

Avaliação do Comportamento de Estudantes em um Ambiente Educacional Ubíquo

Juliete A. R. Costa^{1,2}, Rafael D. Araújo¹, Fabiano A. Dorça¹

¹Faculdade de Computação (FACOM)
Universidade Federal de Uberlândia (UFU)
Uberlândia, MG – Brasil

²Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas
(IFSULDEMINAS)
Carmo de Minas, MG – Brasil

juliete.costa@ifsuldeminas.edu.br, {rafael.araujo, fabianodor}@ufu.br

Abstract. *Computers have been increasingly mixed in daily life. In the educational context, these devices give rise to the so-called Ubiquitous Learning Environments that enhance the learning process in a more dynamic and engaging context. This kind of system generates valuable data that can be exploited by data mining techniques. Therefore, this work analyzes data from a Ubiquitous Learning Environment with the help of the data clustering technique in order to observe students' behavior in learning sessions. Results have shown statistically significant differences in the found clusters and evidence of Self-Regulated Learning in one of the groups.*

Resumo. *Computadores têm se misturado cada vez mais no cotidiano das pessoas. No contexto educacional, esses dispositivos dão origem aos chamados Ambientes Educacionais Ubíquos que potencializam o processo de aprendizagem em um contexto mais dinâmico e engajador. O uso destes sistemas origina dados valiosos que podem ser explorados por técnicas de mineração de dados. Diante disso, este trabalho analisa dados de um Ambiente Educacional Ubíquo com auxílio da técnica de agrupamento de dados a fim de observar o comportamento de estudantes em sessões de aprendizagem. Resultados mostraram diferenças estatisticamente significativas nos grupos encontrados e indícios de Aprendizado Autorregulado em um dos grupos.*

1. Introdução

Dispositivos computacionais e tecnológicos têm sido amplamente utilizados em muitas áreas de conhecimento. A cada dia, esses dispositivos se tornam mais integrados ao cotidiano das pessoas e auxiliam na condução das tarefas diárias. Esse é o conceito de Computação Ubíqua concebido por [Weiser 1991]. Especificamente no contexto educacional, computadores se misturam em salas de aula e são capazes de criar ambientes mais dinâmicos e engajadores e que dão suporte ao processo de ensino e aprendizagem.

Esses ambientes, chamados de Ambientes Educacionais Ubíquos (AEUs), possuem mecanismos que combinam atividades realizadas em espaços físicos e virtuais com o objetivo de produzir artefatos de estudo mais ricos que consideram o contexto em que os

estudantes estão inseridos e suas particularidades [Zhao e Okamoto 2011]. Nestes ambientes, a tarefa de identificar e manter as características particulares de cada indivíduo não é uma tarefa trivial e representa um desafio [Chrysaftadi e Virvou 2013]. Nesse contexto, técnicas de mineração de dados podem ser utilizadas como ferramentas para compreender a forma com que os estudantes aprendem e identificar comportamentos positivos e negativos que influenciam no processo de aprendizagem [Baker et al. 2011]. Assim, este trabalho tem como principal objetivo efetuar uma análise do comportamento dos estudantes em sessões de aprendizagem em um AEU utilizando técnicas de mineração de dados educacionais com intuito de melhor compreender como os estudantes utilizam o ambiente e o seu impacto na aprendizagem.

O restante do artigo está estruturado da seguinte forma: a Seção 2 apresenta conceitos importantes para o entendimento do trabalho; a Seção 3 traz uma discussão acerca de trabalhos relacionados; a Seção 4 apresenta a metodologia seguida; na Seção 5 são discutidos os resultados obtidos; e, por fim, a Seção 6 apresenta as considerações finais.

2. Referencial Teórico

2.1. Classroom eXperience

Classroom eXperience é um AEU que tem como base uma arquitetura fundamentada em conceitos de Captura & Acesso (C&A), com cinco fases propostas por [Abowd et al. 1997] e estendida por [Pimentel et al. 2001]: (i) pré-produção, focado na preparação do conteúdo; (ii) gravação “ao vivo”, para a captura das atividades executadas; (iii) pós-produção, responsável pela sincronização e armazenamento dos fluxos de mídia produzidos em diferentes formatos; (iv) acesso, para a disponibilização do conteúdo aos interessados; e, (v) extensão, quanto o conteúdo já capturado pode ser enriquecido pelos usuários, conforme observa-se em [Araújo et al. 2018]. A Figura 1 apresenta uma visão geral do processo.

