

Os Desafios para Minerar Dados Educacionais de Forma Rápida e Intuitiva: o Caso da DAMICORE e a Caracterização de Alunos em Ambientes de eLearning

Luis Fernando de S. Moro, Aparecida M. Zem Lopes, Alexandre C. B. Delbem, Seiji Isotani

Instituto de Ciências Matemáticas e de Computação – USP
Avenida Trabalhador são-carlense, 400 – Centro - CEP: 13566-590 - São Carlos - SP

{luismoro, cidazem, acbd, sisotani}@icmc.usp.br

***Abstract.** Although much of the interactions between students and teaching support systems are stored, the tools that perform analysis of these data is complex and difficult to use. Researchers have been working on these research data in order to characterize the behavior by some students (gaming the system, off-task behavior, withdrawal from the course, among others). In this context, this article aims to address the challenge of enabling the use of tools and techniques for EDM teachers, so that, from the characterization of students, are entitled to carry out interventions in the process of teaching and learning. Several tools Educational Data Mining were studied, as well as a search is conducted on the Human-Computer Interaction, in order to endow the system with intuitive interfaces and teachers to bring their students on characterizations.*

***Resumo.** Apesar de grande parte das interações entre alunos e sistemas de apoio pedagógicos serem armazenadas, as ferramentas que realizam a análise desses dados são complexas e de difícil utilização. Pesquisadores têm trabalhado na investigação desses dados, visando caracterizar os alunos segundo alguns comportamentos (gaming the system, off-task behavior, desistência do curso, entre outros). Neste contexto, este artigo tem por objetivo abordar o desafio de viabilizar o uso de ferramentas e técnicas de EDM para professores, de forma que, a partir da caracterização dos alunos, estejam habilitados a realizar intervenções no processo de ensino e aprendizagem. Diversas ferramentas de Mineração de Dados Educacionais foram estudadas, bem como uma pesquisa será conduzida sobre a Interação Humano-Computador, com a finalidade de dotar o sistema de interfaces intuitivas e trazer para os professores caracterizações sobre seus alunos.*

1. Introdução

A evolução das tecnologias de informação e comunicação proporciona um aumento expressivo na quantidade de dados gerada e movimentada todos os dias, nos mais diversos campos da ciência e tecnologia (computação, engenharias, medicina etc). Também no processo de ensino e aprendizagem são muitos os dados gerados e armazenados em bancos de dados.

Atualmente, muitos professores utilizam ferramentas computacionais de apoio pedagógico (e.g. ambientes virtuais de aprendizagem - AVA) para ministrar suas aulas, os quais os auxiliam no controle de atividades extraclasse solicitadas aos alunos (disponibilização de enunciados e notas, recebimento de resoluções etc.). Os fóruns, por exemplo, possibilitam a discussão de assuntos tratados em sala de aula. Além disso, é possível efetuar a submissão de conteúdos e/ou exercícios, a disponibilização de lembretes e avisos para os alunos, entre outras atividades.

No entanto, armazenar e analisar os dados provindos destas ferramentas não são tarefas triviais. Apesar disso, em grande parte dos sistemas web, como o Moodle e o Tidia, existe algum tipo de ferramenta para armazenamento de informações das interações dos usuários com o sistema, ou seja, o histórico de utilização. O que ainda precisa ser feito, de forma adequada e planejada, é a análise desses dados em busca de padrões que possam gerar conhecimento sobre estes usuários ou, então, a retratação da carência de informações destes dados armazenados. Para isso, é necessária a utilização de alguma ferramenta que auxilie na extração desses padrões, já que a quantidade de dados é grande e, ao olhar humano, não possui nenhum significado.

Existem diversas ferramentas de Mineração de Dados (MD) (Weka, 2010; Rminer, 2010) e Mineração de Dados Educacionais (EDM) (Tane et al., 2004; Avouris et al., 2005; Mazza e Milani, 2004; Merceron e Yacef, 2005) que podem ser utilizadas com esta finalidade, portanto, escolher a melhor ferramenta diante do conjunto de dados a ser analisado é uma tarefa complexa. Garcia et al. (2011, p. 77-78) retrata isso de forma bem clara em sua pesquisa "*a maioria das ferramentas atuais de EDM são complexas para os educadores utilizarem e suas características vão muito além do escopo do que um educador pode querer fazer*".

