

Análise comparativa de vídeos do YouTube de canais de homens e mulheres na área de computação: uma investigação a partir dos comentários

Mateus Barros Macedo¹, Blenda Gabrielly Souza Morais¹,
Lucas de Almeida Ribeiro², Newarney Torrezão da Costa¹,
Thalia Santos de Santana³, Cleon Pereira Junior¹

¹Instituto Federal Goiano (IF Goiano) – Campus Iporá – Iporá-GO – Brasil

²Instituto Federal de Goiás (IFG) – Campus Luziânia – Luziânia-GO – Brasil

³Universidade Federal de Goiás (UFG) – Goiânia-GO – Brasil

{barrosmateusmacedo, blendagabrielly71, thaliassantana15}@gmail.com

lucas.ribeiro@ifg.edu.br, {newarney.costa, cleon.junior}@ifgoiano.edu.br

Abstract. *This work analyzes YouTube videos from male and female computer science channels, focusing on comments. The research aims to understand possible behavioral and interaction differences between users when consuming content produced by different genders. Sixty videos with at least 50 comments were chosen (30 from women and 30 from men). From the comments, sentiment analysis and other quantitative metrics were carried out. In addition, a qualitative analysis was carried out which made it possible to verify trends in sexist comments on material by women. This research contributes to the understanding of gender dynamics in knowledge sharing videos and can help create more inclusive and equitable strategies in computer science education.*

Resumo. *Este artigo analisa vídeos do YouTube de canais de homens e mulheres na área de computação, com foco nos comentários. A pesquisa visa compreender possíveis diferenças comportamentais e de interação entre os usuários ao consumir conteúdo produzido por diferentes gêneros. Foram escolhidos 60 vídeos com pelo menos 50 comentários (30 de mulheres e 30 de homens). A partir dos comentários, foram realizadas análises de sentimentos e outras métricas quantitativas. Além disso, efetuou-se uma análise qualitativa que permitiu verificar tendências de comentários de cunho sexista em materiais de mulheres. Esta pesquisa contribui para a compreensão das dinâmicas de gênero no compartilhamento de conhecimento em vídeos e pode auxiliar na criação de estratégias mais inclusivas e equânimes na educação em computação.*

1. Introdução

O uso de vídeos para complementar o conhecimento em uma determinada área de formação é algo recorrente na educação [Ebied et al. 2016, Maziriri et al. 2020, Lin 2022]. De fato, a aprendizagem baseada em vídeo, conhecida por *Video Based Learning* (VBL) é bastante explorada e apresenta como vantagens, por exemplo, a liberdade para pausar ou repetir um determinado conteúdo [Sablíć et al. 2021]. Nesse aspecto, a

área de computação tem se beneficiado com o uso de vídeos em plataformas como YouTube¹ e TikTok² para disseminar conhecimentos acerca de tecnologias e programação de computadores [Kadriu et al. 2020, Garcia et al. 2022].

A computação é uma área que apresenta diversas vertentes de trabalho, dentre elas, a empresarial e a acadêmica [Cardoso and de David 2016]. A crescente procura por conteúdo relacionado à computação impulsionou significativamente a produção de materiais didáticos com tal temática, principalmente em formato de vídeo. Tais produções se tornaram um arcabouço de informações acessíveis, promovendo a disseminação de conhecimento de forma dinâmica [Barroso and Antunes 2020]. Entretanto, um dado relevante é que, mesmo em crescimento, a área ainda é composta majoritariamente por pessoas do gênero masculino.

A exemplo disso, dados apresentados pela Sociedade Brasileira de Computação (SBC), considerando uma série temporal entre os anos de 2001 e 2020, demonstram o baixa quantidade de concluintes do gênero feminino em cursos desta área de conhecimento [SBC 2023]. Tal fato pode ser justificado pelo conjunto de estigmas e estereótipos impostos implicitamente às mulheres [Louzada et al. 2014]. Mais, a quantidade reduzida de mulheres na computação, se comparada com a quantidade de homens, pode ser motivada também por fatores culturais, visto que pessoas do gênero feminino desde muito novas são direcionadas a se interessar e/ou gostar de assuntos que envolvem cuidado ou comunicação [Santos et al. 2021].

Além da disparidade de gênero na computação, também se nota uma diferença na quantidade de canais que produzem vídeos voltados à área de tecnologia de pessoas do gênero masculino em relação ao gênero feminino. Com relação à área de Ciência, Tecnologia, Engenharia e Matemática (STEM³), notam-se poucos canais de mulheres dentre os mais populares [Amarasekara and Grant 2019]. Outrossim, estudos apontam que, mesmo em canais de entretenimento, em geral, canais do gênero masculino são maioria entre os mais visualizados e com maior quantidade de inscritos [Regueira et al. 2020]. Em decorrência desses fatores, espera-se uma quantidade ainda menor de mulheres assumindo o protagonismo no compartilhamento de conhecimento por vídeo.

