

Evidência do desânimo de alunos em um ambiente virtual de ensino e aprendizagem: uma proposta a partir da mineração de dados educacionais

Fabírcia Damando Santos, PGIE-UFRGS/ UERGS – fabriciadamando@gmail.com

Magda Bercht, PGIE-UFRGS - bercht@inf.ufrgs.br

Leandro Krug Wives, PGIE-UFRGS - wives@inf.ufrgs.br

***Resumo.** O artigo relata uma pesquisa que busca identificar o aluno desanimado, em um ambiente virtual de ensino e aprendizagem utilizando mineração de dados. Foram aplicadas duas técnicas de classificação de dados utilizando árvore decisão, uma através do Holdout e outra com Cross-validation. Os resultados do modelo preditivo mostraram acerto em cerca de 91% dos dados para o método Holdout e 77% para cross-validation. Novos experimentos serão realizados a fim de validar o modelo e desenvolver ações que subsidiem o professor no apoio a esses alunos*

***Palavras-Chave:** Mineração de Dados Educacionais, Estados de Ânimo, Ambiente Virtual de Aprendizagem.*

Evidence of the student's discouraged in a virtual learning environment: a proposal from the educational data mining

***Abstract.** This paper reports a research that identifies the student discouraged in a virtual learning environment using data mining. Two data sorting techniques were applied using tree decision, through Holdout and another with Cross-validation. The results showed accuracy of the predictive model by about 91% the data for the Holdout and 77% for cross-validation. New experiments will be conducted to validate the model and develop actions that assist the teacher to support these students*

***Keywords:** Educational Data Mining, States of Mood, Virtual Learning Environment.*

1. Introdução

Este trabalho apresenta a descrição de um experimento, parte de uma pesquisa, que pretende investigar e mostrar a importância em considerar a dimensão afetiva no contexto educacional, principalmente relacionado ao aluno desanimado. A pesquisa prevê um experimento com a aplicação de técnicas de Mineração de Dados Educacionais (MDE), com o intuito de apresentar duas formas de classificação de alunos propensos ao desânimo, com tendência a optar pela melhor técnica de classificação. Dessa forma, foi feito um experimento com alunos da disciplina Análise de Demonstrações Contábeis ofertada pelo Departamento de Ciências Contábeis e Atuárias da Universidade Federal do Rio Grande do Sul (UFRGS), a fim de apresentar um modelo de classificação de alunos propensos ao desânimo.

A escolha pelo estado afetivo de desânimo se deu pelo fato dele ser percebido como um estado onde há uma tendência de conduzir o aluno à desistência do curso (Longhi, 2011). Essa situação é mais presente em cursos a distância, devido à quebra da simultaneidade espaço-temporal e a dificuldade do professor em identificar os alunos

desanimados, necessitando de subsídio para essa inferência. Nesse sentido, algumas iniciativas estão sendo realizadas no que diz respeito à análise de dados dos ambientes virtuais de aprendizagem (AVEA), coletando dados comportamentais e aplicando as técnicas de MDE, de modo a gerar subsídios para o apoio ao professor (Rigo *et al.*, 2014; Majadas, Santos e Boticário 2013).

Os aspectos afetivos e comportamentais dos alunos podem ser obtidos através de suas interações no AVEA ao utilizar as ferramentas do ambiente e deixar registradas as ações no ambiente. Ações e dados comportamentais como o tempo de demora em fazer uma atividade, o número de erros que ocorre na execução de uma atividade, o número de solicitações de ajuda, a forma de se expressar em textos, podem apoiar na identificação dos estados afetivos do aluno (Bercht, 2001).

Este trabalho está estruturado da seguinte forma: a seção 2 aborda os aspectos afetivos relacionados à educação, a seção 3 apresenta técnicas da mineração de dados educacionais, na seção 4 os trabalhos relacionados são descritos. A seção 5 descreve o procedimento metodológico e a seção 6 detalha a análise e resultados do experimento. Por fim, a seção 7 apresenta a conclusão.

