



DCC



DEPARTAMENTO DE
CIÊNCIA DA COMPUTAÇÃO

Polarização em Redes Sociais

Bruno Hott, Bruno P. Santos, Túlio C. Loures,
Fabrício Benevenuto e Pedro O.S.
Vaz-de-Melo

7 de novembro de 2022

Outline

1. Introdução
2. Dados no contexto de polarização
3. Viés de entidades dos dados
4. Métricas de polarização
5. Análises sobre polarização
6. Prática: Covid-19 e Hidroxicloroquina (HCQ)
7. Considerações finais

The background consists of two overlapping geometric shapes: a teal triangle in the top-left corner and a light beige triangle in the bottom-left corner, both pointing towards the center. The rest of the background is white.

Introdução

- ▶ A Internet potencializou a disseminação de opiniões.
- ▶ A polarização ganhou enfoque na academia e na indústria.
- ▶ Impacto na sociedade;
 - ▶ Muitos cidadãos não dialogam com familiares.
- ▶ Impacto na política;
 - ▶ Campanhas políticas maliciosas.
- ▶ Impacto na economia.
 - ▶ Incerteza política.

- ▶ Este curso aborda a Polarização por meio de uma perspectiva teórica e prática.
- ▶ Exploraremos a estrutura, organização, desafios e aplicações dos estudos de polarização com enfoque nas mídias sociais.
- ▶ Discutiremos conceitos, motivaremos o aluno com casos recentes e históricos
- ▶ Apresentaremos uma proposta de metodologia para o estudo em polarização.
- ▶ **O que é Polarização?**

- ▶ No Brasil, os próprios candidatos(as) à presidência afirmam¹, que as eleições presidenciais de 2022 estão polarizadas.
- ▶ Na Internet, o Facebook tem sido acusado de promover conteúdos que dividem a população, podendo criar ou acelerar polarizações políticas².
- ▶ Explorar a pré-disposição do cérebro humano em consumir conteúdos que reforçam sua visão de mundo.
- ▶ As plataformas estariam entregando cada vez mais conteúdos extremos com o objetivo de que os usuários fiquem por mais tempo na plataforma?

¹Primeiro debate presidencial 2022

²<https://www.washingtonpost.com/opinions/2020/10/26/facebook-algorithm-conservative-liberal-extremes/>

- ▶ Alguns tópicos são **alvos de debate**: como esportes, política, leis de drogas..
- ▶ Empiricamente, essas discussões tendem a fazer com que as pessoas movam suas opiniões para versões mais **extremas** delas mesmas.
- ▶ As pessoas estão ouvindo suas próprias vozes sendo **ecoadas** por seus semelhantes.
- ▶ Como consequência, se torna mais difícil para que a população possa resolver os problemas que a sociedade enfrenta conjuntamente, como o aquecimento global, por exemplo.

- ▶ As mídias sociais vem sendo apontadas como estimuladoras da polarização ao criarem as chamadas “**câmaras de eco**”.
- ▶ Onde as pessoas são estimuladas por pensamentos análogos, o que as isolam de divergências de opiniões e pensamentos.
- ▶ Cautela! O aumento da polarização pode não estar diretamente ligada ao uso das mídias sociais.
- ▶ Níveis de polarização crescem mais na população idosa.

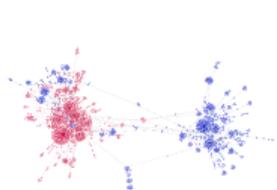
Introdução

Definição

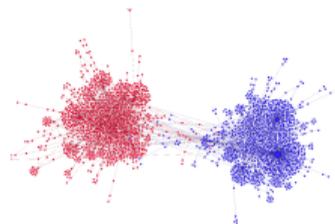
- ▶ “Polarização é a concentração de ideias em um polo que se opõe a outro”
[Priberam, 2022]
- ▶ “Polarização se dá quando membros de grupos da sociedade se movem na direção dos extremos”
[Fiorina et al., 2008]
- ▶ “Polarização é um processo social, onde grupos se dividem em dois sub-grupos opostos cada qual com posições conflitantes umas as outras, com alguns poucos indivíduos neutros”
[Sunstein, 1999]

Introdução

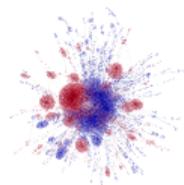
Exemplos de redes polarizadas



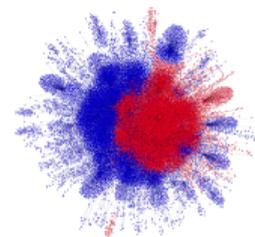
(a)



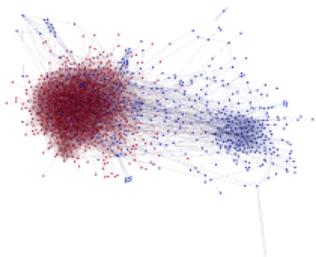
(b)



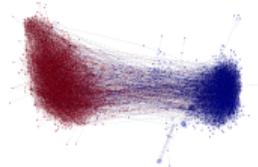
(c)



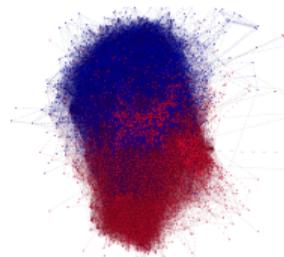
(d)



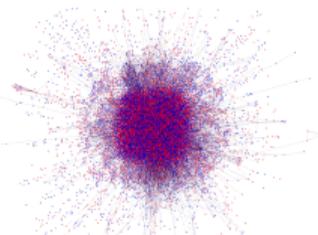
(e)



(f)



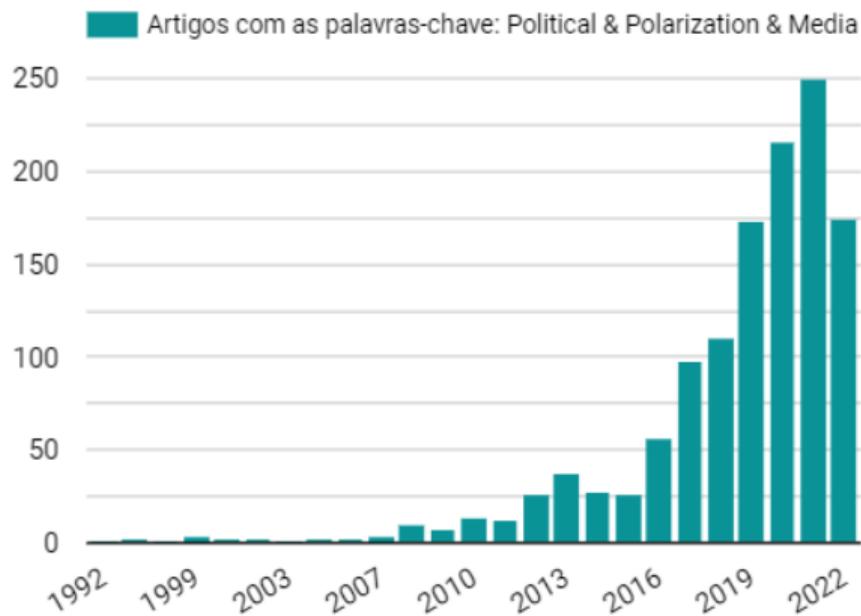
(g)



(h)

Introdução

Perspectiva histórica



Introdução

Perspectiva histórica

O estudo da polarização atrai pesquisadores por diversos motivos.

- ▶ Pelo **valor agregado**.
 - ▶ Exemplo: impacto social na identificação de polarização, no combate a desinformação e informação falsa, na proposta de técnicas para redução da polarização.
- ▶ A temática é inerentemente **multidisciplinar**.
- ▶ Interseção importante entre a academia e a **indústria** que busca por soluções cientes da polarização.

Introdução

Perspectiva histórica

Do ponto de vista da computação, são diversas as áreas de pesquisa:

- ▶ Coleta e processamento de bases de dados;
- ▶ Mineração de opiniões e análise de sentimentos;
- ▶ Desenvolvimento de modelos estatísticos, de teoria dos grafos, de inteligência artificial;
- ▶ Visualização e análise de resultados.

Introdução

Viés vs Polarização

- ▶ A **polarização** pode ser caracterizada pela divisão das entidades de um grupo em duas partições com posições conflitantes.
- ▶ Antes de analisarmos a polarização de uma população, precisamos definir a posição ou o *viés* de cada uma das entidades que a compõe.
- ▶ **Viés** é definido como posicionamento, ou apoio, de um usuário ou declaração com relação a um tópico específico.
- ▶ Por exemplo, ao escolhermos o tópico “legalização das drogas”, teremos um conjunto de indivíduos com viés pró-legalização e outro com viés contrário à legalização.

Introdução

Viés vs Polarização

- ▶ De maneira formal, viés (V) de um indivíduo (X) com relação a um tópico (T).

$$V(X | T) = \{\text{Negativo, Positivo, Neutro}\} \quad (1)$$

- ▶ Posicionamento binário, discreto de múltiplos níveis, ou numérico contínuo, por exemplo, um número real no intervalo entre $[-1; 1]$.

Introdução

Viés vs Polarização

- ▶ O viés pode ser calculado para um indivíduo ou conteúdo.
- ▶ Por exemplo, poderemos encontrar um texto em uma página de notícias com viés claramente favorável ao tema de liberação de drogas, em contraponto a uma notícia de outro jornal que aborda o tema com viés negativo.
- ▶ Neste curso abordaremos somente viés baseado em conteúdos textuais, porém o mesmo se aplica para outros tipos de mídia, como imagens ou vídeos, por exemplo.

Introdução

Uma taxonomia para o estudo da polarização



Introdução

Uma taxonomia para o estudo da polarização

- ▶ Viés:
 - ▶ Questionários (*surveys*);
 - ▶ Extração de *hashtags*;
 - ▶ Análise de sentimentos;
 - ▶ Detecção de posicionamentos.
- ▶ Polarização
 - ▶ Estatísticas;
 - ▶ Teoria dos grafos;
 - ▶ Aprendizado de máquina;
 - ▶ Mineração de informação.

Introdução

Metodologia: um mapa do Curso



Introdução

Dados sobre Polarização

- ▶ Dados correspondem a um tópico de conversação.
 - ▶ Um conjunto de palavras-chave, uma comunidade, etc.
- ▶ Entidades: atores e suas interações.
 - ▶ Conjunto de usuários, postagens, comentários, avaliações, etc.
- ▶ Exemplo:
 - ▶ Tópico: palavra-chave “#ukraine”
 - ▶ Entidades: *tweets* que contêm a *hashtag* ou palavras relacionadas, como “#kyiv” e “#stoprussianaggression”.

Introdução

Viés de entidades dos dados

- ▶ O objetivo aqui é particionar o conjunto de entidades em três conjuntos disjuntos:
 1. O conjunto de entidades que suportam o tópico em questão;
 2. aqueles que são contrários ao tópico em questão e
 3. aqueles que são neutros ao tópico em questão.
- ▶ Em termos didáticos, o resultado desta etapa responde a seguinte questão:

“Assumindo que as entidades se dividem em dois conjuntos opostos de acordo com seu posicionamento com relação ao tópico, quais são esses conjuntos?”

Introdução

Polarização de grupo

- ▶ Já temos os dados pré-processados com o viés das entidades.
- ▶ Intuitivamente, a polarização de um tópico expressa o quão bem separadas as duas partições estão.
- ▶ Isto é, o quanto aqueles dois grupos divergem entre si.
- ▶ Apresentaremos diversas métricas para capturar a polarização.

Introdução

Análise de polarização

- ▶ Análise dos dados e extração de conhecimentos a partir das etapas anteriores.
- ▶ Os tipos de análises e conclusões podem variar bastante a depender do tópico a ser analisado e do objetivo almejado.
- ▶ Alguns temas que podem ser explorados:
 - ▶ Impacto das redes sociais na polarização encontrada na sociedade;
 - ▶ Movimento da polarização ao longo do tempo;
 - ▶ Dentre outros tantos encontrados na literatura.

The background features a diagonal split between a teal upper-left section and a light gray lower-right section, with a white central area where the text is located.

Dados no contexto de polarização

Outline



2. Dados no contexto de polarização

2.1 A gênese dos dados sobre polarização

2.2 Desafios a partir dos dados

2.3 Aquisição dos dados

2.4 Pré-processamento

2.5 Armazenamento

2.6 Extração de conhecimento sobre Polarização

Dados no contexto de polarização

A gênese dos dados sobre polarização

- ▶ A gênese dos dados vem da expressão de **pensamentos** das pessoas em diversas mídias.
- ▶ As **mídias sociais** estão entre os principais veículos onde pessoas se expressam e propagam informações.
- ▶ Esses pensamentos são expressos na forma de postagens, comentários, *likes*, *dislikes*, amizades, etc.
- ▶ Consideraremos as mídias sociais como “agregadores de dados de polarização”, muito embora essa não seja sua finalidade.

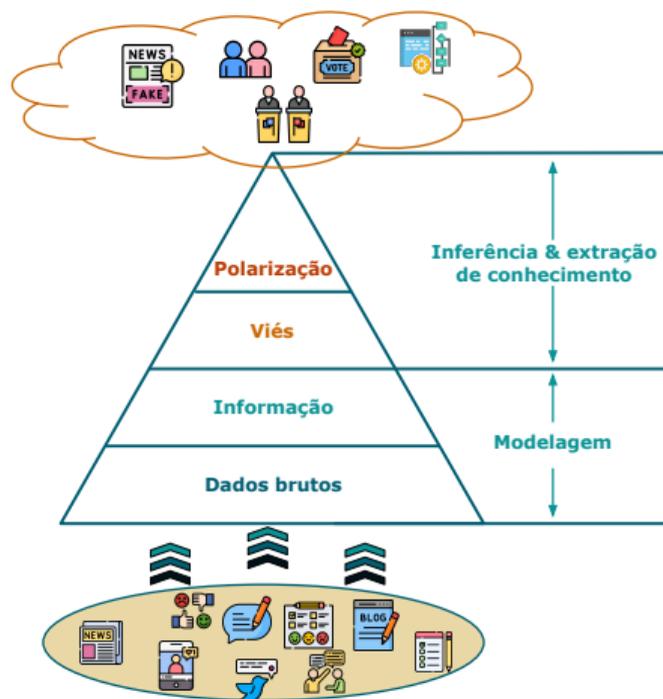
Dados no contexto de polarização

Desafios a partir dos dados

- ▶ Principal problema: **extração de conhecimento** sobre polarização a partir de dados.
- ▶ Modelar e analisar dados a fim de operacionalizar a inferência sobre posicionamentos das entidades.
- ▶ Por exemplo, em [Küçük and Can, 2020], os autores apresentam diferentes abordagens computacionais para indicar posicionamentos de *postagens individuais* em redes sociais ou textos comuns (ex.: blogs).
- ▶ Já em [Ferreira et al., 2021], os autores estão preocupados com o posicionamento de *grupos de indivíduos* que interagem com um tópico de conversação em redes sociais.

