Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Aplicadas
Departamento de Computação e Sistemas

Análise das narrativas sobre mudanças climáticas em comentários de vídeos no YouTube no Brasil

Daniel Ângelo Rosa Morais

TRABALHO DE CONCLUSÃO DE CURSO

ORIENTAÇÃO:

Profª. Drª. Helen de Cássia Sousa da Costa Lima

COORIENTAÇÃO:

Prof. Dr - Carlos Henrique Gomes Ferreira

Setembro, 2025
João Monlevade–MG

Daniel Ângelo Rosa Morais

Análise das narrativas sobre mudanças climáticas em comentários de vídeos no YouTube no Brasil

Orientador: Profª. Drª. Helen de Cássia Sousa da Costa Lima

Coorientador: Prof. Dr - Carlos Henrique Gomes Ferreira

Monografia apresentada ao curso de Sistemas de Informação do Instituto de Ciências Exatas e Aplicadas, da Universidade Federal de Ouro Preto, como requisito parcial para aprovação na Disciplina "Trabalho de Conclusão de Curso II".

Universidade Federal de Ouro Preto

João Monlevade

Setembro de 2025

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

M827a Morais, Daniel Angelo Rosa.
Análise das narrativas sobre mudanças climáticas em comentários de vídeos no YouTube no Brasil. [manuscrito] / Daniel Angelo Rosa Morais. - 2025.
53 f.

Orientadora: Profa. Dra. Helen de Cássia Sousa da Costa Lima.
Coorientador: Prof. Dr. Carlos Henrique Gomes Ferreira.
Monografia (Bacharelado). Universidade Federal de Ouro Preto. Instituto de Ciências Exatas e Aplicadas. Graduação em Sistemas de Informação .

1. Mídia social - Análise. 2. Mineração de dados (Computação). 3. Mudanças climáticas. 4. Processamento de linguagem natural (Computação). 5. Youtube (Recurso eletrônico). I. Lima, Helen de Cássia Sousa da Costa. II. Ferreira, Carlos Henrique Gomes. III. Universidade Federal de Ouro Preto. IV. Título.

CDU 004.8:004.6

Bibliotecário(a) Responsável: Flavia Reis - CRB6-2431

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
INSTITUTO DE CIENCIAS EXATAS E APLICADAS
DEPARTAMENTO DE COMPUTACAO E SISTEMAS

FOLHA DE APROVAÇÃO

Daniel Ângelo Rosa Morais

Análise das narrativas sobre mudanças climáticas em comentários de vídeos no YouTube no Brasil

Monografia apresentada ao Curso de Sistemas de Informação da Universidade Federal
de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação

Aprovada em 05 de setembro de 2025

Membros da banca

Dra. Helen de Cássia Sousa da Costa Lima - Orientadora - (Departamento de Computação e Sistemas - DECSI da Universidade Federal de Ouro Preto - UFOP)
Dr. Carlos Henrique Gomes Ferreira - Coorientador - (Departamento de Computação e Sistemas - DECSI da Universidade Federal de Ouro Preto - UFOP)
Dr. Alexandre Magno de Sousa - (Departamento de Computação e Sistemas - DECSI da Universidade Federal de Ouro Preto - UFOP)
Me. Josemar Alves Caetano - (Departamento de Computação e Sistemas - DECSI da Universidade Federal de Ouro Preto - UFOP)

Helen de Cássia Sousa da Costa Lima, orientadora do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 16/10/2025

Documento assinado eletronicamente por **Helen de Cassia Sousa da Costa Lima**, **PROFESSOR DE MAGISTERIO SUPERIOR**, em 16/10/2025, às 20:52, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0 , informando o código verificador **0998508** e o código CRC **91A91349**.

R. Diogo de Vasconcelos, 122, - Bairro Pilar Ouro Preto/MG, CEP 35402-163
Telefone: (31)3808-0819 - www.ufop.br

Dedico esse trabalho a toda comunidade acadêmica.

Agradecimentos

Em primeiro lugar, agradeço a Deus pela força e pela oportunidade de chegar até aqui.

À minha família, pelo apoio incondicional, paciência nos momentos de ausência e palavras de incentivo que me sustentaram ao longo desta jornada.

Aos meus amigos, que estiveram presentes com apoio, compreensão e boas conversas que aliviaram o peso do caminho.

Um agradecimento especial aos meus orientadores, Helen de Cássia Sousa da Costa Lima e Carlos Henrique Gomes Ferreira, pela disponibilidade em cada etapa do trabalho e pelo vasto conhecimento compartilhado, que enriqueceram profundamente o desenvolvimento desta pesquisa.

À minha instituição e professores, que ao longo do curso compartilharam conhecimento e experiências valiosas, que moldaram minha formação acadêmica e pessoal.

E, por fim, agradeço a todos que, de forma direta ou indireta, contribuíram para que este trabalho fosse possível. Cada gesto, palavra ou ensinamento fez diferença nesta caminhada.

"Science is more than a body of knowledge; it is a way of thinking."

— Carl Sagan (1934 – 1996),
in: The Demon-Haunted World: Science as a Candle in the Dark.

Resumo

O avanço das redes sociais transformou profundamente a forma como temas complexos, como as mudanças climáticas, são debatidos publicamente. O YouTube, em particular, consolidou-se como um espaço central de circulação de narrativas, mas também de desinformação e polarização. Embora existam estudos focados em contextos anglófonos, ainda são raras as análises sistemáticas sobre o ecossistema brasileiro em língua portuguesa. Este trabalho busca preencher essa lacuna ao investigar as narrativas climáticas presentes nos comentários de vídeos publicados no YouTube no Brasil. Foram coletados e pré-processados mais de duzentos mil comentários, analisados em diferentes etapas. Inicialmente, aplicou-se o modelo *BERTopic* (com base no BERTimbau) para identificar tópicos recorrentes e compreender como as discussões se organizam tematicamente. Em seguida, os comentários foram classificados em três categorias de posicionamento — *believers*, *deniers* e *inconclusive* — por meio de um modelo de linguagem (Llama 3.1) ajustado com *LoRA* e combinado a estratégias de *self-training*. Essa abordagem elevou significativamente o desempenho do classificador em relação ao *baseline*, permitindo a rotulação confiável de toda a base. A análise dos resultados revelou um padrão de concentração temática em torno do debate sobre o aquecimento global, coexistindo com nichos mais específicos relacionados a queimadas, Amazônia e disputas político-partidárias. Os dados mostraram forte assimetria nas dinâmicas de engajamento: poucos vídeos concentraram a maior parte das visualizações, curtidas e comentários, reforçando o caráter de cauda longa do debate. A inferência de posicionamentos evidenciou a predominância de comentários *inconclusive*, seguidos pelos *believers* e, em menor proporção, pelos *deniers*. A análise temporal indicou oscilações associadas a eventos climáticos e políticos, enquanto nuvens de palavras e a distribuição condicional das respostas ajudaram a identificar padrões de alinhamento e oposição nos diálogos. Este estudo contribui ao oferecer uma das primeiras análises sistemáticas do discurso climático no YouTube brasileiro, articulando métodos de mineração de texto, aprendizado de máquina e visualização exploratória.

Palavras-chaves: Mudanças Climáticas, PLN, Análise de Mídias Sociais

Abstract

The rise of social media has profoundly transformed the way complex topics, such as climate change, are publicly debated. YouTube, in particular, has established itself as a central space for the circulation of narratives, but also for misinformation and polarization. Although studies focused on English-speaking contexts exist, systematic analyses of the Brazilian ecosystem in Portuguese remain rare. This work seeks to fill this gap by investigating the climate narratives present in the comments of videos published on YouTube in Brazil. More than two hundred thousand comments were collected and pre-processed, and analyzed at different stages. Initially, the BERTopic model (based on BERTimbau) was applied to identify recurring topics and understand how discussions are organized thematically. The comments were then classified into three positioning categories—believers, deniers, and inconclusive—using a language model (Llama 3.1) fine-tuned with LoRA and combined with self-training strategies. This approach significantly improved the classifier's performance compared to the baseline, allowing for reliable labeling of the entire database. Analysis of the results revealed a pattern of thematic concentration around the debate on global warming, coexisting with more specific niches related to fires, the Amazon, and partisan political disputes. The data showed a strong asymmetry in engagement dynamics: a few videos accounted for the majority of views, likes, and comments, reinforcing the long-tail nature of the debate. Position inference revealed the predominance of inconclusive comments, followed by believers and, to a lesser extent, deniers. Temporal analysis indicated fluctuations associated with climate and political events, while word clouds and the conditional distribution of responses helped identify patterns of alignment and opposition in the dialogues. This study contributes by offering one of the first systematic analyses of climate discourse on Brazilian YouTube, combining methods of text mining, machine learning, and exploratory visualization.

Key-words: Climate Change, NLP, Social Media Analysis.

Lista de ilustrações

Figura 1 – Pipeline do BERTopic 22
Figura 2 – *Pipeline* da metodologia utilizada 32
Figura 3 – Caracterização das métricas de popularidade e engajamento ao longo do tempo. 34
Figura 4 – Funções de distribuição cumulativa (CDFs) das métricas de engajamento no conjunto de vídeos analisado. 35
Figura 5 – Distribuição dos tópicos identificados no corpus. 38
Figura 6 – Comparação dos dois modelos finais com intervalos de confiança de 95%. 39
Figura 7 – Análise temporal dos comentários por classe. 42
Figura 8 – CDF do percentual de comentários favoráveis e contrários nos vídeos mais comentados (Top 1%) e nos demais. 43
Figura 9 – Nuvens de palavras dos cinco vídeos com maior percentual de comentários *believers* (a) e *deniers* (b). 44
Figura 10 – Distribuição percentual das classes das respostas condicionada ao comentário pai. 45

Lista de tabelas

Tabela 1 – Palavras-chave utilizadas na coleta de dados no *YouTube* 20
Tabela 2 – Sumário dos dados analisados. 21
Tabela 3 – Interpretação dos valores de Kappa 26
Tabela 4 – Distribuição final dos comentários por classe após rotulação 26
Tabela 5 – Distribuição das amostras. 31
Tabela 6 – Representações de tópicos, suas palavras-chave e descrições interpretativas. 37
Tabela 7 – Desempenho do modelo baseline por *fold*, média com IC (95%) e modelo final. 38
Tabela 8 – Desempenho do modelo com Self-Training por *fold*, média com IC (95%) e modelo final. 39
Tabela 9 – Desempenho do modelo final por classe. 40
Tabela 10 – Distribuição final dos comentários classificados por posicionamento. . . 40

Sumário

1 INTRODUÇÃO 12
1.1 Objetivos 13
1.2 Organização do trabalho 13

2 TRABALHOS RELACIONADOS 14
2.1 Análise de Narrativas sobre Mudanças Climáticas em Mídias Sociais 14
2.2 Análise da Similitude de Comentários sobre Negacionismo Climático no YouTube 15
2.3 Discurso climático no Reddit 16
2.4 Considerações Finais 17

3 METODOLOGIA 19
3.1 Coleta e pré-processamento dos dados 19
3.2 Categorização por Tópicos 21
3.2.1 Pré-processamento dos Comentários e Posts 21
3.2.2 Fluxo de Processamento para Tópicos 22
3.3 Rotulação dos dados 23
3.3.1 Rotulação dos dados 25
3.4 Modelo de Classificação dos Comentários 26
3.4.1 Avaliação e Treinamento 28
3.4.2 Self-training 30
3.4.3 Entropia, Pontuação de Confiança e Critério de Seleção 31
3.4.4 Auto-treinamento com Pseudorótulos 31
3.4.5 Síntese da Abordagem 32
3.5 Inferência dos Posicionamentos 33

4 RESULTADOS 34
4.1 Distribuição dos dados 34
4.2 Categorização por Tópicos 36
4.3 Modelo de Classificação dos Comentários 38
4.4 Análise dos posicionamentos 40

5 CONCLUSÃO 46

REFERÊNCIAS 48

1 Introdução

As plataformas de mídias sociais desempenham um papel central na circulação de informações e na formação da opinião pública contemporânea. No campo ambiental, redes como *Twitter* (X), *YouTube* e *Facebook* possibilitam a rápida disseminação de conteúdos científicos, notícias e narrativas que moldam percepções sociais sobre temas urgentes, como o aquecimento global e as mudanças climáticas (STORANI et al., 2025). Além de funcionarem como canais de popularização científica, essas plataformas também amplificam discursos de contestação e desinformação, influenciando a forma como a sociedade compreende os riscos ambientais e as responsabilidades humanas (LIMA et al., 2024; STORANI et al., 2025).

Entre essas redes, o *YouTube* se destaca como uma das principais fontes de informação ambiental e científica, especialmente no Brasil, onde o consumo de vídeos educativos e jornalísticos cresceu de forma expressiva na última década (FLORES; MEDEIROS, 2018). Pesquisas apontam que a plataforma não apenas amplia o alcance de conteúdos informativos, mas também atua como espaço de disputa discursiva, em que narrativas científicas convivem com posicionamentos políticos e teorias conspiratórias (SALLES et al., 2023). O espaço de comentários, em particular, funciona como arena de interação e engajamento, refletindo tanto apoio à ciência quanto resistência a ela, frequentemente marcada por polarização ideológica (FURTADO; DIAS, 2024).

Diversos estudos já analisaram a comunicação digital sobre mudanças climáticas, abordando desde a representação midiática em jornais e redes sociais até a circulação de desinformação climática (STORANI et al., 2025; TREEN; WILLIAMS; O'NEILL, 2020). Trabalhos recentes também aplicaram técnicas de Processamento de Linguagem Natural (PLN) e aprendizado de máquina para mapear percepções públicas, identificar padrões narrativos e medir a prevalência de negacionismo climático em ambientes digitais (SHAHBAZI; JALALI; SHAHBAZI, 2025; ROJAS et al., 2024). No *YouTube*, estudos destacam tanto o potencial da plataforma para a divulgação científica quanto os desafios impostos pela presença de ruídos, teorias conspiratórias e narrativas politizadas (SALLES et al., 2023).

Apesar dos avanços, a maioria das pesquisas concentra-se em contextos internacionais ou em eventos específicos, como conferências climáticas globais e desastres ambientais. No caso brasileiro, observa-se uma lacuna relevante: ainda são escassos os estudos que caracterizam de forma sistemática como as mudanças climáticas são narradas e disputadas no espaço público digital, particularmente em larga escala e no contexto do *YouTube*.