2. Contexto da afetividade na educação

Atualmente uma nova leitura no contexto educacional tem sido feita sobre as relações entre as dimensões afetivas e cognitivas à qual busca integrá-las e não tratá-las com forma dissociada na construção do conhecimento. Nesse sentido, Piaget (2014) afirmava que não há mecanismo cognitivo sem elementos afetivos e vice-versa, podendo sentimentos de sucesso ou fracasso levar a uma facilitação ou inibição da aprendizagem.

Conceitualmente, o termo afetividade diz respeito à habilidade de experimentar sentimentos e a eles reagir, podendo incluir um conjunto de fenômenos afetivos que possuem diferentes origens, intensidades, duração e reações corporais (Bercht 2001; Longhi, 2011, Scherer, 2005). Picard (2000) preconiza que a afetividade é fator determinante nas escolhas humanas, tendo papel fundamental na tomada de decisão e na aprendizagem de um modo geral. Yang e Damásio (2000) relatam que o nosso repertório de comportamentos tem base em nossos estados afetivos, com ênfase na conexão entre emoção, funcionamento social e tomada de decisões, o que tem contribuído para a compreensão do papel do afeto na educação. Para os autores a aprendizagem, a atenção, memória, a tomada de decisão e funcionamento social são afetados pelos nossos estados afetivos.

Os estados de ânimo são considerados emoções de segundo plano e são as que mais influenciam na tomada de decisão conforme Longhi (2011). Dessa forma, este estudo optou-se pela análise de estados de ânimo dos alunos uma vez que eles persistem mais tempo, estando presente durante as interações dos alunos em ambientes de aprendizagem

Os estados de ânimo, fundamentados a partir do trabalho de Scherer (2005), dividem-se em quatro categorias: satisfeito, insatisfeito, animado e desanimado, sendo que cada um possui termos afetivos associados às famílias afetivas de cada categoria, além de possuir intensidade para cada termo. Os estados de ânimo foram desenvolvidos por Scherer (2005) em uma roda de auto-relatos e sua segunda versão desenvolvida por Sacharin, Schlegel, e Scherer (2013), foi adaptada para o português e é chamada de roda de estados de ânimo (REA) 2.0 (Damando e Bercht, 2014). Apoiada nos trabalhos de

Scherer (2005) e Sacharin, Schlegel, e Scherer (2013, pg. 3), a tabela 1 foi desenvolvida e, apresenta os termos afetivos representativos do estado de ânimo desanimado. O estado de ânimo desanimado é representado por 05 famílias afetivas e há termos afetivos correspondentes para cada uma delas.

Tabela 1 - Termos afetivos empregados para o estado de ânimo desanimado

Estado de Ânimo Desanimado	
Família Afetiva	Termos afetivos
Culpa / Arrependimento	Culpado, arrependido, irresponsável, faltoso, pesaroso.
Constrangimento / Vergonha	Envergonhado, acanhado, humilhado, constrangido, dificuldade, desapontamento.
Preocupação/Medo	Tenso, amedrontado, ansioso, apavorado, assustado, receoso, temeroso, intimidado, aflito, preocupado.
Tristeza / Desamparo	Entristecido, abatido, choroso, magoado, melancólico, sofrido, deprimido, aflito, fracassado, frustrado, sentir-se incompetente, pesaroso, mal humorado, incomodado, abatido, arrependido, desesperança.
Piedade/Compaixão	Com pena, piedoso, compaixão, consentimento, dó, lástima, compreensão.

Estar no estado de ânimo desanimado sugere que o aluno demonstra ou reprime de alguma forma, as emoções de tristeza, medo, vergonha ou culpa, bem como os possíveis termos afetivos que o representa Longhi (2011). Além disso, a autora cita que dependendo do estado de ânimo, o mesmo pode conduzir o aluno à desistência, porém, ao ser identificado, o estado de ânimo desanimado é passível de enfrentamento e mudanças, o que pode fazer com que o aluno retome o curso e sua aprendizagem.

