

Um estudo sobre as variáveis para predição de alunos não concluintes em cursos suportados por Ambientes Virtuais de Ensino e Aprendizagem

Júlia Marques Carvalho da Silva¹, Hazra Imran²

¹Instituto Federal do Rio Grande do Sul – Campus Bento Gonçalves
Av. Osvaldo Aranha, 540 – Bento Gonçalves, RS – Brasil – 95700-000

julia.silva@bento.ifrs.edu.br

²Douglas College – Faculty of Commerce and Business Administration
New Westminister, BC – Canada

imranh@portal.douglas.bc.ca

Resumo. Os alunos não concluintes são aqueles que não apresentam presença ou desempenho suficiente ao final de uma disciplina ou curso. Em sala de aula, cabe ao professor perceber o comportamento do aluno a fim de identificar seu interesse e progresso. Entretanto, os ambientes virtuais de aprendizagem nem sempre dispõem das informações de maneira simplificada ao professor. Com base em uma análise na literatura, este artigo propõe sete variáveis que podem prever a conclusão ou não do aluno. Ainda, dados de 1168 alunos concluintes e não concluintes em 89 disciplinas de 5 cursos de uma instituição de ensino foram estatisticamente comparados a fim de verificar diferenças significativas. Os resultados mostraram que concluintes e não concluintes possuem aproveitamento diferentes, isto é, que as variáveis propostas podem ser utilizadas em um modelo de predição de alunos não concluintes.

Palavras-chave. Evasão, Reprovação, Fatores, Modelo.

Abstract. At-risk students are those who have not attendance or grades enough to successful finish a course. In the classroom, the teacher notices the student's behavior in order to identify their interest and progress. However, in virtual learning environments, this task is not easy, since the teacher needs to look for many information spread. Based on an analysis of the literature, this paper proposes seven variables that can predict the student's outcome. The sample contained data from 1168 students in 89 courses of 5 programs from an educational institute. It was applied the statistical method student t-test in order to determine significant differences between completers and non-completers. The results showed that completers and non-completers have different behavior and performance, that is, the proposed variables can be used in a prediction model of at-risk students.

Keywords. Drop-out, Non-successful completers, Factors, Model.

1 Introdução

Diversas instituições de ensino têm feito uso dos ambientes virtuais de ensino aprendizagem (AVEA), seja como suporte ao ensino presencial ou em atividades a distância, através da disponibilização de conteúdos e atividades aos alunos. A principal

vantagem de sua utilização consiste no armazenamento das interações dos participantes em bases de dados completas (Araque *et al.*, 2009; Macfadyen; Dawson, 2010). Contudo, como consequência, cabe ao professor lidar com toda esta informação a fim de certificar a aprendizagem pelo aluno. Contudo, a tarefa de analisar relatórios de um AVEA não é considerada fácil (Hogo, 2010). Além disto, os atuais AVEAs têm uma variedade limitada de relatórios (Dawson *et al.*, 2008; Mazza; Dimitrova, 2007 citado por Macfadyen; Dawson, 2010).

Várias pesquisas vêm sendo conduzidas a fim de identificar potenciais alunos não concluintes. Alguns deles utilizam o perfil demográfico para determinar a aderência do aluno no curso desejado, tais como: idade do ingressante, tipo de admissão, local de residência, nível escolar e econômicos dos pais, etc. (Aguiar *et al.*, 2014; Araque, 2009; Billings, 1987; Delen, 2010; Dupin-Bryant, 2004; Lauría *et al.*, 2012; Levy, 2007; Lin *et al.*, 2009; Menager-Beeley, 2001; Morris *et al.*, 2005; Pacheto *et al.*, 2009; Pal, 2012; Parker, 1999; Powell *et al.*, 1990; Wolff *et al.*, 2013; Xenos *et al.*, 2002). Com elas, foi possível concluir que os alunos concluintes têm perfil jovem, bom status socioeconômico, e pais com bom nível escolar (Araque, 2009). Billings (1987) e Levy (2007) descrevem que estas variáveis influenciam apenas indiretamente, já que é importante ressaltar que o uso de dados demográficos pode ser relevante quando se observa um contexto escolar específico. Entretanto, seja em cursos a distância ou presencial envolvendo larga abrangência de participantes, as diversidades de perfis tornam mais dificultoso estabelecer um perfil generalizado de concluintes e não concluintes. Desta forma, no presente trabalho optou-se por não utilizar variáveis demográficas na determinação de alunos em risco ou não, uma vez que o algoritmo desenvolvido poderá ser utilizado em diversos pontos geográficos e de diversidade econômica, social e cultural.