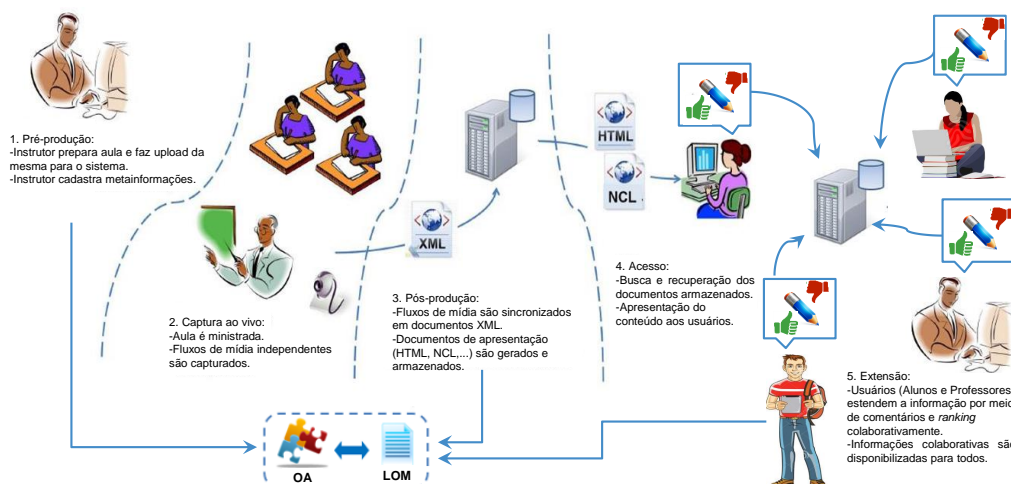


Figura 1. Visão geral das cinco fases do processo de C&A [Araújo et al. 2016].

Neste AEU, existem 23 tipos de interações que os usuários podem exercer sobre a plataforma durante o período de acesso ao conteúdo disponibilizado pelo professor. Essas interações podem ser agrupadas por características que representam funcionalidades do ambiente, por exemplo, acesso às aulas, atividades sociais e colaborativas, autoavaliação,

busca por conteúdo adicional e personalização de conteúdo, conforme descrito na Tabela 1. Além disso, cada acesso possui sete atributos de contexto, sendo quatro identificados automaticamente (data e hora do acesso, tipo de dispositivo utilizado, resolução da tela e largura de banda) e três informados manualmente (motivo do acesso, tempo disponível e local do acesso).

Tabela 1. Tipos de interação armazenados.

Categorias	Interações
Acesso às aulas	Login no ambiente; abertura e fechamento de aula; navegação de slides
Atividades sociais e colaborativas	Criação e exclusão de <i>bookmarks</i> ; <i>like</i> e <i>dislike</i> de <i>bookmarks</i> ; indicação de tipo de recurso de aprendizagem; classificação de slides; comentários e respostas aos comentários (tanto em slides quanto na disciplina)
Autoavaliação	Abertura e respostas aos <i>quizzes</i> ; alteração da visualização do modelo aberto do estudante
Busca por conteúdo adicional	Acesso a conteúdos adicionais recomendados
Personalização de conteúdo	Respostas ao instrumento de aferição de estilos de aprendizagem e alteração da visualização personalizada do conteúdo

