A literatura brasileira sobre mineração de dados educacionais

# Rodrigo Lins Rodrigues<sup>2,3</sup>, Jorge Luis Cavalcanti Ramos<sup>1,3</sup>, João Carlos Sedraz Silva<sup>1,3</sup>, Alex Sandro Gomes<sup>3</sup>

<sup>1</sup>Universidade Federal do Vale do São Francisco-UNIVASF
<sup>2</sup>Departamento de Educação – Universidade Federal Rural de Pernambuco - UFRPE
<sup>3</sup>Centro de Informática – Universidade Federal de Pernambuco - UFPE

rlr@ded.ufrpe.br, jorge.cavalcanti@univasf.edu.br, joão.sedraz@univasf.edu.br, asg@cin.ufpe.br

Abstract. This paper presents an analysis of the research main works in Educational Data Mining (EDM) published in Brazil since 2006. The research was focused on the events and major scientific journals in the area of computer and education, in addition to theses and dissertations related to the topic. The analysis was performed under various classifications of articles, giving interesting dimensions of research in the country.

**Resumo.** O presente trabalho apresenta uma análise dos principais trabalhos de pesquisa na área de Mineração de Dados Educacionais publicados no Brasil desde o ano de 2006. A pesquisa da literatura analisou artigos de eventos científicos e dos principais periódicos na área de informática e educação conforme critérios da CAPES, além de teses e dissertações relacionadas com a temática. Os resultados, feitos a partir de classificações dos artigos, dão conta de dimensões interessantes da pesquisa na área realizada no país.

# 1. Introdução

Já não se discute mais a importância do uso das Tecnologias de Informação e Comunicação (TIC) na educação; seja ela em qualquer modalidade, nível ou dispositivo. As novas áreas de pesquisa sobre a associação das TIC com a educação, o que se busca é encontrar e apresentar novos métodos, técnicas e procedimentos que: aperfeiçoe essa relação, facilite a aprendizagem, permita um melhor planejamento e gestão de cursos, e ofereça alternativas para superar desafios e dificuldades em estruturas educacionais diversas.

Com a evolução da aplicação das TIC na Educação, os ambientes computacionais de ensino e os conteúdos digitais passaram a ser vistos como tecnologias essenciais à condução dos processos educacionais; sejam eles presenciais ou a distância (Galafassi *et al.*, 2013).

O uso crescente de ambientes virtuais de aprendizagem (AVA) e de sistemas de informações educacionais (SIE) tem gerado grandes volumes de dados, provenientes de interações e de registros contínuos de informações dos professores, alunos, gestores e demais atores de sistemas educacionais. Esse volume de dados pode fornecer muito mais que informações gerenciais ou relatórios de desempenho, por exemplo. O uso de técnicas de mineração de dados educacionais (EDM) pode ser útil para se descobrir novos conhecimentos relevantes nesses dados.

O processo de EDM converte os dados brutos de sistemas educacionais em informação útil que pode ser usado por desenvolvedores de software educacionais, professores, pesquisadores educacionais etc. Este processo não difere muito de outras

áreas de aplicação de mineração de dados porque ele baseia-se nos mesmos passos do processo de mineração de dados em geral, conforme Figura 1 (Garcia *et al.*, 2011). Ela tem permitido melhorar os processos de ensino-aprendizagem e os mecanismos de gestão acadêmica e pedagógica das instituições de ensino.

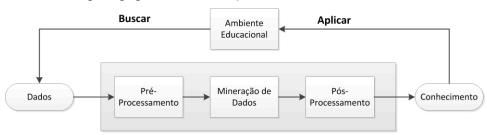


Figura 1 – Etapas da Mineração de Dados Educacionais. (Adaptado de Garcia et *al.*, 2011).

Este trabalho apresenta uma revisão da literatura brasileira sobre em EDM publicadas em diversos eventos, periódicos, teses e dissertações. Destacou-se as principais obras, com ênfase nas: técnicas, modelos, objetivos educacionais, ferramentas utilizadas e um mapeamento das origens das publicações. A revisão teve o propósito de identificar e relatar as principais características de pesquisas na área no país.

