

O ESTUDO DA EVASÃO NO CURSO DE BACHARELADO EM ENGENHARIA AGRÔNOMICA NO IFTO *CAMPUS* DIANÓPOLIS, APLICANDO TÉCNICAS DE MINERAÇÃO DE DADOS

Isaac Pimentel Fernandes Sobrinho¹, Diego Rodrigues²

¹Estudante do Curso Superior de Licenciatura em Computação – IFTO, *campus* Dianópolis. e-mail: <isaacpimentelf@gmail.com>

²Professor do curso de Licenciatura em Computação – IFTO, *Campus* Dianópolis. e-mail: <diego.rodrigues@ifto.edu.br>

Resumo: Este documento consiste em apresentar breves conceitos sobre Evasão no ensino superior e como as técnicas de mineração de dados podem ajudar no combate à evasão no curso de Bacharelado em Engenharia Agrônoma do Instituto Federal de Educação Ciência e Tecnologia do Tocantins, *campus* Dianópolis, ao identificar os padrões dos alunos evadidos de 2014 a 2018. O desenvolvimento do projeto se fundamentou em pesquisas feitas pelos autores deste artigo. Por meio das regras de associação na linguagem Python, R e na ferramenta WEKA obtiveram resultados relevantes para gestores do *campus*.

Palavras-chave: Mineração de dados, Evasão, Ensino Superior, Agronomia.

1 INTRODUÇÃO

É indiscutível que a educação é o principal e mais significativo fator de uma sociedade, ela é responsável por todas as outras demandas que rodeiam e regem a população. Sem a educação, os grandes avanços e melhorias que chegaram a todas as áreas que se presencia hoje, não seriam possíveis. O acesso à educação é direito de todos os cidadãos brasileiros, é o que consta na Constituição Federal (CF) de 1988, art. 6º, onde a educação é definida como um direito social e um dever do estado que deve ser bem desenvolvido para que não haja danos à soberania nacional. Além disso, ela é um direito essencial à formação e ao pleno desenvolvimento da pessoa, seu preparo para o exercício da cidadania e sua qualificação para o trabalho, como consta no artigo 205 da Constituição.

Contudo, para que o indivíduo possa gozar desse direito, o estado deve também assegurá-lo através de políticas públicas que garantam os direitos fundamentais dos cidadãos, como a educação e os benefícios que ela traz para ele, são questões que envolvem não apenas o cumprimento da lei ao ofertar os diferentes níveis de ensino, mas de assegurar os estudantes até o final do percurso.

Na contemporaneidade, ao tratarmos da relação entre educação, instituições de ensino e sociedade, inevitavelmente deparamo-nos com algumas questões conflitantes, dentre elas, a retenção e a evasão merecem destaque. Da educação básica à educação superior, em todos os níveis e modalidades de ensino, esses problemas estão presentes (MINISTÉRIO DA EDUCAÇÃO, 2014).

Entre tantos problemas presentes na educação superior, a evasão merece uma atenção especial e um estudo dedicado, pois se trata de um assunto delicado e preciso, e em comparação com outros

fenômenos, ainda é pouco estudado no sistema educacional brasileiro. Neste contexto, os autores Baggi e Lopes (2011, p.371) consideram que:

A evasão tem múltiplas razões, dependendo do contexto social, cultural, político e econômico em que a instituição está inserida. Pode estar relacionada, por exemplo, a má qualidade de ensino oferecido pela IES, provocando a perda definitiva do aluno.

Diante do exposto, este estudo de evasão foi realizado no curso de Engenharia Agrônômica do Instituto Federal de Tecnologia e Ciências do Tocantins – *Campus* Dianópolis, visto que há poucos estudos acerca da temática no *campus* e principalmente no curso citado, além disso, não só a região, mas como o país necessita de profissionais da engenharia, como defende Borges e Almeida (2013, p.72) “A educação em Engenharia no Brasil constitui um desafio diante de um cenário mundial que demanda uso intensivo de tecnologias e que exige, cada vez mais, um maior número de profissionais altamente qualificados”.