De forma a ilustrar melhor as interações, a Figura 2 mostra uma imagem de uma aula já capturada pelo professor e disponibilizada aos estudantes por meio de uma interface Web. O item (1) indica o componente de navegação dos slides. Cada vez que o estudante troca de slide, a interação é armazenada. O item (2) aponta o componente de *bookmarks*, onde o estudante pode indicar o início de um novo conteúdo naquele conjunto de slides. Após criado, o *bookmark* passa a ser distribuído aos outros estudantes da turma, que podem gostar (*like*), não gostar (*dislike*) ou criar um novo. Somente o dono do *bookmark* pode excluí-lo. O item (3) mostra uma lista suspensa para indicar, colaborativamente, os tipos de recursos educacionais contidos no slide. O item (4) indica o local onde o estudante pode classificar o slide em estrelas (e a média geral é apresentada na estrela localizada no canto superior direito). O item (5) mostra o componente de comentários do slide e o item (6) mostra o componente de *quiz*, onde o estudante pode responder à questão de múltipla-escolha criada pelo professor e relacionado especificamente com aquele slide.

2.2. Mineração de Dados Educacionais

Com o grande crescimento do uso de sistemas educacionais como Ambientes Virtuais de Aprendizagem (AVAs) grandes quantidades de dados têm sido geradas, sejam a partir de interações dos usuários no sistema, fóruns de discussão ou avaliações, por exemplo. Diante deste cenário, surge a linha de pesquisa Mineração de Dados Educacionais (do inglês, *Educational Data Mining* - EDM) que tem como objetivo aplicar e/ou desenvolver métodos para analisar dados gerados em ambientes educacionais. Segundo [Baker et al. 2011], explorar tais dados pode ser importante para compreender de forma mais adequada os estudantes, como eles aprendem, o contexto na qual a aprendizagem ocorre, além de outros fatores que podem influenciar a aprendizagem. A descoberta de conhecimento a partir de dados educacionais possui várias etapas resumidas em pré-processamento, aplicação de técnicas de mineração e pós-processamento.

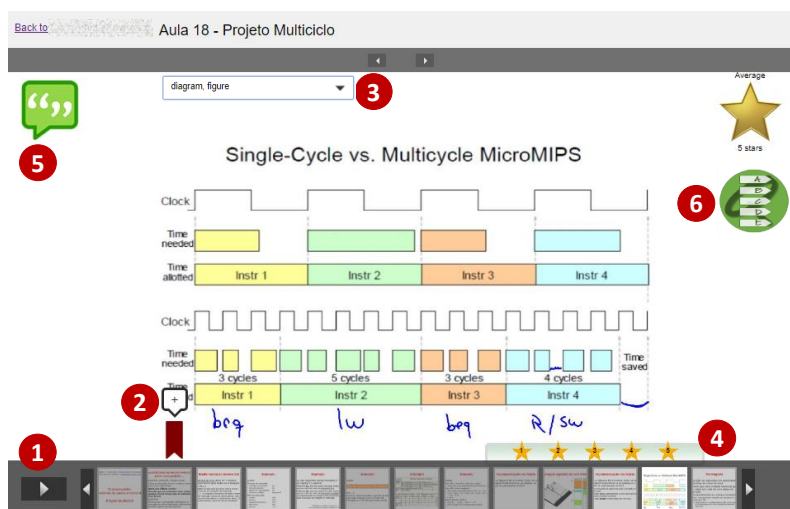


Figura 2. Captura da tela de uma aula disponibilizada no Classroom eXperience.

Uma das etapas mais importantes do processo de descoberta do conhecimento é pré-processamento de dados, que consiste em preparar os dados obtidos em um formato mais adequado para o processo de EDM. Em seguida, a etapa central do processo de descoberta do conhecimento consiste em aplicar técnicas de mineração nos dados pré-processados [García et al. 2011]. Nessa etapa, escolhe-se a técnica mais adequada aos dados pré-processados, como: classificação, regressão, regras de associação, agrupamentos, mineração de padrões sequenciais, mineração de texto, dentre outras. Por fim, a etapa final processo de EDM, e não menos importante, consiste em interpretar os resultados considerando o contexto em que os dados foram coletados a fim de ajudar no processo de tomada de decisão no AEU. A interpretação e análise dos resultados envolvem tanto especialistas da área educacional, quando testes estatísticos para avaliar a sua significância.