Outros estudos especializaram-se em analisar dados de cursos online (Dupin-Bryant, 2004; Levy, 2007; Menager-Beeley, 2001; Morris *et al.*, 2005; Velho e Pinto, 2015; Wolff *et al.*, 2013) ou presenciais (Delen, 2010; Lin *et al.*, 2009). Hu *et al.*, 2014; Macfayden & Dawson, 2010; Morris *et al.*, 2005; Romero *et al.*, 2008; Romero *et al.*, 2009 apresentam pesquisas que utilizam dados de cursos já finalizados para compreender os padrões de comportamento e performance de alunos concluintes e não concluintes. Embora os estudos sejam interessantes, a limitação consiste em apresentar variáveis muito específicas, as quais podem ser aplicadas em qualquer contexto. Percebe-se assim que poucos são os estudos que apresentam modelos concretos que fazem uso de informações contidas nas bases de dados de AVEAs para prever o resultado do aluno, que não considere seu perfil anterior ao ingresso, o tipo de curso (online ou presencial) e o planejamento de uso dos recursos e tarefas (Zielinski e Schmitt, 2015).

Sendo assim, o presente artigo investiga: “Quais são os fatores de performance e comportamento que podem identificar potenciais alunos não concluintes em disciplinas/cursos que utilizam AVEA?”. Para respondê-la, foi necessário seguir quatro passos:

1. Considerando as pesquisas já existentes e as características dos AVEAs mais utilizados, sete variáveis foram selecionadas a fim de investigar se elas podem distinguir alunos concluintes e não concluintes;
2. Foram obtidos dados de alunos matriculados em disciplinas já finalizadas que fizeram uso de AVEA, provenientes tanto do próprio AVEA (arquivos com

- os registros de acesso, que indica a interação do aluno) quanto do sistema acadêmico da instituição (indicando a aprovação, desistência ou reprovação);
3. A partir da base de dados do AVEA, foram calculados os valores das variáveis a fim de realizar os cálculos estatísticos. Tais valores foram normalizados em uma escala de 0 a 1, onde 0 significa baixo aproveitamento em uma tarefa ou baixa quantidade de acesso de um recurso, enquanto 1 significa que o aluno acessou todo o conteúdo ou obteve uma ótima nota.
 4. Por fim, foi necessário analisar os resultados estatísticos.

O artigo encontra-se organizado da seguinte forma: a seção 2 discorre sobre os fatores selecionados para investigação; a seção 3 descreve o conjunto de amostra; a seção 4 mostra os resultados estatísticos; e a seção 5 traz as conclusões.

2 Fatores Selecionados

Com base nas variáveis descritas em estudos anteriores e considerando as informações disponíveis em AVEAs, foram selecionados sete fatores relacionados a performance e comportamento. Eles foram agrupados em três categorias: notas, objetos de aprendizagem e frequência; os quais são descritos nas subseções a seguir.

2.1 Notas: obrigatórias e opcionais

As notas se referem aos valores atribuídos pelos professores às atividades solicitadas e entregues pelos alunos. Tal valor depende da escala a critério do professor, por exemplo, de 0 a 100 ou até mesmo usando rótulos (ex.: “Muito bom”, “Em processo de construção”). Ainda, os professores podem também indicar quais atividades valem nota, ou seja, nem toda tarefa deverá ter uma nota vinculada.