1.1 Objetivos

Este trabalho busca preencher essa lacuna ao analisar as narrativas sobre mudanças climáticas presentes nos comentários de vídeos do *YouTube* do Brasil. A partir da coleta e pré-processamento de milhares de comentários, empregamos técnicas de PLN e aprendizado de máquina para classificar os discursos em três categorias: *Believer*, *Inconclusive* e *Denier*. O objetivo é compreender não apenas a distribuição dessas narrativas, mas também suas dinâmicas de engajamento e possíveis relações com o contexto político-social brasileiro.

De forma mais específica, este estudo se propõe a:

- Coletar e pré-processar um conjunto de comentários do *YouTube* relacionados a vídeos sobre mudanças climáticas no Brasil;
- Aplicar modelos de PLN e aprendizado de máquina para realizar a classificação automática dos comentários em três categorias narrativas;
- Analisar a prevalência de cada narrativa, observando padrões de polarização e engajamento;
- Investigar possíveis mudanças na circulação dessas narrativas ao longo do tempo, em especial no contexto político recente.

1.2 Organização do trabalho

O restante deste documento está estruturado da seguinte forma: o Capítulo 2 apresenta os trabalhos relacionados ao nosso estudo. O Capítulo 3 descreve a metodologia utilizada, incluindo detalhes sobre a coleta, pré-processamento e modelagem. O Capítulo 4 apresenta os resultados obtidos e as análises realizadas. Por fim, o Capítulo 5 discute as principais conclusões e propõe direções para trabalhos futuros.

2 Trabalhos Relacionados

O uso de métodos computacionais para examinar mídias sociais vem se tornando cada vez mais relevante em diferentes áreas do conhecimento, incluindo política, economia, saúde, educação e meio ambiente. Pesquisadores têm explorado dados provenientes de plataformas como *Twitter* e *YouTube* para detectar padrões de comportamento e mapear o posicionamento da população diante de temas de interesse público. Comentários e interações nessas redes representam uma fonte valiosa para compreender percepções coletivas, especialmente em períodos de instabilidade política ou durante grandes eventos. Por serem manifestações espontâneas, esses conteúdos permitem analisar a dinâmica da opinião pública, identificar tendências de polarização e acompanhar variações no engajamento social em tempo quase real. Neste capítulo, apresentamos estudos que se alinham de maneira direta com a abordagem adotada nesta pesquisa.

2.1 Análise de Narrativas sobre Mudanças Climáticas em Mídias Sociais

No contexto da análise de discussões sobre mudanças climáticas antropogênicas em mídias sociais, diversos estudos têm buscado mapear e caracterizar as narrativas predominantes entre usuários online. Um exemplo significativo é apresentado por Elroy, Komendantova e Yosipof (2024), que coletaram 333.635 tweets em inglês relacionados às mudanças climáticas antropogênicas, utilizando a API de pesquisa acadêmica do X (anteriormente Twitter). O estudo aplicou técnicas de Processamento de Linguagem Natural (PLN) e aprendizado de máquina para capturar o significado semântico das mensagens, transformando os textos em representações vetoriais que permitiram o agrupamento das informações. Nesse processo, os tweets foram incorporados por meio do modelo RoBERTa, e os *word embeddings* foram convertidos em *sentence embeddings*, facilitando a análise de similaridade semântica e o posterior agrupamento em clusters.

A aplicação dessa metodologia de *clustering* possibilitou a identificação de quatro narrativas distintas presentes na discussão online: Antropogênica, Científica, Política e Conspiratória. As narrativas Antropogênica e Conspiratória concentram-se principalmente na origem das mudanças climáticas, questionando a existência do fenômeno, o papel do fator humano e outros fatos cientificamente estabelecidos. Em contraste, a narrativa Política enfatiza termos técnicos e medidas regulatórias, oferecendo uma perspectiva mais global e registrando picos de atenção em eventos relevantes, como a divulgação de relatórios do IPCC (*Intergovernmental Panel on Climate Change*). O estudo evidencia que a contestação da

existência das mudanças climáticas, a atribuição de responsabilidade humana e a discussão sobre soluções são pontos que atraem discursos controversos, frequentemente permeados por teorias da conspiração.

Além de destacar o papel persistente dessas narrativas conspiratórias nas redes sociais Elroy, Komendantova e Yosipof (2024), ressaltam a importância de estratégias de comunicação que enfrentem informações falsas, sugerindo que o monitoramento contínuo de múltiplas plataformas seja essencial para a identificação de mudanças nas narrativas ao longo do tempo. Embora o YouTube não tenha sido o foco principal da coleta de dados, a pesquisa observa que ele se constitui como uma fonte relevante de conteúdo frequentemente referenciado, especialmente em tweets que citam céticos do clima inseridos na narrativa conspiratória. Esses achados reforçam a necessidade de integrar análises multicanal para compreender de forma mais abrangente as dinâmicas de disseminação e contestação de informações sobre mudanças climáticas, contribuindo para a elaboração de políticas públicas mais informadas e estratégias de comunicação mais eficazes.

2.2 Análise da Similitude de Comentários sobre Negacionismo Climático no YouTube

No estudo Furtado e Dias (2024), foram empregadas diversas técnicas computacionais para coletar, organizar e analisar os dados textuais, proporcionando uma compreensão aprofundada das interações dos usuários em torno do negacionismo climático. Inicialmente, os comentários do vídeo foram recuperados por meio da API do YouTube, garantindo a coleta de 1.025 dos 1.028 comentários disponíveis no momento da extração. Essa abordagem automatizada permitiu não apenas a obtenção de um conjunto de dados completo, mas também a organização sistemática das informações para análise posterior.

Com o corpus de comentários estruturado, os pesquisadores utilizaram o software Iramuteq, uma ferramenta de análise de dados textuais amplamente empregada em estudos de representação social e linguística computacional. O Iramuteq possibilita diversas análises quantitativas e qualitativas, incluindo a identificação de padrões de coocorrência lexical, frequência de termos e estruturas semânticas. Em particular, foi aplicada a técnica de análise de similitude, que consiste em mapear as conexões entre palavras e expressões recorrentes nos comentários, evidenciando como determinados termos se agrupam e delineando a estrutura temática do debate.

A análise de similitude permitiu identificar seis agrupamentos de assuntos distintos dentro do corpus textual, revelando tanto tópicos diretamente relacionados ao negacionismo climático — como questionamentos sobre a responsabilidade humana e a veracidade do aquecimento global — quanto questões periféricas, incluindo interesses geopolíticos e interpretações divergentes sobre dados científicos. Essa metodologia não apenas destacou os

principais eixos de discussão, mas também forneceu uma representação visual e estruturada das relações entre conceitos, contribuindo para a compreensão de como as narrativas e posições de usuário se organizam e se reforçam no ambiente digital.

Ao combinar a coleta automatizada via API com análises semânticas por meio do Iramuteq, o estudo demonstrou como técnicas computacionais podem ser utilizadas para investigar fenômenos sociais complexos, como a propagação de desinformação ambiental em plataformas de mídia digital, oferecendo insights valiosos para estratégias de comunicação científica e políticas públicas voltadas ao combate à desinformação.

2.3 Discurso climático no Reddit

O estudo Treen et al. (2022) ofereceu uma análise detalhada da discussão climática no *Reddit*, utilizando técnicas computacionais para lidar com a vasta e complexa base de dados da plataforma. Os autores coletaram posts e comentários de *datasets* públicos do *Reddit* (BigQuery, 2019a; BigQuery, 2019b), abrangendo o período de 1º de abril a 30 de junho de 2017, e filtrando os conteúdos por palavras-chave relevantes, como '*climate change*' e '*global warming*'. Essa abordagem permitiu a identificação sistemática de discussões relacionadas ao clima em diversos subreddits, garantindo uma amostra representativa das interações dos usuários sobre o tema.

Para compreender os tópicos predominantes nas discussões, o estudo empregou o modelo *Latent Dirichlet Allocation* (LDA), uma técnica estatística que identifica tópicos como conjuntos de palavras que frequentemente co-ocorrem em documentos. Cada *post* ou comentário foi tratado como uma mistura de tópicos, e cada tópico como uma distribuição probabilística sobre palavras, permitindo capturar a diversidade temática do discurso online. Posteriormente, os 25 tópicos estimados foram interpretados em um processo de duas etapas, correlacionando as distribuições com temas substantivamente significativos. Para aprimorar a compreensão e validar a consistência dos tópicos, calculou-se a distância *Jensen-Shannon* entre eles, visualizando-os em um espaço bidimensional e agrupando-os em seis *clusters* por meio de uma avaliação qualitativa complementada pelo algoritmo *k-means clustering*. Essa abordagem permitiu identificar não apenas tópicos individuais, mas também padrões de sobreposição e proximidade temática entre os debates do *Reddit*.

Além da análise de tópicos, os pesquisadores construíram redes de respostas — em que dois usuários eram conectados se um comentava em um *post* ou comentário do outro — para investigar a estrutura de interações e a formação de comunidades. O *dataset* resultante foi analisado na ferramenta de visualização de grafos *Gephi*, aplicando-se o *layout ForceAtlas2* e utilizando o algoritmo de maximização de modularidade para detectar subcomunidades. Essa abordagem permitiu observar se os debates eram concentrados dentro de subreddits específicos ou se os usuários transitavam entre comunidades, revelando

dinâmicas de polarização e deliberação.

O estudo também explorou as fontes de informação mencionadas pelos usuários, analisando como a referência a links e materiais externos variava entre comunidades e tópicos, oferecendo insights sobre a circulação e credibilidade das informações dentro da plataforma. Complementando as análises qualitativas e estruturais, técnicas estatísticas foram empregadas para avaliar o tamanho e a significância dos padrões observados, fornecendo suporte quantitativo às conclusões sobre polarização e engajamento.

Em síntese, Treen et al. (2022) demonstram como a combinação de processamento de linguagem natural, modelagem de tópicos e análise de redes sociais permite uma compreensão aprofundada das complexas interações em plataformas digitais, revelando tanto a diversidade temática quanto a formação de comunidades no discurso climático do Reddit. Essa metodologia ilustra a aplicabilidade de técnicas computacionais avançadas para estudos de opinião pública e comportamento online, oferecendo subsídios valiosos para pesquisas sobre comunicação científica e desinformação.

2.4 Considerações Finais

A revisão da literatura evidencia que a análise de mídias sociais tem se consolidado como uma abordagem relevante em múltiplos campos do conhecimento. No estudo das mudanças climáticas, a maior parte das pesquisas tem se concentrado na aplicação de técnicas de modelagem de tópicos, capazes de identificar tendências discursivas amplas e mapear o conteúdo geral das interações online. Embora esses trabalhos ofereçam contribuições significativas, eles frequentemente simplificam a diversidade de posicionamentos presentes nas discussões e não capturam as nuances da manifestação pública em contextos específicos, como o brasileiro, marcado por particularidades culturais, políticas e socioeconômicas.

Entre as limitações observadas na literatura, destaca-se a ausência de modelos direcionados à classificação de narrativas, com muitos estudos se restringindo a análises descritivas que identificam temas gerais sem diferenciar os tipos de posicionamento. Ademais, diversos trabalhos concentram-se em contextos internacionais ou em recortes temporais curtos, dificultando a compreensão das dinâmicas discursivas em médio e longo prazo. Essa lacuna torna-se ainda mais relevante quando se considera a polarização política e a circulação intensa de desinformação sobre mudanças climáticas no Brasil, fatores que moldam de maneira singular o debate público no espaço digital.

O presente trabalho se diferencia ao propor um modelo de classificação capaz de organizar os comentários em três categorias bem definidas: *believer*, para comentários que reconhecem a existência e o impacto das mudanças climáticas; *denier*, para comentários que negam ou relativizam o fenômeno; e *inconclusive*, para comentários ambíguos ou sem posicionamento claro. Essa segmentação não apenas permite ir além da simples identificação

de tópicos, mas também oferece uma análise direta sobre os níveis de aceitação, rejeição ou dúvida em relação às mudanças climáticas, revelando padrões de engajamento e polarização que estudos anteriores não capturaram.

Outro diferencial crucial do trabalho é a ênfase no contexto brasileiro, ainda pouco explorado em pesquisas de análise de discurso climático em mídias sociais. Considerando a importância geopolítica do Brasil no debate ambiental global, tanto pelo papel da Amazônia quanto pelas disputas políticas em torno das políticas climáticas, compreender como os usuários brasileiros se posicionam digitalmente permite capturar tensões, contradições e estratégias de argumentação típicas desse contexto.

Adicionalmente, ao utilizar um conjunto de dados amplo proveniente do *YouTube* e aplicar técnicas avançadas de processamento de linguagem natural e aprendizado de máquina, esta pesquisa supera limitações metodológicas observadas na literatura existente. A abordagem adotada possibilita acompanhar como os discursos de crença e negação se articulam, circulam e competem por atenção ao longo do tempo, fornecendo não apenas contribuições metodológicas para o estudo de narrativas em mídias sociais, mas também subsídios para a reflexão crítica sobre o papel dessas plataformas na construção das percepções coletivas sobre mudanças climáticas no Brasil.

Em síntese, este trabalho combina análise quantitativa robusta com sensibilidade contextual, oferecendo uma perspectiva inédita sobre a dinâmica do negacionismo e da aceitação científica em comentários online, reforçando a importância de integrar métodos computacionais à investigação social em contextos complexos e polarizados.

3 Metodologia

Este capítulo apresenta a metodologia adotada nesta pesquisa. Em um primeiro momento, realizamos a coleta de comentários publicados em vídeos do *YouTube* sobre mudanças climáticas, priorizando conteúdos relacionados ao contexto brasileiro. Em seguida, aplicamos técnicas de pré-processamento para assegurar a qualidade dos dados, removendo ruídos e padronizando o material textual para análise. Na etapa seguinte, desenvolvemos uma modelagem de tópicos com o objetivo de explorar as principais narrativas presentes nos comentários e orientar a construção das categorias analíticas. A partir desse diagnóstico inicial, avançamos para a implementação de um modelo supervisionado de classificação, estruturado em três classes — *believer*, *denier* e *inconclusive*. Para otimizar o processo de rotulação e ampliar a robustez do modelo, incorporamos estratégias de *self-training*, que permitiram refinar iterativamente o conjunto de treinamento. Após a validação, o modelo foi aplicado ao restante da base de dados, viabilizando a classificação automática dos comentários e a identificação de padrões e tendências narrativas.

As seções a seguir detalham cada uma dessas etapas em maior profundidade.

