

Um Sistema Proativo para Monitoramento e Avaliação das Atividades de Tutoria a Distância em AVAs

Laysa Mabel de O. Fontes¹, Ricardo Alexandro de M. Valentim²,
Francisco Milton Mendes Neto³, Rafael C. de Souza³

¹Centro Multidisciplinar de Pau dos Ferros (CMPF)
Universidade Federal Rural do Semi-Árido (UFERSA)
BR 226, Km 405 - São Geraldo, Pau dos Ferros, RN | CEP: 59900-000

²Programa de Pós-Graduação em Engenharia Elétrica e de Computação (PPgEEC)
Universidade Federal do Rio Grande do Norte (UFRN)
BR 101 - Lagoa Nova, Natal, RN | CEP: 59078-970

³Programa de Pós-Graduação em Ciência da Computação (PPgCC)
Universidade Federal Rural do Semi-Árido (UFERSA)
BR 110, Km 47 - Presidente Costa e Silva, Mossoró, RN | CEP: 59625-900

mabel.fontes@ufersa.edu.br, ricardo.valentim@ufrnet.br,

{miltonmendes, rafaelcastro}@ufersa.edu.br

Resumo. *O uso dos Ambientes Virtuais de Aprendizagem (AVAs) na Educação a Distância tem gerado um crescente volume de dados provenientes de interações entre os atores desse processo. O tutor a distância é o ator responsável por mediar o processo de aprendizagem dos alunos e por promover a interação nos AVAs. No entanto, nem sempre os tutores a distância desempenham suas atividades de maneira adequada. Diante deste contexto, este trabalho objetiva apresentar um sistema proativo para monitoramento e avaliação das atividades de tutoria a distância em AVAs, baseado em processos de Learning Analytics.*

Abstract. *The use of Virtual Learning Environments (VLEs) in Distance Education has generated a growing volume of data from interactions among the actors in this process. The distance tutor is the actor responsible for mediating the students' learning process and for promoting interaction in the VLEs. However, distance tutors do not always perform their activities properly. Given this context, this work aims to present a proactive system for monitoring and evaluation of distance learning activities in VLEs, based on Learning Analytics processes.*

1. Introdução

O uso dos Ambientes Virtuais de Aprendizagem (AVAs) na Educação a Distância (EaD) tem gerado um crescente volume de dados provenientes de interações entre os atores desse processo. De acordo com o modelo adotado pela Universidade Aberta do Brasil (UAB), o tutor a distância é o ator responsável por mediar o processo de aprendizagem dos alunos e por promover a interação nos AVAs.

O volume de dados gerado a partir dessas interações, se devidamente explorado, pode fornecer o entendimento sobre a relação de influência entre o desempenho dos tutores a distância e a participação efetiva de alunos em AVAs.

Como desdobramento deste problema de pesquisa, tem-se a seguinte indagação: **O desempenho dos tutores a distância influencia a participação efetiva de alunos percententes à modalidade a distância?** Essa indagação, em caso de resposta afirmativa, e considerando a perspectiva de *Design Science* [Hevner et al. 2004, Wieringa 2014], permite definir a seguinte questão do conhecimento central: **É possível desenvolver um sistema proativo¹ que permita monitorar e avaliar o desempenho de tutores a distância em AVAs?**

Diante deste contexto, delimitou-se o objetivo de pesquisa deste trabalho como sendo o desenvolvimento de um sistema proativo para monitoramento e avaliação das atividades de tutoria a distância, baseado em processos de *Learning Analytics* [Siemens et al. 2011].

Este trabalho está organizado da seguinte forma: a Seção 2 apresenta os trabalhos correlatos com o objeto de pesquisa deste trabalho; a Seção 3 apresenta a solução proposta neste trabalho; a Seção 4 descreve os dados utilizados nos experimentos; e, por fim, a Seção 5 apresenta as considerações finais e os trabalhos futuros.

2. Trabalhos Relacionados

A *Learning Analytics* é uma área de pesquisa em expansão, tendo como principais enfoques os trabalhos relacionados com predição, agrupamento, mineração de relações, descoberta com modelos e tratamento de dados para apoio à decisão. Em todos estes casos, em maior ou menor grau, podem ser vislumbradas aplicações ligadas às questões discutidas no âmbito dos modelos teóricos sobre riscos de reprovação, evasão escolar e ligadas aos processos de antecipação de diagnósticos, conforme discutido nos trabalhos a seguir.