É importante que a interface dessas ferramentas de EDM seja mais simples e intuitiva, bem como toda parte de execução e configuração da mesma, de forma que seus usuários sintam-se seguros quando estiverem minerando os dados. Além disso, instrumentos para a colaboração dos resultados devem ser disponibilizados para que todo o conhecimento trabalhado possa ser divulgado auxiliando assim outros pesquisadores (Garcia et al., 2009).

Há que se considerar que os professores devem utilizar seu tempo na preparação das aulas, aperfeiçoamento de avaliações, adaptação de estratégias e técnicas de ensino e aprendizagem, como forma de buscar melhorar a interação que ocorre entre eles e os alunos (Paiva et al., 2012). No entanto, eles precisam corrigir exercícios, provas e realizar outras atividades para descobrir quais as maiores dificuldades dos alunos, o que inviabiliza a aplicação de um *feedback* adequado aos alunos. Para resolver boa parte das dificuldades do processo de ensino e aprendizagem pode-se aplicar um estudo do comportamento dos alunos, com o objetivo de criar um modelo que possa ser utilizado na previsão de comportamentos semelhantes (Baker et al., 2011a).

Para que esses comportamentos sejam identificados, um estudo sobre a utilização de ferramentas de mineração de dados educacionais (EDM) foi realizado e está retratado na seção 5. Por meio desse estudo, a ferramenta de EDM *DATA MINING of COde REpositore* (DAMICORE - Sanches, Cardoso e Delbem, 2011), foi selecionada como sendo a que melhor se encaixa no perfil da pesquisa. A DAMICORE é uma ferramenta de MD desenvolvida e utilizada na Universidade de São Paulo (USP), no

Instituto de Ciências Matemáticas e de Computação (ICMC), já foi utilizada em diversos contextos e, agora, será inserida no meio educacional.

O presente artigo está organizando da seguinte forma: na seção 2 será abordado o desafio dessa pesquisa, o desenvolvimento de uma ferramenta eficiente para mineração de dados educacionais. Na seção 3 serão apresentados alguns artigos relacionados à pesquisa. Na seção 4 será fundamentada a técnica e as ferramentas de mineração de dados educacionais. Na seção 5 será apresentada a proposta de solução ao desafio proposto. Por fim, na seção 6 estão as considerações finais sobre a pesquisa realizada.

2. O desafio da mineração de dados educacionais

Atualmente, apesar das mudanças que podem ser percebidas na área da educação, os professores ainda encontram dificuldades para realizar adequadamente suas tarefas.. Isto porque, além dos desafios da educação, existem diversos fatores que influenciam o desenvolvimento de suas atividades, por exemplo: precariedade de ferramentas e material, instalações inapropriadas, sobrecarga de atividades, falta de material didático atualizado e/ou de qualidade, pouca motivação dos alunos (decorrência da descrença da população nos benefícios da educação), baixos salários, falta de investimento em cursos e especializações, e o grande consumo de tempo com atividades burocráticas (Paiva et al., 2012).

Além disso, entender como acompanhar efetivamente os estudantes, verificar se eles estão interagindo entre si, identificar estudantes com problemas de desempenho e desmotivados ou, então, prestes a abandonar o curso, pode trazer melhorias nos métodos de ensino e aprendizagem utilizados pelos professores (Gottardo et al., 2012). Porém, como os professores precisam investir muito do seu tempo em tarefas de correção de atividades (provas, exercícios, questionários etc.), o que ainda é potencializado quando se envolve grandes quantidades de alunos, sobra para eles pouco tempo para realizar tarefas como a atualização de conteúdos, aperfeiçoamento das avaliações, adaptação de estratégias e táticas de ensino, dentre outras atividades.