Iepsen (2013) e Longhi (2011) citam que os estados emocionais se fazem notar no comportamento não verbal, o que favorece sua inferência ao observar o comportamento do aluno.

É nesse contexto, que se observa a importância da dimensão afetiva no âmbito educacional, sendo discutida de forma a estreitar a dicotomia ainda existente entre cognição e afetividade nos processos de aprendizagem e, no presente trabalho, a possibilidade de sua inferência permitirá um apoio ao professor na identificação do aluno desanimado, possibilitando uma nova mediação do professor junto a esse aluno, podendo ser determinante para evitar desistências em cursos EAD, bem como na retomada dos estudos pelo aluno.

3. Mineração de dados educacionais

Mineração de Dados Educacionais (MDE) é uma área de pesquisa cujo principal foco é desenvolver e/ou adaptar métodos e algoritmos de mineração para explorar conjunto de dados produzidos e coletados em ambientes educacionais (Baker, Isotani e Carvalho, 2011).

Dessa forma, a MDE busca compreender dados em contextos educacionais produzidos por alunos, professores e suas interações, na tentativa de responder perguntas específicas da educação. A MDE pode ser aplicada, por exemplo, para verificar a relação de uma abordagem pedagógica e o aprendizado do aluno, descobrir fatores que influenciam o aprendizado, analisar comportamento do aluno evasor, dentre

outras situações (Baker, Isotani e Carvalho, 2011; Costa *et al.*, 2012; Webber, Zar e Lima, 2013; Rigo *et al.*, 2014).

Apesar de algumas ferramentas educacionais e AVEA oferecerem relatórios aos professores, tutores e gestores o volume de dados gerado é muito grande, tornando difícil extrair informações úteis somente através destes relatórios (Gaudioso e Talavera, 2006). Além disso, os relatórios que podem ser apresentados diretamente pelas ferramentas, apesar de possuírem informações relevantes, podem ser considerados superficiais, diante da quantidade de dados que estão armazenados e, muitas das vezes pode tornar árduo para o professor extrair informação útil através da junção dos relatórios disponíveis (Romero, Ventura e Garcia, 2008). Sendo assim, através da aplicação de MDE obtêm-se dados mais significativos ao propósito desejado.

Porém, para a mineração se efetivar é preciso passar por várias etapas que compreendem o processo de descoberta de conhecimento como pode ser observado na figura 2.



Figura 1 - Processo de descoberta de conhecimento adaptado pelos autores, a partir de (Tan, Steinbach e Kumar, 2009)

Para seguir as etapas supracitadas é preciso saber, inicialmente, qual é o objetivo que se pretende alcançar a fim de escolher técnicas e métodos mais adequados. Pela taxonomia proposta por Baker, Isotani e Carvalho (2011) os métodos são: predição, agrupamento, mineração das relações, destilação de dados para facilitar as decisões humanas e descoberta com modelos. Diante do objetivo da presente pesquisa, ao almejar fazer uma classificação de alunos desanimados, contempla-se o aprendizado de máquina supervisionado, ao obter um modelo para essa classificação e predição.

Cabe ressaltar que a predição é uma técnica usada para prever um valor futuro de um determinado atributo visando desenvolver modelos que deduzam aspectos específicos de dados. Utilizam variáveis da base de dados para prever valores desconhecidos ou valores futuros de outras variáveis de interesse, ou seja, através da predição que se consegue modelos para apoio à inferência a partir da associação dos dados existentes na base (Costa *et al.*, 2012; Rigo *et al.*, 2014). Costa *et al.*, (2012) destacam que através da predição é possível verificar quais dados são mais importantes para o modelo.

As técnicas usadas na predição são classificação e regressão. Para usar o modelo classificador de predição, é preciso trabalhar com duas bases de dados, uma base de treinamento e uma base de teste. Tan, Steinbach e Kumar (2009) explicam que o modelo do conjunto de treinamento é usado para construir um modelo de classificação, que subsequentemente é aplicado ao conjunto de testes, sendo que a avaliação de seu desempenho é baseada na contagem de registros de testes previstos correta e incorretamente. Uma das técnicas de classificação simples e bastante usada é árvore de decisão.