Dados no contexto de polarização

Desafios a partir dos dados



Dados no contexto de polarização

Desafios a partir dos dados

1. Modelagem;

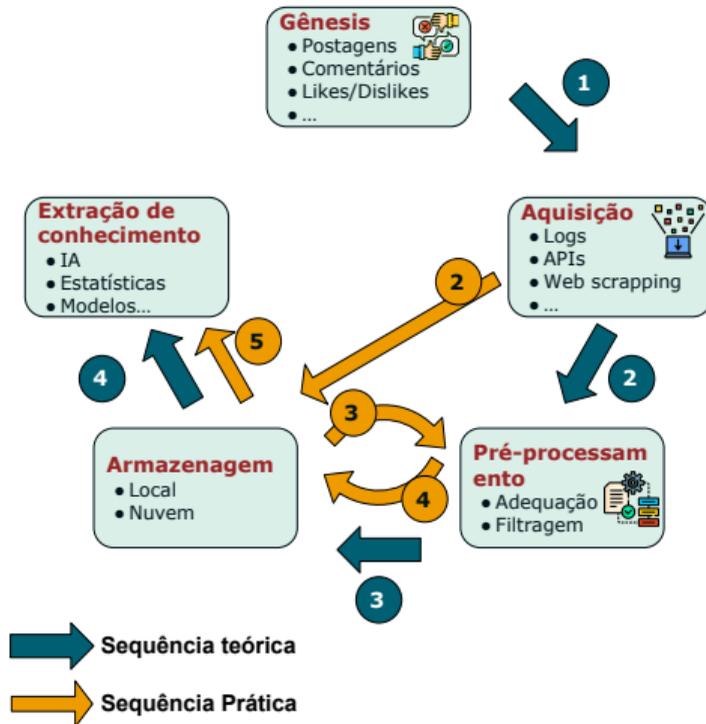
- ▶ O objetivo é adicionar algum nível de semântica e padronização aos dados brutos
- ▶ Dados não estruturados e heterogêneos.
- ▶ Técnicas de pré-processamento, filtragem ou fusão de dados.

2. Análise e inferência sobre polarização.

- ▶ Aplicar interpretações visando delinear o contexto a partir das informações.
- ▶ Extrair conhecimento acerca da situação daquele dado que.

Dados no contexto de polarização

Aquisição dos dados



Dados no contexto de polarização

Aquisição dos dados

- ▶ Nem sempre a sequência teórica é respeitada.
- ▶ Na prática, seguimos a sequência amarela.
- ▶ Obtém-se os dados brutos e ciclos de pré-processamento e armazenamentos são realizados (em uma espécie de tentativas e erros/acertos)
- ▶ Então os dados são encaminhados à algoritmos e técnicas de extração de conhecimento.

Dados no contexto de polarização

Aquisição dos dados

- ▶ Dados sobre polarização surgem de formas distintas
- ▶ As formas de aquisição também serão distintas.
- ▶ Por exemplo, dados de páginas web podem ser coletados através de *Web Scraping*, enquanto *posts* em redes sociais podem ser adquiridos através de APIs.
- ▶ Além disso, podemos adquirir dados por meio de formulários, observação, entrevistas, dentre outras.

Dados no contexto de polarização

Aquisição dos dados: *Web Scrapping*

O processo de *Web Scrapping*, em português “raspagem da web”, é uma técnica utilizada para extrair dados diretamente de páginas da web.

1. Scrapy: Originalmente projetado para *Web Scrapping*, embora possa ser utilizado para outras finalidades como, por exemplo, coletar dados de APIs.
2. BeautifulSoup: Biblioteca Python que fornece meios para analisar e “caminhar” (*tree traversal*) em páginas web.
3. Selenium: Primariamente projetado para testar navegadores.

Dados no contexto de polarização

Aquisição dos dados: APIs

- ▶ É comum que grandes portais web e, especialmente, redes sociais exponham **APIs** que permitam a integração e expansão dos seus conteúdos e serviços.
- ▶ Essas APIs geralmente seguem a arquitetura e orientações do padrão Web REST, que define uma interface para integração e acesso a recursos.
- ▶ APIs REST possuem limitação na quantidade de requisições possíveis.

Dados no contexto de polarização

Aquisição dos dados: APIs

- ▶ Redes sociais como Twitter, Facebook, Instagram, entre outras, expõem APIs para acesso a seus recursos.
- ▶ Razão pela qual os dados dessas redes sociais têm sido amplamente utilizados para o estudo sobre polarização.
- ▶ Através das APIs dessas redes sociais é possível coletar *posts*, *likes/dislikes*, rede de amizade, tópicos/assuntos em alta, *hashtags*, etc.

Dados no contexto de polarização

Pré-processamento

- ▶ As tarefas de modelagem e filtragem dos dados são essenciais para transformar dados brutos em informações úteis.
- ▶ Dados nem sempre possuem uma organização lógica e hierárquica, o que dificulta sua manipulação.
- ▶ **Modelagem dos dados** lida com o problema de representar informação (organizada) a partir de dados brutos.
- ▶ **Filtragem de dados** lida com a eliminação de dados indesejáveis, erros de inserções, informações irrelevantes, etc.

Dados no contexto de polarização

Pré-processamento: Modelagem dos dados

- ▶ **Chave-valor** (*key-value*): Pares de informação relacionadas. Por exemplo, código do usuário e lista de posts.
- ▶ **Markup Schema**: Conjunto de 'tipos' com conjunto de propriedades definidas (*Book: bookEdition, bookFormat, illustrator*).
- ▶ **Grafos**: Construídas a partir de relações entre pessoas, interações com postagens, comentários, *likes*, dentre outras.
- ▶ Outras abordagens também são possíveis, como representação usando ontologias, baseadas em objetos, baseadas em lógicas, etc.

Dados no contexto de polarização

Pré-processamento: Filtragem dos dados

- ▶ **Granularidade dos dados:** medida no nível de detalhes do dados coletados (ex.: frequência em que os dados são coletados).
- ▶ **Dados vagos:** Ocorrem quando os atributos não estão bem definidos (ex.: posicionamento de um tweet).
- ▶ **Outliers** (anomalias): Dados que tipicamente diferem significativamente de outras observações (ex.: bots, spammers).
- ▶ **Dados incompletos:** Observações que possuem um ou mais atributos sem valor (ex.: idade dos usuários).

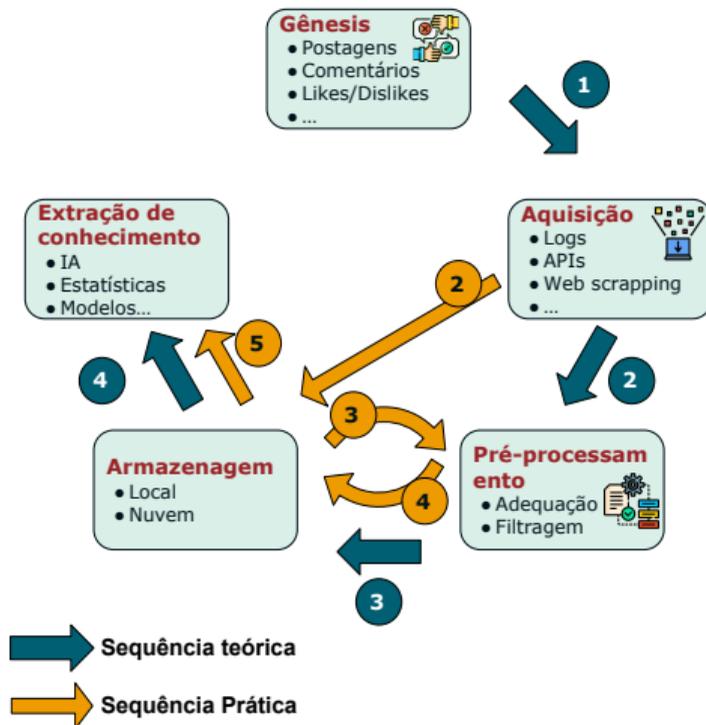
Dados no contexto de polarização

Armazenamento

- ▶ **Armazenamento** aparece na literatura de duas formas principais: **armazenamento local** ou em plataformas na **nuvem**.
- ▶ Em ambas as abordagens é desejável que os dados sejam adequados a um modelo, facilitando a consulta.
- ▶ Entretanto, o que acontece com frequência, é um armazenamento dos dados brutos utilizando um modelo mais simples e,
- ▶ Eventualmente, os dados passam por ciclos de pré-processamento e armazenagem (como mostrado na figura anterior).
- ▶ A escolha de um **SGBD** e esquemas para armazenagem de dados são etapas importantes no processo, porém estão fora do escopo deste curso.

Dados no contexto de polarização

Extração de conhecimento sobre Polarização



Dados no contexto de polarização

Extração de conhecimento sobre Polarização

- ▶ O último nível da hierarquia representa a polarização em si.
- ▶ A etapa de análise e inferência busca definir a polarização e os vieses a partir dos dados coletados.
- ▶ No restante deste curso serão discutidas técnicas para caracterizar viés e polarização.

The background consists of two large, overlapping geometric shapes. A teal-colored shape is in the upper-left corner, and a light gray shape is in the lower-left corner. The rest of the background is white. The text is centered in the white area.

Viés de entidades dos dados



3. Viés de entidades dos dados

3.1 O problema de detecção de posicionamento

3.2 Abordagens de detecção de posicionamento

- ▶ Os dados disponíveis em redes sociais é terreno fértil para diversas análises.
- ▶ Identificar se uma determinada notícia ou postagem contém **conteúdo falso**, por exemplo, é uma tarefa de extrema relevância nos tempos atuais.
- ▶ Verificar se um comentário específico é **ofensivo**, ou
- ▶ Identificar se um determinado usuário é um **robô** (*bot*) ou uma pessoa real.

- ▶ O estudo do viés de dados envolve um vasto conjunto de problemas relacionados.
 - ▶ Análise de sentimentos,
 - ▶ extração de opiniões,
 - ▶ detecção de ironia,
 - ▶ classificação de notícias falsas,
 - ▶ mineração de argumentos, dentre outros.
- ▶ Para este curso, focaremos no viés representado pelo **posicionamento** ou estância de um comentário ou usuário em específico.

Viés de entidades dos dados

O problema de detecção de posicionamento

- ▶ **Definição:** Tarefa de identificar o posicionamento (estância, orientação, apoio) de uma entidade com relação a um ou mais alvos (proposições, temas, tópicos).
- ▶ **Tópicos** podem representar aspectos ideológicos, decisões políticas, organizações, ou indivíduos específicos.
- ▶ **Posicionamento** identifica se a entidade é “*A Favor*” ou “*Contra*” um tópico.
- ▶ Opções como “*Nenhum*” e/ou “*Neutro*” representam conteúdos que não se posicionam claramente a favor ou contra.

Viés de entidades dos dados

[Rathje 2021] Out-group animosity drives engagement on social media

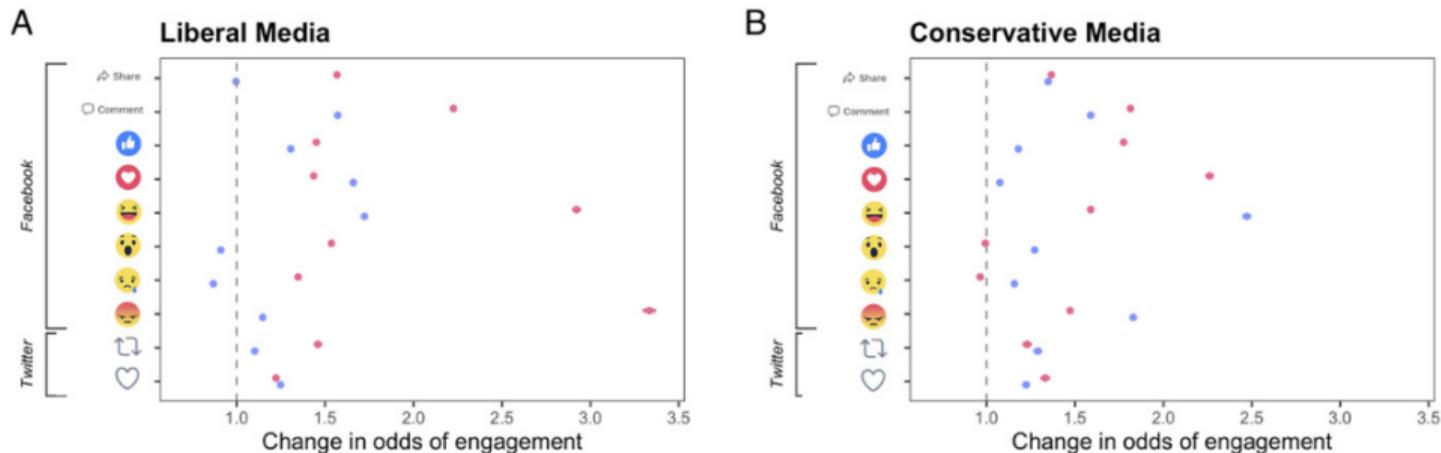


Figura 2: Palavras “In-group” e “Out-group” preveem engajamento em redes liberais e conservadoras.

Viés de entidades dos dados

O problema de detecção de posicionamento: Tipo de entidade

- ▶ Posicionamento de uma **declaração** tem objetivo de identificar a orientação em um determinado texto ou conteúdo.
 - ▶ Tarefa de processamento de linguagem natural.
 - ▶ Ou processamento de imagem/vídeo/ec.
- ▶ Posicionamento de um **usuário** tem objetivo de descobrir se cada usuário é a favor ou contra um tópico.
 - ▶ Considera-se todo o seu comportamento na plataforma digital;
 - ▶ Textos e postagens produzidas por aquele usuário;
 - ▶ Informações de seu perfil (idade, gênero, etc.), comunidades e conteúdos com o qual interage; e
 - ▶ Relações com outros usuários em sua rede social.