Uma das formas utilizadas para conhecer os estudantes e seu desempenho é a análise dos resultados obtidos nas atividades realizadas, o que pode ocupar muito tempo do professor, ou permitir que este só consiga identificar os alunos com deficiências graves tarde demais. Neste contexto, ferramentas que possam analisar essas atividades e fornecer ao professor, de forma simples e rápida, a caracterização dos estudantes, podem trazer muitos benefícios ao processo de ensino e aprendizagem.

Atualmente, o número de professores que dispõem de ferramentas educacionais para auxiliá-los e evitar que o excesso de tarefas se torne um fardo cumulativo, é bastante reduzido. Paiva et al. (2012, p. 2-3) afirmam que:

Mais incomum, ainda, é encontrar ferramentas tecnológicas atualizadas e apropriadas às necessidades dos professores, e que ofereçam considerável e confiável ajuda em suas responsabilidades profissionais. Desta forma, é necessário que soluções para a educação, e para os professores, sejam capazes de lidar com grande quantidade e diversidade de dados, ofereçam soluções ágeis, antecipem a descoberta de problemas, gerem recomendações e, ainda, sejam de fácil uso para professores não especialistas em Tecnologia da Informação e Comunicação (Paiva et al., 2012, p. 2-3).

Existe uma grande variedade de ferramentas e *frameworks* para mineração de dados gerais. No entanto, a maioria delas é complexa demais para que os professores as utilizem, ou seja, não são intuitivas, de simples utilização e, ainda, necessitam de muito tempo para que se obtenha algum resultado efetivo. Desta forma, é importante que essas ferramentas de apoio ao professor possuam uma interface mais intuitiva, com simples utilização e configuração (Garcia et al., 2011).

Com a finalidade de contribuir na resolução destes problemas citados, na área da educação, esse artigo busca apresentar uma ferramenta que consiga caracterizar os estudantes, de acordo com alguns de seus comportamentos mais comuns, como: *gaming the system* (Baker et al., 2010a; Baker et al., 2011a), *off-task behavior* (Baker et al., 2010a; Baker et al., 2011a) e “desistência do curso” (Kampff, 2009). Para isso, foi escolhida uma ferramenta de fácil utilização, que possui uma interface intuitiva e que exibe resultados satisfatórios em um curto período de tempo, oferecendo resultados importantes aos professores, sem que os mesmos necessitem avaliar diversas provas e atividades dos estudantes.

