



SEI-SICITE 2021

Pesquisa e Extensão para um mundo em transformação

# Reconhecimento de padrões com uso de técnicas de machine learning: estudo de caso sobre violência de gênero no ambiente universitário

*Pattern recognition using machine learning techniques: a case study on gender violence in the university environment*

Gracielle Silva Teixeira (orientada)\*, Maria Cristina Cavaleiro†,  
Elisângela Aparecida da Silva Lizzi (orientadora)‡

## RESUMO

A violência de gênero entende-se como qualquer ato de agressão física, psicológica ou sexual, incluindo ameaças e privação da liberdade contra algum indivíduo por conta de sua identidade de gênero, ou orientação sexual. Alguns fatores induzem à prática da violência, estando entre eles as concepções sociais imperceptíveis que foram estruturadas durante séculos. Dessa forma, este estudo tem por objetivo trabalhar com os dados, os quais foram obtidos anteriormente por meio de questionário online, sobre a violência de gênero no ambiente universitário, para compreender como esta pode sofrer interferência de variáveis. A população estudada refere-se aos alunos de graduação de uma universidade paranaense, tratando-se de um estudo transversal, com abordagem quantitativa. Portanto, observa-se que mulheres e minorias de gênero identificam o racismo na sociedade, ademais, pessoas que sofreram com o preconceito na universidade têm mais facilidade para identificar o racismo e o machismo na sociedade, e que a maioria das pessoas que nunca foram acometidas pelo preconceito possuem preconceito LGBT e contra negros. Assim sendo, nota-se que as vivências sociais estão relacionadas com a visão que o indivíduo terá da sociedade e que os métodos estatísticos de *machine learning* serviram para reconhecer e identificar os padrões dentro do estudo.

**Palavras-chave:** Violência de gênero, *machine learning*, estatística aplicada, survey.

## ABSTRACT

Gender violence is understood as any act of physical, psychological or sexual aggression, including threats and deprivation of liberty against an individual on account of their gender identity or sexual orientation. Some factors lead to the practice of violence, among them the imperceptible social conceptions that were structured for centuries. Thus, this study aims to work with data, which were previously obtained through an online questionnaire, on gender violence in the university environment, in order to understand how it can be influenced by variables. The studied population refers to undergraduate students at a university in Paraná, which is a cross-sectional study with a quantitative approach. Therefore, it is observed that women and gender minorities identify racism in society, moreover, people who suffered from prejudice at the university find it easier to identify racism and sexism in society, and that most people who have never been affected because of prejudice, they have prejudice against LGBT and against black people. Therefore, it is noted that social experiences are related to the view that the individual will have of society and that statistical methods of machine learning served to recognize and identify patterns within the study.

**Keywords:** Gender violence, machine learning, applied statistics, survey.

\* Engenharia elétrica, Universidade Tecnológica Federal do Paraná, Cornélio Procópio, Paraná, Brasil; [gracielle@alunos.utfpr.edu.br](mailto:gracielle@alunos.utfpr.edu.br)

† Universidade Estadual do Norte do Paraná-Centro de Ciências Humanas e Educação-UENP, Campus Cornélio Procópio; [mariacristina@uenp.edu.br](mailto:mariacristina@uenp.edu.br); <https://orcid.org/0000-0003-0461-0546>

‡ Universidade Tecnológica Federal do Paraná, Campus Cornélio Procópio; [elisangelalizzi@utfpr.edu.br](mailto:elisangelalizzi@utfpr.edu.br); <https://orcid.org/0000-0001-7064-263X>



