

Marcos Paulo Fontes Feitosa
Orientador: Glauber Dias Gonçalves

**Avaliação de Ferramentas Computacionais para
Análise da Percepção das Pessoas sobre
Violência Urbana**

Picos - PI
3 de março de 2023

Marcos Paulo Fontes Feitosa
Orientador: Glauber Dias Gonçalves

Avaliação de Ferramentas Computacionais para Análise da Percepção das Pessoas sobre Violência Urbana

Monografia submetida ao Curso de Bacharelado em Sistemas de Informação como requisito parcial para obtenção de grau de Bacharel em Sistemas de Informação.

Universidade Federal do Piauí
Campus Senador Heuvídio Nunes de Barros
Bacharelado em Sistemas de Informação

Picos - PI
3 de março de 2023

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

F311a Feitosa, Marcos Paulo Fontes

Avaliação de ferramentas computacionais para análise da percepção das pessoas sobre violência urbana [recurso eletrônico] / Marcos Paulo Fontes Feitosa – 2023.

61 f.

1 Arquivo em PDF

Indexado no catálogo *online* da biblioteca José Albano de Macêdo-CSHNB
Aberto a pesquisadores, com restrições da Biblioteca

Trabalho de Conclusão de Curso (Graduação) – Universidade Federal do Piauí, Bacharelado em Sistemas de Informação, Picos, 2023.

“Orientador: Dr. Glauber Dias Gonçalves”

1. Análise de desempenho. 2. Ferramentas computacionais. 3. Processamento de linguagem natural. 4. Computação social. I. Gonçalves, Glauber Dias. II. Título.

CDD 004.2

AVALIAÇÃO DE FERRAMENTAS COMPUTACIONAIS PARA ANÁLISE DA
PERCEPÇÃO DAS PESSOAS SOBRE VIOLÊNCIA URBANA

MARCOS PAULO FONTES FEITOSA

Monografia apresentada como exigência parcial para obtenção do grau de Bacharel em
Sistemas de Informação.

Data de Aprovação

Picos – PI, 20 de março de 2023

Glauber Dias Gonçalves

Prof. Glauber Dias Gonçalves

Carlos Henrique G. Ferreira

Prof. Carlos Henrique G. Ferreira



Documento assinado digitalmente

RAFAEL TORRES ANCHIETA

Data: 21/03/2023 12:02:21-0300

Verifique em <https://validar.it.gov.br>

Prof. Rafael Torres Anchieta

Agradecimentos

Primeiro agradeço a Deus por permitir que eu tivesse saúde e determinação para não desanimar durante a realização deste trabalho. Não foi uma tarefa simples, mas sempre me mantive firme e determinado a atingir esse objetivo. Hoje me sinto realizado em poder entregar este trabalho pronto, e poder dar orgulho aos meus pais.

Assim deixo meus agradecimentos aos meus pais, Hildete Feitosa Leal e Francisco Fontes Feitosa, que não mediram esforços para que conseguisse atingir meus objetivos. Me incentivaram nos momentos difíceis e compreenderam minha ausência durante a confecção desse trabalho. Agradeço ao meu irmão Francineio Fontes Feitosa, pelo incentivo e amizade. Agradeço a toda minha família e amigos que sempre me deram suporte necessário para atingir os objetivos, e me tornar uma pessoa melhor.

Agradeço ao professor Glauber Dias Gonçalves pela orientação, paciência nos momentos em que não estava bem, e pelos incentivos para continuar trabalhando nessa jornada de pesquisa. Agradeço também ao Carlos Henrique Gomes Ferreira que nos ajudou bastante com todo seu conhecimento na área. Também agradeço a todos os professores da UFPI, todos foram fundamentais para meu aprendizado.

E por último, agradeço aos meus amigos e colegas da UFPI, em especial agradeço ao Pedro Hércules de Sousa Dantas e ao José Maria dos Santos Leal pela amizade incondicional. Pelas conversas descontraídas que faziam esquecer os problemas e as dificuldades, agradeço pela ajuda nas dificuldades da universidade e pelas palavras de incentivos. Esses são os meus mais sinceros agradecimentos.

Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito.

Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes.

Marthin Luther King

Resumo

É cada vez mais frequente o uso de redes sociais virtuais como meio para as pessoas expressarem suas opiniões sobre diferentes aspectos do cotidiano, incluindo violência e insegurança, um problema central a vários centros urbanos. Neste trabalho investigamos o potencial uso de comentários compartilhados em uma rede social bastante popular – o Twitter – para inferir a opinião pública sobre a atuação policial em incidentes de segurança de grande repercussão. Nesse sentido, exploramos atributos extraídos desses comentários e modelos de aprendizado de máquina para inferir o posicionamento e sentimento das pessoas em relação a ações policiais específicas. Nossos experimentos mostram quão desafiante é essa inferência dado grande quantidade de neutralidade e sarcasmo observado em mídias sociais. Não obstante, nossos melhores classificadores alcançaram acurácia e especificidade (macro F1) superiores a 68% para inferir posicionamentos de aprovação, desaprovação e neutralidade da população. Exploramos também a identificação da polaridade de sentimentos das pessoas nesse contexto, e nossos melhores classificadores obtiveram precisão e especificidade (macro F1) superiores a 60%.

Palavras-chaves: Computação Social, Análise de Sentimentos, Análise de Posicionamentos, Processamento de Linguagem Natural, Aprendizagem de Máquina.

Abstract

The use of online social networks is increasingly frequent as a means for people to express their opinions on different aspects of daily life, including violence and security, a central problem in several urban centers. In this work, we investigate the potential use of comments shared on a very popular social network – Twitter – to infer public opinion about police action in violence incidents with great repercussions in digital media. In this sense, we explore features extracted from these comments and machine learning models to infer people’s stances and sentiments in regard to specific actions in this context. Our experiments show how challenging this inference is given the large amount of neutrality and sarcasm observed in social media. Nevertheless, our best classifiers achieved accuracy and specificity (macro F1) greater than 68% to infer the population’s approval, disapproval and neutrality positions. We also explore sentiment polarity identification in this context, our best classifiers obtained precision and specificity (macro F1) superior to 60%.

Lista de ilustrações

Figura 1 – Análise de Posicionamento - Distribuição dos rótulos por notícia. . . .	24
Figura 2 – Análise de Sentimento - Distribuição dos rótulos por notícia.	25
Figura 3 – Análise de posicionamento - Matriz de confusão, modelo BERTimbau com ajuste fino (RN-PtBr).	31
Figura 4 – Análise de sentimento - Matriz de confusão, modelo BERTimbau com ajuste fino (RN-PtBr) e SentiStrength	33

Lista de tabelas

Tabela 1 – Trabalhos Relacionados	20
Tabela 2 – Número de Tweets por Notícia	23
Tabela 3 – Análise de Posicionamentos - Número de Tweets Rotulados por Notícia	24
Tabela 4 – Análise de Posicionamentos - Número de Tweets por Notícia, da amostra de 198 tweets	24
Tabela 5 – Análise de Sentimento - Número de Tweets por Notícia, da amostra de 198 tweets	25
Tabela 6 – Análise de posicionamento - Desempenho dos modelos Multilingual (MultiL) e BERTimbau (PtBr), usando as estratégias de ajuste fino da rede neural (RN) e a da matriz de <i>embedding</i> com os classificadores (SVM e RF). Avaliando com as seguintes métricas: Precisão (P), Revocação (R), F1-score (F1), Acurácia (Acc) e F1-macro.	29
Tabela 7 – Análise de posicionamento - Exemplos de <i>tweets</i> classificados incorretamente pelo modelo BERTimbau com ajuste fino (RN-PtBr).	31
Tabela 8 – Análise de sentimento - Desempenho dos modelos com as métricas: Precisão (P), Revocação (R), F1-score (F1), Acurácia (Acc) e F1-macro. Modelos de rede neural (RN) e classificadores (SVM e RF) utilizam BERT Multilingual (MultiL) e BERTimbau (PtBr).	32
Tabela 9 – Análise de sentimento - Exemplos de <i>tweets</i> classificados incorretamente pelo modelo BERTimbau com ajuste fino (RN-PtBr) e o SentiStrength.	33

Lista de abreviaturas e siglas

PLN	Natural Language Processing (Processamento de Língua Natural)
BERT	Bidirectional Encoder Representations from Transformers (Representações de Codificadores Bidirecionais dos Transformadores)
SVM	Support Vector Machine (Máquina de vetores de suporte)
RF	Random forest (Floresta aleatória)
RN	Neural Network (Rede Neural)
smote	Synthetic Minority Oversampling TEchnique (Técnica de Sobreamostragem Minoritária Sintética)
ACC	Accuracy (Acurácia)
P	Precision (Precisão)
R	Recall (Revocação)
F1	F1 Score

Sumário

1	Introdução	11
1.1	Objetivos	12
2	Referencial Teórico	13
2.1	Análise de Posicionamento	13
2.2	Análise de Sentimento	14
2.3	Métodos Supervisionados	15
2.3.1	BERT	15
2.3.2	SVM	16
2.3.3	RF	17
2.4	Métodos Não Supervisionados	17
2.4.1	SentiStrength	18
3	Trabalhos Relacionados	19
4	Metodologia	21
4.1	Coleta de dados	21
4.2	Pré-processamento e Rotulação	23
4.2.1	Análise de Posicionamentos	23
4.2.2	Análise de Sentimento	25
4.3	Modelos	26
5	Resultados e Discussão	29
5.1	Classificação de Posicionamentos	29
5.2	Análise de Sentimento	31
6	Conclusão	34
7	Publicações	36
	Referências	37
	Apêndices	41
	APÊNDICE A Apêndice	42
	APÊNDICE B Apêndice	56

1 Introdução

A violência causada por altos índices de criminalidade e a sensação de insegurança das pessoas estão dentre os principais problemas dos centros urbanos no mundo. Para se ter uma ideia da gravidade desse problema no Brasil, no ano de 2019, foram registradas 41.726 mortes por crimes violentos, uma taxa média de 20 mortes por mês para cada 100 mil habitantes brasileiros. Essa taxa pode alcançar valores superiores a 40 mortes por mês em estados das regiões norte e nordeste (NEV-USP, 2021).

Diante desses fatos, costumeiramente a população não só brasileira, mas também mundial, vem recorrendo cada vez mais às redes sociais virtuais para manifestar sentimentos e opiniões sobre a atuação dos órgãos e autoridades responsáveis pela segurança pública. Essas manifestações virtuais podem refletir não apenas a percepção negativa sobre a violência urbana, mas também a aprovação ou desaprovação de ações policiais durante o enfrentamento de incidentes de segurança. Além disso, observa-se neutralidade ou mesmo polarização em casos polêmicos que despertam atenção das pessoas, levando a muitos comentários postados e compartilhados publicamente nas redes sociais virtuais (TUCKER et al., 2021; IRANMANESH; ATUN, 2020).

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, principalmente, durante a atuação da polícia. Tipicamente, um ecossistema de opiniões é formado em que, primeiro, as notícias sobre incidentes de segurança envolvendo a polícia são postadas por portais jornalísticos e, posteriormente, são compartilhadas em redes sociais virtuais recebendo milhares de comentários. Esses, por sua vez, geram novos comentários, sejam aprovando, desaprovando ou sem uma posição clara sobre a ação policial (REIS et al., 2015; FERREIRA et al., 2021). Dessa forma, toda essa carga de conteúdo gerada por usuários acerca desses incidentes, potencialmente, possibilita compreender a percepção da população quanto ao esforço dos órgãos de segurança pública no enfrentamento desse problema nas cidades.

Na literatura, esforços anteriores já usaram dados dessa natureza no contexto de segurança para avaliar o sentimento dos comentários (HAND; CHING, 2020; CHAPARRO et al., 2020), e realizar a predição de crimes e taxas criminais (CHEN; CHO; JANG, 2015; TUCKER et al., 2021). Com foco específico na detecção de posicionamentos da população em redes sociais, estudos foram realizados em outros contextos, por exemplo, pandemia de COVID-19 (HOSSAIN et al., 2020; WEINZIERL; HOPFER; HARABAGIU, 2021). Dessa forma, observa-se ainda a falta de estudos e metodologias para inferir como a população reage em face às políticas de segurança pública, em particular, a exploração de conteúdo gerado por usuários em redes sociais sobre um contexto não explorado que são as ações policiais. Especificamente, há a necessidade de entender como conteúdo nesse contexto é

gerado e como ele pode ser quantificado. A partir desse ponto, pode-se investigar o potencial desses dados para inferir posicionamentos e sentimentos, especialmente o quanto a população apoia as ações de segurança pública. No entanto, existem desafios do ponto de vista computacional, principalmente, relacionados ao processamento de linguagem natural. Especificamente, a obtenção de dados rotulados no contexto específico de violência urbana, a representação semântica desses dados considerando o ruído presente em textos de redes sociais virtuais, por exemplo, erros de digitação, uso de gírias e sarcasmo, e o desbalanceamento de opiniões ou posicionamentos das pessoas são alguns dos principais desafios reportados por trabalhos anteriores (WANG et al., 2017; MINAEE et al., 2021).

1.1 Objetivos

Nesse trabalho, investigamos o potencial de uso dos comentários em redes sociais virtuais, em particular a rede social Twitter, para analisar a impressão da população sobre a atuação das entidades responsáveis pela segurança pública. 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 como discutiremos na Seção 3. Contudo, observa-se ainda a falta de metodologias para inferir como a população reage em face às políticas de segurança pública, em particular, ações policiais. Especificamente, há a necessidade de entender como conteúdo nesse contexto é gerado e como ele pode ser quantificado para inferir sentimentos com alta acurácia e especificidade. Nesse sentido oferecemos duas contribuições nesse projeto:

- metodologia para extrair comentários da população sobre ações policiais de interesse;
- análise de posicionamento e sentimento desses comentários comparando resultados de um modelo léxico tradicional e modelos estado da arte de aprendizagem de máquina superficial e profunda;

2 Referencial Teórico

Neste trabalho utilizamos técnicas de Processamento de Linguagem Natural (PLN). PLN é uma área da inteligência artificial cujo objetivo é criar métodos capazes de entender e processar a linguagem humana. Essa área é usada em uma variedade de aplicações, que vão desde assistentes virtuais até chatbots e análise de dados. Dentre as várias aplicações da PLN, neste trabalho focou-se: análise de posicionamento e análise de sentimento, ambas foram utilizadas nesse trabalho. Essas duas técnicas são descritas na sub-seção 2.1 e 2.2 respectivamente. Além disso, explicamos os métodos supervisionados e não supervisionados, nas sub-seções 2.3 e 2.4.

2.1 Análise de Posicionamento

A análise de posicionamento é uma tarefa que busca identificar, de forma automática, a opinião expressa por um autor em relação a um determinado assunto, seja ele uma pessoa, organização, política governamental, movimento, produto ou outro (MOHAMMAD; SOBHANI; KIRITCHENKO, 2017). Por exemplo, é possível inferir a partir do discurso de um político se ele é a favor ou contra o uso de armas por civis. Essa estratégia é amplamente utilizada para analisar postagens em fóruns online, blogs, Twitter, YouTube, Instagram e outras plataformas de mídia social. Essa técnica é especialmente útil para empresas, organizações governamentais e pesquisadores, que podem utilizar os resultados da análise para tomar decisões informadas e desenvolver estratégias mais eficazes. No entanto, é importante ressaltar que a análise de posicionamento não é uma tarefa simples e requer o uso de algoritmos avançados de Processamento de Linguagem Natural (PLN) e aprendizado de máquina. Além disso, é fundamental considerar o contexto em que as opiniões são expressas e a possibilidade de ambiguidade na linguagem utilizada pelos autores.

A tarefa que exploramos é formulada da seguinte forma: dado um texto de *tweet* e uma entidade alvo (pessoa, organização, problema, etc.), os sistemas automáticos de processamento de linguagem natural devem determinar se o *tweet* é a favor do alvo dado, contra o alvo dado, ou se nenhuma das inferências é provável. Por exemplo, o alvo “Direitos humanos” e o *tweet* “Os bandidos também são pessoas com família e amigos. Eles também têm direitos”. Nós humanos podemos inferir que o *tweet* é a favor do dos direitos humanos. Mesmo com a falta de evidências de “favor” ou “contra”, não significa que o *tweet* seja neutro em relação ao alvo. Pode significar apenas que não podemos deduzir a posição do *tweet*. Na verdade, esse é um fenômeno comum. Por outro lado, espera-se que o número de *tweet* dos quais podemos inferir uma postura neutra seja pequena. Como, por exemplo, o alvo “Ação policial” e o *tweet* “O policial usou excesso de força, mas o bandido mereceu”.

Outro ponto importante na tarefa de análise de posicionamento é que o alvo de interesse pode não estar explicitamente mencionado no texto, ou pode não ser alvo de opinião no texto. Por exemplo, o alvo “Ação policial” e o *tweet* “O bandido não teve nem tempo de reação, covardia!”. O alvo da opinião no *tweet* é o bandido, mas o alvo de interesse é a ação policial. No entanto, podemos inferir que o *tweet* provavelmente será desfavorável a ação policial. Observe também que, na detecção de postura, o alvo pode ser expresso de diferentes maneiras, o que afeta se a instância é rotulada como “favor” ou “contra”. Por exemplo, o alvo no primeiro exemplo poderia ter sido formulado como “Ação policial”, caso em que o rótulo correto para essa instância é “contra”. Esses exemplos mostram que a tarefa de identificar o posicionamento das pessoas é extremamente desafiadora.

2.2 Análise de Sentimento

A análise de sentimento, também chamada de mineração de opinião, é o campo de estudo que analisa as opiniões das pessoas, sentimentos, avaliações, atitudes, e emoções em relação a entidades como produtos, serviços, organizações, indivíduos, questões, eventos, tópicos e seus atributos (LIU, 2012). Essa é uma área que tem como tarefa básica a classificação automática de bases de dados contendo opiniões polarizadas (positivas, negativas ou neutras) sobre temas previamente definidos (YU; DUAN; CAO, 2013; PANG; LEE; VAITHYANATHAN, 2002; SOBKOWICZ; KASCHEKY; BOUCHARD, 2012). Com o auxílio da tecnologia da informação, a análise de sentimento reduz drasticamente a necessidade de se lerem grandes quantidades de documentos para extrair opiniões (YU; DUAN; CAO, 2013).

A análise de sentimento possui diversos termos e conceitos na forma de classificar o sentimento. Por exemplo, existe a classificação de polarização, indicando se o texto expressa opiniões positivas, negativas ou neutras em relação ao assunto. Força do sentimento em que apresenta a intensidade do sentimento ou polaridade, sendo também uma forma de saída de alguns métodos. E por fim, existe a classificação de emoção, que indica um sentimento específico presente em um texto (ex.: raiva, surpresa, felicidade, etc.). Para este trabalho escolhemos a classificação de emoção

A Classificação de Emoções expande o conceito de análise de sentimento e tem por objetivo identificar estados afetivos denominados emoções que são projeções de um sentimento, tais como tristeza, angústia ou raiva. No entanto, neste trabalho, associamos as emoções a três classes (Positivo, Neutro, Negativo). Essa modificação pode se assemelhar à classificação de polaridade, entretanto, existem algumas diferenças, por exemplo, a palavra “não” possui polaridade negativa, mas não apresenta uma emoção clara. De maneira que toda palavra pode ter polaridade Positiva, Negativa ou Neutra, mas nem todas representam uma emoção.