Viés de entidades dos dados

O problema de detecção de posicionamento: Outros problemas relacionados

- ▶ **Análise de sentimento.** Determinar a polaridade das emoções expressas em um determinado conteúdo.
- ▶ Exemplo: “estou nervoso!”, indica uma polaridade de sentimento clara, não um posicionamento com relação a um tema .
- ▶ Exemplo: “estou feliz que esse filme não fez sucesso” indica um sentimento positivo, representa um posicionamento negativo com relação ao filme em questão.
- ▶ O que não quer dizer que não haja relação entre sentimento e posicionamento.

Viés de entidades dos dados

O problema de detecção de posicionamento: Outros problemas relacionados

[ALDayel 2021] Stance detection on social media: State of the art and trends

#	Tweet	Target	Sent.	Stance
1	It is so much fun having younger friends who are expecting babies. #beentheredonethat #chooselife	Legalization of abortion	+	-
2	Life is sacred on all levels. Abortion does not compute with my philosophy. (Red on #OITNB)	Legalization of abortion	0	-
3	The biggest terror threat in the World is climate change #drought #floods	Climate change is a real concern	-	+
4	I am sad that Hillary lost this presidential race	Hillary Clinton	-	+

Figura 3: Stance Detection vs Sentiment Analysis

Viés de entidades dos dados

O problema de detecção de posicionamento: Outros problemas relacionados

Outros problemas relacionados, incluem:

- ▶ **Reconhecimento de emoções.** Identifica emoções presentes em um conteúdo dentro um conjunto de classes de emoções;
- ▶ **Detecção de controvérsia.** Identifica tópicos controversos em uma discussão;
- ▶ **Previsão de posicionamento.** Estima como usuários (ou grupos de usuários) se posicionariam com relação a temas dos quais esse posicionamento não foi observado.

Viés de entidades dos dados

Abordagens de detecção de posicionamento: Atributos

- ▶ Para abordar a detecção de posicionamento, há uma multiplicidade de atributos que podem ser utilizados.
- ▶ Atributos extraídos diretamente do conteúdo ou características da rede social em si.
- ▶ **Conteúdo textual.**
 - ▶ Modelagens dos termos em *bag-of-words* ou n-gramas, pontuação e tamanho do texto .
 - ▶ Polaridade de sentimento do conteúdo.
 - ▶ Representações do conteúdo em forma de tópicos latentes.
 - ▶ Análise de vocabulário entre grupos com posicionamentos opostos.

Viés de entidades dos dados

Abordagens de detecção de posicionamento: Atributos

Características e atributos retirados a partir da **rede**.

- ▶ Alguns trabalhos, por exemplo, modelam representações com base no texto do conteúdo em conjunto com características das interações do usuário na rede
- ▶ Atributos indicando as conexões e interações entre usuários na rede,
- ▶ Outras características retiradas das plataformas como hashtags, re-tweets, URLs, e menções a outros usuários.

Viés de entidades dos dados

Abordagens de detecção de posicionamento: Algoritmos

- ▶ **Métodos supervisionados** são frequentemente utilizados para identificar posicionamento.
- ▶ Bases de dados são **anotadas por especialistas** de acordo com rótulos como "A Favor", "Contra" e "Nenhum".
- ▶ A partir desses dados rotulados, algoritmos são treinados para aprender os padrões que indicam cada posicionamento.
 - ▶ [Elfardy and Diab, 2016] Uso de atributos léxicos e semânticos com uma SVM (*Support-Vector Machine*)
 - ▶ [Li and Caragea, 2019] Métodos de aprendizado profundo com uma arquitetura GRU (*Gated Recurrent Unit*).

Viés de entidades dos dados

Abordagens de detecção de posicionamento: Algoritmos

- ▶ Dados rotulados são custosos.
- ▶ Abordagens **semi-supervisionadas** e **não-supervisionadas** também são propostas.
 - ▶ [Ferreira and Vlachos, 2019] Usa técnicas de transferência de aprendizado para reutilizar conhecimento como ponto de partida.
 - ▶ [Darwish et al., 2020] Técnica de agrupamento sobre tweets não-rotulados de diferentes tópicos como ponto inicial para anotação dos dados.
 - ▶ [Rashed et al., 2020] Representações distribuídas de tweets com agrupamento hierárquico para análise de polarização política.

Viés de entidades dos dados

Abordagens de detecção de posicionamento: Algoritmos

As possibilidades de algoritmos para detecção de posicionamentos são muitas.

- ▶ SVMs.
- ▶ Regressão logística.
- ▶ Técnicas de aprendizado profundo.
- ▶ Redes LSTMs (*Long Short-Term Memory*),
- ▶ Redes CNNs (*Convolutional Neural Network*)
- ▶ Modelos baseados BERT.

Métricas de polarização



4. Métricas de polarização

- 4.1 Utilizando análise de sentimentos para medir polarização [Al-Ayyoub et al., 2018]
- 4.2 Momento do dipolo elétrico [Morales et al., 2015]
- 4.3 Conectividade na fronteira entre grupos antagônicos [Guerra et al., 2013]
- 4.4 Centralidade do corte da rede [Garimella et al., 2018]
- 4.5 Random Walk Controversy (RWC) [Garimella et al., 2018]

Métricas de polarização

Viés vs Polarização

- ▶ Como vimos, viés está ligado ao posicionamento de uma entidade.
- ▶ **Polarização** é uma medida de um grupo de usuários (ou conteúdo)
- ▶ Como uma população está distribuída em torno de um tópico?
- ▶ **Exemplo:** A reitora de uma universidade está passando por uma crise de popularidade.
 - ▶ 40% reprovam seu mandato
 - ▶ 45% aprovam seu mandato
 - ▶ 15% estão neutros
- ▶ Podemos dizer que a universidade está polarizada?

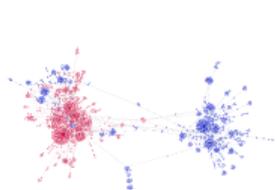
Métricas de polarização

Representação por grafo

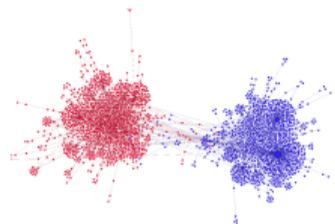
- ▶ Outra forma de visualizar a polarização é por meio do **grafo da rede**.
- ▶ Técnica bastante utilizada no contexto de redes sociais.
- ▶ Cada usuário (ou conteúdo) é um vértice do grafo.
- ▶ Arestas são formadas por suas conexões, como amizade e compartilhamento.

Métricas de polarização

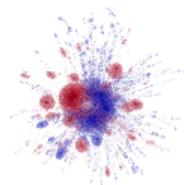
Exemplos de redes polarizadas



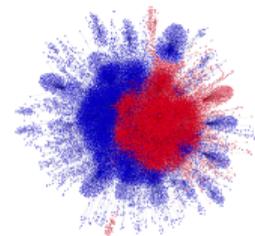
(a)



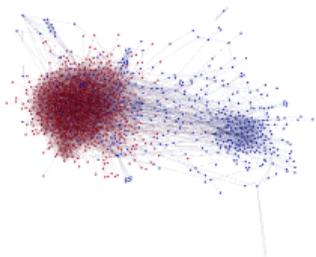
(b)



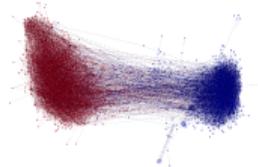
(c)



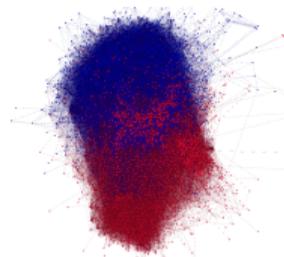
(d)



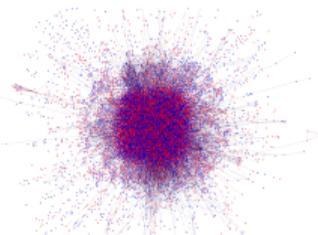
(e)



(f)



(g)



(h)

Métricas de polarização

Qual métrica utilizar?

- ▶ Não existe um consenso na literatura.
 - ▶ Estudos de caso; base de dados específicas; conhecimento de domínio.
1. Métricas independentes da rede (**conteúdo**).
 - ▶ *Vantagem*: Estrutura da rede é desconhecida.
 - ▶ *Exemplo*: Caixa de comentários de uma famosa página de notícias.
 2. Métricas calculadas por meio de sua **rede** de comunicação.
 - ▶ *Vantagem*: Usufrui de informações presentes na rede, como conexões de amizade ou de endosso (retweet)
 - ▶ *Exemplo*: Procura estruturas na rede, como clusters de usuários, fronteiras entre os grupos, entre outros.

Métricas de polarização

Compilado de trabalhos relacionados - Conteúdo

Trabalho	Conteúdo	Rede	Ident.	Quant.	Ferramenta
[Al-Ayyoub et al., 2018]	✓		✓	✓	Matemática
[Akhtar et al., 2019]	✓		✓	✓	Estatística
[Babaei et al., 2018]	✓		✓	✓	Matemática
[Belcastro et al., 2020]	✓		✓		Matemática
[Choi et al., 2010]	✓		✓		Estatística
[Dori-Hacohen and Allan, 2015]	✓		✓		Anotado
[Jang et al., 2016]	✓		✓		Anotado
[Jang and Allan, 2018]	✓		✓		Estatística
[Klenner et al., 2014]	✓		✓		Aprend. superv.
[Mejova et al., 2014]	✓		✓		Aprend. superv.
[Morales et al., 2015]	✓		✓	✓	Estatística
[Popescu and Pennacchiotti, 2010]	✓		✓		Aprend. superv.
[Ribeiro et al., 2019]	✓		✓		Survey
[Roy and Goldwasser, 2020]	✓		✓		Aprend. superv.
[Tsytsarau et al., 2011]	✓		✓		Estatística
[Vicario et al., 2019]	✓		✓	✓	Aprend. superv.
[Waller and Anderson, 2021]	✓		✓	✓	Baixa dimensão
[Yang et al., 2017]	✓		✓	✓	Baixa dimensão

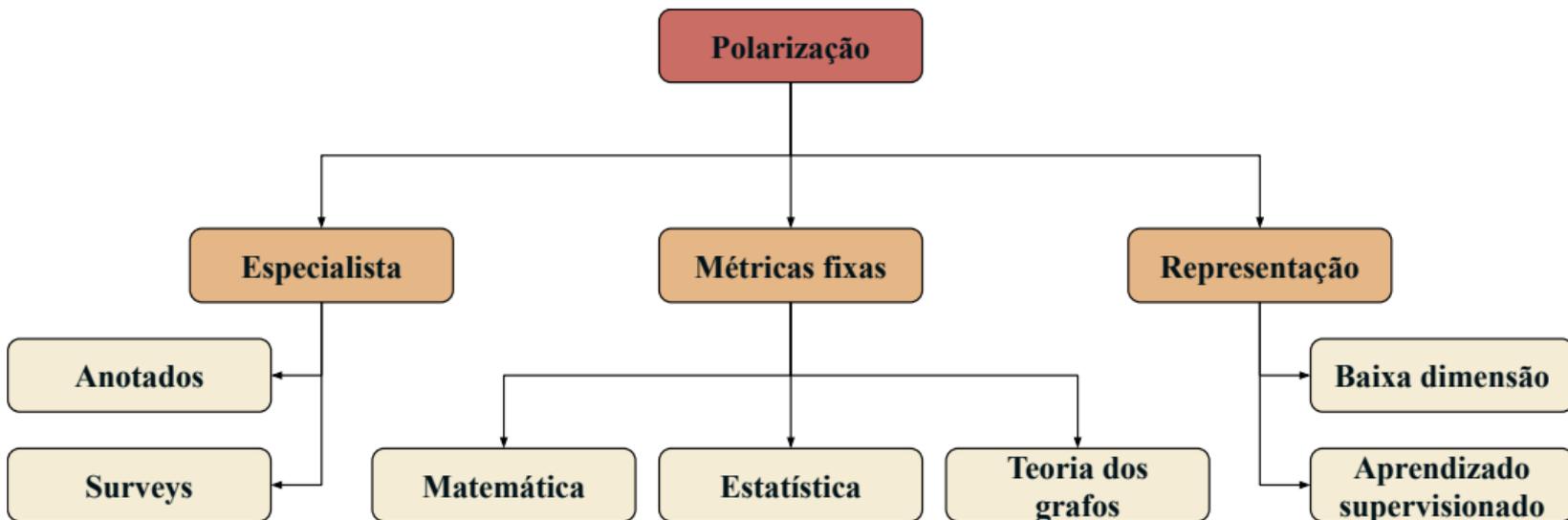
Métricas de polarização

Compilado de trabalhos relacionados - Rede

Trabalho	Conteúdo	Rede	Ident.	Quant.	Ferramenta
[Akoglu, 2014]		✓	✓		Estatística
[Al Amin et al., 2017]		✓	✓		Aprend. superv.
[Coletto et al., 2017]		✓	✓		Teoria dos grafos
[Garimella et al., 2018] (RWC)		✓	✓	✓	Teoria dos grafos
[Garimella et al., 2018] (BCC)		✓	✓	✓	Teoria dos grafos
[Garimella et al., 2018] (EC)		✓	✓	✓	Baixa dimensão
[Garimella et al., 2021]		✓	✓		Teoria dos grafo
[Gillani et al., 2018]		✓	✓		Estatística
[Guerra et al., 2013]		✓	✓	✓	Teoria dos grafos
[Shahrezaye et al., 2019]		✓	✓	✓	Teoria dos grafos
[Tokita et al., 2021]		✓	✓	✓	Teoria dos grafos

Métricas de polarização

Taxonomia das métricas de polarização



Métricas de polarização

Utilizando análise de sentimentos para medir polarização [Al-Ayyoub et al., 2018]

- ▶ A primeira métrica de polarização que vamos abordar foi originalmente concebida para trabalhar com análise de sentimentos de tweets [Al-Ayyoub et al., 2018].
- ▶ Um conjunto de métricas matemáticas simples.
- ▶ Podem ser utilizadas individualmente ou coletivamente para análise de polarização.

