

# PREDIÇÃO DA EVASÃO DOS ALUNOS EM CURSOS DE GRADUAÇÃO PRESENCIAIS

**Gustavo Zanini Kantorski (UFSM)**  
gustavoz@cpd.ufsm.br

**Jader Adiel Schmitt (UFSM)**  
jaderadiel@hotmail.com

**Evandro Gomes Flores (UFSM)**  
evandrosm@gmail.com

**Ivan Londero Hoffmann (UFSM)**  
ilhoffmann@ufsm.br



*As instituições de ensino superior enfrentam desafios para minimizar a evasão dos alunos nos cursos de graduação. O presente artigo propõe a previsão da evasão em cursos de graduação presenciais com a finalidade de visualizar perspectivas que permitam uma ação efetiva de intervenção, mitigando o processo da evasão. A pesquisa desenvolvida neste trabalho realiza uma análise comportamental dos alunos de cursos de graduação com o intuito de prever quais os alunos possuem maior tendência em abandonar o curso. A abordagem utiliza vários métodos de aprendizagem de máquina para a previsão. A intuição é gerar uma lista de prováveis alunos que não realizarão matrícula no período posterior ao que estão cursando. A proposta é genérica e pode ser aplicada a qualquer curso de graduação. A abordagem proposta foi aplicada em dois cursos de graduação da instituição e os resultados demonstram que a abordagem proposta é factível e eficiente. Os experimentos alcançaram uma acurácia média de 96% na previsão, e mais de 73% de sucesso na previsão de alunos que abandonaram o curso.*

*Palavras-chave: evasão, aprendizagem de máquina, previsão*

## 1. Introdução

A evasão de alunos nos cursos de graduação tem sido um dos problemas relevantes para a gestão acadêmica e financeira das Instituições de Ensino Superior (IES), pois cada vez mais são investidos recursos escassos na atração e captação de alunos e os resultados têm sido influenciados negativamente com taxas elevadas de evasão. A desistência do aluno em um curso superior significa prejuízo para si próprio, ao não se diplomar; para o professor, que não atinge sua meta como educador; para a universidade, pelo não atendimento de sua missão; para a sociedade, pelas perdas sociais e econômicas; e, também, para a família, pelo sonho não realizado (SILVA FILHO *et al.*, 2007).

Segundo a abordagem de Santos e Giraffa (2013) é pressuposto que a evasão e a permanência são conceitos intimamente ligados e possuem elementos complementares. Ao buscarmos elucidar as causas da evasão estudantil, teremos condições de avaliar os motivos pelos quais os estudantes desistem de sua carreira universitária, contribuindo para a elaboração e planejamento de ações que privilegiem a permanência dos estudantes, a partir de ações que qualifiquem o período que os estudantes estejam na IES. Desta forma, contribuindo para compor um conjunto de ações que possam auxiliar na permanência, estaremos suavizando as causas da evasão.

A maioria dos recentes trabalhos que trata a evasão no ensino superior tem como finalidade compreender os fatores que causam a evasão (AMARAL 2013; PALACIO, 2012; SANTOS, 2015; VIEIRA JUNIOR, 2014; FERREIRA, 2014). Entre os vários fatores encontrados nesses trabalhos destacam-se a falta de motivação dos alunos/professores, problemas pessoais e socioeconômicos, insatisfação com o curso/instituição, problemas de aprendizagem associados com metodologias de ensino e processos de avaliação, restrições do mercado de trabalho, entre outros.

Outros trabalhos (ABU-ODA & EL-HALEES, 2015; AMAYA *et al.*, 2014; AZOUMANA *et al.*, 2014; JADRIĆ *et al.*, 2010; OREA *et al.*, 2005) abordam a evasão no sentido de prever situações de evasão no ensino superior. Geralmente, esses trabalhos utilizam técnicas de descoberta de conhecimento em bases de dados ou mineração de dados para realizar a previsão de resultados.

O presente trabalho apresenta uma abordagem para previsão da evasão em cursos de graduação presenciais. A abordagem é genérica e pode ser aplicada em qualquer curso de graduação presencial. A intuição é realizar uma análise comportamental dos alunos e prever os prováveis alunos que possuem tendência à evasão. Nesse sentido, uma análise profunda de atributos de alunos é realizada. Essa análise considera informações pessoais, acadêmicas, de caráter sócio econômico e combinam essas informações para criar um modelo preditivo da evasão. A abordagem proposta nesse trabalho não utiliza somente um modelo de previsão, mas faz uma combinação de modelos de previsão para otimizar o resultado da previsão de possíveis evasões.