3.1 Coleta e pré-processamento dos dados

Para a coleta de dados na plataforma *YouTube*, utilizou-se a *Application Programming Interface* (API) de dados do *YouTube*[1], que permite extrair informações de canais, vídeos, comentários e seus respectivos metadados. Essa ferramenta foi fundamental para estruturar a base de dados inicial, garantindo acesso direto e automatizado ao conteúdo público disponível na plataforma.

O processo de coleta partiu da definição de um conjunto de **palavras-chave**, selecionadas a partir de termos amplamente utilizados em debates sobre mudanças climáticas no Brasil. A escolha foi guiada por três critérios principais: (i) relevância no contexto científico e social do tema, (ii) frequência de uso em mídias tradicionais e digitais, e (iii) presença recorrente em estudos anteriores relacionados à comunicação sobre mudanças climáticas. Dessa forma, buscou-se assegurar que os vídeos coletados representassem uma amostra diversificada de narrativas.

A Tabela 1 apresenta as palavras-chave utilizadas na coleta de dados.

[1] <https://developers.google.com/youtube/v3>

Tabela 1 – Palavras-chave utilizadas na coleta de dados no *YouTube*

Categoria	**Palavras-chave**
Mudanças Climáticas	mudanças climáticas, aquecimento global, efeito estufa
Eventos Climáticos	seca, enchente, desmatamento, queimadas
Narrativas Políticas	ambientalismo, sustentabilidade, meio ambiente
Negacionismo	clima é farsa, não existe aquecimento global

Fonte: Produzido pelo autor

A coleta de dados compreendeu o período de janeiro de 2014 a dezembro de 2024, totalizando uma década de observações. Para assegurar relevância e representatividade, foram selecionados apenas vídeos que possuíam mais de 30 comentários, e, dentro desses vídeos, apenas comentários com mais de 15 caracteres foram considerados. Esse recorte temporal e quantitativo foi definido com o objetivo de capturar não apenas debates recentes, mas também a evolução das narrativas e do tema ao longo do tempo. Dessa forma, foi possível considerar transformações políticas, sociais e ambientais ocorridas no Brasil nesse intervalo, refletindo como tais fatores influenciaram o engajamento dos usuários em discussões sobre mudanças climáticas na plataforma.

Como a API do YouTube disponibiliza vídeos em múltiplos idiomas, tornou-se necessário aplicar um processo de filtragem para selecionar apenas os comentários em português. Para essa finalidade, utilizou-se a biblioteca *LangDetect*[2], que emprega algoritmos de aprendizado de máquina e é reconhecida por sua eficácia na detecção de idiomas em textos curtos (ZOLA; RAGNO; CORTEZ, 2020). O uso do *LangDetector* foi essencial para garantir que a análise se concentrasse exclusivamente na perspectiva do público brasileiro, evitando interferências de comentários em outros idiomas que pudessem distorcer os resultados.

Com a base de dados já filtrada por idioma, foi realizada uma etapa adicional de refinamento para garantir que apenas vídeos e comentários diretamente relacionados às mudanças climáticas fossem mantidos na análise. Foram excluídos conteúdos que, apesar de conterem palavras-chave do estudo, não apresentavam pertinência ao tema central. Exemplos notáveis incluem vídeos do jogo FIFA, como *FIFA 17 - MUDANÇA CLIMÁTICA, NOVO MODO CARREIRA E MAIS!*, em que o termo "mudança climática" se referia a alterações climáticas dentro do jogo, e videoclipes musicais, majoritariamente de funk, como *Aquecimento Das Danadas - O Mandrake Feat DJ Xaropinho*, que utilizavam o termo "aquecimento" em outro contexto. Essa filtragem foi conduzida por meio de análise exploratória detalhada, considerando que certos padrões de ruído exigiam atenção especial. Dessa forma, conteúdos que poderiam gerar interpretações equivocadas ou não refletiam a perspectiva do público brasileiro sobre mudanças climáticas foram removidos. O processo assegurou que a base final de dados incluísse apenas vídeos e comentários relevantes ao

[2] <https://pypi.org/project/langdetect>

escopo do estudo, evitando distorções que poderiam comprometer a modelagem de tópicos e a classificação automática subsequente.

Após a coleta e o pré-processamento, foram registrados diversos metadados associados aos comentários, incluindo número de visualizações e curtidas, que funcionam como indicadores de engajamento e popularidade. O conjunto final de dados passou a abranger 478 canais, 1.020 vídeos, 247.514 comentários e 137.585 usuários distintos, conforme apresentado na Tabela 2.

Tabela 2 – Sumário dos dados analisados.

Métrica	**Valor**
Canais	478
Vídeos	1.020
Comentários	247.514
Usuários distintos	137.585

Fonte: Produzido pelo autor

3.2 Categorização por Tópicos

A categorização por tópicos permitiu explorar a diversidade de narrativas presentes nos comentários, fornecendo um panorama inicial sobre os temas mais debatidos e auxiliando na construção do modelo de classificação supervisionada subsequente.

3.2.1 Pré-processamento dos Comentários e Posts

Antes da criação dos tópicos, realizou-se um pré-processamento adicional nos comentários para assegurar a qualidade e a uniformidade dos dados. Textos coletados diretamente da plataforma frequentemente apresentam inconsistências, como diferenças de formatação, termos irrelevantes e ruídos que podem prejudicar a identificação de padrões relevantes. A aplicação de técnicas de pré-processamento permitiu normalizar o conteúdo textual, tornando mais eficiente a extração e a organização dos tópicos durante as etapas subsequentes de análise.

No pré-processamento dos comentários, foram aplicadas diversas etapas para padronizar e limpar o texto: (1) conversão de todo o conteúdo para letras minúsculas; (2) remoção de URLs, menções a usuários (por exemplo, @Nome-do-Usuario) e caracteres especiais, incluindo quebras de linha, tabs, espaços extras, emojis e símbolos; (3) eliminação de números; e (4) filtragem de *stopwords*, ou seja, palavras comuns que não agregam significado relevante ao texto, como “de” e “que”. Esses procedimentos garantem maior consistência na base de dados, facilitando a identificação de padrões semânticos e a modelagem de tópicos.

3.2.2 Fluxo de Processamento para Tópicos

Para realizar a análise de tópicos, adotou-se o modelo *BERTopic* (GROOTENDORST, 2022a), que se baseia em *Sentence Transformers* e combina técnicas de redução de dimensionalidade com métodos de agrupamento não supervisionado, conforme ilustrado na Figura 1. A configuração do *BERTopic* envolveu a definição de diversos parâmetros que impactam diretamente a qualidade e o nível de detalhamento dos tópicos extraídos.

Dentre esses parâmetros, o *min_topic_size* foi definido como 10, de forma a considerar apenas tópicos compostos por pelo menos 10 documentos, evitando a criação de grupos pequenos e pouco significativos. Essa configuração foi associada a outros hiperparâmetros, como a utilização do *HDBSCAN* para agrupamento e o *UMAP* para redução de dimensionalidade, que serão detalhados a seguir.

Antes de aplicar essas técnicas, contudo, foi necessário converter os textos em representações vetoriais, garantindo que os dados pudessem ser processados adequadamente pelo *BERTopic*.

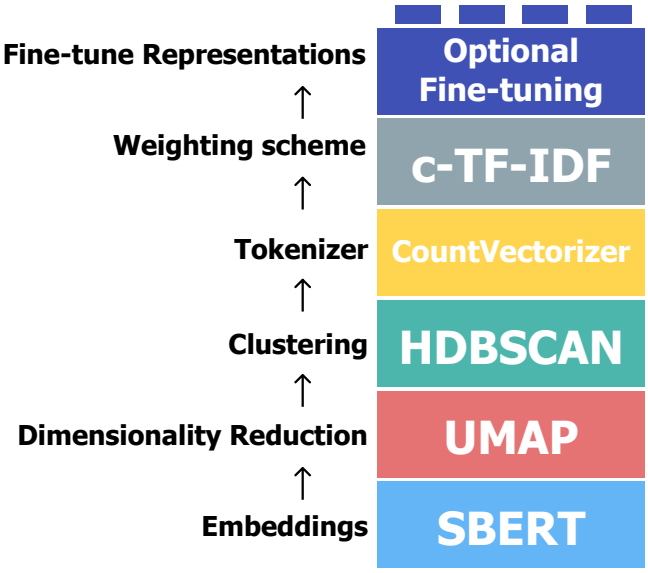

Figura 1 – Pipeline do BERTopic
Fonte: Grootendorst (2022b)

Para este processo, os textos pré-processados foram transformados em *embeddings* utilizando o modelo *BERTimbau Large*[3] , uma versão do BERT adaptada ao português brasileiro (SOUZA; NOGUEIRA; LOTUFO, 2020). Este modelo emprega uma arquitetura de rede neural avançada, treinada em extensos conjuntos de textos na língua portuguesa. O *BERTimbau Large* projeta cada sentença em um espaço vetorial de alta dimensão (1024 dimensões), de modo que comentários com significados semelhantes fiquem representados por vetores próximos entre si.

[3] <https://huggingface.co/neuralmind/bert-large-portuguese-cased>

A elevada dimensionalidade das representações vetoriais apresenta desafios ao agrupamento, tornando-o potencialmente mais custoso computacionalmente e menos eficiente (TANURE et al., 2025). Para contornar esse problema, o *BERTopic* emprega o algoritmo ***UMAP*** (*Uniform Manifold Approximation and Projection for Dimension Reduction*)[4] , que reduz a dimensionalidade do espaço vetorial (MCINNES; HEALY; MELVILLE, 2018), preservando tanto características globais quanto locais dos dados, além da integridade semântica capturada pelos *embeddings*. Neste estudo, o parâmetro `n_neighbors` do *UMAP* foi definido como 5, valor escolhido para enfatizar a preservação de relações locais entre vetores, priorizando agrupamentos mais detalhados em pequenos conjuntos de comentários semanticamente semelhantes.

Em seguida, o algoritmo ***HDBSCAN*** (*Hierarchical Density-Based Spatial Clustering of Applications with Noise*) é aplicado para agrupar os vetores de comentários em conjuntos semanticamente semelhantes, correspondendo a tópicos de discussão (CAMPELLO; MOULAVI; SANDER, 2013). O parâmetro `min_cluster_size` foi definido como 10, garantindo que cada cluster identificado contenha um número mínimo de comentários, evitando a criação de clusters muito pequenos que poderiam representar ruído ou padrões estatisticamente irrelevantes.

Para descrever esses tópicos, o *BERTopic* utiliza uma versão adaptada da técnica TF-IDF, chamada c-TF-IDF (*class-based* TF-IDF) (GROOTENDORST, 2022a), que enfatiza termos frequentes dentro de um cluster, mas raros nos demais, proporcionando uma representação precisa dos tópicos identificados.

Com essa configuração, executamos o nosso *pipeline*, o que resultou na identificação de mais de 2 mil tópicos. Entretanto, grande parte desses tópicos apresentava elevada similaridade entre si, tornando a interpretação dos resultados mais complexa e demandando estratégias adicionais para consolidar e organizar os agrupamentos.

Para reduzir a fragmentação observada nos tópicos identificados, aplicou-se a estratégia de fusão de tópicos por meio do método `reduce_topics` do *BERTopic* (GROOTENDORST, 2022a). Essa técnica combina tópicos que apresentam similaridade semântica, diminuindo a quantidade total de *clusters* e facilitando a interpretação dos resultados. O número final de tópicos foi definido automaticamente pelo modelo, assegurando que apenas agrupamentos semanticamente distintos fossem mantidos. Com isso, foi possível preservar a coerência dos tópicos sem comprometer excessivamente a granularidade da análise.

3.3 Rotulação dos dados

A construção do modelo de detecção de posicionamento seguiu uma abordagem em etapas, organizada para equilibrar qualidade e eficiência na rotulagem dos comentários.

[4] <https://umap-learn.readthedocs.io/>

Inicialmente, realizamos a rotulagem manual a partir de uma amostra proporcional aos tópicos gerados pelo *BERTopic*, garantindo que diferentes categorias estivessem representadas no conjunto de treinamento. Essa estratégia inicial de amostragem proporcional por tópicos foi adotada como uma forma de viabilizar maior diversidade semântica no conjunto rotulado. Ao invés de selecionar comentários de forma aleatória em todo o corpus, utilizamos a distribuição dos agrupamentos temáticos identificados pelo *BERTopic* como critério para a escolha dos exemplos a serem rotulados manualmente.

Em seguida, treinamos um modelo *baseline*, cuja performance apresentou limitações notáveis, sobretudo na capacidade de lidar de forma equilibrada com classes desbalanceadas e captar nuances discursivas mais sutis. Diante desse cenário, adotamos uma estratégia de aprimoramento baseada em *Self-Training*, que busca maximizar o aproveitamento do conjunto de dados anotados e, ao mesmo tempo, reduzir o custo associado à rotulagem manual em larga escala.

Nessa abordagem, o classificador inicial foi utilizado não apenas para prever rótulos, mas também para medir a incerteza de suas próprias previsões. Para isso, calculamos a entropia das distribuições de probabilidade de saída do modelo: amostras com entropia elevada indicam maior incerteza e, menores valores de entropia significam que o modelo possui maior certeza para aquela classe. Dessa forma, previsões de alta confiança puderam ser aproveitadas como *pseudo-rótulos*, ampliando de forma controlada o conjunto de treinamento sem depender exclusivamente de intervenção humana.

Esse ciclo iterativo de seleção ativa e auto-treinamento possibilitou priorizar os exemplos mais relevantes para melhorar a fronteira de decisão do modelo, enquanto expandia gradualmente a base de dados anotada de forma semi-supervisionada. O resultado foi uma melhora consistente nos indicadores de desempenho, aliada a uma economia significativa de esforço de rotulação. Assim, o processo não apenas superou as limitações do *baseline*, como também se mostrou viável e eficaz para a análise de grandes volumes de comentários em contextos de alta variabilidade linguística, como é o caso do YouTube brasileiro.

Por fim, o modelo foi re-treinado combinando os rótulos manuais e os pseudo-rótulos, de modo a expandir a base de treinamento para a totalidade dos comentários. Desse conjunto, aproximadamente 66% corresponde a anotações manuais, enquanto os 33% restantes foram obtidos a partir da estratégia de pseudo-rótulos. Essa composição permitiu aumentar a representatividade sem comprometer a qualidade, garantindo maior robustez na classificação automática do posicionamento em relação às mudanças climáticas.

