



Utilização de modelagem de tópicos para identificar os assuntos mais discutidos sobre o Haiti no Twitter

Jod Fedlet Pierre

fedletpierre15@gmail.com

Denio Duarte

duarte@uffs.edu.br

Roteiro

- **Introdução**
- **Referencial teórico**
- **Trabalhos relacionados**
- **Metodologia**
- **Resultados e discussões**
- **Conclusão**

Introdução

- **Acontecimentos no Haiti**
- **Consequência desses acontecimentos / emigração**
- **Rede social Twitter / Disseminação de informação sobre o Haiti**
- **Agrupar informação por tópicos / Modelagem de tópicos**



Referencial teórico

- **Haiti**



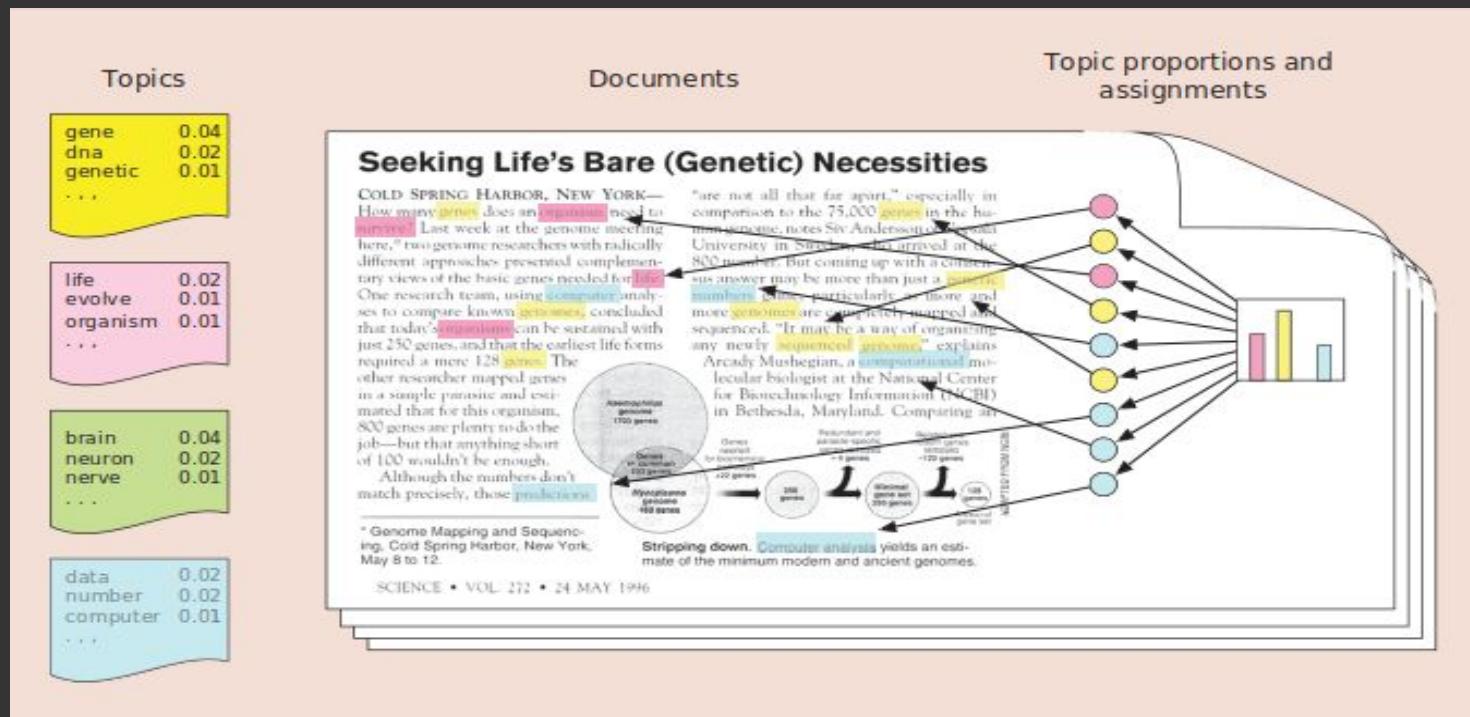
Referencial teórico

- **Twitter**
 - Lançado em 2006
 - 280 caracteres



Referencial teórico

● Modelagem de tópicos



Referencial teórico

- Modelagem de tópicos

Tabela 1 – Três tópicos com as suas top-5 palavras e probabilidades.

