Um modelo preditivo para diagnóstico de evasão baseado nas interações de alunos em fóruns de discussão

Francisco da Conceição Silva[1], Luis Carlos Costa Fonseca[1], Reinaldo de Jesus Silva[1], Josenildo Costa da Silva[2]

[1]Departamento de Engenharia da Computação
Universidade Estadual do Maranhão (UEMA) – São Luis – MA – Brazil

[2]Departamento Acadêmico de Informática
Instituto Federal de Educação, Ciência e Tecnologia do Maranhão (IFMA)
São Luis – MA – Brasil

{fransconcfs, lccfonseca, reinaldosilvarrrb}@gmail.com,
jcsilva@ifma.edu.br

Abstract: *This paper presents the results of a survey that is to create a predictive model of data mining in a Virtual Learning Environment (VLE). The obtained model aims to make the diagnosis of students evasion, based on their interactions in discussion forums, to help stakeholders in decision making. For this, experiments were performed with different sets of data being applied five classification algorithms and compared the performance of each. The results showed that the techniques based on the decision tree has the best performances and in some cases, achieved accuracy rate above 73%.*

Resumo: *Este trabalho apresenta os resultados de uma pesquisa que consiste em criar um modelo preditivo de mineração de dados em um Ambiente Virtual de Aprendizagem (AVA). O modelo obtido visa realizar o diagnóstico de evasão de alunos, baseado nas suas interações em fóruns de discussão, para auxiliar as partes interessadas na tomada de decisão. Para isso, foram realizados experimentos com conjuntos de dados distintos, sendo aplicados cinco algoritmos de classificação e foi comparado o desempenho de cada um. Os resultados obtidos apontaram que as técnicas baseadas em árvore de decisão tiveram os melhores desempenhos e, em alguns casos, alcançaram taxa de precisão acima de 73%.*

1. Introdução

As ferramentas colaborativas e de comunicação tem sido usadas largamente nos contextos educacionais e os Ambientes Virtuais de Aprendizagem (AVA's), que são uma modalidade de Ensino a Distância (EAD), estão cada vez mais sendo inseridos em universidades, escolas e empresas. Essa comunicação ocorre de diversas formas, tais como *chats*, fóruns de discussão, *wikis*, dentre outras. Os fóruns, especialmente, consistem em espaços para

DOI: 10.5753/cbie.sbie.2015.1187

discussões e trocas de ideias sobre assuntos definidos por seus participantes, possibilitando uma experiência de aprendizagem favorável ao processo pedagógico [Abawajy, 2012].

Na EAD existe um problema recorrente e muito desafiador, que é a evasão de alunos. Conforme Abbad *et al*. (2010), as altas taxas de desistência estão preocupando as instituições que oferecem esta modalidade de ensino. Diante disso, cabem alguns questionamentos: quais fatores influenciam, direta ou indiretamente, na evasão e reprovação de alunos? Como deve ser um modelo computacional que se proponha fazer esse diagnóstico?

Uma característica comum em AVA's é sua capacidade de coletar e armazenar uma grande quantidade de dados sobre os alunos, como por exemplo, registros de acesso, interações diversas com o sistema e as mensagens em fóruns. Este grande volume de dados possibilitou, em paralelo ao avanço das técnicas de mineração de dados, um estudo mais acurado do fenômeno da evasão. Identificar as informações relevantes em base de dados como essa não é uma tarefa fácil e o professor que deseja obter informações a respeito do desempenho dos alunos terá à sua disposição tabelas e dados de difícil interpretação. Para um gestor educacional seria de grande relevância poder identificar alunos desmotivados ou prestes a abandonar o curso.

No intuito de possibilitar essa identificação, o presente trabalho apresenta um modelo preditivo para diagnóstico de evasão em AVA's, a partir das interações de alunos em fóruns de discussão, para servir como ponto de partida aos interessados diretos (professores, tutores, gestores, etc.) na tomada de decisão. Para isso, foi realizado um estudo bibliográfico visando ampliar a compreensão sobre evasão na EAD, investigado e definido as técnicas de mineração de dados que foram utilizadas na elaboração do modelo preditivo proposto e foi feito um estudo da base de dados do Moodle para obter os dados necessários aos experimentos realizados.

