

Eleições e mídia digitais:

explorando uma possibilidade de análise

Erlene Avelino
Louise Almeida da Silva
Pedro Luz de Castro
Renan Almeida

Dada as grandes possibilidades de audiência, interação e mobilização das redes sociais, como o twitter, candidatos tem sido cada vez mais atraídos a utilização dessas ferramentas. É possível, por sua vez, observar indícios de suas estratégias políticas em relação aos usuários como eleitores, através do registro das suas postagens. Nesse sentido, propomos aqui observar o comportamento dos principais candidatos à presidência das eleições de 2018, através do efeito do período eleitoral sob a quantidade de postagens e dos principais tópicos abordados.

- ❶ como o período eleitoral impactou o comportamento online dos candidatos à presidência?
- ❷ Há mudanças significativas na frequência de uso e no que é abordado pelos candidatos em relação a datas mais distantes do pleito?

A nossa proposta se insere num corpo de investigações mais amplo, que procura discutir a relação entre o comportamento dos candidatos e os resultados eleitorais (BRAGA e CARLOMAGNO, 2018). Sabemos que a importância das mídias e redes sociais é hoje um dado, mas o seu significado, as mudanças que provoca, e a maneira como ela é de fato incorporada pelos processos políticos precisam ainda ser amplamente discutidas e investigadas. Entendemos, que a partir de estudos empíricos, que levem em conta o uso que os próprios candidatos fazem destas ferramentas, poderemos ter melhores indícios das tendências e conformação desse fenômeno.

- Os dados foram coletados por meio do pacote de Python `getoldtweets`;
- coletamos todos os tuítes dos candidatos: Fernando Haddad, Ciro Gomes, Geraldo Alckimim, Guilherme Boulos, Jair Bolsonaro, Luis Inácio Lula da Silva e Marina Silva, durante o período de 01/01/2018 à 01/01/2020;
- Utilizamos a técnica Bag-of-words (saco de palavras), para extrair as informações do tuítes;
- realizamos um processo de tokenização, onde foi possível vetorizar o texto. Cada frase é agora um vetor e temos as informações do número de palavras e a frequência de cada uma nas frases.
- Após realizar esse processo, fizemos uma limpeza dos dados, removendo palavras repetitivas (“stopwords”);
- Em seguida, calculamos uma matriz de frequência de documentos – DFM, combinando os vetores em uma única matriz dados;
- Aplicamos uma modelagem probabilística de tópicos, o

- Uma vantagem da abordagem do modelo de tópico estrutural (STM) que utilizamos, é que podemos ter um resumo dos dados textuais, algo que seria impossível de se obter manualmente. Por sua vez, essa também é uma ferramenta muito útil para fazer classificações, quando não conhecemos as categorias;
- Contudo, ressaltamos que análises mais aprofundadas devem buscar métodos mais robustos de validação das categorias. Uma alternativa muito interessante foi desenvolvida por Quinn (2010), que propõe associar a distribuição dos tópicos ao longo do tempo com eventos externos;
- Algo muito evidente nos dados que analisamos é que a frequência de postagens dos presidentiáveis aumenta substantivamente durante o período eleitoral. Pudemos observar também exemplos de mudança interessantes nos tópicos abordados pelos candidatos Fernando Haddad e Jair Bolsonaro;
- Na literatura, esse é um recorte empírico interessante, tanto