2 EVASÃO NO ENSINO SUPERIOR

A evasão é considerada um fracasso escolar, ligado aos elementos que resultam no desligamento do aluno do ensino superior seja abandono, desistência, transferência, trancamento etc., tendo como consequência diversos fatores, como afirma Batista, Sousa e Oliveira (2009, p.4):

O abandono à escola é composto então pela conjugação de várias dimensões que interagem e se conflitam no interior dessa problemática. Dimensões estas de ordem política, econômica, cultural e de caráter social. Dessa maneira, o abandono escolar não pode ser compreendido, analisado de forma isolada. Isto porque, as dimensões socioeconômicas, culturais, educacionais, históricas e sociais entre outras, influenciam-se mutuamente.

Os estudos científicos sobre Evasão escolar começaram em 1975, pelas teorias elaboradas com o objetivo de explicar a permanência dos estudantes nas Instituições de Ensino, baseadas em análises voltadas à finalidade estudante-instituição, por meio de variáveis ligadas aos estudantes, instituição e temas como a integração dos estudantes.

Atualmente, a evasão é definida como um termo polissêmico, podendo ser caracterizada como abandono, desistência, fracasso, transferência, saída definitiva e/ou suspensão temporária do curso, da instituição e/ou do sistema escolar. É a situação em que o estudante deixa de renovar a matrícula do curso superior e interrompe o seu ciclo de estudos em um curso.

Para o MEC (1997), a evasão é compreendida como a saída definitiva do estudante do curso de origem sem o concluir. No entanto, essa definição não abrange as transferências de curso interna e nem externas. Além dos diversos conceitos sobre evasão, nem todas as pesquisas se identificam com o termo evasão, passando a adotar o antônimo: permanência. Schimitt (2014) destaca que a relação entre os dois termos leva a uma mesma unidade de pesquisa, pois as questões de

vulnerabilidade, preditivos e preventivos se cruzam. Neste contexto, os autores Baggi e Lopes (2011, p.371) consideram que:

A evasão tem múltiplas razões, dependendo do contexto social, cultural, político e econômico em que a instituição está inserida. Pode estar relacionada, por exemplo, a má qualidade de ensino oferecido pela IES, provocando a perda definitiva do aluno.

2 MINERAÇÃO DE DADOS

O processo de Mineração de Dados auxilia na tomada de decisões úteis e estratégicas para melhoria de qualquer órgão em que se aplica o estudo, pois ao realizar a descoberta de padrões, é também possível obter previsões e realizar análises traduzidas em práticas de melhorias. Esse processo busca uma grande quantidade de dados que são informações desconhecidas.

Todo processo é praticamente inviável de se realizar manualmente, pois é feito na base de dados armazenados eletronicamente nos bancos de dados, sendo algumas informações ocultas em que só o algoritmo consegue extrair de forma útil. Além disso, se realizado de forma manual, seria um processo passível de erros e extremamente custoso de tempo.

Os tipos de padrões que podem ser minerados são: Regras de Associações, Classificação e Predição, Análise de Clusters, Análise de Outliers e Sequências (ou Análise Evolutiva). Podendo o banco de dados ser relacional, transacional, orientado-objeto, relacional-estendido e data warehouse.

As técnicas a serem utilizadas na Mineração de Dados podem ser: Técnicas de aprendizado de máquina (Machine Learning), estatística, redes neurais, algoritmos genéticos e outros.

3 REGRAS DE ASSOCIAÇÃO

No processo de Mineração de Dados, a técnica de Regras de Associação é bastante utilizada, ela serve para encontrar padrões entre os itens de uma base dados por meio da extração de regras geradas pelo algoritmo, no qual os resultados são apresentados de acordo com a co-ocorrência e não casualidade.

Por meio das regras geradas pode se obter resultados que definem padrões e/ou perfis de acordo com as informações da base de dados, esses resultados auxiliam na tomada de decisões de determinadas organizações, pois ao definir esses padrões, facilita nas estratégias dos objetivos que se deseja alcançar para o benefício.