O trabalho está organizado da seguinte forma: na **seção 2** é feita uma revisão de literatura, abordando sobre a evasão na EAD, mineração de dados educacionais e interações de alunos em fóruns. Na **seção 3**, são apresentados alguns trabalhos relacionados a esta pesquisa. Na **seção 4** são apresentados os procedimentos metodológicos para mineração de dados. Na **seção 5**, temos a análise dos resultados obtidos nos experimentos realizados. Na **seção 6** temos as conclusões e trabalhos futuros e, por fim, as referências.

2. Revisão de Literatura

Esta seção apresenta uma visão geral a respeito dos tópicos tratados neste trabalho, mostrando o que há de pesquisas em cada assunto abordado.

2.1 Evasão na EAD

A evasão de alunos é um dos principais problemas na EAD, gerando muitas pesquisas na literatura internacional e nacional [Abbad *et al*., 2010] e refere-se à desistência definitiva do aluno em qualquer etapa do curso. A **tabela 1** mostra a distribuição dos índices de evasão registrados entre 2010 e 2013, segundo o censo da EAD [Abed, 2013], onde no ano de 2013

foi registrado um aumento significativo nos índices de evasão em todos os tipos de cursos apresentados, sendo que o índice médio em cursos regulamentados totalmente a distância foi de 19% e de semipresenciais foi de 14,6%.

Tabela 1. Índices de evasão registrados no período 2010-2013 no Censo EAD-BR [ABED, 2013, p.32].

Tipo de curso	2010	2011	2012	2013
Autorizados pelo MEC	18,6%	20,5%	11,74%	16,94%
Livres não corporativos	22,3%	23,6%	10,05%	17,08%
Livres corporativos	7,6%	20%	3%	14,62%
Disciplinas EAD	—	17,6%	3,10%	10,49%

Para Romero *et al.* (2012), o fracasso escolar é conhecido como "o problema das mil causas", dado que uma série de fatores do tipo pessoal, acadêmico (quanto maior a escolaridade, menor a evasão), físico, familiar, social, dentre outros, podem influenciar no fracasso ou abandono de curso por parte dos alunos.

A identificação de indicadores de evasão em cursos a distância é importante para que seja possível proporcionar as condições necessárias que a reduza ou elimine. Para isso, são necessários métodos e ferramentas de análise de dados a fim de observar o comportamento dos alunos para auxiliar as partes interessadas na tomada de decisão. A Mineração de Dados Educacionais (MDE) é uma abordagem indutiva muito interessante que cria modelos para descobrir automaticamente informações ocultas presentes nos dados dos alunos que podem ser utilizadas na melhoria da aprendizagem [Romero *et al*., 2008].

2.2 Interações em fóruns de discussão

Os fóruns são uma poderosa ferramenta de comunicação assíncrona e consistem em espaços de discussões e trocas de ideias a respeito de temas propostos por seus participantes. Sendo um espaço aberto para alunos e professores, os fóruns devem ser utilizados como estratégia de comunicação e diálogo entre tais atores, fazendo com que eles se movimentem na busca de entendimento e produção do saber [Scherer, 2009].

Pesquisas recentes indicam que, quando bem concebidos, os fóruns motivam e melhoram a experiência de aprendizagem dos participantes, favorecendo o processo pedagógico e possibilitando ao aluno lograr êxito em cursos a distância [Abawajy, 2012]. Conforme este autor, o objetivo final de fóruns de discussão assíncrono é criar um ambiente de aprendizagem online para atingir altos níveis de aprendizagem e para isso, aponta algumas das principais características que podem influenciar e diferenciar os vários tipos de fóruns de discussão:

- grau de interação;
- requisitos de participação;
- volume e frequência de postagens;

- *feedback* do professor e demais participantes.

A discussão é normalmente considerada uma poderosa ferramenta para desenvolver habilidades pedagógicas, como o pensamento crítico, colaboração e reflexão. Os fóruns de discussão oferecem muitas vantagens pedagógicas, como incentivo à reflexão, análise e pensamento de ordem superior.

3. Trabalhos Relacionados

Esta seção descreve alguns trabalhos relacionados ao propósito desta pesquisa e que são de grande relevância na área da MDE.

Kampff (2009) propõe uma arquitetura para sistemas de alertas, tanto pré-definidos como gerados a partir de mineração de dados, para identificar comportamentos e características de alunos com risco de evasão ou reprovação e alertar o professor. Para a autora, a avaliação da arquitetura apontou que as intervenções realizadas contribuíram para a redução dos índices de reprovação e evasão dos alunos na disciplina acompanhada.