Este trabalho está organizado em seis seções. Na segunda seção é feita uma breve descrição da evolução e consolidação da EDM como área de pesquisa. Na seção três define-se o processo de busca usado neste trabalho. Na quarta seção são descritos os resultados do estudo sob várias categorias. Por fim, nas seções cinco e seis são apresentadas as discussões, considerações finais e sugestões de trabalhos futuros.

# 2. Evolução e consolidação da EDM como área de pesquisa

A EDM tem emergido como uma relevante área de pesquisa nos últimos anos em diversas áreas (por exemplo: ciência da computação, educação, psicometria, estatística, sistemas tutores inteligentes, *e-learning* entre outros) ao analisar grandes conjuntos de dados, a fim de resolver as questões de investigação educacional (Baker e Yacef, 2009).

A partir do seu relacionamento com as diversas áreas de conhecimento, Romero e Ventura (2013) afirmaram que a EDM é a combinação de três principais áreas de conhecimento (Figura 2): Computação, Educação e Estatística. A interseção dessas áreas fornece as três subáreas – *E-learning, Data Mining e Machine Learning e a Learning Analytcs* – que estão mais relacionadas com a EDM.

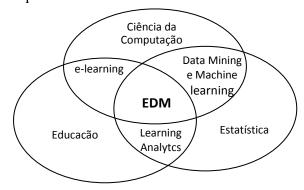


Figura 2 – Principais áreas relacionadas com a EDM. (Adaptado de Romero e Ventura, 2013)

Até o ano de 2005, as pesquisas na área de EDM eram relatadas somente em conferências das grandes áreas ou das subáreas descritas na Figura 1. Naquele ano

\_\_\_\_\_\_

aconteceu o *I Workshop on Educational Data Mining*, em Pittsburgh (EUA), como evento satélite da *20th National Conference on Artificial Intelligence* (AAAI-05). A partir de então, diversos Workshops de EDM aconteceram dentro de eventos maiores, até o ano de 2007 (Romero e Ventura, 2013).

Em 2008, aconteceu o primeiro evento exclusivo para EDM: a *I International Conference on Educational Data Mining*, em Montreal, Canadá. Desde então essa conferência acontece anualmente. A do ano de 2014 aconteceu em Londres, Inglaterra, e a de 2015 acontecerá em Madrid, Espanha.

Da mesma forma que ocorreu com os eventos, nos quais os trabalhos foram inicialmente publicados em eventos de áreas afins, os artigos de pesquisas sobre EDM são ainda muito publicados em periódicos com reconhecimentos internacionais reconhecidos, como: *Internet and Higher Education, Computers and Education, Expert Systems with Applications, Journal of the Learning Sciences*, entre outros. Isso ocorre mesmo depois da criação, em 2009, do *Journal of Educational Data Mining*<sup>1</sup>, mantido pela *EDM Society*.

Para estimular e auxiliar o trabalho de novos pesquisadores na área e para se ter uma percepção mais abrangente do nível de pesquisa mundial em EDM, três *surveys* fizeram um grande levantamento e classificação dos trabalhos internacionais na área, sendo fonte de referências indispensáveis para qualquer pesquisa em EDM: Romero e Ventura (2010), Mohamad e Tasir (2013) e Peña-Ayala (2014).

No Brasil, o primeiro evento com temática relacionada com a análise de dados educacionais foi o *Workshop on Analyzing Student –Tutor Interaction Logs to Improve Educational Outcomes*, que aconteceu em 2004 em Maceió-AL, dentro da programação da 7th International Conference in Intelligent Tutoring Systems (ITS-2004).

