

Mineração de Dados na Identificação de Grupos de Estudantes com Dificuldades de Aprendizagem no Ensino de Programação

Rozelma Soares de França

Universidade de Pernambuco – Garanhuns, PE – Brasil

Centro de Informática – Universidade Federal de Pernambuco – Recife, PE – Brasil

E-mail: rozelma.soares@gmail.com

Haroldo José Costa do Amaral Universidade de

Pernambuco – Garanhuns, PE – Brasil E-mail:

haroldo.amaral@gmail.com

Resumo: Considerar que as salas de aula são compostas por estudantes heterogêneos implica dizer que elas são formadas por aprendizes com diferentes estilos de aprendizagem, que possuem necessidades e preferências de aprendizado particulares. No entanto, a personalização de aulas, no ensino presencial, é algo impraticável. O atendimento personalizado a grupos homogêneos de estudantes é uma possibilidade a ser considerada. Nesse sentido, este artigo tem por objetivo apresentar o uso de técnicas de mineração de dados para a formação de grupos similares de estudantes, com dificuldades de aprendizagem no ensino de Programação. Com isso, espera-se ser possível formular estratégias pedagógicas adequadas a grupos de aprendizes no intuito de melhorar o seu desempenho.

Palavras-chave: Mineração de dados; Taxonomia de Bloom; estratégias de aprendizagem; ensino de programação.

Data Mining in Identifying Groups of Students with Learning Difficulties in Teaching Programming

Abstract: Consider that classrooms are composed of heterogeneous students implies that they are formed by learners with different learning styles, which have needs and preferences of individual learning. However, customizing lessons in classroom teaching is some impractical. The personalized attention to homogeneous groups of students is a possibility to be considered. Accordingly, this article aims to present the use of data mining techniques for the formation of similar groups of students with learning difficulties in teaching programming. Thus, it is expected to be possible to formulate teaching strategies appropriate to groups of learners in order to improve their performance.

Keywords: Data mining; Bloom's Taxonomy; learning strategies; teaching programming.

1. Introdução

Nas salas de aula das universidades brasileiras, temos uma grande quantidade de estudantes com conhecimentos e habilidades heterogêneos. Ainda assim, todos os aprendizes recebem o mesmo tratamento, independente do seu *background* cognitivo

sobre um conteúdo específico. Tal situação condena muitos estudantes a uma falta de aprendizagem adequada, gerando sucessivas reprovações (PIMENTEL et al., 2003).

Perrenoud (2000a) afirma que é preciso diferenciar, rompendo com a pedagogia frontal que impõe a mesma lição, os mesmos exercícios para todos. O autor ainda diz que “praticar uma pedagogia diferenciada é fazer com que, quando necessário, cada aluno seja recolocado ou reorientado para uma atividade fecunda para ele” (PERRENOUD, 2000b, p. 73). Para alcançar isso, deve-se instaurar um diálogo sobre o saber e a aprendizagem.

No ensino presencial, no entanto, a sua personalização é uma atividade impraticável para o professor, ainda que com a diminuição do número de estudantes por turma. Uma proposta a esse dilema está no uso de técnicas computacionais que, quando aplicadas a um conjunto de dados, pode gerar informações úteis para tomada de decisões pedagógicas que favoreçam a aprendizagem dos estudantes. Nesse sentido, este trabalho propõe uma maneira para personalizar o ensino presencial.

A proposta é que o atendimento seja feito a grupos homogêneos de aprendizes que possuam lacunas de aprendizagem similares. Considerando que a identificação desses grupos é uma atividade relativamente árdua para o ser humano, uma vez que pode envolver uma grande massa de dados, propõe-se o uso de técnicas de mineração de dados, clusterização (agrupamento), especificamente. Nessa direção, trabalhos como o de Pimentel et al. (2000) e Correia et al. (2001) foram realizados. No entanto, destaca-se neste trabalho o uso de um conjunto de atributos elaborado por França (2013), baseado na Taxonomia de Bloom (BLOOM et al., 1956; ANDERSON & KRATHWOHL, 2001), que, juntamente com técnicas computacionais, possibilita realizar o mapeamento do conhecimento de estudantes em determinada disciplina.