No contexto de aprendizagem, as médias de notas consistem em uma informação confiável sobre a performance do aluno (Haig *et al.*, 2013). Ela fornece ao professor uma visão sobre o aproveitamento do aluno em um curso. Conseqüentemente, a avaliação não consiste apenas na realização / entrega ou não na atividade, mas na qualidade do trabalho realizado e na assimilação dos conceitos. As pesquisas de Macfadyen; Dawson (2010), Costa *et al.* (2014), e Arnold; Pistilli (2012) utilizam este fator na predição de alunos não-concluintes.

Aqui, o fator nota é distinguido em dois sub-fatores: obrigatório e opcional. A ideia de considerar as notas obrigatórias e opcionais já foi apresentado por Wolff *et al.* (2013). Ele defende que, dentro do processo de ensino e aprendizagem uma avaliação pode ser opcional ou substitutiva por outra. O professor é quem determina quais atividades devem compor o aproveitamento do aluno.

No modelo a ser proposto pelo projeto, os subfatores serão calculados similarmente através da média ponderada, onde os pesos de cada atividade devem ser respeitados. Entretanto, é importante observar que, no caso das notas de atividades obrigatórias, se o estudante não a realiza dentro do prazo, a ele é atribuído o valor zero; e se o professor não atribui a nota, esta atividade não pode ser considerada no cálculo.

2.2 Objetos de Aprendizagem: obrigatório e opcional

Os objetos de aprendizagem contêm recursos e atividades de um curso, os quais permitem alguma interação entre o aluno e o AVEA, como a submissão de uma tarefa avaliativa,

enviar mensagem para um fórum de discussão, resolver um questionário, conversar em uma sala de bate-papo, registrar uma contribuição em *wiki*, adicionar um termo em um glossário, inserir um registro em uma base de dados, ler um texto, assistir um vídeo, etc. Este fator é encontrado em diversos estudos (Billings, 1987; Hu *et al.*, 2014; Macfadyen; Dawso, 2010; Morris *et al.*, 2005; Romero *et al.*, 2008; Romero *et al.*, 2009; Xenos *et al.*, 2002), contudo os pesquisadores o consideram especificamente, por exemplo, a quantidade de textos visualizados, número de mensagens enviadas para o fórum, ou o número de atividades submetidas.

O objetivo deste fator é observar se os alunos estão comprometidos com o curso através da verificação da participação ativa. Ainda, ele determina a ação realizada em cada objeto de aprendizagem, isto é, se o aluno submeteu duas tentativas em um mesmo questionário, apenas uma destas ações (tentativas) é considerada.

Estudos anteriores descrevem que o número total de visitas a um objeto de aprendizagem pode indicar o interesse do aluno (Macfadyen; Dawson, 2010). Outros afirmam que o tempo dispendido em um único recurso pode indicar a qualidade da aprendizagem (Morris *et al.*, 2005). Entretanto, a nova geração de alunos parecem apresentar novos comportamentos no processo de aprendizagem. Normalmente, eles mantêm mais de uma aba do navegador aberta ao mesmo tempo, acessando diversos conteúdos simultaneamente (Thompson, 1998) ou, por vezes preferem salvar uma cópia do objeto de aprendizagem em seu computador, não retornando ao AVEA (Wolff *et al.*, 2013). Consequentemente, a quantidade de vezes que o aluno acessou ou quanto tempo ele despendeu navegando pode ser uma medida discutível.

Desta forma, nesta pesquisa considerou-se apenas o fato do aluno ter interagido ou não com o objeto de aprendizagem. Em outras palavras, assume-se que o aluno deve acessar cada material didático pelo menos uma vez, e não importa se ele voltou a acessá-lo posteriormente. Ainda, tal como nas notas, cabe ao professor indicar se é obrigatório ou opcional o acesso a um objeto de aprendizagem pelo aluno.