Árvore de decisão é um exemplo de modelo estatístico que usa classificação e predição de dados, além de ser um dos algoritmos mais usados para predição. Dentre os

algoritmos de árvore de decisão optou-se pelo uso do J48, visto ser um dos algoritmos que geram melhores resultados para árvores de decisão, além de se mostrar adequado para procedimentos envolvendo variáveis contínuas e discretas conforme mencionado por Librelotto e Mozzaquatro (2013).

4. Trabalhos relacionados

Pesquisas que associam MDE e Computação Afetiva na busca por modelos e/ou aplicações, com finalidades em aprofundar estudos sobre fenômenos afetivos na educação estão em crescimento (Baker, Isotani e Carvalho, 2011; Romero *et al*, 2006).

A pesquisa desenvolvida por Iepsen (2013) utilizou a MDE para detectar padrões de comportamento associados com o estado afetivo de frustração no aluno. O autor analisou variáveis comportamentais do aluno dentro de um ambiente para ensino de algoritmos. A técnica aplicada foi classificação baseada em regras. Na presente pesquisa, o foco está no aluno desanimado e vale ressaltar que, de acordo com Scherer (2005) o estado de ânimo desanimado engloba famílias e termos afetivos aos quais incluem a frustração. Além disso, a técnica escolhida foi classificação a fim de prever o aluno propenso ao desânimo e tomar ações para recuperar esse aluno antecipadamente.

Majadas, Santos e Boticario (2013) também têm dedicado esforços para a predição do estado afetivo positivo ou negativo de alunos em ambientes educacionais. Os autores realizaram um experimento onde os alunos resolveram problemas de matemática e foram monitorados e coletados dados como tempo médio de pressionar e soltar teclas do teclado, detecção de suas expressões faciais com apoio do *kinect* e de um especialista da área, bem como usaram ao final de cada atividade um auto-relato usando questionário SAM(*Self-Assessment Manequin*). Os dados foram minerados usando árvore de decisão e redes bayesianas. Uma-diferença com relação ao trabalho aqui descrito é que o questionário de auto-relato usado não exige um especialista da área de Psicologia, conforme apontado por Scherer(2005), o que propicia maior autonomia no processo de predição, bem como as variáveis analisadas são vindas do *log* do aluno no ambiente, o que a torna menos invasiva por não ter uso de sensores.

5. Procedimento Metodológico

A presente pesquisa é de natureza aplicada ao ensino e aprendizagem, classificada como exploratória com foco em disciplinas que utilizam AVEA o que permite o acesso aos dados das interações do aluno. Com relação aos procedimentos técnicos, foram desenvolvidos dois experimentos além de apresentar uma abordagem quantitativa obtida através da MDE e técnicas estatísticas com apoio do software SPSS (*Statistical Package for Social Sciences*).

Como instrumentos, foram aplicados dois questionários. O primeiro questionário abordou as questões de como o aluno sentia em diversas situações de aprendizagem através da REA 2.0 e, foi implementada utilizando linguagem PHP e banco de dados MySql e encontra-se disponível no link: <http://www.computacaoafetiva.com/pesquisa/>. O segundo, complementar ao primeiro, contou com questões sobre dificuldade em usar ferramentas, formas de apoio que o professor poderia usar com a turma, formas de solicitação de ajuda ao professor, envio de atividades individuais e em grupo, acompanhamento da disciplina.

Foram aplicados métodos estatísticos utilizando o software SPSS como histogramas de frequência, foi feita a discretização dos dados, geração de quartis, tabulações cruzadas, dentre outras, a fim de encontrar as variáveis mais significativas e preparar os dados para a mineração. A *transformação dos dados* dedicou-se a transpor os dados para o tipo de arquivo a ser utilizado pelo minerador WEKA (*Waikato Environment for Knowledge Analysis*), ou seja, os dados foram transformados em arquivos do tipo ARFF (*Attribute Relation File Format*), e adaptados para o algoritmo utilizado. Somente após as etapas supracitadas é que as *etapas de mineração* e de *interpretação* foram realizadas e são descritas abaixo.