Mesmo em minoria e geralmente não sendo exibidos como primeiros resultados de busca no YouTube, é possível encontrar canais de mulheres discutindo temáticas dentro da computação. Nesse sentido, o YouTube permite a coleta de dados que viabilizam análises comportamentais de usuários com alguma forma de interação com os vídeos. Sendo assim, é oportuno realizar comparações com vídeos de canais de pessoas do gênero feminino e do masculino. Especificamente no YouTube, há possibilidade de analisar comentários, curtidas, visualizações, dentre outras variáveis e metadados que podem gerar discussões com relação aos materiais hospedados nessa plataforma [Ferreira et al. 2020].

Ao relacionar questões de gênero, sabe-se que as mídias sociais também refletem as discriminações diretas e indiretas que permeiam a sociedade, evidenciando a capacidade das redes de refletir as problemáticas sociais no meio digital [Moura 2016]. Com

¹<https://www.youtube.com>

²<https://www.tiktok.com>

³Sigla oriunda do inglês (*Science, Technology, Engineering and Mathematics*) para representar o currículo das quatro áreas em uma abordagem interdisciplinar e aplicada.

relação às mulheres, em específico no ambiente virtual, a violência simbólica lhes é direcionada principalmente na forma de menosprezo, assédio e xingamentos/ofensas que se veiculam em mecanismos de interação em ambientes virtuais [Santana et al. 2022]. Com o crescimento de VBL e as questões de gênero presentes na sociedade, é relevante o desenvolvimento de pesquisas com enfoque na análise do comportamento de pessoas a partir da interação com materiais produzidos por homens e por mulheres.

Uma vez que há a presença de mulheres no YouTube desempenhando o papel de disseminar conhecimento sobre tópicos de computação, mesmo que em menor predominância do que os criadores de conteúdo do gênero masculino, e considerando problemas de gênero que permeiam essa área do conhecimento, esta pesquisa visa realizar uma análise comparativa em vídeos de canais de pessoas do gênero masculino e do gênero feminino, com ênfase nos comentários. Mediante um levantamento de 60 vídeos (30 produzidos por mulheres e 30 produzidos por homens), este trabalho traz algumas comparações, tais como análise de sentimentos, a partir dos vídeos e apresenta discussões acerca dos resultados encontrados.

2. Fundamentação teórica

O YouTube é uma plataforma de vídeos *online* que, a partir de uma *Application Programming Interface* (API), permite que sejam extraídos metadados e outras informações relevantes acerca de seus conteúdos disponíveis. Entende-se por metadados como um conjunto de dados que fornecem informações resumidas acerca de alguma entidade. Popularmente, metadados são definidos como dados sobre dados [Pal et al. 2019]. Com a ampliação das possibilidades de interação (a partir de comentários, visualizações e curtidas) em vídeos hospedados no YouTube, atualmente é possível classificar a plataforma como uma mídia social [Hou 2019]. A partir dessa classificação, é pertinente analisar o comportamento de pessoas usuárias que interagem com a plataforma [Alhujaili and Yafooz 2021]. Neste aspecto, técnicas de Inteligência Artificial podem ser aplicadas para automatizar esse processo.

A análise de sentimentos é uma linha de pesquisa investigada por intermédio da área de mineração de textos [Medhat et al. 2014]. Uma das possibilidades desta linha está concentrada na investigação para classificar mensagens no formato de texto como uma mensagem positiva, negativa ou neutra. Tal investigação é comumente utilizada em mídias sociais [El Rahman et al. 2019]. Além disso, a partir da linha de análise de sentimentos, é possível investigar a aprendizagem por transferência, detectar emoções e construir recursos [Medhat et al. 2014].

É possível encontrar diferentes algoritmos e APIs que utilizam esses algoritmos para realizar análise de sentimentos em textos [Abu-Salih et al. 2023]. A partir de um conjunto de textos, pesquisas apresentam a acurácia desses algoritmos e APIs em diferentes domínios [Diaz Jr 2023]. Nesse aspecto, antes de tomar a decisão sobre qual ferramenta utilizar para a classificação, é importante fazer um levantamento da literatura. O trabalho de [Diaz Jr 2023] apresenta um comparativo de diferentes APIs. No domínio analisado, os resultados são positivos para o uso da Komprehend. A Komprehend é uma API desenvolvida para análise de sentimento em documentos. Ela classifica de 0 a 1 os sentimentos positivos, negativos e neutros – sendo que sentimento acima de 0.5 é compreendido como o predominante naquele conjunto de textos.

Outra API que tem sido utilizada com frequência para análise de textos, graças ao seu poder de processamento, é a *Generative Pre-Training* (GPT). Atualmente, a GPT-4⁴ está disponível para pessoas pesquisadoras a partir de pagamento de cotas de uso. A empresa responsável pelo GPT disponibiliza também um chatbot, chamado *chatgpt*⁵ que permite diálogo com uma interface alimentada pelo GPT. Pesquisas recentes apontam o potencial do GPT para análise de sentimentos [Mathew and Bindu 2020, Lammerse et al. 2022].