Métricas de polarização

Utilizando análise de sentimentos para medir polarização [Al-Ayyoub et al., 2018]

Razão da quantidade de tweets positivos e negativos

Premissa: Em um tópico polarizado um dos grupos utiliza mensagens que demonstram sentimentos de aprovação ao tópico contra mensagens de reprovação.

$$PN = \frac{|\text{tweets positivos}|}{|\text{tweets negativos}|} \quad (2)$$

Razão entre a quantidade de tweets positivos e negativos

A RPN é uma melhoria da PN. Teremos sempre um valor entre 0 e 1.

$$RPN = \frac{\min\{|\text{tweets positivos}|, |\text{tweets negativos}|\}}{\max\{|\text{tweets positivos}|, |\text{tweets negativos}|\}} \quad (3)$$

Métricas de polarização

Utilizando análise de sentimentos para medir polarização [Al-Ayyoub et al., 2018]

Razão entre tweets neutros e tweets positivos e negativos

Premissa: Em um tópico polarizado temos um número pequeno de mensagens neutras em comparação com mensagens com viés claro.

$$NPN = \frac{|\text{tweets neutros}|}{|\text{tweets positivos}| + |\text{tweets negativos}|} \quad (4)$$

Razão entre a soma dos tweets positivos e negativos com a quantidade total de tweets

Premissa: Tópicos polarizados obtêm um maior número de comentários com viés claro.

$$PNT = \frac{|\text{tweets positivos}| + |\text{tweets negativos}|}{|\text{tweets totais}|} \quad (5)$$

Métricas de polarização

Utilizando análise de sentimentos para medir polarização [Al-Ayyoub et al., 2018]

Métrica PN ponderada pela métrica PNT

Essa métrica é uma combinação entre a métrica PN e a métrica PNT.

$$\text{PNPNT} = \text{PN} \times \text{PNT} \quad (6)$$

Razão entre os valores de sentimentos positivos e negativos

A RPNV é uma modificação da RPN e leva em consideração os valores de viés calculados para cada tweet.

$$\text{RPNV} = \frac{\min\{\sum \text{valores positivos}, \sum \text{valores negativos}\}}{\max\{\sum \text{valores positivos}, \sum \text{valores negativos}\}} \quad (7)$$

Métricas de polarização

Utilizando análise de sentimentos para medir polarização [Al-Ayyoub et al., 2018]

- ▶ Métricas calculadas com baixíssimo custo computacional; o que as torna uma ótima opção para serem utilizadas como **features** para um algoritmo de aprendizado supervisionado, por exemplo.
- ▶ O trabalho de [Al-Ayyoub et al., 2018] utiliza esse conjunto de métricas como entrada para um classificador SVM de tópicos de discussão do Twitter.
- ▶ Métricas podem ser facilmente adaptadas para outros cenários, como postagens em uma página de jornal ou votação de políticos na câmara de deputados.
- ▶ (Abordaremos a implementação da métrica na parte prática deste minicurso).

Métricas de polarização

Momento do dipolo elétrico [Morales et al., 2015]

- ▶ Métrica inspirada no momento do dipolo elétrico (*Dipole Moment* – DM)
- ▶ Tem como objetivo capturar o quão dividido encontram-se os membros de uma população.
- ▶ **Premissa:** Uma população é perfeitamente polarizada se ela pode ser dividida em dois grupos de mesmo tamanho e com as opiniões de seus indivíduos concentradas nos extremos.

Métricas de polarização

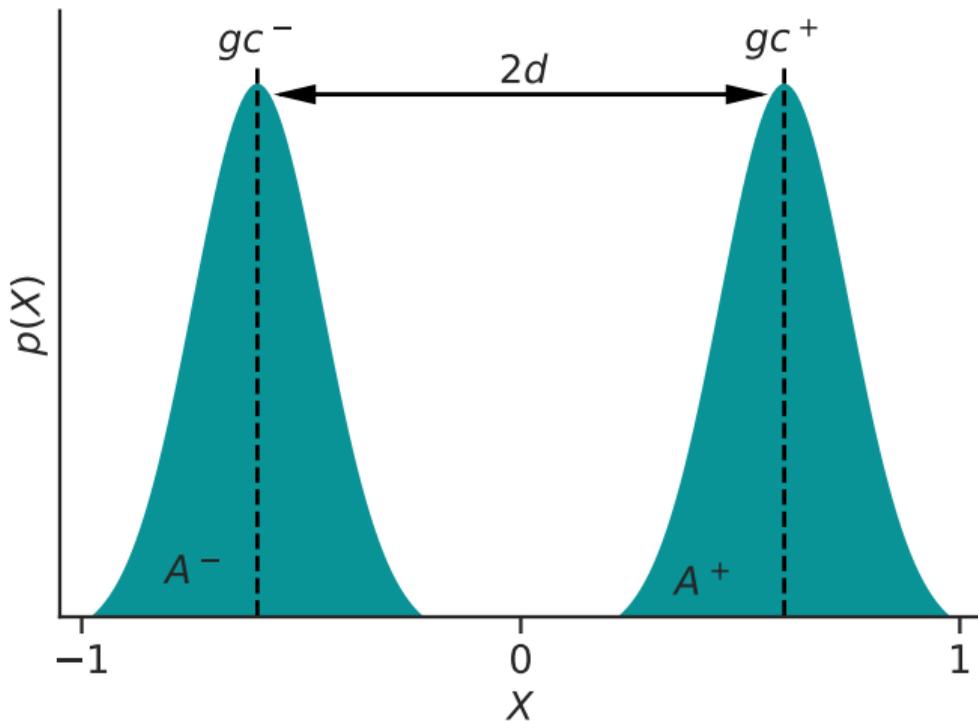
Momento do dipolo elétrico [Morales et al., 2015]

TODO: botar uma figurinha de dois elétrons aqui.

- ▶ Vem do cálculo de polarização de cargas de um sistema eletromagnético.
- ▶ Calcula o grau de separação de cargas positivas e negativas de um sistema.
- ▶ Quando temos somente duas cargas, uma positiva e uma negativa ($-q$ e $+q$), o momento do dipolo elétrico é proporcional à distância entre essas duas cargas.
- ▶ Cenário análogo onde temos duas pessoas de opiniões contrárias entre si.

Métricas de polarização

Momento do dipolo elétrico [Morales et al., 2015]



Métricas de polarização

Momento do dipolo elétrico [Morales et al., 2015]

- ▶ Seja X uma variável aleatória que modela a distribuição do viés de uma população de um tópico e X_i o viés de um usuário i de modo que $-1 \leq X_i \leq +1$.
- ▶ Temos $p(X)$ como uma função de densidade da opinião dos usuários.

Métricas de polarização

Momento do dipolo elétrico [Morales et al., 2015]

- ▶ Primeiramente iremos calcular o tamanho das populações associadas a cada viés (negativo e positivo).
- ▶ Seja A^- a população com viés negativo ($X < 0$), calculamos seu tamanho como a área sob a curva da função de densidade $p(X)$ no intervalo $[-1, 0]$
- ▶ De maneira análoga, calculamos A^+ como a área sob a curva de $p(X)$ no intervalo $[0, +1]$.

$$A^- = \int_{-1}^0 p(X) dX = P(X < 0), \quad (8)$$

$$A^+ = \int_0^1 p(X) dX = P(X > 0) \quad (9)$$

Métricas de polarização

Momento do dipolo elétrico [Morales et al., 2015]

- ▶ De posse do tamanho dos grupos, calculamos a diferença de tamanho entre suas populações.
- ▶ Esta fórmula dá como resultado $\Delta A = 0$ quando a população está perfeitamente dividida em dois grupos de tamanhos iguais ($A^- = A^+$).
- ▶ Por outro lado, $\Delta A = 1$ quando todos os elementos da população concordam entre si ($A^- = 1$ ou $A^+ = 1$).

$$\Delta A = |A^+ - A^-| = |P(X > 0) - P(X < 0)| \quad (10)$$

Métricas de polarização

Momento do dipolo elétrico [Morales et al., 2015]

- ▶ Em seguida, calculamos o quão distante estão as opiniões de ambos os grupos.
- ▶ Calculamos o centro de gravidade dos vieses negativos (gc^-) e positivos (gc^+).

$$gc^- = \frac{\int_{-1}^0 p(X)XdX}{\int_{-1}^0 p(X)dX}, \quad gc^+ = \frac{\int_0^1 p(X)XdX}{\int_0^1 p(X)dX} \quad (11)$$

Métricas de polarização

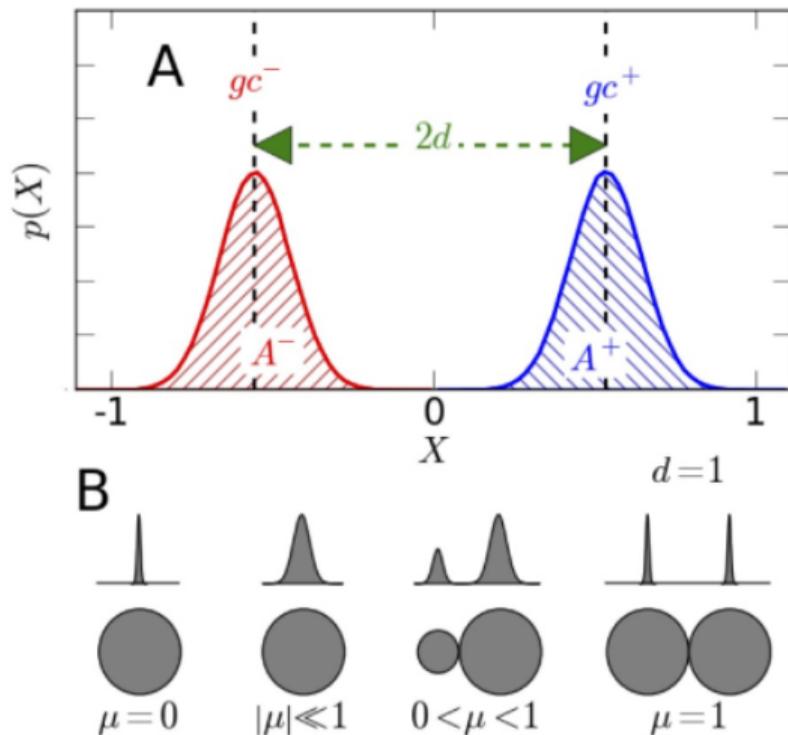
Momento do dipolo elétrico [Morales et al., 2015]

- ▶ E então calculamos a distância entre as opiniões dos grupos como a diferença absoluta dos centros gravitacionais gc^- e gc^+ .
- ▶ Esta fórmula dá como resultado $d = 0$ quando ambos os grupos concordam integralmente entre si.
- ▶ Por outro lado, teremos $d = 1$ quando ambos os grupos discordam entre si e suas opiniões se concentram nos extremos.

$$d = \frac{|gc^+ - gc^-|}{|X_{\max} - X_{\min}|} = \frac{|gc^+ - gc^-|}{2} \quad (12)$$

Métricas de polarização

Momento do dipolo elétrico [Morales et al., 2015]



Métricas de polarização

Momento do dipolo elétrico [Morales et al., 2015]

- ▶ Por fim, calculamos o índice de polarização DM à partir de ΔA e d .
- ▶ DM é inversamente proporcional a ΔA e diretamente proporcional a d .
- ▶ O resultado de DM se encontra no intervalo $[0, 1]$.
- ▶ Teremos polarização máxima ($DM = 1$) quando a população estiver perfeitamente dividida ($\Delta A = 0$) e as opiniões discordantes estiverem nos extremos ($d = 1$).
- ▶ Teremos polarização mínima ($DM = 0$) quando a opinião da população como um todo estiver concentrada em um único ponto, ou seja, quando não há discordância.

$$DM = (1 - \Delta A)d \quad (13)$$

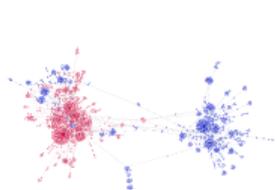
Métricas de polarização

Momento do dipolo elétrico [Morales et al., 2015]

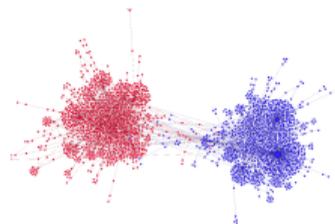
- ▶ A métrica do momento do dipolo elétrico é útil quando o viés das entidades puderem ser quantificados (valores entre -1 e $+1$) e o grafo da rede não for conhecido.
- ▶ A métrica é de fácil implementação e baixo custo computacional.
- ▶ A função de densidade de opinião pode ser uma ótima visualização da polarização da população.
- ▶ (Abordaremos a implementação da métrica na parte prática deste minicurso)

Métricas de polarização

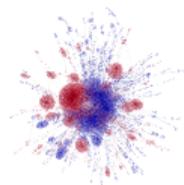
Exemplos de redes polarizadas



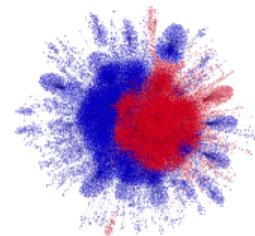
(a)



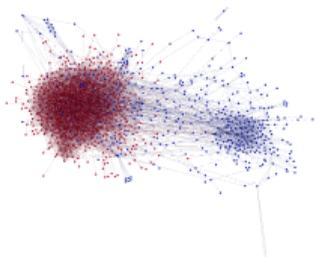
(b)



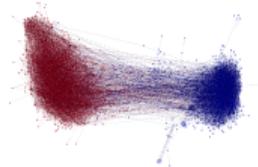
(c)



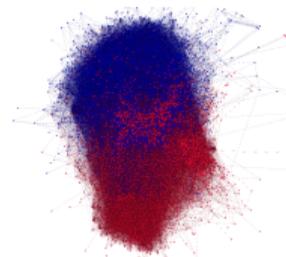
(d)



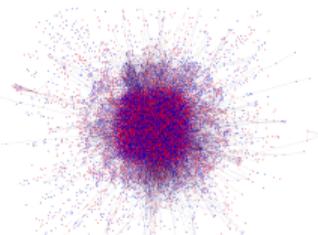
(e)



(f)



(g)



(h)

Métricas de polarização

Conectividade na fronteira entre grupos antagônicos [Guerra et al., 2013]

- ▶ A estrutura de uma rede social é afetada pelo comportamento dos usuários.
- ▶ Padrões de comportamento como **homofilia**, alteram a probabilidade que dois usuários se conectem.
- ▶ Em uma rede polarizada é esperado encontrar padrões que representem a divisão de uma população.
- ▶ Um desses padrões é o **antagonismo**, isto é, conjuntos de usuários que não apresentarão laços entre si.