Os dois experimentos foram realizados com uma amostra de 44 alunos, os quais participaram dos dois questionários e que puderam ter seus dados de *logs* analisados durante o semestre de 2014/2. Foram coletadas 26.611 interações nas mais diversas funcionalidades do Moodle para as turmas A e B.

Esse experimento teve como foco o componente de *log* fórum, o qual inclui todos os fóruns criados, com um total de 2622 interações. Ao analisar os *logs* dos alunos com relação ao fórum, constatou-se que os *logs* dos fóruns se dividiram em 3 segmentos (Fórum avisos da professora; Fórum Contatos com a Monitoria e Fórum dos grupos de trabalho), onde somaram respectivamente 404, 149 e 1973 acessos. O *log* de cada fórum é composto por: *fórum_add_post*, *fórum_view_discussion*, *fórum_view_forum*, *fórum_view_discussion*, *fórum_add_discussion*, dentre outros, o que gerou mais de 15 atributos somente relacionados aos fóruns.

Com relação aos questionários aplicados, o questionário que usou a REA 2.0 foi composto por 7 questões, porém foram utilizadas na mineração as três questões iniciais q1: *O que você sente quando envia atividades individuais em atraso*; q2: *O que você sente quando envia atividades em grupo com atraso?*; q3 - *O que você se sente quanto tem dificuldades em realizar atividades individuais?* Em cada questão o aluno escolheu a família afetiva e a intensidade sentimento.

Com relação ao segundo questionário, as questões selecionadas para mineração utilizaram a escala *Likert* (nunca, raramente, algumas vezes, quase sempre e sempre) e abordavam a frequência com que o aluno enviava atividades individuais/grupo em atraso e se sentia dificuldades em realizá-las. A escolha das questões se deu pelas tabulações cruzadas encontradas entre elas através do SPSS às quais mostraram-se relevantes para o estudo.

6. Análise e resultados do experimento

O método usado foi classificação e foi feito utilizando aprendizado supervisionado através do algoritmo de árvore de decisão J48. Foram usados dois métodos, no primeiro experimento foi aplicado o método *Holdout* e no segundo experimento o método *Cross-validation*.

• Método *Holdout*

Para a aplicação do método *Holdout*, os dados foram separados em dois conjuntos de dados sendo um conjunto de treinamento (20 instâncias) e um conjunto de testes (14 instâncias).

A etapa de *interpretação* foi feita pelo especialista da área e usada para suporte de tomadas de decisão na disciplina em questão. Sendo assim, com o conjunto de

treinamento foi possível classificar corretamente 95,45% dos dados. O coeficiente Kappa obtido teve escore de 0,92%. O coeficiente Kappa é usado como uma medida para descrever e testar a confiabilidade e precisão gerando o grau de concordância na classificação. Valores Kappa acima de 0,75 representam excelente concordância (Perroca e Gaidzinski, 2003).