Em [Silva et al. 2016], foi proposto o uso de técnicas de *Learning Analytics* para analisar a situação de 807 alunos do Curso Técnico Semipresencial de Tecnologia da Informação do Instituto Metrópole Digital (IMD) da Universidade Federal do Rio Grande do Norte (UFRN) referente à contribuição da utilização do ambiente virtual em colaboração às atividades presenciais. Segundo os autores, os resultados encontrados permitiram identificar um padrão de comportamento entre os alunos capaz de fornecer uma base para a proposta de novos formatos de acompanhamento virtual do aluno.

Em [Srilekshmi et al. 2016], foi proposto um modelo para identificar os alunos que estão em situação de risco, usando regras de associação e agrupamento. Uma vez que um aluno é identificado em situação de risco, o modelo define o seu perfil: (i) aluno visual, isto é, aquele que gosta de aprender assistindo vídeos ou vendo imagens; (ii) aluno auditivo, ou seja, aquele que gosta de aprender ouvindo áudios ou palestras; e (iii) aluno cinestésico, isto é, aquele que gosta de aprender fazendo coisas. Dependendo do perfil do aluno, o modelo sugere métodos para melhorar o seu desempenho. Segundo os autores, diversos experimentos foram conduzidos e os resultados mostraram que o modelo proposto identificou 90% de alunos em risco.

Em [Iseppon et al. 2016], foi apresentado um estudo que verificou a relação entre as ações do professor e a participação dos alunos em AVAs, através da técnica de Análise de Correlação Canônica [Tabachnick 2012]. Segundo os autores, os resultados

¹Neste trabalho, os termos ‘sistema proativo’ foram empregados no sentido de sistema automatizado, isto é, aquele que não precisa de intervenção humana para realizar seus serviços.

obtidos mostraram que as ações do professor possuem bastante influência na interação de seus alunos no AVA. Além disso, os autores apresentaram um conjunto de variáveis que representam as ações do professor e as interações dos alunos no ambiente virtual.

Embora o presente trabalho esteja no mesmo campo de pesquisa que os trabalhos supracitados, este apresenta uma perspectiva diferente. O sistema apresentado neste artigo está voltado para o monitoramento e avaliação dos tutores a distância em AVAs.

3. Solução Proposta

O sistema proposto neste trabalho foi desenvolvido para monitorar e avaliar o desempenho dos tutores a distância, fazer recomendações de ações que possam promover a participação efetiva dos alunos e servir como subsídio para os gestores dos referidos cursos. A Figura 1 ilustra o fluxo do sistema proposto, denominado de MONITUM.