Métricas de polarização

Conectividade na fronteira entre grupos antagônicos [Guerra et al., 2013]

- ▶ Partindo dessa premissa, foram elaboradas métricas de polarização que utilizam informações da rede de modo a extrair informações topológicas do grafo da rede.
- ▶ O primeiro exemplo desse tipo de métrica de polarização é a métrica de **Conectividade de Fronteira** (*Boundary Connectivity* – BC).
- ▶ Essa métrica foca sua análise nos nós que possuem alguma interação com nós do suposto grupo oposto, que aqui denominamos de **nós de fronteira**.

Métricas de polarização

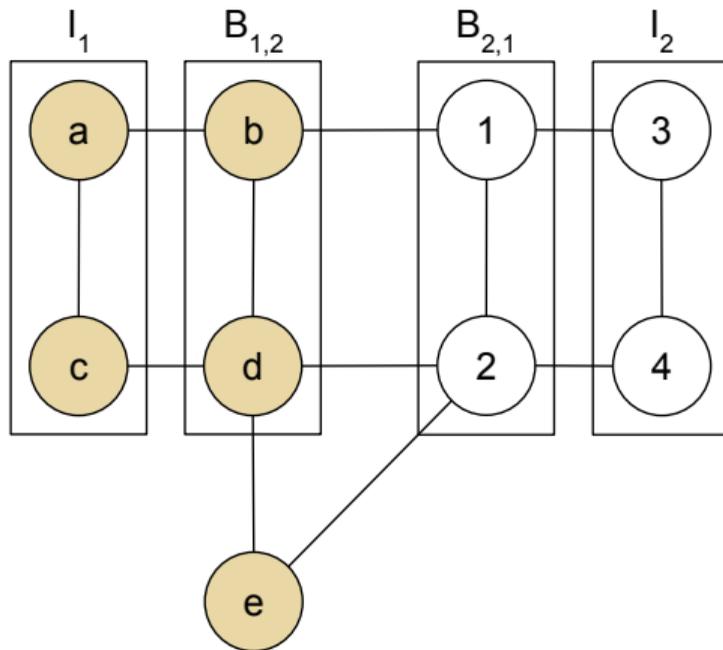
Conectividade na fronteira entre grupos antagônicos [Guerra et al., 2013]

Seja $G = G_i \cup G_j$ o grafo que representa uma rede dividida em dois conjuntos G_i e G_j ($G_i \cap G_j = \emptyset$). Definimos fronteira de comunidade do grupo G_i , como o conjunto de nós $B_{i,j}$ que satisfaz duas condições:

1. $v \in G_i$ possui ao menos uma conexão com um nó do grupo G_j ;
2. $v \in G_i$ possui ao menos uma conexão com um nó do grupo G_i que não possua conexão com os nós de G_j .

Métricas de polarização

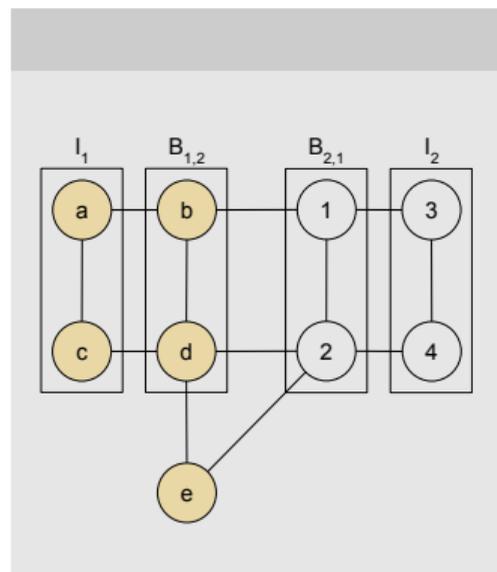
Conectividade na fronteira entre grupos antagônicos [Guerra et al., 2013]



Métricas de polarização

Conectividade na fronteira entre grupos antagônicos [Guerra et al., 2013]

- ▶ Atenemos ao nó b.
 1. (b,1) é uma aresta externa (liga nós de fronteiras opostas).
 2. (b,a) é uma aresta interna (liga nós de fronteiras opostas).
 3. (b,d) não é nem uma aresta interna nem externa.
- ▶ Olhando para as conexões do nó b, ele não nos parece apresentar nenhum tipo de antagonismo.
- ▶ Essa mesma análise pode ser extrapolada para os demais nós da fronteira (nós d, 1 e 2).
- ▶ Logo, o exemplo não possui polarização.



Métricas de polarização

Conectividade na fronteira entre grupos antagônicos [Guerra et al., 2013]

- ▶ A equação BC leva em consideração as escolhas que nós de $B_{i,j}$ fazem ao se conectar com nós de I_i ou $B_{j,i}$.
- ▶ $v \in B$: nó v pertencente à fronteira $B = B_{i,j} \cup B_{j,i}$.
- ▶ $d_i(v)$: número de arestas internas do nó v .
- ▶ $d_b(v)$: número de arestas externas do nó v .

$$BC = \frac{1}{|B|} \sum_{v \in B} \left[\frac{d_i(v)}{d_b(v) + d_i(v)} - 0.5 \right] \quad (14)$$

Métricas de polarização

Conectividade na fronteira entre grupos antagônicos [Guerra et al., 2013]

- ▶ A métrica BC poderá apresentar valores entre $-1/2$ e $+1/2$.
- ▶ Um valor menor que 0 indica não somente a falta de polarização mas também que um nó da fronteira é mais provável de se conectar com nós do grupo oposto.
- ▶ No exemplo anterior temos $BC = 0$ uma vez que os nós da fronteira possuem o mesmo número de arestas com nós internos e com nós do grupo oposto.
- ▶ A métrica BC tem a vantagem de utilizar informações de interação e comunicação entre os usuários.
- ▶ Seu foco na fronteira mensura a interação entre os antagônicos.
- ▶ Traz uma visão sobre o nível de troca de informações entre os diferentes grupos.

Métricas de polarização

Centralidade do corte da rede [Garimella et al., 2018]

- ▶ A métrica de centralidade da rede vai analisar o conjunto de arestas presentes no corte que se forma ao particionar um grafo em dois grupos opostos X e Y .
- ▶ Para isso vamos utilizar o conceito de **centralidade de arestas** (*edge betweenness*).

Métricas de polarização

Centralidade do corte da rede [Garimella et al., 2018]

- ▶ A centralidade de uma aresta ‘e’ é calculada pelo que chamamos de “intermediação” (*shortest-path betweenness centrality* – $bc(e)$).
- ▶ $\sigma_{s,t}$: menor caminho entre os nós s e t .
- ▶ $\sigma_{s,t}(e)$: menor caminho entre s e t que necessariamente passa pela aresta e .

$$bc(e) = \sum_{s \neq t \in V} \frac{\sigma_{s,t}(e)}{\sigma_{s,t}}, \quad (15)$$

Métricas de polarização

Centralidade do corte da rede [Garimella et al., 2018]

- ▶ Consideremos um grafo $G = (V, E)$ polarizado com grupos X e Y opostos e bem definidos ($X \cup Y = G$ e $X \cap Y = \emptyset$).
- ▶ O conjunto de arestas $C \subseteq E$ do corte $C = (X, Y)$ age como uma “ponte” levando informação entre os grupos.
- ▶ Neste caso, o caminho mínimo entre pares de vértices de grupos opostos devem conter alguma aresta do corte C .
- ▶ O que leva a valores altos de intermediação ($bc(e)$) das arestas em C .
- ▶ Se considerarmos um grafo onde os grupos não estejam tão bem separados, existirão outras arestas pelas quais a informação pode passar.
- ▶ O que leva a valores de intermediação das arestas em C que sejam similares ao restante das arestas do grafo.

Métricas de polarização

Centralidade do corte da rede [Garimella et al., 2018]

- ▶ Vamos transformar a ideia em uma métrica que compara a distribuição do cálculo de centralidade das arestas do corte com a centralidade das demais arestas do grafo.
- ▶ Para isso é computada a divergência de Kullback-Leibler (d_{KL}) entre a distribuição de centralidade dos dois conjuntos de arestas.
- ▶ A divergência de KL é uma medida de distância entre duas distribuições de probabilidade (mais detalhes estão fora do escopo deste minicurso).

$$BCC = 1 - e^{-d_{KL}}, \quad (16)$$

Métricas de polarização

Random Walk Controversy (RWC) [Garimella et al., 2018]

- ▶ A última métrica que iremos abordar almeja mensurar a facilidade do acesso da informação por usuários de grupos opostos.
- ▶ Supondo uma rede polarizada com poucas ligações entre os grupos, o acesso ao grupo oposto será dificultado.
- ▶ Utilizando caminhamento aleatório, a métrica *Random Walk Controversy (RWC)* calcula a probabilidade de um usuário acessar informações do grupo oposto.

Métricas de polarização

Random Walk Controversy (RWC) [Garimella et al., 2018]

Considerando dois usuários que irão caminhar aleatoriamente sobre a rede, a métrica *RWC* é definida como a diferença de probabilidade de que ambos terminem no grupo em que começaram e a probabilidade de que ambos terminem em grupos opostos.

$$RWC = P_{XX}P_{YY} - P_{XY}P_{YX}, \quad (17)$$

onde $P_{AB}, A, B \in \{X, Y\}$ é a seguinte probabilidade condicional:

$$P_{AB} = P[\text{começou na partição A} | \text{terminou na partição B}]. \quad (18)$$

Métricas de polarização

Random Walk Controversy (RWC) [Garimella et al., 2018]

- ▶ O resultado da métrica *RWC* ficará próximo a 1 quando a probabilidade de cruzar os grupos é baixa, o que implica uma alta polarização.
- ▶ Por outro lado, o resultado da métrica ficará próximo a 0 quando a probabilidade de cruzar os grupos é comparável a de terminar do mesmo lado, o que implica uma baixa polarização.

Análises sobre polarização



5. Análises sobre polarização

- 5.1 *Baselines* das métricas de polarização
- 5.2 Polarização como *features*
- 5.3 Impactos de eventos nas mídias sociais
- 5.4 Previsão de resultados de eleições
- 5.5 Análise do comportamento dos usuários
- 5.6 Influência dos algoritmos na polarização
- 5.7 Abusos do uso da polarização
- 5.8 Contenção da polarização

- ▶ Chegamos no último estágio da metodologia que é a análise da polarização à partir dos dados coletados e processados.
- ▶ A análise da polarização deve levar em consideração as particularidades das áreas e os objetivos a serem alcançados.
- ▶ Vamos discutir alguns tópicos de pesquisa encontrados na literatura.

Análises sobre polarização

1. *Baselines* para validação e comparação de métricas de polarização.
2. Utilização de polarização como *features* em áreas relacionadas, como hate-speech, toxicidade, fake-news.
3. Análises do impacto nas redes sociais de fatos que acontecem na vida real, por exemplo a morte de uma celebridade ou a corrida eleitoral em um país.
4. Previsão de resultados de eleições.
5. Análises de comportamento dos usuários em torno da polarização.
6. Impacto (ou influência) dos algoritmos das mídias sociais na sua polarização.
7. Abusos do uso da polarização de forma deliberada para ganho próprio.
8. Ferramentas e soluções para conter a polarização.

Análises sobre polarização

Baselines das métricas de polarização

- ▶ Trabalhos validaram e compararam diversas métricas de polarização.
- ▶ Métricas de polarização que utilizam o grafo da rede.
- ▶ Métricas baseadas em conteúdo.
- ▶ Métricas que são independentes do conteúdo.
- ▶ Métricas ainda demandam de um especialista.
- ▶ Estudos nesta área se mostraram importantes pois ainda não existe um consenso sobre como medir polarização.

Análises sobre polarização

[Garimella 2018] Quantifying Controversy on Social Media.

Dataset	$ V $	$ E $	C?	RWC	BCC	EC	GMCK	MBLB
Political blogs	1222	16 714	✓	0.42	0.53	0.49	0.18	0.45
Twitter politics	18 470	48 053	✓	0.77	0.79	0.62	0.28	0.34
Gun control	33 254	349 782	✓	0.70	0.68	0.55	0.24	0.81
Brazil soccer	20 594	82 421	✓	0.67	0.48	0.68	0.17	0.75
Karate club	34	78	✓	0.11	0.64	0.51	0.17	0.11
Facebook university	281	4389	✗	0.35	0.26	0.38	0.01	0.27
NYC teams	95 924	176 249	✗	0.34	0.24	0.17	0.01	0.19

Figura 6: Comparação entre métricas de polarização em diferentes base de dados

Análises sobre polarização

Polarização como *features*

- ▶ Métricas de polarização não são um fim em si mesmas.
- ▶ Ferramenta para discurso de ódio, toxicidade, fake-news.
- ▶ Exemplo: Tópicos com discurso de ódio e população dividida (polarizada)
- ▶ Busca tópicos polarizados como ponto de partida.

Análises sobre polarização

Polarização como *features*

- ▶ Exemplo: Nível de discordância (polarização) entre especialistas utilizado juntamente com conjunto de treino para gerar melhores classificadores.
- ▶ Exemplo: Métricas de polarização melhoram previsões de mensagens tóxicas.
- ▶ Exemplo: Tópicos alvos de notícias falsas.
- ▶ Todos estes são exemplos de classificadores onde a polarização não é o único aspecto a ser considerado, porém utilizar essas informação trouxe ganho aos classificadores.

Análises sobre polarização

Impactos de eventos nas mídias sociais

- ▶ Polarização nas mídias sociais e eventos do mundo real.
- ▶ Exemplo: Polarização quando da morte do Hugo Chávez e,
- ▶ A crise política instaurada na Venezuela em 2019.
- ▶ O processo de Impeachment da presidenta Dilma Roussef no Brasil,
- ▶ Corrida eleitoral de Donald Trump em 2016, dentre outros.

Análises sobre polarização

Impactos de eventos nas mídias sociais

- ▶ Cada trabalho abordando como as mídias sociais reagiram aos eventos da vida real levando suas particularidades em consideração.
- ▶ [Moreira et al., 2020] comparou polarização entre “elite” (classe política) e “massa” (população comum).
- ▶ Viés da “elite” foi calculado à partir das votações ocorridas na câmara dos deputados.
- ▶ [Morales et al., 2015] Analisou a polarização dividindo a população em diferentes regiões geográficas.