## 2. Metodologia

A metodologia proposta utiliza conceitos de aprendizagem de máquina para prever as informações relacionadas à evasão. Para análise da evasão a metodologia proposta é uma adaptação da metodologia CRISP-DM (*CRoss Industry Standard Process for Data Mining*) (CHAPMAN, 2000). A metodologia é geral e não se restringe a uma ferramenta ou tecnologia específica. A primeira etapa é o entendimento da evasão e tem como finalidade identificar os objetivos sob o ponto de vista de descoberta de conhecimento em bases de dados. Esses objetivos incluem o tipo de tarefa a ser executada, por exemplo, previsão e os critérios de avaliação dos modelos utilizados.

A próxima etapa, entendimento dos dados dos alunos, consiste na coleta inicial dos dados e na descrição e exploração dos dados verificando suas propriedades e qualidade. O processo exploratório compreende a análise dos dados e a utilização de técnicas de visualização. A descrição envolve a avaliação de características, tais como, faixas de valores, número de atributos, significado de cada atributo e sua importância para alcançar o objetivo.

Na etapa de preparação dos dados um conjunto é gerado em conformidade com os modelos de descoberta de conhecimento que serão utilizados. A geração desse conjunto envolve uma série de passos incluindo seleção, limpeza, construção, integração e formatação. A seleção dos dados e das instâncias pode ser realizada de maneira manual ou por meio de algoritmos. A limpeza tem como finalidade melhorar a qualidade. Dentre as atividades de limpeza pode-se citar a eliminação de dados com erro, a padronização de dados, por exemplo, abreviaturas, formatos de data. A construção dos dados concentra-se na normalização, se for o caso, na

transformação de valores simbólicos para numéricos e na discretização dos atributos. A integração serve para combinar múltiplas fontes de dados e a formatação pode realizar alterações sintáticas nos dados sem modificar o seu significado.

A etapa de modelagem é a fase de descoberta de conhecimento propriamente dita. Nessa etapa são selecionados os algoritmos que serão aplicados para a solução da tarefa. A etapa de avaliação interpreta os resultados em relação aos objetivos definidos na etapa de entendimento do negócio. Um ciclo existente entre a etapa de modelagem e avaliação possibilita a comparação entre diversos modelos de aprendizagem. Finalmente, a etapa de disponibilização entrega os resultados alcançados por meio da comparação entre os diversos modelos gerados.

### 3. Experimentação

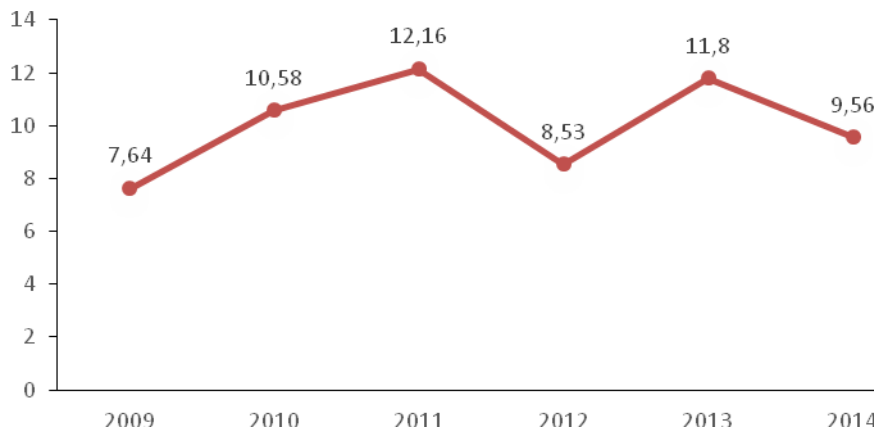
A metodologia foi aplicada em dois cursos de graduação presencial da Instituição: Curso de Zootecnia e Administração. Por questões de espaço, serão abordados os resultados de um dos cursos - Zootecnia. O curso de Zootecnia foi selecionado baseado na avaliação de indicadores do Ministério da Educação (MEC) do Brasil. Uma das variáveis utilizadas para a escolha do curso foi o peso. O peso determina uma área de custo do curso para que seja possível diferenciar cursos de maiores custos em relação aos demais. Por exemplo, um dos fatores predominantes é a utilização de laboratórios que exigem maiores recursos para seu funcionamento. O curso de Zootecnia, juntamente com Medicina, Medicina Veterinária e Odontologia, é um dos cursos que possui maior peso. Na perspectiva de evasão, nota-se que o maior percentual de evasão entre os cursos com maior peso, é o curso de Zootecnia. Para o cálculo da evasão foi utilizada a fórmula definida por Leal (2007), a qual é baseada numa comparação entre o número de alunos que estavam matriculados num curso num determinado ano, subtraídos os concluintes, com a quantidade de alunos matriculados no ano seguinte, subtraindo deste último o total de ingressantes desse ano, dada pela equação 1. A Figura 1 apresenta a taxa de evasão do curso de Zootecnia, desde 2009 até 2014.