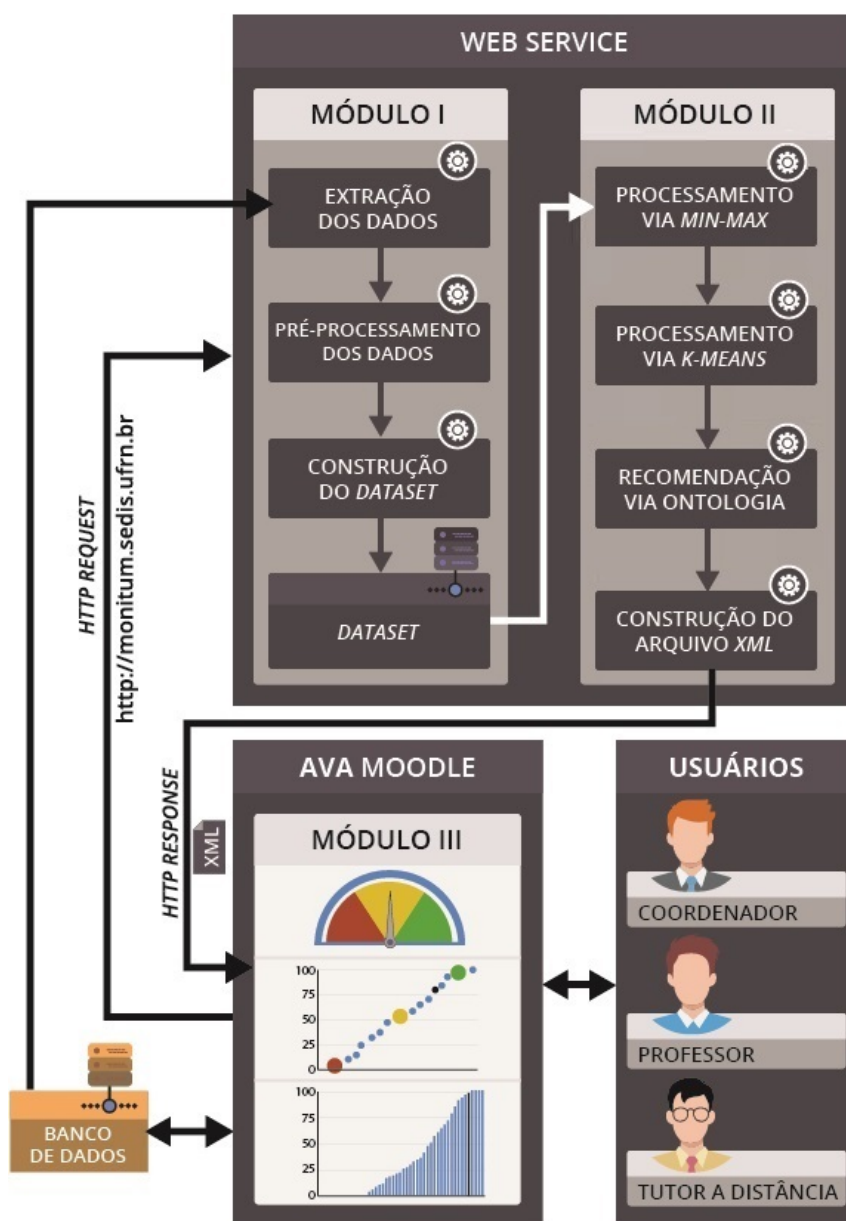


Figura 1. Fluxo do sistema MONITUM

Conforme pode ser visto na Figura 1, o sistema MONITUM é composto por três módulos. Os Módulos I e II compõem um *Web Service*, sendo este responsável por todo o processamento e análise dos dados. O Módulo III é um *plugin* e está integrado ao Moodle. Ele é responsável por consumir os resultados do *Web Service* e por apresentar aos coordenadores de curso, professores e aos próprios tutores a distância formas distintas de visualizar os desempenhos obtidos por cada tutor a distância em suas respectivas turmas.

O fluxo do sistema MONITUM inicia-se no Moodle. Para isso, foi desenvolvida uma rotina que solicita o serviço do *Web Service* para cada tutor a distância cadastrado no semestre corrente, sendo que essas solicitações são efetuadas mensalmente, a partir do início do semestre letivo. Cada solicitação é realizada através do protocolo *HTTP Request*. A partir de cada solicitação, o Módulo I executa os seus três processos, isto é, a extração dos dados, o pré-processamento dos dados e a construção do *dataset*, conforme ilustrado na Figura 1.

O Módulo II, por sua vez, realiza o processamento dos dados advindos do *dataset*, através das técnicas *Min-Max* [Marquesone 2014] e *k-Means* [MacQueen 1967], seleciona as recomendações a partir de uma ontologia e constrói um arquivo *xml* com o resultado de todos esses processamentos, conforme ilustrado na Figura 1. Por fim, o Módulo III consome os resultados do *Web Service*, isto é, processa o arquivo *xml* e, através de diferentes formas, exibe os resultados dos desempenhos obtidos pelos tutores a distância.

Os usuários do Moodle que têm acesso ao Módulo III são: (i) coordenadores de curso; (ii) professores; e (iii) os próprios tutores a distância, conforme ilustrado na Figura 1. As subseções a seguir apresentam os detalhes dos três módulos do sistema MONITUM.

3.1. Módulo I

O Módulo I é responsável por três processos: (i) extração dos dados; (ii) pré-processamento dos dados; e (iii) construção do *dataset*. Inicialmente, o Módulo I realiza a extração dos atributos ilustrados na Tabela 1.