Análises sobre polarização

[Moreira 2020] Elite versus mass polarization on the Brazilian impeachment

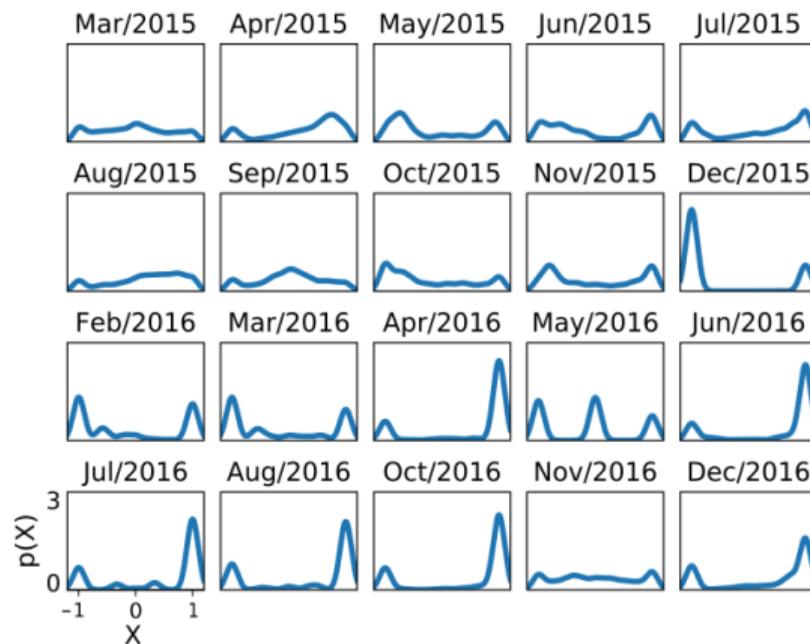


Figura 7: Evolução da polarização dos deputados federais.

Análises sobre polarização

[Moreira 2020] Elite versus mass polarization on the Brazilian impeachment

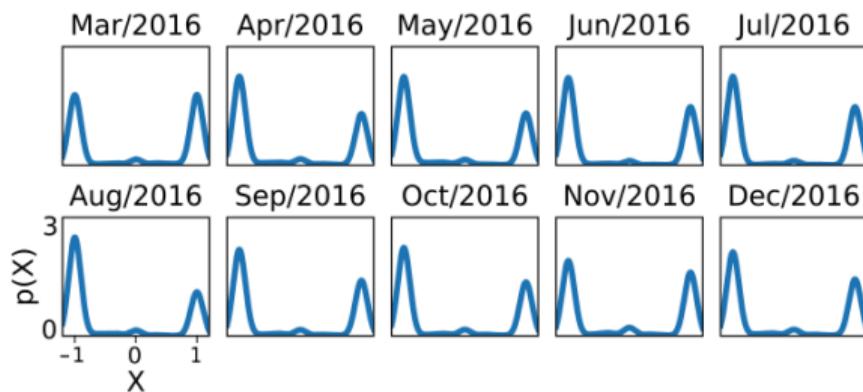


Figura 8: Evolução da polarização no Twitter.

Análises sobre polarização

[Moreira 2020] Elite versus mass polarization on the Brazilian impeachment

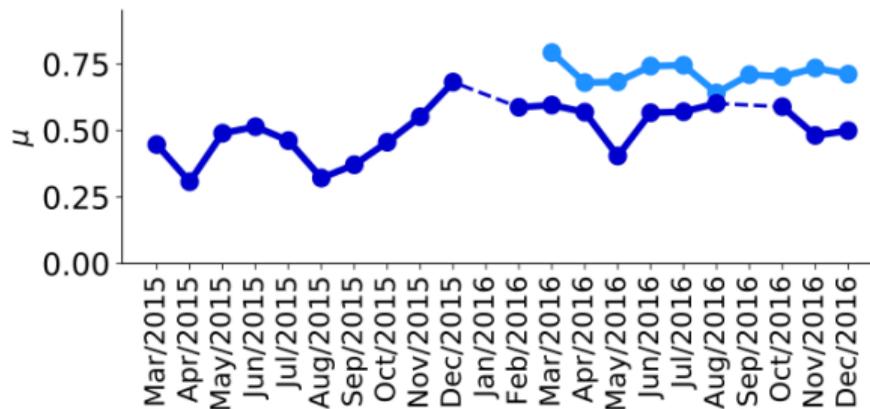


Figura 9: Evolução da polarização

Análises sobre polarização

Previsão de resultados de eleições

- ▶ Métricas de polarização classificam e quantificam grupos antagônicos da população.
- ▶ Quando o tópico escolhido é aborto, por exemplo, temos uma ideia do tamanho da população pró-escolha e o tamanho da população pró-vida.
- ▶ Quando o tópico é um jogo de futebol, quantificamos o tamanho da população que torce para cada um dos times.
- ▶ E, quando o tópico é uma **corrida eleitoral**, a análise da polarização em torno de um candidato nos trará o tamanho da população que o apoia.

Análises sobre polarização

Previsão de resultados de eleições

- ▶ [Belcastro et al., 2020] Classificação dos usuários do Twitter durante as eleições italianas de 2016.
- ▶ Viés dos usuários foi calculado como a razão entre o número de tweets compartilhados em apoio a cada candidato.
- ▶ Resultado: Previsão da votação mais acurada do que as pesquisas tradicionais.
- ▶ O mesmo resultado preditivo foi visualizado na eleição de Donald Trump em 2016.

Análises sobre polarização

[Belcastro 2020] Learning political polarization on social media using neural networks.

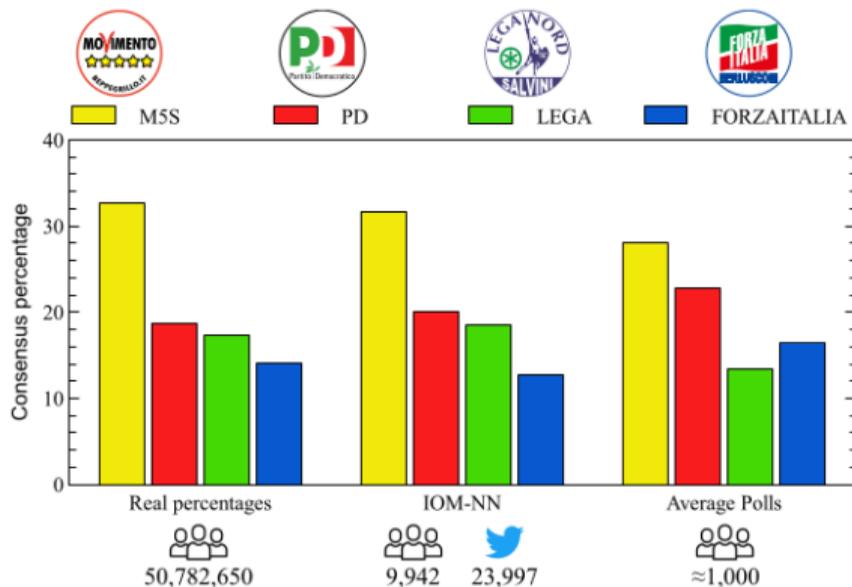


Figura 10: Comparação entre resultado da eleição, pesquisas eleitorais e IOM-NN

Análises sobre polarização

[Belcastro 2020] Learning political polarization on social media using neural networks.

	Colorado	Florida	Iowa	Michigan	New Hampshire	North Carolina	Ohio	Pennsylvania	Virginia	Wisconsin
Real										
IOM-NN										
Opinion polls										

Figura 11: Comparação entre resultado da eleição, pesquisas eleitorais e IOM-NN

Análises sobre polarização

Previsão de resultados de eleições

- ▶ Outras aplicações:
- ▶ Uma empresa pode estar interessada em prever a recepção de um novo produto, ou um político, em saber o nível de aprovação da população a um novo projeto, dentre outras aplicações.

Análises sobre polarização

Análise do comportamento dos usuários

- ▶ Entendimento do comportamento dos usuários em uma rede polarizada e suas consequências.
- ▶ Pontos a serem abordados:
 - ▶ Caracterização da rede e dos seus usuários,
 - ▶ Melhor entendimento do processo de homofilia e das câmaras de eco,
 - ▶ dos conteúdos que tem maior alcance em compartilhamento
 - ▶ ou tempo de permanência dos leitores em páginas com diferentes vieses.

Análises sobre polarização

Análise do comportamento dos usuários

Simulações das redes e do comportamento dos usuários.

- ▶ Avaliar do impacto da polarização na estrutura da rede.
- ▶ Entendimento do papel das decisões dos usuários ou do algoritmo da rede social.
- ▶ As diferentes particularidade entre as redes sociais, como o Twitter ou o Reddit, e seus impactos na polarização.

Análises sobre polarização

Análise do comportamento dos usuários

- ▶ [Waller and Anderson, 2021] Análise de polarização política em comunidades do Reddit desde sua criação
- ▶ Observaram que a polarização começou a aumentar durante o ano de 2016 com as eleições do Trump e não voltaram a diminuir.
- ▶ Observaram que a polarização se dá pelos novos usuários que entram na plataforma primariamente.

Análises sobre polarização

[Waller 2021] Quantifying social organization and political polarization in online platforms.

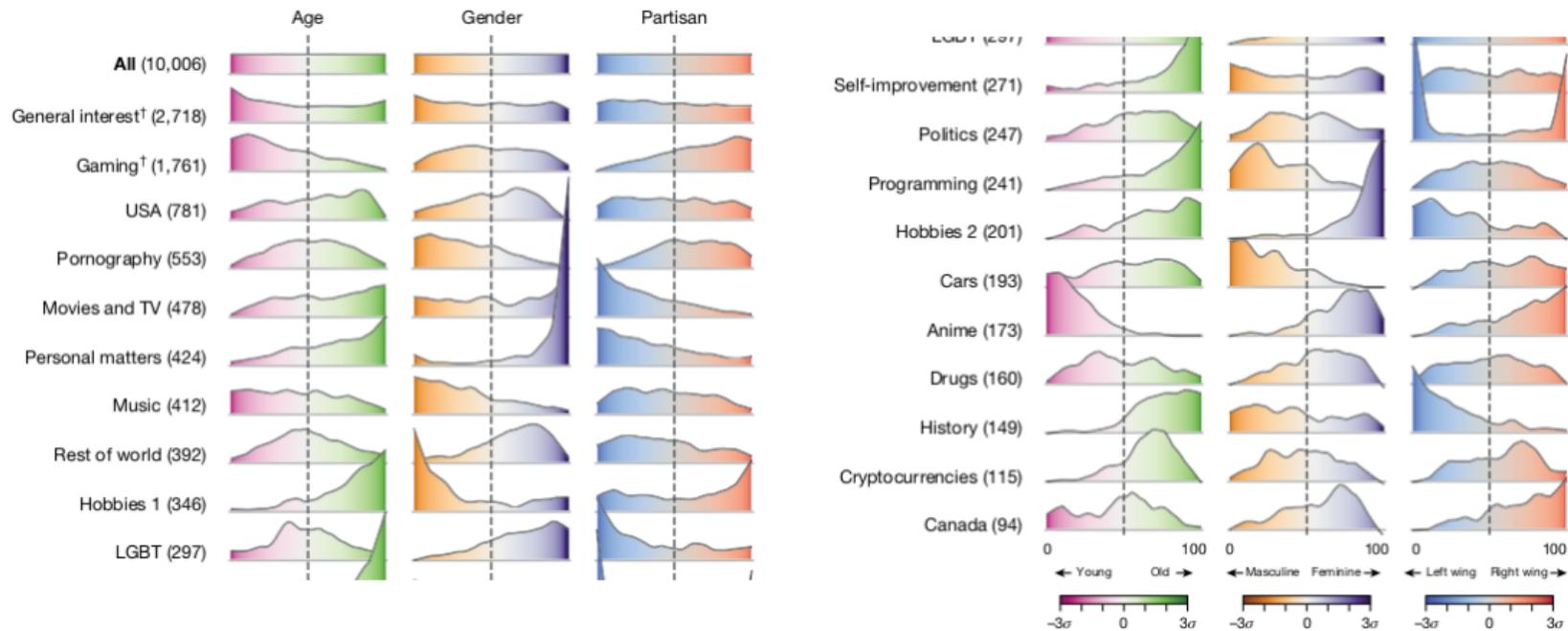
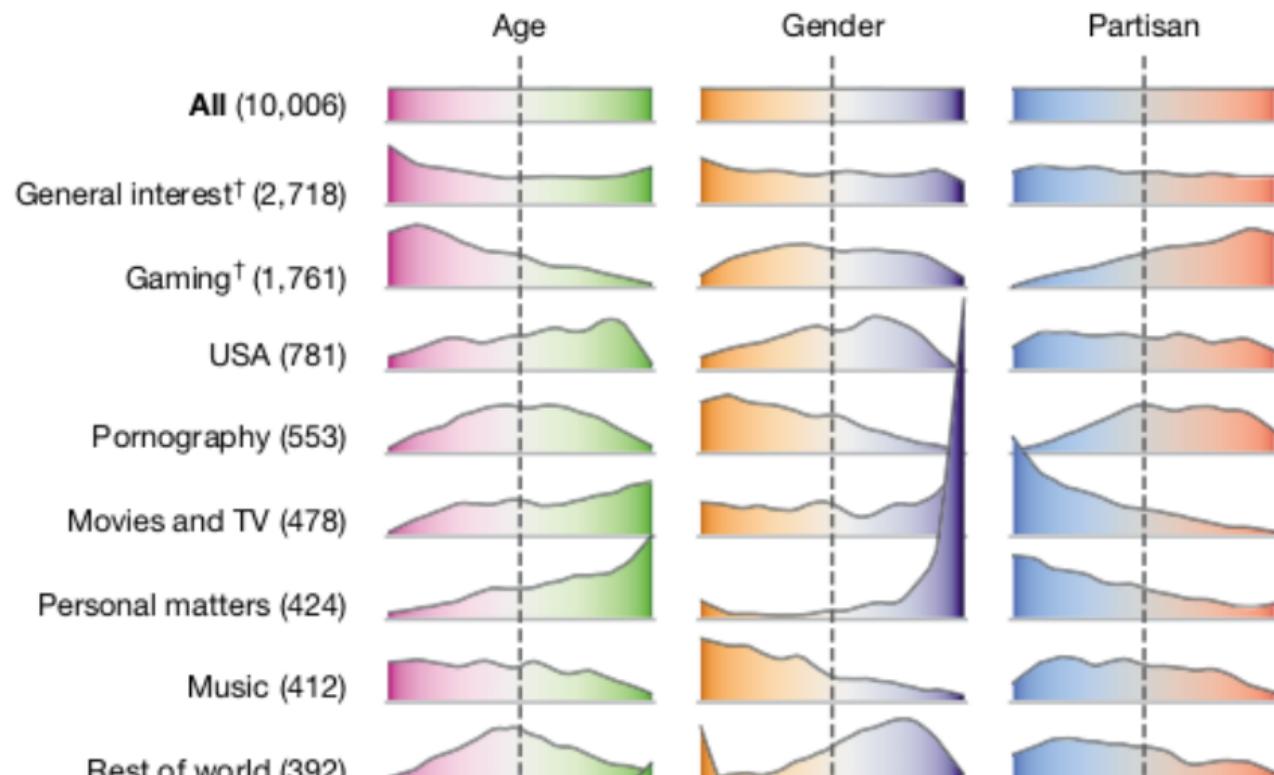


Figura 12: Organização social das comunidades do Reddit

Análises sobre polarização

[Waller 2021] Quantifying social organization and political polarization in online platforms.