O Simpósio Brasileiro de Informática na Educação (SBIE), o Workshop de Desafios da Computação Aplicada à Educação (DESAFIE) e eventos de tecnologias aplicadas à educação são eventos anuais e contínuos que recebem pesquisadores e estudantes brasileiros para relatar suas pesquisas sobre EDM. As publicações em periódicos ocorrem na Revista Brasileira de Informática na Educação (RBIE), na Revista de Novas Tecnologias na Educação (RENOTE), na INFORMÁTICA NA EDUCAÇÃO: teoria & prática e na Colabor@. Estas revistas têm apresentado artigos relevantes com pesquisas com essa temática.

Dentre as publicações brasileiras, merece destaque o artigo de Baker *et al.* (2011) como marco importante para o desenvolvimento de pesquisas em EDM no Brasil. No entanto, já existiam trabalhos anteriores na área, como os de Pimentel e Omar (2006), Kampff *et al.* (2008), Dias *et al.* (2008) e Marinho *et al.* (2009). O artigo de Baker et al. (*up cit.*) listou uma série de técnicas e apontou caminhos interessantes para a EDM no país e é o mais citado nos trabalhos analisados nesse estudo.

#### 3. Processo de busca de obras relacionadas

Para o desenvolvimento desse trabalho, foi considerado como período de publicação os últimos oito anos. A busca foi feita em páginas de conferências e periódicos da área de Informática na Educação. Utilizaram-se termos como: "mineração de dados educacionais", "EDM", "mineração de dados + ambientes virtuais" e "mineração de

<sup>&</sup>lt;sup>1</sup> http://www.educationaldatamining.org/JEDM/

textos". A busca no Portal Periódicos da CAPES apresentou resultados bastante limitados, daí a opção pela ferramenta Google Scholar.

A busca foi toda feita a partir da digitação dos termos nos campos de busca. A análise dos resultados das buscas também foi feita manualmente, não havendo nenhuma outra ferramenta para tratamento e refinamento dos resultados, a não ser as próprias opções de busca avançada nas fontes, quando disponíveis.

## 3.1 Objetivo

O principal objetivo deste trabalho foi catalogar e analisar a maior quantidade possível de publicações cuja temática predominante fosse o uso da EDM como alvo da publicação, obtendo uma real dimensão das pesquisas na área no país.

## 3.2 Fluxo de trabalho da pesquisa

Para a execução deste estudo foi necessário estabelecer etapas com tarefas definidas, caracterizando um fluxo de atividades sequenciais e possibilitando um trabalho contínuo e abrangente. Para tanto, foi adotado o modelo de fluxo de trabalho descrito no trabalho de Peña-Ayala (2014), com as devidas adaptações para esta pesquisa. Este fluxo adaptado é mostrado na Figura 3.

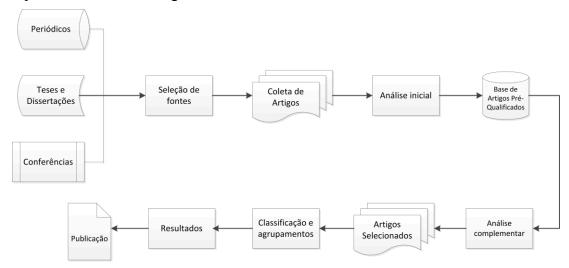


Figura 3: Fluxo de trabalho para analisar, classificar, representar dados das obras EDM publicadas no Brasil (Adaptado de Peña-Ayala, 2014).

O fluxo apresenta inicialmente uma seleção de fontes baseadas em três categorias: periódicos regulares, teses e dissertações e conferências. Nesta última foram incluídos todos os tipos e abrangências, desde que realizadas no Brasil.

Após a coleta inicial, foi feita uma análise de todos os artigos obtidos, para verificar sua aderência ao tema proposto e ao período dos últimos oito anos. Assim, obteve-se uma base de artigos pré-qualificados. Uma nova análise foi feita para detectar artigos indevidamente pré-qualificados, verificar novamente a relevância da EDM no estudo apresentado e para checar a origem ou fonte do artigo. Buscou-se assim eliminar trabalhos com local de publicação indefinido. Artigos não enquadrados nas três categorias e que fossem relevantes para a área foram classificados como 'outros' e incluído na análise posterior.