Em Romero *et al.* (2013) há o desenvolvimento de uma ferramenta de mineração específica para o Moodle, com o objetivo de predizer notas finais de alunos em cursos nesse ambiente. Alguns experimentos foram realizados usando dados disponíveis (numéricos) e filtrados (filtro por linha, filtro por coluna e balanceados) de 438 alunos de 7 cursos de engenharia no Moodle da Universidade de Cordova. Para o autor, os resultados dos modelos obtidos usando dados categóricos e algoritmos de árvores de decisão são mais compreensíveis que aqueles que usam dados numéricos porque favorecem a compreensão e permitem ao professor fazer uma interpretação do modelo obtido para a tomada de decisão.

Santana (2014) propôs a aplicação de técnicas de classificação em um conjunto de dados de alunos de um curso na modalidade a distância para obter resultados como forma de apoio à tomada de decisão. Para isso, comparou algumas técnicas de classificação na interação Perfil de Uso do AVA, tendo como variável alvo o desempenho do aluno. Os resultados que foram considerados satisfatórios pelo autor foram gerados pela aplicação da técnica de árvore de decisão J48, que alcançou taxa de 74% de precisão.

Nesse sentido, o trabalho apresentado aqui busca ampliar as contribuições dos trabalhos correlatos, uma vez que, além do modelo preditivo, resultante da comparação de desempenho de algumas técnicas de mineração, propõe como trabalho futuro a implementação de uma ferramenta para exibição do diagnóstico gerado pelo modelo, com base nas técnicas de árvores de decisão, para auxiliar as partes interessadas na tomada de decisão, sendo que esta ferramenta já está em desenvolvimento e tem uma versão funcional para exibição de diagnóstico.

4. Procedimentos Metodológicos para os Experimentos de Mineração de Dados

Para este trabalho foram realizados experimentos utilizando dois conjuntos de dados com características diferentes, obtidos de um curso técnico em um Ava de uma instituição de ensino. O curso escolhido foi o de Técnico em Informática, composto de 25 disciplinas e com um total de 1171 alunos. Dentre essas disciplinas, foi selecionada a de Sistemas

Operacionais, oferecida no segundo módulo. Ela disponibilizou três fóruns de discussão e do quantitativo de alunos do curso, considerou-se apenas 502 para a criação dos conjuntos de dados, porque eles participaram em algum momento da disciplina.

A **figura 1** mostra a arquitetura do modelo computacional proposto neste trabalho e os parágrafos seguintes fazem sua descrição.

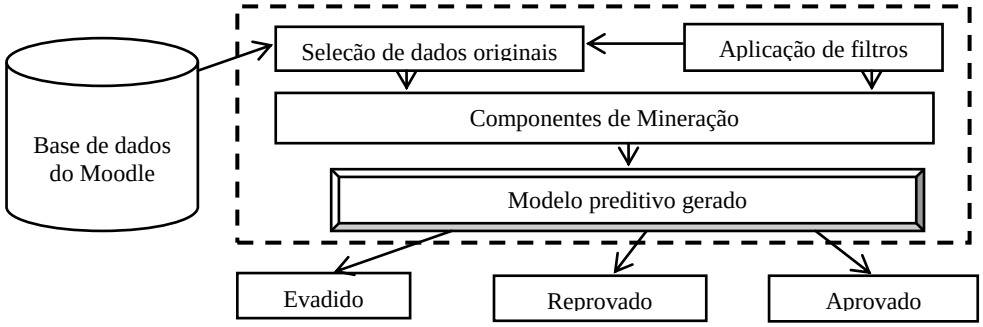

Figura 1. Arquitetura do modelo preditivo para diagnóstico de evasão no Moodle.

O primeiro conjunto de dados, com 502 alunos, foi obtido pela seleção realizada a partir das tabelas e atributos da base de dados do Moodle que registram as interações dos alunos nos fóruns de discussão, bem como no próprio AVA da instituição de ensino. A definição dos atributos de cada tupla foi realizada manualmente e considerou o quantitativo de participação em fóruns e as postagens nas discussões, conforme sugerido por Abawajy (2012), por representarem fortes indicadores de interação nesta atividade, sendo que a tabela de sumarização ficou com 9 atributos (ver **tabela 2**).