$$E_n = 1 - \frac{[M_n - I_n]}{[M_{n-1} - C_{n-1}]} \quad (1)$$

em que:

- E = taxa de evasão anual
- M = número de matriculados
- C = número de concluintes
- I = número de ingressantes
- n = ano em estudo
- (n-1) = ano anterior

Figura 1 - Taxa de Evasão do Curso de Zootecnia



Fonte: Sistema de Informação da Instituição (2014)

A coleta de dados levou em consideração vários tipos de atributos, entre eles, atributos pessoais, acadêmicos, sociais, econômicos e financeiros. Essas categorias foram selecionadas conforme vários trabalhos (TIGRINHO, 2008; SILVA FILHO, 2007; SOBRINHO, 2008), existentes na literatura sobre o tema. O Anexo I mostra as categorias e os atributos selecionados para análise. Cabe salientar que uma limitação na definição dos atributos está relacionada na disponibilidade de outras informações confiáveis nos sistemas de informação projetados para atender as necessidades de controle das atividades desenvolvidas pelo acadêmico na Instituição. Exemplos de atributos que, geralmente, não são registrados, dizem respeito aos aspectos sociais e emocionais.

A metodologia foi aplicada em três simulações. Uma simulação possui um conjunto de atributos e um conjunto de algoritmos. O que diferencia uma simulação de outra são os atributos selecionados para o processo preditivo e os algoritmos aplicados a esses atributos. A primeira simulação consiste de 25 atributos, a segunda simulação contém 32 atributos e a terceira simulação é composta de 34 atributos. Além dos atributos selecionados, existe um atributo especial que é o atributo classe. O atributo classe possui dois valores para avaliação: “Aluno Regular” e “Evadido”. Além disso, em uma mesma simulação, foram utilizadas

técnicas para normalização dos atributos. A normalização pode ser necessária, pois alguns atributos assumem uma faixa de valores mais ampla do que outros e, determinados modelos podem ser influenciados por atributos com valores mais altos ou mais baixos. A normalização faz com que todos os atributos fiquem na mesma faixa de valores.

As informações de alunos com ingresso no ensino superior de 2009 a 2014 foram consideradas para análise. No período indicado foram selecionados 394 instâncias de treinamento. Do ponto de vista da categoria pessoal, do total de instâncias de treinamento, 217 são do sexo feminino e 177 do sexo masculino. Ainda, 292 instâncias não ingressaram por sistema de cotas, enquanto 102 cotistas estão no conjunto. A análise sócio econômica mostra que a maioria dos alunos é proveniente de escola pública (252), assim como a maioria dos alunos irá se manter com recursos familiares durante o curso (273). A maioria das informações da categoria Institucional é faltante (em torno de 90%). Isso ocorre porque essas informações são geradas no processo de avaliação institucional, em que a maioria dos alunos não participa ativamente. Um fato observado é que 90% das instâncias previstas não responderam ao processo de avaliação institucional.