Com os artigos selecionados, foi feito o detalhamento para classificação e agrupamento dos mesmos, usando uma planilha com os cabeçalhos das colunas

adequadamente definidos para facilitar aplicação dos filtros e análises posteriores, cujos principais resultados são apresentados na próxima seção.

#### 4 Resultados

A busca inicial resultou em 75 trabalhos listados e o posterior refinamento resultou numa seleção final de 68 trabalhos. A lista final dos trabalhos selecionados pode ser obtida no endereço: http://bit.ly/mdeBrasil.

Como primeiro resultado das análises feitas, a Tabela 1 apresenta a distribuição dos artigos pelos tipos das fontes dos trabalhos.

Tipo da Fonte	Quantidade	
Conferências	44	
Periódicos	14	
Dissertações	4	
Teses	2	
Outros	4	
Total	68	

Tabela 1 - Distribuição dos estudos por categoria.

No total de 'outros' encontram-se dois Trabalhos de Conclusão de Curso (TCC) e dois relatórios de pesquisa considerados relevantes para este levantamento. O detalhamento dos eventos e periódicos encontra-se na Tabela 2. Entre as conferências, destaque para o Simpósio Brasileiro de Informática na Educação (SBIE), com 18 publicações. Em relação aos periódicos, o maior número de trabalhos foi identificado na revista Novas Tecnologias na Educação (RENOTE), coordenada pela Universidade Federal do Rio Grande do Sul (UFRGS), com 8 trabalhos publicados.

Tabela 2 - Distribuição dos estudos por fonte.

	Fonte	Quantidade
Conferências	SBIE	18
	LACLO	4
	DESAFIE	4
	TISE	2
	SBSI	2
	ERBASE	2
	Outras	12
Periódicos Rev Info	Novas Tecnologias na Educação (RENOTE)	8
	Revista Brasileira de Informática na Educação (RBIE)	3
	Informática na Educação: Teoria & Prática	2
	Colabor@	1
Outros (teses, dissertações, TCC, relatórios)		10
	Total	68

A frequência anual das publicações é apresentada no gráfico da Figura 4. Observa-se que a partir de 2011 houve um aumento considerável de publicações, concentrando (69,11%) dos artigos analisados.

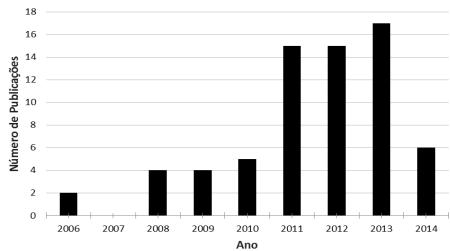


Figura 4 - Gráfico de distribuição dos trabalhos por ano.

O número de publicações em 2014, com artigos publicados até maio desse ano, é ainda inferior à média recente. Isso se deve ao fato de as principais conferências alvo acontecerem no segundo semestre de cada ano. No presente ano ainda não houve edição da RENOTE e apenas uma edição da RBIE, principal periódico na área de Informática e Educação no Brasil, foi publicada.

Em relação às três grandes áreas da EDM, a análise (Tabela 3), mostra uma maior predominância da mineração de dados com (70,59%) dos trabalhos, seguido de mineração de textos (23, 56%) e visualização com (5,85%) dos trabalhos.

Dentre as técnicas utilizadas no desenvolvimento dos trabalhos (Figura 5), a maior predominância (35,13%) se deu na utilização da técnica de classificação, seguida por trabalhos que não deixaram clara qual técnica utilizaram no desenvolvimento do mesmo (22,97%), mencionando apenas que foi utilizada a mineração de dados.

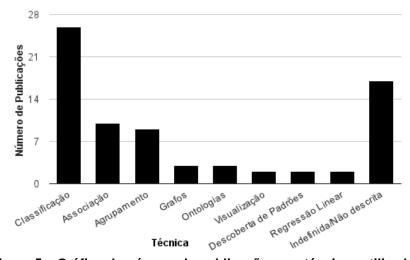


Figura 5 – Gráfico do número de publicações por técnicas utilizadas.