Assim, este trabalho apresenta os resultados da aplicação de uma técnica de clusterização sobre dados coletados a partir de avaliações de aprendizagem, na tentativa de agrupar estudantes com dificuldades de aprendizagem similares em uma disciplina de Programação. A partir do uso do conjunto de atributos proposto por França (2013), foi possível perceber que conteúdos e em que nível cognitivo (Lembrar, Entender, Aplicar, Analisar, Avaliar, Criar) eles foram assimilados pelos aprendizes, permitindo, assim, que estratégias pedagógicas sejam adequadamente planejadas para grupos de estudantes que apresentem as dificuldades identificadas.

O restante do artigo está organizado como segue: a Seção 2 apresenta conceitos sobre descoberta de conhecimento em bases de dados e descreve a técnica de agrupamento em mineração de dados utilizada neste trabalho. A Seção 3 descreve alguns trabalhos que aplicaram técnicas de mineração de dados no âmbito da educação. Nas Seções 4 e 5, descreve-se, respectivamente, o estudo de caso realizado e os resultados obtidos. Por fim, na Seção 6, são feitas algumas considerações acerca deste trabalho.

2. Descoberta de Conhecimento e Mineração de Dados

Com o avanço das tecnologias computacionais que permitem o armazenamento e processamento de um grande volume de dados, novas tecnologias têm sido desenvolvidas para auxiliar a extração de informações dessas bases de dados, destacando-se o *Knowledge Discovery in Database* (KDD) e a *Data Mining* (DM).

Segundo Fayyad (1996), KDD consiste num processo de várias etapas, não trivial, interativo e iterativo, para identificação de padrões compreensíveis, válidos,

novos e potencialmente úteis a partir de grandes conjuntos de dados. Ele é um processo abrangente de descoberta de conhecimento útil, a partir de dados; por outro lado, mineração de dados restringe-se à aplicação de algoritmos de extração de padrões (FAYYAD et al., 2002). Este trabalho debruça-se sobre a aplicação de KDD na área de Educação, na tentativa de descobrir padrões em dados de avaliações da aprendizagem, focando-se na etapa de Mineração de Dados.

O processo de descoberta do conhecimento envolve várias fases e etapas. Neste trabalho, o processo de KDD é classificado em três grandes fases, tendo em vista as abordagens descritas em Rezende et al., (2003) e Pimentel et al. (2006): Preparação, Extração de Padrões e Pós-Processamento. A descrição de cada fase é apresentada na Subseção 2.1, tendo como referencial os autores supramencionados.

2.1. Fases do KDD

Na fase de **Preparação**, ocorre a identificação e o entendimento do problema, considerando aspectos como objetivos e as fontes de dados da qual se pretende extrair o conhecimento. O passo seguinte consiste na seleção dos dados a partir das fontes, de acordo com os objetivos do processo, e tratamento dos dados para poderem ser submetidos aos métodos e ferramentas na etapa de extração de padrões.

Na fase de **Extração de Padrões**, escolhe-se inicialmente um ou mais algoritmos para a extração de conhecimentos. Essa fase pode ser executada diversas vezes para o ajuste de parâmetros de maneira a se obter resultados adequados aos objetivos estabelecidos.

Na fase de **Pós-Processamento**, o conhecimento extraído é avaliado quanto a sua qualidade e/ou utilidade para que, em caso positivo, possa ser utilizado para apoiar algum processo de tomada de decisão, seja por um especialista humano ou por um sistema especialista.

2.2. Tarefas e Técnicas de Mineração de Dados

No processo de KDD, a mineração de dados é a etapa central que executa a descoberta de conhecimento propriamente dita, pois são os algoritmos de mineração de dados que produzem, de maneira semi-automática, o conhecimento a partir de dados existentes.

Segundo Pimentel et al. (2006), as tarefas de mineração de dados são as classes de problemas determinadas de acordo com o tipo de conhecimento a ser minerado e também dos objetivos almejavéis para a solução. A escolha da técnica de mineração e do algoritmo a ser utilizado, contudo, depende da tarefa que se quer executar.