O viés de assistência estudantil descreve que existem mais alunos bolsistas (56%), que não bolsistas. Na perspectiva de bolsas de monitoria a situação é inversa. A maioria dos alunos não possui bolsa de monitoria (95%), assim como, a maioria dos alunos não reside na casa de estudante universitário (96%). A análise de desempenho acadêmico mostra que a média geral de notas dos alunos está em 5,472 apresentando um desvio padrão de 2,414. A análise da posição do aluno considera a carga horária do curso, ou seja, a análise da carga horária cumprida pelo aluno e o semestre cronológico que o aluno se encontra atualmente no curso. Portanto, esse atributo analisa se o aluno está em dia com o que dele se espera, evitando-se assim a retenção acadêmica. O valor médio desse atributo é -2,548 e varia de -14 a +4 entre os 394 alunos do curso.

Para a fase de modelagem foi utilizada a ferramenta Weka (HALL, 2009). Essa ferramenta consiste de um conjunto de algoritmos e modelos de aprendizagem de máquina. Cada simulação é composta por um conjunto de modelos. Além disso, em uma mesma simulação, um determinado modelo pode ser aplicado várias vezes com configurações e parametrizações diferentes. Os algoritmos aplicados na etapa de modelagem foram J48, *nearest neighbor* (kNN), CART, *Naïve Bayes*, *Multilayer Perceptron* e *OneR*. Os algoritmos foram aplicados

em todas as matrículas realizadas no primeiro semestre de 2015, totalizando 302 alunos regulares no curso de Zootecnia.

Todas as simulações trabalharam com atributos normalizados e não normalizados. A simulação 1 utilizou cinco algoritmos (*kNN*, *CART*, *J48*, *Naïve Bayes*, *Multilayer Perceptron*), mas somente um – *kNN* – apresentou resultados mensuráveis. A Tabela 1 apresenta os resultados encontrados para a simulação 1. A primeira coluna apresenta o algoritmo utilizado, a segunda coluna descreve se os atributos foram normalizados ou não e a terceira coluna mostra o número total previsto e o percentual da previsão no total analisado. Os valores de *k* na primeira coluna da Tabela 1 representam o número de instâncias que serão comparadas para a previsão. Por exemplo, se  $k=3$  serão analisadas as três instâncias mais próximas para determinar a previsão. Os valores previstos de potenciais evasões variam de 1 a 30 e não são exclusivos. Existe uma sobreposição no total previsto. A previsão obteve 14 matrículas presentes em quatro técnicas, quatro matrículas em três técnicas, oito matrículas em duas técnicas e, finalmente, 12 matrículas em um modelo. A relação final, considerando todos os modelos, possuem 38 matrículas, o que representa 12,5% do total de matriculados.

Tabela 1 - Previsão da Simulação 1

<b>Algoritmo</b>	<b>Atributos normalizados</b>	<b>Previsão</b>
<i>kNN</i> – $k=1$	Não	2 (0,6%)
<i>kNN</i> – $k=3$	Não	26 (9%)
<i>kNN</i> – $k=5$	Não	30 (10%)
<i>kNN</i> – $k=1$	Sim	1 (0,3%)
<i>kNN</i> – $k=3$	Sim	19 (6%)
<i>kNN</i> – $k=5$	Sim	18 (6%)

A simulação 2 foi aplicada nos mesmos cinco modelos da simulação 1 porém os resultados alcançados foram diferentes. Entre os cinco modelos aplicados, três – *kNN*, *MultiLayer Perceptron* e *J48* – produziram resultados interessantes. A Tabela 2 ilustra os resultados alcançados. O modelo *kNN* foi aplicado com três configurações diferentes. Os valores previstos nessa simulação variam de 1 a 11 e, também, não são exclusivos. Os experimentos mostraram uma relação de 15 matrículas diferentes entre os três modelos aplicados. O detalhamento da previsão mostra que uma matrícula foi prevista em cinco modelos, seis matrículas foram previstas em quatro modelos, cinco matrículas em dois modelos e três matrículas em somente em um modelo. A simulação 2 obteve um número menor de previsões

porque atributos que consideram o desempenho acadêmico forma adicionados ao conjunto de treinamento e teste. Mais especificamente, atributos tais como número de disciplinas reprovadas, trancamentos de disciplinas, trancamentos de semestres e disciplinas reprovadas por frequência.