Em Fan, Che e Chen (2017), a análise de sentimento pode ser uma ferramenta poderosa

para melhorar o relacionamento com os clientes. O estudo mostrou que as empresas que utilizam a análise de sentimento para monitorar as opiniões dos clientes em relação aos seus produtos e serviços conseguem responder de forma mais rápida e eficaz às suas necessidades. Além disso, a análise de sentimento também pode ser útil na identificação de problemas e oportunidades de melhoria, permitindo que as empresas ajustem seus processos e produtos conforme as demandas dos clientes. No geral, a análise de sentimento é uma área promissora que pode trazer muitos benefícios para as empresas que desejam se manter competitivas em um mercado cada vez mais exigente e dinâmico.

2.3 Métodos Supervisionados

Os métodos supervisionados empregam o termo supervisionado justamente pelo fato de exigir uma etapa de treinamento de um modelo com amostras previamente classificadas. Em outras palavras, essa abordagem utiliza aprendizado de máquina para classificação automática (PANG; LEE; VAITHYANATHAN, 2002). A ideia por trás dessa técnica é que os algoritmos possam aprender com dados, assim identificam padrões que podem ser replicados a novas observações, com isso é possível tomar decisões com o mínimo de intervenção humana. O procedimento para realizar a aprendizagem de máquina compreende quatro etapas principais: (1) obtenção de dados rotulados que serão utilizados para treino e para teste; (2) definição das features ou características que permitam a distinção entre os dados; (3) treinamento de um modelo computacional com um algoritmo de aprendizagem; e (4) aplicação do modelo.

2.3.1 BERT

O BERT (*Bidirectional Encoder Representations from Transformers*) (DEVLIN et al., 2019) é uma arquitetura de aprendizado profundo (do, inglês *Deep Learning*) desenvolvido pelo *Google* para processamento de linguagem natural. BERT possui um modelo pré-treinado com conteúdo do Wikipédia e do Book Corpus, ambos em língua inglesa. Um recurso importante do BERT é sua construção com representações contextuais de modelos pré-treinados, baseado na arquitetura de *transformers*. Através desse recurso, comunidades de pesquisadores ou corporações podem treinar novos modelos adaptados a alguma linguagem no contexto local e distribuí-los como modelos pré-treinados BERT.

O BERT possui inúmeros modelos pré-treinados para os mais diversos tipos de aplicações, nesse trabalho utilizamos os modelos *Multilingual* e o *BERTimbau*. O BERT *Multilingual* é uma versão do BERT treinada em um corpus de texto que contém dados de vários idiomas (DEVLIN et al., 2019). Ele consegue processar texto em vários idiomas, portanto, pode ser usado para tarefas de processamento de linguagem natural em diferentes idiomas. O *BERTimbau*, por outro lado, é uma versão do BERT desenvolvida por pesquisadores da Unicamp, que foi treinada especificamente para a língua portuguesa. Ele

consegue realizar tarefas de processamento de linguagem natural em português e pode ser usado em diversas aplicações, como análise de sentimentos, classificação de texto e geração de texto, entre outras (SOUZA; NOGUEIRA; LOTUFO, 2020).

Esses modelos são pré-treinados para um idioma específico, para utilizá-los em nosso trabalho precisamos realizar uma técnica chamada Fine-tuning. O Fine-tuning BERT envolve o ajuste dos pesos do modelo pré-treinado para acomodar uma tarefa específica, como classificação de sentimento ou análise de posicionamento. O processo de fine-tuning envolve a adição de camadas de classificação ao modelo pré-treinado BERT e treiná-las com dados rotulados de uma tarefa específica. Durante o treinamento, os pesos do modelo pré-treinado e as camadas de classificação são ajustados simultaneamente para minimizar a perda da tarefa. Com essa técnica podemos reduzir significativamente o tempo e recursos necessários para treinar um modelo do zero. Além disso, os modelos fine-tuned com BERT tendem a superar os modelos que não utilizam a técnica, pois o pré-treinamento permite que o modelo tenha uma melhor compreensão do contexto e da semântica da linguagem (SUN et al., 2019).

2.3.2 SVM

SVM (*Support Vector Machine*) é um algoritmo de aprendizado de máquina frequentemente utilizado em tarefas de classificação e regressão proposto por Vapnik em 1995 (CORTES; VAPNIK, 1995). Ele é amplamente aplicado em áreas como análise de imagens, bioinformática e processamento de linguagem natural. A ideia básica do SVM é encontrar um hiperplano que separe as amostras em duas classes, de tal forma que a margem entre os pontos de cada classe seja maximizada. Um estudo interessante sobre SVM é BholaneSavita e Gore (2016). O estudo propõe um modelo de análise de sentimentos baseado em SVM para análise de dados do Twitter. O modelo obteve resultados promissores na classificação de *tweets* como positivos, negativos ou neutros, com uma precisão média de 80,86%.

Para treinar o SVM, primeiro é feita a preparação dos dados: as amostras são preparadas e os recursos são extraídos. Segundo, é feita a seleção das características mais importantes para obter a melhor precisão possível. Terceiro, é iniciado o treinamento do modelo. O SVM treina um modelo com base em um conjunto de treinamento rotulado. O modelo tenta encontrar o hiperplano que maximiza a margem entre as amostras das diferentes classes. Quarto, uma vez que o modelo está treinado, pode ser usado para classificar novos dados de entrada. O modelo classifica uma nova amostra com base na posição da amostra em relação ao hiperplano encontrado no treinamento. E, por fim, é feita a avaliação. O SVM é avaliado em um conjunto de teste. O desempenho é medido usando uma métrica como acurácia, precision, recall ou F1-score.

2.3.3 RF

O RF (*Random Forest*) é um algoritmo de aprendizado de máquina usado para problemas de classificação e regressão, introduzidos por Leo Breiman (BREIMAN, 2001). Ele é uma técnica de ensemble, que combina várias árvores de decisão individuais para produzir um modelo mais robusto e preciso. Um estudo interessante sobre *Random Forest* é o Li et al. (2013). O estudo apresenta uma análise detalhada do algoritmo *Random Forest* e sua aplicação em tarefas de classificação e regressão. Ele também discute várias técnicas para melhorar o desempenho do algoritmo, incluindo ajuste de parâmetros e seleção de recursos. O estudo conclui que *Random Forest* é um algoritmo poderoso e versátil que pode ser usado em uma ampla variedade de aplicações de aprendizado de máquina.

Para treinar o RF primeiro é feito a preparação dos dados, as amostras são preparadas e os recursos são extraídos. Segundo o RF cria várias árvores de decisão aleatórias, onde cada árvore é treinada com um subconjunto aleatório dos dados de treinamento e um subconjunto aleatório dos recursos. Terceiro classificação, o RF usa a técnica de voto para decidir a classe de cada amostra. Cada árvore individual vota em uma classe e a classe com mais votos é escolhida como a classe final. E por fim é feito a avaliação, o modelo RF é avaliado em um conjunto de teste, usando as métricas acurácia, precision, recall ou F1-score.

2.4 Métodos Não Supervisionados

Os métodos de análise de texto não supervisionados geralmente utilizam um dicionário léxico, composto por palavras e seus respectivos sentimentos. Ao contrário das técnicas supervisionadas, que exigem um processo de rotulação e treinamento de dados para criar um modelo, as técnicas não supervisionadas são vantajosas, pois não requerem esse processo, economizando tempo e reduzindo os custos (LIU, 2012). Além disso, a abordagem não supervisionada permite que o modelo não fique preso ao contexto em que foi criado. Existem algumas variações importantes de dicionários léxicos, como aqueles que possuem polaridade do sentimento (negativo, neutro ou positivo, representados por -1, 0 e 1, respectivamente) e aqueles que incluem a intensidade do sentimento (negativo ou positivo, representados por um número entre -1 e 1). Em suma, as técnicas não supervisionadas, como a análise léxica, são uma abordagem eficaz para análise de texto, permitindo a identificação de sentimentos e opiniões expressas em um corpus sem a necessidade de rotulação e treinamento de dados, além de permitir a criação de dicionários léxicos mais precisos e refinados. Nesse trabalho utilizamos o *SentiStrength*(THELWALL, 2017) para análise de sentimento usando a abordagem léxica.

2.4.1 SentiStrength

SentiStrength é uma ferramenta de análise de sentimento de texto que utiliza técnicas de análise léxica para determinar a polaridade dos sentimentos expressos em um texto. Ele atribui um valor de intensidade emocional de -4 a +4 para cada palavra em um texto, dependendo da polaridade do sentimento expresso (THELWALL, 2017). O *SentiStrength* utiliza dois dicionários léxicos, um para polaridade positiva e outro para negativa, e cada palavra é analisada individualmente para determinar sua polaridade e intensidade. Após a análise de cada palavra, o *SentiStrength* computa uma pontuação global que representa a polaridade do sentimento expresso no texto. Essa pontuação pode variar de -4 a +4, dependendo da quantidade e intensidade das palavras com polaridade positiva e negativa encontradas, e se for 0 indica neutralidade. O *SentiStrength* pode ser treinado em diferentes idiomas e contextos, e os usuários também podem personalizar os dicionários léxicos para se adequarem às suas necessidades específicas.

Ao analisar um texto, o *SentiStrength* calcula a polaridade e intensidade do sentimento presente em cada palavra individualmente, e então soma esses valores para obter uma medida geral do sentimento presente no texto. Por exemplo, considere a frase “Eu adorei o filme, foi incrível!”. Ao analisar essa frase, o *SentiStrength* identificaria as palavras “adorei” e “incrível” como sendo positivas, portanto, atribuiria a cada uma delas um valor de sentimento positivo e uma intensidade alta. Por outro lado, a palavra “filme” é neutra e não contribuiria para o cálculo da polaridade do sentimento. Ao somar os valores de sentimento e intensidade de cada palavra, o *SentiStrength* concluiria que a frase tem um sentimento geral, positivo e intenso.

Outro exemplo seria a frase “Não gostei do atendimento na loja, foi péssimo”. Nesse caso, o *SentiStrength* identificaria as palavras “não gostei” e “péssimo” como negativas e atribuiria a cada uma delas um valor de sentimento negativo e uma intensidade alta. A palavra “atendimento” é neutra, mas a expressão “na loja” pode indicar um contexto específico que modifica o sentimento geral do texto. Ao somar os valores de sentimento e intensidade de cada palavra, o *SentiStrength* concluiria que a frase tem um sentimento geral, negativo e intenso.

3 Trabalhos Relacionados

As plataformas de mídia social são atualmente um importante fórum para as pessoas expressarem suas opiniões e trocarem informações. Ao interagir entre si por meio de postagens, re-postagens, respostas ou menções, os usuários favorecem a disseminação de informações (AL-GARADI et al., 2018). Assim, o crescente uso dessas plataformas chama a atenção de vários pesquisadores que visam modelar e analisar o comportamento dos usuários em face de fenômenos reais. Em particular, análises utilizando ferramentas de processamento natural de linguagem tem auxiliado na compreensão de diversos aspectos, por exemplo, relacionados a política (FERREIRA et al., 2021; NOBRE; FERREIRA; ALMEIDA, 2022), pandemia (MALAGOLI et al., 2021) e percepção da segurança urbana (HAND; CHING, 2020; OGLESBY-NEAL; TIRY; KIM, 2019). Nesta seção nós discutimos alguns destes esforços mais próximos deste trabalho.

Especificamente no contexto da percepção de segurança pelas pessoas manifestada em plataformas de mídias sociais, técnicas para análise de sentimentos vem sendo até então a principal forma de análise nesse tema. Em Hand e Ching (2020) foram analisados os sentimentos dos comentários das pessoas às postagens das agências de polícia em suas páginas no Facebook antes e após incidentes com tiros envolvendo policiais. Os autores observaram tendência à neutralidade logo após o incidente, o que pode indicar pouca eficiência dessa técnica para analisar percepção de segurança. Os autores de Oglesby-Neal, Tiry e Kim (2019) focaram em uma ação policial que resultou na morte do afro-americano Freddie Gray na cidade de Baltimores em 2015. Por sua vez, os autores de Chaparro et al. (2020) utilizaram *tweets* georreferenciados em Bogotá, na Colômbia para identificar o sentimento da população sobre a segurança nessa cidade. Cada *tweet* foi rotulado como um sentimento positivo ou negativo por um grupo de especialistas e os autores observaram que técnicas de aprendizagem de máquina supervisionada obtiveram acurácia e especificidade melhores que técnicas baseadas em regras léxicas.

Outro contexto bastante explorado é a predição de crimes a partir de características extraídas do conteúdo gerado por usuários em mídias sociais. Por exemplo, em Wang, Gerber e Brown (2012) *tweets* postados por agências de notícias foram utilizados para identificar possíveis crimes. Já em Chen, Cho e Jang (2015), os autores usam *tweets* geo-localizados com foco na reincidência de crimes em uma dada região. Alguns trabalhos mesclam fontes de dados oficiais a dados de redes sociais para analisar o quanto, postagens do Twitter estão correlacionados com a violência urbana (TUCKER et al., 2021; IRANMANESH; ATUN, 2020). Em Wang et al. (2017), Sousa, Feitosa e Gonçalves (2021) foi investigado o potencial de pontos de interesse gerados por usuários em serviços de mapeamento para predição das taxas criminais.

Mais recentemente, várias das tarefas supramencionadas têm sido realizadas por estra-

tégias mais avançadas para o processamento de linguagem natural, como BERT (DEVLIN et al., 2019), SentenceBERT (REIMERS et al., 2019), e outras variações de arquiteturas *transformers*. De fato, estes modelos fornecem representações mais ricas, portanto, com resultados potencialmente melhores. Focando no contexto do Twitter, alguns esforços devem ser mencionados. Em Hossain et al. (2020), os autores utilizaram vários modelos pré-treinados para identificar o posicionamento dos usuários quanto a conceitos errôneos sobre a COVID-19 no Twitter. Os resultados alcançados que, no melhor modelo, apresenta um *F1-score* de 50.2 evidenciam os desafios dessa tarefa. Considerando o mesmo conjunto de dados e tarefa, o trabalho de Weinzierl, Hopfer e Harabagiu (2021) conseguiu aumentar o valor do *F1-score* para 74.3. Ainda no contexto da pandemia, Glandt et al. (2021) realizaram a análise do desempenho dos classificados por instância, enfatizando que os resultados variam conforme a instância. Em geral, os autores observaram *F1-score* que variam de 0.53 a 0.83. Assim, eles reforçam o argumento de que é difícil ter um único modelo que seja melhor para todos os tipos de instâncias analisados, ou seja, que capture todas as nuances presentes em textos de redes sociais online. Esses resultados refletem o atual estado da arte, como mostram algumas revisões recentes sobre tal tarefa em dados de redes sociais (MINAEE et al., 2021). A Tabela 1 apresenta uma comparação entre os trabalhos, com o contexto usado em cada um deles e a técnica utilizada, análise de posicionamento e de sentimento.

Tabela 1 – Trabalhos Relacionados

Trabalhos Relacionados	Análise de Posicionamento	Análise de Sentimento	Contexto
Nobre, Ferreira e Almeida (2022)	Não	Não	Política
Weinzierl, Hopfer e Harabagiu (2021)	Sim	Não	Covid19
Ferreira et al. (2021)	Não	Sim	Política
Glandt et al. (2021)	Sim	Sim	Covid19
Malagoli et al. (2021)	Não	Sim	Covid19
Tucker et al. (2021)	Não	Não	Taxas de Crimes
Sousa, Feitosa e Gonçalves (2021)	Não	Não	Taxas de Crimes
Chaparro et al. (2020)	Não	Sim	Segurança
Hossain et al. (2020)	Sim	Não	Covid19
Hand e Ching (2020)	Não	Sim	Segurança
Oglesby-Neal, Tiry e Kim (2019)	Sim	Sim	Segurança
Chen, Cho e Jang (2015)	Não	Sim	Previsão do Crimes
Wang, Gerber e Brown (2012)	Não	Sim	Previsão do Crimes
Este Trabalho	Sim	Sim	Percepção sobre ações policiais

4 Metodologia

Nesta seção descrevemos a metodologia para o desenvolvimento desse trabalho. Primeiramente, descrevemos a metodologia de coleta de dados. A seguir, descrevemos os passos para o seu pré-processamento e rotulação. Por fim, descrevemos os modelos utilizados para inferência de posicionamentos e sentimentos, assim como as suas respectivas estratégias de treinamento.

4.1 Coleta de dados

Realizamos a coleta de dados em duas etapas. Primeiramente, selecionamos notícias de grande repercussão sobre o contexto desse trabalho em portais de mídia digital. Em seguida, coletamos comentários das pessoas em redes sociais (*tweets*) sobre essas notícias como descrito a seguir.

Inicialmente, selecionamos notícias sobre incidentes de segurança com intervenção policial entre o período de 04/2019 até 12/2021 que foram amplamente divulgadas e repercutidas em mídia digital. Nesse sentido, pesquisamos por notícias nas seções sobre violência e crimes em três relevantes portais de notícias nacionais: UOL¹, Folha de São Paulo², e G1³. Utilizamos como critério de seleção a participação dos leitores via comentários de texto no próprio portal de notícia. Especificamente, selecionamos notícias com mais de 700 comentários, marca notável e acima da média usual de comentários e intuitivamente indica uma grande repercussão em redes sociais. Ao todo foram selecionadas 8 notícias, que listamos a seguir com um identificador, e as informações que permitem encontrá-las, isto é, o portal de origem e o título da notícia.

- **N_1**, G1: “Sequestrador de ônibus é morto por atirador de elite na Ponte Rio-Niterói, os 39 reféns passam bem”; UOL: “Sequestrador é morto, polícia libera reféns e encerra sequestro de ônibus”; Folha de São Paulo: “PM mata homem que manteve passageiros de ônibus reféns na ponte Rio-Niterói”.
- **N_2**, G1: “Menina de 8 anos morre baleada no Complexo do Alemão”; UOL: “Menina de 8 anos morre baleada após operação policial no Complexo do Alemão”; Folha de São Paulo: “Menina de oito anos morre baleada no Rio de Janeiro”.
- **N_3**, G1: “Mulher é imobilizada por PMs com bebê no colo em Itabira, MG”; UOL: “Policial imobiliza mulher com criança no colo em Itabira (MG)”; Folha de São Paulo: “PMs derrubam e imobilizam mulher com bebê no colo em Itabira (MG)”.

¹ <https://www.uol.com.br/>

² <https://www.folha.uol.com.br/>

³ <https://g1.globo.com/>

- N_4, G1: “Lázaro Barbosa morre após ser preso em Goiás”; UOL: “Lázaro é morto em Goiás; policiais comemoram após carregar corpo”; Folha de São Paulo: “Lázaro Barbosa, o serial killer do DF, é morto pela polícia após 20 dias de buscas”;
- N_5, G1: “Menino de 14 anos morre durante operação das polícias Federal e Civil no Complexo do Salgueiro, RJ”; UOL: “Adolescente João Pedro é morto em operação no Rio, família critica polícia”; Folha de São Paulo: “Menino de 14 anos é morto em casa durante ação da PF no Rio”.
- N_6, G1: “Menino de 7 anos morre após ser baleado na porta de casa na Baixada Fluminense”; UOL: “Menino de 7 anos morre após ser baleado enquanto brincava na porta de casa”; Folha de São Paulo: “Menino de 7 anos morre na porta de casa após tiroteio na Baixada Fluminense”.
- N_7, G1: “Homem morre após ser baleado em ação do Exército na Zona Oeste do Rio”; UOL: “Exército dispara 80 tiros em carro de família no Rio e mata músico”; Folha de São Paulo: “Exército dispara 80 tiros em carro de família no Rio e mata músico”.
- N_8, G1: “Operação no Jacarezinho deixa 28 mortos, provoca intenso tiroteio e tem fuga de bandidos”; UOL: “Operação com 25 mortos no Jacarezinho é a mais letal da história do Rio”; Folha de São Paulo: “Polícia faz operação mais letal da história do RJ, com ao menos 25 mortos”.