2.3 Frequência: último acesso, acesso semanal, e carga de atividades

A importância do aluno manter uma regularidade em sua hora de estudo é discutida por Powell *et al.* (1999). Shin; Kim, 1999; Verdigis; Panagiotakopoulos, 2002; Xenos *et al.*, 2002, discorrem sobre como a falta de autorregulação de tempo necessária para os estudos afetam diretamente na conclusão de um curso. Ainda, há estudos que descrevem como a frequência semanal em cursos ofertados via AVEA auxilia no comprometimento do aluno (Er, 2012; Zhang *et al.*, 2007), e portanto, pode ser um indicador confiável de aprovação do estudante (Wolff *et al.*, 2013).

Aqui, o fator frequência foi dividido em três subfatores: “frequência semanal”, “último acesso”, e “carga de atividades”. A “frequência semanal” é utilizada para verificar se os alunos acessaram o AVEA ao menos uma vez por semana. Morris *et al.* (2005) confirmam que a frequência de participação é importante para uma aprendizagem com sucesso.

O subfator “último acesso” observa se o aluno realizou alguma ação recentemente. Para o cálculo, verifica-se a duração total do curso e então se extrai o período correspondente a 20% mais recente. Por exemplo, se um curso em andamento tem a duração de cinco meses e o aluno não têm visitado o AVEA há mais de um mês, isto representa que a sua frequência é aquém da desejada e ele está propenso a não concluir.

Finalmente, o subfator “carga de atividades” mede quantas ações o aluno realizou no período correspondente aos últimos 20% do curso, e então compara este valor com o período prévio. Novamente, assumindo que o curso dura cinco meses, compare-se o número de atividades realizadas no último mês com o número de atividades realizadas no penúltimo mês. Tal verificação se faz importante, uma vez que há alunos que continuam visitando o AVEA, mas realiza menos ações que ele realizada anteriormente. Logo, isto pode indicar uma diminuição no comprometimento com os estudos, e então, tornando-se um potencial aluno não concluinte (Casey *et al.*, 2010 citado por Haig, 2012; Wolff *et al.*, 2013; Zhang *et al.*, 2007).

Ainda, é importante criar um mecanismo de filtro na base de dados de registros de atividades do aluno, pois conforme Wolff *et al.* (2013), alguns tipos de acesso (ex.: página inicial, página dos colegas) não afetam o processo de aprendizagem.

3 Descrição da Amostra

Para verificar os fatores propostos, foi conduzida uma análise estatística. A amostra foi composta por 89 disciplinas de cinco cursos de uma instituição de ensino. Elas ocorreram entre os anos 2012 e 2014, nos níveis de ensino médio, pós-médio e graduação, os quais oferecem disciplinas presenciais e a distância. Nas turmas de pós-médio e graduação, os alunos apresentam idades a partir de 17 anos, enquanto no ensino médio, eles têm entre 13 e 19 anos. A duração de cada disciplina também varia: no pós-médio e graduação dura 20 semanas, enquanto no ensino médio são 40 semanas. Outra diferença dá-se pela distribuição das disciplinas em cada período letivo:

- Os estudantes de graduação têm uma média de cinco disciplinas por semestre, sendo cada uma delas de 4 horas/aula, sendo que a maioria das disciplinas ocorre em um dia da semana; as disciplinas são presenciais e o uso do AVEA se dá como apoio das atividades.
- Os estudantes do pós-médio têm cerca de 12 disciplinas de 10, 20 ou 30 horas/aula, mas cada disciplina ocorre de forma contínua, isto é, a disciplina é lecionada compactamente, sem haver outra concomitantemente sendo ofertada; as disciplinas são a distância, utilizando o AVEA para todas as atividades (exceto avaliação presencial final).
- Os estudantes do ensino médio têm uma média de 17 disciplinas de forma concomitante e ofertadas semanalmente em encontros de 1, 2 ou 3 horas/aulas; as disciplinas são presenciais, e o AVEA é usado como apoio.