O trabalho de [Bringula et al. 2020] apresenta *clusters* e sentimentos de *feedback* dos usuários do YouTube ao aprender a programar em Python e C++. A pesquisa analisou 2583 comentários. Os resultados das análises de *cluster k-means* mostraram que os grupos de *feedback* são semelhantes tanto para os vídeos que ensinam C++ quanto para os vídeos que ensinam Python. Foram expressos pelos usuários sentimentos positivos em relação aos vídeos tutoriais. Além disso, as palavras que *thank* e *thanks* (que significam agradecimentos) foram as mais frequentes. Na mesma linha, [Miranda and Martin 2020] apresentam uma análise de sentimento acerca dos comentários em vídeos que envolvem aprendizagem de artes marciais.

Considerando aspectos relacionados a gênero, a pesquisa de [Amarasekara and Grant 2019] analisou vídeos de STEM e diferentes comportamentos a partir do gênero da pessoa que estava realizando a comunicação no vídeo. Nesta pesquisa, verificou-se que os canais cujas mulheres eram responsáveis pela comunicação acumulavam mais comentários por visualização e proporções significativamente mais altas de comentários hostis ou acerca da aparência. Entretanto, tais canais também tiveram mais comentários críticos/negativos e de cunho sexistas/sexuais.

Baseado nos resultados existentes na literatura e visando avançar o estado da arte que discute gênero, computação, e o relacionamento entre o comportamento de pessoas usuárias em mídias sociais, esta pesquisa visa investigar comportamentos apresentados em vídeos produzidos por pessoas do gênero masculino com relação ao gênero feminino. Para isso, técnicas de análise de sentimentos foram aplicadas a fim de tecer discussões que permitem ir além dos resultados quantitativos. A seguir, é apresentada a metodologia aplicada para este estudo.

3. Materiais e métodos

Esta pesquisa visa realizar análises quantitativa e qualitativa a partir das questões de gênero e ensino de programação utilizando VBL. Para isso, foram estabelecidas etapas para alcançar os objetivos propostos. A Figura 1 apresenta as etapas da pesquisa.

Inicialmente, foram definidos os parâmetros de busca na plataforma YouTube com o intuito de delimitar o escopo da pesquisa. Conforme apresentado nas Seções 1 e 2, os canais cujas pessoas apresentadoras são do gênero feminino possuem, em geral, menos inscrições. Neste aspecto, foi feita uma busca exploratória por canais de mulheres que falavam sobre tecnologia⁶. A partir desses canais, foi possível obter a quantidade de inscrições por canal. Com isso, foram selecionados os 5 canais produzidos por pessoas do gênero feminino com o maior quantidade de inscrições. Outros critérios de inclusão foram

⁴<https://openai.com/gpt-4>

⁵<https://chat.openai.com>

⁶A lista completa dos canais analisados pode ser conferida clicando aqui

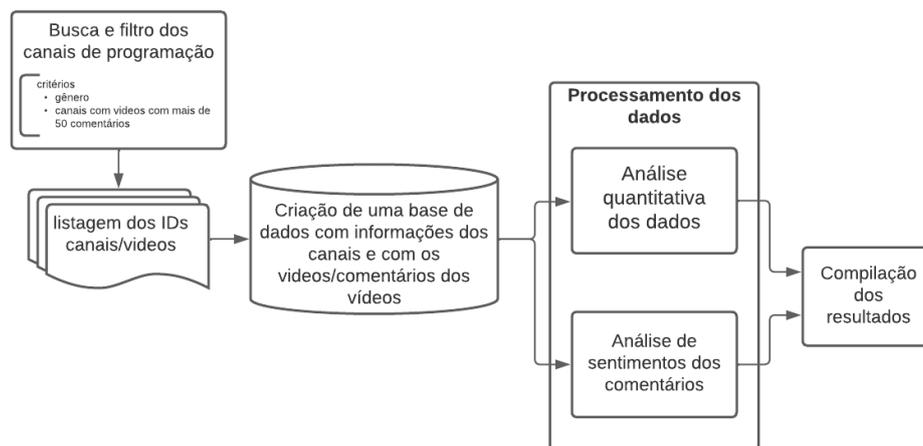


Figura 1. Etapas da pesquisa.

estabelecidos, sendo eles: i) os canais deveriam possuir no mínimo 6 vídeos com mais de 50 comentários em cada vídeo, e ii) os vídeos deveriam ter como conteúdo principal o ensino de conteúdos de programação. Após a aplicação dos critérios, foi realizada a busca por vídeos em canais de pessoas do gênero masculino, visando manter quantidades próximas as de pessoas inscritas e que cumpriam os critérios de inclusão. Também foram selecionados 5 canais.