Tabela 1. Atributos dos tutores a distância
Número de Questionários Criados
Número de Tópicos Criados nos Fóruns
Média de Postagens em Tópicos dos Fóruns
Taxa de Visualização em Fóruns
Taxa de Visualização em Tópicos dos Fóruns
Número de Tarefas Criadas
Número de Tarefas Avaliadas
Média de Postagens em <i>Chats</i>
Total de Cliques
Número de URLs Criadas
Número de Arquivos Criados

Após realizar a extração desses onze atributos dos tutores a distância, realiza-se o cálculo do atributo ‘desempenho’. Esse atributo representa o desempenho geral de cada tutor a distância. O atributo ‘desempenho’ é calculado por meio da Equação 3.1.

$$d = \frac{NTC + 2 \cdot (NTCF + MPTF + TVF + TVTF + NTA + TC + NAC) + 4 \cdot (NQC + MPC + NUC)}{7} \quad (3.1)$$

Onde,

NTC representa o número de tarefas criadas;

NTCF representa o número de tópicos criados nos fóruns;

MPTF representa a média de postagens em tópicos dos fóruns;

TVF representa a taxa de visualização em fóruns;

TVTF representa a taxa de visualização em tópicos dos fóruns;

NTA representa o número de tarefas avaliadas;

TC representa o total de cliques;

NAC representa o número de arquivos criados;

NQC representa o número de questionários criados;

MPC representa a média de postagens em *chats*;

NUC representa o número de URLs criadas.

A equação anterior trata-se de uma média ponderada que foi definida através de correlações encontradas em experimentos realizados anteriormente (detalhados em [Fontes 2017]), por meio do Coeficiente de Correlação de Pearson [LeBlanc 2004, Sharma 2012].

Os pesos dessa média foram utilizados para representar o impacto que os atributos analisados possuem na participação efetiva dos alunos. Vale ressaltar que esses pesos foram definidos de forma empírica. Para cada atributo, verificou-se a sua correlação mais relevante e, a partir disso, definiu-se por qual peso tal atributo seria multiplicado. Por exemplo, o atributo ‘número de questionários criados’ apresentou três correlações: uma correlação forte positiva e duas correlações moderadas positivas. Portanto, a correlação mais relevante desse atributo é a correlação forte positiva. Dessa forma, definiu-se que esse atributo deveria ser multiplicado pelo peso 4. Esse processo foi repetido para os demais atributos.

De um modo geral, o atributo que está sendo multiplicado pelo peso 1 impacta levemente, os atributos que estão sendo multiplicados pelo peso 2 impactam moderadamente e os atributos que estão sendo multiplicados pelo peso 4 impactam fortemente na participação efetiva dos alunos.

Após a extração de todos os atributos dos tutores a distância, o Módulo I realiza o pré-processamento dos dados, que consistiu basicamente na eliminação de dados inconsistentes. Por fim, o Módulo I constrói e gera o *dataset* no formato *arff*, que servirá de entrada para o Módulo II, detalhado na subseção a seguir.

3.2. Módulos II e III

A análise e interpretação dos resultados se dão através da atuação em conjunto do módulo de processamento de dados (Módulo II) e do módulo de apresentação dos resultados (Módulo III).

O Módulo II tem o objetivo de processar as informações que foram extraídas pelo Módulo I, conforme ilustrado na Figura 1. Ao todo, são realizados dois tipos de processamento no Módulo II: (i) via *Min-Max*; e (ii) via *k-Means*.

O Módulo II também é responsável por selecionar recomendações a partir de uma ontologia (detalhada em [Fontes 2017]). Para cada atributo analisado, faz-se uma consulta na ontologia para recuperar quais tipos de impactos tal atributo causa na participação efetiva dos alunos e para elaborar o texto da recomendação. Por fim, o Módulo II constrói um arquivo *xml* com o resultado de todos esses processamentos.

Já o Módulo III é responsável por consumir os resultados do *Web Service*, isto é, processar todas as informações do arquivo *xml* e apresentar aos coordenadores de curso, professores e aos próprios tutores a distância formas distintas de visualizar os desempenhos obtidos por cada tutor a distância em suas respectivas turmas.

O Módulo III foi construído seguindo as especificações de uma estrutura do Moodle conhecida como bloco, que nada mais é que um *plugin*. A Figura 2 ilustra o *plugin* integrado ao Moodle.