A próxima etapa da coleta de dados consiste em obter comentários das pessoas sobre as notícias acima relacionadas em redes sociais. Especificamente, coletamos dados na rede social Twitter. Optamos por essa plataforma devido às suas ferramentas disponíveis para pesquisadores, que permitem a coleta eficiente e eficaz de dados. Além disso, o Twitter é uma rede social baseada em texto, e a maioria dos *tweets* contém conteúdo escrito em até 280 caracteres atualmente. Isso torna a coleta de dados mais fácil e mais direta, pois podemos usar os títulos e links de cada notícia para procurar e coletar *tweets* relevantes.

Para coletar *tweets*, utilizamos a API do Twitter versão 2 com um nível de acesso acadêmico, o Twitter Academic Research. Esse nível de acesso oferece aos pesquisadores um conjunto mais amplo de dados do Twitter para fins de pesquisa, porém existem algumas limitações importantes a serem consideradas. Por exemplo, é possível buscar apenas 250 *tweets* por segundo e até 10 milhões de *tweets* por mês. Além disso, o Twitter impõe restrições de privacidade, uso e geográficas que devem ser respeitadas pelos pesquisadores. Essas limitações foram cuidadosamente consideradas durante a confecção deste trabalho para garantir que todas as práticas fossem éticas e dentro das diretrizes estabelecidas pelo Twitter.

A API do Twitter funciona por requisições HTTP, entretanto é um pouco esgotante fazer as requisições de forma manual, para isso utilizamos a biblioteca Tweepy⁴. O Tweepy é uma biblioteca em Python que permite aos desenvolvedores acessar e utilizar a API do Twitter de maneira mais fácil e eficiente. Através do Tweepy, é possível realizar operações como autenticação, pesquisa, coleta e publicação de *tweets*. Com o Tweepy, é possível buscar *tweets* com base em palavras-chave, filtrar *tweets* por localização geográfica, coletar informações de perfil de usuários, postar *tweets*, entre outras funcionalidades. Com essa ferramenta coletamos um total de 16 276 *tweets* sobre todas as notícias, esse total já com os *tweets* pré-processado. A Tabela 2, mostrar o número de tweets para cada notícia.

Tabela 2 – Número de Tweets por Notícia

Notícia	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8
Número de Tweets	731	521	997	642	5872	2092	1211	4209

4.2 Pré-processamento e Rotulação

Conduzimos o mínimo de pré-processamentos de textos nos *tweets* coletados, seguindo recomendações de outros trabalhos que lidaram com inferência via processamento de linguagem natural (MOZAFARI; FARAHBAKHS; CRESPI, 2019). Assim, removemos dados que dificultam esse processamento como *links* e quebras de linha. Adicionalmente, desconsideramos *tweets* com apenas uma palavra, dado o menor potencial de inferência semântica desses. Por outro lado, mantivemos todas as *stop words* para treinar modelos de inferência com a mesma sequência em que as palavras aparecem nos *tweets*. Todas essas tarefas de pré-processamento foram realizadas com a biblioteca NLTK⁵ na linguagem Python.

4.2.1 Análise de Posicionamentos

Após o pré-processamento, realizamos a etapa de rotulação das instâncias de dados. O rótulo consiste na classificação de um *tweet* como *Aprova*, *Desaprova* ou *Neutro* considerando o posicionamento do usuário que o postou sobre a ação policial no incidente ao qual a notícia se refere. Em outras palavras, o *tweet* com o rótulo *Aprova* significa que o usuário aprova a ação policial, *Desaprova* significa que o usuário não aprova a ação, *Neutro* expressa que o usuário é indiferente ao assunto ou não manifesta claramente seu posicionamento. A rotulação serve de base para treinamento de modelos de inferência realizada por humanos, conforme usual na literatura. Nesse sentido, contamos com três pessoas (não especialistas⁶) para rotular um subconjunto de *tweets*, dado a inviabilidade

⁴ <https://www.tweepy.org/>

⁵ <https://www.nltk.org/>

⁶ Entendemos que o contexto das notícias é conhecido pela população amplamente, sendo desnecessária rotulação por um especialista.

de rotulação da base de dados completa. Esse subconjunto é composto por 4467 *tweets* definidos aleatoriamente e proporcionalmente ao total de *tweets* por notícia. A Tabela 3, mostrar o número de *tweets* para cada notícia.

Tabela 3 – Análise de Posicionamentos - Número de Tweets Rotulados por Notícia

Notícia	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8
Número de Tweets	1432	636	470	468	461	375	375	250

As informações sobre a distribuição dos rótulos para todos os *tweets* rotulados por notícia são mostradas na Figura 1. A distribuição das classes é fortemente desbalanceada, contendo principalmente *tweets* com postura neutra que representam 49% do conjunto total, seguidos por *tweets* da classe *desaprova* representando 27% do conjunto total, e por fim, é a classe *aprova* com 24%. Esse desbalanceamento, principalmente, em direção a classe neutra torna a inferência de posicionamentos mais difícil como já observado em outros trabalhos da literatura (MOZAFARI; FARAHBAKHS; CRESPI, 2019; HAND; CHING, 2020). Além disso, nossa proposta é ainda mais desafiadora, pois lidamos com comentários em português brasileiro em contextos locais envolvendo incidentes de violência envolvendo ação policial, em que nem sempre as ferramentas e modelos apresentam o melhor resultado.

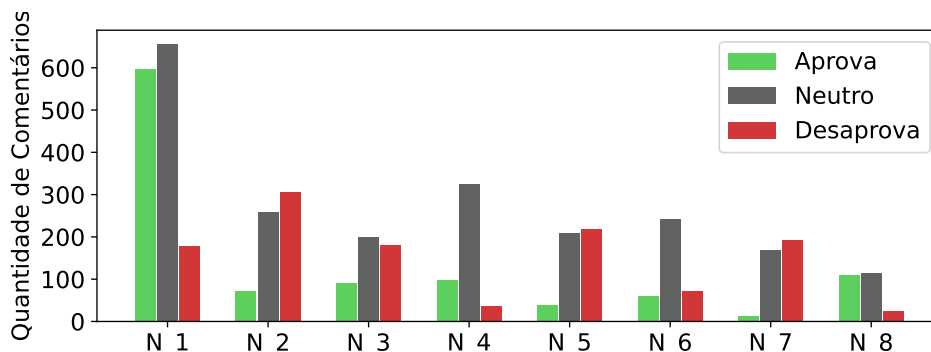


Figura 1 – Análise de Posicionamento - Distribuição dos rótulos por notícia.

Avaliamos a qualidade da rotulação via métricas de concordância entre os três rotuladores. Nesse sentido, selecionamos outro subconjunto de 198 *tweets* aleatórios com rotulação das três pessoas, a Tabela 4 apresenta a distribuição dos *tweets* por notícia para essa amostra. O percentual de concordância entre os rotuladores para essa amostra de 198 *tweets* foi de 70,2%, ao passo que o índice *Fleiss Kappa* foi de 0,67, indicando uma concordância substancial entre os pesquisadores em comparação a outros trabalhos da literatura (HOSSAIN et al., 2020).

Tabela 4 – Análise de Posicionamentos - Número de Tweets por Notícia, da amostra de 198 tweets

Notícia	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8
Número de Tweets	17	22	29	23	32	29	24	22

4.2.2 Análise de Sentimento

A etapa de rotulação foi aplicada também para analisar o sentimento, entretanto a rotulação muda em relação à análise de posicionamento. O rótulo consiste na classificação de um *tweet* como Positivo, Negativo ou Neutro, considerando o sentimento do usuário que o postou sobre a ação policial no incidente ao qual a notícia se refere. Seguimos a estratégia proposta, onde o *tweet* Positivo usa uma linguagem positiva em geral relacionada a otimismo, felicidade, orgulho e apoio. *Tweet* negativo usa linguagem de tristeza, angústia, medo, incluindo ridicularização, sarcasmo e ódio. No que lhe concerne, *tweet* neutro não tem sentimento claro, sendo nenhuma ou ambas opções anteriores.

Na rotulação do *tweets* para análise de sentimento usamos um conjunto de dados composto por 3000 *tweets* definidos aleatoriamente e proporcionalmente ao total de *tweets* por notícia (375 *tweets* por notícia). Cada voluntário rotulou 1000 *tweets*, obtendo o total de 3000 *tweets* rotulados. As informações sobre a distribuição dos rótulos por notícia e por *tweet* são apresentadas na Figura 2.

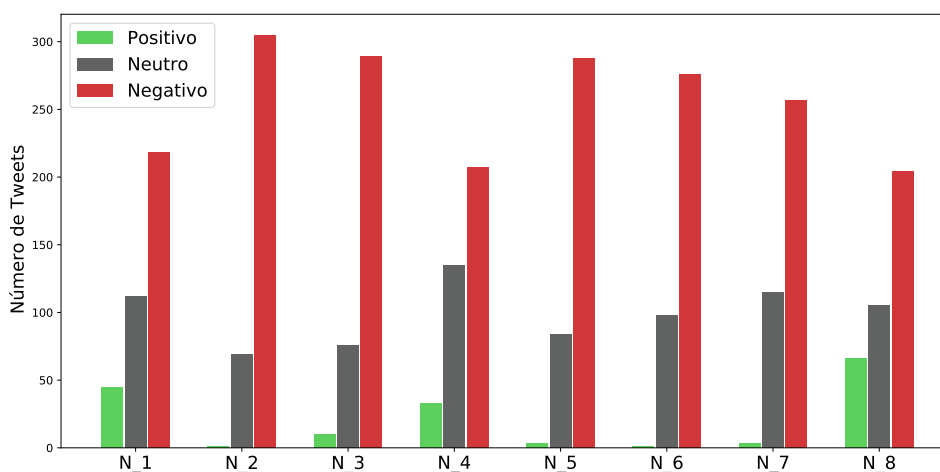


Figura 2 – Análise de Sentimento - Distribuição dos rótulos por notícia.

Avaliamos a qualidade da rotulação via métricas de concordância entre os três rotuladores. Nesse sentido, selecionamos uma amostra de 198 *tweets* a partir de um subconjunto aleatório de *tweets* rotulados. Voluntários foram solicitados a rotular cada instância dessa amostra, resultando em um percentual de concordância entre os rotuladores de 70,2% e um índice Fleiss Kappa de 0,58, indicando uma concordância moderada entre os pesquisadores.

Tabela 5 – Análise de Sentimento - Número de Tweets por Notícia, da amostra de 198 tweets

Notícia	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8
Número de Tweets	24	23	25	22	28	29	24	23

4.3 Modelos

Neste trabalho utilizamos dois modelos pré-treinados que são o *Multilingual* e o *BERT-Timbau*. O modelo *Multilingual* é pré-treinado em 102 idiomas, e possui 12 camadas (blocos de *transformers*), 12 cabeçotes de atenção e 110 milhões de parâmetros. No que lhe concerne, BERTimbau é pré-treinado na língua portuguesa do Brasil, e em nossas avaliações utilizamos a versão desse modelo com a maior quantidade de codificadores (*large*), contendo 24 camadas, 16 cabeçotes de atenção e 335 milhões de parâmetros. O BERTimbau apesar de ser treinado em uma língua complexa como portuguesa, já se mostrou eficiente em tarefas de processamento de linguagem natural que requerem reconhecimento de entidade nomeada, semelhança textual de sentença e reconhecimento de enlace textual (SOUZA; NOGUEIRA; LOTUFO, 2020).

Utilizamos matrizes de *word embeddings* geradas pelos modelos baseados em BERT para treinar classificadores supervisionados na inferência de posicionamento de *tweets*. Em suma, *Word Embedding* é uma forma de representar palavras por números para tarefas de processamento de linguagem natural. Essa representação é normalmente na forma de um vetor de valores reais, que representa o significado das palavras conforme o contexto e o significado da sentença, i.e., o *tweet* em que a palavra está inserida. Para gerar essa representação, utilizamos os modelos pré-treinados acima descritos. O BERTimbau *large* gera um vetor de tamanho 1024 para cada sentença, ao passo que o Multilingual gera um vetor de tamanho 768. Ao fim temos uma matriz de *word embeddings* em que cada linha representa um *tweet* e cada coluna representa uma característica do *tweet*.

A matriz de *words embeddings* foi utilizada como as características de um conjunto de *tweets* rotulados para treinar dois classificadores populares que são *Random Forest* (RF) e *Support Vector Machine* (SVM). Para treinar os classificadores dividimos a matriz de embeddings em treino e teste de forma aleatória, com 80% e 20% dos *tweets* respectivamente, e treinamos os classificadores sem e com balanceamento de classes utilizando a técnica *smote* (MAHESH, 2020).

Além dos modelos de classificação acima, avaliamos outra abordagem baseada em ajustes finos diretamente sob os modelos BERT. Especificamente, adicionamos uma nova camada de neurônios aos modelos BERT pré-treinados e a treinamos para nossa tarefa de classificação com os *tweets* rotulados. Esse ajuste tem a vantagem de acrescentar novas informações aos modelos, que já codificam muitas informações sobre a língua em que foram pré-treinados. Dessa forma pode-se obter melhores resultados para inferência de posicionamentos com menor quantidade de dados rotulados para treinamento, apesar do tempo e complexidade para construção de mais uma camada na rede neural do BERT. Os códigos e a base de dados referentes aos classificadores e os modelos BERT estão disponíveis no GitHub, tanto para análise de posicionamento ⁷, quanto para análise de

⁷ https://github.com/LABPAAD/crimes_stance

sentimento ⁸.

Para implementar a técnica de ajuste fino, usamos a biblioteca Pytorch⁹, com ela podemos adicionar a camada não treinada de neurônios e treinarmos o novo modelo. Como ambiente de implementação utilizamos uma máquina virtual da AWS do tipo G4dn que são instâncias que usam GPUs e são otimizadas para inferência de *machine learning*, a máquina utilizada possui 1 GPU, 8 vCPUs, 32G de RAM e 225 NVMe SSD. Com base em nossos experimentos, treinamos o BERTimbau Large com um Batch Size de 8 um Learning rate de 1e-5, e o Multilingual Cased com um Batch Size de 64 um Learning rate de 3e-5, para cada modelo utilizamos 10 épocas e o otimizador Adam com os respectivos Learning rate.

Para análise de sentimento existem vários métodos usando a abordagem não supervisionada, como: LIWC (PENNEBAKER; BOOTH; FRANCIS, 2001), SenticNet (CAMBRIA; HUSSAIN; HAVASI, 2010), Sentwordnet (ESULI; SEBASTIANI, 2006), Panas (WATSON; CLARK, 1988), *SentiStrength*(THELWALL, 2017), AFINN (NIELSEN, 2011), Vader (HUTTO; GILBERT, 2014), Sentiment140 (GO; BHAYANI; HUANG, 2009), Emolex (CROWDFLOWER, 2015), Opinionlexicon (HU; LIU, 2004). Entretanto, nesse trabalho utilizamos apenas o *SentiStrength*, pois em nossos experimentos esse método foi o que teve o melhor desempenho. *SentiStrength* usa uma abordagem lexical que explora uma lista de itens relacionados ao sentimento, com regras para lidar com métodos linguísticos e sociais da *web* padrão para expressar sentimento, como emoticons, pontuação exagerada e erros ortográficos deliberados (THELWALL, 2017). Para cada texto, o *SentiStrength* emite um valor entre -4 a 4, entretanto, não consideramos a intensidade do sentimento, definimos apenas que valores de 1 até 4 indica sentimento positivo, e valores de -1 até -4 indica sentimento negativo, e 0 o sentimento neutro.

E por fim, como nossos dados rotulados apresentados na seção 4.2, temos um problema de desbalanceamento de classes, com isso os métodos supervisionados não funcionam bem com dados desequilibrados. Para tentar solucionar esse problema testamos o *smote*. O *smote* significa Synthetic Minority Over-sampling Technique e é uma técnica usada em problemas de desequilíbrio de classes em aprendizado de máquina (CHAWLA et al., 2002). O desequilíbrio de classe refere-se a um problema em que uma das classes em um conjunto de dados é significativamente menor do que a outra. A técnica *smote* trabalha sinteticamente gerando novos exemplos da classe minoritária, para que a distribuição das classes no conjunto de dados fique mais equilibrada. O *smote* funciona selecionando exemplos da classe minoritária e criando exemplos sintéticos que são uma combinação dos exemplos selecionados. O processo começa selecionando aleatoriamente um exemplo da classe minoritária e encontrando seus k vizinhos mais próximos. A partir desses vizinhos, o *smote* cria exemplos sintéticos interpolando entre o exemplo original e seus vizinhos, ou

⁸ https://github.com/LABPAAD/sentiment_analysis

⁹ <https://pytorch.org/>

seja, gerando novos pontos entre eles.

5 Resultados e Discussão

Nesta seção analisamos o desempenho de modelos BERT, usando tanto a estratégia de ajuste fino com os modelos pré-treinados BERTimbau e Multilingual, como a matriz de Embedding com os classificadores SVM (Suport Vector Machine) e RF (*Random Forest*), para inferências sobre o posicionamentos e sentimentos da população sobre atuação policial em incidentes de violência.

5.1 Classificação de Posicionamentos

Na Tabela 6, apresentamos os resultados para a avaliação de desempenho dos classificadores treinados sem e com o tratamento para desbalanceamento de classes (*smote*), além da rede neural BERT com ajuste fino. Para cada um desses modelos, apresentamos os resultados via cinco métricas de desempenho. Nota-se que a acurácia (*acc*) é a métrica, majoritariamente, com maior desempenho para as classes desaprova e aprova. Contudo, é importante observar que os dados analisados são desbalanceados e essas classes são minoritárias na maior parte das notícias analisadas (i.e., N_1, N_3, N_4, N_6 e N_8). Nesse caso, as métricas que capturam a especificidade (precisão, revocação e F1) dos modelos em identificar corretamente a classes minoritárias ganham importância, e essas métricas são o foco da avaliação de desempenho nessa seção.

Tabela 6 – Análise de posicionamento - Desempenho dos modelos Multilingual (MultiL) e BERTimbau (PtBr), usando as estratégias de ajuste fino da rede neural (RN) e a da matriz de *embedding* com os classificadores (SVM e RF). Avaliando com as seguintes métricas: Precisão (P), Revocação (R), F1-score (F1), Acurácia (Acc) e F1-macro.