Análises sobre polarização

[Waller 2021] Quantifying social organization and political polarization in online platforms.

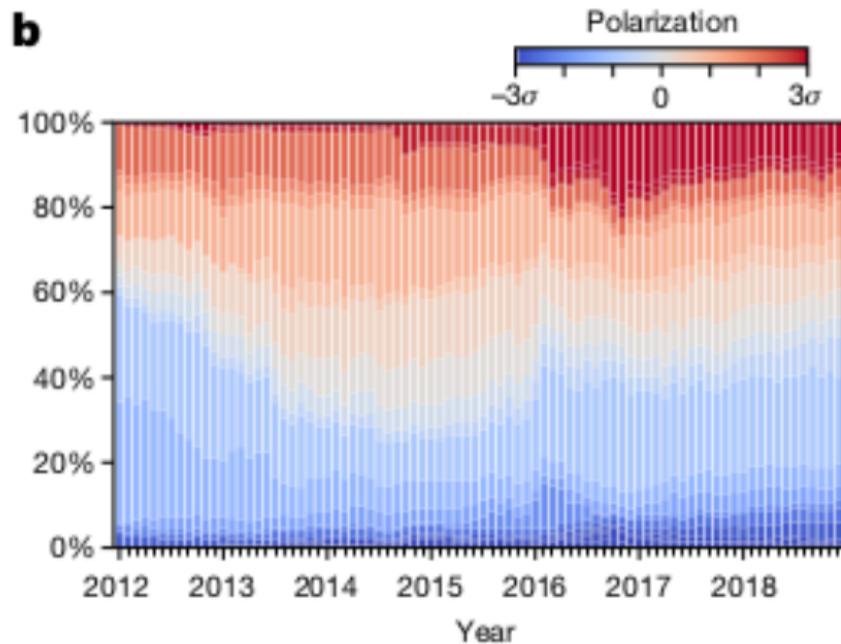


Figura 14: Distribuição de viés no Reddit

Análises sobre polarização

[Waller 2021] Quantifying social organization and political polarization in online platforms.

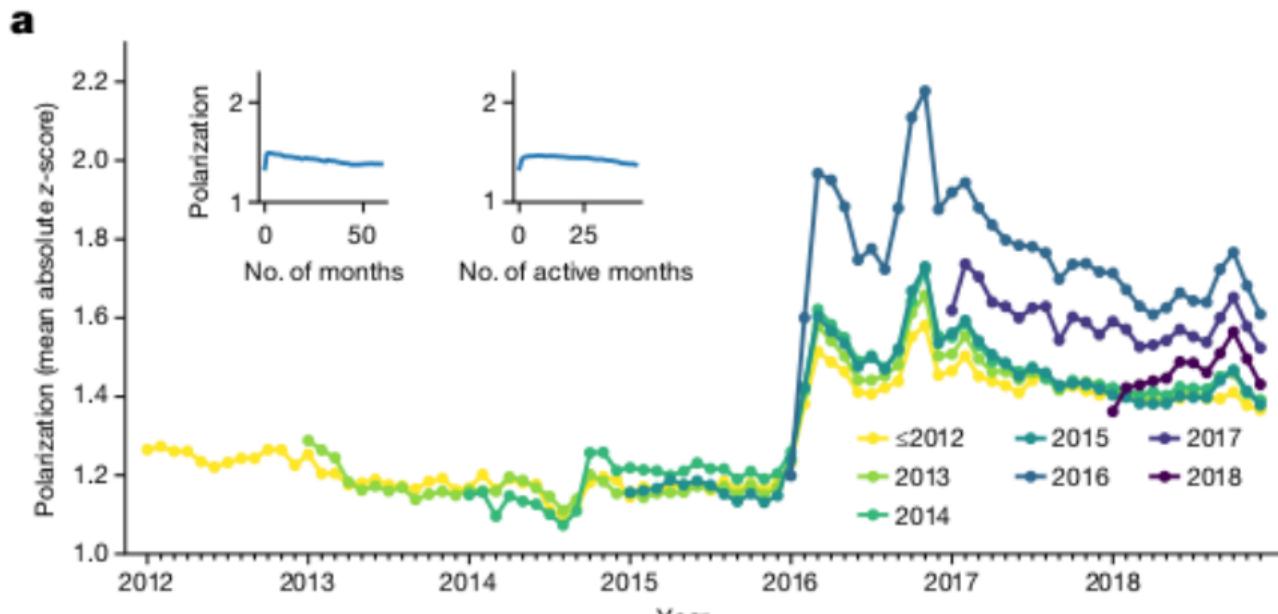


Figura 15: Polarização política de novos e antigos usuários

Análises sobre polarização

[Waller 2021] Quantifying social organization and political polarization in online platforms.

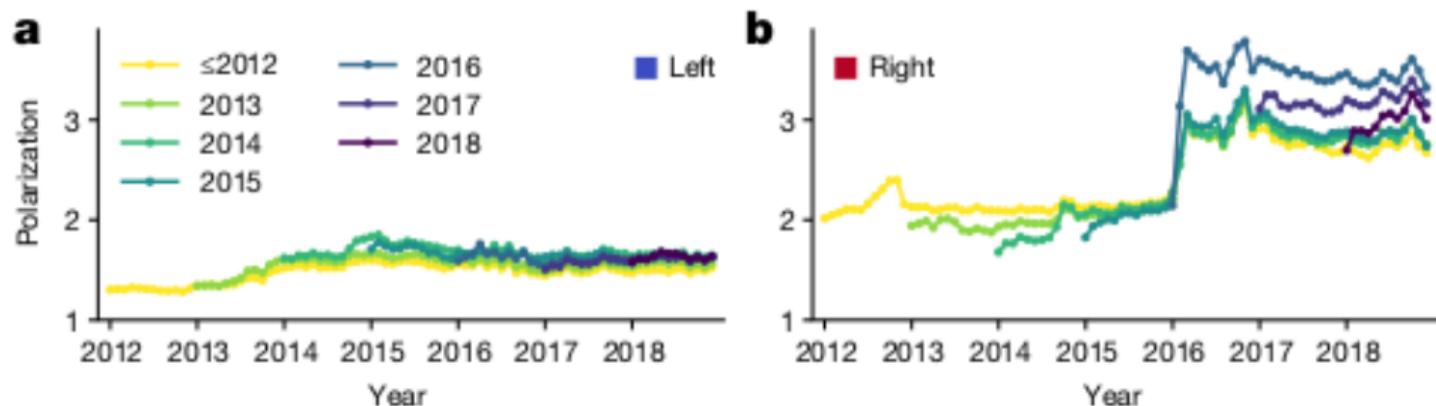


Figura 16: Assimetria ideológica na polarização

Análises sobre polarização

Influência dos algoritmos na polarização

- ▶ Indícios apontam um aumento da polarização ao longo dos anos.
- ▶ As redes sociais possuem sua parcela de responsabilidade?
- ▶ [Kulshrestha et al., 2017] focou em calcular o viés do algoritmo de busca do Twitter.
- ▶ Calcula os vieses dos tweets que retornam como respostas a determinadas consultas.
- ▶ [Valensise et al., 2022] Simulação para entender o impacto da escolha dos usuários e do algoritmo das mídias sociais.
- ▶ O trabalho concluiu que o algoritmo tem um grande papel na polarização dos usuários.

Análises sobre polarização

Abusos do uso da polarização

- ▶ Existem elementos que utilizam a polarização de forma deliberada?
- ▶ [Ribeiro et al., 2019] investigou uma série de anúncios políticos disparados por uma agência russa ao povo americano durante as eleições de 2016.
- ▶ Anúncios foram disparados com o intuito de explorar a polarização da rede.
- ▶ Direcionados para perfis específicos de usuários, possuindo um alcance 10 vezes maior do que a média de anúncios na plataforma.
- ▶ Atacantes procuraram perfis de usuários pertencentes a populações específicas (principalmente liberais e negros) com o objetivo de criar discórdia.

Análises sobre polarização

Contenção da polarização

- ▶ Indícios de que a polarização vem aumentando.
- ▶ Quais seriam as soluções existentes para mudar este cenário?
- ▶ Recomendação de usuários e conteúdos com viés diferente
- ▶ Elaboração de *feeds* de notícias que não sejam enviesados.
- ▶ Mudanças na rede de comunicação ao adicionar conexões entre antagônicos ou adicionar novos nós na rede.
- ▶ De modo geral, o objetivo dessas soluções é diminuir as câmaras de eco, promovendo o contato com visões opostas.

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Outline

6. Prática: Covid-19 e Hidroxicloroquina (HCQ)

6.1 Base de dados

6.2 Viés dos tweets

6.3 Polarização do grupo

6.4 Análise da polarização

Prática: Covid-19 e Hidroxicloroquina (HCQ)

- ▶ Apresentaremos um exemplo prático completo desde a coleta até a medição e análise de polarização em rede social.
- ▶ O código fonte utilizado nesta seção podem ser acessados em:
`https://github.com/brhott/webmedia2022-polarization`.
- ▶ Para este estudo prático utilizaremos ferramentas e APIs desenvolvidas na linguagem Python. As ferramentas serão apresentadas no desenvolver da seção.

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Base de dados

- ▶ A primeira fase é a escolha e obtenção da base de dados a ser estudada.
- ▶ Uso do medicamento Hidroxicloroquina (HCQ) para o tratamento da Covid-19.
- ▶ A base de dados será composta por um conjunto de tweets acerca do tema.
- ▶ Para coleta, utilizamos a ferramenta *Tweepy*³ para acesso à API do Twitter⁴.

³<https://tweepy.org>

⁴<https://developer.twitter.com>

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Base de dados

```
1 import tweepy
2 import pandas as pd
3
4 client = tweepy.Client(Bearer_Token)
5 query = "hydroxychloroquine chloroquine HCQ -is:retweet"
6 tweets = client.search_recent_tweets(query=query, max_results=100)
7 df = pd.DataFrame(tweets.data).set_index('id')
```

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Base de dados

- ▶ Aprendemos como realizar uma consulta utilizando a biblioteca *Tweepy*.
- ▶ Vamos utilizar a base de dados encontrada em [Mutlu et al., 2020].
- ▶ 14.374 *tweets* sobre o uso da Hidroxicloroquina como medicamento para a covid-19.
- ▶ Viés classificados manualmente:
- ▶ “cloroquina/hidroxicloroquina é a cura para o novo coronavírus?”.
- ▶ 9.117 *tweets* continuavam online, sendo 3.732 *tweets* classificados como negativos, 3.385, positivos, e, 2.000, neutros.

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Viés dos tweets

- ▶ A segunda fase da metodologia é a de cálculo dos vieses dos dados.
- ▶ Optamos por utilizar uma base de dados anotada manualmente.
- ▶ Porém, vamos abordar também a análise de sentimentos destes tweets.
- ▶ Utilizaremos a ferramenta *Vader*⁵ para extrair o sentimento dos tweets.

⁵<https://nltk.org>

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Viés dos tweets

```
1 from nltk.sentiment.vader import SentimentIntensityAnalyzer
2 sid = SentimentIntensityAnalyzer() #inicializacao
3
4 scores = df['text'].apply(lambda text : sid.polarity_scores(text))
5 df['compound'] = scores.apply(lambda score : score['compound'])
```

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Viés dos tweets

Segundo a documentação, os valores são interpretados da seguinte maneira:

- ▶ valores no intervalo $[-1.0, -0.05)$ denominam um tweet com sentimento negativo;
- ▶ valores no intervalo $[-0.05, 0.05]$, sentimento neutro; e
- ▶ valores no intervalo $(0.05, 1.0]$, sentimento positivo.

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Polarização do grupo

- ▶ A terceira fase compreende na escolha e aplicação das métricas de polarização.
- ▶ Como estamos trabalhando com o conteúdo dos tweets e não com a estrutura da rede, vamos implementar as métricas de **análise de sentimentos** e do **dipolo elétrico**.

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Polarização do grupo

```
1 # Separacao entre tweets positivos e negativos.
2 g = df['compound']
3 gn = df[df['compound'] <= -0.05]['compound']
4 gp = df[df['compound'] >= 0.05]['compound']
5
6 A = g.count() # num de tweets totais
7 An = gn.count() # populacao de tweets negativos
8 Ap = gp.count() # populacao de tweets positivos
9 A0 = A - An - Ap # populacao de tweets neutros
10 Sn = abs(gn.sum()) # soma dos valores de sentimento negativo
11 Sp = gp.sum() # soma dos valores de sentimento positivo
12 gcp = gp.mean() # centroide dos tweets positivos
13 gcn = gn.mean() # centroide dos tweets negativos
```

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Polarização do grupo

```
1 PN      = Ap / An                # metrica PN
2 RPN     = min(An, Ap) / max(An, Ap) # metrica RPN
3 NPN     = AO / (An + Ap)         # metrica NPN
4 PNT     = (Ap + An) / A          # metrica PNT
5 PNPNT   = PN * PNT              # metrica PNPNT
6 RPNV    = min(Sn, Sp) / max(Sn, Sp) # metrica RPNV
7
8 dA      = abs(Ap - An) / A      # diferenca do tamanho das populacoes
9 d       = (gcp - gcn) / 2       # distancia entre centroides
10 m      = (1 - dA) * d          # metrica do dipolo eletrico
```

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Polarização do grupo

- ▶ O código utilizou os dados de análise de sentimento calculados pelo *Vader*.
- ▶ Porém, é fácil observar que as linhas 2-4 podem ser facilmente adaptadas para carregar os valores de viés obtidos por outros meios.
- ▶ Também iremos apresentar os resultados de cálculo das métricas utilizando os vieses classificados manualmente.