Ao final da etapa de seleção, baseada nos critérios explicitados anteriormente, foi criada uma base com 5 canais apresentados por homens e 5 canais apresentados por mulheres. Para cada canal, foram selecionados manualmente 6 vídeos considerando os critérios de inclusão. Após definir os canais que atendiam os critérios do estudo, foram pesquisados os IDs dos canais, posteriormente utilizados como parâmetro de requisição na API do YouTube (V3). Para a extração dos comentários dos vídeos, foi utilizada a API do YouTube, onde, para cada um dos vídeos selecionados, efetuou-se uma requisição na API com o parâmetro de no máximo 50 comentários, e os resultados dessas requisições formularam a base de dados utilizada. Vale ressaltar que em alguns casos, mesmo o vídeo possuindo mais de 50 comentários, a requisição retornou menos de 50 comentários.

A etapa seguinte consistiu em aspectos de análise quantitativa. A partir dos comentários extraídos, foi calculada a média de quantidade de caracteres por comentário, a quantidade de perguntas realizadas em vídeos de homens e em vídeo de mulheres. Essas métricas têm por objetivo quantificar a interação em vídeos por meio de comentários.

Com base na literatura, a princípio foram selecionadas a API Komprehend e a API GPT (a partir da interface com o ChatGPT – por ser uma versão gratuita). Para auxiliar na decisão sobre qual API seria melhor para a análise do domínio português e comentários de vídeos no YouTube, algumas análises estatísticas foram realizadas. Para verificar a eficácia dos modelos utilizados, foram selecionados os 10 últimos comentários do retorno de cada vídeo e efetuada uma análise manual desta base. A base foi composta por 300 comentários retirados dos vídeos de homens e 300 comentários de vídeos de mulheres. Nessa etapa, duas pessoas classificaram os sentimentos em positivo, negativo ou neutro. Posteriormente, foram utilizadas métricas de verificação estatística de acurácia, *recall* e

F1-score, dos resultados dos modelos, tomando como base de comparação a classificação das pessoas avaliadoras. A Tabela 1 mostra o resultado da análise. Com relação ao recorte de base de 600 comentários, notou-se que o GPT, a partir do uso do ChatGPT, realizou uma classificação superior ao Komprehend. Deste modo, optou-se pelo uso do GPT para realizar o percentual dos sentimentos.

Tabela 1. Comparativo de API para análise de sentimento.

	acurácia	precisão	recall	f1-score
GPT (ChatGPT)	73,33%	79,74%	73,33%	71,33%
Komprehend (komprehend.io)	50,00%	55,85%	50,00%	49,28%

Visando a melhor compreensão dos resultados, também foram realizadas análises qualitativas. Ainda com o apoio do GPT, a partir do ChatGPT, foi solicitada a extração dos adjetivos dos comentários dos vídeos da base de comentários selecionada para este trabalho. Os adjetivos foram coletados e processados para gerar nuvens de palavras e permitir análises acerca dos diferentes termos utilizados nas bases. Por fim, manualmente, foram selecionados alguns comentários que permitem uma discussão acerca de gênero nos vídeos analisados. O uso das ferramentas para a execução da pesquisa foram o Google Collaboratory⁷ como ambiente de desenvolvimento e a linguagem de programação Python⁸. A escolha dessas ferramentas se direcionou pelo fato de a linguagem Python possuir bibliotecas voltadas para a análise de dados, e o Collaboratory ser de fácil acesso e uso. Além da análise qualitativa baseada nos sentimentos, foi calculado um percentual de comentários positivos, negativos e neutros para cada um dos vídeos. A seguir, os resultados do trabalho.

4. Resultados

A pesquisa desenvolvida obteve resultados quantitativos e qualitativos com relação aos comentários de vídeos de canais de programação de homens e mulheres. Na análise quantitativa, foi inicialmente realizada uma média da quantidade de caracteres dos comentários dos vídeos. A Tabela 2 apresenta valores estatísticos sobre a quantidade de caracteres em comentários pelo gênero do autor do vídeo. São apresentados os valores de média, mediana, desvio padrão, máximo e mínimo de caracteres.

Tabela 2. Valores estatísticos sobre a quantidade de caracteres em comentários separados pelo gênero do autor do vídeo.

	média	desvio padrão	mediana	máximo	mínimo
Masculino	122,57	275,74	71	9059	1
Feminino	121,42	165,23	73	2083	1

Conforme o cenário apresentado na Tabela 2, embora a média e a mediana da quantidade de caracteres em comentários para amostras de vídeo de pessoas do gênero masculino e feminino sejam próximas, a quantidade máxima de caracteres e o desvio padrão das amostras para vídeos de pessoas do gênero masculino são substancialmente

⁷<https://colab.google>

⁸<https://www.python.org>

maiores que os das amostras para o feminino. Dessa forma, os comentários usuais na plataforma não apresentam uma tendência com relação ao tamanho com base no gênero da pessoa apresentadora, porém os comentários com maior quantidade de caracteres estão presentes nos vídeos de canais de pessoas do gênero masculino.