Tabela 2 - Previsão da Simulação 2

Algoritmo	Atributos Normalizados	Previsão
J48	Não	3 (1%)
<u>kNN – k=3</u>	Não	10 (3%)
<u>kNN – k=5</u>	Não	8 (3%)
<u>kNN – k=3</u>	Sim	11 (4%)
<u>kNN – k=5</u>	Sim	9 (3%)
<u>Multilayer Perceptron</u>	Sim	1 (0,3%)

Finalmente, na simulação 3 foram adicionados dois novos atributos. Os atributos posição no período e desempenho acadêmico (sequências 19 e 28 no Anexo I, respectivamente). A Tabela 3 descreve os resultados obtidos com a simulação 3. Na simulação 3 foram aplicados vários modelos (J48, kNN, CART, Naïve Bayes, *OneR* e *Multilayer Perceptron*), mas somente dois – J48 e CART – apresentaram resultados interessantes. O algoritmo J48 foi utilizado com duas configurações diferentes. A primeira configuração utiliza árvores de decisão não binárias, enquanto que a segunda configuração gera o modelo com árvores binárias. De maneira similar às demais simulações, atributos normalizados e não normalizados foram usados. O número total previsto para a simulação 3 foi de 37, e a variação dos modelos ficou entre 4 e 23. Entre as 37 previsões, uma delas foi prevista em seis modelos, uma em cinco modelos, duas em quatro modelos, oito em três modelos, seis em dois modelos e 19 em um modelo.

Tabela 3 - Previsão da Simulação 3

Algoritmo	Atributos Normalizados	Previsão
J48– Binária	Não	5 (2%)
J48	Não	4 (1%)
CART	Não	4 (1%)
J48 – Binária	Sim	23 (8%)
J48	Sim	22 (7%)
CART	Sim	16 (5%)

A avaliação das três simulações permite uma visão holística e cartesiana das previsões. A visão holística ilustra todas as simulações em um modelo único e integrado. A visão cartesiana estuda cada modelo como uma coleção de partes dissociadas. O viés cartesiano visualiza três simulações isoladas, em que a primeira simulação é composta por 38 matrículas, a simulação II é formada por 15 matrículas e, finalmente, a simulação III é composta de 37 matrículas. Por sua vez, o viés holístico mostra uma visão interconectada das simulações, na

qual a totalidade das matrículas não se dá pela simples soma das simulações. Esta visão integrada apresenta uma relação de 65 matrículas possíveis de evasão.

#### 4. Resultados

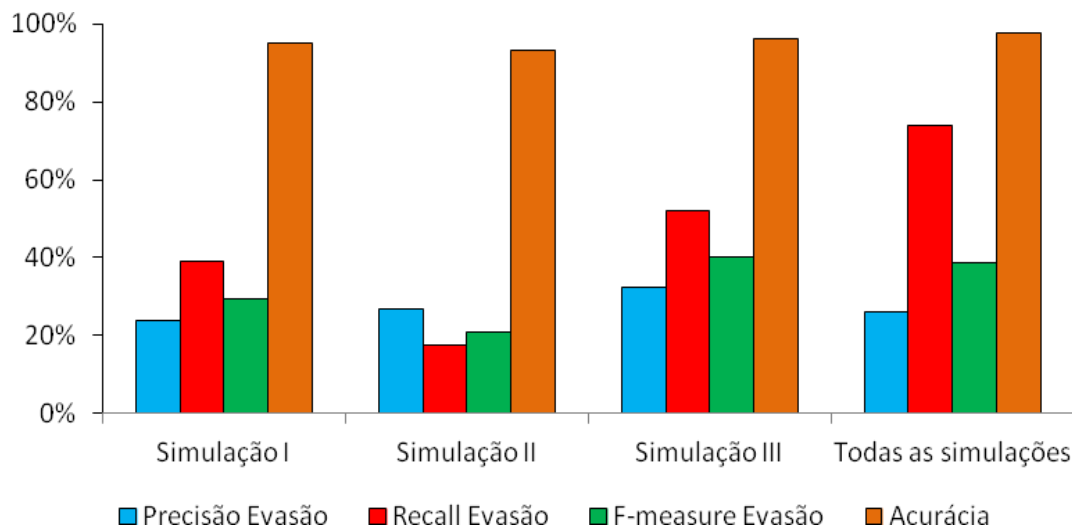
Esta seção apresenta os resultados encontrados após o processo de matrícula ser realizado na Instituição. O número de alunos regulares matriculados no período 2015/1 foi 302. Dos 302 alunos regulares matriculados, os alunos que ingressaram no semestre corrente, no total de 36 alunos, foram desconsiderados para previsão. Esse descarte foi realizado, pois os alunos ingressantes ainda possuem um perfil comportamental que permita a previsão de possível evasão. O número de alunos aptos a realizar a matrícula no semestre de 2015/2 foram 299 alunos. Após o prazo de realização de matrículas foram identificados 23 abandonos no curso de Zootecnia. A mineração dos dados acertou 17 dos 23 alunos que abandonaram o curso. Os 17 alunos estavam na relação de 65 matrículas, as quais foram previstas pela consolidação das três simulações realizadas.