Para classificar os trabalhos quanto às abordagens adotadas (Tabela 3), foram utilizadas três categorias: (i) análise, quando o texto apresentava técnicas de EDM para analisar dados e apresentar resultados; (ii) desenvolvimento, para caracterizar que o texto propunha o desenvolvimento de alguma aplicação para mineração, e, (iii) teórico, quando o texto fazia apenas uma abordagem teórica e conceitual da EDM. Alguns trabalhos continham mais de uma abordagem, por isso o número total é superior ao número de artigos analisados.

Tabela 3 – Distribuição dos trabalhos por área e por abordagem adotada.

Área	Quantidade	Abordagem	Quantidade
Mineração de dados	48	Análise	46
Mineração de textos	16	Desenvolvimento	13
Visualização	4	Teórico	12
Total	68	Total	71

Na análise dos principais objetivos educacionais descritos nos objetivos gerais ou específicos dos trabalhos analisados, verificou-se que (17,64%) dos trabalhos focavam na estimativa ou modelagem do desempenho de estudantes. Em seguida tivemos um percentual de (14,70%) que tinham como objetivo a modelagem de grupos de aprendizagem. A lista dos objetivos identificados está na Tabela 4.

Tabela 4 – Lista de objetivos educacionais verificados.

Objetivos Educacionais	Quantidade
Estimativa ou Modelagem de desempenho de estudante	12
Modelagem de Grupos ou Aprendizagem Colaborativa	10
Mediação ou Recomendação Pedagógica	9
Apoio ao estudante e feedback	8
Detecção ou Previsão de Evasão	7
Avaliação ou Modelagem do Estudante	6
Teórico	5
Objetos de Aprendizagem	2
Apoio ou Gestão Institucional	2
Detecção de plágio	1
Avaliação de sentimento	1
Indefinido/Não descrito	5
Total	68

Levando em consideração as ferramentas computacionais utilizadas para a mineração de dados, verificou-se que o WEKA foi a de maior utilização (35,44%), conforme a Tabela 5. Outra classificação que mereceu destaque corresponde aos trabalhos que não utilizaram ferramentas de análises, já que 27,84% dos estudos desenvolveram análises por meio de do uso de linguagens de programação. Mais uma vez, o número total superior ao número de artigos é devido à utilização de mais de uma ferramenta em alguns trabalhos. A principal fonte de dados foi o *Moodle*, com utilização em (42,65%) dos trabalhos.

Tabela 5 - Ferramentas utilizadas nos trabalhos.

Ferramentas	Quantidade
WEKA	28
RapidMiner	7
SOBEK	5
Mediatec	3
R	2
Linguagens	12
Outras ferramentas	22
Total	79

A distribuição dos trabalhos por IES é apresentada na Tabela 6 abaixo, onde somente as IES com 4 ou mais publicações são identificadas. Importante destacar que para publicações onde mais de um autor pertencia à mesma IES, foi contada apenas uma publicação para a instituição.

Tabela 6 – Distribuição dos trabalhos por IES.

IES	Quantidade	IES	Quantidade
UFRGS	21	IFF	4
UFAL	8	UFPEL	4
UNISINOS	8	UFRJ	4
UTFPR	6	USP	4
IFRS	5	Outras	64
ŗ	<b>Fotal</b>		128

Por fim, foi coletado também no *Google Scholar*, em agosto de 2014, o número de citações de cada obra, para estabelecer um *ranking* das obras mais referenciadas dentre as analisadas. Assim gerou-se a Tabela 7 abaixo com esse *ranking*, levando-se em conta os trabalhos com quatro ou mais citações.

Tabela 7 - Trabalhos mais citados.