Um segundo aspecto analisado foi com relação a quantidade de perguntas que existiam nos comentários de cada uma das categorias. É importante ressaltar que o método de detecção de perguntas feito na base de dados completa utiliza a verificação da ocorrência do ponto de interrogação (“?”) para definir se o comentário é ou não uma pergunta. Deste modo, não foi contabilizado, portanto, comentários com mais de uma pergunta e nem mesmo comentários que não continham o caractere. Nessa análise, pode-se verificar que aproximadamente 23,09% dos comentários de vídeos de homens apresentavam pelo menos uma pergunta, enquanto os vídeos de mulheres apresentavam uma porcentagem ligeiramente menor com 20,74% dos comentários sendo classificados como pergunta. Quanto à análise de sentimentos expressos nos comentários, classificados pelo GPT, utilizando a interface ChatGPT, as diferenças foram pequenas. Os gráficos da Figura 2 trazem essa comparação.

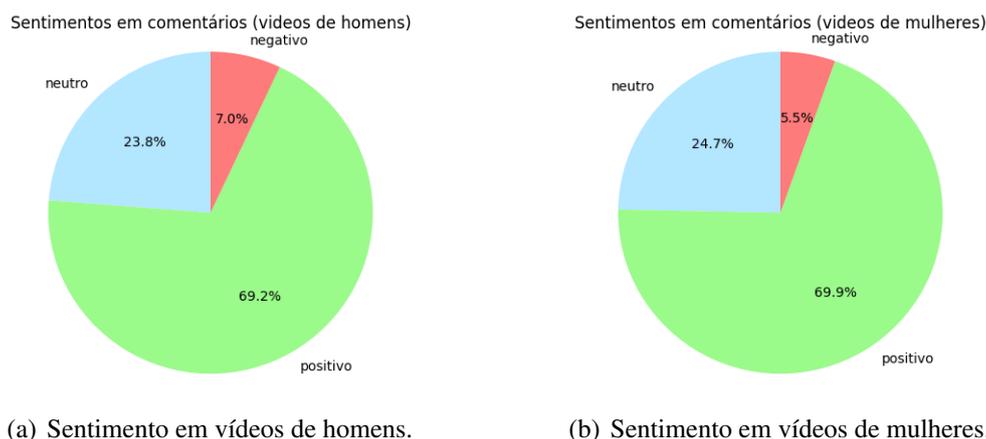


Figura 2. Análise de sentimentos dos comentários (Homens X Mulheres).

Em geral, os vídeos trazem comentários positivos, seguido por comentários neutros e uma mínima parcela de comentários negativos. Nesse caso, os vídeos produzidos por mulheres tiveram diferença mínima de comentários positivos a mais que o vídeo produzido por homens.

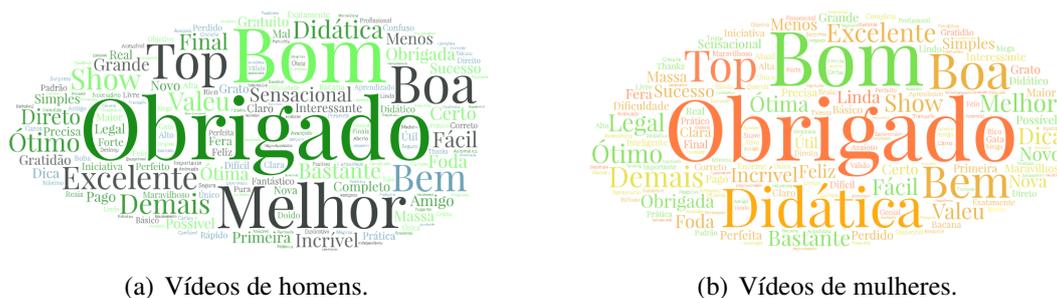


Figura 3. Palavras frequentes em vídeos de homens e mulheres.

Com relação à parte qualitativa, a primeira análise realizada foi referente aos adjetivos presentes nos comentários. A Figura 3 apresenta as nuvens de palavras construídas a partir dos comentários. A palavra de agradecimento “Obrigado” é vista com maior destaque tanto nos vídeos de mulheres quanto nos vídeos de homens. Além disso, palavras como “Bom” e “Didática” (Figura 3(b)) apareceram também em destaque em vídeos de mulheres, enquanto nos vídeos de homens as palavras “Bom” e “Melhor” (Figura 3(a)) estavam em destaque na nuvem de palavras.

Ainda com relação a uma análise qualitativa dos dados, foi feita uma busca com algumas palavras, a saber: i) linda ou lindo; ii) gata ou gato; iii) bonita ou bonito; iv) namorar; v) cabelo; vi) olhos e; vi) voz. As palavras dizem respeito a adjetivos e características físicas, além do verbo que indica relacionamento. Tal busca foi feita para verificar se os comentários, independente de transmitir um sentimento positivo ou negativo, trazia cunho sexista. Das palavras buscadas, somente um comentário com relação à aparência dos homens (“Que prof. GATO hahaha”) e pelo menos 19 comentários que envolvem alguma característica da mulher ou que indicam interesse em relacionamento. A Tabela 3 apresenta esses comentários.