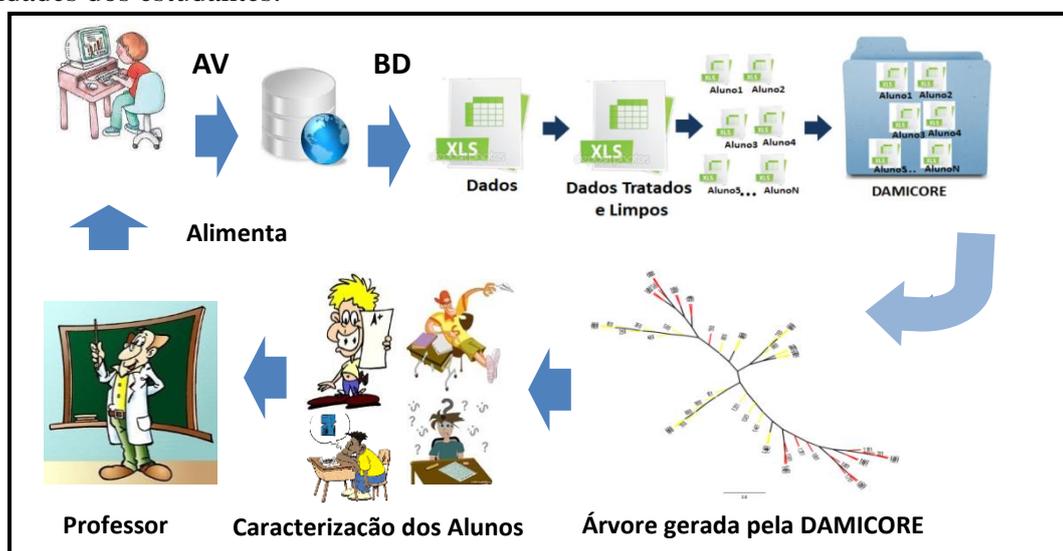


Figura 1. Ciclo da mineração de dados com a DAMICORE

Desta forma, o desafio tratado nesse artigo é viabilizar o uso de ferramentas e técnicas de EDM simples e intuitivas para professores, de forma que, a partir da caracterização dos alunos por meio dos comportamentos comuns apontados pela literatura, os mesmos estejam habilitados a realizar intervenções, em tempo hábil, no processo de ensino e aprendizagem. Isso pode ser observado na Figura 1, onde, por meio da utilização de um AVA, os alunos tem suas interações com o sistema armazenadas e, após o pré-processamento desses dados, eles são minerados pela DAMICORE, que gera uma classificação dos alunos. A partir disso um professor pode analisar essa classificação, gerando uma caracterização dos alunos, intervindo, assim, no processo de ensino e realimentando o AVA, de acordo com as dificuldades apresentadas pela turma.

3. Trabalhos relacionados

Para fundamentar esta pesquisa, alguns experimentos foram estudados, buscando entender o estado da arte na área de EDM (suas vantagens e limitações) e identificar como outros pesquisadores têm utilizado a EDM para lidar com problemas

educacionais. Também é importante entender as metodologias e técnicas utilizadas juntamente com os resultados obtidos.

De acordo com Romero e Ventura (2010), apesar da similaridade existente entre a aplicação de técnicas de MD em ambientes tradicionais (medicina, genética, comércio eletrônico etc.) e em ambientes educacionais, algumas características particulares relacionadas a estes últimos exigem que se façam adaptações nas técnicas utilizadas para tratar um problema específico. Segundo Baker, Isotani e Carvalho (2011), as técnicas de EDM têm sido utilizadas com a finalidade de oferecer: suporte e *feedback* aos professores, recomendações a estudantes, identificação de grupos de estudantes com características comuns, risco de evasão ou previsão de desempenho.

Gottardo et al. (2012) propõem um conjunto de atributos que possam ser extraídos de um AVA para definir e estruturar um modelo de processo de previsão de desempenho. Os autores agruparam estes atributos propostos em três dimensões: i) perfil geral de uso do AVA, cujo objetivo é a identificação dos dados relacionados ao planejamento, organização e gestão do tempo do estudante para a realização do curso; ii) interação estudante-estudante, com a finalidade de verificação da interação dos estudantes entre si, por meio de ferramentas como fóruns, *chats* e mensagens trocadas, além da identificação da existência de colaboração e cooperação; e iii) interação estudante-professor, que objetiva verificar a forma de interação entre professores ou tutores com os estudantes no ambiente (AVA). Para averiguar se o conjunto de atributos propostos estava adequado, os autores realizaram experimentos com uma base de dados Moodle, que continha informações de estudantes em um curso realizado a distância. Os resultados obtidos confirmaram a viabilidade da utilização de um conjunto amplo de atributos para representação de estudantes, com taxas de acurácia superiores às obtidas em outros trabalhos.

Estudar o comportamento dos alunos, dentro da sala de aula ou em ambientes que utilizem sistemas computacionais, tem sido um dos grandes desafios na área de pesquisa em ciência da aprendizagem (Koedinger et al., 2012). Ao conseguir prever o comportamento dos alunos, professores poderão intervir antes que eles sejam reprovados, evitar notas baixas e até mesmo proporcionar maior interesse dos mesmos pela disciplina. Com isso, para conseguir identificar os diferentes tipos de comportamento, a comunidade de pesquisa da área vem desenvolvendo modelos cada vez mais precisos para classificar os alunos (Cen et al., 2007).

Já Garcia et al. (2011), critica a complexidade das atuais ferramentas disponíveis para o descobrimento de conhecimento sobre dados educacionais, e propõe uma ferramenta que utiliza mineração de regras de associação e filtragem colaborativa, afim de realizar recomendações para os instrutores sobre como melhorar cursos *e-learning*. O grupo obteve boas experiências na utilização da ferramenta.

Por meio dos trabalhos pesquisados, pode-se comprovar a carência de ferramentas com interface simples, que apresentem resultados de forma rápida e eficiente aos professores. Para alcançar tais objetivos, será utilizada a EDM, a qual será detalhada a seguir.