Neste estudo, a média de alunos por disciplina foi de 15,03, enquanto a média de disciplinas cursadas por aluno através do AVEA nos três anos foi de 3,97. O total de alunos analisados foi de 1168, sendo 965 (83%) concluintes e 203 (17%) não concluintes. Quanto ao perfil das disciplinas analisadas, a média de objetos de aprendizagem por disciplina foi de 29,53.

4 Resultados Estatísticos

A análise estatística tem sido usada para comparar alunos concluintes e não concluintes por diversos pesquisadores (Aguiar *et al.*, 2014; Dupin-Bryant, 2004; Morris *et al.*, 2005; Parker, 1999). De acordo com Romero *et al.* (2008), a estatística é considerada como um ponto de início para a avaliação de dados educacionais.

Inicialmente, foi necessário obter os dados e calcular os fatores para cada aluno matriculado em uma disciplina. Durante este processo, percebeu-se que apenas 56 (4,8%) registros apresentavam valores correspondentes as sete variáveis propostas. Este fato ocorre devido a não utilização de, por exemplo, avaliações com notas e tarefas obrigatórias. Como consequência, cada fator apresentou um tamanho de amostra diferente.

O valor da média e do desvio padrão foram de cada fator calculados para cada grupo de estudantes: concluintes e não concluintes. A Figura 1 apresenta o valor da média de cada fator para os concluintes, não concluintes, e para todo o grupo. Considerando que os valores encontram-se entre 0 e 1, percebe-se que a média tem uma diferença significativa entre os concluintes e não concluintes.

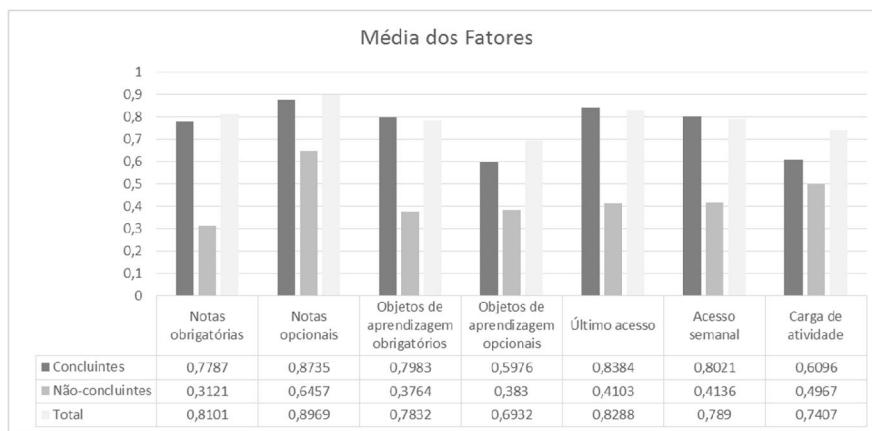


Figura 1– Valor médio de cada fator.

Entretanto, a média deve ser analisada juntamente com o desvio padrão, que é apresentado na Figura 2. Nota-se que o fator de último acesso, para os concluintes, tende a ser a informação mais homogênea da amostra; enquanto a qualidade de objetos de aprendizagem com visualização obrigatória e as notas opcionais dos alunos não concluintes, possuem as diferenças mais altas da amostra.

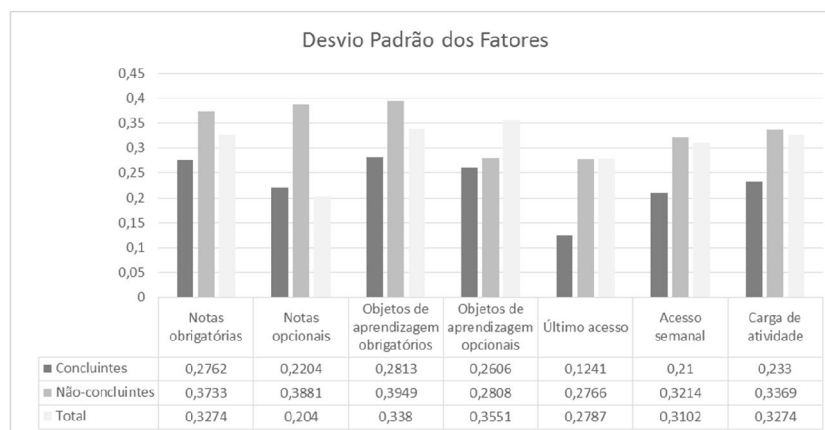


Figura 2– Desvio padrão de cada fator.

