

Uma modelagem computacional para análise de dados sobre Fake News

Giovanna Dias Cardoso¹, Luiz Henrique da Cruz Silvestrini²

Departamento de Matemática, FC/UNESP, Bauru, SP

As fake news, ou desordem informativa, são notícias falsas, conteúdos inventados ou descontextualizados, com a intenção de parecerem notícias verídicas, que são difundidas amplamente para influenciar um determinado público, fazendo com que divulguem aquele conteúdo que foi recebido. Uma fake news possui 70% mais chances de ser propagada do que notícias verídicas (Vosoughi, Roy e Aral [1]), pois o compartilhamento de informações falsas possui mais velocidade e alcance do que fatos e quando o tema é política, as notícias falsas se propagam três vezes mais rápidas.

Não é fácil dar uma caracterização completa das fake news, pois estas podem conter informações verdadeiras, manipuladas para enganar, ou um conteúdo verídico, mas compartilhado em um contexto falso. Especialistas defendem que uma das maneiras de combate às fake news é por meio do desenvolvimentos de habilidades voltadas ao pensamento crítico. Nesse sentido, a famosa “caixa preta” de cada indivíduo, segundo Barros e Bassanezi [2], precisa receber novos parâmetros para levar o indivíduo a tomar boas decisões com base nas interpretações de tais informações recebidas.

A teoria dos conjuntos fuzzy, proposta na década de 60 por Lotfi Zadeh [3], é uma teoria alternativa a dos conjuntos usuais da matemática e menos rígida do que o habitual. Em sua proposta, a mudança da pertinência para não pertinência de um elemento do universo em um conjunto, acontece de forma gradual e não abrupta. Seja X um universo de discurso, um conjunto fuzzy A em X é definido por uma função μ_A , tal que $[0, 1]$ é um intervalo de números reais e a função μ_A indica o grau de pertinência de x no conjunto fuzzy A . Quanto mais próximo de 1, maior é o grau de pertinência. Uma função de um subconjunto fuzzy pode assumir como contradomínio estruturas mais gerais, como reticulados, semigrupos, entre outros. Podemos indicar um subconjunto fuzzy A de X do seguinte modo:

$$A = \{(x, \mu_A(x)) : x \in X\} \quad \text{em que} \quad \mu_A : X \rightarrow [0, 1]$$

O uso da teoria dos conjuntos fuzzy nos possibilita incorporar uma certa imprecisão a um problema, a um sistema. Nesta teoria podemos fazer uso de variáveis reais como: temperatura, pressão, velocidade, entre outras, e a essas variáveis são associados termos como: alto, baixo, pouco baixo. Na lógica fuzzy, a expressão “ x é F ”, é uma relação que associa cada elemento x do domínio V a um termo predicado fuzzy F , o qual tem seu valor dado por um conjunto fuzzy, com universo de discurso X , possivelmente distinto de V , tal expressão é chamada proposição fuzzy.

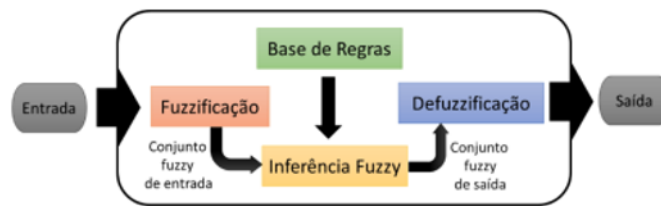
Com a utilização dos conjuntos fuzzy podemos interpretar fenômenos imprecisos e não quantitativos e buscar mecanismos de inferências através desses dados. Um sistema fuzzy é qualquer sistema que faz uso da teoria dos conjuntos fuzzy para representar suas variáveis e suas interações. Neste trabalho, abordaremos os sistemas de inferência baseados em regras fuzzy (SBRF). As partes essenciais de um SBRF são a base de conhecimento e o seu mecanismo de inferência.

Um SBRF consiste em quatro módulos conectados como na Figura 1.

¹ giovanna.dias@unesp.br

² lh.silvestrini@unesp.br

Figura 1: Sistema Baseado em Regras Fuzzy



Fonte: Oficina “Conjuntos nebulosos, notícias falsas e modelagem computacional” promovida pelo Grupo Interdisciplinar de Divulgação e Educação em Lógica e Inteligência Artificial - IdeIA+ (2021).

A Base de Regras descreve um conjunto de regras fuzzy associadas ao problema em função das variáveis e dos termos linguísticos da base de dados. Ela desempenha um papel chave nos SBRF, pois é por meio das regras que o conhecimento é representado no sistema. É a etapa em que são estabelecidas regras do tipo “*se um conjunto de condições são satisfeitas, então um conjunto de consequências podem ser inferidas*”.

Para a presente pesquisa, utilizaremos o método Mamdani, introduzido por Mamdani e Assilian em 1975. Seus passos básicos compreendem a *fuzzificação*, ou seja, o momento em que as entradas, mesmo que sejam valores precisos, são mapeadas pelas funções de pertinência estabelecidas durante a construção do sistema. Em seguida, temos a *Inferência fuzzy*, em que cada proposição fuzzy é traduzida matematicamente através de técnicas da lógica fuzzy, a qual produzirá uma saída fuzzy, a partir de cada entrada. E por fim, temos a *defuzzificação*, em que o valor fuzzy, que foi obtido como resposta, passa por um processo para obtermos novamente uma saída precisa (crisp).

Nesta investigação sobre o funcionamento e compartilhamento das fake news, a modelagem será dada por meio da extração de conhecimento diretamente do especialista de domínio. A abordagem manterá o especialista no controle das decisões nos níveis linguísticos e semântico, além das definições das regras. Optamos por priorizar a interoperabilidade do modelo, o que a participação do especialista promove benefícios, em detrimento à extração exclusiva dos dados, seja na promoção de uma semântica correta, uma vez que adotaremos as regras com os consequentes também formados por termos linguísticos, ou de Mamdani, seja nos julgamentos sobre o comportamento das saídas inferidas, o que acreditamos ser crucial na aplicação do modelo resultante.

Ademais, buscamos a construção de um SBRF a fim de se estabelecer uma modelagem adequada para fornecer o quão suscetível no compartilhamento de desinformação alguém pode ser, a partir das variáveis de entrada do sistema.

Agradecimentos

Agradecemos à Unesp pela bolsa PIBIC concedida.

Referências

- [1] S. Vosoughi; D. Roy; S. Aral. “The spread of true and false news online”. Em: **Science** 6380 (2018), pp. 1146–1151.
- [2] L. Barros; R. C. Bassanezi. **Tópicos de Lógica Fuzzy e Biomatemática**. Vol. 5. Coleção IMECC. Textos Didáticos. Campinas: Unicamp, 2006.
- [3] L. A. Zadeh. “Fuzzy Sets”. Em: **Information and Control** 8 (1965), pp. 338–353.