| word | P(w) | word | P(w) | word | P(w) |
|---------|-------|------------|-------|-----------|-------|
| court | 0.025 | ventilator | 0.028 | italy | 0.028 |
| police | 0.018 | mask | 0.028 | infection | 0.015 |
| law | 0.015 | supply | 0.023 | lockdown | 0.013 |
| prison | 0.010 | care | 0.021 | europe | 0.012 |
| justice | 0.009 | doctor | 0.019 | authority | 0.012 |

Referencial teórico

- **Modelagem de tópicos**
 - **Latent Dirichlet Allocation (LDA)**

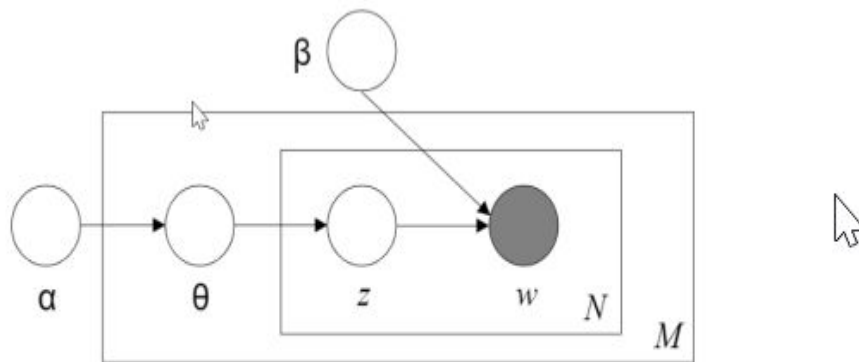


Figura 4 – Representação gráfica do modelo LDA.

Trabalhos relacionados

PEREIRA, Mariana. Análise exploratória de tweets utilizando modelagem de tópicos para textos curtos: caso Olimpíadas Rio 2016. Universidade Federal da Fronteira Sul, 2019.

- **Extração dos dados no Twitter**
 - rio2016, olimpiadas, olimpiada, ceremoniaderura, ceremoniadeencerramento, jogosolimpicos, olympics, olympicgames, openingceremony, closingceremony
 - 02/08/2016 a 24/08/2016
- **Pré-processamento dos tweets / BTM (Bitrem Topic Model) /Quantidade de tópicos: 15**

Trabalhos relacionados

ASGHARI, Mohsen; SIERRA-ROSA, Daniel; ELMAGHRABY, Adel. Trends on health in social media: Analysis using twitter topic modeling. In: IEEE. 2018 IEEE international symposium on signal processing and information technology (ISSPIT). [S.l.: s.n.], 2018. p. 558-563.

- **Extração dos dados no Twitter**

- **healthcare, health, doctors, homecare, digitalhealth and digital health.**

- **Pré-processamento dos tweets / LDA e Word2vect / Quantidade de tópicos: 5**
- **Resultados**

Tópico: Care / Estados mais interessados: NY e California (80%) / menos: Virgínia (41%)

Trabalhos relacionados

OH, Onook; KWON, Kyounghee Hazel; RAO, H Raghav. An exploration of social media in extreme events: Rumor theory and twitter during the haiti earthquake 2010, 2010.

- **Extração dos dados no Twitter**

- **haitiearthquake**

- **12/01/2010 a 21/01/2010**

- **Pré-processamento dos tweets / ANOVA de Friedman**

- **Resultados**

blog, CNN, picture, list, info, report e update -> representam declarações de autenticação.

Metodologia

- **Definição de datas e palavras-chave**
 - 01/01/2018 a 31/05/2022; haiti, haitiano, haitiana
- **Coleta dos dados**
 - Extração dos tweets usando a biblioteca Snsrape (185.348).
- **Pré-processamento dos dados**
 - Limpeza dos tweets extraídos (150.174)

| |
|---|
| Antes do pré-processamento |
| FUGA DO BOZO: 'Presidente, o senhor é racista?', pergunta repórter a Trump sobre polêmica envolvendo Haiti e países africanos ele saiu sem responder. https://t.co/AUprDvBNT0 |
| Depois do pré-processamento |
| fuga bozo presidente senhor racista perguntar reporter trump sobre polemica envolver haiti pais africano sair responder |

Tabela 2 – Exemplo de *tweet* antes e depois do pré-processamento dos dados.