Modelos	Desaprova			Aprova			Neutro			Acc	F1-macro
	P	R	F1	P	R	F1	P	R	F1		
SVM-MultiL-smote	0,46	0,53	0,49	0,48	0,59	0,53	0,76	0,63	0,69	0,59	0,57
SVM-MultiL	0,51	0,49	0,50	0,55	0,47	0,51	0,71	0,77	0,74	0,63	0,58
RF-MultiL-smote	0,50	0,48	0,49	0,57	0,50	0,54	0,72	0,77	0,74	0,63	0,59
RF-MultiL	0,52	0,34	0,41	0,61	0,38	0,47	0,64	0,86	0,73	0,62	0,54
SVM-PtBr-smote	0,52	0,65	0,58	0,55	0,66	0,60	0,84	0,66	0,74	0,66	0,64
SVM-PtBr	0,54	0,58	0,56	0,66	0,61	0,63	0,78	0,77	0,77	0,69	0,66
RF-PtBr-smote	0,55	0,56	0,56	0,68	0,56	0,61	0,74	0,79	0,77	0,68	0,65
RF-PtBr	0,55	0,44	0,49	0,75	0,49	0,59	0,69	0,86	0,76	0,67	0,61
RN-PtBr	0,58	0,63	0,61	0,69	0,68	0,69	0,77	0,74	0,76	0,70	0,68
RN-MultiL	0,49	0,55	0,52	0,61	0,53	0,56	0,71	0,71	0,71	0,63	0,60

Primeiramente, discutimos os resultados do tratamento de desbalanceamento das classes para os classificadores RF e SVM. Sem esse tratamento, há um viés para precisão nas classes minoritárias, pois apenas *tweets* com posicionamentos claros de aprovação ou desaprovação serão classificadas como tal. Isso pode ser observado na Tabela 6, com valores de precisão majoritariamente maiores que revocação, levando a baixa especificidade e tendência a classificar posicionamentos como neutros. Ao tratarmos o desbalanceamento com

smote, observamos que SVM e RF, em geral, aumentam o desempenho para revocação e, por consequência, aumentam o acerto para as posturas de aprovação e desaprovação. A métrica F1 é a referência para o melhor compromisso entre precisão e revocação, e obtivemos F1 maiores para os classificadores com *smote* (exceção apenas de SVM-MultiL para Aprova e SVM-PtBR para Desaprova). Observa-se também que o SVM, em geral, tem o melhor desempenho alcançando F1 melhores (49-63%) com ou sem balanceamento, superando as marcas de F1 do classificador RF (41-61%).

Agora focamos no aspecto da língua em que os modelos BERT são pré-treinados, i.e., multi-lingual (MultiL) e português brasileiro (PtBr). Para isso utilizamos a métrica F1-macro que considera a média do F1 para as três classes de posicionamentos. Observa-se que o pré-treinamento na língua local dos *tweets* (português brasileiro) tem ganhos de desempenho notáveis para os três modelos avaliados. Ao comparar os ganhos dos modelos PtBr em relação a MultiL, os classificadores RF e SVM alcançam aumentos em F1-macro de até 7 e 8 pontos percentuais respectivamente, ao passo que a rede neural ajustada (RN) tem um aumento de 8 pontos percentuais. Isso evidencia que o treinamento de modelos com textos no domínio e língua específicas do contexto em investigação são aspectos fundamentais para inferências de posicionamentos. Essa observação corrobora com outros trabalhos recentes que também utilizam *tweets* para inferências de posicionamentos, mas em contextos diferentes ao investigado nesse trabalho (HOSSAIN et al., 2020).

Finalmente, discutimos o uso da rede neural BERT com ajuste fino (RN). Como esperado, esse modelo obteve o melhor resultado. Especificamente, o modelo pré-treinado em português brasileiro (RN-PtBr) alcançou o F1-macro de 68% e uma acurácia de 70%. É importante mencionar, todavia, que o desempenho da rede neural não é tão superior ao melhor classificador (SVM-PtBr). O ganho do RN-PtBr corresponde a apenas 2 pontos percentuais sobre o F1-macro do SVM-PtBr. Considerando o custo computacional para ajustar a rede neural BERT (i.e., o dobro do tempo dos classificadores), o modelo SVM apresenta um bom compromisso entre desempenho e custo de treinamento.

A Figura 3 reporta a matriz de confusão para o modelo RN-PtBr que obteve o melhor desempenho, mostrado na Tabela 6. Cada linha representa os *tweets* em uma classe real, enquanto cada coluna representa os *tweets* em uma classe prevista, o que nos permite analisar onde o classificador mais erra e como isso acontece em função da classe.

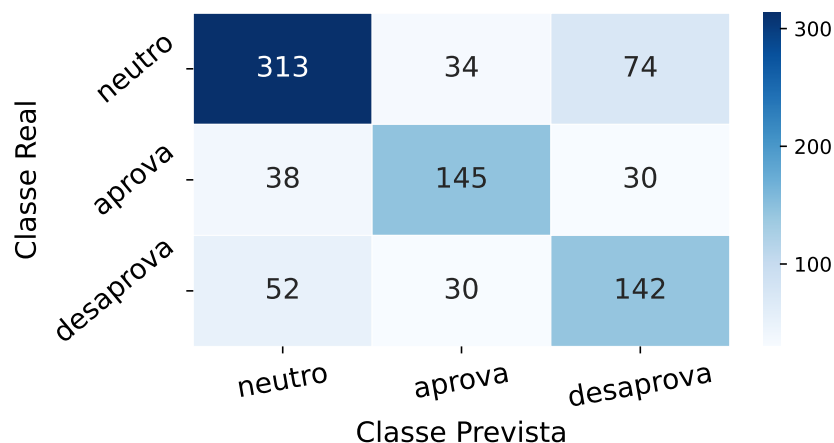


Figura 3 – Análise de posicionamento - Matriz de confusão, modelo BERTimbau com ajuste fino (RN-PtBr).

De forma geral, é possível notar que, em termos absolutos, o maior desafio na tarefa aqui endereçada está em diferenciar as classes *Neutro* da *Desaprova* e vice-versa. Para exemplificarmos a complexidade da tarefa, nós selecionamos alguns casos a partir da matriz de confusão focando especificamente nessas duas classes e apresentamos na Tabela 7.

Tabela 7 – Análise de posicionamento - Exemplos de *tweets* classificados incorretamente pelo modelo BERTimbau com ajuste fino (RN-PtBr).

Tweet	Real	Predito
Eu já vi isso mano, é osso	Neutro (0)	Desaprova (-1)
Meu Deus que realidade cruel, que os passageiros fiquem bem	Neutro (0)	Desaprova (-1)
Isso foi de propósito?	Neutro (0)	Desaprova (-1)
É esse o nível de quem vai nos salvar do comunismo quero saber quem vai nos salvar deles	Desaprova (-1)	Neutro (0)
#ficaemcasa e leve um tiro	Desaprova (-1)	Neutro (0)
Não é saudável e nem natural se acostumar com tragédias, as pessoas estão doentes!	Desaprova (-1)	Neutro (0)

Os exemplos apresentados mostram como, *tweets* curtos, contendo algum tipo de sarcasmo ou que não apontam um posicionamento claro em relação à ação policial, dificultam a tarefa de inferência aqui endereçada. Nós observamos que, de fato, esses casos representam uma larga fração dos *tweets* classificados incorretamente.

5.2 Análise de Sentimento

Na Tabela 8, apresentamos os desempenhos dos modelos para análise de sentimento avaliados, iniciando do modelo léxico *SentiStrength*, a seguir os classificadores treinados com e sem balanceamento de classes (*smote*), e por fim as redes neurais. Para cada um desses modelos, apresentamos as cinco métricas de desempenho. Nota-se que a acurácia (*acc*) é a métrica, majoritariamente, com maior desempenho para os modelos apresentados. Contudo, a base de dados é desbalanceada, i.e., a classe negativa é majoritária, logo

Tabela 8 – Análise de sentimento - Desempenho dos modelos com as métricas: Precisão (P), Revocação (R), F1-score (F1), Acurácia (Acc) e F1-macro. Modelos de rede neural (RN) e classificadores (SVM e RF) utilizam BERT Multilingual (MultiL) e BERTimbau (PtBr).

Modelos	Negativo			Neutro			Positivo			Acc	F1-macro
	P	R	F1	P	R	F1	P	R	F1		
SentiStrength	0,78	0,56	0,65	0,37	0,52	0,43	0,37	0,52	0,43	0,54	0,50
SVM-MultiL-smote	0,78	0,80	0,79	0,53	0,50	0,51	0,43	0,35	0,38	0,70	0,56
SVM-MultiL	0,76	0,77	0,77	0,46	0,43	0,44	0,26	0,31	0,28	0,66	0,50
RF-MultiL-smote	0,72	0,88	0,79	0,48	0,26	0,34	0,50	0,15	0,24	0,68	0,45
RF-MultiL	0,72	0,97	0,83	0,74	0,20	0,32	1,00	0,04	0,07	0,72	0,41
SVM-PtBr-smote	0,81	0,80	0,80	0,54	0,59	0,56	0,56	0,35	0,43	0,72	0,60
SVM-PtBr	0,78	0,76	0,77	0,47	0,49	0,48	0,30	0,38	0,34	0,67	0,53
RF-PtBr-smote	0,75	0,88	0,81	0,55	0,36	0,43	0,57	0,15	0,24	0,71	0,50
RF-PtBr	0,73	0,95	0,83	0,63	0,28	0,38	1,00	0,04	0,07	0,72	0,43
RN-MultiL (ajustado)	0,77	0,85	0,81	0,55	0,39	0,46	0,23	0,27	0,25	0,70	0,50
RN-PtBr (ajustado)	0,81	0,86	0,83	0,61	0,53	0,57	0,50	0,46	0,48	0,75	0,63

as métricas que capturam a especificidade (precisão, revocação e F1) dos modelos em identificar corretamente cada classe ganham importância, e predominam nas discussões nessa seção.

Iniciamos observando que *SentiStrength* obtém desempenho notavelmente inferior aos outros modelos tanto em acurácia como especificidade. Isso ocorre porque ele tende a neutralizar comentários negativos como gírias, ironias e sarcasmos típicos do contexto, como discutiremos a seguir. Logo, analisamos o desempenho dos modelos treinados na base de dados para identificar aquele com o melhor desempenho. Primeiramente, observamos que o tratamento de desbalanceamento para os classificadores RF e SVM aumentaram suas especificidades (F1) para as classes minoritárias neutro e positivo, aumentando por conseguinte a especificidade geral (F1-macro). Nesse caso, SVM tem o melhor desempenho alcançando F1-macro melhores (50-60%) que RF (41-50%).

A seguir, observamos que a língua em que os modelos BERT são pré-treinados, i.e., multi-lingual (MultiL) e português brasileiro (PtBr), é o fator mais importante para o desempenho dos modelos treinados com a base de dados, i.e., classificadores e redes neurais. O pré-treinamento na língua local dos *tweets* (português brasileiro) leva a ganhos de desempenho notáveis para esses modelos. Ao comparar os ganhos dos modelos PtBr em relação a MultiL, os classificadores RF e SVM alcançam aumentos em F1-macro de até 4 pontos percentuais, ao passo que a rede neural (RN) tem um aumento de 12 pontos percentuais. Isso evidencia que o treinamento de modelos com textos no domínio e língua específicas do contexto são aspectos fundamentais para análise de sentimentos. Nesse sentido, a rede neural RN-PtBr é o modelo com o melhor desempenho, alcançando F1-macro de 63% e uma acurácia de 75%. No que lhe concerne, o classificador SVM-PtBr se destaca com o segundo melhor desempenho, apresentando F1-macro de 60% e acurácia de 73%, se tornando também um modelo elegível para análise de sentimento na ausência de redes neurais.

Agora, analisamos os erros de classificação dos sentimentos, considerando o melhor modelo treinado para a base de dados (RN-PtBr) e o *SentiStrength* que é o modelo de

análise léxica tipicamente utilizada para língua portuguesa. A Figura 4 reporta a matriz de confusão para ambos os modelos. Cada linha representa os *tweets* em uma classe real, enquanto cada coluna representa os *tweets* em uma classe prevista, o que nos permite analisar onde cada modelo mais erra e como isso acontece em função da classe.

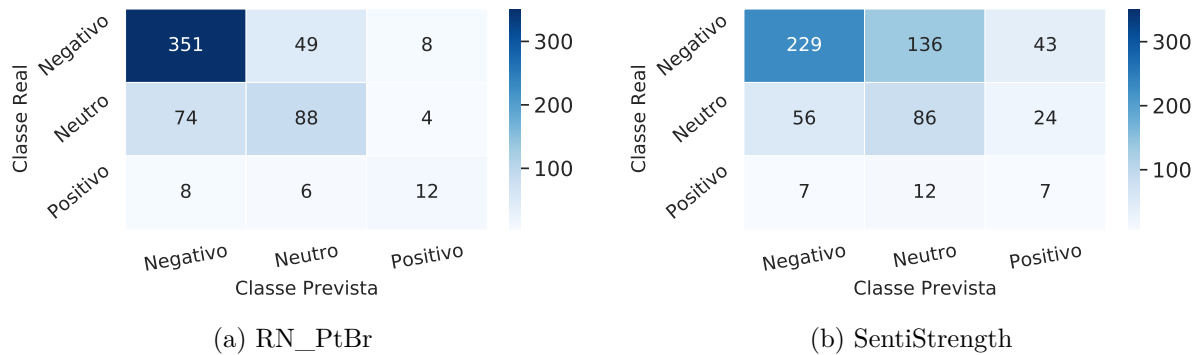


Figura 4 – Análise de sentimento - Matriz de confusão, modelo BERTimbau com ajuste fino (RN-PtBr) e SentiStrength

De forma geral, é possível notar que, em termos absolutos, o maior desafio na tarefa aqui endereçada está em diferenciar as classes *Negativo* da *Neutro* e vice-versa. Nota-se que *SentiStrength* tende a neutralizar *tweets* com sentimento negativo, sendo a vasta maioria, devido à especificidade de expressões e palavras negativas no contexto analisado. Para exemplificarmos a complexidade da tarefa, nós selecionamos alguns casos a partir da matriz de confusão focando especificamente nessas duas classes e apresentamos na Tabela 9.

Tabela 9 – Análise de sentimento - Exemplos de *tweets* classificados incorretamente pelo modelo BERTimbau com ajuste fino (RN-PtBr) e o SentiStrength.

Tweet	Real	RN-PtBr	SentiStrength
defensora de bandidos !!!!	-1	-1	1
bandido bom é bandido morto	-1	-1	0
vai ter gente falando que é so um caso isolado, caso isolado o caralho	-1	0	0
os jornais não podem chamar de bandido, sequestrador, ou nominar qualquer termo, antes da pessoa ser julgada, pois eles podem ser processados	0	-1	-1
extrema emoção, é a nova era	0	0	1
isso mesmo, roger, fale a verdade pois a sua voz tem eco	0	0	1

Os exemplos apresentados mostram como *tweets* com algum tipo de sarcasmo ou um sentimento expresso por palavras do contexto em relação a ações policiais, dificultam a tarefa de inferência aqui endereçada. Nós observamos que, de fato, esses casos representam uma larga fração dos *tweets* classificados incorretamente por ambos os métodos, em especial por *SentiStrength*.

6 Conclusão

Violência e insegurança estão entre os principais problemas dos centros urbanos. No Brasil em especial, estima-se uma taxa média de 20 mortes por mês para cada 100 mil habitantes em decorrência da violência. Redes sociais virtuais 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 inferir posicionamentos e sentimentos de usuários da rede social Twitter sobre a atuação policial em incidentes de segurança de grande repercussão em mídias de comunicação. Nesse sentido, construímos modelos de classificação baseados em linguagem natural para inferência que são aplicáveis não apenas ao Twitter, mas a outros sistemas Web baseados em comentários de usuários sobre ações policiais. Nossos experimentos mostram quão desafiante é essa inferência dada à alta polarização, neutralidade e sarcasmo observado em conteúdo gerado por pessoas.

Quanto à inferência de posicionamentos, nossos resultados mostram que modelos BERT pré-treinados em língua local obtiveram melhor desempenho para inferir posicionamentos da população sobre atuação policial expressas em *tweets*, ao passo que técnicas de balanceamento de classes tendem a melhorar esse desempenho para classificadores treinados com word embeddings. Adicionalmente, classificadores como SVM tem uma boa relação entre custo e benefício se comparado a um modelo BERT com rede neural ajustada especificamente para os dados. Nossos melhores classificadores alcançaram acurácia e especificidade (macro F1) superiores a 68% para inferir posições neutras, negativas ou positivas da população.

Em seguida, analisamos o sentimento dos usuários do Twitter em relação a ações policiais, um contexto específico que compreende um dos principais desafios da sociedade. Comparamos sentimentos inferidos via o modelo baseado em dicionário léxico *SentiStrength* e modelos baseados em BERT para processamento de linguagem natural, especificamente, redes neurais ajustadas e classificadores treinados com representações semânticas (embeddings). Os resultados mostram que modelos BERT, em especial as redes neurais e o classificador SVM com *smote*, pré-treinados em língua local, obtém melhor desempenho para inferir sentimentos nesse contexto, ao passo que *SentiStrength* tende a inferir sentimentos como neutro. Nossos melhores classificadores alcançaram acurácia e especificidade (macro F1) superiores a 60% para identificar a polaridade dos sentimentos, indicando uma metodologia promissora para inferir automaticamente a opinião pública sobre as operações policiais.

Em suma, observamos a necessidade de treinar modelos com dados no contexto (comentários repercutindo ações policiais) para inferir adequadamente posicionamentos e polaridades dos sentimentos. Isso é devido às características peculiares desse contexto, i.e.,

majoritariamente expressões com gírias de tom negativo e sarcástico. Trabalhos futuros incluem ampliar nossa metodologia de coleta de dados com fontes jornalísticas diversificadas em estilos conservador e liberal e avaliar novos modelos classificadores e redes neurais com mais camadas de neurônios. Adicionalmente, planejamos explorar características sobre sentimentos de usuários em *tweets* via modelos BERT para predição de taxas de crimes.

7 Publicações

Artigo publicado no evento BRASNAM (Anais Do Brazilian Workshop on Social Network Analysis And Mining):

FEITOSA, Marcos Paulo Fontes et al. Análise da Percepção das Pessoas no Twitter Sobre Ações Policiais. In: Anais do XI Brazilian Workshop on Social Network Analysis and Mining. SBC, 2022. p. 73-84.

Artigo publicado no XVIII Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia 2022):

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.

Referências

- AL-GARADI, M. A. et al. Analysis of online social network connections for identification of influential users: Survey and open research issues. *ACM Computing Surveys (CSUR)*, ACM New York, NY, USA, v. 51, n. 1, p. 1–37, 2018. Citado na página 19.
- BHOLANESAVITA, D.; GORE, D. Sentiment analysis on twitter data using support vector machine. *International Journal of Computer Science Trends and Technology (IJCST)–Volume*, v. 4, p. 365, 2016. Citado na página 16.
- BREIMAN, L. Random forests. *Machine learning*, Springer, v. 45, p. 5–32, 2001. Citado na página 17.
- CAMBRIA, E.; HUSSAIN, A.; HAVASI, C. Senticnet: A publicly available semantic resource for opinion mining. In: IEEE. *Proceedings of the International Conference on Semantic Computing*. [S.l.], 2010. p. 446–451. Citado na página 27.
- CHAPARRO, L. F. et al. Sentiment analysis of social network content to characterize the perception of security. In: IEEE. *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. [S.l.], 2020. p. 685–691. Citado 3 vezes nas páginas 11, 19 e 20.
- CHAWLA, N. V. et al. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, v. 16, p. 321–357, 2002. Citado na página 27.
- CHEN, X.; CHO, Y.; JANG, S. Y. Crime prediction using twitter sentiment and weather. In: IEEE. *Systems and Information Engineering Design Symposium*. [S.l.], 2015. p. 63–68. Citado 3 vezes nas páginas 11, 19 e 20.
- CORTES, C.; VAPNIK, V. Support-vector networks. *Machine learning*, Springer, v. 20, p. 273–297, 1995. Citado na página 16.
- CROWDFLOWER. Data for everyone! an introduction to the emotion lexicon. 2015. Disponível em: <<https://www.crowdfLOWER.com/wp-content/uploads/2016/07/Emotion-Lexicon-CrowdFlower.pdf>>. Citado na página 27.
- DEVLIN, J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding. In: *Association for Computational Linguistics: Human Language Technologies*. [S.l.: s.n.], 2019. p. 4171–4186. Citado 2 vezes nas páginas 15 e 20.
- ESULI, A.; SEBASTIANI, F. Sentiwordnet: A publicly available lexical resource for opinion mining. *Proceedings of the 5th Conference on Language Resources and Evaluation (LREC-2006)*, p. 417–422, 2006. Citado na página 27.
- FAN, Z.-P.; CHE, Y.-J.; CHEN, Z.-Y. Product sales forecasting using online reviews and historical sales data: A method combining the bass model and sentiment analysis. *Journal of Business Research*, v. 74, p. 90–100, 2017. ISSN 0148-2963. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0148296317300231>>. Citado na página 14.