3.3.1 Rotulação dos dados

Para o desenvolvimento do modelo, foi realizada uma rotulagem manual parcial dos comentários, classificando-os em três categorias: *Believers*, *Deniers* e *Inconclusive*. Comentários classificados como *Believers* indicam que o autor reconhece a realidade das mudanças climáticas e manifesta concordância com o consenso científico. Os comentários *Deniers* expressam ceticismo ou negam a ocorrência e os impactos das mudanças climáticas. Por fim, a categoria *Inconclusive* inclui comentários que não apresentam um posicionamento claro ou contêm ambiguidade quanto à opinião do autor, sendo insuficientes para atribuir uma das categorias anteriores.

Como a rotulagem manual de toda a base seria inviável, optou-se por uma amostragem proporcional aos tópicos gerados pelo *BERTopic*, garantindo que todas as categorias identificadas estivessem representadas no conjunto rotulado. Para manter a diversidade dos comentários e evitar redundâncias, aplicou-se um filtro de remoção de duplicatas, eliminando textos idênticos. Após esse processo, a base final consistiu em 4077 comentários únicos, que foram utilizados para o treinamento e validação do modelo de detecção de posicionamento sobre mudanças climáticas.

A rotulação dos comentários foi realizada por dois avaliadores independentes, com idades entre 22 e 23 anos, com o objetivo de capturar diferentes interpretações e reduzir vieses individuais. Ambos possuem escolaridade em nível superior em andamento. Não há qualquer vínculo familiar ou conjugal entre os avaliadores.

Para avaliar a consistência das rotulações, utilizou-se o coeficiente *Kappa de Fleiss* (FLEISS; LEVIN; PAIK, 2013), uma métrica amplamente utilizada para mensurar o grau de concordância entre múltiplos avaliadores categóricos. Essa métrica corrige a concordância observada pelo acaso e fornece um valor que varia entre -1 e 1, permitindo quantificar objetivamente a confiabilidade das rotulações.

Na equação, P_o representa a proporção de concordância observada entre os avaliadores, calculada a partir do número de classificações em que houve concordância total. Já P_e indica a proporção de concordância esperada pelo acaso, considerando a distribuição das categorias atribuídas. O numerador expressa a diferença entre o acordo real e o esperado, enquanto o denominador normaliza essa diferença, garantindo que o valor de κ varie entre -1 e 1. A equação do coeficiente Kappa de Fleiss é dada por:

$$\kappa = \frac{P_o - P_e}{1 - P_e} \tag{3.1}$$

A Tabela 3 apresenta a interpretação dos valores de Kappa, conforme sugerido na literatura (DIAS et al., 2024).

Tabela 3 – Interpretação dos valores de Kappa

Valor de κ	Interpretação
< 0	Concordância menor que o acaso
0.01 – 0.20	Concordância fraca
0.21 – 0.40	Concordância razoável
0.41 – 0.60	Concordância moderada
0.61 – 0.80	Concordância substancial
0.81 – 1.00	Concordância quase perfeita

Fonte: Produzido pelo autor

O estudo alcançou um coeficiente Kappa de 0.88, o que demonstra um nível de concordância considerado quase perfeito (LANDIS; KOCH, 1977). Esse elevado grau de acordo entre os avaliadores confirma a robustez do processo de rotulação, garantindo que os comentários selecionados para o treinamento do modelo sejam representativos e consistentes em relação às opiniões manifestadas pelos autores.

Após a rotulação individual pelos dois avaliadores principais, foram identificados 284 comentários em que houve empate entre eles. Para resolver esses casos, os comentários foram avaliados por uma terceira avaliadora, com aproximadamente 40 anos e doutorado, que atuou como desempate. Mesmo após essa intervenção, 69 comentários permaneceram sem definição clara e, por isso, foram descartados do conjunto final.

Dessa forma, todos os demais comentários receberam rótulos confiáveis para as classes *Believers*, *Deniers* e *Inconclusive*, garantindo a consistência da base utilizada para o treinamento e validação do modelo. A distribuição final por categoria está apresentada na Tabela 4.

Tabela 4 – Distribuição final dos comentários por classe após rotulação

Classe	Número de comentários
Believers	1097
Deniers	670
Inconclusive	2241
Total de comentários utilizados	4008

Fonte: Produzido pelo autor

3.4 Modelo de Classificação dos Comentários

O modelo desenvolvido para a classificação de comentários sobre mudanças climáticas neste trabalho é baseado no Llama 3.1 8B[5], um *Large Language Model* (LLM) de última geração (CHEN, 2024; WEERAWARDHENA et al., 2025). LLMs são redes neurais profundas baseadas na arquitetura *Transformer*, composta por múltiplos blocos de

[5] <https://huggingface.co/meta-llama/Llama-3.1-8B>

autoatenção, normalização e projeções lineares, que possibilitam capturar dependências de longo alcance em textos. O pré-treinamento foi realizado pela Meta em extensos corpora de textos não rotulados, permitindo que o modelo adquirisse conhecimento linguístico amplo, útil para tarefas gerais de PLN (DAVILA; COLAN; HASEGAWA, 2025). A designação "8B" refere-se aos 8 bilhões de parâmetros treináveis no pré-treinamento, conferindo ao modelo a capacidade de lidar com nuances discursivas sutis, especialmente relevantes para diferenciar narrativas em debates socioambientais. Uma das principais vantagens de modelos baseados em *Transformers*, como o Llama, é a capacidade de gerar representações contextuais dinâmicas. Diferentemente de *embeddings* estáticos, que atribuem o mesmo vetor fixo para cada palavra, o Llama ajusta as representações conforme o contexto em que os termos aparecem. Essa consciência contextual enriquece a compreensão linguística e se mostra particularmente eficaz para lidar com os desafios de textos oriundos de mídias sociais, como uso de ironia, coloquialismos e construções gramaticais não canônicas. O pré-treinamento em grandes volumes de dados textuais fornece ao modelo um repertório robusto de padrões linguísticos, que pode ser adaptado para tarefas específicas por meio de ajuste fino.

Para tornar o ajuste fino do modelo mais eficiente e viável em termos de recursos computacionais, foi empregada a técnica de *Low-Rank Adaptation* (LoRA) (HU et al., 2021; RENDUCHINTALA; KONUK; KUCHAIEV, 2023). Essa técnica atua da seguinte forma:

1. **Congelamento dos pesos originais**: os parâmetros do modelo base são preservados, mantendo o conhecimento adquirido no pré-treinamento.
2. **Inserção de matrizes de baixa patente**: pequenas camadas adicionais são introduzidas em pontos estratégicos, como as projeções de atenção (q_proj, k_proj, v_proj). Essas matrizes têm posto reduzido ($r = 16$ no nosso caso), o que significa que apenas uma fração dos parâmetros é efetivamente ajustada.
3. **Treinamento eficiente**: apenas as matrizes de baixa patente são atualizadas, o que reduz em até duas ordens de magnitude o número de parâmetros treináveis, resultando em menor consumo de memória e maior velocidade de convergência.

Esse processo é particularmente adequado em cenários como o deste trabalho, em que o conjunto de dados é específico e não suficientemente grande para justificar um *full fine-tuning*.

Para viabilizar o treinamento em GPUs com memória limitada, aplicou-se a quantização de 4 bits utilizando a biblioteca BitsAndBytes[6], adotando o formato nf4 e computação em bfloat16 (LANG; GUO; HUANG, 2024; YOUNG, 2024; PATIL; KHOT; GUDIVADA,

[6] <https://github.com/bitsandbytes-foundation/bitsandbytes>

2025). A quantização reduz a precisão dos pesos do modelo de 16 ou 32 bits para 4 bits, enquanto a computação em bfloat16 mantém a estabilidade numérica. Essa combinação permite executar modelos maiores de maneira eficiente, preservando a capacidade de generalização.

Além disso, foram empregadas técnicas adicionais para reduzir custo computacional sem perda significativa de performance:

- **Gradient Checkpointing**: armazena apenas parte dos gradientes intermediários e recalcula os demais durante o *backpropagation*, diminuindo o consumo de memória.
- **Mixed Precision Training**: combina cálculos em 16 e 32 bits, acelerando o treino sem comprometer a estabilidade.
- **Gradient Accumulation**: simula o uso de *batch sizes* maiores, acumulando gradientes antes da atualização dos pesos.
- **Scheduler com warmup e decaimento cosseno**: garante que a taxa de aprendizado aumente suavemente no início do treinamento, reduzindo riscos de instabilidade, e decaia de forma gradual para favorecer convergência.

3.4.1 Avaliação e Treinamento

A avaliação do modelo foi conduzida utilizando validação cruzada cíclica de 5 *folds*, uma estratégia que permite estimar o desempenho do classificador em diferentes subconjuntos de dados e reduzir vieses na medição da performance (KRSTAJIć et al., 2014). Em cada rodada r da validação, um fold é utilizado como conjunto de validação, o fold seguinte $(r + 1)$ mod 5 como conjunto de teste, e os três folds restantes como treinamento. Dessa forma, cada subconjunto é avaliado em diferentes papéis ao longo das rodadas, garantindo robustez e generalização do modelo.

O processo de treinamento incluiu técnicas de regularização e calibração. O *early stopping* foi utilizado para prevenir sobreajuste, interrompendo o treinamento quando a *validation loss* não melhorava por três épocas consecutivas (HUSSEIN; SHAREEF, 2024; PRECHELT, 1998). Para selecionar o melhor modelo de cada *fold*, utilizou-se o Macro F1 Score, métrica que calcula o F1 Score individualmente para cada classe e, em seguida, realiza a média não ponderada, equilibrando a influência de classes minoritárias e majoritárias (FARHADPOUR; WARNER; MAXWELL, 2024; LIPTON; ELKAN; BALAKRISHNAN, 2014; OPITZ, 2024).

A função de perda utilizada foi a *cross-entropy*, modificada com duas estratégias para lidar com desbalanceamento de classes: (i) aplicação de pesos inversos à frequência de cada classe, de forma a valorizar as minoritárias no cálculo do gradiente (CUI et al.,

2019; QIU; SONG, 2018); e (ii) uso de *label smoothing*, técnica que suaviza os rótulos verdadeiros a valores como $(1 - \epsilon)$ para a classe correta e $\epsilon/(K - 1)$ para as demais, reduzindo *overconfidence* e melhorando a generalização (PEREYRA et al., 2017).

As métricas de avaliação utilizadas para medir o desempenho do modelo foram definidas da seguinte forma:

Precisão (*Precision*)

A precisão mede a proporção de exemplos corretamente classificados como positivos em relação ao total de exemplos que o modelo previu como positivos. Para uma classe i, é calculada por:

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i} \tag{3.2}$$

onde TP_i representa o número de verdadeiros positivos e FP_i o número de falsos positivos da classe i. Essa métrica indica quão confiável é o modelo quando ele prediz uma determinada classe.

Sensibilidade (*Recall*)

O *recall*, também chamado de sensibilidade ou taxa de verdadeiros positivos, avalia a capacidade do modelo de identificar corretamente todas as instâncias positivas de uma classe i:

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i} \tag{3.3}$$

onde FN_i representa os falsos negativos da classe i. O recall mostra a eficiência do modelo em não deixar exemplos positivos passarem despercebidos.

F1-Score

O F1-score combina precisão e *recall* em uma média harmônica, equilibrando ambos os aspectos e penalizando desvios extremos de uma das métricas:

$$F1_i = 2 \cdot \frac{\text{Precision}_i \cdot \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i} \tag{3.4}$$

Essa métrica é especialmente útil em problemas com classes desbalanceadas, pois garante que o modelo seja avaliado tanto pelo acerto de suas previsões quanto pela capacidade de encontrar todos os exemplos positivos.

Macro F1-Score

Para problemas multiclasse, calcula-se o F1-score para cada classe individualmente e, em seguida, obtém-se a média não ponderada das classes:

$$\text{Macro F1} = \frac{1}{K} \sum_{i=1}^{K} F1_i \tag{3.5}$$

onde K é o número total de classes. Essa abordagem assegura que todas as classes, independentemente de sua frequência, tenham peso igual na avaliação final do modelo.

Após a validação cruzada, o modelo final foi treinado em uma divisão estratificada de 80% dos dados para treino e 20% para validação. Para calibrar as probabilidades de saída, aplicou-se *temperature scaling*, ajustando os logits z_i do modelo pela temperatura T:

$$\hat{p}_i = \frac{\exp(z_i/T)}{\sum_j \exp(z_j/T)} \tag{3.6}$$

Essa técnica suaviza a distribuição de probabilidades, tornando-as mais confiáveis e representativas da incerteza real do modelo (GUO et al., 2017; KULL et al., 2019).

Diversos aprimoramentos técnicos adicionais foram incorporados, incluindo o uso do otimizador AdamW, agendamento da taxa de aprendizado em cosseno com fase de warmup, *gradient accumulation*, precisão mista (*mixed precision*) e *gradient checkpointing* para permitir treinamento de modelos grandes de forma eficiente (LOSHCHILOV; HUTTER, 2018; LEWKOWYCZ, 2021). A reprodutibilidade foi assegurada por meio da fixação de sementes aleatórias (*random seeds*) e registro da configuração de hardware utilizada (GPU com suporte a operações em bfloat16 e quantização de 4 bits).

3.4.2 Self-training

No campo do aprendizado de máquina aplicado ao Processamento de Linguagem Natural (PLN) com modelos de grande porte, como o *Llama 3.1*, a disponibilidade de dados rotulados de alta qualidade é determinante para o desempenho (DOSSOU et al., 2022; ZHANG; TAKADA, 2025). Contudo, a anotação manual é onerosa, lenta e dependente de conhecimento de domínio. Neste estudo, a estratégia adotada foi híbrida: combinou-se um núcleo rotulado manualmente com técnicas semi-supervisionadas para alavancar o grande volume de dados não rotulados.

A fração rotulada serviu tanto para treinar o modelo de base quanto para ancorar as iterações seguintes. Diferentemente do uso canônico de *Active Learning* que prioriza amostras de maior incerteza (GUPTA et al., 2016; CITOVSKY et al., 2021; TAKEZOE et al., 2023), optou-se aqui por um desenho orientado à **alta confiança** do modelo: a seleção das amostras foi composta por duas metades complementares, de modo a equilibrar qualidade de sinal e diversidade. Metade do lote foi preenchida com exemplos de **baixa entropia** (i.e., previsões mais nítidas), e a outra metade foi amostrada **aleatoriamente** entre instâncias cuja probabilidade máxima prevista superava um limiar de confiança, fixado em $\tau = 0{,}6$. Esse arranjo impôs diversidade sem abrir mão de uma qualidade mínima do sinal probabilístico, mitigando a concentração do treinamento em vizinhanças já bem separadas pelo classificador e reduzindo o risco de reforço de ruído.

