

ABORDAGENS DE MACHINE LEARNING PARA ANÁLISE DE SENTIMENTOS EM COMENTÁRIOS DE APLICATIVOS

Kalidsa B. de Oliveira¹(IC); Gabriel M. Lunardi¹(O); Vitor X. Siqueira²; Williamson Silva²

¹ Departamento de Eletrônica e Computação, Universidade Federal de Santa Maria (UFSM);

² Universidade Federal do Pampa (UNIPAMPA);

A análise de sentimentos em comentários de aplicativos atualmente é necessária para a compreensão da percepção dos usuários. A precisão na interpretação desses *feedbacks* permite que as empresas ajustem suas estratégias de mercado, resultando em impactos econômicos significativos, como aumento no número de *downloads*, fidelização do cliente e otimização contínua do produto. Nesse sentido, este trabalho utiliza diferentes técnicas de aprendizado de máquina para automatizar a classificação de sentimentos e polaridade em 3.000 comentários de usuários de 10 aplicativos populares da *Google Play Store*. Com esse propósito, foram empregadas abordagens clássicas de estruturação de textos, como *Bag of Words* (BOW) e TF-IDF, em conjunto com classificadores como *Naive Bayes*, *K-Nearest Neighbors* (KNN) e regressão logística. Além dessas abordagens clássicas, também foram utilizadas técnicas mais recentes, como o Word2Vec e o BERT (*Bidirectional Encoder Representations from Transformers*). O objetivo principal deste trabalho é comparar a eficácia das estratégias tradicionais com essas abordagens mais recentes, avaliando como os modelos de linguagem podem superar os métodos clássicos na análise de sentimentos. O BERT, em especial, foi utilizado como representante dessa nova geração de modelos de linguagem. Os melhores resultados foram obtidos da associação de Word2Vec com Regressão logística, que capta relações semânticas entre palavras por meio de representações vetoriais, alcançando uma acurácia de 61,5% na predição de sentimentos e 82,4% na classificação de polaridade. Já o BERTimbau, uma versão do BERT pré-treinada para português, atingiu resultados superiores, atingindo 71,6% na predição de sentimentos e 87,4% na classificação de polaridade. A comparação entre os modelos mostra que o BERTimbau foi mais eficaz na captura de nuances semânticas e na análise de textos complexos, especialmente em contextos que exigem maior sensibilidade, como comentários de usuários. O estudo destaca, assim, a importância dos métodos de aprendizado profundo na interpretação de opiniões em plataformas digitais.