A aplicação do algoritmo para regras de associação pode ser utilizada em diversas organizações como mercados, lojas físicas e online, planos de vendas, combate à evasão escolar (estudo deste trabalho) e afins. Para ajudar a entender a força da associação, existem métricas a serem definidas, sendo elas: Suporte e Confiança. Suporte é a frequência em que o conjunto de itens aparece nas transações, servindo para ajudar a identificar as regras válidas; Confiança é a probabilidade em

que os itens ocorrem, é encontrado de acordo com a quantidade dos itens que aparecem em conjunto dividido pelo número total de transações.

4 METODOLOGIA

Esta pesquisa estatística exploratória tem como fundamentação a observação, pois os elementos utilizados para o estudo não são afetados, no qual consiste nas características gerais dos alunos do curso de Bacharelado em Engenharia Agrônômica – *Campus* Dianópolis.

As variáveis utilizadas são qualitativas e quantitativas, a primeira pois os dados são de categoria nominal sem hierarquia, e a segunda, os resultados assumem quantidades contínuas ao apresentar a Confiança dos algoritmos.

A população da pesquisa são os dados relacionados à matrícula de todos os acadêmicos do curso Bacharelado em Engenharia Agrônômica – *Campus* Dianópolis no período de 2014 a 2019. Os dados foram fornecidos pela Coordenação de Registros Escolares (CORES) do *campus* em estudo, neles constam o total de 321 registros de matrículas desde a implementação do curso. A amostra utilizada foi de registros de ingresso de 2014 a 2018, reduzindo para 283 registros, o que representa 88% da população. A amostra foi selecionada pela objetividade de identificar os padrões/perfis dos alunos evadidos ingressantes do período de 2014 a 2018 por meio do algoritmo Apriori na Linguagem Python (3.7). Os dados de 2019 não foram utilizados, pois a coleta foi realizada no primeiro semestre do ano de 2019, portanto não continham os dados significativos para identificação. Os seguintes dados foram solicitados para realização da pesquisa: Raça, Sexo, Renda, Ano de Nascimento (Faixa Etária), Percentual de Conclusão e Status (Evadido, em Curso, Concluído). A escolha se deu dentre os dados brutos fornecidos, foi realizada uma seleção e escolha dos registros de 2014 a 2018.

4.1 REGRAS DE ASSOCIAÇÃO NO PYTHON

Para gerar as regras em um ambiente gráfico interativo, foi utilizado o ambiente científico Spyder – The Scientific Python Development Environment com a linguagem Python, foi utilizado o algoritmo apriori, importado em forma de biblioteca apyori no Python, consiste em gerar padrões associativos de acordo com os parâmetros definidos no script e identificar relações subjacentes entre itens diferentes.

Os parâmetros Suporte e Confiança são importantes no processo. O suporte dá uma ideia do quão frequente é um conjunto de itens em todas as transações, o seu valor ajuda a identificar as regras que valerão a pena considerar. A confiança define a probabilidade de ocorrência. Os parâmetros utilizados no algoritmo foram os seguintes: Suporte ($\text{min_support} = 0.3$), e Confiança ($\text{min_confidence} = 0.8$). Após a execução do algoritmo apriori no Python, foram geradas 136 regras de associação. Para análise, foram consideradas apenas as regras que possuíam como consequente

“Evadido”, sendo o total de 26 regras selecionadas, analisadas e expostas na **seção** resultados. As regras selecionadas foram separadas em três grupos com base no sexo (Masculino e Feminino e Indefinido). O item “Indefinido” foi considerado para os indivíduos cujas regras não definiram o sexo específico, ou o campo vazio. Para definir o padrão dos alunos evadidos gerados pelas regras na linguagem Python foi utilizado a média que é uma medida estatística que representa um conjunto com informações de determinada base de dados, usado para medir a confiança dos resultados.

4.2 REGRAS DE ASSOCIAÇÃO EM R

A segunda etapa do processo foi gerar as regras de associação na linguagem R, no ambiente RStudio por meio do algoritmo Apriori, importado pelo pacote Aures, no qual oferece um ambiente para exploração de técnicas para regras de associação. Como na linguagem Python, no R também foram definidos os mesmos parâmetros para Suporte ($\text{supp}=0.03$) Confiança ($\text{conf}=0.8$), e definido na estrutura do código o $\text{minlen}=6$ para a utilização de todos os campos da base de dados selecionada, gerando o total de 11 regras em torno dos parâmetros definidos, mas apenas uma com o consequente Evadido, que será apresentado na **seção** de Resultados.