O resultado obtido para o conjunto de teste, representado pela figura 2, mostra que com esse conjunto foi possível classificar corretamente cerca de 91% dos casos. O coeficiente Kappa apresentou escore de 0,85. Juntamente a esses dados, o minerador gerou a matriz de confusão, que auxilia na avaliação do desempenho de um modelo de classificação e é baseada nas contagens de registros de testes previstos correta e incorretamente pelo modelo e, essas contagens são apresentadas através da matriz de confusão (Tan, Steinbach e Kumar, 2009).

```

Correctly Classified Instances      21          91.3043 %
Incorrectly Classified Instances    2           8.6957 %
Kappa statistic                    0.8502
Mean absolute error                0.0348
Root mean squared error            0.1374
Relative absolute error            18.9641 %
Root relative squared error        46.1037 %
Total Number of Instances         23

=== Detailed Accuracy By Class ===
      TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
      0.0000  0.0000  0.0000    0.000  0.0000    0.0000  Satisfeito
      0.5000  0.0000  1.0000    0.500  0.6667    0.9760  Animado
      1.0000  0.0000  1.0000    1.000  1.0000    1.0000  Desanimado
      1.0000  0.1538  0.8333    1.000  0.9091    0.9540  Insatisfeito
      0.0000  0.0000  0.0000    0.000  0.0000    0.9090  Nenhuma
      0.0000  0.0000  0.0000    0.000  0.0000    0.0000  Outra
      0.0000  0.0000  0.0000    0.000  0.0000    0.0000  N_respondeu
Weighted Avg.   0.9130  0.0670  0.8840  0.913  0.8880  0.9740

=== Confusion Matrix ===
 a  b  c  d  e  f  g  <-- classified as
0  0  0  0  0  0  0 | a = Satisfeito
0  1  0  1  0  0  0 | b = Animado
0  0  10  0  0  0  0 | c = Desanimado
0  0  0  10  0  0  0 | d = Insatisfeito
0  0  0  1  0  0  0 | e = Nenhuma
0  0  0  0  0  0  0 | f = Outra
0  0  0  0  0  0  0 | g = N_respondeu

```

Figura 2 - Mineração J48 – experimento 1

A partir da matriz de confusão, das 23 instancias do conjunto de teste 21 foram classificadas corretamente, que inclui as classes Desanimado, Insatisfeito e Nenhuma, às quais obtiveram valores verdadeiro/positivos. Porém, a classe Animado apresentou 01 instância falso/positivo, o que demonstra que um registro foi classificado incorretamente para a classe d(Insatisfeito). Ressalta-se que foram classificados corretamente todos os alunos desanimados.

Porém, nesse caso, a árvore de decisão gerada ficou muito grande, dificultando a sua interpretação, devido à quantidade de atributos (38) utilizados. Apesar de ter obtido bons resultados, uma árvore de decisão extensa, inviabiliza a sua utilização em meio acadêmico para fins de apoio ao professor.

- Método *cross-validation*

A finalidade desse experimento é feita com uma proposta reduzida de atributos, sendo que foram utilizados 19 atributos buscando o enfoque somente nos atributos relacionados ao fórum de grupos, por ter sido o fórum com maior interação dos alunos. Sendo assim, foram minerados atributos da REA 2.0 e do questionário que incluem: q1, q2, q3, quartis das visualizações das discussões, das visualizações nos fóruns e dos posts adicionados aos fóruns de grupos.

Nesse exemplo, foi usado o método *Cross-Validation 10 Folds*, bem como, considerou como atributo classe o q2 – como você se sente ao enviar atividades em grupo atrasadas.

Dessa forma, a *interpretação* do resultado minerado apresentou 72% de corretitude na classificação dos dados, o que pode ser observado através da figura 3. O coeficiente Kappa foi de 0,38, o que indica baixa concordância. A precisão da classificação ao considerar o desânimo, para esse método foi de 0.735, o que indicando uma concordância considerável.

```

=== Stratified cross-validation ===
=== Summary ===

```

Correctly Classified Instances	32	72.7273 %
Incorrectly Classified Instances	12	27.2727 %
Kappa statistic	0.3817	
Mean absolute error	0.1083	
Root mean squared error	0.2671	
Relative absolute error	67.4457 %	
Root relative squared error	97.8016 %	
Total Number of Instances	44	

```

=== Detailed Accuracy By Class ===

```

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0	0	0	0	0	0.221	Satisfeito
	0	0	0	0	0	?	Animado
	0.893	0.563	0.735	0.893	0.806	0.604	Desanimado
	0.538	0.097	0.7	0.538	0.609	0.597	Insatisfeito
	0	0	0	0	0	0.488	Nenhuma
	0	0	0	0	0	?	Outra
	0	0	0	0	0	0.477	N_respondeu
Weighted Avg.	0.727	0.387	0.675	0.727	0.693	0.588	

```

=== Confusion Matrix ===