3.4.3 Entropia, Pontuação de Confiança e Critério de Seleção

Seja $P(y \mid x)$ a distribuição predita para uma entrada x sobre C classes. A incerteza foi quantificada pela entropia de Shannon

$$H(P(y \mid x)) = -\sum_{i=1}^{C} P(y_i \mid x)\, \log_2 P(y_i \mid x), \tag{3.7}$$

e a confiança pontual foi resumida por $\hat{p}(x) = \max_i P(y_i \mid x)$. No procedimento adotado, *baixa* entropia (pequenos valores de H) e *alta* confiança (grandes valores de $\hat{p}$) operaram como critérios centrais. Para a parcela "determinística" do lote de baixa entropia selecionaram-se os elementos com menores valores de $H(P(y \mid x))$; para a parcela "aleatória controlada", sorteou-se uniformemente entre as instâncias com $\hat{p}(x) > \tau = 0{,}6$. Em termos práticos, definiram-se dois subconjuntos candidatos,

$$\mathrm{S}_{\text{low-}H} = \{x \in \mathcal{U} : H(P(y \mid x)) \leq \eta\} \quad \text{e} \quad \mathcal{S}_{\text{rand-}\tau} = \{x \in \mathcal{U} : \hat{p}(x) > \tau\}, \tag{3.8}$$

dos quais se amostraram metades iguais por lote, com η escolhido empiricamente a partir do percentil inferior da distribuição de entropias no *pool* não rotulado $\mathcal{U}$. Esse desenho favoreceu a cobertura do espaço de entradas mantendo um piso de confiança probabilística.

Adicionalmente, a amostragem incorporou um fator de balanceamento inverso: as classes sobrerrepresentadas no conjunto original tiveram menor probabilidade de serem selecionadas, enquanto as classes minoritárias foram proporcionalmente privilegiadas. Esse ajuste mitigou o viés de frequência do *dataset* inicial, permitindo que a expansão do conjunto rotulado refletisse melhor a diversidade discursiva presente no corpus.

Tabela 5 – Distribuição das amostras.

Fonte da Rotulagem	**Believers**	**Inclusive**	**Deniers**	**Total**
Manual	1097	2241	670	4008
Pseudo-rótulos	641	314	1049	2004
Total	**1738**	**2555**	**1719**	**6012**

Fonte: Produzido pelo autor

3.4.4 Auto-treinamento com Pseudorótulos

A etapa semi-supervisionada seguiu o *self-training* clássico (KARISANI, 2023; PAVLINEK; PODGORELEC, 2017; RIZVE et al., 2021): o modelo treinado no núcleo

rotulado gerou previsões sobre o conjunto não rotulado, e apenas instâncias do **self learning** foram convertidas em pseudorrótulos e agregadas ao conjunto de treinamento.

Os exemplos selecionados foram então incorporados ao conjunto de treinamento ampliado, passando novamente por todas as etapas de ajuste do modelo detalhadas na seção anterior. Dessa forma, assegurou-se a continuidade do ciclo de refinamento, no qual o modelo é progressivamente fortalecido ao integrar previsões confiáveis e representativas do espaço de dados não rotulado.

3.4.5 Síntese da Abordagem

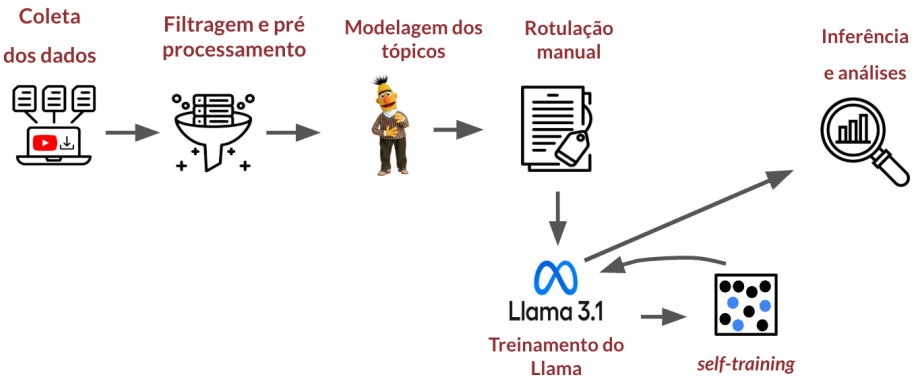

Figura 2 – *Pipeline* da metodologia utilizada
Fonte: Produzido pelo Autor.

O pipeline, conforme visto na Figura 2, combinou: (i) um núcleo inicial formado por 4008 comentários rotulados manualmente; (ii) uma etapa de *self learning* responsável por adicionar metade desse volume (2004 instâncias), selecionadas a partir de uma estratégia que **não** privilegiou exemplos de alta incerteza, mas sim uma mistura de instâncias de **baixa entropia** e exemplos **aleatórios** condicionados a $\hat{p} > 0{,}6$; e (iii) um processo de auto-treinamento baseado na fusão entre o conjunto rotulado manualmente e os comentários complementados via amostragem. Esse arranjo privilegiou sinais mais confiáveis para expandir a base de treinamento, ao mesmo tempo em que injetou diversidade controlada, resultando em maior capacidade de generalização do modelo.

Na prática, essa configuração mostrou-se adequada para capturar as nuances discursivas dos comentários sobre mudanças climáticas e reduzir o risco de vieses decorrentes de erros sistemáticos. Além disso, manteve a coerência estatística das distribuições de classe ao longo das iterações, fortalecendo o modelo sem recorrer a novas rodadas de anotação manual e resultando em ganhos consistentes de desempenho em todas as métricas avaliadas (BAYER, 2025; RAJEEV et al., 2025; CHI et al., 2023).

3.5 Inferência dos Posicionamentos

Com o modelo de classificação devidamente treinado e validado, a etapa seguinte consistiu em aplicá-lo ao conjunto de comentários que não havia sido utilizado no processo de treinamento. Essa fase de inferência possibilitou a categorização automática de um volume expressivo de interações nas três classes definidas.

O objetivo foi ampliar o alcance da análise, permitindo avaliar em larga escala como os usuários do YouTube se posicionam diante das discussões sobre mudanças climáticas. A partir dessa classificação, investigaram-se as distribuições dos comentários em nível global e também segmentadas por recortes temporais, de modo a identificar variações nos posicionamentos e observar como determinados eventos sociais, políticos ou climáticos impactaram a percepção dos usuários.

Além da análise geral das frequências, foram conduzidas investigações condicionais que relacionaram o posicionamento dos comentários originais com o das respostas subsequentes. Esse procedimento permitiu identificar padrões de reforço ou contraposição de opiniões, revelando dinâmicas de engajamento, alinhamento discursivo e potenciais indícios de polarização dentro das conversas.

4 Resultados

Este capítulo apresenta os principais resultados obtidos ao longo do estudo, organizados da seguinte forma. Na Seção 4.1, discutimos a distribuição dos dados coletados, destacando padrões temporais e a análise cumulativa por meio das funções de distribuição cumulativa (CDFs). Em seguida, a Seção 4.2 apresenta os resultados da categorização por tópicos, obtida a partir do modelo *BERTopic*, incluindo a análise da distribuição dos comentários entre os tópicos identificados. Na sequência, a Seção 4.3 descreve a avaliação comparativa entre o modelo *baseline* e a versão aprimorada com *self-training*, detalhando as métricas de desempenho e justificando a escolha do classificador final utilizado para a inferência em toda a base de comentários. Por fim, a Seção 4.4 aprofunda a investigação sobre a evolução temporal desses posicionamentos, relacionando-os a eventos de destaque e explorando padrões de engajamento discursivo ao longo do tempo.

4.1 Distribuição dos dados

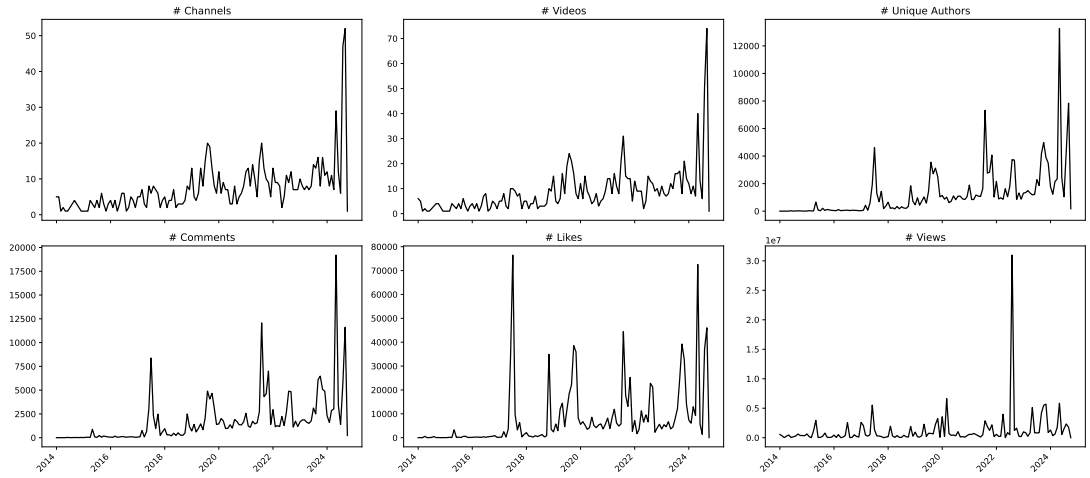

Figura 3 – Caracterização das métricas de popularidade e engajamento ao longo do tempo.
Fonte: Produzido pelo Autor.

A análise inicial da distribuição temporal do conjunto de dados revela padrões claros de variação nas interações do YouTube ao longo do período estudado. Conforme visto na Figura 3, os gráficos ilustram a evolução de seis métricas principais: número de *channels*, *videos*, *unique authors*, *comments*, *likes* e *views*.

Observa-se que tanto o número de *channels* quanto o de *videos* manteve uma tendência relativamente estável na maior parte do período, com picos notáveis em datas específicas. Esses picos sugerem momentos de maior produção de conteúdo possivelmente relacionados a eventos ou publicações relevantes sobre mudanças climáticas. O número de

unique authors apresenta um padrão semelhante, indicando que os momentos de maior atividade na plataforma também atraíram um aumento no engajamento de usuários distintos.

Quanto aos comentários, observa-se uma oscilação mais acentuada, com picos significativos correspondentes aos períodos de maior visualização dos vídeos. Essa variação indica que a discussão sobre o tema é fortemente dependente do alcance do conteúdo, evidenciando a importância de vídeos com grande audiência para impulsionar interações. O número de *likes* segue parcialmente o mesmo padrão, refletindo a correlação entre engajamento e o interesse geral pelo conteúdo.

Finalmente, as métricas de *views*, cujo o eixo vertical está representado em escala logarítmica, exibem os maiores picos, demonstrando que o consumo de vídeo precede e potencialmente impulsiona a produção de comentários e reações. A relação temporal entre visualizações e comentários sugere que eventos de destaque ou vídeos com grande repercussão têm papel central na formação e circulação de opiniões, sendo fundamentais para a análise de padrões de posicionamento sobre mudanças climáticas.

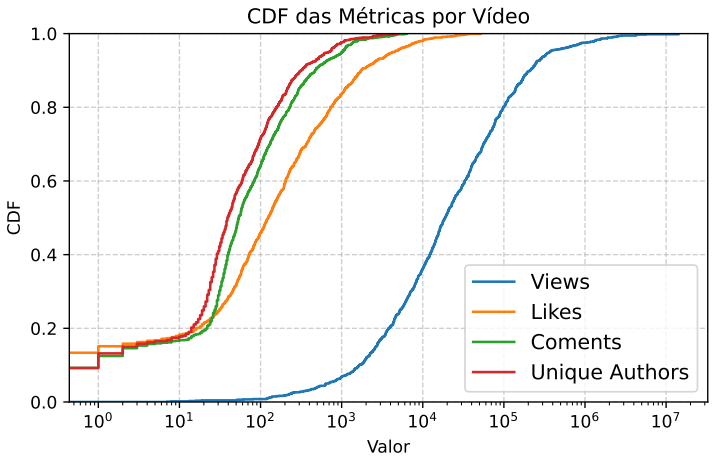

Figura 4 – Funções de distribuição cumulativa (CDFs) das métricas de engajamento no conjunto de vídeos analisado.
Fonte: Produzido pelo Autor.

A Figura 4 apresenta as funções de distribuição cumulativa (CDFs) das métricas de engajamento *views*, *likes*, *comments* e *unique authors* por vídeo. Essas distribuições permitem observar como as interações estão concentradas entre os conteúdos analisados.

Os resultados mostram que a distribuição de *views* é a mais desigual entre todas as métricas. A curva cresce lentamente até a faixa de 10^3 visualizações e apenas então acelera, indicando que a maior parte dos vídeos acumula valores muito baixos, enquanto poucos conteúdos concentram a quase totalidade da audiência, alcançando a ordem de milhões de visualizações. Esse comportamento reflete claramente o padrão de *cauda longa* característico do YouTube.

A distribuição de *likes* também apresenta assimetria, mas menos acentuada do que as visualizações. A curva cresce mais cedo, estabilizando-se por volta de 10^3 curtidas, o que sugere que a maioria dos vídeos atinge um patamar modesto de engajamento positivo, embora poucos vídeos se destaquem com valores muito acima da média.

Já as distribuições de *comments* e *unique authors* apresentam padrões muito próximos, revelando que a dinâmica de comentários está fortemente associada ao número de usuários distintos que interagem com os vídeos. Em ambos os casos, a maioria dos conteúdos apresenta menos de cem comentários ou autores únicos, mas existe uma pequena fração de vídeos que sustenta discussões mais amplas, atingindo milhares de interações.

De modo geral, as CDFs confirmam a concentração das interações: o debate climático no YouTube depende fortemente de um conjunto restrito de vídeos de grande alcance, enquanto a maioria permanece com baixo nível de engajamento. Esse resultado reforça a importância de analisar não apenas a presença do conteúdo, mas também sua capacidade de atrair visualizações e interações significativas.