Considerando que o propósito deste trabalho é formar grupos homogêneos de aprendizes, a tarefa de Agrupamento foi escolhida, visando agrupar um conjunto de dados utilizando algum critério de similaridade. Diferentemente da tarefa de Classificação, esses dados não devem ter sido previamente classificados (LOPES, 2003).

A técnica de particionamento implementa este tipo de tarefa, sendo o algoritmo *K-means* o mais conhecido (WITTEN et al., 2011). Esse algoritmo, basicamente, divide o conjunto de dados em grupos (*clusters*), sendo necessário definir inicialmente o número de *clusters* que serão criados. Esse número é chamado de K, razão para o nome *K-means*.

3. Mineração de Dados Educacionais

Técnicas de mineração de dados podem ser aplicadas a uma variedade de contextos de tomada de decisão como *marketing*, finanças, manufatura, saúde. Analogamente, é possível minerar dados de alunos para verificar a relação entre uma abordagem pedagógica e o seu aprendizado e, através dessa informação, o professor poder compreender se sua abordagem está realmente ajudando o estudante e desenvolver métodos de ensino mais eficazes (BAKER et al., 2011).

Em sistemas educacionais, o conhecimento descoberto pode ser útil tanto para os educadores e profissionais responsáveis pela concepção, planejamento, construção e manutenção dos sistemas, quanto para os próprios estudantes que interagem com a aplicação. Romero & Ventura (2007) afirmam que a descoberta do conhecimento pode ser orientada a diferentes atores de acordo com os interesses particulares de cada um:

- **Orientada para estudantes:** técnicas de mineração de dados são utilizadas para recomendar atividades, recursos e tarefas que favoreçam e melhorem a aprendizagem do aluno. Isso pode ser feito por meio de sugestões de links, baseadas nas experiências de aprendizagem de outros estudantes, por exemplo;
- **Orientada para educadores:** objetiva fornecer *feedback* aos instrutores, avaliar a estrutura do conteúdo do curso e a sua eficácia no processo de aprendizagem. Também é foco de interesse classificar grupos de estudantes baseado em suas necessidades, descobrir padrões de aprendizagem regulares e irregulares, descobrir os erros frequentes, descobrir atividades que são mais eficazes, organizar o conteúdo de forma mais eficiente para o progresso do aluno;
- **Orientada para gestores educacionais:** o objetivo é ter parâmetros sobre como melhorar a eficiência do sistema e adaptá-lo ao comportamento de seus usuários, ter medidas sobre como melhor organizar os recursos institucionais (humanos e materiais), melhorar a oferta de programas educacionais, entre outros.

Atualmente, a Mineração de Dados Educacionais – *Educational Data Mining* (EDM) – vem se estabelecendo como uma forte e consolidada linha de pesquisa tanto nos EUA, quanto na Europa; contudo, no Brasil, essa área ainda é pouco explorada, apesar de crescer no país o potencial para a pesquisa, o desenvolvimento e a aplicação da mineração em ambientes educacionais (BAKER et al., 2011).

Há na literatura, diversos trabalhos sobre mineração de dados educacionais. Kampff et al. (2008) utilizaram técnicas de mineração para identificar perfis de alunos com risco de evasão ou reprovação em ambientes virtuais de aprendizagem (AVA). Ao verificar que um aluno possui tais características/comportamentos, o sistema alerta o professor que poderá tomar decisões pedagógicas necessárias para que o aluno fique mais motivado, se engaje, se sinta apoiado e seguro na sua aprendizagem.

Na pesquisa de Dias et al. (2008), os autores objetivaram analisar a utilização do ambiente de ensino-aprendizagem LabSQL, a partir da aplicação de métodos provenientes da mineração de dados, denominados Árvore de Decisão e Redes Bayesianas, e descobrir informações relevantes sobre o perfil do aluno com relação à utilização dessa tecnologia.