- FERREIRA, C. H. et al. On the dynamics of political discussions on instagram: A network perspective. *Online Social Networks and Media*, Elsevier, v. 25, p. 100155, 2021. Citado 3 vezes nas páginas 11, 19 e 20.
- GLANDT, K. et al. Stance detection in covid-19 tweets. In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*. [S.l.: s.n.], 2021. v. 1. Citado na página 20.
- GO, A.; BHAYANI, R.; HUANG, L. Twitter sentiment classification using distant supervision. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. *Proceedings of the 23rd International Conference on Computational Linguistics*. [S.l.], 2009. p. 1–9. Citado na página 27.
- HAND, L. C.; CHING, B. 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*, Elsevier, v. 37, n. 1, p. 101420, 2020. Citado 4 vezes nas páginas 11, 19, 20 e 24.
- HOSSAIN, T. et al. Covidlies: Detecting covid-19 misinformation on social media. In: *Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020*. [S.l.: s.n.], 2020. Citado 4 vezes nas páginas 11, 20, 24 e 30.
- HU, M.; LIU, B. Mining and summarizing customer reviews. In: ACM. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. [S.l.], 2004. p. 168–177. Citado na página 27.
- HUTTO, C. J.; GILBERT, E. E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In: THE AAAI PRESS. *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*. [S.l.], 2014. Citado na página 27.
- IRANMANESH, A.; ATUN, R. A. Reading the urban socio-spatial network through space syntax and geo-tagged twitter data. *Journal of Urban Design*, Taylor & Francis, v. 25, n. 6, p. 738–757, 2020. Citado 2 vezes nas páginas 11 e 19.
- LI, X. et al. Using "random forest" for classification and regression. *Chinese Journal of Applied Entomology*, Institute of Zoology, v. 50, n. 4, p. 1190–1197, 2013. Citado na página 17.
- LIU, B. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, Morgan & Claypool Publishers, v. 5, n. 1, p. 1–167, 2012. Citado 2 vezes nas páginas 14 e 17.
- MAHESH, B. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*. [Internet], v. 9, p. 381–386, 2020. Citado na página 26.
- MALAGOLI, L. G. et al. A look into covid-19 vaccination debate on twitter. In: *ACM Web Science Conference 2021*. [S.l.: s.n.], 2021. Citado 2 vezes nas páginas 19 e 20.
- MINAEE, S. et al. Deep learning–based text classification: a comprehensive review. *ACM Computing Surveys (CSUR)*, ACM New York, NY, USA, v. 54, n. 3, p. 1–40, 2021. Citado 2 vezes nas páginas 12 e 20.

- MOHAMMAD, S. M.; SOBHANI, P.; KIRITCHENKO, S. Stance and sentiment in tweets. *ACM Transactions on Internet Technology (TOIT)*, ACM New York, NY, USA, v. 17, n. 3, p. 1–23, 2017. Citado na página 13.
- MOZAFARI, M.; FARAHBAKHS, R.; CRESPI, N. A bert-based transfer learning approach for hate speech detection in online social media. In: SPRINGER. *International Conference on Complex Networks and Their Applications*. [S.l.], 2019. p. 928–940. Citado 2 vezes nas páginas 23 e 24.
- NEV-USP. *Monitor da violência*. 2021. Disponível em: <https://nev.prp.usp.br/projetos/projetos-especiais/monitor-da-violencia/>. Acesso em 07 de jun. 2021. Citado na página 11.
- NIELSEN, F. Å. A new anew: Evaluation of a word list for sentiment analysis in microblogs. In: CITESEER. *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages*. [S.l.], 2011. p. 93–98. Citado na página 27.
- NOBRE, G. P.; FERREIRA, C. H.; ALMEIDA, J. M. A hierarchical network-oriented analysis of user participation in misinformation spread on whatsapp. *Information Processing & Management*, Elsevier, v. 59, n. 1, p. 102757, 2022. Citado 2 vezes nas páginas 19 e 20.
- OGLESBY-NEAL, A.; TIRY, E.; KIM, K. Public perceptions of police on social media. *Washington, DC: Urban Institute*, 2019. Citado 2 vezes nas páginas 19 e 20.
- PANG, B.; LEE, L.; VAITHYANATHAN, S. Thumbs up? sentiment classification using machine learning techniques. *arXiv preprint cs/0205070*, 2002. Citado 2 vezes nas páginas 14 e 15.
- PENNEBAKER, J.; BOOTH, R.; FRANCIS, M. *Linguistic Inquiry and Word Count (LIWC): A computerized text analysis program*. [S.l.]: Lawrence Erlbaum Associates, 2001. Citado na página 27.
- REIMERS, N. et al. Sentence-bert: Sentence embeddings using siamese bert-networks. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. [S.l.], 2019. Citado na página 20.
- REIS, J. et al. Breaking the news: First impressions matter on online news. In: *Proceedings of the International AAAI Conference on Web and Social Media*. [S.l.: s.n.], 2015. v. 9, n. 1, p. 357–366. Citado na página 11.
- SOBKOWICZ, P.; KASCHEKY, M.; BOUCHARD, G. Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web. *Government information quarterly*, Elsevier, v. 29, n. 4, p. 470–479, 2012. Citado na página 14.
- SOUSA, D. d. S.; FEITOSA, M. P. F.; GONÇALVES, G. D. Relações entre crimes e o espaço urbano: Um estudo de caso baseado em pontos de interesses extraídos da web. In: SBC. *Anais do V Workshop de Computação Urbana*. [S.l.], 2021. p. 196–208. Citado 2 vezes nas páginas 19 e 20.

- SOUZA, F.; NOGUEIRA, R.; LOTUFO, R. BERTimbau: pretrained BERT models for Brazilian Portuguese. In: *Brazilian Conference on Intelligent Systems*. [S.l.: s.n.], 2020. Citado 2 vezes nas páginas 16 e 26.
- SUN, C. et al. How to fine-tune bert for text classification? In: SPRINGER. *Chinese Computational Linguistics: 18th China National Conference, CCL 2019, Kunming, China, October 18–20, 2019, Proceedings 18*. [S.l.], 2019. p. 194–206. Citado na página 16.
- THELWALL, M. The heart and soul of the web? sentiment strength detection in the social web with sentistrength. In: *Cyberemotions*. [S.l.]: Springer, 2017. p. 119–134. Citado 3 vezes nas páginas 17, 18 e 27.
- TUCKER, R. et al. Who ‘tweets’ where and when, and how does it help understand crime rates at places? measuring the presence of tourists and commuters in ambient populations. *Journal of Quantitative Criminology*, Springer, v. 37, n. 2, p. 333–359, 2021. Citado 3 vezes nas páginas 11, 19 e 20.
- WANG, H. et al. Non-stationary model for crime rate inference using modern urban data. *IEEE transactions on big data*, IEEE, v. 5, n. 2, p. 180–194, 2017. Citado 2 vezes nas páginas 12 e 19.
- WANG, X.; GERBER, M. S.; BROWN, D. E. Automatic crime prediction using events extracted from twitter posts. In: SPRINGER. *International conference on social computing, behavioral-cultural modeling, and prediction*. [S.l.], 2012. p. 231–238. Citado 2 vezes nas páginas 19 e 20.
- WATSON, D.; CLARK, L. A. Development and validation of brief measures of positive and negative affect: The panas scales. *Journal of personality and social psychology*, American Psychological Association, v. 54, n. 6, p. 1063–1070, 1988. Citado na página 27.
- WEINZIERL, M.; HOPFER, S.; HARABAGIU, S. Misinformation adoption or rejection in the era of covid-19. In: *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM), AAAI Press*. [S.l.: s.n.], 2021. Citado 2 vezes nas páginas 11 e 20.
- YU, Y.; DUAN, W.; CAO, Q. The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision support systems*, Elsevier, v. 55, n. 4, p. 919–926, 2013. Citado na página 14.

Apêndices

APÊNDICE A – Apêndice

Artigo submetido no evento BRASNAM (Anais Do Brazilian Workshop on Social Network Analysis And Mining)

Análise da Percepção das Pessoas no Twitter Sobre Ações Policiais*

Marcos Paulo Fontes Feitosa¹, Carlos H. G. Ferreira^{2,3}
Glauber Dias Gonçalves¹, Jussara Marques de Almeida²

¹Universidade Federal do Piauí (UFPI) – CSHNB

²Universidade Federal de Minas Gerais (UFMG) – DCC

³Universidade Federal de Ouro Preto (UFOP) – DECSI

Resumo. *É cada vez mais frequente o uso de redes sociais online como meio para as pessoas trocarem ideias e expressarem suas opiniões sobre diferentes aspectos do cotidiano, incluindo violência e insegurança, um problema central a vários centros urbanos. Neste artigo investigamos o potencial uso de comentários compartilhados em uma rede social bastante popular – o Twitter – para inferir a opinião pública sobre a atuação policial em incidentes de segurança de grande repercussão. Nesse sentido, exploramos atributos extraídos desses comentários e modelos de aprendizado de máquina para inferir o posicionamento das pessoas em relação a ações policiais específicas. Nossos experimentos mostram quão desafiante é essa inferência dado grande quantidade de neutralidade e sarcasmo observado em mídias sociais. Não obstante, nossos melhores classificadores alcançaram acurácia e especificidade (macro F1) superiores a 68% para inferir posicionamentos de aprovação, desaprovação e neutralidade da população.*

1. Introdução

A violência causada por altos índices de criminalidade e a sensação de insegurança das pessoas estão dentre os principais problemas dos centros urbanos no mundo. Para se ter uma ideia da gravidade desse problema no Brasil, no ano de 2019, foram registradas 41.726 mortes por crimes violentos, uma taxa média de 20 mortes por mês para cada 100 mil habitantes brasileiros. Essa taxa pode alcançar valores superiores a 40 mortes por mês em estados das regiões norte e nordeste [NEV-USP 2021].

Diante desses fatos, costumeiramente a população não só brasileira, mas também mundial, vem recorrendo cada vez mais às redes sociais virtuais para manifestar sentimentos e opiniões sobre a atuação dos órgãos e autoridades responsáveis pela segurança pública. Essas manifestações virtuais podem refletir não apenas a percepção negativa sobre a violência urbana, mas também a aprovação ou desaprovação de ações policiais durante o enfrentamento de incidentes de segurança. Além disso, observa-se neutralidade ou mesmo polarização em casos polêmicos que despertam atenção das pessoas, levando a muitos comentários postados e compartilhados publicamente nas redes sociais virtuais [Tucker et al. 2021, Iranmanesh and Alpar Atun 2020].

Ao passo que esses comentários são uma das principais formas de manifestação da população, eles oferecem uma oportunidade única para compreender a percepção da

*Esta pesquisa é financiada pelo CNPq processo no. 402194/2021-7.

população, principalmente, durante a atuação da polícia. Tipicamente, um ecossistema é formado em que, primeiro, as notícias sobre incidentes de segurança envolvendo a polícia são postadas por portais jornalísticos e, posteriormente, são compartilhadas em redes sociais virtuais recebendo milhares de comentários. Esses, por sua vez, geram novos comentários sejam aprovando, desaprovando ou sem uma posição clara sobre a ação policial [Reis et al. 2015, Ferreira et al. 2021]. Dessa forma, toda essa carga de conteúdo gerada por usuários acerca desses incidentes, potencialmente, possibilita compreender a percepção da população quanto ao esforço dos órgãos de segurança pública no enfrentamento desse problema nas cidades.

Na literatura, esforços anteriores já usaram dados dessa natureza no contexto da atuação da polícia para avaliar o sentimento dos comentários [Hand and Ching 2020, Chaparro et al. 2020], e realizar a predição de crimes e taxas criminais [Chen et al. 2015, Tucker et al. 2021]. Com foco específico na detecção de posicionamentos da população em redes sociais, estudos foram realizados em outros contextos, por exemplo, pandemia de COVID-19 [Hossain et al. 2020, Weinzierl et al. 2021]. Dessa forma, observa-se ainda a falta de estudos e metodologias para inferir como a população reage em face às políticas de segurança pública, em particular, a exploração de conteúdo gerado por usuários em redes sociais sobre ações policiais. Especificamente, há a necessidade de entender como conteúdo nesse contexto é gerado e como ele pode ser quantificado. A partir desse ponto, pode-se investigar o potencial desses dados para inferir posicionamentos, especialmente o quanto a população apoia as ações de segurança pública. No entanto, existem desafios do ponto de vista computacional, principalmente, relacionados ao processamento de linguagem natural. Especificamente, a obtenção de dados rotulados no contexto específico de violência urbana, a representação semântica desses dados considerando o ruído presente em textos de redes sociais online, por exemplo, erros de digitação, uso de gírias e sarcasmo, e o desbalanceamento de opiniões ou posicionamentos das pessoas são alguns dos principais desafios reportados por trabalhos anteriores [Wang et al. 2017, Minaee et al. 2021].

Este trabalho oferece uma primeira tentativa em direção a inferência da opinião pública sobre a atuação policial em incidentes de segurança de grande repercussão em mídias de comunicação. Especificamente, nós investigamos o potencial de uso de dados de redes sociais para inferir a opinião pública do ponto de vista se elas aprovam ou não as ações realizadas pela polícia durante uma determinada ação policial. Para isso, nós determinamos um conjunto de notícias alvo divulgadas por três diferentes portais visando mitigar o viés ideológico, coletamos milhares de comentários sobre elas no *Twitter*, e aplicamos uma metodologia de rotulação acerca do posicionamento dos usuários sobre aquela notícia. Em seguida, nós exploramos modelos de *deep learning* baseados em arquiteturas *transformers*, que compreendem o estado da arte, para obter uma representação semântica mais fiel ao contexto. Por fim, nós avaliamos os principais classificadores e redes neurais como uma tarefa de classificação para analisar a capacidade da nossa representação em capturar a aprovação ou não a operações policiais. Nossos resultados mostram que o uso de modelos dessa natureza oferecem um potencial enorme para inferência da percepção da população neste contexto.

As próximas seções desse artigo estão organizadas da seguinte forma: A Seção 2 apresenta os trabalhos relacionados. Em seguida, a Seção 3 detalha a coleta, o processa-

mento e a rotulação dos dados. Nossa avaliação e resultados são discutidos nas Seções 4 e 5 ao passo que nossas considerações finais são apresentadas na Seção 6.

2. Trabalhos Relacionados

As plataformas de mídia social são atualmente um importante fórum para as pessoas expressarem suas opiniões e trocarem informações. Ao interagir uns com os outros por meio de postagens, re-postagens, respostas ou menções, os usuários favorecem a disseminação de informações [Al-Garadi et al. 2018]. Assim, o crescente uso dessas plataformas tem chamado a atenção de vários pesquisadores que visam modelar e analisar o comportamento dos usuários em face de fenômenos reais. Em particular, análises utilizando ferramentas de processamento natural de linguagem tem auxiliado na compreensão de diversos aspectos, por exemplo, relacionados a política [Ferreira et al. 2021, Nobre et al. 2022], pandemia [Malagoli et al. 2021] e percepção da segurança urbana [Hand and Ching 2020, Oglesby-Neal et al. 2019]. Nesta seção nós discutimos alguns destes esforços mais próximos deste trabalho.

Especificamente no contexto da percepção de segurança pelas pessoas manifestada em plataformas de mídias sociais, técnicas para análise de sentimentos vem sendo até então a principal forma de análise nesse tema. Em [Hand and Ching 2020] foram analisados os sentimentos dos comentários das pessoas às postagens das agências de polícia em suas páginas no Facebook antes e após incidentes com tiros envolvendo policiais. Os autores observaram tendência à neutralidade logo após o incidente, o que pode indicar pouca eficiência dessa técnica para analisar percepção de segurança. Os autores de [Oglesby-Neal et al. 2019] focaram em uma ação policial que resultou na morte do afro-americano Freddie Gray na cidade de Baltimore em 2015. Por sua vez, os autores de [Chaparro et al. 2020] utilizaram *tweets* georreferenciados em Bogotá na Colômbia para identificar o sentimento da população sobre a segurança nessa cidade. Cada *tweet* foi rotulado como um sentimento positivo ou negativo por um grupo de especialistas e os autores observaram que técnicas de aprendizagem de máquina supervisionada obtiveram acurácia e especificidade melhores que técnicas baseadas em regras léxicas.

Outro contexto bastante explorado é a predição de crimes a partir de características extraídas do conteúdo gerado por usuários em mídias sociais. Por exemplo, em [Wang et al. 2012] *tweets* postados por agências de notícias foram utilizados para identificar possíveis crimes. Já em [Chen et al. 2015], os autores usam *tweets* geolocalizados com foco na reincidência de crimes em uma dada região. Alguns trabalhos mesclam fontes de dados oficiais a dados de redes sociais para analisar o quanto, postagens do *Twitter* estão correlacionados com a violência urbana [Tucker et al. 2021, Iranmanesh and Alpar Atun 2020]. Em [Wang et al. 2017, Sousa et al. 2021] foi investigado o potencial de pontos de interesse gerados por usuários em serviços de mapeamento para predição das taxas criminais.

Mais recentemente, várias das tarefas supramencionadas têm sido realizadas por estratégias mais avançadas para o processamento de linguagem natural, como BERT [Devlin et al. 2019], SentenceBERT [Reimers et al. 2019], e outras variações de arquiteturas *transformers*. De fato, estes modelos fornecem representações mais ricas, portanto, com resultados potencialmente melhores. Focando no contexto do *Twitter*, alguns esforços devem ser mencionados. Em [Hossain et al. 2020], os autores utilizaram vários

modelos pré-treinados para identificar o posicionamento dos usuários quanto a conceitos errôneos sobre o COVID-19 no Twitter. Os resultados alcançados que, no melhor modelo, apresenta um *F1-score* de 50.2 evidenciam os desafios dessa tarefa. Considerando o mesmo conjunto de dados e tarefa, o trabalho de [Weinzierl et al. 2021] conseguiu aumentar o valor do *F1-score* para 74.3. Ainda no contexto da pandemia, Glandt et al. [Glandt et al. 2021] realizaram a análise do desempenho dos classificados por instância, enfatizando que os resultados variam de acordo com a instância. Em geral, os autores observaram *F1-score* que variam de 0.53 a 0.83. Assim, eles reforçam o argumento de que é difícil ter um único modelo que seja melhor para todos os tipos de instâncias analisados, ou seja, que capture todas as nuances presentes em textos de redes sociais online. Esses resultados refletem o atual estado da arte, como mostram algumas revisões recentes sobre tal tarefa em dados de redes sociais [Minaee et al. 2021].