Metodologia

- **Aplicação do LDA**

- Treinamento com $K = 3, 5, 7, 10, 15$ e 25 / α e $\beta = \text{auto}$
- Grupo mensal (53 meses)
- Melhor $K = 5$, usando C_v .
- Extração dos tópicos / identificar os mais frequentes

- **Rotulação dos tópicos**

- Tópico: terremoto, morto, passar, sofrer, grande, situacao, numero_morto, taliba, noticia, tristeza
- Rótulo: terremoto

Resultados e discussões

● 2018

| Mês | Palavras | Assunto |
|------|---|----------------------|
| jan. | haiti, trump, brasil, eua, africa, pessoa, brasileiro, chamar, onu, terremoto | imigração e racismo |
| fev. | exercito, militar, general, brasileiro, povo, rio_janeiro, trabalho, usar, coisa, sexual | intervenção militar |
| mar. | haiti, brasil, rio, gente, exercito, brasileiro, militar, sírio, ajudar, venezuela | intervenção militar |
| abr. | haiti, brasil, exercito, usar, militar, vida, missao_paz, terremoto, mandar, missao | intervenção militar |
| maio | contra, messi, selecao, jogo, grande, cuba, argentino, amistoso, fazer_gol, hat_trick | jogo de futebol |
| jun. | haiti, brasil, brasileiro, selecao, povo, ficar, ajudar, futebol, pessoa, venezuela | jogo de futebol |
| jul. | haiti, gente, brasileiro, jamaica, ficar, selecao, chorar, pessoa, mundo, pais | jogo de futebol |
| ago. | haiti, brasil, militar, brasileiro, ficar, chegar, bolsonaro, missao, mundo, grande | intervenção militar |
| set. | haiti, brasileiro, passar, pessoa, missao_paz, contra, governo, mundo, missao, militar | intervenção militar |
| out. | haiti, povo, cuba, brasileiro, passar, militar, presidente, paz, conhecer, guerra | comparação econômica |
| nov. | haiti, brasil, grande, ajudar, mundo, contra, eua, usar, coisa, deixar | imigração |
| dez. | haiti, brasil, militar, brasileiro, imigrante, bolsonaro, haitiana_cega, governo, povo, aprovar_oab | imigração e educação |

Tabela 4 – Tópico predominante de cada mês de 2018.

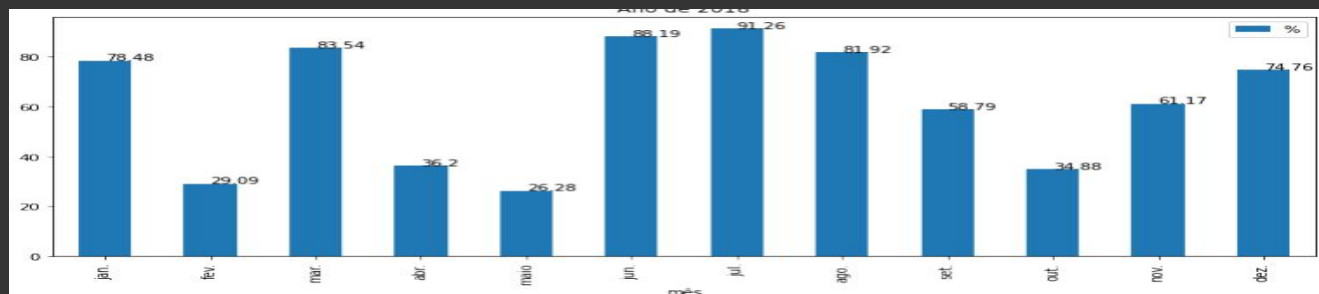


Figura 10 – Frequência dos tópicos predominantes de cada mês de 2018.

