

Projeto Final: Detector de Fake News

Allan de M. Silva¹, Eliaquim dos S. Costa²

Bacharelado Interdisciplinar em Tecnologia da Informação – UFRN - Brasil

{allandemiranda, eliaquimdossantos}@gmail.com

Abstract. *Fake News (...) is literally the spreading of false news through social networks that, with the popularization of the Internet, can have negative impacts on society (CELI, 2019). To avoid such problems, it is necessary to create a news verification tool. The project is willing to present a tool capable of identifying such problems.*

Resumo. *Fake News (...) trata-se literalmente da divulgação de notícias falsas por meio das redes sociais que, com a popularização da internet, podem causar impactos negativos na sociedade (CELI, 2019). Para evitar tais problemas, é necessário a criação de uma ferramenta de verificação de notícias. O projeto está disposto a apresentar uma ferramenta capaz de identificar tal problemática.*

1. Introdução

Em diversos momentos da história da humanidade pode-se observar a divulgação de informações falsas tidas como verdade (CELI, 2019). Contudo, esse fenômeno conhecido como fake news vem ganhando força após a popularização da internet e o domínio das mídias sociais (CELI, 2019).

“As fake news são um perigo para a população, sendo eles desde problemas mais imediatos, como acusações falsas sobre alguém — e que podem causar algum dano físico ou psicológico —, como problemas mais amplos, como influenciar a vitória de um candidato em uma corrida presidencial.” (CELI, 2019)

Assim, para que o erro não acabe se tornando um grande pesadelo, o nosso projeto conta com a elaboração de um software, desenvolvido em java, para que usuários possam verificar se aquela notícia é falsa, ou o quanto falsa ela é possível ser, utilizando banco de dados de notícias falsas e algoritmos de similaridade.

2. Descrição

2.1. Funcionamento do software

O programa baseia-se na comparação entre a notícia inserida pelo usuário para verificação e notícias falsas que estão armazenadas em uma determinada base de dados.

Por meio deste modelo, foi aplicado às seguintes regras na notícia (contidas na classe `HandlingTextController`):

- Remover palavras com menos de 3 caracteres;
- Remover acentuação
- Remover pontuações
- Remover caracteres especiais
- Colocar todas as palavras em minúsculo;
- Eliminar palavras repetidas;
- Colocar palavras em ordem alfabética;

Para comparar quantidade de palavras semelhantes e traçar sua interação a fim de obter uma similaridade entre a notícia do usuário e as milhares de notícias falsas na base de dados do sistema, é exigido um custo de processamento relativamente alto, e isto pode requerer muito tempo. Em criptografia, SHA-1 é uma função de dispersão criptográfica (ou função hash criptográfica) projetada pela Agência de Segurança Nacional dos Estados Unidos e é um Padrão Federal de Processamento de Informação dos Estados Unidos publicado pelo Instituto Nacional de Padrões e Tecnologia (NIST) (FIPS, 2012). SHA-1 produz um valor de dispersão de 160 bits (20 bytes) conhecido como resumo da mensagem. Um valor de dispersão SHA-1 é normalmente tratado como um número hexadecimal de 40 dígitos (Wikipedia, 2019). A fim de diminuir o impacto de processamento, existe uma proposta de obter um texto tratado e aplicar uma função de hash sobre este texto (SHA-1), para que seja realizada a comparação dos hashes da base dados com o da notícia que foi gerado, caso o valor do hash seja igual, uma notícia falsa será detectada sem o alto custo dos algoritmos de semelhança de strings, como podemos identificar, na classe abstrata `NewsController` do projeto, a criação da hash a partir de um parágrafo e depois seu uso de comparação na classe `SimilaritySystemController` no método `startTests`.

Com a notícia informada pelo usuário, já pré tratada e explicada anteriormente, iremos fazer a análise de similaridade com as demais notícias falsas do sistema. Nesta proposta de solução, o sistema é configurado para trabalhar com os seguintes algoritmos de semelhança entre strings:

- Trigram:

“Algoritmo Trigram (Dunning, T., 1994): no campo da linguística computacional e da probabilidade, um n-grama é uma sequência contígua de n itens de uma dada amostra de texto ou fala. Um algoritmo de Trigram é um caso de n-grama, uma sequência contígua de n (três, neste caso) itens de uma dada amostra.”