3. Trabalhos Relacionados

O trabalho [do Carmo et al. 2019], por exemplo, mostra uma estratégia baseada em mineração de padrões sequenciais para encontrar padrões de comportamento nos registros de *logs* de um sistema Adaptativo na Web. Com essa análise, os autores descobriram padrões de comportamentos semelhantes em turmas diferentes, os recursos mais utilizados, os tipos de caminhos de navegação mais preferidos e porcentagem de acesso aos recursos. No entanto, os autores destacam que não foi possível observar um padrão entre os recursos mais utilizados durante a trajetória do estudante e o desempenho final destes.

Em [El-Halees 2009], são utilizadas quatro técnicas de mineração de dados (regras de associação, classificação, agrupamento, detecção de *outliers*) sobre o mesmo conjunto de dados para analisar o comportamento dos estudantes e a relação com o desempenho final. Com a aplicação das técnicas, eles conseguem encontrar os atributos que influenciam os resultados, estudantes com comportamento que tendem a reprovação, grupos de estudantes com comportamento semelhantes e irregularidades nos comportamentos.

O trabalho de [Lallé e Conati 2020] utiliza duas técnicas de mineração de dados em sua estrutura orientada a dados (agrupamento e regras de associação) para analisar o comportamento de estudantes no contexto de visualização de vídeos em Curso Online Aberto e Massivo (MOOCs, do inglês *Massive Open Online Course*). Com a técnica de

agrupamento aplicada em dados de interações, os autores obtiveram grupos de comportamento dos estudantes mais e menos propícios ao aprendizado efetivo e, posteriormente, aplicam regras de associação em cada grupo para extrair os padrões comportamentais.

Já o trabalho apresentado por [Jie et al. 2017] destaca um experimento de análise de dados em uma Plataforma de Aprendizagem Autônoma. São coletados dados básicos dos estudantes, *logs* de pouso, navegação nos recursos da plataforma e experiência do aprendizado. A partir destes dados, os autores utilizam a técnica de classificação para encontrar perfis de comportamento nas aterrissagens feitas no curso e testes estatísticos para encontrar fatores que influenciam na navegação dos recursos disponíveis no sistema. A partir da análise destes trabalhos e, considerando seus aspectos particulares, foi possível elaborar uma metodologia para análise dos dados específicos associados a este trabalho, conforme apresentado na próxima seção.

4. Metodologia

O processo metodológico utilizado neste trabalho se divide em cinco etapas, destacadas na Figura 3. Em um primeiro momento, foram coletadas as interações do AEU e, em seguida, foram agrupadas em sessões de aprendizagem utilizando os *logs* de acesso. Em seguida, foi realizada uma análise exploratória dos dados com objetivo de selecionar os atributos do conjunto de dados que seriam utilizados na próxima etapa, de agrupamento de dados. Por fim, os resultados foram analisados de maneira a identificar o comportamento dos estudantes em sessões de aprendizagem e seu impacto.



Figura 3. Visão geral da metodologia do trabalho.

Coleta dos Dados: Foram coletados dados de acessos ocorridos durante os anos de 2017 e 2018 em 22 disciplinas ministradas nos semestres 2017/01, 2017/02 e 2018/01 em uma instituição pública brasileira de ensino superior. Utilizou-se consultas SQL para recuperar os 3.187 registros de acesso e 193.034 interações na plataforma. Cada interação é categorizada como um tipo específico a depender da ação do usuário, como apresentado na Subseção 2.1.

Agrupamento de sessões de aprendizagem: Após a coleta de dados, foi realizada uma agregação das interações por sessão de aprendizagem para compor o conjunto de dados. Uma sessão de aprendizagem representa um acesso do usuário no sistema considerando o intervalo de tempo entre o *login* e o *logout*, e, em cada acesso o usuário pode realizar vários tipos de interações. A navegação no sistema é livre, não sendo necessário visualizar todo o conteúdo dos slides, por exemplo. Diante deste contexto, foi realizado um agrupamento para obter a quantidade de tipos de interações distintas em cada sessão, como mostrado na Figura 4. Ao final desta etapa, obteve-se o conjunto de dados (*dataset*) no formato CSV (*Comma Separated Values*) com 3.187 registros, sendo cada registro uma sessão de aprendizagem com 19 atributos representando as interações (as outras 4 interações não tiveram nenhum dado no período selecionado) e 7 atributos relacionados ao contexto de acesso.