Resultados e discussões

● 2019

| Mês | Palavras | Assunto |
|------|--|------------------------------------|
| jan. | haiti, terremoto, ajudar, brasileiro, exercito_brasileiro, militar, brumadinho, eua, atuar, precisar | terremoto de 12 de jan. de 2010 |
| fev. | eua, petroleo, venezuela, humanitario, ajuda_humanitaria, ajudar_humanitaria, mundo, trump, mandar, pais | ajuda humanitária discriminativa |
| mar. | haiti, brasil, venezuela, eua, ajudar, governo, petroleo, cuba, grande, mundo | ajuda humanitária discriminativa |
| abr. | haiti, pessoa, gente, ajudar, ficar, mundo, brasileiro, mocambique, venezuela, coisa | ajuda humanitária discriminativa |
| maio | haiti, brasil, brasileiro, venezuela, ficar, jovem, ajudar, grande, gente, governo | ajuda humanitária discriminativa |
| jun. | brasil, costa_rica, copa_ouro, grupo, jogo, eua, nicaragua, copa_america, curacao, contra | jogo de futebol |
| jul. | haiti, brasil, mexico, virar, jogo, eua, copa_ouro, ouro, grande, perder | jogo de futebol |
| ago. | haiti, brasil, brasileiro, chamar, presidente, mundo, governo, pobre, povo, general | comparação sociopolítico-econômica |
| set. | haiti, general, brasileiro, pessoa, povo, deixar, contra, mundo, pais, coisa | intervenção militar |
| out. | haiti, brasil, equador, protesto, presidente, mundo, chile, gente, rua, hong_kong | protesto contra o governo haitiano |
| nov. | haiti, heleno, contra, eua, negro, cuba, mundo, pensar, pais, pobre | atuação do general Heleno no Haiti |
| dez. | haiti, brasil, onu, soldado, filho, mulher, missao, ajudar, pessoa, vida | intervenção militar |

Tabela 5 – Tópico predominante de cada mês de 2019.

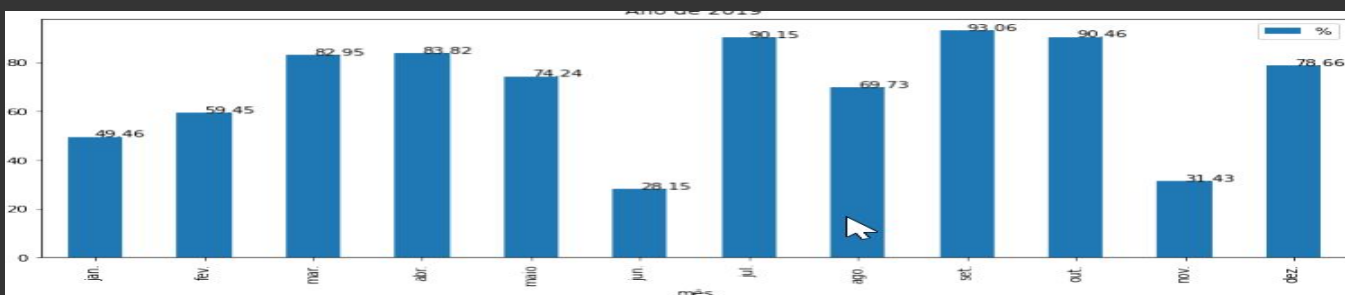


Figura 11 – Frequência dos tópicos predominantes de cada mês de 2019.

Resultados e discussões

● 2020

| Mês | Palavras | Assunto |
|------|---|-------------------------|
| jan. | general, massacre, missao, filme, governo, general_helene, filme_sobre, terror, precisar, haitiana | genocídio |
| fev. | lula, helene, ficar, general_helene, criminoso, povo, brasileiro, militar, deixar, gente | genocídio |
| mar. | haitiano, presidente, bolsonaro, brasil, brasileiro, mandar, ficar, governo, bolsonaro_acabar, sair | Intromissão política |
| abr. | haitiano, presidente, bolsonaro, acabar, governo, bolsonaro_acabar, vir_futuro, futuro, brasileiro, razao | Intromissão política |
| maio | haiti, haitiano, presidente, bolsonaro, governo, brasileiro, povo, lembrar, revolucao, tempo | Intromissão política |
| jun. | haiti, brasil, republica_dominicana, panama, colombia_equador, trinidad_tobago, pais, suriname_trinidad, banco_mundial, grupo | países do Banco Mundial |
| jul. | haiti, ficar, eua, brasileiro, grande, povo, governo, general, matar, ajudar | genocídio |
| ago. | haiti, brasil, cuba, pessoa, mundo, eua, gente, republica_dominicana, grande, senhor | tempestade Laura |
| set. | haiti, brasil, negro, brasileiro, gente, matar, grande, militar, pessoa, pobre | genocídio |
| out. | haiti, passar, eua, deixar, ficar, povo, olhar, antes, tornar, acontecer | imigração |
| nov. | haiti, brasil, eua, brasileiro, mundo, chegar, pais, pessoa, venezuela, pensar | imigração |
| dez. | haiti, brasil, gente, general, governo, helene, bolsonaro, pais, negro, venezuela | genocídio |