4.2 Categorização por Tópicos

A aplicação do modelo *BERTopic* sobre o corpus de comentários resultou na formação de 19 tópicos distintos. Essa etapa foi central não apenas para condensar o conjunto massivo de interações em grupos temáticos coesos, mas também porque serviu de base inicial para a amostragem manual de comentários que subsidiou o treinamento do classificador supervisionado.

O algoritmo combina técnicas de redução de dimensionalidade e agrupamento com a extração de termos característicos a partir da métrica c-TF-IDF (GROOTENDORST, 2022a). Essa abordagem garante que os termos destacados sejam mais representativos de cada tópico em particular, evitando que palavras genéricas comprometam a interpretação. A Tabela 6 apresenta os tópicos identificados, suas palavras-chave mais relevantes e uma breve descrição interpretativa de cada grupo.

Tópico	Palavras-chave	Descrição
1	global, aquecimento, planeta, climate, brasil, co2, pessoas	Macro-discussão ampla sobre mudanças climáticas, englobando termos gerais do debate científico e social.
2	obrigado, professor, aula, parabéns, vídeo, sempre	Comentários de agradecimento, elogios e feedback positivo ao conteúdo.
3	queimadas, amazônia, desmatamento, vulcões, ciclos	Associação direta das mudanças climáticas ao contexto ambiental brasileiro, com foco em Amazônia e queimadas.
4	média, pessoas, maior, hoje, planeta, bem	Discussões genéricas ou dispersas, frequentemente ligadas a comparações e generalizações.
5	usp, professor, ricardo, felício, pirula	Embate entre Ricardo Felício e Pirula, polarizando ciência e negacionismo.
6	brasil, água, meio, agro, produção, ambiente	Debate sobre recursos naturais, agronegócio e sustentabilidade no Brasil.
7	presidente, bolsonaro, culpa, governo, país	Politização explícita do debate climático, vinculando-o a disputas partidárias.
8	pessoas, climáticas, mudanças, muitos, agora	Interações genéricas sobre mudança climática, sem foco temático específico.
9	papo, sensacional, conversa, maravilhoso, ver	Discussões em tom de opinião ou conversas informais, geralmente pró-clima.
10	erro, humano, único, deus, pessoas, então	Narrativas de responsabilização do ser humano, em alguns casos misturadas a argumentos religiosos.
11	ah, oh, deus, sim, cara, aqui	Expressões curtas ou conversacionais, muitas vezes sem conteúdo informativo relevante.
12	conhecimento, poder, parabéns, canal, obrigado	Ênfase no valor do conhecimento e aprendizado, com destaque a canais e conteúdos educativos.
13	baixo, água, ruim, ficou, fala, nível	Discussões sobre impactos práticos (ex.: enchentes, nível do mar, eventos climáticos extremos).
14	globo, mídia, carro, brasil, propaganda	Críticas à mídia, associadas a narrativas de manipulação e teorias conspiratórias.
15	sendo, bem, cada, vez, todo, acho	Comentários opinativos vagos, com baixa coesão semântica.
16	precisamos, mudar, cuidar, planeta, pessoas	Chamados à ação coletiva e discursos normativos em defesa do meio ambiente.
17	comentários, fontes, comida, parabéns, vídeo	Interações metacomunicativas (falando sobre o próprio debate ou sobre fontes).
18	basta, ricardo, ouvir, excelente, global	Cluster vinculado a Ricardo Felício, reforçando o debate contra especialistas.

Tabela 6 – Representações de tópicos, suas palavras-chave e descrições interpretativas.
Fonte: Elaborado pelo Autor.

A distribuição dos tópicos não foi uniforme, conforme ilustrado na Figura 5. O Tópico 1, de caráter central, concentrou sozinho mais de 95% de todos os comentários, revelando que o debate climático no YouTube tende a se articular em torno de uma macro-narrativa ampla sobre o aquecimento global. Em contrapartida, tópicos menores, como o Tópico 3 (queimadas e Amazônia) e o Tópico 7 (polarização política), revelam nichos discursivos específicos que introduzem nuances nacionais e ideológicas. Outros tópicos, ainda que residuais, atuam como "subcorrentes" discursivas: alguns vinculam a mudança climática a fenômenos religiosos ou conspiratórios, enquanto outros mobilizam aspectos técnicos e científicos de forma mais restrita.

Essa heterogeneidade mostra que, embora a macro-discussão sobre o clima domine o espaço, micro-discursos desempenham papel importante na fragmentação e no enrique-

cimento da narrativa pública. Em termos analíticos, isso indica que a arena discursiva sobre mudanças climáticas no YouTube não é monolítica, mas composta por uma camada centralizada de consenso ou negação, acompanhada de bolsões narrativos que introduzem elementos nacionais, ideológicos e identitários ao debate.

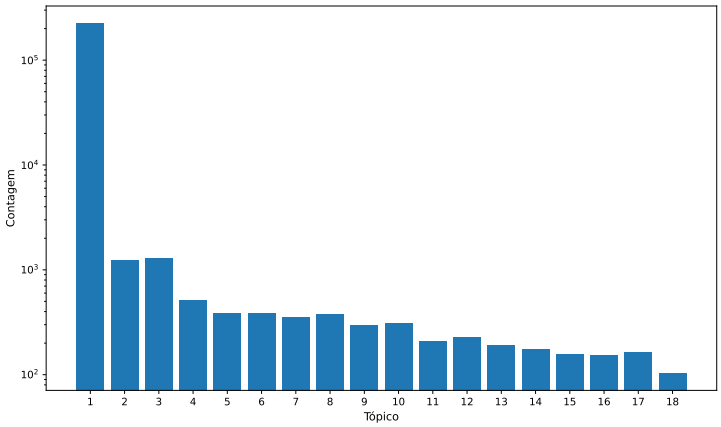

Figura 5 – Distribuição dos tópicos identificados no corpus.
Fonte: Produzido pelo Autor.

4.3 Modelo de Classificação dos Comentários

Os experimentos de classificação foram conduzidos em dois estágios: primeiro, um modelo *baseline* supervisionado, e depois uma versão aprimorada que incorporou *self-training*.

No caso do *baseline*, os resultados estão apresentados na Tabela 7. O desempenho médio obtido na validação cruzada foi de 71,9% de acurácia, com acurácia balanceada de 0,687 e um *Macro-F1* de 0,677. Embora esses valores mostrem que o classificador é capaz de capturar padrões relevantes, observa-se que o equilíbrio entre as classes não foi plenamente alcançado, especialmente para a categoria minoritária (*denier*), que apresentou maior instabilidade entre os *folds*.

Fold	**Acc**	**BalAcc**	**Macro-P**	**Macro-R**	**Macro-F1**
r0	0.700	0.681	0.651	0.681	0.661
r1	0.696	0.650	0.656	0.650	0.648
r2	0.732	0.683	0.689	0.683	0.675
r3	0.739	0.718	0.699	0.718	0.707
r4	0.729	0.703	0.685	0.703	0.692
Média ± 95%	0.719 ± 0.025	0.687 ± 0.032	0.676 ± 0.026	0.687 ± 0.032	0.677 ± 0.029
Modelo Final	0.768	0.749	0.627	0.900	0.651

Tabela 7 – Desempenho do modelo baseline por *fold*, média com IC (95%) e modelo final.
Fonte: Elaborado pelo Autor.

A versão aprimorada, com *active learning* e *self-training*, apresentou ganhos consis-

tentes em todas as métricas (Tabela 8). Nesse caso, a acurácia média subiu para 79,5% e o *Macro-F1* para 0,789, com intervalos de confiança mais estreitos, sugerindo maior robustez e estabilidade entre os *folds*. Esse desempenho confirma a efetividade da estratégia de pseudorrótulos de baixa entropia aliada à seleção ativa de amostras informativas.

Fold	**Acc**	**BalAcc**	**Macro-P**	**Macro-R**	**Macro-F1**
r0	0.784	0.781	0.777	0.781	0.779
r1	0.791	0.786	0.785	0.786	0.782
r2	0.790	0.794	0.786	0.794	0.784
r3	0.823	0.810	0.829	0.810	0.817
r4	0.789	0.785	0.783	0.785	0.784
Média ± 95%	0.795 ± 0.019	0.791 ± 0.014	0.792 ± 0.026	0.791 ± 0.014	0.789 ± 0.020
Modelo Final	0.796	0.786	0.814	0.791	0.789

Tabela 8 – Desempenho do modelo com Self-Training por *fold*, média com IC (95%) e modelo final.
Fonte: Elaborado pelo Autor.

A Figura 6 sintetiza essa evolução, comparando diretamente os modelos finais. Nota-se que, enquanto o baseline obteve resultados mais modestos, o modelo com *active learning* superou sistematicamente as métricas de referência, em especial no *Macro-F1* e na precisão macro. Apesar de uma leve queda em *Macro-Recall*, o ganho global indica maior equilíbrio na classificação das três categorias, tornando-o mais adequado para capturar as nuances dos discursos analisados.

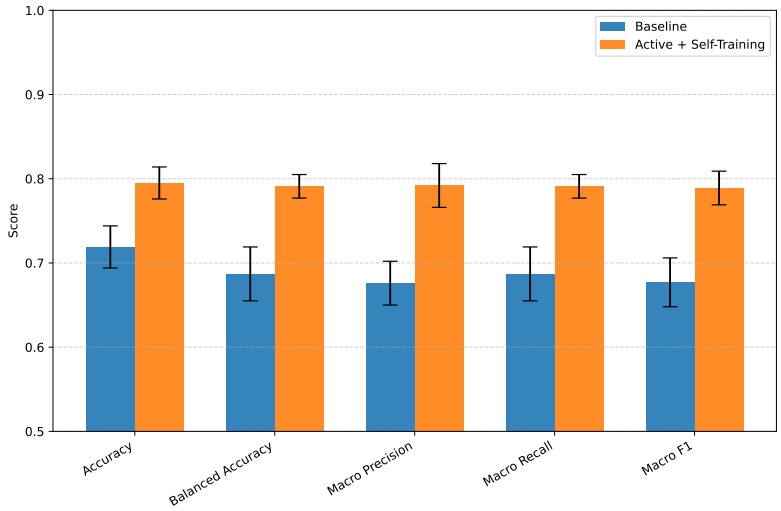

Figura 6 – Comparação dos dois modelos finais com intervalos de confiança de 95%.
Fonte: Produzido pelo Autor.

Com base nesses resultados, optamos por adotar o modelo treinado de maneira hibrida, com *self-training* e rótulos manuais, como classificador final. Essa escolha se justifica não apenas pelos ganhos expressivos em métricas globais, mas também pela maior

estabilidade observada entre os *folds*. Esse modelo, portanto, foi utilizado para realizar a inferência sobre toda a base de comentários e fundamentar as análises.

Para tornar mais clara a performance do classificador em cada categoria, a Tabela 9 apresenta as métricas de *precision*, *recall* e F_1 para as três classes de interesse. Observa-se que, embora o desempenho seja superior nas classes mais representadas, o modelo alcança valores satisfatórios mesmo na categoria mais difícil (*denier*), reduzindo o risco de vieses acentuados na etapa de análise.

Classe	**Precision**	**Recall**	**F1**
Believer	0.861	0.878	0.878
Inconclusive	0.703	0.825	0.831
Denier	0.798	0.786	0.786

Tabela 9 – Desempenho do modelo final por classe.
Fonte: Elaborado pelo Autor.

4.4 Análise dos posicionamentos

Após o treinamento e a validação do classificador, o modelo selecionado foi aplicado a todo o conjunto de comentários válidos — totalizando 242.367 — com o objetivo de inferir o posicionamento dos usuários em larga escala. Essa etapa corresponde à aplicação prática da metodologia desenvolvida, permitindo a rotulação automática de uma massa expressiva de dados e possibilitando análises aprofundadas sobre a percepção pública em torno das mudanças climáticas no contexto brasileiro.

A Tabela 10 apresenta a distribuição final dos comentários por classe, organizados segundo as três categorias estabelecidas: *believer* (aceitação das mudanças climáticas), *denier* (negação do fenômeno) e *inconclusive* (comentários ambíguos ou sem posicionamento claro).

Classe	**Total de Comentários**	**Proporção (%)**
Believer	46.867	19.3%
Inconclusive	157.414	64.9%
Denier	38.086	15.7%
Total	242.367	100%

Tabela 10 – Distribuição final dos comentários classificados por posicionamento.
Fonte: Elaborado pelo Autor.

A análise da distribuição revela um quadro complexo. A classe *inconclusive*, que agrupa comentários ambíguos, irônicos ou sem posicionamento explícito, concentra quase dois terços de todas as interações (64,9%). Esse predomínio sugere que, embora o debate climático esteja presente na plataforma, grande parte dos usuários não se compromete de

forma clara com uma narrativa de aceitação ou de negação, seja pela superficialidade dos comentários, seja pelo uso de linguagem coloquial pouco informativa.

Em contrapartida, a categoria *believer* corresponde a 19,3% do corpus, refletindo a presença de um grupo expressivo de usuários que reconhecem a existência das mudanças climáticas e suas causas antropogênicas. Ainda que minoritário em relação ao total de comentários, esse contingente demonstra que há uma base ativa de discursos alinhados ao consenso científico, reforçando a circulação de narrativas pró-ciência dentro do espaço digital brasileiro.

Já os comentários classificados como *denier*, que representam 15,7% do conjunto, evidenciam a persistência do negacionismo climático. Embora numericamente inferiores aos *believers*, esses discursos são suficientemente numerosos para influenciar a percepção pública, sobretudo em um ambiente como o YouTube, em que o engajamento se concentra em poucos conteúdos de grande alcance. A presença dessa fração considerável de negacionistas confirma a hipótese de que a plataforma funciona não apenas como espaço de divulgação científica, mas também como terreno fértil para a circulação de desinformação e teorias conspiratórias.

Assim, os resultados da Tabela 10 demonstram que o debate climático no YouTube é caracterizado por um predomínio de interações ambíguas, acompanhado por um equilíbrio relativo entre narrativas de aceitação e negação. Esse panorama destaca tanto os limites quanto as oportunidades de se utilizar a plataforma como termômetro da opinião pública: por um lado, a fragmentação do discurso e a força da desinformação; por outro, a persistência de vozes alinhadas à ciência, ainda que em proporção reduzida.

A Figura 7 apresenta a evolução temporal cumulativa dos comentários classificados. O gráfico permite observar a trajetória do debate climático no YouTube ao longo dos anos, destacando tanto o crescimento absoluto das interações quanto as diferenças relativas entre os posicionamentos.