4.3 REGRAS DE ASSOCIAÇÃO NO WEKA

A terceira etapa foi a mineração dos dados para a regras de associação na ferramenta WEKA, no qual foi utilizado também o algoritmo apriori. No Weka Explorer, os dados foram carregados no formato ARFF, o processo para gerar as regras foi na sessão Associate no qual foi escolhido o algoritmo apriori.

Para gerar as regras de associação, foi utilizado o parâmetro $\text{metricType}=\text{Confidence}$, que define o padrão confiança para gerar as regras, e a confiança foi o mesmo dos scripts anteriores ($\text{minMetric}=0.8$). Na ferramenta é possível definir o número de regras a serem geradas, para apresentar o máximo possível, foi definido o total de 283 regras, esse número é total de linhas de informações da base de dados, porém, de acordo com o número de confiança definido, a ferramenta apresentou 40 regras.

Com o consequente Evadido, o WEKA gerou seis regras que serão expostas e analisadas na **sessão** seguinte.

5. RESULTADOS

5.1 RESULTADOS PYTHON

Na análise dos dados coletados e nas regras geradas pela linguagem Python, alguns critérios foram levados em consideração, nas 26 regras selecionadas os resultados obtidos foram: o campo “Raça” obteve resultado “Parda” em todos os casos. Já o campo “Sexo” foi dividido nas três categorias: Masculino, Feminino e Indefinido. A “Renda” obteve “Não Declarado” em todos os seus campos, a

“Faixa Etária” foi “Jovem”, e o “Percentual de Conclusão” em todos os casos foram “Menos de 20%”.

Dessa forma, foi possível esquematizar os dados para definir o perfil dos alunos evadidos com base na recorrência de cada característica, como apresentado nas tabelas abaixo:

Tabela 1 – Padrões gerados em Python

ANTECEDENTE					CONSEQUENTE		
Raça	Sexo	Renda	Faixa etária	% conclusão	=	Status	Confiança
Padrão 1							
Parda	Feminino	Não declarado	Jovem	Menos de 20	=	Evadido	68%
Padrão 2							
Parda	Masculino	Não declarado	Jovem	Menos de 20	=	Evadido	68%
Padrão 3							
Parda	Indefinido	Não declarado	Jovem	Menos de 20	=	Evadido	82%

Fonte: o autor.

5.2 RESULTADOS EM R

Nas 11 regras geradas pela Linguagem R, apenas uma apresentou o conseqüente como “Evadido, com isso, obtém se apenas um padrão de alunos evadidos do curso de Bacharelado em Engenharia Agrônômica, *campus* Dianópolis, como apresentado na tabela abaixo: Tabela 2 – Padrão

Tabela 2 – Padrões gerados em R

ANTECEDENTE					CONSEQUENTE		
Raça	Sexo	Renda	Faixa Etária	Conclusão (%)	=	Status	Confiança
Padrão 1							
Parda	Masculino	Não declarado	Jovem	Menos de 20	=	Evadido	100%

Fonte: o autor.

No padrão único gerado em R, pode-se observar que os alunos de cor “Parda”, do sexo “Masculino”, “Jovem” e com percentual de conclusão “Menos de 20%” se encontram com o Status de “Evadido”, a confiança para essa regra é de 100% gerada pela própria regra.