4. Mineração de dados educacionais

Baker e Yacef (2009, p. 2) definem a Mineração de Dados Educacionais como sendo “*uma das áreas de pesquisa emergentes e interdisciplinares no contexto educacional*”.

Pode ser definida, ainda, como a aplicação de técnicas de mineração de dados para dados educacionais (Romero et al., 2008). A principal preocupação da EDM é desenvolver métodos que possam explorar as informações extraídas de ambientes educacionais e empregar os métodos para entender melhor os estudantes e os contextos nos quais eles aprendem (Abdullah et al., 2011).

Assim, para que os métodos de EDM possam ser aplicados é necessário que alguma ferramenta de apoio seja utilizada, pois as quantidades de dados são extremamente grandes, o que, sem nenhum tipo de auxílio, seria uma tarefa inviável. Algumas dessas ferramentas de apoio são apresentadas na seção a seguir.

4.1. Ferramentas para mineração de dados educacionais

Com o surgimento do *e-learning*, educação flexível e o aumento do número de alunos nas salas de aula, ferramentas de ensino *on-line* estão se tornando cada vez mais importantes. Elas oferecem um ambiente personalizado, no qual os alunos podem realizar as atividades em seu próprio ritmo, ter acesso às lições por meio de tutoriais, além de explicações e comentários sobre seu desempenho. Conforme Foster (2002), estes benefícios são valiosos para os alunos, de forma que é possível assistir a uma “revolução silenciosa ocorrendo nas salas de aula”.

Para oferecer esse suporte aos alunos e professores, são utilizadas ferramentas e técnicas de EDM. De acordo com o que foi apresentado em seções anteriores, o processo de EDM (Figura 2) converte os dados brutos de sistemas educacionais em informação útil que pode ser utilizada por desenvolvedores de software educacionais, professores, pesquisadores etc. Este processo não difere muito de outras áreas de aplicação da mineração de dados, como negócios, genética, medicina etc. (Romero et al., 2004).

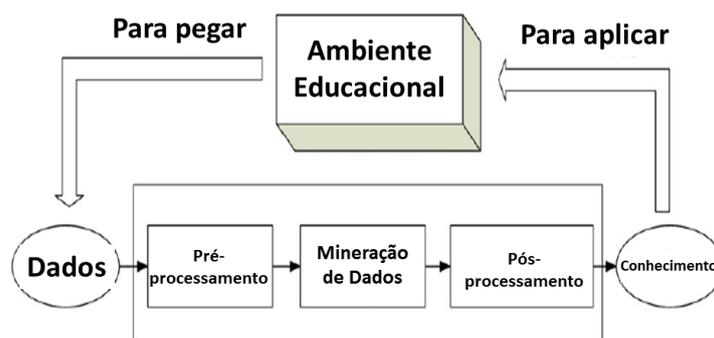


Figura 2 - Processo geral de Mineração de Dados adaptada de Romero et al., (2004)

Atualmente, existe uma grande variedade de ferramentas de mineração de dados gerais e *frameworks* para usuários de MD. Alguns exemplos de ferramentas de mineração de dados são: comerciais: DBMiner (DBMiner, 2010); e de domínio público: Weka (Weka, 2010), Keel (Keel, 2010) e Rapid Miner (Rminer, 2010).

5. Proposta: um caminho para vencer o desafio

Apesar das diversas ferramentas existentes, uma ferramenta de interface simples e de rápida utilização ainda não foi desenvolvida. Um exemplo interessante é a ferramenta Weka que atualmente é uma das mais utilizadas para se executar tarefas de MD. A

Weka possui diversas interfaces, ferramentas de pré-processamento e algoritmos de mineração de dados. No entanto, como afirmado por Garcia et al., (2011) essa robustez da ferramenta vai muito além do escopo do que um professor pode querer fazer. Isso torna a ferramenta complexa para professores não especialistas em Tecnologia da Informação e Comunicação (Paiva et al., 2012, p. 2-3).

