

# **2º Encontro da SBPC em MS/ XI ENEPEX / XIX ENEPE/ 22ª SNCT - UEMS / UFGD 2025**

## **SOLUÇÕES DE ENGENHARIA DE SOFTWARE PARA IDENTIFICAÇÃO DE FAKE NEWS**

**Instituição:** UEMS – Universidade Estadual do Mato Grosso do Sul

**Área temática:** Ciências Exatas e da Terra

**Mello, Lucas Sarri<sup>1</sup>** ([rgm43975@comp.uems.br](mailto:rgm43975@comp.uems.br)); **VASCONCELOS, Ana Júlia Furtado<sup>2</sup>** ([rgm51696@comp.uems.br](mailto:rgm51696@comp.uems.br)); **PRATES, Jorge Marques<sup>3</sup>** ([terceiro.autor@uems.br](mailto:terceiro.autor@uems.br));

<sup>1</sup> – Discente do Curso de Sistemas de Informação - Dourados;

<sup>2</sup> – Discente do Curso de Sistemas de Informação - Dourados;

<sup>3</sup> – Docente do Curso de Sistemas de Informação - Dourados;

Devido ao crescente acesso às redes sociais, a divulgação de notícias falsas se agravou, como em períodos críticos como a eleição presidencial brasileira de 2018 e americana em 2016, e recentemente no período pandêmico da COVID-19. A solução viável para verificar informações se dá pelo uso da engenharia de software inteligente. A inteligência artificial com suas subcategorias de processamento de linguagem natural (PLN) e análise de sentimento (AS) se torna potencialmente uma resolução ou mitigação da problemática. O PLN tem a capacidade de interpretar e categorizar conteúdos textuais do cotidiano de forma automatizada, enquanto a AS trata de identificar a polaridade do texto, se tornando tangível visto que diz respeito a opinião e sensação daquele que transmite desinformação. Diante desse contexto, a pesquisa tem como objetivo explorar aplicações da análise de sentimento na detecção de fake news, utilizando técnicas de engenharia de software, mapeando o cenário atual das implementações e suas lacunas para estudos futuros. Na primeira etapa foram utilizados textos fundamentais para o tema de inteligência artificial para compreender os conceitos gerais. A partir disso, foram considerados trabalhos com implementações da tecnologia de PLN e técnicas de AS para identificar notícias falsas. No período final, foi analisada a execução de um algoritmo em linguagem python com a função de classificar notícias, com o entendimento das bibliotecas e funções disponíveis. Sobre os estudos examinados, foram identificados processos correspondentes na implementação da abordagem de PLN. Pode-se destacar, o pré-processamento de dados corresponde a tratar os dados para sua análise posterior, são consideradas algumas técnicas como: tokenização, remoção de stop-words, análise de sequência, POS-tags. Em sequência, há o processo de representação de palavras, que pode ser feito por meio de vetorização, alocação em matrizes de valores binários, ou vetores compactos, ou seja, embeddings. As diferentes formas são aplicadas de acordo com a necessidade, por exemplo, uma representação pode armazenar ou não informações semânticas das palavras. Os algoritmos voltados para PLN mais abordados e consideráveis foram: Term Frequency-Inverse Document Frequency, Bag of Words, baseados em frequência; com base em word embeddings: Word2Vec e FastText; SkipThought e Universal Sentence Encoder (USE) embasado por sentence embeddings; por inferência de linguagem natural: InferSent; SentenceBERT (SBERT) baseado na arquitetura Transformer e sentence embeddings. As principais considerações sobre AS foram: pode ser classificado por níveis de frase, documento ou aspecto; seus métodos são baseados em léxicos, como EmoLex e o LIWC, aprendizado de máquina, como Perspective API, e híbridos. Além disso, foram exploradas diversas facetas da engenharia de software inteligente, incluindo mecanismos de atenção, aprendizagem profunda e redes neurais. A dificuldade em destaque, além de manusear dados textuais, é provar a veracidade dos algoritmos e testá-los, retratando a relevância dos testes metamórficos, já que a identificação de notícias falsas não tem uma resposta exata, e assim são observadas as relações de entrada e saída, as relações metamórficas. Em geral, os estudos consultados relatam que esta área é promissora e há novas direções que podem orientar trabalhos futuros, com muitas lacunas a serem resolvidas.

**PALAVRAS-CHAVE:** Processamento de Linguagem Natural, Análise de Sentimento, *Fake News*.

**AGRADECIMENTOS:** Agradecemos pelos apoio da Universidade Estadual do Mato Grosso do Sul (UEMS)