Tabela 3. Comentários que envolvem características da apresentadora ou indicam interesse em relacionamento.

Comentários em vídeos de mulheres
Ela é tao linda q tenho q ver o video 2x pra prestar atenção.
Além de linda e simpática ensina muito!
Nem sei o que vc ta falandokkkk mais fico te assistindo por que te acho linda d+ tão linda apaixonado por vc kakaka ganhou 1 escrito kk
Voz meiga, inteligente e linda, tô apaixonado !????
LINDA E INTELIGENTE
Meu Deus, que mulher linda..... Infelizmente o meu modo Gado foi ativado n tem como prestar atenção na aula vei, ela é linda demais ;3
Linda e programadora...Ta usando cheater no jogo da vida UHAUHAUHAUHAU
Nossa que gata é essa
Sempre vejo varios videos seus, muito boa a didatica.. ps: vc tá mt gata nesse video mds
Cabelos lindos
Playlist maravilhosa!!! Adorei sua voz!!
Voz meiga, inteligente e linda, tô apaixonado !????
Muito bom, sua voz é muito boa de se ouvir, ela não vai me extressar
Parabéns você ensina muitos bem com uma voz bem netidez que faz com que a pessoa entenda bem a tonalidade da sua voz
Além de bonita muito inteligente
num entendi nada..... professora bonita demais, tive que deixar só o audio.....aí entendi tudo..... p.s. Brincadeira! nasceu pra ser professora!!!!
Ta bonita la profe

5. Discussão dos resultados

O artigo trouxe uma discussão acerca das questões de gênero em vídeos que trazem como temática o ensino de programação. A pesquisa é motivada, uma vez que a literatura apresenta várias questões de cunho sexista na área de computação. A análise quantitativa a partir da base de dados levantada para este trabalho apresentou resultados semelhantes para homens e mulheres. Em um determinado momento aparentou até motivador e contraditório com resultados anteriores. Ao analisar dados dos comentários, como a possibilidade do comentário ser uma pergunta ou até mesmo a quantidade média de caracteres

por comentários, não é notada uma diferença no comparativo de gênero, uma vez que os resultados são praticamente iguais.

Caminhando para a análise de sentimentos, os percentuais obtidos nas duas bases foram novamente semelhantes. Nota-se que, em geral, comentários de vídeos que envolvem programação de computadores trazem um sentimento positivo ou neutro. Tal resultado é motivador no aspecto de VBL, uma vez que são raros os comentários negativos e que pessoas que interagem nestes vídeos tem, geralmente, o intuito de tirar dúvidas ou fazer ponderações positivas acerca daquele conhecimento.

Por outro lado, na análise qualitativa, notam-se diferenças que são omitidas a partir dos resultados quantitativos. Com relação às palavras, o termo “Didática” é repetidamente colocado em vídeos de pessoas do gênero feminino. Não é possível afirmar, mas talvez cause uma certa surpresa de ter mulheres explicando sobre assuntos classificados como *hard skills* e vistos mais frequentemente ministrados por homens. De fato, as pesquisas apresentam que os canais de homens são mais populares [Amarasekara and Grant 2019], porém não há dados que comprovem serem superiores em qualidade que os canais de mulheres.

Não obstante, termos como “simpática”, “linda”, “bonita”, “gata” também figuram entre a nuvem de palavras dos vídeos de mulheres. Tais adjetivos são frequentemente empregados ao se referir a imagem feminina idealizada socialmente. Assim, é perceptível a existência de vieses implícitos nos papéis de gênero relacionados às figuras femininas e masculinas, i.e, os adjetivos expressam a forma com que as pessoas vislumbram o que é feminino e o que é masculino, de modo que para o público masculino, por exemplo, o adjetivo “lindo” não foi pontuado no aspecto de beleza física.

Considerando as manifestações de pessoas usuárias por meio dos comentários atribuídos às apresentadoras do gênero feminino (como apresentado na Tabela 3), o teor destes corroboram para evidenciar a existência de violências simbólicas e sexistas nos canais de comunicação, mesmo que aconteçam de maneira velada. Nesse sentido, apesar de adjetivos relacionados à didática terem sido evidenciados nos vídeos femininos, há comentários que sinalizam uma possível deslegitimação e desqualificação do conhecimento disseminado em vídeo, a exemplo do comentário “*Nem sei o que vc ta falandokkkk mais fico te assistindo por que te acho linda d+*”. Portanto, isso suscita discussões quanto à objetificação e vulnerabilidade de mulheres, mesmo em mídias sociais, em que adjetivos como “linda”, “gata”, “bonita” são usuais para denotar pontos positivos e/ou neutros relacionados a um vídeo. Conforme [Costa and Carvalho 2020], esta posição de vulnerabilidade no YouTube está atrelada a preconceitos e estereótipos decorrentes de processos socioculturais, econômicos e políticos que também possuem efeito na esfera da Internet.