Em suma, neste trabalho apresentamos um estudo ortogonal aos esforços anteriores. Nosso objetivo aqui é inferir posicionamentos das pessoas quanto incidentes de violência envolvendo atuação policial através de técnicas para processamento de linguagem natural baseada na arquitetura de *transformers*, que fornecem uma representação semântica estado da arte para esse tipo de processamento. Para isso, nós exploramos fontes de informações geradas por diversas agências de notícias similares aos trabalhos de [Wang et al. 2012, Chen et al. 2015] e as usamos para quantificar o posicionamento dos usuários em um contexto específico e não explorado dessa forma anteriormente, que é a sobre a atuação policial.

3. Bases de Dados

Nesta seção descrevemos as bases de dados e a metodologia de processamento desses dados para o uso em inferências sobre posicionamentos da população sobre atuação policial em incidentes de violência.

3.1. Coleta de dados

Realizamos a coleta de dados em duas etapas. Primeiramente, selecionamos notícias de grande repercussão sobre o contexto desse trabalho em portais de mídia digital. Em seguida, coletamos comentários das pessoas em redes sociais (*tweets*) sobre essas notícias como descrito a seguir.

Inicialmente, selecionamos notícias sobre incidentes de segurança com intervenção policial entre o período de 04/2019 até 12/2021 que foram amplamente divulgadas e repercutidas em mídia digital. Nesse sentido, pesquisamos por notícias nas seções sobre violência e crimes em três relevantes portais de notícias nacionais: UOL¹, Folha de São Paulo², e G1³. Utilizamos como critério de seleção a participação dos leitores via comentários de texto no próprio portal de notícia. Especificamente, selecionamos notícias com mais de 700 comentários, marca notável e acima da média usual de comentários e intuitivamente indica uma grande repercussão em redes sociais. Ao todo foram selecionadas 8 notícias, que listamos a seguir com um identificador, e as informações que permitem encontrá-las, isto é, o portal de origem e o título da notícia.

¹<https://www.uol.com.br/>

²<https://www.folha.uol.com.br/>

³<https://g1.globo.com/>

- N_1, G1: "Sequestrador de ônibus é morto por atirador de elite na Ponte Rio-Niterói, os 39 reféns passam bem"; UOL: "Sequestrador é morto, polícia libera reféns e encerra sequestro de ônibus"; Folha de São Paulo: "PM mata homem que manteve passageiros de ônibus reféns na ponte Rio-Niterói".
- N_2, G1: "Menina de 8 anos morre baleada no Complexo do Alemão"; UOL: "Menina de 8 anos morre baleada após operação policial no Complexo do Alemão"; Folha de São Paulo: "Menina de oito anos morre baleada no Rio de Janeiro".
- N_3, G1: "Mulher é imobilizada por PMs com bebê no colo em Itabira, MG"; UOL: "Policial imobiliza mulher com criança no colo em Itabira (MG)"; Folha de São Paulo: "PMs derrubam e imobilizam mulher com bebê no colo em Itabira (MG)".
- N_4, G1: "Lázaro Barbosa morre após ser preso em Goiás"; UOL: "Lázaro é morto em Goiás; policiais comemoram após carregar corpo"; Folha de São Paulo: "Lázaro Barbosa, o serial killer do DF, é morto pela polícia após 20 dias de buscas";
- N_5, G1: "Menino de 14 anos morre durante operação das polícias Federal e Civil no Complexo do Salgueiro, RJ"; UOL: "Adolescente João Pedro é morto em operação no Rio, família critica polícia"; Folha de São Paulo: "Menino de 14 anos é morto em casa durante ação da PF no Rio".
- N_6, G1: "Menino de 7 anos morre após ser baleado na porta de casa na Baixada Fluminense"; UOL: "Menino de 7 anos morre após ser baleado enquanto brincava na porta de casa"; Folha de São Paulo: "Menino de 7 anos morre na porta de casa após tiroteio na Baixada Fluminense".
- N_7, G1: "Homem morre após ser baleado em ação do Exército na Zona Oeste do Rio"; Folha de São Paulo: "Exército dispara 80 tiros em carro de família no Rio e mata músico".
- N_8, G1: "Operação no Jacarezinho deixa 28 mortos, provoca intenso tiroteio e tem fuga de bandidos"; UOL: "Operação com 25 mortos no Jacarezinho é a mais letal da história do Rio"; Folha de São Paulo: "Polícia faz operação mais letal da história do RJ, com ao menos 25 mortos".

A próxima etapa da coleta de dados consiste em obter comentários das pessoas sobre as notícias acima relacionadas em redes sociais. Utilizamos os títulos e os links de cada notícia para coletas na rede social *Twitter*. Optamos por essa rede devido aos recursos oferecidos à pesquisadores para coleta de dados e o foco em conteúdo de texto, i.e., a maioria dos *tweets* são textos curtos limitados em 280 caracteres atualmente. Para coletar *tweets* utilizamos a API do Twitter versão 2 com a biblioteca *Tweepy* na linguagem Python⁴. Essa biblioteca funciona como um facilitador para acessar a API, tornando possível buscar *tweets* por palavras chaves (i.e., título) e links das notícias. Ao fim dessa etapa, coletamos um total de 16 276 *tweets* sobre todas as notícias.

3.2. Pré-processamento e Rotulação

Conduzimos o mínimo de pré-processamentos de textos nos *tweets* coletados, seguindo recomendações de outros trabalhos que lidaram com inferência via processamento de linguagem natural [Mozafari et al. 2019]. Assim, removemos dados que dificultam esse processamento como *links* e quebras de linha. Adicionalmente, desconsideramos *tweets* com apenas uma palavra, dado o menor potencial de inferência semântica desses. Por outro lado, mantivemos todas as *stop words* para treinar modelos de inferência com a

⁴<https://docs.tweepy.org/en/stable/>

mesma sequência em que as palavras aparecem nos *tweets*. Todas essas tarefas de pré-processamento foram realizadas com a biblioteca NLTK⁵ na linguagem Python.

Após o pré-processamento, realizamos a etapa de rotulação das instâncias de dados. O rótulo consiste na classificação de um *tweet* como *Aprova*, *Desaprova* ou *Neutro* considerando o posicionamento do usuário que o postou sobre a ação policial no incidente ao qual a notícia se refere. Em outras palavras, o *tweet* com o rótulo *Aprova* significa que o usuário aprova a ação policial, *Desaprova* significa que o usuário não aprova a ação, por sua vez, *Neutro* expressa que o usuário é indiferente ao assunto ou não manifesta claramente seu posicionamento. A rotulação serve de base para treinamento de modelos de inferência sendo realizada por humanos, como é usual na literatura. Nesse sentido, contamos com três pessoas (não especialistas⁶) para rotular um subconjunto de *tweets*, dado a inviabilidade de rotulação da base de dados completa. Esse subconjunto é composto por 4467 *tweets* definidos aleatoriamente e proporcionalmente ao total de *tweets* por notícia.

Avaliamos a qualidade da rotulação via métricas de concordância entre os três rotuladores. Nesse sentido, selecionamos outro subconjunto de 198 *tweets* aleatórios com rotulação das três pessoas. O percentual de concordância entre os rotuladores foi de 70,2%, ao passo que o índice *Fleiss Kappa* foi de 0,67, indicando uma concordância substancial entre os pesquisadores em comparação a outros trabalhos da literatura [Hossain et al. 2020].

As informações sobre a distribuição dos rótulos por notícia são mostradas na Figura 1. A distribuição das classes é fortemente desbalanceada, contendo principalmente *tweets* com postura neutra que representam 49% do conjunto total, seguidos por *tweets* da classe *desaprova* representando 27% do conjunto total, e por fim, é a classe *aprova* com 24%. Esse desbalanceamento, principalmente, em direção a classe neutra torna a inferência de posicionamentos mais difícil como já observado em outros trabalhos da literatura [Mozafari et al. 2019, Hand and Ching 2020]. Além disso, nossa proposta é ainda mais desafiadora, pois lidamos com comentários em português brasileiro em contextos locais envolvendo incidentes de violência envolvendo ação policial, em que nem sempre as ferramentas e modelos apresentam o melhor resultado.

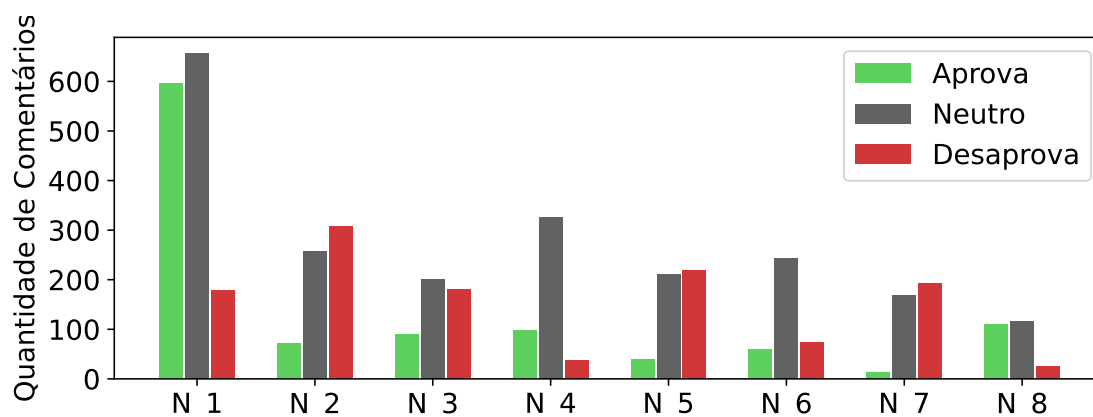


Figura 1. Distribuição dos rótulos por notícia.

⁵<https://www.nltk.org/>

⁶Entendemos que o contexto das notícias é conhecido pela população amplamente, sendo desnecessário rotulação por um especialista.

4. Metodologia

Nesta seção apresentamos, primeiramente, os modelos para inferência de posicionamento das pessoas em relação à atuação policial. A seguir descrevemos a metodologia e métricas utilizadas para avaliar o desempenho desses modelos.

4.1. Modelos

O BERT (*Bidirectional Encoder Representations from Transformers*) [Devlin et al. 2019] é um algoritmo de aprendizado profundo (do, inglês *Deep Learning*) desenvolvido pelo *Google* para processamento de linguagem natural. BERT é um modelo pré-treinado com conteúdo do Wikipédia e o Book Corpus, ambos em língua inglesa, contendo 2.500 milhões e 800 milhões de palavras respectivamente. Um recurso importante do BERT é sua construção com representações contextuais de modelos pré-treinados, baseado na arquitetura de *transformers*. Através desse recurso, comunidades de pesquisadores ou corporações podem treinar novos modelos adaptados a alguma linguagem no contexto local e distribuí-los como modelos pré-treinados BERT.

Neste trabalho utilizamos dois modelos pré-treinados que são o *Multilingual* [Devlin et al. 2019] e o *BERTimbau* [Souza et al. 2020]. O modelo *Multilingual* é pré-treinado em 102 idiomas, e possui 12 camadas (blocos de *transformers*), 12 cabeçotes de atenção e 110 milhões de parâmetros. No que lhe concerne, *BERTimbau* é pré-treinado na língua portuguesa do Brasil, e em nossas avaliações utilizamos a versão desse modelo com a maior quantidade de codificadores (*large*), contendo 24 camadas, 16 cabeçotes de atenção e 335 milhões de parâmetros. Embora treinado para uma língua específica, *BERTimbau* já se mostrou eficiente em tarefas de processamento de linguagem natural que requerem reconhecimento de entidade nomeada, semelhança textual de sentença e reconhecimento de enlace textual [Souza et al. 2020].

Utilizamos matrizes de *word embeddings* extraídas do BERT para treinar classificadores supervisionados na inferência de posicionamento de *tweets*. Em suma, *Word Embedding* é uma forma de representar palavras através de números para processamento de linguagem natural. Essa representação é normalmente na forma de um vetor de valores reais, que representa o significado das palavras conforme o contexto e o significado da sentença, i.e., o *tweet* em que a palavra está inserida. Para gerar essa representação, utilizamos os modelos pré-treinados acima descritos. O *BERTimbau large* gera um vetor de tamanho 1024 para cada sentença, ao passo que o *Multilingual* gera um vetor de tamanho 768. Ao fim temos uma matriz de *word embeddings* em que cada linha representa um *tweet* e cada coluna representa uma característica do *tweet*.

A matriz de *words embeddings* foi utilizada como as características de um conjunto de *tweets* rotulados para treinar dois classificadores populares que são *Random Forest* (RF) e *Support Vector Machine* (SVM) [Mahesh 2020]. RF combina um conjunto de preditores de árvores de modo que cada árvore depende dos valores de um vetor aleatório da amostragem com a mesma distribuição. SVR busca prever um valor real após traçar duas retas paralelas, chamadas limites. O modelo ainda traça uma reta linear entre as duas outras retas retas de modo a ajustar seus valores.

Além dos modelos de classificação acima, avaliamos outra abordagem baseada em ajustes finos diretamente sob os modelos BERT. Especificamente, adicionamos uma nova

camada de neurônios aos modelos BERT pré-treinados e a treinamos para nossa tarefa de classificação com os *tweets* rotulados. Esse ajuste tem a vantagem de acrescentar novas informações aos modelos, que já codificam muitas informações sobre a língua em que foram pré-treinados. Dessa forma pode-se obter melhores resultados para inferência de posicionamentos com menor quantidade de dados rotulados para treinamento, apesar do tempo e complexidade para construção de mais uma camada na rede neural do BERT.

4.2. Configurações, Treinamento e Métricas de avaliação

Como ambiente de experimentação utilizamos uma máquina virtual da AWS do tipo G4dn que são instâncias que usam GPUs otimizadas para aprendizagem profunda. Logo, ajustamos os modelos pré-treinados BERTimbau e Multilingual com 10 épocas e otimizador Adam com os respectivos *learning rate* e *batch size* de cada modelo. Para a avaliação da rede neural ajustada, consideramos 80% dos *tweets* rotulados para treino e 20% para teste, treinamos esse modelo em cada época, e ao fim de cada época avaliamos o seu desempenho. Para a abordagem com os classificadores, geramos a matriz de *word embedding* com os modelos BERT pré-treinados, em seguida dividimos a matriz em treino e teste com 80% e 20% dos *tweets* respectivamente, e treinamos os classificadores sem e com balanceamento de classes utilizando a técnica *smote* [Mahesh 2020]. O código fonte, bem como a base de dados, deste trabalho estão disponibilizados no GitHub ⁷.

Avaliamos os modelos com o conjunto de teste utilizando cinco métricas, que são acurácia, precisão, revocação, f1-score e f1-macro, para cada classe (Desaprova, Aprova, Neutro). A acurácia indica o percentual de *tweets* corretamente classificados, isto é, a soma acertos de todas as classes dividido pelo número total de *tweets* classificados. Já a precisão é calculada para cada classe individualmente e evidencia o percentual de *tweets* corretamente classificados para aquela classe. A revocação é calculada justamente pelo total de *tweets* corretamente classificados para uma classe sobre o total de *tweets* dessa classe. F1-score é a média harmônica entre precisão e revocação para cada classe, ao passo que o F1-macro é a média do F1-score considerando todas as classes.

5. Resultados

Nesta seção apresentamos os resultados obtidos pelos modelos descritos anteriormente na inferência de posicionamentos sobre ações policiais no conjunto de teste dos *tweets*. Adicionalmente, analisamos os erros desses modelos para mostrar a dificuldade da inferência e possíveis estratégias para lidar com os erros.

5.1. Classificação de Posicionamentos

Na Tabela 1, apresentamos os resultados para a avaliação de desempenho dos classificadores treinados sem e com o tratamento para desbalanceamento de classes (*smote*), além da rede neural BERT com ajuste fino. Para cada um desses modelos, apresentamos os resultados via cinco métricas de desempenho. Nota-se que a acurácia (*acc*) é a métrica, majoritariamente, com maior desempenho para as classes desaprova e aprova. Contudo, é importante observar que os dados analisados são desbalanceados e essas classes são minoritárias na maior parte das notícias analisadas (i.e., N_1, N_3, N_4, N_6 e N_8). Nesse

⁷https://github.com/LABPAAD/crimes_stance

caso, as métricas que capturam a especificidade (precisão, revocação e f1) dos modelos em identificar corretamente a classes minoritárias ganham importância, e essas métricas são o foco da avaliação de desempenho nessa seção.

Tabela 1. Desempenho dos modelos Multilingual (MultiL) e BERTimbau (PtBr), usando as estratégias de ajuste fino da rede neural (RN) e a da matriz de *embedding* com os classificadores (SVM e RF). Avaliando com as seguintes métricas: Precisão (P), Revocação (R), F1-score (F1), Acurácia (Acc) e F1-macro.

Modelos	Desaprova			Aprova			Neutro			Acc	F1-macro
	P	R	F1	P	R	F1	P	R	F1		
SVM-MultiL-smote	0,46	0,53	0,49	0,48	0,59	0,53	0,76	0,63	0,69	0,59	0,57
SVM-MultiL	0,51	0,49	0,50	0,55	0,47	0,51	0,71	0,77	0,74	0,63	0,58
RF-MultiL-smote	0,50	0,48	0,49	0,57	0,50	0,54	0,72	0,77	0,74	0,63	0,59
RF-MultiL	0,52	0,34	0,41	0,61	0,38	0,47	0,64	0,86	0,73	0,62	0,54
SVM-PtBr-smote	0,52	0,65	0,58	0,55	0,66	0,60	0,84	0,66	0,74	0,66	0,64
SVM-PtBr	0,54	0,58	0,56	0,66	0,61	0,63	0,78	0,77	0,77	0,69	0,66
RF-PtBr-smote	0,55	0,56	0,56	0,68	0,56	0,61	0,74	0,79	0,77	0,68	0,65
RF-PtBr	0,55	0,44	0,49	0,75	0,49	0,59	0,69	0,86	0,76	0,67	0,61
RN-PtBr	0,58	0,63	0,61	0,69	0,68	0,69	0,77	0,74	0,76	0,70	0,68
RN-MultiL	0,49	0,55	0,52	0,61	0,53	0,56	0,71	0,71	0,71	0,63	0,60

Primeiramente, discutimos os resultados do tratamento de desbalanceamento das classes para os classificadores RF e SVM. Sem esse tratamento, há um viés para precisão nas classes minoritárias, pois apenas *tweets* com posicionamentos claros de aprovação ou desaprovação serão classificadas como tal. Isso pode ser observado na Tabela 1, com valores de precisão majoritariamente maiores que revocação, levando a baixa especificidade e tendência a classificar posicionamentos como neutros. Ao tratarmos o desbalanceamento com *smote*, observamos que SVM e RF, em geral, aumentam o desempenho para revocação e, por consequência, aumentam o acerto para as posturas de aprovação e desaprovação. A métrica F1 é a referência para o melhor compromisso entre precisão e revocação, e obtemos F1 maiores para os classificadores com *smote* (exceção apenas de SVM-MultiL para Aprova e SVM-PtBR para Desaprova). Observa-se também que o SVM, em geral, tem o melhor desempenho alcançando F1 melhores (49-63%) com ou sem balanceamento, superando as marcas de F1 do classificador RF (41-61%).