## 1 INTRODUÇÃO

A violência de gênero entende-se como qualquer ato de agressão física, psicológica ou sexual, incluindo ameaças e privação da liberdade contra algum indivíduo por conta de sua identidade de gênero, ou orientação sexual (POLITIZE, 2021). Por sua vez, o conceito de violência de gênero também recebe diferentes definições, evidenciando a complexidade do fenômeno. De acordo com a definição de Scott (1990, p.16) “o gênero é o primeiro campo no seio do qual, ou por meio do qual, o poder é articulado”. Nessa perspectiva, o uso da expressão violência de gênero “designa a produção da violência em um contexto de relações produzidas socialmente” (ALMEIDA, 2007, p.24), sustentada em um quadro que expõe múltiplas opressões, que “vão se configurando a partir de disputas simbólicas e materiais, processadas, dentre outros espaços, na (...) família, escola, igreja e os meios de comunicação – e materializadas, ainda, nas relações de trabalho, no quadro político-partidário, nas relações sindicais e na divisão sexual do trabalho” (ALMEIDA, 2007, p.28). Tal violência possui maior ocorrência em mulheres, porém, acontece também com homens, minorias sexuais e de gênero (POLITIZE, 2021).

Alguns fatores induzem à prática da violência, estando entre eles as concepções sociais imperceptíveis que foram estruturadas durante séculos, como a dominação masculina e a sua necessidade da representação de força, superioridade e controle sobre mulheres e minorias (POLITIZE, 2021). Para que este mal diminua a cada dia é preciso estratégias de prevenção primária, isto é, formas que impeçam a violência de acontecer, as quais combinam o empoderamento da mulher na economia à formação em igualdade de gênero, e a comunicação dentro da comunidade que ela esteja inserida (BRASIL, 2021). Desta maneira, fez-se necessário compreender a gravidade da violência de gênero e fazer deste assunto pauta para a pesquisa. Para isso, optou-se por utilizar técnicas de *machine learning*, as quais são bastante usadas para reconhecer padrões, para que fosse possível atingir a resposta do seguinte problema, “como várias características interferem na violência de gênero no ambiente universitário?”, pois este será o assunto tratado e estudado no decorrer deste trabalho. Portanto este estudo, tem por objetivo trabalhar com os dados obtidos sobre a violência de gênero no ambiente universitário, compreender como esta violência pode sofrer interferência de variáveis, e mostrar tabelas e gráficos para a exposição dos resultados obtidos.

## 2 MÉTODO

A população estudada refere-se aos alunos de graduação da uma universidade regional do norte paranaense, somente foram incluídos estudantes regulares, maiores de 18 anos e que estivessem matriculados até o terceiro ano ou quinto período do curso.

Trata-se de um estudo transversal, com abordagem quantitativa. Os dados utilizados foram anteriormente coletados, no período de setembro a novembro de 2019, por meio de um questionário desenvolvido por pesquisadores da UENP e UTFPR, o qual trata sobre a violência de gênero no âmbito universitário dos alunos da UENP. Esta pesquisa foi aprovada pelo Comitê de Ética em Pesquisas (CEP). Após a coleta de dados, fez-se uma consolidação do banco de dados obtidos em planilhas eletrônicas e posterior implementação computacional desenvolvida dentro do ambiente de programação do programa R (versão 3.9), desta maneira, o tamanho amostral analítico ficou com 647 respondentes. Em todas as análises utilizou-se um nível de significância de 5%.



Nesta pesquisa foi utilizado os métodos de *machine learning* supervisionados, os quais englobaram a regressão logística e a árvore de regressão.

O modelo de regressão logística busca estimar a probabilidade da variável dependente, neste caso a violência de gênero no ambiente universitário, em uma categoria em função do conhecimento das variáveis independentes. Para descrever o modelo da regressão logística deve-se saber que tudo é baseado na função logística, onde a probabilidade de ocorrência de um evento pode ser estimada diretamente. Para estimar tais coeficientes é utilizado o método da Máxima Verossimilhança, ou seja, encontra-se uma combinação de coeficientes que maximiza a probabilidade de algo ter sido realmente observado (AMARAL, 2021).