Figura 2. Tela inicial do *plugin* integrado ao Moodle

Conforme pode ser visto na Figura 2, no canto inferior esquerdo, existe uma imagem no formato de um velocímetro. O ponteiro do velocímetro e o texto em cima dele representam o desempenho geral obtido pelo tutor a distância analisado, obtido por meio da classificação, via *k-Means*, do atributo ‘desempenho’. Essa é uma forma rápida dos usuários saberem o desempenho geral obtido pelo tutor a distância. No entanto, para obter o relatório completo do desempenho do tutor a distância, o usuário deve clicar no *link* abaixo do velocímetro.

O relatório completo é composto por duas partes: (i) dois gráficos que mostram o comparativo entre o desempenho obtido pelo tutor a distância analisado em relação aos demais; e (ii) uma tabela com a lista de atributos e seus respectivos desempenhos e recomendações.

Inicialmente, o relatório apresenta a seção de gráficos, que é composta por dois gráficos: (i) valores relativos; e (ii) gráfico *k-Means*. A Figura 3(a) ilustra o gráfico de valores relativos e a Figura 3(b) ilustra o gráfico *k-Means*.

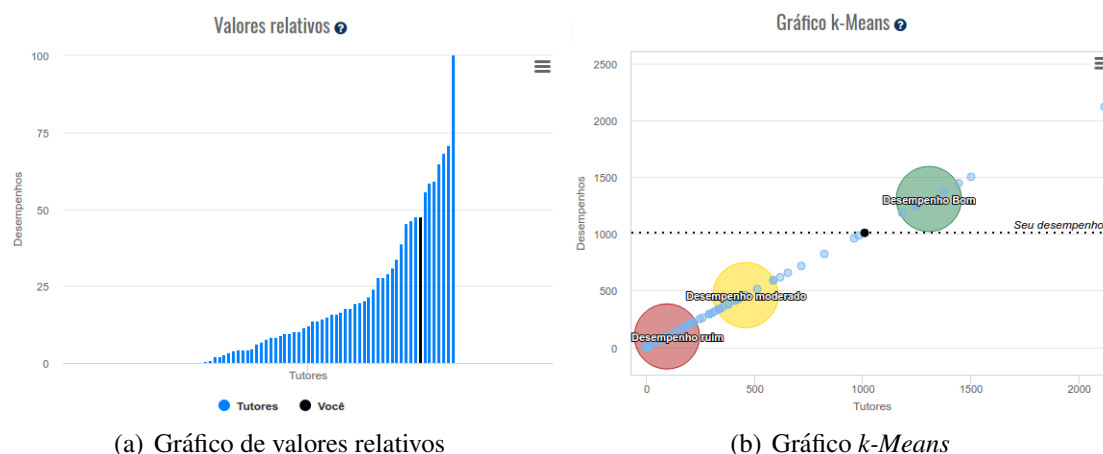


Figura 3. Seção de gráficos do relatório

O gráfico de valores relativos, ilustrado na Figura 3(a), é um gráfico de barras que apresenta um comparativo entre o desempenho do tutor a distância analisado em relação aos desempenhos dos demais tutores a distância. Nesse gráfico, os valores obtidos pelo atributo ‘desempenho’ de cada tutor a distância são normalizados numa escala de 0 a 100 antes de serem plotados no gráfico. Para isso, utiliza-se a técnica de normalização *Min-Max*.

Conforme pode ser visto na Figura 3(a), existe uma barra na cor preta. Essa barra representa o valor relativo obtido pelo atributo ‘desempenho’ do tutor a distância analisado. As outras barras representam os desempenhos dos demais tutores a distância. O usuário pode passar o *mouse* em cima de cada barra para saber o seu respectivo valor. O usuário também pode obter informações sobre o gráfico, clicando no ícone ao lado direito do título do gráfico.

O gráfico ilustrado na Figura 3(b) apresenta os desempenhos dos tutores a distância a partir de uma técnica de aprendizado não supervisionado. Neste caso, a classificação do atributo ‘desempenho’ é realizada por meio da técnica *k-Means*, utilizando-se do critério de distância euclidiana entre os centroides, tendo sido definido

$k = 3$, ou seja, foram criados três agrupamentos: (i) desempenho ruim; (ii) desempenho moderado; e (iii) desempenho bom.