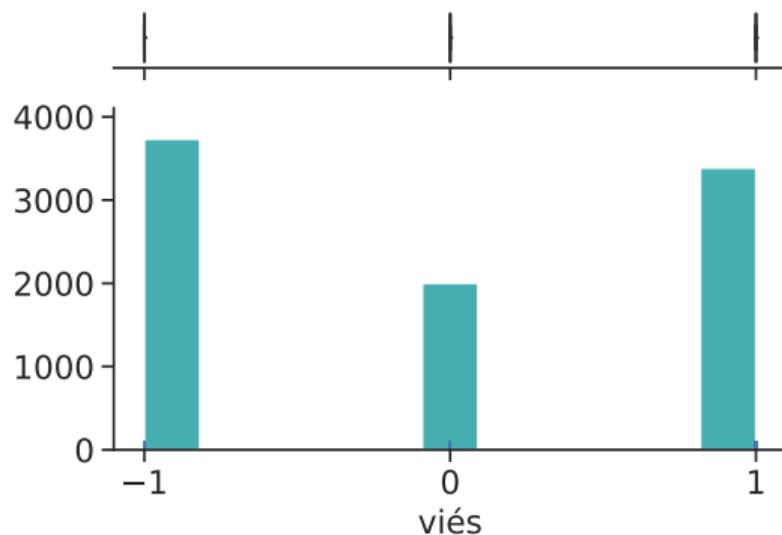
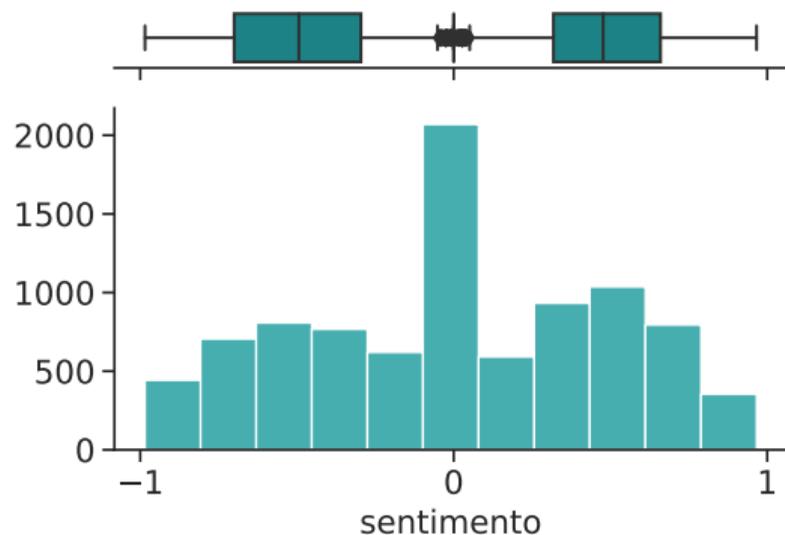
Prática: Covid-19 e Hidroxicloroquina (HCQ)

Análise da polarização

- ▶ A última etapa consiste na análise e discussão dos resultados.
- ▶ Pela visualização de ambos os histogramas, podemos esperar uma alta polarização dessa população (população dividida entre dois grupos de posições opostas).

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Análise da polarização



Prática: Covid-19 e Hidroxicloroquina (HCQ)

Análise da polarização

	PN	RPN	NPN	PNT	PNPNT	RPNV	DM
S	1.11	0.90	0.25	0.80	0.89	0.94	0.47
M	0.91	0.91	0.28	0.78	0.70	0.91	0.96

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Análise da polarização

- ▶ A polarização encontrada foi alta na grande maioria das métricas.
- ▶ A métrica NPN é a única que é inversamente proporcional à polarização.
- ▶ A métrica DM obteve resultados diferentes entre as bases de dados (0.47 e 0.96).
- ▶ A métrica utiliza o distanciamento das opiniões de cada um dos grupos.
- ▶ O viés anotado manualmente terá distância máxima entre os grupos.
- ▶ Não é prudente comparar os resultados com vieses calculados de maneira diferente.

Prática: Covid-19 e Hidroxicloroquina (HCQ)

Análise da polarização

- ▶ Os resultados das métricas, juntamente com a visualização dos histogramas dos *tweets* nos trazem valiosas informações acerca da polarização do tópico em questão.
- ▶ Vimos que as populações antagônicas tem tamanhos próximos (PN, RPN),
- ▶ o número de tweets neutros é baixo (NPN, PNT) e
- ▶ a razão entre o somatório dos vieses de cada grupo fica próxima de 1 (RPNV), ou seja, que os grupos estão enviesados (distância até o centro) de maneira similar entre eles.
- ▶ Todas estas são características de uma população polarizada.

Considerações finais

- ▶ Identificar polarização nas redes sociais ainda é uma tarefa dependente do contexto.
- ▶ Apresentamos uma revisão bibliográfica com o objetivo de contextualizar e apresentar ao leitor o atual cenário da pesquisa em polarização.
- ▶ Apresentamos os principais conceitos e definições da área
- ▶ E provemos ao leitor o ferramental necessário para elaborar suas próprias análises na área de polarização.

Propusemos e apresentamos uma metodologia que passa

1. pela coleta e processamento de dados no contexto da polarização;
2. a classificação do viés das entidades presentes nos dados, com enfoque em texto;
3. a escolha e aplicação de métricas de identificação e, a depender da técnica utilizada quantificação da polarização do tópico em questão;
4. por fim, a análise e interpretação dos resultados da polarização.

- ▶ Foi proposta uma taxonomia das métricas e técnicas de polarização em redes sociais.
- ▶ Abordamos desde métricas estatísticas a métricas que utilizam técnicas de teoria dos grafos, o que evidencia a falta um consenso na literatura sobre como operacionalizar a identificação e a polarização de uma população.
- ▶ Apresentamos diversos exemplos de trabalhos correlatos, seus resultados e suas implicações.
- ▶ Apresentamos um exemplo prático de análise de polarização no contexto da pandemia Covid-19 e da discussão em torno do medicamento Hidroxicloroquina.

Muito Obrigado!

Bruno Hott

`brunohott@dcc.ufmg.br`

`https://github.com/brhott/webmedia2022-polarization`

-  Akhtar, S., Basile, V., and Patti, V. (2019).
A new measure of polarization in the annotation of hate speech.
In *International Conference of the Italian Association for Artificial Intelligence*, pages 588–603. Springer.
-  Akoglu, L. (2014).
Quantifying political polarity based on bipartite opinion networks.
In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 8, pages 2–11.
-  Al Amin, M. T., Aggarwal, C., Yao, S., Abdelzaher, T., and Kaplan, L. (2017).
Unveiling polarization in social networks: A matrix factorization approach.
In *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, pages 1–9. IEEE.
-  Al-Ayyoub, M., Rabab'ah, A., Jararweh, Y., Al-Kabi, M. N., and Gupta, B. B. (2018).
Studying the controversy in online crowds' interactions.

-  Babaei, M., Kulshrestha, J., Chakraborty, A., Benevenuto, F., Gummadi, K. P., and  Weller, A. (2018).
Purple feed: Identifying high consensus news posts on social media.
In Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, pages 10–16.
-  Belcastro, L., Cantini, R., Marozzo, F., Talia, D., and Trunfio, P. (2020).
Learning political polarization on social media using neural networks.
IEEE Access, 8:47177–47187.
-  Choi, Y., Jung, Y., and Myaeng, S.-H. (2010).
Identifying controversial issues and their sub-topics in news articles.
In Pacific-Asia Workshop on Intelligence and Security Informatics, pages 140–153.
Springer.
-  Coletto, M., Garimella, K., Gionis, A., and Lucchese, C. (2017).

Automatic controversy detection in social media: a content-independent motif-based approach.

Online Social Networks and Media, 3:22–31.



Darwish, K., Stefanov, P., Aupetit, M., and Nakov, P. (2020).

Unsupervised user stance detection on twitter.

Proceedings of the International AAAI Conference on Web and Social Media, 14(1):141–152.



Dori-Hacohen, S. and Allan, J. (2015).

Automated controversy detection on the web.

In *European Conference on Information Retrieval*, pages 423–434. Springer.



Elfardy, H. and Diab, M. (2016).

CU-GWU perspective at SemEval-2016 task 6: Ideological stance detection in informal text.

In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 434–439, San Diego, California. Association for Computational Linguistics.

 Ferreira, C. H., Murai, F., Silva, A. P., Almeida, J. M., Trevisan, M., Vassio, L., Mellia, M., and Drago, I. (2021).

On the dynamics of political discussions on instagram: A network perspective.
Online Social Networks and Media, 25:100155.

 Ferreira, W. and Vlachos, A. (2019).

Incorporating label dependencies in multilabel stance detection.

In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 6350–6354, Hong Kong, China. Association for Computational Linguistics.

 Fiorina, M. P., Abrams, S. A., and Pope, J. C. (2008).

Polarization in the american public: Misconceptions and misreadings.

The Journal of Politics.

 Garimella, K., Morales, G. D. F., Gionis, A., and Mathioudakis, M. (2018).
Quantifying controversy on social media.

ACM Transactions on Social Computing, 1(1):1–27.

 Garimella, K., Smith, T., Weiss, R., and West, R. (2021).
Political polarization in online news consumption.

In *Proceedings of the International AAAI Conference on Web and Social Media*,
volume 15, pages 152–162.

 Gillani, N., Yuan, A., Saveski, M., Vosoughi, S., and Roy, D. (2018).
Me, my echo chamber, and i: Introspection on social media polarization.

The Web Conference 2018 - Proceedings of the World Wide Web Conference, WWW 2018,
pages 823–831.

-  Guerra, P., Meira Jr, W., Cardie, C., and Kleinberg, R. (2013).
A measure of polarization on social media networks based on community boundaries. 
In *Proceedings of the International AAAI Conference on Web and Social Media*.
-  Jang, M. and Allan, J. (2018).
Explaining controversy on social media via stance summarization.
41st International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2018, pages 1221–1224.
-  Jang, M., Foley, J., Dori-Hacohen, S., and Allan, J. (2016).
Probabilistic approaches to controversy detection.
In *Proceedings of the 25th ACM international on conference on information and knowledge management*, pages 2069–2072.
-  Klenner, M., Amsler, M., and Hollenstein, N. (2014).

Verb polarity frames: a new resource and its application in target-specific polarity classification.

In *KONVENS*, pages 106–115.



Küçük, D. and Can, F. (2020).

Stance detection: A survey.

ACM Comput. Surv., 53(1).



Kulshrestha, J., Eslami, M., Messias, J., Zafar, M. B., Ghosh, S., Gummadi, K. P., and Karahalios, K. (2017).

Quantifying search bias: Investigating sources of bias for political searches in social media.

In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 417–432.



Li, Y. and Caragea, C. (2019).

Multi-task stance detection with sentiment and stance lexicons.

In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 6299–6305, Hong Kong, China. Association for Computational Linguistics. DCC

 Mejova, Y., Zhang, A. X., Diakopoulos, N., and Castillo, C. (2014).

Controversy and sentiment in online news.

arXiv preprint arXiv:1409.8152.

 Morales, A. J., Borondo, J., Losada, J. C., and Benito, R. M. (2015).

Measuring political polarization: Twitter shows the two sides of venezuela.

Chaos: An Interdisciplinary Journal of Nonlinear Science, 25(3):033114.

 Moreira, R. C., Vaz-de Melo, P. O., and Pappa, G. L. (2020).

Elite versus mass polarization on the brazilian impeachment proceedings of 2016.

Social Network Analysis and Mining.

 Mutlu, E. C., Oghaz, T., Jasser, J., Tutunculer, E., Rajabi, A., Tayebi, A., Ozmen, O., and Garibay, I. (2020).

A stance data set on polarized conversations on twitter about the efficacy of hydroxychloroquine as a treatment for covid-19.

Data in brief, 33:106401.

 Popescu, A.-M. and Pennacchiotti, M. (2010).

Detecting controversial events from twitter.

In Proceedings of the 19th ACM international conference on Information and knowledge management, pages 1873–1876.

 Rashed, A., Kutlu, M., Darwish, K., Elsayed, T., and Bayrak, C. (2020).

Embeddings-based clustering for target specific stances: The case of a polarized turkey.

CoRR, abs/2005.09649.

 Ribeiro, F. N., Saha, K., Babaei, M., Henrique, L., Messias, J., Benevenuto, F., Goga, O., Gummadi, K. P., and Redmiles, E. M. (2019). **DCC**

On microtargeting socially divisive ads: A case study of russia-linked ad campaigns on facebook.

In Proceedings of the Conference on Fairness, Accountability, and Transparency, pages 140–149.

 Roy, S. and Goldwasser, D. (2020).

Weakly supervised learning of nuanced frames for analyzing polarization in news media.

In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 7698–7716, Online. Association for Computational Linguistics.

 Shahrezaye, M., Papakyriakopoulos, O., Serrano, J. C. M., and Hegelich, S. (2019).

Measuring the ease of communication in bipartite social endorsement networks: A proxy to study the dynamics of political polarization.

ACM International Conference Proceeding Series, pages 158–165.



Sunstein, C. R. (1999).

The law of group polarization.

University of Chicago Law School, John M. Olin Law & Economics Working Paper.



Tokita, C. K., Guess, A. M., and Tarnita, C. E. (2021).

Polarized information ecosystems can reorganize social networks via information cascades.

Proceedings of the National Academy of Sciences of the United States of America, 118(50).



Tsytsarau, M., Palpanas, T., and Denecke, K. (2011).

Scalable detection of sentiment-based contradictions.

DiversiWeb, WWW, 1:9–16.

-  Valensise, C. M., Cinelli, M., and Quattrociocchi, W. (2022).
The dynamics of online polarization.
arXiv preprint arXiv:2205.15958.
-  Vicario, M. D., Quattrociocchi, W., Scala, A., and Zollo, F. (2019).
Polarization and fake news: Early warning of potential misinformation targets.
ACM Transactions on the Web (TWEB), 13(2):1–22.
-  Waller, I. and Anderson, A. (2021).
Quantifying social organization and political polarization in online platforms.
Nature, 600(7888):264–268.
-  Yang, M., Wen, X., Lin, Y.-R., and Deng, L. (2017).
Quantifying content polarization on twitter.
In *2017 IEEE 3rd international conference on collaboration and internet computing (CIC)*, pages 299–308. IEEE.