A árvore de decisão se trata de uma ferramenta com uma variável resposta pré-definida, ela ajuda na tomada de decisão, retornando um gráfico parecido com uma árvore. Nela, é dividido a amostra em dois ou mais conjuntos homogêneos, esta ferramenta se divide em árvores de classificação, quando a variável dependente é categórica, e de regressão, quando a variável dependente é contínua. A árvore possui uma estrutura típica, a qual é, os nós internos, as folhas e os ramos, que permite uma leitura visual bastante intuitiva sobre a resposta em estudo (BRESSAN, AZEVEDO e LIZZI, 2017).

### 3 RESULTADOS

Neste capítulo serão exibidos os resultados, de forma resumida, obtidos em cada hipótese de estudos estruturada previamente.

#### 3.1 Descrição geral e tabulação cruzada

Nesta seção será exibido a tabulação, a qual cruza quantas pessoas se identificam com cada uma das alternativas propostas nas questões respondidas, e o quadro de regressão logística. A sigla “I.S.” utilizada no quadro de regressão logística refere-se à informação insuficiente, percebe-se que como se trata de uma divisão, em caselas que há presença de zero não é possível proceder com os cálculos.

**Quadro 1- Resultados da regressão logística para as variáveis de interesse.**

<b>Resultados da regressão logística para as variáveis de interesse.</b>						
<b>Variável resposta: Sociedade brasileira racista</b>						
<b>Variáveis</b>	<b>Sociedade brasileira racista</b>		<b>Odds Ratio-Bruto</b>	<b>I.C. (95%)</b>	<b>Odds Ratio-Ajustado</b>	<b>I.C. (95%)</b>
	<b>Não</b>	<b>Sim</b>				
<b>Sexo</b>						
Feminino	14(0,44)	487(0,79)	Ref.	Ref.	Ref.	Ref.
Masculino	18(0,56)	128(0,21)	0,20	(0,10;0,42)	0,15	(0,07;3,30)



Sexualidade						
Heterossexual	30(0,94)	432(0,70)	Ref.	Ref.	Ref.	Ref.
Homossexual	1(0,03)	33(0,05)	2,29	(0,47;41,44)	6,30	(1,12;1,19)
Bissexual	1(0,03)	119(0,19)	8,26	(1,74;147,93)	6,35	(1,28;1,15)
Lésbica	0(0)	17(0,03)	I.S.	I.S.	I.S.	(0,00;6,09)
Não quer declarar	0(0)	5(0,01)	I.S.	I.S.	I.S.	I.S.
Não sabe	0(0)	9(0,01)	I.S.	I.S.	I.S.	I.S.

Fonte: Autoria própria (2021).

No Quadro 1 observa-se que os homens possuem menor chance de identificarem o racismo na sociedade em comparação com as mulheres, e pessoas bissexuais têm 8 vezes mais chance de identificarem o racismo na sociedade em relação aos heterossexuais. Infere-se que a causa destes resultados são os fatores e vivências sociais, pois tanto mulheres quanto bissexuais se tratam de uma população mais exposta a preconceitos, dessa forma, conseguem enxergar claramente o racismo presente na sociedade brasileira.

**Quadro 2 - Tabulação cruzada do sofrimento ou não pela violência de gênero e as variáveis de interesse.**

		Não sofreu violência de gênero	Sofreu violência de gênero	Valor p
Preconceito LGBT	Não tenho preconceito	457(91,40%)	134(91,16%)	0,99
	Tenho preconceito	43(8,60%)	13(8,84%)	
Total		500(100,00%)	147(100,00%)	
Identificou violência ocasionada por fatores sexuais	Não	290(58,00%)	48(32,65%)	<0,001
	Sim	210(42,00%)	99(67,35%)	
Total		500(100,00%)	147(100,00%)	
Sociedade Machista	Não	25(5,00%)	4(2,72%)	0,34
	Sim	475(95,00%)	143(97,28%)	
Total		500(100,00%)	147(100,00%)	
Sociedade Racista	Não	29(5,80%)	3(2,04%)	0,10
	Sim	471(94,20%)	144(97,96%)	
Total		500(100,00%)	147(100,00%)	
Preconceito contra negros	Não tenho preconceito	474(94,80%)	140(95,24%)	0,99
	Tenho preconceito	26(5,20%)	7(4,76%)	
Total		500(100,00%)	147(100,00%)	

