

Aluna: Marília Melo Favalesso

Resenha crítica do artigo:

Yousefinaghani, S., Dara, R., Poljak, Z. et al. The Assessment of Twitter's Potential for Outbreak Detection: Avian Influenza Case Study. Scientific Reports 9, 18147 (2019) doi:10.1038/s41598-019-54388-4.

Área: Ciência de dados; Epidemiologia.

A raspagem de dados (do inglês “data scraping”) é uma técnica utilizada para extrair informações de uma interface moldada para leitura do ser humano. Na área da saúde, principalmente na epidemiologia, essa técnica vem sendo aplicada para a extração de informações de documentos que não estão organizados de maneira estrutura (por exemplo: estão em PDF e não em um arquivo .csv e precisam ser extraídos para uso) ou monitoramento de enfermidades via postagens em redes sociais. Segundo a Organização Mundial da Saúde, a vigilância em saúde é definida como a coleta, análise e interpretação contínua e sistemática de dados relacionados a saúde. Neste aspecto, a vigilância de doenças via redes sociais pode servir como ferramenta de alerta precoce para surto de doenças. Esses sistemas podem ser utilizados para o planejamento, implementação e avaliação de práticas na área da epidemiologia. Isso ocorre porque as informações podem ser transmitidas pelas redes sociais mais rapidamente do que os canais tradicionais (exemplo, os sistemas de informações governamentais), ajudando na oportunização de tomadas de decisão.

Diante do exposto, Yousefinaghani et al. (2019) construíram um sistema de vigilância epidemiológica no Twitter com o objetivo de avaliar a possibilidade de utilizar a rede social como fonte complementar de relatórios oficiais de gripe aviária (GA) amostrados pelo OIE (“World Organization For Animal Health”). Os autores desenvolveram uma estrutura de análise de dados baseada no Twitter para monitorar automaticamente os surtos de GA em tempo real, além de criar modelos para a previsão da doença (figura 1). A estrutura foi implementada para encontrar publicações preocupantes na rede social, além de filtrar publicações irrelevantes e detectar o aparecimento de surtos em vários países. Através das técnicas de raspagem de dados, o sistema coletou e analisou mais de 209 mil postagens entre o período de julho de 2017 e novembro de 2018. Os autores examinaram o potencial dos dados

Aluna: Marília Melo Favalesso

do Twitter para representar a data, a gravidade e o tipo de vírus dos relatórios oficiais da OIE através de “modelo de aprendizado de sentimentos semi-supervisionado”.

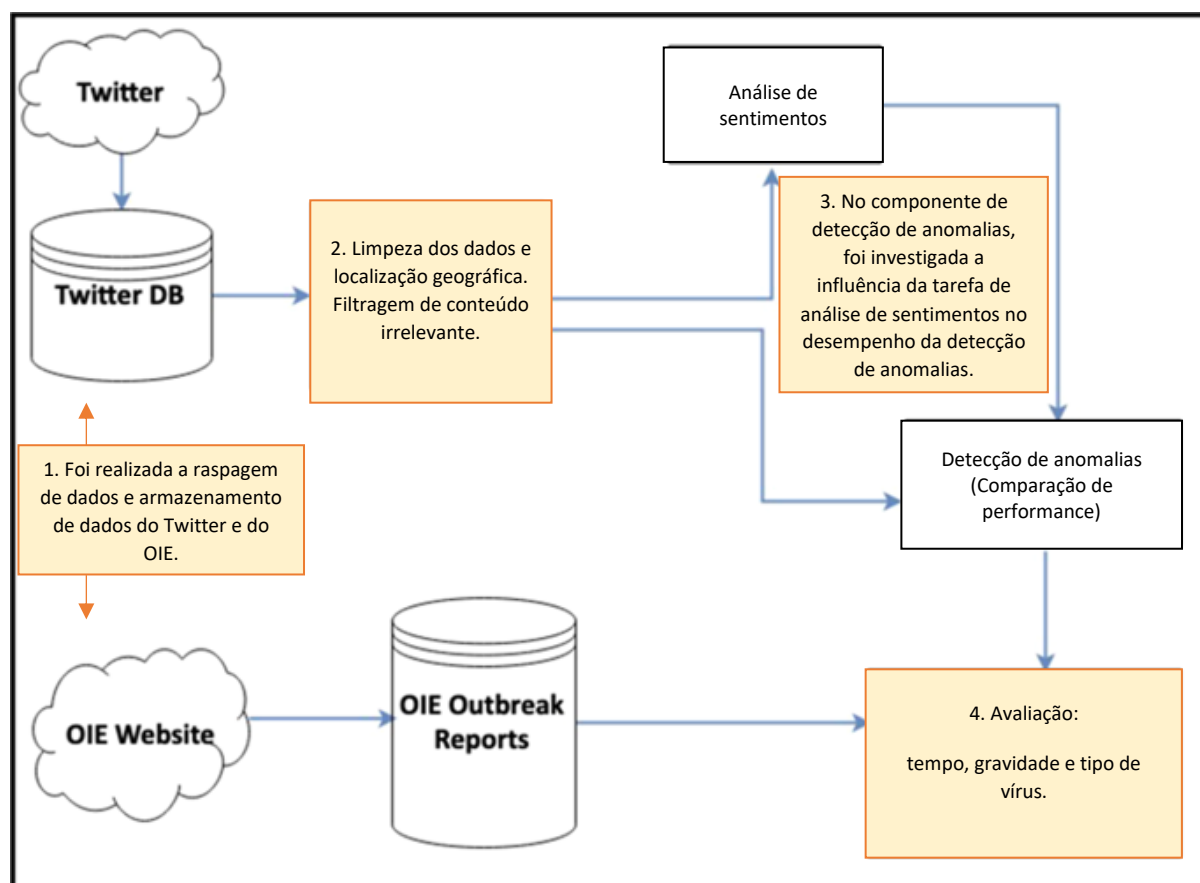


Figura 1. Arquitetura do sistema proposto para amostragem de dados de gripe aviária: os dados são coletados do Twitter (apenas conteúdo recente em inglês, seguindo um rol de palavras-chave pré-definidos por profissionais da saúde) e da OIE (“*World Organization For Animal Health*”). Após o pré-processamento, as anomalias são detectadas e avaliadas.

A ferramenta criada por Yousefinaghani et al. (2019), através de técnicas de raspagem de dados e modelagem, conseguiu identificar 75% das notificações de surto de GA do mundo real, segundo dados da OIE, a partir de postagens na rede social Twitter. Esse resultado demonstra a capacidade do sistema de servir como uma abordagem complementar aos métodos tradicionais de informação epidemiológica. Os autores relataram que conseguiram identificar

Aluna: Marília Melo Favalesso

um terço (1/3) das notificações de surto a partir das postagens (“tweets”) antes dos relatórios oficiais. Esse recurso pode auxiliar os sistemas de vigilância tradicionais fornecendo a possibilidade de detecção precoce de surtos de GA. Precisamente, o sistema obteve os mesmos resultados relativos ao tempo, gravidade e subtipo de vírus que os casos informados nos relatórios oficiais. No entanto, a ferramenta não conseguiu identificar a patogenicidade do vírus, deixando uma brecha para estudos futuros relacionados com a rede social Twitter. Além do colocado, o estudo pode servir como base para o monitoramento de outras doenças, ao mesmo tempo que utilizar a metodologia dos autores com outros patógenos pode auxiliar no aprimoramento dessa ferramenta.