Prata et al. (2009) aplicaram técnicas de mineração de dados em atividades colaborativas realizadas em ambientes de educação a distância. No trabalho, os autores estudaram a relação entre os atos de colaboração e a aprendizagem dos estudantes. Os resultados obtidos apontaram que os discentes, que aprendem satisfatoriamente o

conteúdo apresentado pelo professor, possuem maior probabilidade de mostrar comportamentos inapropriados específicos, como insultos, durante o andamento do curso. Tal resultado é interessante, pois muitos educadores acreditam que os estudantes com baixo desempenho escolar são os maiores responsáveis por atrapalhar a aprendizagem dos demais alunos, contradizendo os dados obtidos na pesquisa de Prata et al. (2009).

Correia et al. (2011) utilizaram a tarefa de clusterização para formar turmas similares de recuperação paralela. O estudo permitiu agrupar os estudantes pelos problemas de aprendizagem apresentados, inclusive, separando em um grupo específico aqueles com dificuldades de interpretação, podendo, assim, formar uma turma interdisciplinar.

O trabalho de Sheard (2010) concentra-se no uso de técnicas estatísticas e explicita como elas podem ser aplicadas na análise de *logs* de arquivos recolhidos de ambientes de aprendizagem baseados na Web. Já Mazza (2010), faz uso da técnica de visualização em sistemas educacionais Web, permitindo um rápido descobrimento de padrões e regularidades de aprendizagem para representar modelos de usuários.

Há ainda trabalhos relacionados à construção de ferramentas de mineração de dados que sejam fáceis de serem utilizadas por educadores e não especialistas na área de mineração. Nesse contexto, destacam-se as pesquisas de Pedraza-Perez et al. (2011) e Merceron et al. (2005).

Pesquisas na área de EDM também proporcionaram modelos automatizados que podem ser utilizados durante a interação dos alunos com a interface do sistema, possibilitando verificar automaticamente quando ocorrem comportamentos inadequados. Assim, essa funcionalidade permite que sistemas educacionais apresentem comportamentos inteligentes. Baker et al. (2006), por exemplo, desenvolveram um personagem que reage de acordo com o comportamento apresentado pelo aluno no ambiente. Quando o estudante interage de maneira adequada, o personagem faz sinal de positivo. Já quando tenta trapacear, por exemplo, solicitando ajuda ao sistema diversas vezes, na tentativa de obter a resposta de determinado exercício sem ao menos tentar resolvê-lo, o personagem altera seu comportamento e tenta diagnosticar o conhecimento do aluno através de uma sequência de perguntas adicionais. Segundo os autores, o comportamento do personagem desenvolvido auxiliou o professor a identificar os estudantes que estavam aprendendo "corretamente" e também incentivou os discentes a manterem um comportamento adequado no processo de aprendizagem.

4. Estudo de Caso

Para a realização deste trabalho, utilizou-se os dados provenientes de avaliações da aprendizagem da disciplina Programação Orientada a Objetos, ministrada no primeiro semestre de 2010, no curso de Licenciatura em Computação da Universidade de Pernambuco, onde havia 33 estudantes matriculados.

A abordagem adotada na disciplina foi além da avaliação convencional, baseada no rendimento alcançado pelo aluno em poucas provas. Adotaram-se, assim, avaliações formativas, ocorridas ao longo do processo de aprendizagem. Dentro deste escopo, neste trabalho, foram considerados os dados obtidos a partir de 11 sessões de avaliação, a saber: 5 listas de exercícios, 4 miniprovas e 2 projetos. As atividades ocorreram durante todo o semestre e contemplaram diferentes conceitos. Cada sessão era composta por “n”

enunciados de problemas, totalizando 890 exercícios a serem avaliados pelo professor titular da disciplina. Todas essas atividades antecederam as duas avaliações somativas, ocorridas ao fim de cada bimestre, e a análise e interpretação desses dados podem ser exploradas a fim de fornecer ao docente um melhor discernimento sobre o conhecimento e dificuldades dos estudantes na aprendizagem de programação. As avaliações foram formuladas de modo a obter uma medida cognitiva, que deveria retratar o real desempenho do estudante na resolução de cada problema e seria resultado do processo de correção do professor.