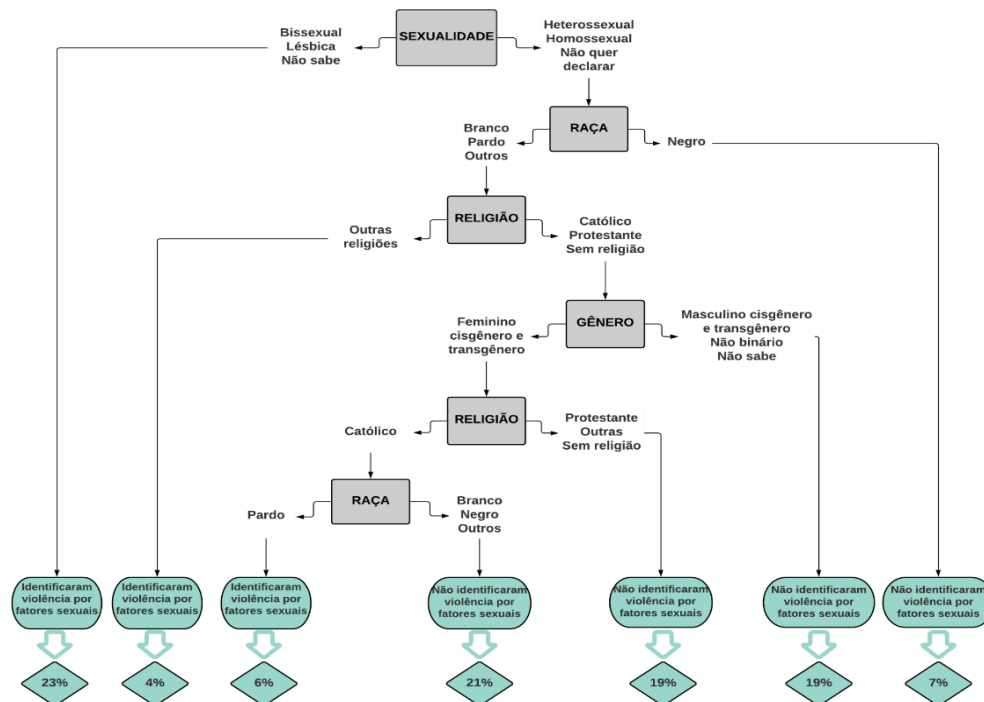
Fonte: Autoria própria (2021).

No Quadro 2 observa-se que a maioria das pessoas que têm preconceito LGBT (43 pessoas) e preconceito contra negros (26 pessoas) nunca sofreram violência de gênero na universidade. Subentende-se que tais indivíduos persistam com ações e pensamentos preconceituosos pois nunca sofreram com as dificuldades geradas pelo preconceito e, além disso, não demonstram empatia. Entre as pessoas que sofreram violência de gênero na universidade nota-se que os indivíduos que identificaram violência ocasionada por fatores sexuais na universidade (67,35%), que acham a sociedade machista (97,28%) e que acham a sociedade racista (97,96%) são a maioria. Isto se deve, pois, estas pessoas que sofreram com o preconceito têm mais facilidade para identificá-lo na sociedade.

### 3.2 Modelos gerados

Nesta seção será exibida a Árvores de Regressão, a qual demonstra como diferentes questões respondidas associam-se.

Gráfico 1 - Árvore de regressão logística - identificação da violência ocasionada por fatores sexuais na universidade.



Fonte: Autoria própria (2021).