Pode-se observar na Figura 3(b) que existem três agrupamentos, sendo o agrupamento de cor vermelha referente ao desempenho ruim, o agrupamento de cor amarela referente ao desempenho moderado e o agrupamento de cor verde referente ao desempenho bom. O ponto preto refere-se ao valor do atributo ‘desempenho’ do tutor a distância analisado. Também foi adicionada uma linha tracejada sob o ponto preto para indicar isso. Os demais pontos azuis referem-se aos valores do atributo ‘desempenho’ dos demais tutores a distância.

Ainda de acordo com a Figura 3(b), pode-se perceber que o ponto preto está mais próximo do centroide do agrupamento de desempenho bom. Desse modo, a indicação é de desempenho bom para o tutor a distância analisado, segundo a técnica *k-Means*.

Por fim, é exibido um relatório individual dos atributos analisados, exceto o atributo ‘desempenho’, já que esse atributo é calculado com base nos demais atributos e por ele já ter sido exibido por meio do velocímetro e dos gráficos. A Figura 4 ilustra um trecho do relatório individual.

Critério analisado	Desempenho	Recomendação
Total de arquivos criados	bom	O aumento de arquivos criados pelos tutores tende a: <ul style="list-style-type: none"> Aumentar moderadamente a participação dos alunos nos questionários.
Total de URLs criadas	ruim	O aumento de URLs criadas pelos tutores tende a: <ul style="list-style-type: none"> Aumentar fortemente a visualização dos alunos nas páginas; Aumentar moderadamente a participação dos alunos nos questionários; Aumentar moderadamente a submissão de tarefas por parte dos alunos; e Aumentar levemente a visualização dos alunos nos tópicos dos fóruns.
Total de cliques	bom	O aumento de cliques dos tutores na disciplina tende a: <ul style="list-style-type: none"> Aumentar moderadamente a participação dos alunos nos questionários; Aumentar moderadamente a visualização dos alunos nas páginas; e Aumentar levemente a submissão de tarefas por parte dos alunos.
Média de postagens nos chats	moderado	O aumento de postagens dos tutores nos chats tende a: <ul style="list-style-type: none"> Aumentar fortemente as postagens dos alunos nos chats.
Total de tarefas avaliadas	ruim	O aumento de tarefas avaliadas pelos tutores tende a: <ul style="list-style-type: none"> Aumentar moderadamente a participação dos alunos nos questionários; e Aumentar moderadamente a visualização dos alunos nas páginas.

Figura 4. Relatório individual dos atributos analisados

Conforme pode ser visto na Figura 4, o relatório possui três colunas: (i) critério analisado; (ii) desempenho; e (iii) recomendação.

A coluna ‘critério analisado’ lista os atributos analisados e as colunas ‘desempenho’ e ‘recomendação’ listam os desempenhos obtidos por cada atributo e as suas recomendações, respectivamente.

Para cada atributo, foi criada sua própria rede de classificação, onde os atributos são classificados por meio da técnica *k-Means*. Já as recomendações são geradas a partir da ontologia, detalhada em [Fontes 2017]. Essas recomendações são importantes para os tutores a distância saberem quais ações promovem a participação efetiva dos alunos.

Vale ressaltar que, em [Fontes 2017], são apresentadas as justificativas de escolha das técnicas utilizadas neste trabalho.

4. Dados Utilizados nos Experimentos

Os experimentos realizados neste trabalho utilizaram uma amostra. Os dados dessa amostra foram extraídos de uma base de dados histórica, correspondente aos anos de 2013 a 2016, cedida pela Secretaria de Educação a Distância (SEDIS) da Universidade Federal do Rio Grande do Norte (UFRN). Essa instituição usa o Moodle como AVA.

Os dados da amostra pertencem a 11 cursos de graduação, que são: (i) Bacharelado em Administração Pública; (ii) Licenciatura em Ciências Biológicas; (iii) Licenciatura em Educação Física; (iv) Licenciatura em Física; (v) Licenciatura em Geografia; (vi) Licenciatura em Letras; (vii) Licenciatura em Matemática; (viii) Licenciatura em Pedagogia; (ix) Licenciatura em Química; (x) Licenciatura em História; e (xi) Tecnólogo em Gestão Pública.