Título	Nº Citações
Mineração de Dados Educacionais: Oportunidades para o Brasil (Baker et al., 2011).	
Previsão de Estudantes com Risco de Evasão Utilizando Técnicas de Mineração de Dados (Manhães <i>et al.</i> , 2011).	
Mineração de dados educacionais para a construção de alertas em ambientes virtuais de aprendizagem como apoio à prática docente (Kampff <i>et al.</i> , 2018).	
Mineração de dados educacionais para geração de alertas em ambientes virtuais de aprendizagem como apoio à prática docente (Kampff, 2009).	5
Aplicação de Técnicas de Mineração de Dados no Processo de Aprendizagem na Educação a Distância (Dias <i>et al.</i> , 2008)	4
Descobrindo Conhecimentos em Dados de Avaliação da Aprendizagem com Técnicas de Mineração de Dados (Pimentel e Omar, 2006).	4
Uma ferramenta de apoio à mediação pedagógica para o ambiente Moodle (Severo et al., 2011).	4
Conceitos de Mineração de Dados na Web (Santos, 2009).	4
A Mediação Pedagógica em Ambientes Virtuais de Ensino-Aprendizagem Através de Agentes de Mineração de Dados Educacionais (Severo <i>et al.</i> , 2011).	4
Evasão no ensino superior: um estudo utilizando a mineração de dados como ferramenta de gestão do conhecimento em um banco de dados referente à graduação de engenharia (de Souza, 2018).	4

Dentre os 18 trabalhos que citam Baker *et al.* (2011), 14 foram também analisados neste estudo. Os demais foram publicados em eventos ou periódicos internacionais (03 trabalhos) ou não tinha o foco principal em EDM (01 trabalho).

#### 5 Discussão

O aumento no desenvolvimento e utilização de plataformas educacionais tem proporcionado a possibilidade de armazenamento de grandes bases de dados educacionais. Diversos trabalhos vêm sendo desenvolvidos por pesquisadores brasileiros fortalecendo a área de mineração de dados educacionais no país.

Foi uma boa surpresa verificar o número de publicações brasileiras em EDM e ainda melhor constatar o seu crescimento recente, inclusive com pesquisadores de todas as regiões do Brasil. Isso fornece evidências da consolidação da área também no país, a exemplo do que acontece internacionalmente.

Grande parte destes trabalhos tratou de análises em bases de dados de sistemas do tipo ambientes virtuais de aprendizagem, onde predomina o ambiente *Moodle*. A ferramenta de mineração de dados mais utilizada para estes estudos foi a ferramenta WEKA, pois trata-se de uma ferramenta já consolidada mundialmente.

Embora o crescimento de trabalhos na área seja evidente, poucos trabalhos têm direcionado para o desenvolvimento de ferramentas próprias utilizando mineração de dados nestes ambientes. A maioria dos trabalhos foi realizada através da aplicação direta das técnicas de mineração e posterior processamento das informações utilizando plataforma como o WEKA.

Percebemos que há uma tendência crescente na utilização da EDM para modelagem e previsão de desempenho de estudantes de forma individual ou coletiva, embora outras abordagens como detecção e previsão de evasão também estejam em crescimento.

#### 6 Conclusões

Este trabalho teve como objetivo apresentar a revisão da literatura brasileira sobre EDM nos últimos oito anos. Através de revisão *ad hoc*, pudemos evidenciar diversas caraterísticas de trabalhos desenvolvidos no Brasil e confirmar um aumento significativo de trabalhos nos últimos quatro anos.

Como limitações para o desenvolvimento desta pesquisa, tivemos a possibilidade de artigos não terem sido localizados e relacionados, definição de técnicas e objetivos educacionais as vezes não muito claros, o que dificultou a classificação de alguns artigos.

A criação de um evento específico de EDM dentro do CBIE vai contribuir significativamente para o aumento do número de pesquisas e publicações na área.

Como sugestões de trabalhos futuros temos a divulgação perante a comunidade científica e o compartilhamento da planilha de coleta de dados (*http://bit.ly/mdeBrasil*), onde esperamos uma colaboração para atualização contínua da lista de artigos, que beneficiará diversos pesquisadores que tem interesse em entender a evolução da área e suas características no decorrer dos próximos anos.