Após a análise dos enunciados dos problemas, foi possível classificá-los de acordo com os níveis da Taxonomia de Bloom revisada: Lembrar, Entender, Aplicar, Analisar, Avaliar, Criar (ANDERSON & KRATHWOHL, 2001). Pelo fato de as questões não terem sido formuladas, tendo como referência uma taxonomia de objetivos educacionais, uma limitação deste trabalho está na classificação. Assim, o professor-autor dessas tarefas poderia ter um objetivo diferente do apresentado nos enunciados e ter equivocadamente utilizado um verbete de um nível cognitivo que não desejava avaliar naquele momento nos estudantes, levando, assim, à classificação, de uma questão em um nível superior ou inferior ao desejado pelo docente.

Ao final de cada bimestre, após a realização de avaliações formativas, foram realizadas avaliações somativas. Os 33 estudantes matriculados na disciplina realizaram tais avaliações, sendo que os valores possíveis estavam na faixa entre 0 e 10. Para efeitos de análise, os estudantes foram agrupados em dois grupos distintos. Os que alcançaram uma média igual ou superior a 7, média requerida para aprovação, foram rotulados como aprovados, e os demais estudantes, que obtiveram uma média inferior a 7 ou um número de faltas superior a 25% , foram rotulados como reprovados.

A análise das médias das avaliações somativas revelou uma taxa de sucesso de apenas 24,24% dentre os estudantes da disciplina. Tal taxa de sucesso demonstra claramente a dificuldade encontrada por esses estudantes. A partir dessa realidade, passou-se a tentar agrupar os alunos com características similares, considerando os dados disponíveis com o histórico das avaliações formativas, visando o planejamento de estratégias pedagógicas que busquem recuperar os grupos de aprendizes com dificuldades na aprendizagem de conteúdos da disciplina Programação Orientados a Objetos.

A escolha dos atributos para compor o conjunto de dados experimental sobre os 33 alunos selecionados foi baseada no modelo para mapeamento do conhecimento de estudantes, proposto por França (2013). O Quadro 1 reapresenta esses atributos, com sua descrição, tipo e o domínio de valores que eles podem assumir.

O ambiente utilizado neste estudo de caso para aplicação das técnicas de mineração de dados foi o WEKA - *Waikato Environment for Knowledge Analysis*, desenvolvido pela Universidade de Waikato da Nova Zelândia (WITTEN & FRANK, 2005).

Quadro 1 – Atributos para o mapeamento do conhecimento do estudante

Atributo	Descrição	Tipo de Dado	Domínio
IdAluno	Código que identifica o aluno participante da sessão de avaliação	Numeric	[1, n]
IdSessao	Código que identifica o número da sessão de avaliação e a sequência temporal	Numeric	[1, n]

Analisando os *clusters* da Figura 1 é possível realizar algumas interpretações:

- *Cluster 0*: Neste *cluster* com 113 ocorrências de itens (13%), foram agrupados os estudantes que alcançaram conceito “C” na avaliação formativa, composta por itens que tratavam sobre *Arrays*, abordados no nível Entender da Taxonomia de Bloom, e que obtiveram rótulo APR (aprovado) na disciplina;
- *Cluster 1*: Neste *cluster* com 144 ocorrências de itens (16%), foram agrupados os estudantes que alcançaram conceito “B” na avaliação formativa, composta por itens que abordavam o conteúdo *Arrays*, tratado no nível Criar da Taxonomia de Bloom, e que obtiveram rótulo APR (aprovado) na disciplina;
- *Cluster 2*: Este *cluster* agrupa os estudantes reprovados, que não responderam às questões (351 itens, perfazendo 39%) das avaliações formativas que tratavam sobre Estruturas de Repetição, as quais foram classificadas no nível Criar da Taxonomia de Bloom;
- *Cluster 3*: Este *cluster* agrupa os estudantes que realizaram satisfatoriamente as questões sobre Componentes Gráficos do Pacote GUI e Objetos e Classes (14% do total de itens), trabalhadas no nível Aplicar, mas que, no entanto, foram reprovados;
- *Cluster 4*: Foram agrupados, neste *cluster*, estudantes que obtiveram baixo desempenho na resolução de itens (8% de 890) que tratavam sobre Estruturação de Sistemas em Camadas que foram trabalhados no nível Entender. No final da disciplina, os alunos deste grupo foram rotulados como reprovados.
- *Cluster 5*: Representa aqueles estudantes com baixo desempenho na resolução de itens (10% de 890) de avaliações formativas que tratavam sobre *Arrays*, abordados a partir do nível Criar. Os estudantes deste grupo foram rotulados como reprovados.

