

## Aquisição automática de conhecimento sobre o diagnóstico discente do IFRS Campus Veranópolis: uma abordagem baseada em mineração de dados

Adilson de Quadros Coutinho<sup>1</sup>, Gabriel Rottini<sup>1</sup>, Denise Genari<sup>1</sup>, Marcos Vinícios Luft<sup>1</sup>, Roger Sá da Silva<sup>1\*</sup>

<sup>1</sup>Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS) - *Campus Veranópolis*. Veranópolis, RS.

\*Orientador(a)

Apesar dos esforços institucionais, e considerando os reflexos do período de pandemia, sabe-se que os Planos Estratégicos de Permanência e Êxito dos campi do IFRS, de responsabilidade das Comissões Internas de Acompanhamento das Ações de Permanência e Êxito dos Discentes (CIAAPE), estão em construção e podem usufruir de informações e conhecimentos extraídos a partir dos dados já obtidos anualmente nos questionários de inscrição do processo seletivo e de diagnóstico discente para que possam ser implementados na prática. Nesse cenário, este projeto fornecerá à CIAAPE do Campus Veranópolis mais uma possibilidade de obtenção de conhecimentos sobre a realidade educacional dos estudantes do campus, por meio de uma abordagem baseada em mineração de dados. O presente projeto justifica-se, ainda, pois com o crescimento do volume de dados produzidos e armazenados, aumenta a dificuldade do ser humano de analisar, interpretar e extrair informações desses, surgindo a necessidade do uso de ferramentas computacionais que auxiliem nessas tarefas, tal como o processo de Descoberta de Conhecimento em Bases de Dados. Diante desse contexto, o presente projeto de pesquisa tem por objetivo geral descobrir e identificar automaticamente relações implícitas em dados de diagnóstico discente por meio do processo de mineração de dados. Para tanto, o projeto é desenvolvido a partir de um estudo sobre o processo de KDD (Knowledge Discovery in Data Bases), principalmente sobre mineração de dados e suas técnicas e algoritmos, com destaque para algoritmos de regras de associação. Até o momento, partindo dos dados das respostas ao questionário de diagnóstico discente aplicado aos estudantes do Campus Veranópolis no ano de 2022, analisou-se a base de dados para um melhor entendimento dos tipos de dados à disposição. Em sua maioria, são dados categóricos, pois o questionário possui perguntas fechadas com múltiplas opções de escolha para cada uma das mais de 80 questões que refletem a condição socioeconômica dos discentes e seu entorno. Logo após, iniciou-se a etapa de pré-processamento do processo de KDD, com o intuito de formatá-los corretamente em categorias numéricas, para posterior aplicação do algoritmo A Priori. Tal algoritmo é responsável por extrair as regras de associação entre os diversos atributos da base de dados, sendo que cada atributo representa uma das questões do questionário. Dessa forma, o resultado esperado ao final do projeto é a obtenção de modelos de dados de forma automatizada, capazes de fornecer conhecimentos relacionados às associações existentes entre os atributos da base de dados, com vistas a apoiar o desenvolvimento e a implementação de ações de permanência e êxito dos estudantes.

Palavras-chave: Mineração de Dados; Diagnóstico discente; Descoberta de conhecimento; Permanência e êxito.