### Agradecimentos

Os autores agradecem o apoio da Universidade Federal do Vale do São Francisco (UNIVASF/PRPPGI - Edital 18/2013). Alex Sandro Gomes é bolsista DT Nível 2 pelo CNPq apoiado pelos processos nº 310466/2012-1 e nº 475634/2013-6.

#### Referências

- Baker, R. S. J. D. & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. Journal of Educational Data Mining, 1(1), 3-17.
- Baker, R. S. J. D., Isotani, S., & Carvalho, A. (2011). Mineração de Dados Educacionais: Oportunidades para o Brasil. Revista Brasileira de Informática na Educação, 19(02), 03.
- de Souza, S. L. (2008). Evasão no ensino superior: um estudo utilizando a mineração de dados como ferramenta de gestão do conhecimento em um banco de dados referente à graduação de engenharia (Tese de Doutorado, UFRJ).
- Dias, M. M., Lino, A. D. P., Favero, E. L., & Ramos, E. M. L. S. (2008). Aplicação de Técnicas de Mineração de dados no Processo de Aprendizagem na Educação a

- Distância. Anais do Simpósio Brasileiro de Informática na Educação (Vol. 1, No. 1, pp. 105-114).
- Galafassi, F. P., Gluz, J. C., & Galafassi, C. (2013). Análise Crítica das Pesquisas Recentes sobre as Tecnologias de Objetos de Aprendizagem e Ambientes Virtuais de Aprendizagem. Revista Brasileira de Informática na Educação, 21(03), 100.
- García, E., Romero, C., Ventura, S., & de Castro, C. (2011). A collaborative educational association rule mining tool. The Internet and Higher Education, 14(2), 77-88.
- Kampff, A. J. C. (2009). Mineração de dados educacionais para geração de alertas em ambientes virtuais de aprendizagem como apoio à prática docente. (Tese de Doutorado UFRGS).
- Kampff, A. J. C., Reategui, E. B., & de Lima, J. V. (2008). Mineração de dados educacionais para a construção de alertas em ambientes virtuais de aprendizagem como apoio à prática docente. RENOTE, 6(1).
- Manhães, L. M. B., da Cruz, S. M. S., Macário Costa, R. J., Zavaleta, J., & Zimbrão, G. (2011). Previsao de estudantes com risco de evasão utilizando técnicas de mineração de dados. Anais do Simpósio Brasileiro de Informática na Educação (Vol. 1, No. 1).
- Marinho, T., Dermeval, D., Ferreira, R., Braz, L. M., Bittencourt, I. I., Costa, E. B., & L Luna, H. P. (2009). Um Framework para Mineração de Dados Educacionais Basedo em Serviçõs Semânticos. Anais do Simpósio Brasileiro de Informática na Educação (Vol. 1, No. 1).
- Mohamad, S. K., & Tasir, Z. (2013). Educational Data Mining: A Review. Procedia-Social and Behavioral Sciences, 97, 320-324.
- Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. Expert systems with applications, 41(4), 1432-1462.
- Pimentel, E. P., & Omar, N. (2006). Descobrindo Conhecimentos em Dados de Avaliação da Aprendizagem com Técnicas de Mineração de Dados. Anais do Workshop de Informática na Escola (Vol. 1, No. 1).
- Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 40(6), 601-618.
- Romero, C., & Ventura, S. (2013). Data mining in education. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 3(1), 12-27.
- Santos, R. (2009) "Conceitos de Mineração de Dados na Web." Anais do XV Simpósio Brasileiro de Sistemas Multimídia e Web e VI Simpósio Brasileiro de Sistemas Colaborativos. 81-124.
- Severo, C. E. P., Passerino, L. M., & Lima, J. V. (2011). Uma ferramenta de apoio à mediação pedagógica para o ambiente Moodle. Primer MoodleMoot Uruguay, 1-5.
- Severo, C. E. P., Passerino, L. M., Gluz, J. C., & Raminelli, Á. (2011). Mediação pedagógica em ambientes virtuais de ensino-aprendizagem através de agentes de mineração de dados educacionais. Informática na educação: teoria & prática, 14(2).