Nesse sentido, foi realizada uma pesquisa sobre as ferramentas de MD, e foi possível identificar uma ferramenta de simples utilização, a DATA Mining of CODE REpositore (DAMICORE), desenvolvida com base na combinação de algoritmos já utilizados, com sucesso, em outros campos da ciência: o Normalized Compression Distance (NCD) da Teoria da Informação (Cilibrasi e Vitanyi, 2005), Neighbor Joining (NJ) da filogenética (Felsenstein, 2003) e FastNewman (FN) de Redes Complexas (McLaren et al., 2008) (Carvalho, 2012).

A DAMICORE é uma ferramenta de MD que possui utilização bastante simples. Basta colocar em um repositório (pasta) arquivos do mesmo tipo, cujo conteúdo são os dados que necessitam ser minerados. Isso torna a tarefa de pré-processamento dos dados muito mais simples do que na maioria das outras ferramentas que, por sua vez, utilizam alguns tipos fechados de arquivos, como no caso da Weka, que trabalha com Arff, Csv, LibSvm, Dat e Bsi (Sanches et al., 2011).

Em seguida, basta que cada indivíduo esteja representado num arquivo diferente e, a partir destes arquivos, a DAMICORE aplica o NCD para o cálculo da distância, por meio da compactação dos dados analisados. Depois disso, o NJ constrói uma rede filogenética e, por fim, o FN identifica e forma os clusters (Ag1, Ag2, Ag3 e Ag4) dos dados de entrada, conforme a Figura 3 (Sanches, Cardoso e Delbem, 2011). Desta forma, a DAMICORE gera, como resultado da mineração, uma árvore que facilita a visualização para seus usuários, ao invés de gerar apenas matrizes e valores numéricos, o que dificultaria bastante a identificação dos grupos dentre os indivíduos analisados.

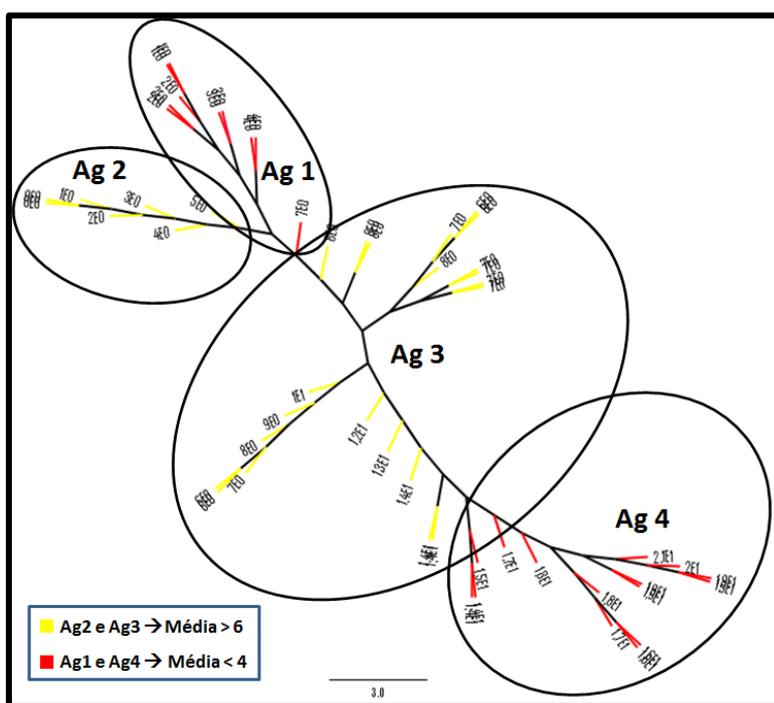


Figura 3 - Arvore gerada pela DAMICORE

Neste trabalho, foram realizados alguns testes da DAMICORE, com a finalidade de descobrir se a ferramenta estaria apta a trabalhar com os dados educacionais. Os resultados foram satisfatórios, como se pode observar na Figura 3, em que a ferramenta conseguiu agrupar os alunos que participaram da pesquisa. Resta, ainda, refinar essa mineração de forma que resultados mais conclusivos possam ser obtidos. A partir da definição da ferramenta a ser utilizada, ainda existe o desafio de desenvolver uma interface, intuitiva e simples, que disponibilize para o professor resultados que possibilitem a caracterização dos alunos, para que, dessa forma, o educador possa intervir nos casos de alunos com comportamentos inadequados evitando que eles obtenham baixo desempenho ou até mesmo abandonem o curso.

