



# CONEPE 2021

## 8.º CONGRESSO DE ENSINO, PESQUISA E EXTENSÃO

ENSINO, SAÚDE E MEIO AMBIENTE: O IMPACTO DAS INOVAÇÕES TECNOLÓGICAS

*de 22 a 26 de novembro de 2021*



ISSN 2525-975X

## **MINERAÇÃO DE DADOS PARA O AUXÍLIO DE IDENTIFICAÇÃO DE POSSÍVEIS NOVOS CASOS DE COVID-19**

R.S.L. de Carvalho<sup>1\*</sup>, L. Evangelista<sup>1</sup>, R. Colpani<sup>1</sup>

<sup>1</sup>Centro Universitário das Faculdades Associadas

\*[renan.carvalho@sou.fae.br](mailto:renan.carvalho@sou.fae.br)

A mineração de dados é uma técnica que pode ser usada para se obter informações e padrões importantes a partir de dados, e se mostra cada vez mais necessária nas mais diversas áreas de atuação, inclusive na de saúde, onde foram desenvolvidos modelos de mineração de dados para ajudar na recuperação da COVID-19. Assim, este trabalho tem como objetivo propor um modelo de mineração de dados para auxiliar na identificação de novos casos de COVID-19 com base nos sintomas. Para isso, foi utilizada a base de dados disponibilizada pelo openDataSUS, que contém informações da triagem de pacientes com síndrome respiratória, incluindo COVID-19, aplicando a metodologia Knowledge Discovery in Databases (KDD), que se refere no processo de extrair conhecimento a partir de uma base de dados, juntamente dos algoritmos K-Nearest Neighbors (KNN) e Árvore de Decisão, que são algoritmos de classificação nos quais é possível classificar as amostras como positivo ou negativo para COVID-19. A metodologia para ambos algoritmos baseou-se em dividir a base de dados em 70% para treinamento e 30% para testagem e, no caso do KNN, foi necessário validar qual seria o melhor valor para K entre 1 até 20. Foi utilizada a linguagem de programação Python com as bibliotecas Pandas e Sklearn. Após concluir todo o processamento, o resultado obtido para o KNN com o melhor valor de K, que foi 5, demonstra que o algoritmo possui uma acurácia de 75.77%, enquanto o algoritmo de Árvore de Decisão, que teve como atributo-alvo a perda de paladar, apresentou 79.87% de acurácia. Portanto, conclui-se que a aplicação da mineração de dados pode contribuir significativamente para auxiliar no diagnóstico de COVID-19. Para trabalhos futuros, pretende-se realizar um estudo de comparação com outros algoritmos de aprendizado de máquina.

**Palavras-chave:** Mineração de dados, Aprendizado de máquina, COVID-19.