No Gráfico 1 observa-se que 33% dos indivíduos identificaram a violência causada por fatores sexuais dentro da universidade. Destes citados, 23% são bissexuais, lésbicas ou não sabem sua sexualidade. 66% não identificaram a violência causada por fatores sexuais dentro da universidade. Destes citados, 21% são heterossexuais, homossexuais ou que não querem declarar sua sexualidade, são brancos, pardos ou outros (amarelos e indígenas), são católicos, protestantes, e sem religiões, são mulheres cisgêneros e transgêneros, são católicos, e são brancos, negros ou outros (amarelos e indígenas). Esta leitura do gráfico é feita de forma análoga para as outras ramificações do mesmo.



## 4 CONCLUSÃO

De acordo com os resultados obtidos sobre a violência de gênero no ambiente universitário dos estudantes da universidade e a análise realizada, pode-se observar que entre os achados apresentados destacam-se, de forma geral, que mulheres e minorias de gênero identificam e possuem maior chance de identificarem o racismo na sociedade, ademais, pessoas que sofreram com o preconceito na universidade têm mais facilidade para identificar o racismo e o machismo na sociedade, e que a maioria das pessoas que nunca foram acometidas pelo preconceito possuem preconceito LGBT e contra negros. Assim sendo, nota-se que as vivências sociais, uma maior exposição ao preconceito, estão relacionadas com a visão que o indivíduo terá da sociedade e consequentemente influenciará nas suas ações.

Desta forma, verificou-se que realmente há violência de gênero nas universidades, e checkou-se que os métodos estatísticos de *machine learning*, especificamente a regressão logística e a árvore de regressão, serviram para reconhecer e identificar os padrões dentro do estudo, ou seja, quais indivíduos identificam a violência e porquê, fatos citados acima. Portanto, este estudo atingiu os objetivos traçados e as análises dos dados colaboram para indagações de possíveis medidas que podem ser tomadas para a luta contra os preconceitos, bem como agregar para conhecimento pessoal ou contribuição científica na área.

## AGRADECIMENTOS

Agradeço à Professora Elisângela Lizzi, pela dedicação e paciência com as quais me instruiu para o desenvolvimento deste trabalho. A minha família. E principalmente a Deus, o qual é a fonte da minha vida.

## REFERÊNCIAS

ALMEIDA, Suely de S. Essa Violência mal-dita. In: ALMEIDA, Suely de S. (Org.). **Violência de gênero e políticas públicas**. Rio de Janeiro: Editora da UFRJ, 2007.

AMARAL, Ernesto F. L. **Aula extra-Análise de Regressão Logística**. 2012. Disponível em: <<http://ernestoamaral.com/docs/dcp854b-122/Aula33.pdf>> Acesso em: 20 abr. 2021.

BRASIL; ORGANIZAÇÃO PAN-AMERICANA DA SAUDE; ORGANIZAÇÃO MUNDIAL DA SAUDE. **Violência contra as mulheres**. Disponível em: <<https://www.paho.org/pt/topics/violence-against-women>> Acesso em: 20 abr. 2021.

BRESSAN, Glaucia M.; DE AZEVEDO, Beatriz CF; LIZZI, E. A. S. **A decision tree approach for the musical genres classification**. *Applied Mathematics & Information Sciences*, v. 11, n. 6, p. 1703-1713, 2017.

POLITIZE. **O que é violência de gênero e como se manifesta?**. 2020. Disponível em: <<https://www.politize.com.br/violencia-de-genero-2/>> Acesso em: 20 abr. 2021.

SCOTT, Joan W. **Gênero: uma categoria útil de análise histórica**. *Educação e Realidade*, Porto Alegre, n. 2, p. 5-22, jul./dez. 1990. Disponível em: <<https://seer.ufrgs.br/index.php/educacaoerealidade/article/view/71721>> Acesso em: 20 abr. 2021.