```

a	b	c	d	e	f	g	<-- classified as
0	0	1	0	0	0	0	a = Satisfeito
0	0	0	0	0	0	0	b = Animado
0	0	25	3	0	0	0	c = Desanimado
0	0	6	7	0	0	0	d = Insatisfeito
0	0	1	0	0	0	0	e = Nenhuma
0	0	0	0	0	0	0	f = Outra

Figura 3 - Mineração J48 – experimento 2

Ao observar a matriz de confusão, das 44 instâncias, 32 foram classificadas corretamente, consideradas verdadeiro/positivos. A classe Desanimado, objeto de estudo, apresentou 25 instâncias classificadas corretamente, porém houve maior número de falsos positivos e classificações incorretas para esta classe, e que pertenciam a outras classes. Porém, a árvore de decisão gerada, apresentada pela figura 4, ficou menor e com uma interpretação melhor, além de serem obtidos resultados que apóia na composição o modelo de predição do desânimo.

```

q1 = Satisfeito: Insatisfeito (1.0)
q1 = Animado: N_respondeu (1.0)
q1 = Desanimado: Desanimado (23.0/4.0)
q1 = Insatisfeito
|   FGrup_viewDiscussion_Quartil = Baixo: Desanimado (7.0/1.0)
|   FGrup_viewDiscussion_Quartil = BaixoMedio: Insatisfeito (2.0)
|   FGrup_viewDiscussion_Quartil = Medio: Desanimado (1.0)
|   FGrup_viewDiscussion_Quartil = Alta: Insatisfeito (8.0/2.0)
q1 = Nenhuma: Nenhuma (1.0)
q1 = Outra: Desanimado (0.0)
q1 = N_respondeu: Desanimado (0.0)

```

Figura 4 - Resultado da árvore em formato de regras - experimento 2

7. Conclusão

O presente trabalho teve como objetivo prever o estado de desânimo de alunos em interação em um AVEA a partir da classificação dos seus dados e interações,

fomentando a importância em considerar estados afetivos que podem levar à desistência do aluno no curso.

A mineração foi feita com dois experimentos, sendo que no método *Holdout* apresentou uma taxa de acerto de 91% dos alunos propensos ao desânimo, o segundo experimento, utilizou o método *Cross-Validation* e obteve 72% de acurácia. Apesar de o primeiro experimento apresentar uma acurácia maior, a árvore de decisão gerada foi de difícil interpretação pelo especialista. Já no segundo método, diminuído a quantidade de atributos, a árvore de decisão apresentada e analisada através de suas regras mostrou-se significativa, mesmo apresentando uma acurácia menor comparada com o primeiro método, porém pode ser interpretado mais facilmente pelo professor ou implementadas computacionalmente, possibilitando que o mesmo identifique um aluno propenso ao desânimo.

Apesar da limitação da base de dados, seus resultados mostraram-se representativos nesta amostra, para a predição do desânimo, o que não descarta a necessidade do modelo ser reavaliado com um conjunto maior de dados. Dessa forma, sugere-se as regras apresentadas pela árvore de decisão possam ser usadas para compor um modelo computacional genérico que identifique alunos com propensão ao estado de ânimo desanimado e, ao final, apresente os alunos identificados nesse modelo computacional para o professor, a fim de que ele possa acompanhar esses alunos apropriadamente.

Pretende-se, em trabalhos futuros, desenvolver o modelo computacional genérico para a identificação do desânimo e, ao mesmo tempo, apresentar ao professor os alunos desanimados utilizando modelos gráficos em forma de *dashboards*.