O passo seguinte foi aplicar o teste *t* de *student* a fim de verificar se os alunos concluintes apresentam melhores resultados que os não concluintes em cada fator. Para tanto, foram definidas sete hipóteses:

- H_a : “Alunos concluintes apresentam notas maiores em atividades obrigatórias que os alunos não concluintes.”
- H_b : “Alunos concluintes apresentam notas maiores em atividades opcionais que os alunos não concluintes.”
- H_c : “Alunos concluintes apresentam maior quantidade de visualização de objetos de aprendizagem obrigatórios que os alunos não concluintes.”
- H_d : “Alunos concluintes apresentam maior quantidade de visualização de objetos de aprendizagem opcionais que os alunos não concluintes.”
- H_e : “Alunos concluintes apresentam data de último acesso mais recente que os alunos não concluintes.”
- H_f : “Alunos concluintes apresentam maior quantidade de acesso semanal que os alunos não concluintes.”
- H_g : “Alunos concluintes apresentam maior carga de atividade que os alunos não concluintes.”

Para todas as hipóteses foi definida a hipótese nula e alternativa:

$H_0: \mu_1 = \mu_2$ e $H_1: \mu_1 > \mu_2$, onde μ_1 consiste em alunos concluintes e μ_2 em alunos não concluintes, considerando $\alpha = 0,05$

Os resultados foram sumarizados e apresentados na Tabela 1.

Tabela 2– Resultados do teste *t* de *student*.

Hipótese	Resultado
a	$t(393) = 10,02$
b	$t(106) = 2,45$
c	$t(924) = 15,64$
d	$t(1166) = 10,5$
e	$t(872) = 4,64$
f	$t(615) = 6,75$
g	$t(922) = 5,02$

De acordo com os resultados, a conclusão é que H_0 pode ser rejeitada, pois $t > t_{\text{critical}}$ e $p < 0,05$, e H_1 é verdadeira. Portanto, o teste *t* de *student* prova que os alunos concluintes têm melhores resultados que os não concluintes. Como já discutido por Morris *et al.* (2005), este resultado não é surpreendente, pois os alunos precisam ter boas notas e frequência para serem aprovados. Mesmo assim, este resultado permite validar as crenças sobre o comportamento e performance do aluno. A partir de agora surge a questão: “Como usar estes fatores para distinguir os alunos concluintes e não concluintes?”; e mais além: “Como usar estes fatores para prever alunos concluintes e não-concluintes?”.

5 Conclusão

O uso de AVEAs em disciplinas a distância ou como suporte as presenciais tem se tornado cada vez mais popular. Entretanto, mesmo com o expressivo volume de informações armazenadas nas suas bases de dados, cabe ao professor analisá-las a fim de identificar os alunos que têm propensão a não conclusão. E que estas informações poderiam ser utilizadas para compreender o perfil do aluno, seja pelo comportamento ou aproveitamento; podendo até mesmo prever a sua conclusão ou não do curso.

Contudo, a fim de se estabelecer um modelo de predição de alunos não concluintes, faz-se necessário identificar quais são os fatores que podem identificar tais alunos. Diversos estudos já foram realizados a fim de compreender estes fatores, porém ou estão vinculados ao contexto geográfico-social do estudante, ou a práticas docente e AVEAs específicos, impossibilitando a generalização.