5.3 RESULTADOS EM WEKA

Na ferramenta WEKA das 40 regras geradas, apenas 7 constavam resultados de interesse para esta pesquisa, por apresentar como conseqüentes “Evadido”. O campo “Cor” obteve em todas as regras selecionadas o resultado “Parda”, o “Sexo” obteve resultado “Indefinido”, a “Renda” apresentou “Não Declarado” em todos as regras e a “Faixa Etária” foi “Jovem”, o “Percentual de Conclusão” apresentou em todos os casos “Menos de 20%”, dessa forma, foi definido a média entre os resultados para apresentar a confiança, como exposto na tabela abaixo:

Tabela 2 – Padrões gerados no WEKA

Raça	Sexo	ANTECEDENTE			=	CONSEQUENTE	
		Renda	Faixa Etária	% conclusão		Status	Confiança
Padrão 1							
Parda	Indefinido	Não Declarada	Jovem	Menos de 20	=	Evadido	89%

Fonte: o autor.

A média foi calculada de acordo a confiança das 7 regras com consequente evadido, obtendo o resultado de 89%.

6 CONSIDERAÇÕES FINAIS

Com a realização desse estudo, desde o processo de pesquisa que englobou os seus diversos tipos, aos resultados, foi notório o enriquecimento do saber que foi obtido. Com ele foi possível refletir e analisar sobre o perfil dos alunos evadidos do curso de Bacharelado em Engenharia Agrônômica do IFTO *campus* Dianópolis.

Todas as etapas de mineração de dados foram essenciais para se obter os resultados baseados nos objetivos. O processo de KDD se fez muito válido para encontrar as regras de associação e fazer uma análise reflexiva sobre elas, o que acarretou resultados satisfatórios à pesquisa.

Por meio do algoritmo apriori utilizado na linguagem Python (3.7), R e WEKA foi possível analisar as regras obtidas pelos dados de 2014 a 2018 do curso supracitado.

Os resultados servem para nortear os gestores do *campus* a continuar o trabalho de combate à evasão, pois, esse fenômeno traz resultados negativos em diversas vertentes, tanto institucional quanto social. Sendo assim, esse processo é voltado para todos os envolvidos na Instituição de Ensino, sendo um papel didático acadêmico que visa a permanência do aluno, pois a educação é a única arma capaz de mudar os mais diversos âmbitos existentes, pois “A sociedade e cada meio social particular estabelece o ideal que educação desenvolve” – Émile Durkheim.

REFERÊNCIAS

- BRASIL. CONSTITUIÇÃO FEDERAL (1988). Título II, capítulo II, **dos direitos sociais**.
- BORGES, M. N.; ALMEIDA, N. N. **Perspectivas para a Engenharia Nacional - Desafios e Oportunidades**. Disponível em: <<http://ptdocz.com/doc/330551/baixar-este-arquivo-pdf>> Acesso em: 08 jun. 2019.
- MEC - MINISTÉRIO DA EDUCAÇÃO. **Documento Orientador para a Superação da Evasão e Retenção na Rede Federal de Educação Profissional, Científica e Tecnológica**. 52 p. Brasília, 2014.
- BAGGI, Cristiane A. S.; LOPES, Doraci, A. L. **Evasão e avaliação institucional no ensino superior: uma discussão bibliográfica**. *Avaliação*, Campinas, v.16, n.2, p.355-374, julho, 2011.
- BATISTA, S. D.; SOUZA, A. M.; OLIVEIRA, J. M da S. **A Evasão escolar no ensino médio: um estudo de caso**. *Revista Profissão Docente*, Uberaba, v.9, n.19, 2009.

MEC, Ministério da Educação. **Comissão Especial de Estudos sobre a Evasão nas Universidades Públicas Brasileiras. *Diplomação, retenção e evasão nos cursos de graduação em instituições de ensino superior públicas.*** Associação Nacional dos Dirigentes das Instituições Federais de Ensino Superior (ANDIFES), Associação Brasileira dos Reitores das Universidades Estaduais e Municipais (ABRUEM), Ministério de Educação e Cultura. Secretaria de Ensino Superior. Brasília, 1997. 152 p.

SCHIMITT, Rafael E.. **A evasão na educação superior: uma compreensão ecológica do fenômeno como estratégia para a gestão da permanência estudantil.** In: X Anped Sul -Reunião Científica da ANPED, 2014. Textos Completos. Florianópolis: UDESC, Outubro, 2014 p 1-21. Disponível em: http://xanpedsul.faed.udesc.br/arq_pdf/690-0.pdf. Acesso em: set. 2019.