Análise Exploratória e Seleção dos Atributos: A partir do arquivo CSV gerado na etapa anterior, realizou-se a análise exploratória dos dados com utilização de recursos

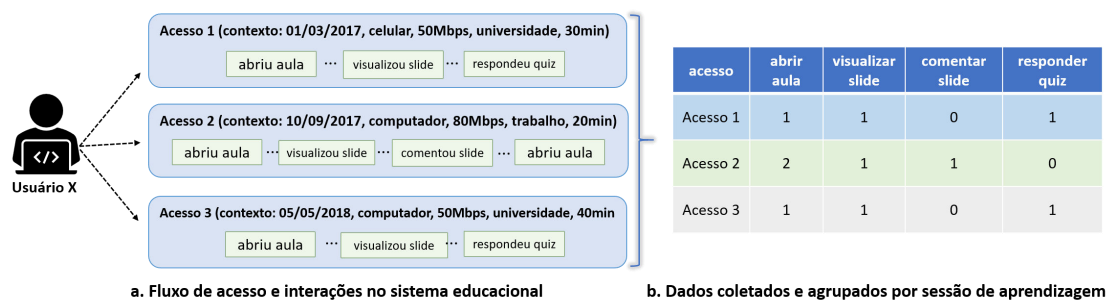


Figura 4. Exemplos de interações agrupadas em sessões de aprendizagem.

disponíveis na biblioteca *Pandas*¹, escrita na linguagem de programação Python² para trabalhar com manipulação e análise de dados. A análise exploratória, feita a partir da interface *Jupyter Notebook*³, foi utilizada para realizar uma descrição do conjunto de dados, verificar valores faltantes, bem como a correlação entre os atributos existentes. Com os recursos e métodos da biblioteca *Pandas*, foram calculados os coeficientes de correlação de Spearman [Urdan 2010] entre os atributos e, posteriormente, foram selecionados os atributos de interesse para a próxima etapa.

Foram removidos nesta etapa atributos representados por variáveis nominais com muitos valores faltantes, algumas interações com poucos acessos no período analisado e algumas interações relacionados a colaboração do estudante (conforme Tabela 1) foram agrupados em um único atributo. Dessa forma, obteve-se um conjunto de dados com 9 atributos: *device* (dispositivo utilizado na sessão - Celular ou Computador), *bandwidth* (largura de banda), *accesstime* (tempo de acesso da sessão em segundos), *lectureopen* (quantidade de aberturas de aula na sessão), *slidevisualization* (quantidade de visualizações de slides), *changechart* (quantidade de vezes que o estudante trocou a visualização do modelo aberto do estudante), *collaboration* (atributo agrupado que indica a quantidade de vezes que o estudante colaborou na sessão de aprendizagem), *quizanswer* (quantidade de vezes que o estudante respondeu ao *quiz*) e *cquizanswer* (quantidade de vezes que o estudante acertou o *quiz*). Como algumas sessões de aprendizagem (N=5) possuíam tempo de acesso muito grande (entre 6 e 11 horas) e poucas interações, entendeu-se que esses casos aconteceram por estudantes que acessaram o sistema e deixaram a sessão aberta indefinidamente. Assim, optou-se por atualizar o tempo dessas sessões para o valor médio das outras sessões.

Agrupamento de Dados: Nesta etapa foram utilizados recursos das bibliotecas *pandas* e *scikit-learn*⁴, para pré-processamento dos dados e aplicação do algoritmo, respectivamente. Para o agrupamento escolheu-se o algoritmo K-means [Aggarwal 2015] e para definir o número de *clusters* (*k*), utilizou-se o coeficiente de silhueta [Aggarwal 2015] calculado a partir da aplicação do K-means com diferentes valores de *k*.