Tabela 6 – Tópico predominante de cada mês de 2020.

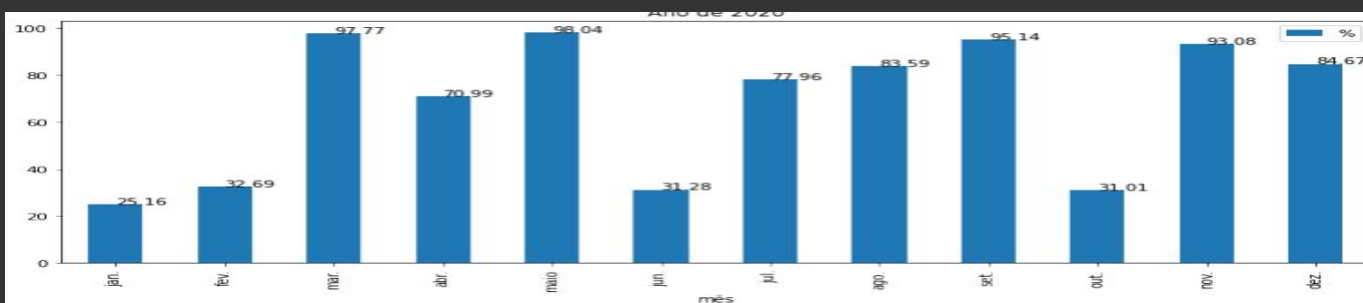


Figura 12 – Frequência dos tópicos predominantes de cada mês de 2020.

Resultados e discussões

● 2021

| Mês | Palavras | Assunto |
|------|---|---|
| jan. | haiti, brasileiro, grande, povo, independencia, ficar, primeiro, cuba, militar, governo | independência haitiana |
| fev. | haiti, brasil, governo, presidente, eua, ficar, contra, bolsonaro, grande, tentar | violação constitucional |
| mar. | haiti, brasil, governo, brasileiro, ficar, eua, grande, mundo, cuba, venezuela | interferência política |
| abr. | haiti, brasil, brasileiro, ficar, povo, pais, eua, cuba, missao, deixar | intervenção militar |
| maio | haiti, revolucao, grande, povo, ajudar, general, passar, matar, pensar, historia | revolução haitiana |
| jun. | haiti, brasil, contra, eua, grande, brasileiro, gente, revolucao, povo, chegar | imigração e xenofobia |
| jul. | presidente, brasil, assassinar, jovenel_moise, brasileiro, matar_presidente, politico, bolsonaro, casa, matar | assassinato do presidente |
| ago. | terremoto, morto, passar, sofrer, grande, situacao, numero_morto, taliba, noticia, tristeza | terremoto |
| set. | haiti, presidente, ajudar, terremoto, grande, pessoa, gente, venezuela, deixar, passar | assassinato do presidente e terremoto |
| out. | haiti, brasil, eua, governo, pessoa, presidente, venezuela, americano, mundo, povo | imigração e refúgio |
| nov. | haiti, brasil, brasileiro, grande, mundo, ficar, povo, lula, negro, pais | imigração e refúgio |
| dez. | haitiano, brasil, lula, brasileiro, governo, presidente, mundo, cuba, revolucao, acontecer | lembração do atrevimento haitiano de 2020 |

Tabela 7 – Tópico predominante de cada mês de 2021.