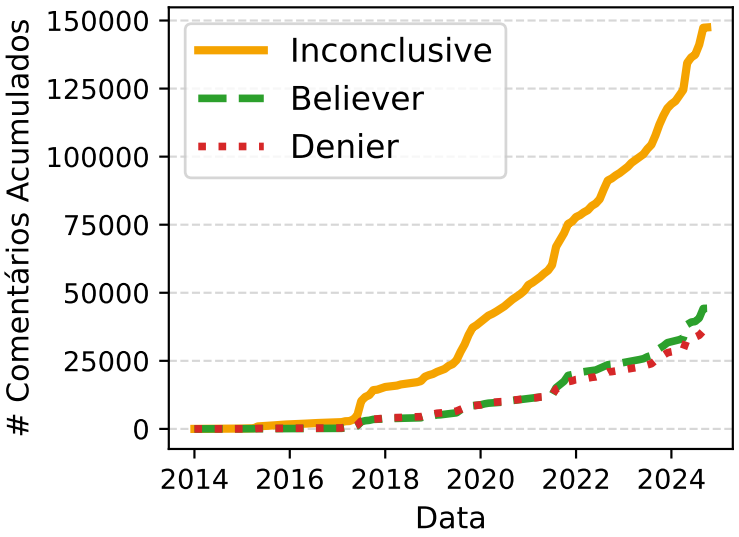

Figura 7 – Análise temporal dos comentários por classe.
Fonte: Produzido pelo Autor.

De maneira geral, nota-se uma expansão acentuada do volume de comentários a partir de 2021, com forte aceleração entre 2022 e 2024. Esse período coincide com a intensificação da cobertura midiática sobre eventos climáticos extremos, como queimadas, enchentes e ondas de calor, o que provavelmente impulsionou a participação dos usuários. Antes desse ponto, o crescimento era modesto e relativamente equilibrado entre as classes.

Os comentários classificados como *inconclusives* representam a maior parcela do corpus, acumulando mais de 150 mil registros até o final de 2024. A predominância dessa categoria sugere que, embora o tema das mudanças climáticas seja amplamente discutido, grande parte das interações não expressa posicionamentos claros de aceitação ou negação, mas se manifesta de forma ambígua ou tangencial.

Já os *believers* apresentam um crescimento expressivo a partir de 2022, quase alcançando a marca de 50 mil comentários acumulados. Esse resultado indica que uma parcela significativa do público engaja de maneira explícita com a narrativa de reconhecimento do aquecimento global, especialmente em períodos de maior visibilidade de eventos relacionados ao tema.

Por sua vez, os *deniers* aparecem em menor escala, somando pouco mais de 25 mil comentários ao final da série temporal. Apesar do volume inferior, sua presença contínua evidencia a persistência de discursos negacionistas no espaço digital, ainda que restritos a uma fração minoritária da base analisada.

Para investigar como os posicionamentos se distribuem entre vídeos mais e menos populares, calculamos a Função de Distribuição Acumulada (CDF) do percentual de comentários favoráveis e contrários em cada vídeo. Os resultados estão apresentados na

Figura 8.

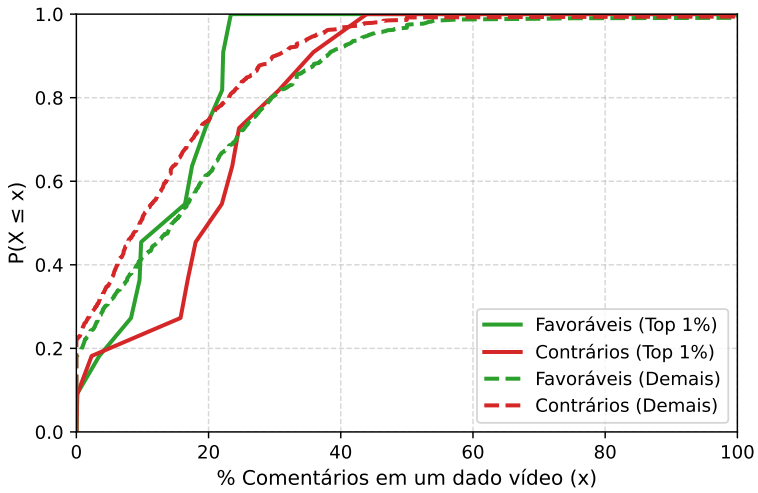

Figura 8 – CDF do percentual de comentários favoráveis e contrários nos vídeos mais comentados (Top 1%) e nos demais.
Fonte: Produzido pelo Autor.

A análise revela que os vídeos mais comentados (Top 1%) atuam como polos de debate, reunindo tanto comentários favoráveis quanto contrários em proporções mais próximas. Embora a aceitação das mudanças climáticas ainda apareça de forma relevante, a contestação se distribui de maneira comparável nesses conteúdos de maior visibilidade, indicando uma dinâmica de polarização acentuada.

Nos demais vídeos, que representam a vasta maioria da base, a distribuição é menos polarizada: comentários contrários tendem a se concentrar em percentuais menores, enquanto os favoráveis aparecem mais dispersos ao longo dos valores possíveis. Isso sugere que, fora dos grandes polos de repercussão, o discurso de aceitação mantém certa predominância, ainda que não absoluta.

Em síntese, os resultados indicam que os vídeos de maior visibilidade funcionam como arenas de embate mais equilibradas entre narrativas, enquanto o restante da plataforma apresenta um viés levemente mais favorável à aceitação, mas sem excluir a presença de contestação.

Para compreender os termos mais recorrentes no discurso dos usuários, foram geradas nuvens de palavras a partir dos cinco vídeos com maior percentual de comentários classificados como *believers* e *deniers*. Esse recorte permite capturar o vocabulário característico dos contextos mais polarizados, evidenciando tanto a frequência das palavras quanto a ênfase semântica atribuída por cada grupo. A Figura 9 apresenta o resultado da análise, destacando em verde os termos mais comuns entre os comentários *believers* e em vermelho aqueles predominantes nos comentários *deniers*.

Figura 9 – Nuvens de palavras dos cinco vídeos com maior percentual de comentários *believers* (a) e *deniers* (b).

Fonte: Produzido pelo Autor.

Observa-se que, entre os comentários favoráveis, emergem termos como *"agro"*, *"governo"*, *"natureza"*, *"meio ambiente"* e referências a atores políticos como *"Lula"*. Essa composição lexical sugere que a defesa da mudança climática está frequentemente articulada a pautas ambientais e políticas nacionais, reforçando o papel das políticas públicas e do agronegócio no debate.

Por outro lado, a nuvem dos comentários contrários é dominada por expressões como *"aquecimento global"*, *"globalistas"*, *"projeto"*, *"teoria"*, *"controle"* e *"HAARP"*, refletindo narrativas de contestação ao fenômeno, muitas vezes atravessadas por teorias conspiratórias ou questionamentos sobre a legitimidade científica. A presença de termos associados a *globalismo* e a discursos de descrédito evidencia uma dimensão política e ideológica da negação.

Até este ponto, a investigação concentrou-se nos comentários de forma agregada, observando padrões gerais de posicionamento ao longo do tempo e entre os vídeos. Entretanto, considerando a dinâmica interativa das discussões no YouTube, é relevante avaliar também como o posicionamento de um comentário inicial (*comentário pai*) pode influenciar a natureza das respostas que ele recebe. Em outras palavras, buscou-se verificar se comentários classificados como *believers* ao reconhecimento das mudanças climáticas tendem a atrair respostas alinhadas ao mesmo ponto de vista, ou se geram maior oposição. O mesmo raciocínio aplica-se aos comentários da classe *deniers*.

Essa análise condicional permite captar de forma mais clara os mecanismos de engajamento e contraposição dentro das conversas, indo além da distribuição agregada. A Figura 10 apresenta os resultados, mostrando a distribuição percentual das classes de respostas em função da classe do comentário original. Com isso, é possível observar se determinados posicionamentos atuam como polos de concordância ou como gatilhos de contestação.

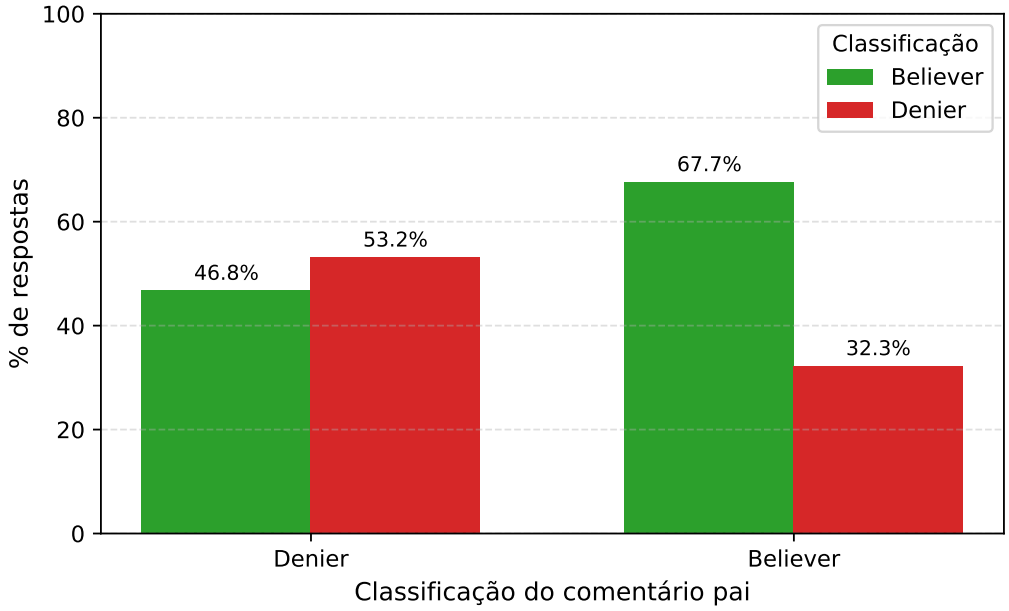

Figura 10 – Distribuição percentual das classes das respostas condicionada ao comentário pai.

Fonte: Produzido pelo Autor.

Os resultados mostram um padrão claro de alinhamento: aproximadamente 67,7% das respostas a comentários *believers* também foram classificadas como *believers*, indicando uma tendência de reforço discursivo. Da mesma forma, quando o comentário inicial é *denier*, observa-se que 53,2% das respostas seguem essa mesma orientação (*denier*). Esse comportamento sugere que a seção de comentários funciona, em grande medida, como um espaço de eco e reforço ideológico, em que respostas tendem a reproduzir a posição já estabelecida pelo comentário pai.

Apesar desse predomínio do alinhamento, nota-se que existe também um percentual considerável de respostas que assumem a posição oposta, especialmente em relação aos comentários contrários, nos quais quase metade das réplicas (46,8%) são favoráveis. Esse equilíbrio relativo sugere que comentários de negação climática despertam maior contestação e oposição explícita do público em comparação com comentários favoráveis, que atraem uma proporção menor de discordâncias (apenas 32,3% de respostas contrárias).

Portanto, os dados indicam que, embora haja um efeito de "câmara de eco" nos comentários do YouTube, ele não é absoluto: o debate permanece permeado por interações de conflito, especialmente em torno de narrativas negacionistas. Esse padrão é relevante para compreender a dinâmica de propagação das narrativas climáticas, pois aponta que comentários contrários ao consenso científico tendem a gerar mais polarização nas respostas do que aqueles que expressam aceitação das mudanças climáticas.

5 Conclusão

Neste trabalho, investigamos as narrativas sobre mudanças climáticas em comentários publicados no YouTube no Brasil, com o objetivo de compreender como os usuários se posicionam diante do tema e de que forma essas interações se organizam ao longo do tempo e em torno de eventos específicos. Diferentemente de estudos anteriores, que frequentemente se limitam a análises manuais em pequenas amostras ou a métricas descritivas simples, adotamos uma abordagem baseada em aprendizado de máquina e técnicas de mineração de texto para lidar com um volume massivo de dados. Para isso, desenvolvemos um *pipeline* que abrange desde o pré-processamento até a aplicação de modelos de classificação e análise temporal, combinando métodos quantitativos e qualitativos.

A análise revelou padrões importantes na organização discursiva. A aplicação do *BERTopic* mostrou que a maior parte dos comentários se concentra em um núcleo amplo e genérico sobre o aquecimento global, ainda que tópicos menores tragam nuances importantes, como menções a queimadas, à Amazônia e à política nacional. O classificador de posicionamento, aprimorado com *self-training*, obteve desempenho superior em relação ao modelo baseline, alcançando um *Macro-F1* próximo de 0,79. Esse resultado permitiu aplicar o modelo a toda a base de comentários e, com isso, estimar que cerca de 19,3% dos comentários são favoráveis às mudanças climáticas, 15,7% são contrários e 64,9% apresentam posicionamento inconclusivo. Também foi observado que comentários iniciais favoráveis atraem respostas de mesmo alinhamento, enquanto comentários contrários tendem a gerar maior contestação, evidenciando a dinâmica de reforço e conflito nas interações.

A principal contribuição deste trabalho está na integração de métodos de aprendizado de máquina com análises discursivas em larga escala, oferecendo uma visão abrangente sobre como o debate climático se estrutura no YouTube brasileiro. Essa combinação permitiu ir além da descrição de frequências, possibilitando compreender relações entre popularidade dos vídeos, alinhamentos discursivos e padrões de interação entre usuários. A originalidade da abordagem está justamente no uso de técnicas modernas de modelagem de tópicos e de classificação ativa para enfrentar os desafios de um corpus altamente heterogêneo e desbalanceado.

Apesar dos avanços, algumas limitações devem ser reconhecidas. A base analisada restringiu-se a vídeos em língua portuguesa e a uma única plataforma, o que limita a generalização dos resultados para outros contextos culturais e mídias sociais. Além disso, a predominância de comentários inconclusivos sugere que parte significativa da discussão ocorre de forma ambígua, sem expressar claramente um posicionamento, o que

pode comprometer a precisão interpretativa. Outro ponto é que, embora o modelo tenha alcançado bons índices de desempenho, ainda há dificuldades em capturar adequadamente as classes minoritárias, reflexo do desbalanceamento inerente ao corpus.