6. Considerações Finais

Os resultados apresentados confirmam que técnicas de clusterização são bastante úteis para a formação de grupos homogêneos de estudantes. A identificação de grupos similares em turmas no ensino presencial é um mecanismo importante que o educador pode utilizar para planejar suas atividades pedagógicas considerando as especificidades dos aprendizes.

No experimento descrito, o *K-means* foi utilizado. O uso desse algoritmo permitiu agrupar alunos pelas suas dificuldades de aprendizagem na disciplina Programação Orientada a Objetos. Foi possível, por exemplo, identificar um grupo de aprendizes com dificuldades em **Entender** Estruturação de Sistemas em Camadas e outro composto por estudantes com problemas em **Criar** *Arrays*. Ambos os grupos foram reprovados ao final do semestre.

Vale ressaltar que ainda não é possível concluir se os resultados obtidos com essa técnica são significativos. Entretanto, as informações descobertas poderão ser utilizadas para apoiar as reuniões pedagógicas e a tomada de decisão na formação de grupos de estudantes mais homogêneos.

Espera-se que, com os resultados dos padrões encontrados, estratégias pedagógicas sejam formuladas, a exemplo de atividades de monitoria dirigidas a grupos de aprendizes com tendência a não alcançar o desempenho requerido.

Referências Bibliográficas

ANDERSON, L. W.; KRATHWOHL, D. R. (Ed.). **A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives**. New York: Addison Wesley Longman, 2001.

BAKER, R. S. J. D.; CORBETT, A. T.; KOEDINGER, K. R.; EVENSON, S. E.; ROLL, I.; WAGNER, A. Z.; NAIM, M.; RASPAT, J.; BAKER, D. J.; BECK, J. **Adapting to When Students Game an Intelligent Tutoring System**. In: Proceedings of the International Conference on Intelligent Tutoring Systems, 2006, p. 392-401.

BAKER, R.; ISOTANI, S.; CARVALHO, A. **Mineração de Dados Educacionais: Oportunidades para o Brasil**. Revista Brasileira de Informática na Educação, v. 19, n. 02, p. 3-13, 2011.

CORREIA, C. F; PIMENTEL, E. P. . **Mineração de Dados na Formação de Turmas para a Recuperação Paralela na Educação Básica**. In: XXII Simpósio Brasileiro de Informática na Educação, 2011, Aracaju. Anais do XXII SBIE - XVII WIE, 2011. p. 172-175.

DIAS, M. M. ; FILHO., L. A. DA S. ; LINO, A. D. P. ; FAVERO, E. L. ; RAMOS, E. M. L. S. **Aplicação de Técnicas de Mineração de Dados no Processo de Aprendizagem na Educação a Distância**. In: XIX Simpósio Brasileiro de Informática na Educação, 2008, Fortaleza - CE. Anais do XIX SBIE, 2008. p. 105-114.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P.. **From Data Mining to Knowledge Discovery**. American Association for Artificial Intelligence. 1996.

FAYYAD, U.; GRINSTEIN, G. G.; WIERSE, A. . Information Visualization. In: **Data Mining And Knowledge Discovery**. Morgan Kaufmann Publishers Inc, San Francisco, CA, USA, 2002.

FRANÇA, R. S. de. **Aplicação de Técnicas de Mineração de Dados, Baseada na Taxonomia de Bloom, para o Mapeamento do Conhecimento na Aprendizagem de Programação**. Garanhuns: Universidade de Pernambuco, 2013. 45p. Monografia de Graduação.