- Distância Levenshtein:

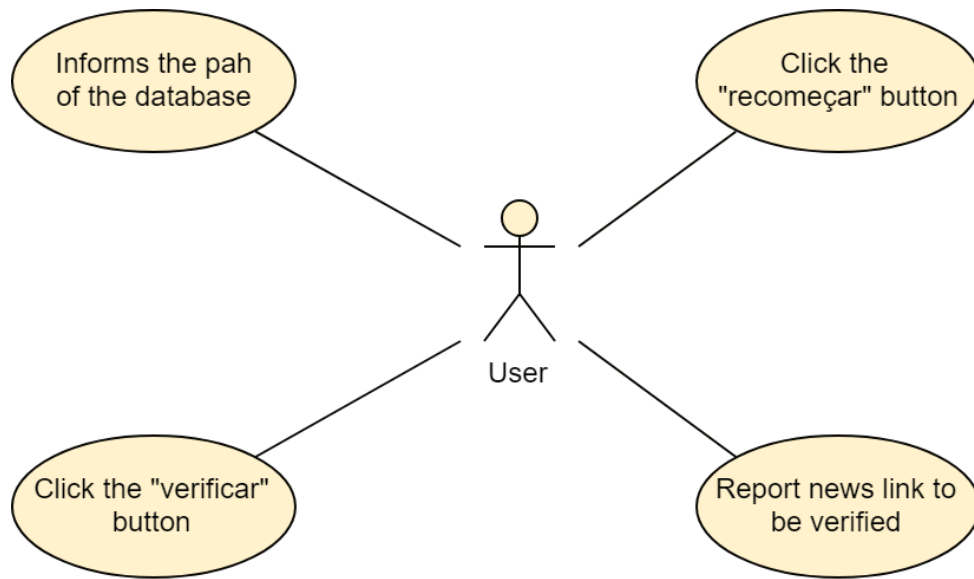
“Distância Levenshtein (Navarro, G., 2001; Wagner, Robert A. and Fischer, Michael J., 1974): em teoria da informação, a distância Levenshtein ou distância de edição entre duas strings (duas sequências de caracteres) é dada pelo número mínimo de operações necessárias para transformar uma string na outra. Entendemos por "operações" a inserção, deleção ou substituição de um carácter. O nome advém do cientista russo Vladimir Levenshtein, que considerou esta distância já em 1965. É muito útil para aplicações que precisam determinar quão semelhantes duas strings são.”

Para a coleta das informações de notícias para poder verificá-las se são ou não *fake news*, foi utilizada uma técnica denominada *Web Scrapping*. Por meio de processos automatizados, implementados usando um rastreador *bot*, esse tipo de “raspagem” de informações é uma forma de realizar cópias de dados em que informações específicas são coletadas e copiadas da *web*, tipicamente em um banco de dados ou planilha local central, para posterior recuperação ou análise (Westcon, 2019). A partir do *link* da notícia, é possível coletar as informações, onde, na implementação desta técnica, o *software* foi configurado para tratar a página HTML (HyperText Markup Language) buscando informações de parágrafos – indicados no HTML com a tag “<p>”, contudo, apenas se tiverem 200 caracteres ou mais, para evitar a coleta de informações que não fazem parte do conteúdo da notícia coletada, como podemos identificar na classe *WebScraping*.

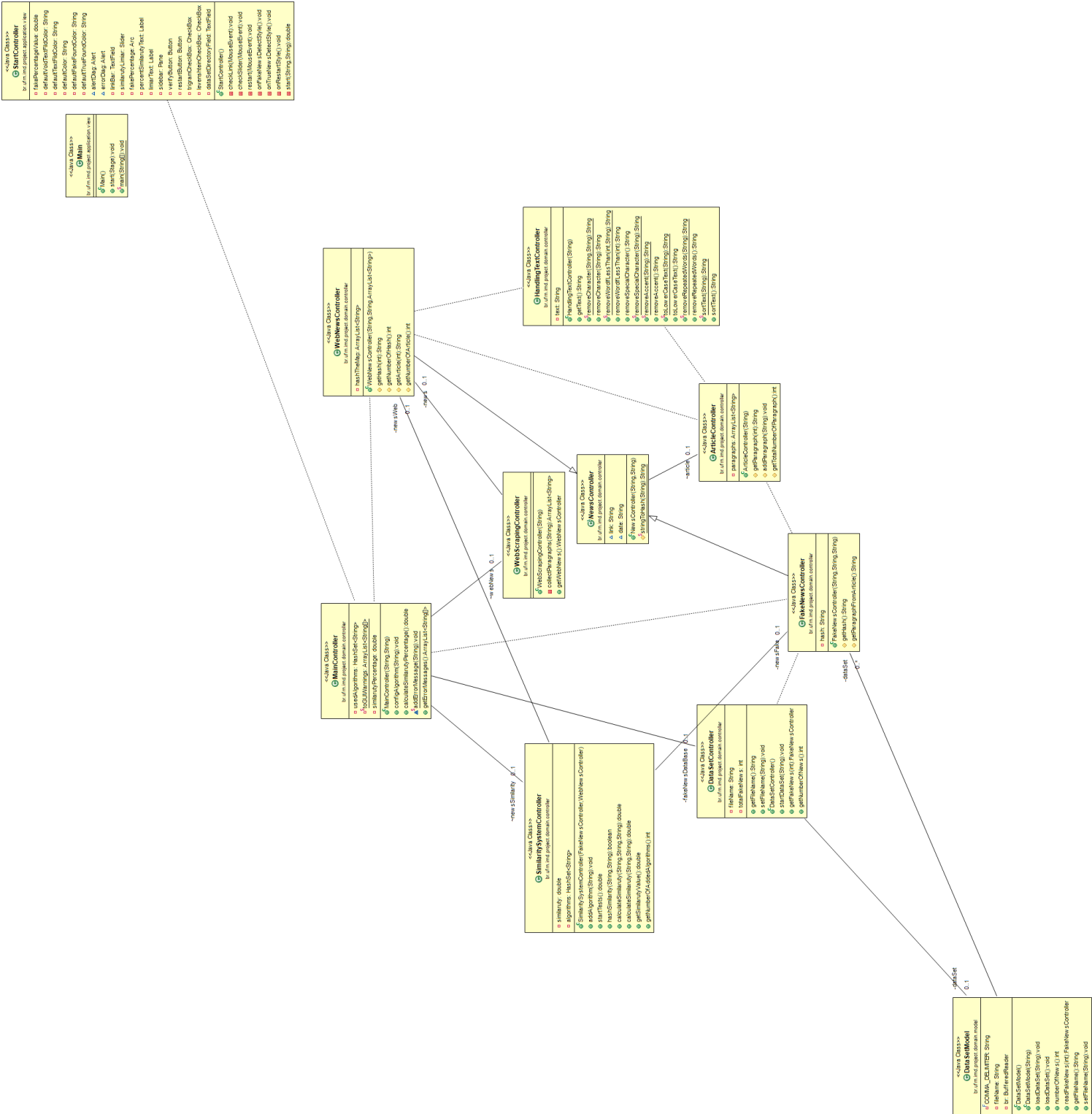
A base de dados a ser utilizada é nativamente um arquivo CSV (*Comma Separated Values*). A base de dados contará com as informações principais de uma notícia: Data, conteúdo e endereço na *web*. O arquivo de dados pode ser especificado através da interface gráfica do sistema.