É importante, portanto, que se invista em uma pesquisa na área de Interação Humano-Computador (HCI - Human Computer Interaction), buscando gerar interfaces intuitivas para os usuários (professores) do sistema. Também é necessário estudar o nível de conhecimento que os professores (candidatos a utilizar o sistema) possuem e quais as suas aspirações a respeito deste sistema.

Desta forma, este artigo aborda o desafio de viabilizar o uso de ferramentas e técnicas de EDM para professores, de forma que, a partir da caracterização dos alunos, estejam habilitados a realizar intervenções no processo de ensino e aprendizagem de forma mais efetiva e antecipada, gerando assim um ciclo (Figura 1) que traga melhores resultados no processo de ensino e aprendizagem.

6. Considerações finais

A meta dessa pesquisa é suprir uma necessidade básica no sistema de ensino, uma ferramenta que auxilie o professor, caracterizando os alunos e trazendo para o educador os indícios necessários para realizar uma intervenção efetiva e de forma rápida, diminuindo assim a taxa de reprovação da disciplina, ou até mesmo de abandono do curso. Para isso é importante que se invista em técnicas de EDM, bem como nas ferramentas de *e-learning*, como, por exemplo, os AVAs, que intermediam o acesso dos dados entre os alunos e a ferramenta de EDM. Assim, com boas ferramentas de *e-learning* e uma ferramenta eficiente de EDM, pode-se, então, investir em uma interface intuitiva para a ferramenta de EDM. É importante que a interface dessa ferramenta possua simples utilização, como foi discutido nessa pesquisa, trazendo para os professores resultados no tempo devido, tornando assim a intervenção realizada pelo professor mais efetiva, o que ofereça uma maior possibilidade de sucesso no ensino de seus alunos.

Após a definição da ferramenta a abordagem de EDM adequadas é necessário desenvolver interfaces intuitivas para conseguir comprovar a eficiência da EDM junto aos professores.

Agradecimentos

Este trabalho foi parcialmente apoiado pela CAPES e CNPq.

Referências

Abdullaha, Z., Herawanb, T., Ahmadb, N., Derisc, M. M. (2011). Mining significant association rules from educational data using critical relative support approach. *Porcoecdeida - Social and Behavioral Science* 28, p. 97–101.

- Avouris, N. Komis, V. Fiotakis, G. Margaritis, M. Voyiatzaki, E. (2005). Why logging of fingertip actions is not enough for analysis of learning activities. Workshop on usage analysis in learning systems, p. 1-8.
- Baker, R. S. J. D. e Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining*, 1 (1), p. 3-17.
- Baker, R. S. J. d., Carvalho, A. M. J. B., Raspat, J., Alevan, V., Corbett, A. T., Koedinger, K. R., Cocea, M. e HersHKovitz, A. (2010a). Educational Data Mining Methods For Studying Student Behaviors Minute by Minute Across an Entire School Year. *International Conference for the Learning Sciences*.
- Baker, R.S.J.D., Isotani, S., Carvalho, A.M.J.B.D. (2011). Mineração de Dados Educacionais: Oportunidades para o Brasil. *Revista Brasileira de Informática na Educação*, 19(2), p. 2-13.
- Baker, R. S. J. d., Goldstein, A. B. e Heffernan, N. T. (2011a). Detecting Learning Moment-by-Moment. *International Journal of Artificial Intelligence in Education*, 21 (1-2), p. 5-25.
- Carvalho, H. L. (2012). Método de análise para a coordenação dos processos de produção sob a ótica de redes de inovação colaborativas apoiado por agente inteligente evolutivo. Tese de doutorado apresentada para a Escola de Engenharia de São Carlos, São Carlos, 2012.
- Cen, H., Koedinger, K.R., Junker, B. (2007). Is Over Practice Necessary? Improving Learning Efficiency with the Cognitive Tutor. In: *Proceedings of the 13th International Conference on Artificial Intelligence and Education*.
- Cilibrasi, A. R. e Vitanyi A. P. (2005). Clustering by compression. *IEEE Trans. Information Theory*, 51 (4), p. 1523-1545.
- DBMiner. (2010). Disponível em: <<http://www.dbminer.com>>. Acesso em 21.Jan.2013.
- eSchool. (2005). eLearning giants to merge. Disponível em: <<http://www.eschoolnews.com/news/top-news/index.cfm?i=36643&CFID=22179215&CFTOKEN=81170764>>. Acesso em 18.Fev.2013.
- Felsenstein, J. (2003). Inferring phylogenies.
- Forster, A. (2002). Online teaching and learning, Discussion paper, SYNERGY, I. 18.
- García, E., Romero, C., Ventura, S., Castro, C. (2011). A collaborative educational association rule mining tool *Internet and Higher Education* 14, p. 77–88.
- Gottardo, E., Kaestner, C., Noronha, R. V. (2012) Avaliação de desempenho de estudantes em cursos de educação a distância utilizando mineração de dados. I Workshop de Desafios da Computação Aplicada à Educação, Curitiba, PR. Disponível em: <<http://www.lbd.dcc.ufmg.br/colecoes/desafie!/2012/003.pdf>>. Acesso em 05.Abr.2013.