Logo, este artigo buscou identificar fatores que possam distinguir perfil do aluno concluinte e não concluinte. Para isto, utilizou-se o método estatístico teste *t* de *student*, onde se concluiu que dados os fatores expostos, há sim diferenças significativas; sendo então possível a utilização destes. Portanto, a pergunta: “Quais são os fatores de performance e comportamento que podem identificar potenciais alunos não concluintes em disciplinas/cursos que utilizam AVEA?”, pode ser respondida como: as notas e acesso aos objetos de aprendizagem, sejam obrigatórios ou opcionais, aliados a frequência do aluno (averiguada pela data de “último acesso”, “acesso semanal”, e “carga de atividades”) podem ser usados como fatores para identificar possíveis alunos não concluintes dentro de um AVEA.

Como característica limitante da pesquisa, observou-se que apenas 56 registros continham valores para todos os fatores. Isto é justificável pelo *course design* escolhido pelo professor, que nem sempre opta pela atribuição de notas no AVEA ou pelo uso de tarefas obrigatórias. Logo, como trabalhos futuros, acredita-se que a utilização de técnicas de mineração de dados pode auxiliar na extração de modelos mais precisos, onde a ausência de dados pode ser configurada.

Referências

- AGUIAR, E., *et al.* Engagement vs performance: using electronic portfolios to predict first semester engineering student retention. In: **Proceedings of the Fourth International Conference on Learning Analytics and Knowledge**. ACM. 2014.
- ARAQUE, F.; ROLDÁN, C.; SALGUERO, A. Factors influencing university drop out rates. In: **Computers & Education**, 53.3, p. 563-574, 2009.
- ARNOLD, K. E.; PISTILLI, M. D. Course signals at Purdue: Using learning analytics to increase student success. In: **Proceedings of the 2nd International Conference on Learning Analytics and Knowledge**. ACM. 2012.
- BILLINGS, D. M. Factors related to progress towards completion of correspondence courses in a baccalaureate nursing programme. In: **Journal of Advanced Nursing**, 12.6, 1987, pp. 743-750.
- COSTA, S. S.; CAZELLA, S.; RIGO, S. J. Minerando dados sobre o desempenho de alunos de cursos de educação permanente em modalidade EAD: Um estudo de caso sobre evasão escolar na UNA-SUS. In: **Revista Novas Tecnologias na Educação**, 12.2, 2012.
- DAWSON, S. P.; McWILLIAM, E.; TAN, J. P. L. Teaching smarter: How mining ICT data can inform and improve learning and teaching practice. In: ASCILITE, 2008, Melbourne. Australia. Disponível em: <http://www.ascilite.org.au/conferences/melbourne08/procs/dawson.pdf>>.
- DELEN, D. A comparative analysis of machine learning techniques for student retention management. In: **Decision Support Systems**, 49.4, 2010, pp. 498-506.