A amostra foi utilizada durante o desenvolvimento do Módulo I do sistema proposto (Subseção 3.1). Essa amostra foi composta pelas informações extraídas dos seguintes atributos dos tutores a distância: (i) número de questionários criados; (ii) número de tópicos criados nos fóruns; (iii) média de postagens em tópicos dos fóruns; (iv) taxa de visualização em fóruns; (v) taxa de visualização em tópicos dos fóruns; (vi) número de tarefas criadas; (vii) número de tarefas avaliadas; (viii) média de postagens em *chats*; (ix) total de cliques; (x) número de URLs criadas; (xi) número de arquivos criados; e (xii) desempenho.

No total, a amostra foi composta por informações de 353 tutores a distância, pertencentes a 1.281 turmas.

5. Considerações Finais e Trabalhos Futuros

Este trabalho possibilitou o desenvolvimento de um arcabouço tecnológico para monitoramento e avaliação das atividades de tutoria a distância em AVAs. Os resultados encontrados nas análises via técnica de agrupamento permitem uma inferência de desempenho comportamental de um tutor a distância em relação aos desempenhos dos demais tutores a distância.

Dessa forma, esses resultados respondem à questão de pesquisa deste trabalho, cuja conclusão é: **Sim, foi possível desenvolver um sistema capaz de monitorar e avaliar o desempenho de tutores a distância em AVAs.**

Como principais contribuições deste trabalho, tem-se: (i) construção de um *Web Service* capaz de processar as informações relacionadas aos atributos dos tutores a distância, classificar os seus desempenhos por meio da técnica *k-Means* e fazer recomendações com base em uma ontologia; e (ii) criação de um *plugin* para o Moodle que permite a visualização dos dados processados pelo *Web Service* e que serve como subsídio para os gestores dos referidos cursos.

Como possibilidades de trabalhos futuros, que podem ser explorados a partir desse estudo, tem-se: (i) definição dos pesos da Equação 3.1 por reforço de máquina; (ii) definição de estratégias pedagógicas associadas às recomendações que, além de informarem aos tutores a distância quais ações promovem a participação efetiva dos alunos, definam como esse objetivo pode ser alcançado; e (iii) realização de um estudo de caso para validar a eficácia do sistema proposto.

Referências

- Fontes, L. M. O. (2017). *MONITUM: Um Sistema Proativo para Monitoramento e Avaliação das Atividades de Tutoria a Distância em AVAs*. (Tese de Doutorado). Universidade Federal do Rio Grande do Norte.
- Hevner, A., March, S., Park, J., and Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1).
- Iseppon, D. B., Gomes, A. S., Rodrigues, R. L., Ramos, J. L. C., and Silva, J. C. S. (2016). Impact of teaching action on student interaction in virtual learning environments: canonical correlation analysis. In *Anais do XXVII Simpósio Brasileiro de Informática na Educação (SBIE)*.
- LeBlanc, D. (2004). *Statistics: Concepts and Applications for Science*. Jones & Bartlett Publishers.
- MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*.
- Marquesone, R. (2014). *Big Data: técnicas e tecnologias para extração de valor dos dados*. Casa do Código.
- Sharma, J. K. (2012). *Business Statistics*. Pearson Education India.
- Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Shum, S. B., Ferguson, R., Duval, E., Verbert, K., and d. Baker, R. S. J. (2011). *Open Learning Analytics: an integrated & modularized platform*. Proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques.
- Silva, W. R. M., Silva, J. L. D., Ugulino, F. R., Lucena, M., and Nunes, I. D. (2016). Learning Analytics como ferramenta para a análise do desempenho dos alunos em Cursos Semipresenciais. In *Anais do XXVII Simpósio Brasileiro de Informática na Educação (SBIE)*.
- Srilekshmi, M., S., S., Chatterjee, S., and Bijlani, K. (2016). Learning Analytics to Identify Students at-risk in MOOCs. In *Proceedings of the 8th International Conference on Technology for Education*. IEEE.
- Tabachnick, B. G. (2012). *Using Multivariate Statistics*. Pearson, 6 edition.
- Wieringa, R. (2014). *Design Science Methodology for Information Systems and Software Engineering*. Springer.