Agora focamos no aspecto da língua em que os modelos BERT são pré-treinados, i.e., multi-lingual (MultiL) e português brasileiro (PtBr). Para isso utilizamos a métrica F1-macro que considera a média do F1 para as três classes de posicionamentos. Observa-se que o pré-treinamento na língua local dos *tweets* (português brasileiro) tem ganhos de desempenho notáveis para os três modelos avaliados. Ao comparar os ganhos dos modelos PtBr em relação a MultiL, os classificadores RF e SVM alcançam aumentos em F1-macro de até 7 e 8 pontos percentuais respectivamente, ao passo que a rede neural ajustada (RN) tem um aumento de 8 pontos percentuais. Isso evidencia que o treinamento de modelos com textos no domínio e língua específicas do contexto em investigação são aspectos fundamentais para inferências de posicionamentos. Essa observação corrobora com outros trabalhos recentes que também utilizam *tweets* para inferências de posicionamentos, mas em contextos diferentes ao investigado nesse trabalho [Hossain et al. 2020].

Finalmente, discutimos o uso da rede neural BERT com ajuste fino (RN). Como esperado, esse modelo obteve o melhor resultado. Especificamente, o modelo pré-treinado em português brasileiro (RN-PtBr) alcançou o F1-macro de 68% e uma acurácia de 70%. É importante mencionar, todavia, que o desempenho da rede neural não é tão superior ao

melhor classificador (SVM-PtBr). O ganho do RN-PtBr corresponde a apenas 2 pontos percentuais sobre o F1-macro do SVM-PtBr. Considerando o custo computacional para ajustar a rede neural BERT (i.e., o dobro do tempo dos classificadores), o modelo SVM apresenta um bom compromisso entre desempenho e custo de treinamento.

5.2. Análise dos Erros de Classificação

A Figura 2 reporta a matriz de confusão para o modelo RN-PtBr que obteve o melhor desempenho, mostrado na Tabela 1. Cada linha representa os *tweets* em uma classe real, enquanto cada coluna representa os *tweets* em uma classe prevista, o que nos permite analisar onde o classificador mais erra e como isso acontece em função da classe.

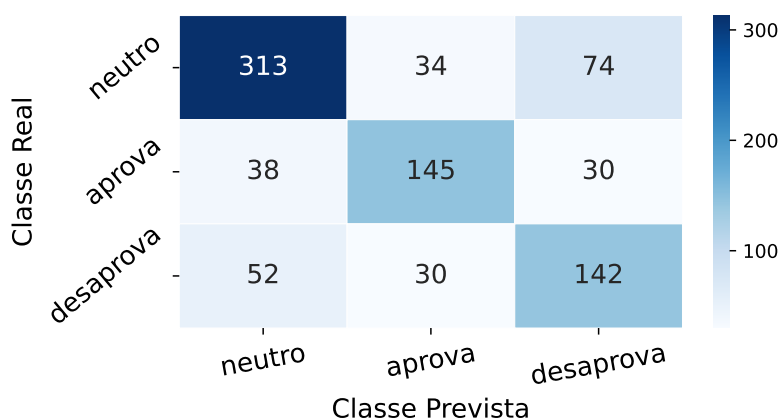


Figura 2. Matriz de confusão, modelo BERTimbau com ajuste fino (RN-PtBr).

De forma geral, é possível notar que, em termos absolutos, o maior desafio na tarefa aqui endereçada está em diferenciar as classes *Neutro* da *Desaprova* e vice-versa. Para exemplificarmos a complexidade da tarefa, nós selecionamos alguns casos a partir da matriz de confusão focando especificamente nessas duas classes e apresentamos na Tabela 2.

Tabela 2. Exemplos de *tweets* classificados incorretamente pelo modelo BERTimbau com ajuste fino (RN-PtBr).

Tweet	Real	Predito
Eu já vi isso mano, é osso	Neutro (0)	Desaprova (-1)
Meu Deus que realidade cruel, que os passageiros fiquem bem	Neutro (0)	Desaprova (-1)
Isso foi de propósito?	Neutro (0)	Desaprova (-1)
É esse o nível de quem vai nos salvar do comunismo quero saber quem vai nos salvar deles	Desaprova (-1)	Neutro (0)
#ficaemcasa e leve um tiro	Desaprova (-1)	Neutro (0)
Não é saudável e nem natural se acostumar com tragédias, as pessoas estão doentes!	Desaprova (-1)	Neutro (0)

Os exemplos apresentados mostram como, *tweets* curtos, contendo algum tipo de sarcasmo ou que não apontam um posicionamento claro em relação à ação policial, dificultam a tarefa de inferência aqui endereçada. Nós observamos que, de fato, esses casos representam uma larga fração dos *tweets* classificados incorretamente. [Meta and community 2015]

6. Conclusões e Trabalhos Futuros

Neste trabalho, nós propomos inferir o posicionamento das pessoas em relação a operações policiais, um contexto específico dentro do tema de segurança pública e que compreende um dos principais desafios da sociedade. Para isso, nós definimos um conjunto de notícias que tomaram dimensões relativamente grandes devida à atuação da polícia no Brasil a fim de avaliar a percepção das pessoas por meio de dados de redes sociais, especificamente, do *Twitter*.

Considerando os desafios presentes em atividades de processamento natural de linguagem, nós propomos o uso das arquiteturas *transformers* para obter uma representação semântica de baixa dimensão (*embeddings*) e endereçamos o problema por meio da tarefa de classificação. Dessa forma, nossa ideia é compreender como o posicionamento dos usuários podem ser capturados e, futuramente, generalizados para outras aplicações. Nossos resultados que modelos BERT pré-treinados em língua local obtiveram melhor desempenho para inferir posicionamentos da população sobre atuação policial expressas em tweets, ao passo que técnicas de balanceamento de classes tendem a melhorar esse desempenho para classificadores treinados com *word embeddings*. Adicionalmente, classificadores como SVM tem uma boa relação entre custo e benefício se comparado a um modelo BERT com rede neural ajustada especificamente para os dados.

Como trabalhos futuros, pretendemos ampliar nossa avaliação incluindo outros classificadores e redes neurais ajustadas com mais camadas de neurônios. Adicionalmente, planejamos explorar características extraídas de tweets via modelos BERT também para predição de taxas de crimes.

Referências

- Al-Garadi, M. A., Varathan, K. D., Ravana, S. D., Ahmed, E., Mujtaba, G., Khan, M. U. S., and Khan, S. U. (2018). Analysis of online social network connections for identification of influential users: Survey and open research issues. *ACM Computing Surveys (CSUR)*, 51(1):1–37.
- Chaparro, L. F., Pulido, C., Rudas, J., Reyes, A. M., Victorino, J., Narváez, L. Á., Gómez, F., and Martínez, D. (2020). Sentiment analysis of social network content to characterize the perception of security. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASO-NAM)*, pages 685–691. IEEE.
- Chen, X., Cho, Y., and Jang, S. Y. (2015). Crime prediction using twitter sentiment and weather. In *Systems and Information Engineering Design Symposium*, pages 63–68. IEEE.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Association for Computational Linguistics: Human Language Technologies*, pages 4171–4186.
- Ferreira, C. H., Murai, F., Silva, A. P., Almeida, J. M., Trevisan, M., Vassio, L., Mellia, M., and Drago, I. (2021). On the dynamics of political discussions on instagram: A network perspective. *Online Social Networks and Media*, 25:100155.
- Glandt, K., Khanal, S., Li, Y., Caragea, D., and Caragea, C. (2021). Stance detection in covid-19 tweets. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, volume 1.
- Hand, L. C. and Ching, B. D. (2020). Maintaining neutrality: A sentiment analysis of police agency facebook pages before and after a fatal officer-involved shooting of a citizen. *Government Information Quarterly*, 37(1):101420.

- Hossain, T., Logan IV, R. L., Ugarte, A., Matsubara, Y., Young, S., and Singh, S. (2020). Covidlies: Detecting covid-19 misinformation on social media. In *Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020*.
- Iranmanesh, A. and Alpar Atun, R. (2020). Reading the urban socio-spatial network through space syntax and geo-tagged twitter data. *Journal of Urban Design*, 25(6):738–757.
- Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]*, 9:381–386.
- Malagoli, L. G., Stancioli, J., Ferreira, C. H., Vasconcelos, M., Couto da Silva, A. P., and Almeida, J. M. (2021). A look into covid-19 vaccination debate on twitter. In *ACM Web Science Conference 2021*.
- Meta and community (2015). React native. Disponível em: <https://reactnative.dev/>. Acesso em: 2023-01-20.
- Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., and Gao, J. (2021). Deep learning-based text classification: a comprehensive review. *ACM Computing Surveys (CSUR)*, 54(3):1–40.
- Mozafari, M., Farahbakhsh, R., and Crespi, N. (2019). A bert-based transfer learning approach for hate speech detection in online social media. In *International Conference on Complex Networks and Their Applications*, pages 928–940. Springer.
- NEV-USP (2021). Monitor da violência. Disponível em: <https://nev.prp.usp.br/projetos/projetos-especiais/monitor-da-violencia/>. Acesso em 07 de jun. 2021.
- Nobre, G. P., Ferreira, C. H., and Almeida, J. M. (2022). A hierarchical network-oriented analysis of user participation in misinformation spread on whatsapp. *Information Processing & Management*, 59(1):102757.
- Oglesby-Neal, A., Tiry, E., and Kim, K. (2019). Public perceptions of police on social media. *Washington, DC: Urban Institute*.
- Reimers, N., Gurevych, I., Reimers, N., Gurevych, I., Thakur, N., Reimers, N., Daxenberger, J., Gurevych, I., Reimers, N., Gurevych, I., et al. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.
- Reis, J., de Souza, F., de Melo, P. V., Prates, R., Kwak, H., and An, J. (2015). Breaking the news: First impressions matter on online news. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 9, pages 357–366.
- Sousa, D. d. S., Feitosa, M. P. F., and Gonçalves, G. D. (2021). Relações entre crimes e o espaço urbano: Um estudo de caso baseado em pontos de interesses extraídos da web. In *Anais do V Workshop de Computação Urbana*, pages 196–208. SBC.
- Souza, F., Nogueira, R., and Lotufo, R. (2020). BERTimbau: pretrained BERT models for Brazilian Portuguese. In *Brazilian Conference on Intelligent Systems*.
- Tucker, R., O'Brien, D. T., Ciomek, A., Castro, E., Wang, Q., and Phillips, N. E. (2021). Who 'tweets' where and when, and how does it help understand crime rates at places? measuring the presence of tourists and commuters in ambient populations. *Journal of Quantitative Criminology*, 37(2):333–359.
- Wang, H., Yao, H., Kifer, D., Graif, C., and Li, Z. (2017). Non-stationary model for crime rate inference using modern urban data. *IEEE transactions on big data*, 5(2):180–194.
- Wang, X., Gerber, M. S., and Brown, D. E. (2012). Automatic crime prediction using events extracted from twitter posts. In *International conference on social computing, behavioral-cultural modeling, and prediction*, pages 231–238. Springer.

Weinzierl, M., Hopfer, S., and Harabagiu, S. (2021). Misinformation adoption or rejection in the era of covid-19. In *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, AAAI Press.

APÊNDICE B – Apêndice

Artigo submetido no XVIII Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia 2022)



Sentiment Analysis on Twitter Repercussion of Police Operations

Marcos Fontes Feitosa
CSHNB, UFPI
Brazil
marcosmpff@ufpi.edu.br

Saul Rocha
CSHNB, UFPI
Brazil
saul.rocha2001@ufpi.edu.br

Glauber Dias Gonçalves
CSHNB, UFPI
Brazil
ggoncalves@ufpi.edu.br

Carlos Henrique Ferreira
DCC, UFMG
Brazil
chgferreira@dcc.ufmg.br

Jussara Marques Almeida
DCC, UFMG
Brazil
jussara@dcc.ufmg.br

ABSTRACT

Violence and a sense of insecurity are among the main problems in urban centres. In Brazil, an average rate of 20 deaths per month is estimated for every 100,000 inhabitants due to violence. Virtual social networks are increasingly used as a means for users to express their opinions or indignation about this problem. In this article, we analyze the sentiment of users in comments shared on Twitter about police operations with great repercussions in news portals in Brazil. In this sense, we explore lexicon and machine learning models to understand the emotion in which users discuss public safety on social networks and their opinion about the work of government agencies to reduce violence in cities. Our experiments show how challenging this inference is given peculiar characteristics of the context, such as mostly negative and sarcastic expressions. Nevertheless, our best classifiers achieved accuracy and specificity (macro F1) greater than 60% for identifying sentiments polarity, indicating a promising methodology for automatically inferring public opinion about police operations.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in collaborative and social computing.**

KEYWORDS

Sentiment Analysis, Twitter, Natural Language Processing, Police Operations

ACM Reference Format:

Marcos Fontes Feitosa, Saul Rocha, Glauber Dias Gonçalves, Carlos Henrique Ferreira, and Jussara Marques Almeida. 2022. Sentiment Analysis on Twitter Repercussion of Police Operations. In *Brazilian Symposium on Multimedia and Web (WebMedia '22)*, November 7–11, 2022, Curitiba, Brazil. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3539637.3558050>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WebMedia '22, November 7–11, 2022, Curitiba, Brazil

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-9409-3/22/11... \$15.00

<https://doi.org/10.1145/3539637.3558050>

1 INTRODUÇÃO

Dentre os diferentes desafios dos centros urbanos, destaca-se os altos índices criminalidade. No Brasil em especial, estima-se uma taxa média de 20 mortes por mês para cada 100 mil habitantes, alcançando taxas superiores a 40 mortes por mês nas regiões norte e nordeste [8]. Diante desses fatos, costumeiramente a população não só brasileira, mas também mundial, vem recorrendo cada vez mais às redes sociais virtuais para manifestar sentimentos e opiniões sobre a atuação dos órgãos responsáveis pela segurança pública via comentários de texto. Por sua vez, esses órgãos têm a oportunidade de se beneficiar desses comentários e desenvolver ferramentas para avaliar os anseios da população sobre suas gestões no enfrentamento da violência nas cidades.

Nesse artigo, investigamos o potencial de uso dos comentários na rede social *Twitter*, i.e., *tweets*, para analisar a impressão da população sobre a atuação das entidades responsáveis pela segurança pública. 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 como discutimos na Seção 2. Contudo, observa-se ainda a falta de metodologias para inferir como a população reage em face às políticas de segurança pública, em particular, ações policiais. Especificamente, há a necessidade de entender como conteúdo nesse contexto é gerado e como ele pode ser quantificado para inferir sentimentos com alta acurácia e especificidade. Nesse sentido oferecemos duas contribuições nesse trabalho: (i) *metodologia para extrair comentários da população sobre ações policiais de interesse nas Seções 3 e 4 e*, (ii) *análise de sentimento desses comentários comparando resultados de um modelo léxico tradicional e modelos de aprendizagem de máquina e profunda treinados com esses comentários na Seções 5*. Um resumo dessas contribuições e principais descobertas desse trabalho, com os apontamentos para os próximos passos, são apresentados na Seção 6.

2 TRABALHOS RELACIONADOS

As plataformas de mídia social são atualmente um importante fórum para as pessoas expressarem suas opiniões e trocarem informações. Especificamente no contexto da percepção de segurança manifestada em plataformas de mídias sociais, modelos baseados em dicionário léxico para análise de sentimentos vem sendo até então a principal abordagem. Em [3] foram analisados os sentimentos dos comentários das pessoas às postagens das agências de polícia em

suas páginas no Facebook antes e após incidentes com tiros envolvendo policiais. Os autores observaram tendência à neutralidade logo após o incidente, o que pode indicar pouca eficiência do modelo para analisar percepção de segurança. Os autores de [9] focaram em uma ação policial que resultou na morte do afro-americano Freddie Gray na cidade de Baltimore em 2015. Por sua vez, os autores de [1] utilizaram *tweets* georreferenciados em Bogotá na Colômbia para identificar o sentimento da população sobre a segurança nessa cidade. Os autores observaram que modelos de aprendizagem de máquina supervisionada obtiveram acurácia e especificidade melhores que modelos baseados em dicionários léxicos.

Mais recentemente, várias análises de textos e comentários têm sido realizadas por estratégias mais avançadas para o processamento de linguagem natural, como BERT [2], e outras variações de arquiteturas *transformers*. De fato, estes modelos fornecem representações mais ricas, portanto, com resultados potencialmente melhores. Focando em comentários do *Twitter*, alguns esforços devem ser mencionados. Em [4], os autores utilizaram vários modelos pré-treinados para identificar o posicionamento dos usuários quanto a conceitos errôneos sobre o COVID-19 no *Twitter*. Os resultados alcançados que, no melhor modelo, apresenta um *F1-score* de 50.2 evidenciam os desafios dessa tarefa. Considerando o mesmo conjunto de dados e tarefa, o trabalho de [12] conseguiu aumentar o valor do *F1-score* para 74.3. Esses resultados refletem o atual estado da arte, como mostram algumas revisões recentes sobre tal tarefa em dados de redes sociais [6].

Neste trabalho apresentamos um estudo ortogonal aos esforços anteriores. Nosso objetivo aqui é analisar sentimentos sobre incidentes e violência envolvendo atuação policial comparando a abordagem tradicional baseada em modelos léxicos e modelos para processamento de linguagem natural usando aprendizado de máquina e profunda. Para isso, exploramos notícias de diferentes fontes e as usamos para coletar *tweets* e inferir sentimentos em um contexto ou incidente específico envolvendo ações policiais, ainda não explorado pelos trabalhos mencionados [4, 12]. Adicionalmente avançamos em relação aos trabalhos anteriores nesse contexto específico [1, 3, 9] com o uso de BERT para inferência de sentimentos, e diferente dessa nossa metodologia de coleta de *tweets* permite analisar opinião da população sobre um incidente de segurança ou ação policial de interesse específico.

3 BASES DE DADOS

Selecionamos notícias sobre incidentes de segurança com intervenção policial entre o período de 04/2019 até 12/2021 que foram amplamente divulgadas e repercutidas em mídia digital. Nesse sentido, pesquisamos por notícias nas seções sobre violência e crimes em três relevantes portais de notícias nacionais: UOL, Folha de São Paulo, e G1. Utilizamos como critério de seleção a participação dos leitores via comentários de texto no próprio portal de notícia. Especificamente, selecionamos notícias com mais de 700 comentários, marca notável e acima da média usual de comentários, o que intuitivamente indica uma grande repercussão em redes sociais. Ao todo foram selecionadas 8 notícias, e após a seleção dessas, realizamos a busca por comentários na rede social *Twitter*, através da API do *Twitter*, onde buscamos *tweets* relacionado aos links e os títulos das

notícias. Ao fim dessa etapa, coletamos um total de 16276 *tweets* sobre as notícias.

Após a etapa de coleta dos *tweets*, conduzimos o mínimo de pré-processamento de texto nos *tweets* coletados, removendo dados que dificultam esse processamento como *links*, quebras de linha, emojis, *tweets* curtos com somente uma palavra. E por fim realizamos a etapa de rotulação das instâncias de dados. O rótulo consiste na classificação de um *tweet* como *Positivo*, *Negativo* ou *Neutro* considerando o sentimento do usuário que o postou sobre a ação policial no incidente ao qual a notícia se refere. Seguimos a estratégia proposta em [7], onde *tweet Positivo* usa linguagem positiva em geral relacionada a otimismo, felicidade, orgulho e apoio. *Tweet Negativo* usa linguagem de tristeza, angústia, medo, incluindo ridicularização, sarcasmo e ódio. Por sua vez, *Tweet Neutro* não tem sentimento claro, sendo nenhuma ou ambas opções anteriores.