Tabela 2. Atributos da tabela de sumarização.

Atributo	Descrição
nota_media_em_foruns	nota média do aluno nos fóruns
qtd_posts_forum1	quantidade de postagens no fórum 1
qtd_posts_forum2	quantidade de postagens no fórum 2
qtd_posts_forum3	quantidade de postagens no fórum 3
total_foruns_por_aluno	total de fóruns que o aluno participou
total_posts_em_todos_os_foruns	total de postagens em todos os fóruns
media_posts_por_forum	média de postagens por fórum
desempenho	desempenho do aluno nos fóruns [Ruim, Regular, Bom e Excelente]
resultado	resultado final do aluno na disciplina (atributo classe)

A distribuição das classes referentes ao resultado do primeiro conjunto de dados está assim apresentada: Aprovado Por Média = 288 (57,37%), Aprovado Por Final = 148 (29,48%), Reprovado Por Média = 63 (12,55%) e Reprovado Por Final = 3 (0,6%). Estas classes foram transformadas em apenas duas, sendo que foi mantida a classe Aprovado Por

Média (doravante chamada de **A.P.M)**, enquanto que as demais foram transformadas na classe Tendência a Baixo Desempenho (doravante chamada de **T.B.D**). Com isso, uma nova distribuição foi obtida, conforme mostrado na **tabela 3**.

Tabela 3. Distribuição das classes referentes ao resultado do primeiro conjunto de dados.

Resultado	Total de Alunos	Percentual
A.P.M	288	57,37%
T.B.D	214	42,63%
TOTAL	**502**	**100%**

Tal procedimento foi adotado porque a aprovação por média é o que se espera de um aluno em uma situação ótima, no entanto, quando isso não acontece, significa que alguns fatores influenciaram para que o mesmo tivesse desempenho abaixo do esperado. Desta forma, as três classes foram consideradas uma só por elas representarem os alunos nesta condição e também porque o baixo desempenho é um forte indício de evasão.

O segundo conjunto de dados foi obtido a partir do primeiro através de um filtro por linha, que consistiu em manter apenas os alunos que participaram de pelo menos um fórum da disciplina, o que resultou em apenas 353 alunos nessa condição, sendo mantidos os 9 atributos do conjunto de dados anterior e a distribuição das classes referentes ao resultado ficou assim: A.P.M = 247 (69,97%) e T.B.D = 106 (30,03%).

Para a mineração de dados deste trabalho fez-se uso da ferramenta Weka (*Waikato Environment for Knowledge Analyis*), que é um software que disponibiliza uma variedade de algoritmos para as mais diversas tarefas de aprendizado de máquina. Os modelos gerados a partir de árvores de decisão utilizaram os algoritmos **J48**, **BFTree** e **SimpleCart**, e os baseados em estatística utilizaram os algoritmos **BayesNet** e **NaiveBayes**.

Foi aplicada ainda uma técnica de validação de dados, chamada *cross validation 10-folds*. Esta técnica consiste em dividir os dados em dez partições aleatórias, onde nove dessas partições são destinadas ao conjunto de treinamento e uma partição para o conjunto de testes [Romero *et al*., 2013], possibilitando melhor avaliação do modelo obtido.

O objetivo dos experimentos foi verificar se um modelo preditivo com maior precisão para o diagnóstico de evasão seria obtido usando dados filtrados ou dados originais. Além disso, buscou-se verificar quais indicadores das interações em fóruns de discussão influenciaram no resultado final do aluno na disciplina.

5. Análise dos Ressultados

Esta seção apresenta a análise dos resultados obtidos nos experimentos realizados com os conjuntos de dados e as técnicas de mineração definidos em seções anteriores, onde é feita a comparação das taxas de precisão dos algoritmos nos conjuntos de dados usados, cujos índices são apresentados na **tabela 4**, é definida a técnica de mineração que obteve melhor desempenho e são apresentados os principais indicadores que podem influenciar nas tendências tratadas neste trabalho (evasão, reprovação e aprovação).

Tabela 4. Precisão dos classificadores em cada conjunto de dados (% de acerto).