Para avaliação dos resultados da previsão foram utilizadas as métricas de acurácia, precisão, *recall* e *f-measure*. Essas métricas foram utilizadas para avaliar a qualidade da previsão. A acurácia mede o total de acerto comparado com o total de instâncias avaliadas. A precisão avalia a fração relevante comparada com o total previsto. O *recall* avalia a fração que foi prevista comparada com o total relevante. A métrica *f-measure* combina a precisão e o *recall* em uma só medida. Seja  $R$  uma relação de todas as matrículas previstas. Seja  $S$  um conjunto de todas as matrículas que evadiram durante o processo de matrícula. As definições de precisão( $P$ ), *recall* ( $R$ ) e *f-measure*( $F$ ) são dadas por:  $P = \frac{|R \cap S|}{|R|}$ ,  $R = \frac{|R \cap S|}{|S|}$ ,  $F = 2 \times \frac{P \times R}{P + R}$ .

A avaliação foi realizada para cada simulação e também com a combinação das três simulações. A Figura 4 apresenta os valores alcançados para cada uma das métricas em cada simulação e na combinação das simulações. A Figura 4 mostra algumas informações interessantes. Todas as simulações alcançaram um valor alto de acurácia. Isso aconteceu porque a maioria dos alunos foram previstos como alunos regulares. O valor máximo de acurácia foi de 98% e a média entre as simulações alcançou 95%. As simulações I e II apresentaram resultados inferiores comparados com a simulação III e a combinação das três simulações. A precisão teve o melhor desempenho na simulação III, alcançando o valor

percentual de 33%. O *recall*, por sua vez, teve o melhor desempenho na combinação das três simulações, com o valor percentual de 74%. A *f-measure* alcançou o melhor resultado para a simulação III totalizando 0,4. A combinação das simulações teve resultado semelhante (0,39). A *f-measure* obteve o melhor resultado para a simulação III considerando-se que a precisão e o *recall* possuem a mesma importância. Caso o objetivo seja acertar mais alunos que evadem, a combinação das três simulações é melhor que a simulação III.

Figura 2 - Métricas de Avaliação



## 5. Conclusão

Este artigo propõe a abordagem inovadora para previsão de possíveis evasões de cursos de graduação presenciais em Universidades públicas. A abordagem extrai informações pessoais, acadêmicas, sociais e econômicas e, por meio das informações extraídas, são construídos modelos de previsão utilizando-se métodos de mineração de dados. A vantagem da abordagem proposta é a combinação de vários modelos de mineração de dados em uma única previsão, otimizando o resultado do processo. Essa visão holística, na qual os modelos são combinados em um todo único e integrado, permite um resultado mais abrangente e preciso. Os experimentos demonstraram que a abordagem proposta nesse trabalho é apropriada para prever a evasão no curso de graduação em Zootecnia da Instituição. As métricas alcançaram 98% na previsão das classes envolvidas (aluno regular e aluno evadido) e uma taxa de sucesso em 74% na previsão dos alunos que abandonaram o curso. A metodologia foi aplicada também para o curso de Administração e apresentou resultados semelhantes, alcançando uma

taxa de 73% de sucesso na previsão dos alunos que abandonaram o curso. Isso demonstra que a abordagem é genérica e pode ser aplicada a outros cursos de graduação. Outra contribuição é a possibilidade de avaliação dos modelos, a qual permite identificar os principais atributos que ajudam na previsão das classes. Essa identificação de atributos influi diretamente nos fatores associados ao processo de evasão e podem servir de estudo para mitigar o abandono nos cursos de graduação.