2.2. Diagramas de estrutura

2.2.1. Diagrama de casos de uso



2.2.2. Diagrama de classes



3. Conclusão

Para que o propósito de uma ferramenta de detecção de *Fake News* seja atingido com perfeição, e seguindo a lógica do projeto desenvolvido, temos de adquirir um banco de dados tipo “*real time*” com notícias já taxadas falsas, o que não é uma tarefa tão simples, visto que identificar se algo é falso ou não pode ser bastante subjetivo, incluindo-se o fator humano (personalidade, caráter, ideologia, política, etc.). Além desta existem também a dificuldade da coleta de informações de sites de notícias. Por não seguirem o padrão de *tags* HTML, é possível que textos que não fazem parte da notícia gere ruídos nas análises. O aperfeiçoamento destes dois pontos precisam ser aprimorados.

Concluimos que é possível, através de algoritmo de similaridade, obter resultado viável de similaridade e detectar o quanto ela é parecida com notícias falsas já identificadas e armazenadas em um banco de dados próprio, sendo assim possível distinguir se ela é falsa ou não (ainda precisa-se incluir o fator humano na decisão final). Porém é necessário que nos aprofundemos cientificamente neste tipo de análise, tanto buscando soluções em TI (para similaridade textual) quando em Psicologia (visto que isso é um dilema social de natureza psíquico), para não taxarmos uma notícia de falsa, simplesmente pela sua similaridade com uma outra já identificada falsa, caído no erro e críticas semelhantes ao Sistema Compas (Maybin, 2016) dos Estados Unidos.

Referencias

- CELI, Renata. Fake news: o que é, consequências e redação!. Stood. Disponível em: <<https://www.stoodi.com.br/blog/2019/01/03/fake-news-o-que-e/>>. Acesso em: 23 mai. 2019.
- Ted Dunning. Statistical Identification of Language. Technical Report MCCA. New Mexico State University: 94–273, 1994.
- FIPS. FEDERAL INFORMATION PROCESSING STANDARDS PUBLICATION. PUB 180-4. Secure Hash Standard (SHS). Mar. 2012. Disponível em: <<https://csrc.nist.gov/csrc/media/publications/fips/180/4/final/documents/fips180-4-draft-aug2014.pdf>>. Acesso em: 01 jun. 2019.
- Maybin, S. Sistema de algoritmo que determina pena de condenados cria polêmica nos EUA. BBC Brasil. 31 out. 2016. Disponível em: <<https://www.bbc.com/portuguese/brasil-37677421>>. Acesso em: 06 jun. 2019.
- Navarro, Gonzalo. A guided tour to approximate string matching. ACM Computing Surveys. 33 (1): 31–88, 2001.
- Westcon, Blog Brasil. O que é web scraping?. Disponível em: <<https://blogbrasil.westcon.com/o-que-e-web-scraping>>. Acesso em: 01 jun. 2019.
- Wikipedia. SHA-1. Disponível em: <<https://pt.wikipedia.org/wiki/SHA-1>>. Acesso em: 01 jun. 2019.