		Conjuntos de dados	
Método	Algoritmo	originais	filtrados por linha
Árvores de Decisão	J48	**72,1116**	**73,3711**
	BFTree	70,5179	72,5212
	SimpleCart	71,5139	71,9547
Bayesianos	NaiveBayes	**72,7092**	70,8215
	BayesNet	71,9124	68,8385

O primeiro experimento usou dados originais e a taxa de acerto dos algoritmos ficou acima de 70%, sendo que o maior índice de precisão no experimento foi registrado pelo algoritmo NaiveBayes (72,7092%), seguido pelo algoritmo J48 (72,1116%). Os índices obtidos mostram que para dados originais, as duas técnicas tiveram desempenho muito semelhantes, com uma leve vantagem das técnicas bayesianas em relação às baseadas em árvore de decisão. No segundo experimento, usando dados filtrados por linha, percebe-se uma inversão em relação ao experimento anterior, visto que os melhores desempenhos foram obtidos pelas técnicas baseadas em árvores de decisão (todas acima de 71%), sendo que as técnicas bayesianas ficaram abaixo de 71% e a maior taxa de acerto no experimento foi registrada pelo algoritmo J48, que ultrapassou os 73%.

Outra informação essencial para a análise dos experimentos realizados refere-se à matriz de confusão. Ela é uma característica que trata da qualidade da classificação realizada, mostrando quais foram os casos em que o algoritmo se confundiu na classificação. A diagonal principal da matriz apresenta as instâncias classificadas corretamente para cada classe e corresponde à precisão do classificador. As **figuras 2 e 3** mostram a matriz de confusão do algoritmo J48 em cada experimento realizado.

```
   a    b   <-- classified as
 133   81 |   a = T.B.D
  59  229 |   b = A.P.M
```

Figura 2. Matriz de confusão do algoritmo J48 no experimento com dados originais.

Ao verificar a matriz de confusão gerada pelo algoritmo J48 no experimento com dados originais (**figura 2**) percebe-se que 81 instâncias foram classificadas erroneamente como sendo de **b**, sendo que eram de **a**. Da mesma forma, 59 instâncias foram classificadas erroneamente como sendo de **a**, quando, na realidade, eram de **b**.

```
   a    b   <-- classified as
  26   80 |   a = T.B.D
  14  233 |   b = A.P.M
```

Figura 3. Matriz de confusão do algoritmo J48 no experimento com dados filtrados por linha.

A **figura 3** apresenta a matriz de confusão gerada pelo algoritmo J48 em dados filtrados por linha, onde vê-se que a grande maioria das instâncias de **a** foram classificadas erroneamente como sendo de **b**, correspondendo a 80 instâncias e apenas 26 foram

classificadas corretamente; por outro lado, apenas 14 instâncias foram classificadas como sendo de **a**, quando, na verdade, eram de **b**.

A matriz de confusão mostra se o algoritmo está errando muito mais em uma classe do que na outra, principalmente quando se trata de dados desbalanceados, como é o caso do conjunto de dados filtrados por linha (**figura 3**), que apresentou um desbalanceamento de suas classes e isso refletiu no grande número de instâncias classificadas erroneamente como **b**, quando eram **a**. Além disso, a precisão do classificador fica menor quando há menos dados disponíveis [Romero *et al*., 2013].

As técnicas baseadas em árvores de decisão obtiveram índices significativos de precisão nos conjuntos de dados utilizados neste trabalho, sendo que o algoritmo J48 foi o que obteve melhor desempenho, apesar de uma pequena diferença com o algoritmo NaiveBayes. Através das regras de classificação geradas por ele é possível apontar quais fatores são mais indicativos para diagnosticar alunos com tendência à evasão e reprovação.

```
nota_media_em_foruns <= 4.67
|   total_foruns_por_aluno <= 0: T.B.D (149.0/41.0)
|   total_foruns_por_aluno > 0
|   |   total_foruns_por_aluno <= 2
|   |   |   qtd_posts_forum3 <= 0
|   |   |   |   qtd_posts_forum2 <= 0
|   |   |   |   |   nota_media_em_foruns <= 2.75: A.P.M (18.0/6.0)
|   |   |   |   |   nota_media_em_foruns > 2.75
|   |   |   |   |   |   qtd_posts_forum1 <= 1: T.B.D (22.0/5.0)
|   |   |   |   |   |   qtd_posts_forum1 > 1
|   |   |   |   |   |   |   nota_media_em_foruns <= 3.08: T.B.D (3.0/1.0)
|   |   |   |   |   |   |   nota_media_em_foruns > 3.08: A.P.M (2.0)
|   |   |   |   qtd_posts_forum2 > 0
|   |   |   |   |   nota_media_em_foruns <= 4: A.P.M (4.0)
|   |   |   |   |   nota_media_em_foruns > 4: T.B.D (5.0)
|   |   |   qtd_posts_forum3 > 0: A.P.M (3.0)
|   |   total_foruns_por_aluno > 2: T.B.D (7.0)
nota_media_em_foruns > 4.67: A.P.M (289.0/69.0)
```