Como trabalhos futuros pretende-se realizar o estudo dos atributos que determinaram o sucesso do processo para filtrar somente aqueles que interessam à previsão. Dessa forma, são identificados os principais fatores que levaram ao processo de evasão. Além disso, a coleta de novos atributos pode ser realizada pela mineração em redes sociais (por exemplo, *facebook*, *twitter*, entre outras) e redes profissionais (por exemplo, *linkedin*), as quais contêm muitas informações sobre as atividades comportamentais dos alunos. Outra perspectiva tem como foco a generalização desse processo para os demais cursos de graduação, uma vez que o perfil de comportamento dos alunos em diferentes cursos é distinto. Espera-se que esse processo contribua na valorização de novas ações no sentido de minimizar a evasão de alunos dos cursos de graduação. Acredita-se que a possibilidade de gerar uma lista menor de potenciais alunos, permita ao curso realizar uma melhor gestão de permanência do aluno junto ao curso.

## REFERÊNCIAS

- ABU-ODA, Ghadeer S.; EL-HALEES, Alaa M. DATA MINING IN HIGHER EDUCATION: UNIVERSITY STUDENT DROPOUT CASE STUDY. **International Journal of Data Mining & Knowledge Management Process**, v. 5, n. 1, p. 15, 2015.
- AMARAL, João Batista do. Evasão discente no ensino superior: estudo de caso no Instituto Federal de Educação, Ciência e Tecnologia do Ceará (Campus Sobral). 2013.
- AMAYA TORRADO, Yegny Karina; BARRIENTOS AVENDAÑO, E.; HEREDIA VIZCAÍNO, D.J. Modelo predictivo de deserción estudiantil utilizando técnicas de minería de datos. 2014.
- AZOUMANA, Kamagate. Análisis de la deserción estudiantil en la Universidad Simón Bolívar, facultad Ingeniería de Sistemas, con técnicas de minería de datos. **Revista Pensamiento Americano**, v.6, n.10, 2014.
- CHAPMAN, Pete; CLINTON, J.; KERBER, R.; KHABAZA, T.; REINARTZ, T.; SHEARER, C.; WIRTH, R. CRISP-DM 1.0 Step-by-step data mining guide. 2000.
- FERREIRA, G.S.S.; AMARO, T.C.F. Evasão Na Educação A Distância: Um Estudo Exploratório Das Causas Da Evasão Em Uma Instituição De Ensino Superior . ENEGEP. Curitiba, PR. 2014.
- HALL, Mark et al. The WEKA data mining software: an update. **ACM SIGKDD explorations newsletter**, v. 11, n.1, p.10-18, 2009.
- JADRIĆ, Mario; GARAČA, Željko; ČUKUŠIĆ, Maja. Student dropout analysis with application of data mining methods. **Management: Journal of Contemporary Management Issues**, v.15, n.1, p.31-46, 2010.

LEAL, R.; MOTEJUNAS, P. R., HIPÓLITO, O., LOBO, M. B. V. M. A evasão no ensino superior brasileiro. **Cadernos de Pesquisa**, vol.9, n.132, pp. 641-659, set/dez. 2007.

OREA, S.V.; VARGAS, A.S.; ALONSO, M.G. Minería de datos: predicción de la deserción escolar mediante el algoritmo de árboles de decisión y el algoritmo de los k vecinos más cercanos. **Ene**, v.779, n.73, p.33, 2005.

PALÁCIO, Paula da Paz. Políticas de acesso e permanência do estudante da Universidade Federal do Ceara. 123p. Dissertação. UFC, Fortaleza, 2012.

SANTOS, N. V. M. , JUNIOR,M.L., RIBEIRO,M.L.L. Evasão no Curso de Engenharia de Produção da Universidade Federal de Goiás – Regional Catalão. ENEGEP. Fortaleza, CE. 2015.

SANTOS, P., GIRAFFA, L. M. (2013). Evasão na Educação Superior: um estudo sobre o Censo da Educação Superior no Brasil. In: CLABES, III. Anais. Disponível:  
[http://www.clabes2011alfaguia.org.pa/docs/0\\_LIBRO\\_COMPLETO.pdf](http://www.clabes2011alfaguia.org.pa/docs/0_LIBRO_COMPLETO.pdf)

SILVA FILHO, Roberto Leal Lobo; MONTEJUNAS, P.R.; HIPÓLITO, O. LOBO, M.B.C.M. A evasão no ensino superior brasileiro. **Cadernos de Pesquisa**, v.37, n.132, p 641-659, 2007.