Deste modo, ainda que o assunto de um vídeo seja algo de interesse da comunidade, discutindo, muitas das vezes, aspectos técnicos relacionados ao ensino de programação, parte das pessoas comentadoras atribuem pouca seriedade ao tema. São enfatizados aspectos como voz, cabelo e/ou outros adjetivos empregados na sociedade para representar características tidas como “femininas” no imaginário social, consequentemente, resultando em discussões mais rasas sobre o tópico abordado nos vídeos.

Os resultados também permitem trazer como discussão que assuntos que envolvem gênero nem sempre é possível de observar se há ou não diferenças somente a partir

de dados quantitativos. Tal análise deve servir como reflexão para que trabalhos futuros se atentem às possíveis violências simbólicas e/ou de mensagens de cunho sexistas e discursos de ódio que podem ser interpretadas por classificadores como mensagens positivas.

6. Considerações finais

Os resultados da pesquisa indicam que existem semelhanças e diferenças nos comentários em vídeos de ensino de computação produzidos por homens e por mulheres. Em termos de análise quantitativa, não houve diferenças significativas entre os gêneros em relação à média de caracteres nos comentários, presença de perguntas nos comentários e sentimento geral expresso nos comentários. Tanto homens quanto mulheres receberam predominantemente comentários positivos e neutros.

No entanto, a análise qualitativa revelou alguns padrões distintos. A análise da nuvem de palavras mostrou que as palavras em destaque nos vídeos produzidos por homens e mulheres se assemelham, porém as palavras que possuem menos frequência trazem questões de gênero. A análise qualitativa também revelou comentários que focavam na aparência e características físicas das apresentadoras, indicando a presença de comentários objetificadores e sexistas. Esses resultados sugerem que, embora haja semelhanças na análise de sentimento e no engajamento nos comentários de vídeos de ensino de programação em geral, também existem dinâmicas específicas de gênero. As mulheres podem enfrentar desafios adicionais relacionados à objetificação e estereótipos de gênero, que podem se manifestar na forma de comentários que focam em sua aparência no lugar de suas habilidades com o conhecimento. Os resultados destacam a importância de refletir acerca de preconceitos de gênero e promover ambientes inclusivos no campo da educação em programação.

É importante ressaltar que essa pesquisa representa uma análise específica com base em uma amostra limitada de vídeos e comentários. Os resultados devem ser interpretados no contexto do estudo e podem não ser generalizáveis para todos os vídeos de ensino de programação ou canais do YouTube. Outra limitação é que a classificação de sentimentos foi feita com a interface ChatGPT, sendo que a técnica que alimenta a interface é classificada como não determinística. Pesquisas futuras com conjuntos de dados maiores e mais diversos podem fornecer uma compreensão mais abrangente das dinâmicas entre gênero, comentários e educação em programação no YouTube.

Agradecimentos

As pessoas autoras agradecem o IF Goiano e o IFG pelo apoio recebido para a produção e apresentação deste trabalho.

Referências

- [Abu-Salih et al. 2023] Abu-Salih, B., Alhabashneh, M., Zhu, D., Awajan, A., Alshamaileh, Y., Al-Shboul, B., and Alshraideh, M. (2023). Emotion detection of social data: Apis comparative study. *Heliyon*, 9(5).
- [Alhujaili and Yafooz 2021] Alhujaili, R. F. and Yafooz, W. M. (2021). Sentiment analysis for youtube videos with user comments. In *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, pages 814–820. IEEE.