- Kampff, A. J. C. (2009). Mineração de dados educacionais para geração de alertas em ambientes virtuais de aprendizagem como apoio a prática docente. Tese de Doutorado, 2009.
- Keel. (2010). Disponível em: <<http://www.keel.es/>>. Acesso em 21.Jan.2013.
- McLaren, B. M. , Adams, D., Durkin, K., Gogvadze, G., Mayer, R. E., Rittle-Johnson, B., Sosnovsky, S., Isotani, S. e Velsen, M. (2012). To Err Is Human, to Explain and Correct Is Divine: A Study of Interactive Erroneous Examples with Middle School Math Students, LNCS 7563, p. 222–235.
- Mazza, R., Milani, C. (2004). GISMO: A graphical interactive student monitoring tool for course management systems. International Conference on Technology Enhanced Learning. (p. 1-8).
- Merceron, A., e Yacef, K. (2005). Educational data mining: A case study. International Conference on Artificial Intelligence in Education, (p. 1-8).
- Muehlenbrock, M. (2005). Automatic action analysis in an interactive learning environment. Proceedings of the workshop on Usage Analysis in Learning Systems at AIED-2005.
- Paiva, R., Bittencourt, I.I., Pacheco, H., Silva, A; P. da, Jacques, P., Isotani, S. (2012). Mineração de dados e a gestão inteligente da aprendizagem: desafios e direcionamentos. I Workshop de Desafios da Computação Aplicada à Educação, Curitiba, PR. Disponível em: < <http://www.lbd.dcc.ufmg.br/colecoes/desafie!/2012/0024.pdf>>. Acesso em 02.Abr.2013.
- RMiner. (2010) Disponível em: <<http://www.rapidminer.com>>. Acesso em 21.Jan.2013.
- Romero, C., Ventura, S., Bra, P. D. (2004). Knowledge discovery with genetic programming for providing feedback to courseware author. User Modeling and User-Adapted Interaction, 14(5), p. 425–464).
- Romero, C. Ventura, S., García, E. (2008). Data Mining in Course Management Systems: Moodle Case Study and Tutorial. Computers & Education, n. 51, p. 368-384.
- Romero, C. Ventura, S. (2010). Educational Data Mining: A Review of the State of Art. IEEE Transactions on Systems, Man, and Cybernetics – Part C: Application and Reviews, 40 (6), p. 601-618.
- Sanches, A. Cardoso, J. M. P. e Delbem, A. C. B. (2011). Identifying Merge-Beneficial Software Kernels for Hardware Implementation. International Conference on Reconfigurable Computing and FPGAs.
- Tane, J., Schmitz, C., Stumme, G. (2004). Semantic resourcemanagement for theweb: An *e-learning* application. Proceedings of the Conference, p. 1-10.
- Weka. (2010). Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>. Acesso em 21.Jan.2013.