SOBRINHO, José Dias. Avaliação educativa: produção de sentidos com valor de formação. **Avaliação, Campinas**, v. 13, n. 1, p. 193-207, 2008.

TIGRINHO, Luiz Maurício Valente. Evasão escolar nas instituições de ensino superior. **Revista Gestão Universitária**, v.173, p.01-14, 2008.

VIEIRA JUNIOR, J.E.S. A Análise Da Evasão Dos Alunos Do Proeja (Educação De Jovens E Adultos) No Campus De Porto E No Campus Palmas Do Ifto Com Enfoque Na Teoria Das Restrições. ENEGEP. Curitiba. PR. 2014.

## ANEXO I

Seq	Categoria	Atributo	Descrição
1	Pessoal	Sexo	Masculino/Feminino
2	Pessoal	Idade	18 a 61 (anos)
3	Pessoal	Etnia	Branca/Outra
4	Pessoal	Estado Civil	Solteiro/Casado/Divorciado/Outro
5	Sócio-Econômico	Cotista	Sim/Não
6	Acadêmico	Formação no ensino médio	Escola pública/Escola Privada/Não Informado
7	Acadêmico	Iniciou curso superior	Indicador se o aluno já iniciou curso superior
8	Sócio-Econômico	Como se manter durante o curso	Recursos familiares Bolsa de estudos Recursos do trabalho Recursos não definidos
9	Sócio-Econômico	Motivo que optou pelo curso	Atendimento de aptidões e interesses Aperfeiçoamento do exercício profissional Prestígio social da profissão Outro motivo
10	Institucional	Avaliação do plano de ensino	Boa/Excelente/Satisfatória/Muito Boa/Regular/Insatisfatória
11	Institucional	Orientação no uso	Bons/Desconheço/Excelente/Muito

Seq	Categoria	Atributo	Descrição
		de laboratórios	bons/Regulares/Insatisfatórios
12	Institucional	Atuação do coordenador	Boa/Excelente/Satisfatória/Muito Boa/Regular/Desconheço
13	Institucional	Práticas de estágio	Bons/Desconheço/Excelente/Muito bons/Regulares/Insatisfatórios
14	Institucional	Avaliação do corpo docente	Boa/Excelente/Satisfatória/Muito Boa/Regular/Desconheço
15	Acadêmico	Ano de ingresso	2008 a 2015
16	Acadêmico	Período de ingresso	1/2
17	Acadêmico	Tempo no curso	Tempo que o aluno está no curso
18	Acadêmico	Período atual	Período ideal que o aluno está no curso
19	Acadêmico	Posição no período	Razão entre carga horária total do curso e o número de semestres que o aluno ainda deve cursar. 0 – aluno em dia com a relação entre a carga horária vencida e a que deveria -n – aluno com n semestres atrasados +n – aluno com n semestres antecipados
20	Acadêmico	Disciplinas aprovadas	Número total
21	Acadêmico	Disciplinas reprovadas	Número total
22	Acadêmico	Disciplinas reprovadas por frequência	Número total
23	Acadêmico	Disciplinas dispensadas	Número total (cursadas em outro curso)
24	Acadêmico	Trancamentos totais	Número total: o aluno cancela o período
25	Acadêmico	Trancamentos parciais	Número total: o aluno cancela disciplina
26	Acadêmico	Média das notas	Valor médio
27	Acadêmico	Total	Número total de disciplinas matriculadas
28	Acadêmico	Desempenho acadêmico	-1 – Alunos que tiverem média de desempenho inferior a 5 1 – Alunos que tiverem desempenho superior a 7 0 – Demais alunos
29	Acadêmico	Médias de disciplinas por período	Valor médio de disciplinas por período (considera: ano atual – ano de ingresso)
30	Sócio-Econômico	Bolsista	Sim/Não
31	Sócio-Econômico	Moradia Estudantil	Sim/Não
32	Acadêmico	Monitoria	Sim/Não
33	Sócio-Econômico	Refeições RU	Número médio de refeições realizadas
34	Acadêmico	Retiradas Biblioteca	Taxa de utilização da biblioteca



<b>Seq</b>	<b>Categoria</b>	<b>Atributo</b>	<b>Descrição</b>
35	Classe	Classe	Aluno Regular/Evadido