Figura 4 - Árvore de decisão gerada pelo algoritmo J48 com dados originais.

No primeiro experimento o algoritmo J48 gerou uma árvore de decisão composta por 10 regras de classificação (ver **figura 4**) e elas indicam que quando a **nota media em fóruns** foi maior que 4.67, o aluno foi aprovado por média sem depender de outras variáveis. Abaixo dessa média, o algoritmo identificou, por exemplo, que quando não houve participação em fórum (mais de 20% dos casos), ocorreu a classe T.B.D.

Conforme tratado por Abawajy (2012), os fóruns podem motivar e melhorar a experiência de aprendizagem dos participantes, favorecendo o processo pedagógico, além de possibilitar ao aluno lograr êxito em cursos a distância. Quando isso não ocorre, pode acontecer da não participação em fórum produzir o efeito contrário ao que diz este autor, como foi o caso apontado pelo algoritmo nesse experimento.

Variáveis como o **número de fóruns que o aluno participou** e a **quantidade de postagens nos fóruns** foram aspectos importantes para predizer o resultado final na disciplina, quando a nota média em fóruns foi menor ou igual a 4.67.

Em relação ao segundo experimento, o algoritmo J48 gerou uma árvore de decisão composta por 7 regras (ver **figura 5**) e elas indicam, principalmente, que a **nota média em fóruns**, **o total de fóruns por aluno**, **quantidade de postagens nos fóruns 2 e 3**, foram fatores importantes para definir o resultado do aluno na disciplina.

```
nota_media_em_foruns <= 5.33
|   total_foruns_por_aluno <= 2
|   |   qtd_posts_forum3 <= 0
|   |   |   qtd_posts_forum2 <= 0
|   |   |   |   nota_media_em_foruns <= 2.75: A.P.M (18.0/6.0)
|   |   |   |   nota_media_em_foruns > 2.75: T.B.D (27.0/8.0)
|   |   |   qtd_posts_forum2 > 0
|   |   |   |   qtd_posts_forum2 <= 1: A.P.M (13.0/4.0)
|   |   |   |   qtd_posts_forum2 > 1: T.B.D (4.0/1.0)
|   |   qtd_posts_forum3 > 0: A.P.M (6.0/1.0)
|   total_foruns_por_aluno > 2: T.B.D (10.0)
nota_media_em_foruns > 5.33: A.P.M (275.0/63.0)
```

Figura 5 - Árvore de decisão gerada pelo algoritmo J48 com dados filtrados por linha.

Uma observação exclusiva para o contexto desses experimentos é que o algoritmo considerou a participação em mais de dois fóruns, em alguns casos, como sendo um indício de evasão e/ou reprovação, tanto que quando o total de fóruns por aluno foi superior a 2, esses alunos foram definidos como sendo da classe T.B.D.

6. Conclusões e Trabalhos Futuros

Este trabalho investigou como os dados armazenados em um AVA podem ser transformados em informações potencialmente úteis para apoiar o acompanhamento de alunos em cursos EAD. Através dos experimentos realizados buscou-se obter um modelo preditivo com alta precisão que fosse capaz de predizer quando um aluno apresenta características tendenciosas à evasão a partir de suas interações em fóruns de discussão.

As técnicas baseadas em árvores de decisão são recomendadas no contexto educacional, uma vez que elas geram um resultado mais compreensível e fácil de interpretar ao usuário que as utilizar para a tomada de decisão [Romero *et al*., 2013] e nos experimentos deste trabalho observou-se que elas obtiveram os melhores desempenhos. Desta forma, são as mais indicadas, dentre as testadas, para a geração de um diagnóstico mais preciso das tendências de evasão e reprovação de alunos.