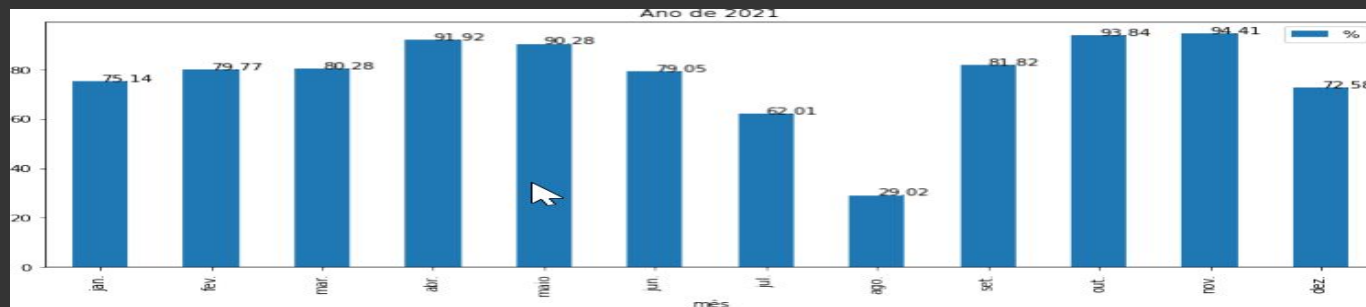


Figura 13 – Frequência dos tópicos predominantes de cada mês de 2021.

Resultados e discussões

● 2022

| Mês | Palavras | Assunto |
|------|---|------------------------------------|
| jan. | haiti, brasil, eua, presidente, ficar, pessoa, mundo, terremoto, governo, independencia | comparação sociopolítico-econômica |
| fev. | brasil, eua, cuba, guerra, mundo, pais, venezuela, lembrar, congo, ucrania | racismo |
| mar. | haiti, brasil, guerra, brasileiro, eua, cuba, ucrania, pessoa, grande, gente | racismo |
| abr. | haiti, brasileiro, ficar, pessoa, povo, coisa, passar, venezuela, conseguir, servir | racismo |
| maio | brasil, venezuela, bolsonaro, cuba, pais, mau, franca, gente, grande, governo | comparação sociopolítico-econômica |



Tabela 8 – Tópico predominante de cada mês de 2022.

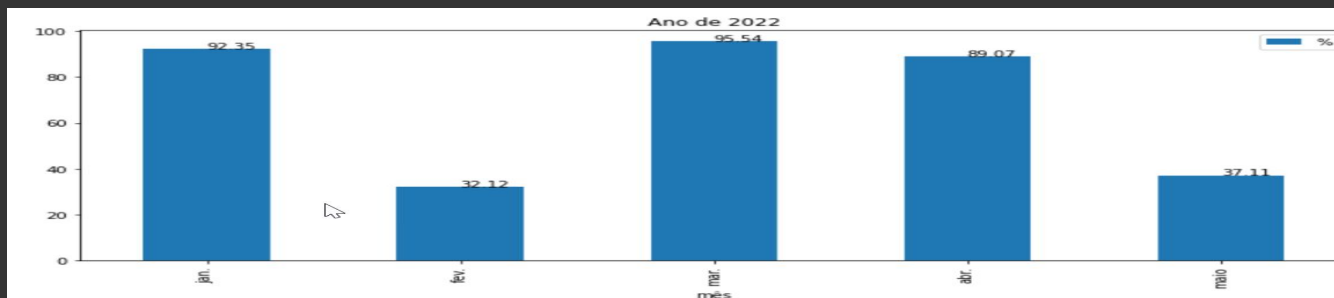


Figura 14 – Frequência dos tópicos predominantes dos cinco meses de 2022.

Conclusão

- **Contribuições**
 - **Concepção dos Internautas**
 - **Base de estudo sobre o Haiti**
- **Discussões no Twitter x Fatos**

Conclusão

- **Trabalhos futuros**
 - **Rotular todos os tópicos**
 - **Analisar a evolução dos assuntos**



Utilização de modelagem de tópicos para identificar os assuntos mais discutidos sobre o Haiti no Twitter

Obrigado!

Créditos

Orientação

Denio Duarte
duarte@uffrs.edu.br

Ícones

mangsaabguru
flaticon.com/authors/mangsaabguru

Freepik
flaticon.com/authors/freepik

Smashicons
flaticon.com/authors/smashicons

Icons made by mangsaabguru, Freepik and Smashicons
from www.flaticon.com