KAMPPFF, A. J. C.; REATEGUI, E. B.; LIMA, J. V. DE. **Mineração de dados educacionais para a construção de alertas em ambientes virtuais de aprendizagem como apoio à prática docente**. Novas Tecnologias na Educação. v. 6, Nº 2, Dezembro, 2008.

KRATHWOHL, D. R. **A revision of Bloom's taxonomy: an overview**. Theory in Practice, v. 41, n. 4, p. 212-218, 2002.

LOPES, C. C. **Um Sistema de Apoio à Tomada de Decisão no Acompanhamento do Aprendizado em Educação a Distância**. Campina Grande: UFCG, 2003. Dissertação de Mestrado.

MAZZA, R. Visualization in Educational Environments. In: ROMERO, C., VENTURA, S., PECHENIZKY, M., BAKER, R.. **Handbook of Educational Data**

Mining. Editorial Chapman and Hall/CRC Press, Taylor & Francis Group. Data Mining and Knowledge Discovery Series, 2010.

MERCERON, A.; YACEF, K. **TADA-Ed for Educational Data Mining**. Interactive Multimedia Electronic Journal of Computer-Enhanced Learning. Volume 7, Number 1, May, 2005.

PEDRAZA-PEREZ, R.; ROMERO, C.; VENTURA, S. **A Java desktop tool for mining Moodle data**. In: Proceedings of the 4th International Conference on Educational Data, 2011.

PERRENOUD, P.. **Dez Novas Competências para Ensinar**. Porto Alegre: Artes Médicas Sul, 2000a.

PERRENOUD, P. **Pedagogia diferenciada: das intenções à ação**. Porto Alegre: Artes Médicas, 2000b.

PIMENTEL, E. P.; FRANÇA, V. F.; OMAR, N.. **A identificação de grupos de aprendizes no ensino presencial utilizando técnicas de clusterização**. In: XIV Simpósio Brasileiro de Informática na Educação, 2003, Rio de Janeiro. Anais do XIV SBIE, 2003.

PIMENTEL, E. P.; OMAR, N.. **Descobrendo Conhecimentos em Dados de Avaliação da Aprendizagem com Técnicas de Mineração de Dados**. In: XII Workshop de Informática na Escola, 2006, Campo Grande, MS. Anais do XXVI CSBC, 2006.

PRATA, D. N.; BAKER, R. S. J. D.; COSTA, E., ROSÉ, C. P.; CUI, Y.; CARVALHO, A. M. J. B de. **Detecting and Understanding the Impact of Cognitive and Interpersonal Conflict in Computer Supported Collaborative Learning Environments**. In: Proceedings of the International Conference on Educational Data Mining, 2009. p. 131-140.

REZENDE, S. O.; PUGLIESI, J. B.; MELANDA, E. A.; PAULA, M. F. de. Mineração de dados. In: REZENDE, S. O. . **Sistemas Inteligentes: Fundamentos e Aplicações**. Barueri, SP: Manole, 2003.

ROMERO, C; VENTURA, S.. **Educational Data Mining: A Survey from 1995 to 2005**. Expert Systems with Applications 33, 2007, p. 125-146.

SHEARD, J. Basics of Statistical Analysis of Interactions Data from Web-Based Learning Environments. In: ROMERO, C., VENTURA, S., PECHENIZKY, M., BAKER, R. **Handbook of Educational Data Mining**. Editorial Chapman and Hall/CRC Press, Taylor & Francis Group. Data Mining and Knowledge Discovery Series, 2010.

WITTEN, I. H.; FRANK, E. . **Data Mining: Practical Machine Learning Tools and Techniques**. 2nd edition, Morgan Kaufmann Publishers, San Francisco, CA, 2005.

WITTEN, I. H.; FRANK, E.; HALL, M. A.. **Data Mining: Practical Machine Learning Tools and Techniques**, 3rd edition, Morgan Kaufmann, Burlington, MA, 2011.