A rotulação serve de base para treinamento de modelos de inferência sendo realizada por humanos, como é usual na literatura. Nesse sentido, contamos com três voluntários (não especialistas¹) para rotular um subconjunto de *tweets*, dado a inviabilidade de rotulação da base de dados completa. Esse subconjunto é composto por 3000 *tweets* definidos aleatoriamente e proporcionalmente ao total de *tweets* por notícia. Cada voluntário rotulou 1000 *tweets*, obtendo o total de 3000 *tweets* rotulados. Avaliamos a qualidade da rotulação via métricas de concordância entre os três rotuladores. Nesse sentido, selecionamos outro subconjunto de 198 *tweets* aleatórios dentro do conjunto de *tweets* rotulados. Requisitamos então aos voluntários que rotulasse cada instância desse subconjunto, se não rotulada anteriormente. Dessa forma obtivemos a amostra de 198 *tweets* (7% dos dados) com cada instância rotulada por todos os voluntários. O percentual de concordância entre os rotuladores foi de 70,2%, ao passo que o índice *Fleiss Kappa* foi de 0,58, indicando uma concordância moderado entre os pesquisadores em comparação a outros trabalhos da literatura [4]. As informações sobre a distribuição dos rótulos por notícia e por *tweet* são mostradas na Figura 1.

4 METODOLOGIA

Nesta seção apresentamos os modelos utilizados nesse trabalho para análise de sentimento em relação a ações policiais e a metodologia para aplicá-los à base de dados. O primeiro é o SentiStrength [11], que se baseia em um dicionário léxico com milhares de palavras rotuladas em intensidade de sentimento por linguistas. O SentiStrength classifica o sentimento de uma sentença como positivo ou negativo em uma escala de -4 (muito negativo) a +4 (muito positivo) e 0 indica sentimento neutro. No uso do SentiStrength, consideramos intensidades -4 a -1 negativo, 0 neutro, +1 a +4 positivo, como ocorre no modo ternário dessa ferramenta.

Os demais modelos utilizam BERT (*Bidirectional Encoder Representations from Transformers*) [2], que é um algoritmo de aprendizado profundo desenvolvido pelo Google para processamento de linguagem natural. Um recurso importante do BERT é sua construção com representações contextuais de modelos pré-treinados, baseado na arquitetura de *transformers*. Neste trabalho utilizamos

¹Entendemos que o contexto das notícias é conhecido pela população amplamente, sendo desnecessário rotulação por um especialista.

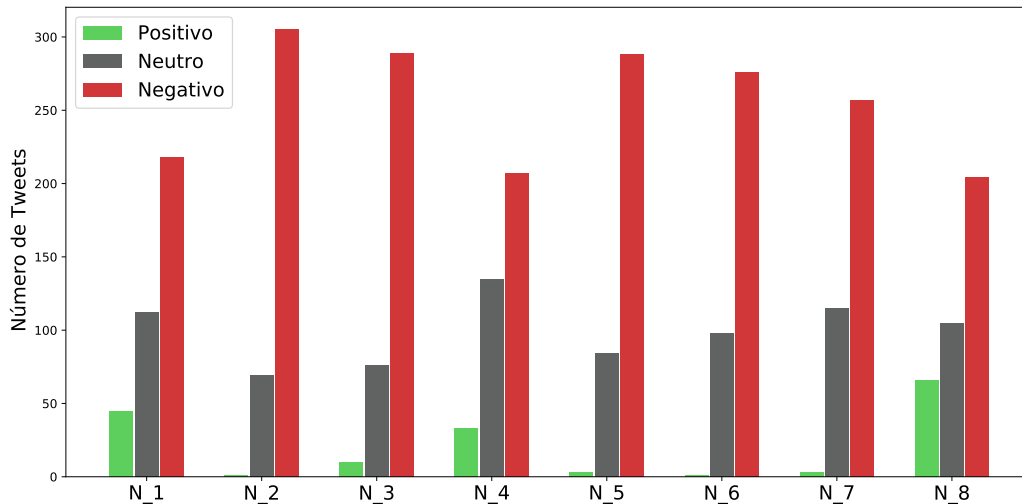


Figure 1: Distribuição dos rótulos por notícia.

dois modelos pré-treinados que são o *Multilingual* [2] e o *BERTimbau* [10]. Nesse caso, desenvolvemos uma rede neural, onde ambos os modelos pré-treinados foram retreinados com a base de dados como uma nova camada de neurônios na rede, o que denominamos como *modelos ajustados*, i.e., *fine tuning* para o contexto. Outra forma de aplicar o BERT é gerando uma matriz de *word embeddings*, que é uma representação quantitativa das características das palavras contidas em um comentário via processamento de linguagem natural. Utilizamos essa matriz para treinar dois modelos de classificação populares, baseados em aprendizagem de máquina: *Random Forest* (RF) e *Support Vector Machine* (SVM) com e sem balanceamento de classes via *smote* [5]. A parametrização padrão para os modelos de aprendizagem de máquina RF e SVM foi utilizada.

Avaliamos todos os modelos utilizando 20% dos *tweets* rotulados selecionados aleatoriamente, ao passo que os outros 80% foram utilizados para treinar os modelos. A avaliação consiste em medir o desempenho da inferência dos sentimentos via cinco métricas: acurácia, precisão, revocação, f1-score e f1-macro. A acurácia indica o percentual de *tweets* corretamente classificados, isto é, a soma acertos de todas as classes dividido pelo número total de *tweets* classificados. Já a precisão é calculada para cada classe individualmente (Negativo, Neutro e Positivo) e evidencia o percentual de *tweets* corretamente classificados para aquela classe. A revocação é calculada justamente pelo total de *tweets* corretamente classificados para uma classe sobre o total de *tweets* dessa classe. F1-score é a média harmônica entre precisão e revocação para cada classe, ao passo que o F1-macro é a média do F1-score considerando todas as classes.

Os códigos fontes, bem como a base de dados, deste trabalho estão disponíveis em repositório público.² Convidamos a comunidade científica e entusiastas sobre tema a acessá-los e comunicar com os autores em casos de dúvidas e sugestões para o trabalho.

5 RESULTADOS

Na Tabela 1, apresentamos os desempenhos dos modelos para análise de sentimento avaliados, iniciando do modelo léxico *SentiStrength*, a seguir os classificadores treinados com e sem balanceamento de classes (*smote*), e por fim as redes neurais. Para cada um desses modelos, apresentamos as cinco métricas de desempenho. Nota-se que a acurácia (*acc*) é a métrica, majoritariamente, com maior desempenho para os modelos apresentados. Contudo, a base de dados é desbalanceada, i.e., a classe negativa é majoritária, logo as métricas que capturam a especificidade (precisão, revocação e f1) dos modelos em identificar corretamente cada classe ganham importância, e predominam nas discussões nessa seção.

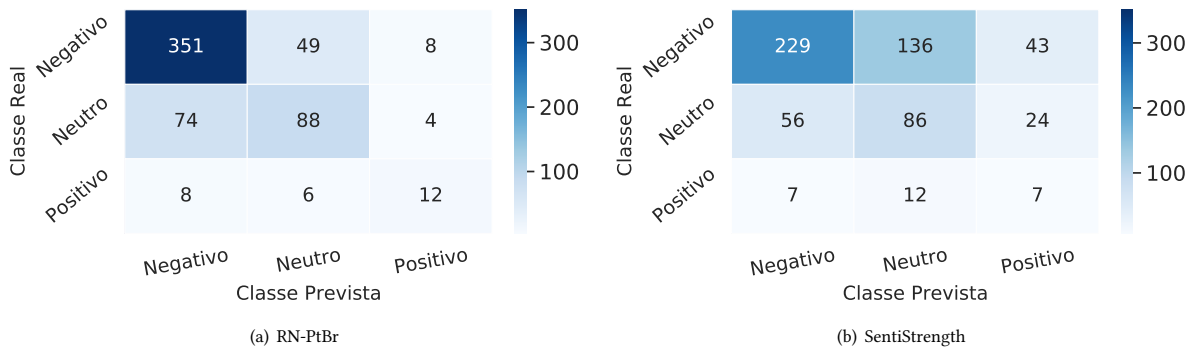
Iniciamos observando que *SentiStrength* obtém desempenho notavelmente inferior aos outros modelos tanto em acurácia como especificidade. Isso ocorre porque ele tende a neutralizar comentários negativos como gírias, ironias e sarcasmos típicos do contexto, como discutiremos a seguir. Logo, analisamos o desempenho dos modelos treinados na base de dados para identificar aquele com o melhor desempenho. Primeiramente, observamos que o tratamento de desbalanceamento para os classificadores RF e SVM aumentaram suas especificidades (f1) para as classes minoritárias neutro e positivo, aumentando por conseguinte a especificidade geral (F1-macro). Nesse caso, SVM tem o melhor desempenho alcançando F1-macro melhores (50-60%) que RF (41-50%).

A seguir, observamos que a língua em que os modelos BERT são pré-treinados, i.e., multi-lingual (MultiL) e português brasileiro (PtBr), é o fator mais importante para o desempenho dos modelos treinados com a base de dados, i.e., classificadores e redes neurais. O pré-treinamento na língua local dos *tweets* (português brasileiro) leva a ganhos de desempenho notáveis para esses modelos. Ao comparar os ganhos dos modelos PtBr em relação a MultiL, os classificadores RF e SVM alcançam aumentos em F1-macro de até 4

²https://github.com/LABPAAD/sentiment_analysis

Table 1: Desempenho dos modelos com as métricas: Precisão (P), Revocação (R), F1-score (F1), Acurácia (Acc) e F1-macro. Modelos de rede neural (RN) e classificadores (SVM e RF) utilizam BERT Multilingual (MultiL) e BERTimbau (PtBr).

Modelos	Negativo			Neutro			Positivo			Acc	f1-macro
	P	R	F1	P	R	F1	P	R	F1		
SentiStrength	0,78	0,56	0,65	0,37	0,52	0,43	0,37	0,52	0,43	0,54	0,50
SVM-MultiL-smote	0,78	0,80	0,79	0,53	0,50	0,51	0,43	0,35	0,38	0,70	0,56
SVM-MultiL	0,76	0,77	0,77	0,46	0,43	0,44	0,26	0,31	0,28	0,66	0,50
RF-MultiL-smote	0,72	0,88	0,79	0,48	0,26	0,34	0,50	0,15	0,24	0,68	0,45
RF-MultiL	0,72	0,97	0,83	0,74	0,20	0,32	1,00	0,04	0,07	0,72	0,41
SVM-PtBr-smote	0,81	0,80	0,80	0,54	0,59	0,56	0,56	0,35	0,43	0,72	0,60
SVM-PtBr	0,78	0,76	0,77	0,47	0,49	0,48	0,30	0,38	0,34	0,67	0,53
RF-PtBr-smote	0,75	0,88	0,81	0,55	0,36	0,43	0,57	0,15	0,24	0,71	0,50
RF-PtBr	0,73	0,95	0,83	0,63	0,28	0,38	1,00	0,04	0,07	0,72	0,43
RN-MultiL (ajustado)	0,77	0,85	0,81	0,55	0,39	0,46	0,23	0,27	0,25	0,70	0,50
RN-PtBr (ajustado)	0,81	0,86	0,83	0,61	0,53	0,57	0,50	0,46	0,48	0,75	0,63

**Figure 2: Matriz de confusão com erros de modelos.**

pontos percentuais, ao passo que a rede neural (RN) tem um aumento de 12 pontos percentuais. Isso evidencia que o treinamento de modelos com textos no domínio e língua específicas do contexto são aspectos fundamentais para análise de sentimentos. Nesse sentido, a rede neural RN-PtBr é o modelo com o melhor desempenho, alcançando F1-macro de 63% e uma acurácia de 75%. Por sua vez, o classificador SVM-PtBr se destaca com o segundo melhor desempenho, apresentando F1-macro de 60% e acurácia de 73%, se tornando também um modelo elegível para análise de sentimento na ausência de redes neurais.

Agora, analisamos os erros de classificação dos sentimentos, considerando o melhor modelo treinado para a base de dados (RN-PtBr) e o *SentiStrength* que é o modelo de análise léxica tipicamente utilizada para língua portuguesa. A Figura 2 reporta a matriz de confusão para ambos os modelos. Cada linha representa os *tweets* em uma classe real, enquanto cada coluna representa os *tweets* em uma classe prevista, o que nos permite analisar onde cada modelo mais erra e como isso acontece em função da classe.

De forma geral, é possível notar que, em termos absolutos, o maior desafio na tarefa aqui endereçada está em diferenciar as classes *Negativo* da *Neutro* e vice-versa. Nota-se que *SentiStrength* tende a neutralizar *tweets* com sentimento negativo, que são a vasta maioria, devido a especificidade de expressões e palavras negativas no contexto analisado. Para exemplificarmos a complexidade da tarefa, nós selecionamos alguns casos a partir da matriz de confusão focando especificamente nessas duas classes e apresentamos na Tabela 2.

Os exemplos apresentados mostram como *tweets* com algum tipo de sarcasmo ou um sentimento expresso por palavras do contexto em relação à ações policial, dificultam a tarefa de inferência aqui endereçada. Nós observamos que, de fato, esses casos representam uma larga fração dos *tweets* classificados incorretamente por ambos os métodos em especial por *SentiStrength*.

6 CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho, analisamos o sentimento das pessoas em relação a ações policiais, um contexto específico que compreende um dos principais desafios da sociedade. Para isso, definimos um conjunto de notícias com grande repercussão jornalística devido à atuação da polícia no Brasil a fim de avaliar o sentimento das pessoas via *tweets*. Comparamos sentimentos inferidos via o modelo baseado em dicionário léxico *SentiStrength* e modelos baseados em BERT para processamento de linguagem natural, especificamente, redes neurais ajustadas e classificadores treinados com representações semânticas (*embeddings*). Os resultados mostram que modelos BERT, em especial as redes neurais e o classificador SVM com *smote*, pré-treinados em língua local obtém melhor desempenho para inferir sentimentos nesse contexto, ao passo que *SentiStrength* tende a inferir sentimentos como neutro. Em suma, observamos a necessidade de treinar modelos com dados no contexto (comentários repercutindo ações policiais) para inferir adequadamente polaridade dos sentimentos. Isso é devido as características peculiares desse contexto, i.e., majoritariamente expressões com gírias de tom negativo e sarcástico

Table 2: Exemplos de tweets classificados incorretamente pelo modelo BERTimbau com ajuste fino (RN-PtBr) e o SentiStrength.

Tweet	Real	RN-PtBr	SentiStrength
defensora de bandidos !!!!	-1	-1	1
bandido bom é bandido morto	-1	-1	0
vai ter gente falando que é so um caso isolado, caso isolado o caralho	-1	0	0
os jornais não podem chamar de bandido, sequestrador, ou nominar qualquer termo, antes da pessoa ser julgada, pois eles podem ser processados	0	-1	-1
extrema emoção, é a nova era	0	0	1
isso mesmo, roger, fale a verdade pois a sua voz tem eco	0	0	1

como exemplificado na Tabela 2. Trabalhos futuros incluem ampliar nossa metodologia de coleta de dados com fontes jornalísticas diversificadas em estilos conservador e liberal e avaliar novos modelos classificadores e redes neurais com mais camadas de neurônios. Adicionalmente, planejamos explorar características sobre sentimentos de usuários em tweets via modelos BERT para predição de taxas de crimes.

ACKNOWLEDGMENTS

Este trabalho foi realizado com apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) processo no. 402194/2021-7.

REFERENCES

- [1] Luisa Fernanda Chaparro, Cristian Pulido, Jorge Rudas, Ana Maria Reyes, Jorge Victorino, Luz Ángela Narváez, Francisco Gómez, and Darwin Martinez. 2020. Sentiment analysis of social network content to characterize the perception of security. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE, The Hague, Netherlands, 685–691.
- [2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* 1 (2019). <https://doi.org/10.48550/arXiv.1810.04805>
- [3] Laura C Hand and Brandon D Ching. 2020. Maintaining neutrality: A sentiment analysis of police agency Facebook pages before and after a fatal officer-involved shooting of a citizen. *Government Information Quarterly* 37, 1 (2020), 101420. <https://doi.org/10.1016/j.giq.2019.101420>
- [4] Tamanna Hossain, Robert L Logan IV, Arjuna Ugarte, Yoshitomo Matsubara, Sean Young, and Sameer Singh. 2020. COVIDLies: Detecting COVID-19 Misinformation on Social Media. In *Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020*. NLP COVID-19 Workshop, Event Online. <https://doi.org/10.21275/ART20203995>
- [5] Batta Mahesh. 2020. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*, [Internet] 9 (2020), 381–386. <https://doi.org/10.21275/ART20203995>
- [6] Shervin Minaee, Nal Kalchbrenner, Erik Cambria, Narjes Nikzad, Meysam Chenaghlu, and Jianfeng Gao. 2021. Deep learning-based text classification: a comprehensive review. *ACM Computing Surveys (CSUR)* 54, 3 (2021), 1–40. <https://doi.org/10.1145/3439726>
- [7] Saif M Mohammad, Parinaz Sobhani, and Svetlana Kiritchenko. 2017. Stance and sentiment in tweets. *ACM Transactions on Internet Technology (TOIT)* 17, 3 (2017), 1–23. <https://doi.org/10.1145/3003433>
- [8] NEV-USP. 2022. Monitor da violência. (2022). Available in: <https://nev.prp.usp.br/projetos/projetos-especiais/monitor-da-violencia>. Access: August 2022.
- [9] Ashlin Oglesby-Neal, Emily Tiry, and K Kim. 2019. Public perceptions of police on social media. *Washington, DC: Urban Institute* 12 (2019).
- [10] Fábio Souza, Rodrigo Nogueira, and Roberto Lotufo. 2020. BERTimbau: pretrained BERT models for Brazilian Portuguese. In *Brazilian Conference on Intelligent Systems*. Springer International Publishing, Cham, 403–417. https://doi.org/10.1007/978-3-030-61377-8_28
- [11] Mike Thelwall. 2017. The Heart and soul of the web? Sentiment strength detection in the social web with SentiStrength. In *Cyberemotions*. Springer International Publishing, Cham, 119–134. https://doi.org/10.1007/978-3-319-43639-5_7
- [12] Maxwell Weinzierl, Suellen Hopfer, and S Harabagiu. 2021. Misinformation adoption or rejection in the era of covid-19. In *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, AAAI Press. ICWSM, Held Virtually, 787–795.



**TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DIGITAL NA BIBLIOTECA
“JOSÉ ALBANO DE MACEDO”**

Identificação do Tipo de Documento

- () Tese
() Dissertação
(X) Monografia
() Artigo

Eu, Marcos Paulo Fontes Feitosa, autorizo com base na Lei Federal nº 9.610 de 19 de Fevereiro de 1998 e na Lei nº 10.973 de 02 de dezembro de 2004, a biblioteca da Universidade Federal do Piauí a divulgar, gratuitamente, sem ressarcimento de direitos autorais, o texto integral da publicação Avaliação de Ferramentas Computacionais para Análise da Percepção das Pessoas sobre Violência Urbana de minha autoria, em formato PDF, para fins de leitura e/ou impressão, pela internet a título de divulgação da produção científica gerada pela Universidade.

Picos-PI 29 de Março de 2023.

Marcos Paulo Fontes Feitosa

Assinatura

Assinatura