- DUPIN-BRYANT, P. A. Pre-entry variables related to retention in online distance education. In: **The American Journal of Distance Education**, 18.4, 2004, pp. 199-206.
- ER, E. Identifying At-Risk Students Using Machine Learning Techniques: A Case Study with IS 100. In: **International Journal of Machine Learning and Computing**, 2012, 2(4)
- HAIG, T.; FALKNER, K.; FALKNER, N.. Visualisation of learning management system usage for detecting student behaviour patterns. In: **Proceedings of the Fifteenth Australasian Computing Education Conference**-Volume 136. Australian Computer Society, Inc., 2013. p. 107-115.
- HOGO, M. A. Evaluation of e-learners behaviour using different fuzzy clustering models: a comparative study. In: **arXiv preprint arXiv:1003.1499**, 2010.
- HU, Y.; LO, C; SHIH, S. Developing early warning systems to predict students' online learning performance. In: **Computers in Human Behavior**, 36, 2014, pp. 469-478.
- LAURÍA, E. J. M., *et al.* Mining academic data to improve college student retention: An open source perspective. In: **Proceedings of the 2nd International Conference on Learning Analytics and Knowledge**. ACM, 2012.
- LEVY, Y. Comparing dropouts and persistence in e-learning courses. In: **Computers & education**, 48.2, 2007, pp. 185-204.
- LIN, J. J. J.; IMBRIE, P. K.; REID, K. Student retention modelling: An evaluation of different methods and their impact on prediction results. In: **Research in Engineering Education Symposium**, 2009, pp. 1-6.
- MACFADYEN, L. P.; DAWSON S. Mining LMS data to develop an “early warning system” for educators: A proof of concept. In: **Computers & Education**, 54.2, 2010, pp. 588-599.
- MENAGER-BEELEY, R. Student Success in Web Based Distance Learning: Measuring Motivation To Identify at Risk Students and Improve Retention in Online Classes. In: **Proceedings of World Conference on the WWW and Internet**, Orlando, FL, 2001.
- MORRIS, L. V.; FINNEGAN, C; WU, S. Tracking student behavior, persistence, and achievement in online courses. In: **The Internet and Higher Education**, 8.3, 2005, pp. 221-231.
- PACHECO, A. S. V, et al. Fatores que Influenciam na Evasão dos Alunos em um Curso Livre. In: **Revista Novas Tecnologias na Educação**, 7.1, 2009.
- PAL, S. Mining Educational Data Using Classification to Decrease Dropout Rate of Students. In: **arXiv preprint arXiv:1206.3078**, 2012.
- PARKER, A. A study of variables that predict dropout from distance education. In: **International journal of educational technology**, 1.2, 1999, pp. 1-10.
- POWELL, R; CONWAY, C; ROSS, L. Effects of student predisposing characteristics on student success. In: **The Journal of Distance Education/Revue de l'Éducation à Distance** 5.1, 1990, pp. 5-19.
- ROMERO, C.; VENTURA, S; GARCÍA, E. Data mining in course management systems: Moodle case study and tutorial. In: **Computers & Education**, 51.1, 2008, pp. 368-384.

- ROMERO, C., *et al.* Evolutionary algorithms for subgroup discovery in e-learning: A practical application using Moodle data. In: **Expert Systems with Applications**, 36.2, 2009, pp. 1632-1644.
- SHIN, N.; KIM, J. An exploration of learner progress and drop-out in Korea National Open University. In: **Distance education**, 20.1, 1999, pp. 81-95.
- THOMPSON, M. M. Distance learners in higher education. In C.C. Gibson (Ed.), **Distance learners in higher education**, Madison, WI: Atwood Publishing, 1998, pp. 9-24.
- VELHO, D. S.; PINTO, A. S. As Competências EaD de alunos concluintes do Ensino Médio: resultados do instrumento de coleta de dados. In: **Revista Novas Tecnologias na Educação**, 13.1, 2015.
- VERGIDIS, D.; PANAGIOTAKOPOULOS, C. Student Dropout at the Hellenic Open University: Evaluation of the graduate program, Studies in Education. In: **The International Review of Research in Open and Distance Learning**. 3.2, 2002.
- WOLFF, A. *et al.* Improving retention: predicting at-risk students by analysing clicking behaviour in a virtual learning environment. In: **Proceedings of the Third International Conference on Learning Analytics and Knowledge**. ACM, 2013.
- XENOS, M.; PIERRAKEAS, C.; PINTELAS P. A survey on student dropout rates and dropout causes concerning the students in the Course of Informatics of the Hellenic Open University. In: **Computers & Education** 39.4, 2002, pp. 361-377.
- ZHANG, H.; ALMEROOTH, K.; KNIGHT, A.; BULGER, M.; MAYER, R. Moodog: tracking students online learning activities. In: **Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications**, 2007, pp. 4415-4422.
- ZIELINSKI, F. D. C; SCHMITT, M. A. R. Uma ferramenta gráfica para suporte a atividade docente no Moodle. In: **Revista Novas Tecnologias na Educação**, 13.1, 2015.