- [Amarasekara and Grant 2019] Amarasekara, I. and Grant, W. J. (2019). Exploring the youtube science communication gender gap: A sentiment analysis. *Public Understanding of Science*, 28(1):68–84.
- [Barroso and Antunes 2020] Barroso, F. and Antunes, M. (2020). Tecnologia na educação: ferramentas digitais facilitadoras da prática docente. *Pesquisa e Debate em Educação*, 5(1):124–131.
- [Bringula et al. 2020] Bringula, R. P., Victorino, J. N., De Leon, M. M., and Regina Estuar, M. (2020). Cluster and sentiment analyses of youtube textual feedback of programming language learners to enhance learning in programming. In Arai, K., Bhatia, R., and Kapoor, S., editors, *Proceedings of the Future Technologies Conference (FTC) 2019*, pages 913–924, Cham. Springer International Publishing.
- [Cardoso and de David 2016] Cardoso, É. E. C. and de David, T. (2016). A falta de profissionais de tecnologia de informação no mercado de trabalho. *Uma Nova Pedagogia para a Sociedade Futura*, pages 697–700.
- [Costa and Carvalho 2020] Costa, V. S. d. and Carvalho, C. A. (2020). Mulheres não podem falar de ciência? análise de comentários sexistas em vídeo do canal nerdologia. *Em Questão*, 26(1):42–64.
- [Diaz Jr 2023] Diaz Jr, M. O. (2023). A domain-specific evaluation of the performance of selected web-based sentiment analysis platforms. *International Journal of Software Engineering and Computer Systems*, 9(1):01–09.
- [Ebied et al. 2016] Ebied, M. M. A., Kahouf, S., and Rahman, S. A. A. (2016). Effectiveness of using youtube in enhance the learning of computer in education skills in najran university. *International Interdisciplinary Journal of Education*, 5(3):619–625.
- [El Rahman et al. 2019] El Rahman, S. A., AlOtaibi, F. A., and AlShehri, W. A. (2019). Sentiment analysis of twitter data. In *2019 international conference on computer and information sciences (ICCIS)*, pages 1–4. IEEE.
- [Ferreira et al. 2020] Ferreira, J. P. B., de Torres Maschio, P., de Santana, T. S., da Costa, N. T., and Junior, C. P. (2020). Análise de vídeos como recurso educacional em plataforma não formal de aprendizagem. In *Anais do XXXI Simpósio Brasileiro de Informática na Educação*, pages 1733–1742. SBC.
- [Garcia et al. 2022] Garcia, M. B., Juanatas, I. C., and Juanatas, R. A. (2022). Tiktok as a knowledge source for programming learners: a new form of nanolearning? In *2022 10th International Conference on Information and Education Technology (ICIET)*, pages 219–223. IEEE.
- [Hou 2019] Hou, M. (2019). Social media celebrity and the institutionalization of youtube. *Convergence*, 25(3):534–553.
- [Kadriu et al. 2020] Kadriu, A., Abazi-Bexheti, L., Abazi-Alili, H., and Ramadani, V. (2020). Investigating trends in learning programming using youtube tutorials. *International Journal of Learning and Change*, 12(2):190–208.
- [Lammerse et al. 2022] Lammerse, M., Hassan, S. Z., Sabet, S. S., Riegler, M. A., and Halvorsen, P. (2022). Human vs. gpt-3: The challenges of extracting emotions from child

- responses. In *2022 14th International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–4.
- [Lin 2022] Lin, P. (2022). Developing an intelligent tool for computer-assisted formulaic language learning from youtube videos. *ReCALL*, 34(2):185–200.
- [Louzada et al. 2014] Louzada, C. S., Gomes, W. F., Nunes, M., Salgueiro, E. M., Andrade, B. T., and Lima, P. (2014). Um mapeamento das publicações sobre o ingresso das mulheres na computação. In *CLEI 2014: Conferência Latino-americana em Informática-VI Congresso da Mulher Latino-americana na Computação*. Montevideu, page 16.
- [Mathew and Bindu 2020] Mathew, L. and Bindu, V. R. (2020). A review of natural language processing techniques for sentiment analysis using pre-trained models. In *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, pages 340–345.
- [Maziriri et al. 2020] Maziriri, E. T., Gapa, P., and Chuchu, T. (2020). Student perceptions towards the use of youtube as an educational tool for learning and tutorials. *International Journal of Instruction*, 13(2):119–138.
- [Medhat et al. 2014] Medhat, W., Hassan, A., and Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4):1093–1113.
- [Miranda and Martin 2020] Miranda, J. P. P. and Martin, J. T. (2020). Topic modeling and sentiment analysis of martial arts learning textual feedback on youtube. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(3).
- [Moura 2016] Moura, M. A. (2016). *O discurso do ódio em redes sociais*. Lura Editorial (Lura Editoração Eletrônica LTDA-ME).
- [Pal et al. 2019] Pal, S., Pramanik, P. K. D., Majumdar, T., and Choudhury, P. (2019). A semi-automatic metadata extraction model and method for video-based e-learning contents. *Education and Information Technologies*, 24(6):3243–3268.
- [Regueira et al. 2020] Regueira, U., Alonso-Ferreiro, A., and Da-Vila, S. (2020). Women on youtube: Representation and participation through the web scraping technique. *Comunicar*, 28(63):31–40.
- [Sablić et al. 2021] Sablić, M., Miroslavljević, A., and Škugor, A. (2021). Video-based learning (vbl)—past, present and future: An overview of the research published from 2008 to 2019. *Technology, Knowledge and Learning*, 26(4):1061–1077.
- [Santana et al. 2022] Santana, F. C., de Carvalho Júnior, O. F., Bernardo, R., Vicente, E. F., and Bernardo, C. H. C. (2022). Mulheres gamers: uma análise da hostilidade enfrentada pelas mulheres em jogos virtuais. *Cadernos de Gênero e Tecnologia*, 15(46):120–138.
- [Santos et al. 2021] Santos, V. L., Carvalho, T. F., and do Socorro Barreto, M. (2021). Mulheres na tecnologia da informação: Histórico e cenário atual nos cursos superiores. In *Anais do XV Women in Information Technology*.
- [SBC 2023] SBC (2023). Educação Superior em Computação: Estatísticas 2020. <https://www.sbc.org.br/documentos-da-sbc/send/133-estatisticas/1420-educacao-superior-em-computacao-estatisticas-2020>.