Referencias Bibliográficas

- Baker, R. S. J.; Isotani, S.; Carvalho, A. M. J. B. (2011). Mineração de Dados Educacionais: Oportunidades para o Brasil. *Revista Brasileira de Informática na Educação*. V.19, N.02.
- Bercht, M. (2001). Em Direção a Agentes Pedagógicos com Dimensões Afetivas. Tese de doutorado. Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação. Tese de doutorado.
- Costa, E.; Baker, R.; Amorin, L.; Magalhães, J.; Marinho, T. (2012). Mineração de Dados Educacionais: Conceitos, Técnicas, Ferramentas e Aplicações. *Jornada de Atualização em Informática na Educação – JAIE*.
- Damando, F. S.; Bercht, M. (2014) Análise da disposição afetiva do aluno em um AVEA: descoberta do desânimo. *XXV Simpósio Brasileiro de Informática na Educação*.
- Gaudioso, E.; Talavera, L. (2006). Data mining to support tutoring in virtual learning communities: experiences and challenges. In *Data mining in e-learning*. WIT Press.
- Iepsen, E. F. (2013). Ensino de Algoritmos: Detecção do Estado Afetivo de Frustração para Apoio no Processo de Aprendizagem. Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Informática na Educação. Tese de Doutorado.
- Inventado, P. S.; Legaspi P. Cabredo, R.; Numa, M. (2013) Modeling Affect in Student-driven Learning Scenarios. *The 6th International Conference on Educational Data Mining*.

- Jaques, P. A.; Nunes, M. A. S. N. (2012). Ambientes inteligentes de aprendizagem que inferem, expressam e possuem emoções e personalidade. Jornada de Atualização em Informática na Educação. 23º Congresso Brasileiro de Informática na Educação.
- Librelotto, S. R.; Mozzaquatro, P. M. (2013). Análise dos Algoritmos de Mineração J48 e Apriori Aplicados na Detecção de Indicadores de Qualidade de Vida. *Revint – Revista Interdisciplinar de Ensino, Pesquisa e Extensão* V.01, N.01.
- Longhi, M. T. (2011). Mapeamento de Aspectos Afetivos em um Ambiente Virtual de Aprendizagem. Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Informática na Educação. Tese de doutorado.
- Majadas, S. S.; Santos, O.; Boticario, J. G. (2013) Affective state detection in educational systems through mining multimodal data sources. The 6th International Conference on Educational Data Mining.
- Pardos, Z, A.; Baker, R. S. J. D.; Pedro, M.S.; Gowda, Sujith. M.; Gowda, Supreeth, M. (2014). Affective States and State Tests: Investigating How Affect and Engagement during the School Year Predict End-of-Year Learning Outcomes. In *Journal of Learning Analytics*. V.01, N.01.
- Perroca, M. G.; Gaidzinski, R.R. (2003). Avaliando a Confiança Interavaliadores de um Instrumento para Classificação de Pacientes - Coeficiente Kappa. *Revista da Escola de Enfermagem USP*. V. 37, N. 1. Scielo.
- Piaget. J. (2014). *Relações Entre a Afetividade e a Inteligência no Desenvolvimento Mental da Criança*. Organização e Tradução: Saltini e Cavenaghi. Rio de Janeiro. Wak Editora.
- Picard, R. W. (2000). *Affective Computing*. The MIT Press.
- Rigo, S. J.; Cambuzzi, W.; Barbosa, J. L. V.; Cazella, S. C. (2014). Aplicações de Mineração de Dados Educacionais e Learning Analytics com foco na evasão escolar: oportunidades e desafios. *Revista Brasileira de Informática na Educação*. V. 22, N.01.
- Romero, C. M.; Ventura, S.; García, E. (2008). Data Mining in Course Management Systems: Moodle Case Study and Tutorial. *Computers & Education* 51, pg. 368-384, Elsevier.
- Sacharin, V.; Schlegel, K.; Scherer, K. R. (2013). Geneva Emotion Wheel Rating Study. Geneva, Switzerland: University of Geneva, Swiss Center for Affective Sciences.
- Scherer, K. R. (2005). What are the emotions? And how can they measured? In: *Social Science Information*.
- Tan, Pang-Ning; Steinbach, M.; Kumar, V. (2009). *Introdução ao DataMining Mineração de Dados*. Rio de Janeiro: Editora Ciência Moderna Ltda.
- Webber, C. G.; Zat, D.; Lima, M. F. W. P. (2013). Utilização de algoritmos de agrupamento na mineração de dados educacionais. *Revista Renote: Novas tecnologias na educação*. V.11, Nº 1.