Sentiment analysis of social network content to characterize the perception of security. In: Proceedings of the IEEE/ACM International Conference ASONAM. IEEE, 2020. p. 685-691.
- FEITOSA, Marcos Fontes et al. Análise da percepção das pessoas no twitter sobre ações policiais. In: Anais do XI BRASNAM. SBC, 2022. p. 73-84.
- FEITOSA, Marcos Fontes et al. Sentiment Analysis on Twitter Repercussion of Police Operations. In: Proceedings of the Brazilian Symposium on Multimedia and the Web. 2022. p. 84-88.
- HAND, Laura C.; CHING, Brandon D. Maintaining neutrality: A sentiment analysis of police agency Facebook pages before and after a fatal officer-involved shooting of a citizen. Government Information Quarterly, v. 37, n. 1, p. 101420, 2020.



## 6. APOIO

---

Bolsa de Iniciação Tecnológica PIBITI UFPI