Como trabalhos futuros, sugerimos expandir a coleta de dados para múltiplas plataformas e idiomas, de modo a enriquecer as comparações entre diferentes públicos e contextos. Também seria relevante explorar a aplicação de modelos hierárquicos de classificação, capazes de organizar o posicionamento em diferentes níveis de granularidade (por exemplo, distinções gerais entre favoráveis e contrários e, em seguida, subdivisões mais específicas de cada grupo). Esse tipo de abordagem pode capturar nuances discursivas que não emergem em classificações planas, permitindo identificar variações internas relevantes em cada classe. Finalmente, análises qualitativas mais detalhadas podem aprofundar a compreensão das narrativas identificadas, revelando estratégias retóricas e mecanismos de desinformação que nem sempre emergem em métricas quantitativas.

Assim, este trabalho oferece uma contribuição significativa para o entendimento das dinâmicas discursivas sobre mudanças climáticas no espaço digital, destacando tanto a centralidade de narrativas amplas e genéricas quanto a influência de eventos externos e da polarização política na formação do debate. Ao propor uma metodologia escalável e adaptável, abre caminho para futuras investigações que pretendam mapear de forma sistemática como a opinião pública se constrói e se transforma diante da crise climática.

Referências

BAYER, M. Activellm: Large language model-based active learning for textual few-shot scenarios. In: _____. Springer Nature, 2025. p. 89–112. Disponível em: <https://doi.org/10.1007/978-3-658-48778-2_7>. Citado na página 32.

CAMPELLO, R. J.; MOULAVI, D.; SANDER, J. Density-based clustering based on hierarchical density estimates. In: SPRINGER. *Pacific-Asia conference on knowledge discovery and data mining*. [S.l.], 2013. p. 160–172. Citado na página 23.

CHEN, L. Research on code generation technology based on llm pre-training. *Frontiers in Computing and Intelligent Systems*, v. 10, p. 69–75, 10 2024. ISSN 2832-6024. Disponível em: <https://doi.org/10.54097/scrwpt34>. Citado na página 26.

CHI, T.-Y. et al. Wc-sbert: Zero-shot text classification via sbert with self-training for wikipedia categories. *arXiv (Cornell University)*, Cornell University, 01 2023. Disponível em: <https://arxiv.org/abs/2307.15293>. Citado na página 32.

CITOVSKY, G. et al. Batch active learning at scale. *arXiv (Cornell University)*, Cornell University, 01 2021. Disponível em: <https://arxiv.org/abs/2107.14263>. Citado na página 30.

CUI, Y. et al. Class-balanced loss based on effective number of samples. In: . [s.n.], 2019. Disponível em: <https://doi.org/10.1109/cvpr.2019.00949>. Citado na página 29.

DAVILA, A.; COLAN, J.; HASEGAWA, Y. Beyond single models: Enhancing llm detection of ambiguity in requests through debate. arXiv, 01 2025. Disponível em: <https://arxiv.org/abs/2507.12370>. Citado na página 27.

DIAS, A. et al. Análise da percepção do uso de cigarros eletrônicos no brasil por meio de comentários no youtube. In: SBC. *Brazilian Symposium on Multimedia and the Web (WebMedia)*. [S.l.], 2024. p. 45–53. Citado na página 25.

DOSSOU, B. F. P. et al. Afrolm: A self-active learning-based multilingual pretrained language model for 23 african languages. p. 52–64, 01 2022. Disponível em: <https://doi.org/10.18653/v1/2022.sustainlp-1.11>. Citado na página 30.

ELROY, O.; KOMENDANTOVA, N.; YOSIPOF, A. Cyber-echoes of climate crisis: Unraveling anthropogenic climate change narratives on social media. *Current Research in Environmental Sustainability*, Elsevier BV, v. 7, p. 100256–100256, 01 2024. ISSN 2666-0490. Disponível em: <https://doi.org/10.1016/j.crsust.2024.100256>. Citado 2 vezes nas páginas 14 e 15.

FARHADPOUR, S.; WARNER, T. A.; MAXWELL, A. E. Selecting and interpreting multiclass loss and accuracy assessment metrics for classifications with class imbalance: Guidance and best practices. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 16, p. 533–533, 01 2024. ISSN 2072-4292. Disponível em: <https://doi.org/10.3390/rs16030533>. Citado na página 28.

FLEISS, J. L.; LEVIN, B.; PAIK, M. C. *Statistical methods for rates and proportions*. [S.l.]: john wiley & sons, 2013. Citado na página 25.

FLORES, N. M.; MEDEIROS, P. Muniz de. Science on youtube: legitimation strategies of brazilian science youtubers. *Revue française des sciences de l'information et de la communication*, Société Française de Sciences de l'Information et de la Communication, n. 15, 2018. Citado na página 12.

FURTADO, C. de M.; DIAS, T. M. R. O combate à desinformação no youtube: uma análise de similitude dos comentários em um vídeo sobre negacionismo climático publicado pela bbc news brasil. 01 2024. Disponível em: <https://doi.org/10.22477/vii.widat.203>. Citado 2 vezes nas páginas 12 e 15.

GROOTENDORST, M. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*, 2022. Citado 3 vezes nas páginas 22, 23 e 36.

GROOTENDORST, M. P. *The Algorithm - BERTopic — maartengr.github.io*. 2022. <https://maartengr.github.io/BERTopic/algorithm/algorithm.html>. Acesso em 15 de Agosto de 2025]. Citado na página 22.

GUO, C. et al. On calibration of modern neural networks. *arXiv (Cornell University)*, Cornell University, 01 2017. Disponível em: <https://arxiv.org/abs/1706.04599>. Citado na página 30.

GUPTA, A. et al. Font identification in historical documents using active learning. *arXiv (Cornell University)*, Cornell University, 01 2016. Disponível em: <https://arxiv.org/pdf/1601.07252.pdf>. Citado na página 30.

HU, J. E. et al. Lora: Low-rank adaptation of large language models. *arXiv (Cornell University)*, Cornell University, 01 2021. Disponível em: <https://arxiv.org/abs/2106.09685>. Citado na página 27.

HUSSEIN, B. M.; SHAREEF, S. M. An empirical study on the correlation between early stopping patience and epochs in deep learning. *ITM Web of Conferences*, EDP Sciences, v. 64, p. 1003–1003, 01 2024. ISSN 2271-2097, 2431-7578. Disponível em: <https://doi.org/10.1051/itmconf/20246401003>. Citado na página 28.

KARISANI, P. Neural networks against (and for) self-training: Classification with small labeled and large unlabeled sets. In: . [s.n.], 2023. p. 12148–12162. Disponível em: <https://doi.org/10.18653/v1/2023.findings-acl.769>. Citado na página 31.

KRSTAJIć, D. et al. Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of Cheminformatics*, BioMed Central, v. 6, 03 2014. ISSN 1758-2946. Disponível em: <https://doi.org/10.1186/1758-2946-6-10>. Citado na página 28.

KULL, M. et al. Beyond temperature scaling: Obtaining well-calibrated multiclass probabilities with dirichlet calibration. *arXiv (Cornell University)*, Cornell University, 01 2019. Disponível em: <https://arxiv.org/abs/1910.12656>. Citado na página 30.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *biometrics*, JSTOR, p. 159–174, 1977. Citado na página 26.

LANG, J. C.; GUO, Z.; HUANG, S. Y. A comprehensive study on quantization techniques for large language models. p. 224–231, 12 2024. Disponível em: <https://doi.org/10.1109/icairc64177.2024.10899941>. Citado 2 vezes nas páginas 27 e 28.

LEWKOWYCZ, A. How to decay your learning rate. *arXiv (Cornell University)*, Cornell University, 01 2021. Disponível em: <https://arxiv.org/abs/2103.12682>. Citado na página 30.

LIMA, R. O. et al. (un) certainty in science and climate change: a longitudinal analysis (2014–2022) of narratives about climate science on social media in brazil (instagram, facebook, and twitter). *Andre and Lycarião, Diógenes and Oliveira, Thaiane and Evangelista, Simone and Massarani, Luisa and Alves, Marcelo,(Un) certainty in Science and Climate Change: a Longitudinal Analysis (2014–2022) of Narratives About Climate Science on Social Media in Brazil (Instagram, Facebook, and Twitter)(march 30, 2024)*, 2024. Citado na página 12.

LIPTON, Z. C.; ELKAN, C.; BALAKRISHNAN, N. Thresholding classifiers to maximize f1 score. *arXiv (Cornell University)*, Cornell University, 01 2014. Disponível em: <https://arxiv.org/abs/1402.1892>. Citado na página 28.

LOSHCHILOV, I.; HUTTER, F. Fixing weight decay regularization in adam. 02 2018. Disponível em: <https://openreview.net/pdf?id=rk6qdGgCZ>. Citado na página 30.

MCINNES, L.; HEALY, J.; MELVILLE, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018. Citado na página 23.

OPITZ, J. A closer look at classification evaluation metrics and a critical reflection of common evaluation practice. *Transactions of the Association for Computational Linguistics*, Association for Computational Linguistics, v. 12, p. 820–836, 01 2024. ISSN 2307-387X. Disponível em: <https://doi.org/10.1162/tacl_a_00675>. Citado na página 28.

PATIL, R.; KHOT, P.; GUDIVADA, V. N. Analyzing llama3 performance on classification task using lora and qlora techniques. *Applied Sciences*, Multidisciplinary Digital Publishing Institute, v. 15, p. 3087–3087, 03 2025. ISSN 2076-3417. Disponível em: <https://doi.org/10.3390/app15063087>. Citado 2 vezes nas páginas 27 e 28.

PAVLINEK, M.; PODGORELEC, V. Text classification method based on self-training and lda topic models. *Expert Systems with Applications*, Elsevier BV, v. 80, p. 83–93, 03 2017. ISSN 0957-4174, 1873-6793. Disponível em: <https://doi.org/10.1016/j.eswa.2017.03.020>. Citado na página 31.

PEREYRA, G. et al. Regularizing neural networks by penalizing confident output distributions. arXiv, 01 2017. Disponível em: <https://arxiv.org/abs/1701.06548>. Citado na página 29.

PRECHELT, L. Automatic early stopping using cross validation: quantifying the criteria. *Neural Networks*, Elsevier BV, v. 11, p. 761–767, 06 1998. ISSN 0893-6080, 1879-2782. Disponível em: <https://doi.org/10.1016/s0893-6080(98)00010-0>. Citado na página 28.

QIU, Q.; SONG, Z. A nonuniform weighted loss function for imbalanced image classification. p. 78–82, 02 2018. Disponível em: <https://doi.org/10.1145/3191442.3191458>. Citado na página 29.

RAJEEV, A. et al. Small sample-based adaptive text classification through iterative and contrastive description refinement. arXiv, 01 2025. Disponível em: <https://arxiv.org/abs/2508.00957>. Citado na página 32.

RENDUCHINTALA, A.; KONUK, T.; KUCHAIEV, O. Tied-lora: Enhacing parameter efficiency of lora with weight tying. *arXiv (Cornell University)*, Cornell University, 01 2023. Disponível em: <https://arxiv.org/abs/2311.09578>. Citado na página 27.

RIZVE, M. N. et al. In defense of pseudo-labeling: An uncertainty-aware pseudo-label selection framework for semi-supervised learning. *arXiv (Cornell University)*, Cornell University, 01 2021. Disponível em: <https://arxiv.org/abs/2101.06329>. Citado na página 31.

ROJAS, C. et al. Hierarchical machine learning models can identify stimuli of climate change misinformation on social media. *Communications Earth & Environment*, Nature Publishing Group UK London, v. 5, n. 1, p. 436, 2024. Citado na página 12.

SALLES, D. et al. The far-right smokescreen: Environmental conspiracy and culture wars on brazilian youtube. *Social Media+ Society*, SAGE Publications Sage UK: London, England, v. 9, n. 3, p. 20563051231196876, 2023. Citado na página 12.

SHAHBAZI, Z.; JALALI, R.; SHAHBAZI, Z. Ai-driven framework for evaluating climate misinformation and data quality on social media. *Future Internet*, MDPI, v. 17, n. 6, p. 231, 2025. Citado na página 12.

SOUZA, F.; NOGUEIRA, R.; LOTUFO, R. Bertimbau: pretrained bert models for brazilian portuguese. In: SPRINGER. *Brazilian conference on intelligent systems*. [S.l.], 2020. p. 403–417. Citado na página 22.

STORANI, S. et al. Relative engagement with sources of climate misinformation is growing across social media platforms. *Scientific Reports*, Nature Publishing Group UK London, v. 15, n. 1, p. 18629, 2025. Citado na página 12.

TAKEZOE, R. et al. Deep active learning for computer vision: Past and future. *APSIPA Transactions on Signal and Information Processing*, Cambridge University Press, v. 12, 01 2023. ISSN 2048-7703. Disponível em: <https://doi.org/10.1561/116.00000057>. Citado na página 30.

TANURE, R. R. et al. Caracterização do debate online sobre cigarro eletrônico no brasil: Uma análise de tópicos de discussão no youtube. In: SBC. *Brazilian Workshop on Social Network Analysis and Mining (BraSNAM)*. [S.l.], 2025. p. 54–64. Citado na página 23.

TREEN, K. M. d.; WILLIAMS, H. T.; O'NEILL, S. J. Online misinformation about climate change. *Wiley Interdisciplinary Reviews: Climate Change*, Wiley Online Library, v. 11, n. 5, p. e665, 2020. Citado na página 12.

TREEN, K. M. d'I. et al. Discussion of climate change on reddit: Polarized discourse or deliberative debate? *Environmental Communication*, Taylor & Francis, v. 16, p. 680–698, 04 2022. ISSN 1752-4032, 1752-4040. Disponível em: <https://doi.org/10.1080/17524032.2022.2050776>. Citado 2 vezes nas páginas 16 e 17.

WEERAWARDHENA, S. et al. Llama-3.1-foundationai-securityllm-8b-instruct technical report. arXiv, 01 2025. Disponível em: <https://arxiv.org/abs/2508.01059>. Citado na página 26.

YOUNG, S. I. Foundations of large language model compression – part 1: Weight quantization. arXiv, 01 2024. Disponível em: <https://arxiv.org/abs/2409.02026>. Citado 2 vezes nas páginas 27 e 28.

ZHANG, Y.; TAKADA, S. Applying llms to active learning: Towards cost-efficient cross-task text classification without manually labeled data. arXiv, 01 2025. Disponível em: <https://arxiv.org/abs/2502.16892>. Citado na página 30.

ZOLA, P.; RAGNO, C.; CORTEZ, P. A google trends spatial clustering approach for a worldwide twitter user geolocation. *Information Processing & Management*, Elsevier, v. 57, n. 6, p. 102312, 2020. Citado na página 20.