Embora o desempenho do algoritmo J48 tenha sido melhor com dados filtrados, em termos de precisão, o modelo errou muito ao classificar instâncias rotuladas como T.B.D. Em relação aos dados originais, o modelo conseguiu classificar de modo mais equilibrado exemplos das duas classes.

Como trabalhos futuros, pretende-se possibilitar melhor compreensão e interpretação dos resultados do modelo obtido. Para isso, apresentam-se alguns desafios: criação de uma ferramenta para mostrar o diagnóstico ao público interessado para cada um dos resultados trabalhados neste estudo; apresentar a situação individual de cada aluno e integrar essa ferramenta como um módulo do Moodle. Vale ressaltar ainda que esta ferramenta já está em desenvolvimento e tem uma versão funcional que utiliza as regras de classificação geradas por modelos baseados em árvores de decisão. O foco deste trabalho foi especificamente nas interações em fóruns de discussão e, como meta futura, pretende-se ainda unir dados de diversos recursos do Moodle para uma análise de caráter mais amplo.

Através da mineração de dados em AVA's é possível verificar a relação entre uma abordagem pedagógica e o aprendizado do aluno, a fim de que o professor avalie se sua abordagem realmente está ajudando ou não o aluno a ter um bom desempenho nas atividades propostas.

Agradecimentos

Este trabalho foi financiado pela Fundação de Amparo à Pesquisa do Maranhão - FAPEMA e Fundação de Apoio ao Ensino, Pesquisa e Extensão - FAPEAD. O trabalho teve ainda como co-autor Lanylldo Araujo dos Santos (Universidade Estadual do Maranhão - UEMA).

Referências

ABAWAJY, J. Analysis of Asynchronous Online Discussion Forums for Collaborative Learning. **International Journal of Education and Learning, v. 1, n. 2, p**. 2012.

ABBAD, Gardênia da Silva; ZERBINI, Thaís; SOUZA, Daniela Borges Lima de. **Panorama das pesquisas em educação a distância no Brasil**. In: Estudos de Psicologia, 15(3), setembro-dezembro/2010. Disponível em: < http://www.scielo.br/pdf/epsic/v15n3/a09v15n3.pdf >. Acesso em: 15 out. 2015.

ASSOCIAÇÃO BRASILEIRA DE EDUCAÇÃO A DISTÂNCIA. **Censo EAD.BR. Relatório Analítico da aprendizagem a distância no Brasil 2013**. Disponível em: http://www.abed.org.br/censoead2013/CENSO_EAD_2013_PORTUGUES.pdf. Acesso em 24 abr 2015.

KAMPFF, A. J. C. **Mineração de Dados Educacionais para Geração de Alertas em Ambientes Virtuais de Aprendizagem como Apoio à Prática Docente**. 2009. Tese (Doutorado) – PPGIE/UFRGS.

MOODLE Org. Community driven, globally supported (2015). Disponível em: <https://moodle.org/>. Acesso em: 20 abr. 2015.

ROMERO, C., VENTURA, S., G, ENRIQUE. (2008). **Data mining in course management systems:** Moodle case study and tutorial. Computers & Education, 51(1): 368–384.

_____, VENTURA, S., PECHENIZKIY, M., BAKER, R.S.J.d. **Handbook of Educational Data Mining,** Ed. C R C, 2011, 503p**.**

_____, VERA, C., M., SOTO, S., V. Predicción del Fracaso Escolar mediante Técnicas de Minería de Datos. **IEEE-RITA Vol. 7, Núm. 3, Nov. 2012**.

_____. ESPEJO, P., G, P. G., ZAFRA . A. (2013). **Web usage mining for predicting final marks of students that use Moodle courses**. Disponível em: <https://www.deepdyve.com/lp/wiley/web-usage-mining-for-predicting-final-marks-of-students-that-use-moodle-courses-P23vC7HY6h/1>. Acesso em 23 de abril de 2015.

SANTANA, Leandro, C. (2014). Avaliação do Perfil de Uso no Ambiente Moodle Utilizando Técnicas de Mineração de Dados. **Revista Brasileira de Informática na Educação. Pág. 269**.

SCHERER, Suely. **Educação bimodal: habitantes, visitantes ou transeuntes?** In: VALENTE, J. A. e BUSTAMANTE, S. B. V. Educação a Distância: prática e formação do profissional reflexivo. São Paulo: Avercamp, 2009. 259 p.