Análise dos Resultados: Para analisar os resultados obtidos pela aplicação do K-means primeiramente observou-se as diferenças médias dos atributos em cada *cluster*, e para verificar se as diferenças são válidas ou não, usou-se testes de normalidade e significância

¹<https://pandas.pydata.org/>

²<https://www.python.org/>

³<https://jupyter.org/>

⁴<https://scikit-learn.org/stable/>

estatística disponíveis na biblioteca *SciPy*⁵ do Python. Antes da aplicação do teste de significância estatística, foi necessário analisar se os dados de cada *cluster* são provenientes de uma distribuição normal e, assim, definir qual teste de significância estatística utilizar (paramétrico ou não-paramétrico). Para essa análise, usou-se o teste de normalidade Kolmogorov-Smirnov [Razali et al. 2011], dado o tamanho do conjunto de dados. Além disso, foi executado o teste *U* de Mann Whitney [Urdan 2010] para analisar se existem diferenças significativas entre os atributos dos *clusters*.

5. Resultados e Discussão

A partir das tabelas geradas na etapa de coleta de dados foi realizado o agrupamento de sessões utilizando consultas SQL. Esta etapa consistiu em agrupar cada sessão com as respectivas quantidades de interações realizadas nesta. Ao final desta etapa, um arquivo no formato CSV foi gerado com 26 atributos representando as 3187 sessões do sistema. Este arquivo contém 7 atributos que destacam o contexto do acesso (razão por estar acessando o sistema, localização, dispositivo, largura de banda, tempo disponível, data e tempo de acesso) e os 19 tipos de interações utilizadas no período observado com os respectivos valores de quantidade para cada tipo.

Após obter o conjunto de dados, na etapa anterior, foi realizada uma análise exploratória dos dados para observar a porcentagem de valores vazios em cada atributo e a correlação entre os atributos. Para a análise de correlação utilizou-se coeficiente de Spearman. Com essa análise percebeu-se que muitos atributos nominais relacionados ao acesso possuíam correlação mediana e muitos registros com valores vazios e optou-se por removê-los do conjunto de dados. Além disso, alguns atributos referentes aos tipos de interação possuíam muitos valores zero, por exemplo as interações relacionadas a categoria de busca por conteúdo adicionais, representando que o recurso foi pouco utilizado durante o período observado, portanto optou-se também por removê-los. Os atributos que representam a categoria “Atividades Sociais e Colaborativas” foram agrupados em um único atributo denominado “collaboration”, pois analisando-os separadamente também se observou que possuíam muitos registros com valores nulos. Ao final dessa etapa, o conjunto de dados com 26 atributos foi reduzido para um arquivo com 9 atributos conforme amostra ilustrada pela Figura 5.

Amostra do conjunto de dados com 26 atributos

idcontextlog	reason	location	device	bandwidth	availabletime	accesstime	lectureopen	slidevisualization	...	Interação N
17703	ordinary study	university	desktop	1	121	36	1.0	2.0	...	0.0
16123	quick review	university	desktop	5	30	2988	2.0	1.0	...	1.0
12359			desktop	6	NULL	298	1.0	2.0	...	3.0

↓
Análise Exploratória e Seleção de Atributos

Amostra do conjunto de dados com 9 atributos

device	bandwidth	accesstime	lectureopen	slidevisualization	changechart	collaboration	quizanswer	cquizanswer
smartphone	77	2790	1.0	2.0	0.0	0	0.0	0.0
desktop	122	1811	1.0	114.0	6.0	148	2.0	2.0
desktop	1230	1638	1.0	20.0	4.0	0	10.0	3.0
smartphone	516	483	6.0	13.0	0.0	0	0.0	0.0
desktop	210	904	4.0	176.0	0.0	0	19.0	15.0

Figura 5. Amostra do conjunto de dados gerado a partir da Análise Exploratória e Seleção de Atributos.

Após a análise exploratória e redução da dimensionalidade do conjunto de dados,

⁵<https://www.scipy.org/>

aplicou-se o algoritmo de agrupamento K-means. Para obter número de *clusters* ideal, foi realizada uma análise do coeficiente de silhueta com 18 valores diferentes de k , sendo que o valor $k = 2$ obteve o melhor resultado ($S_i = 0.75$). A Tabela 2 apresenta dados da estatística descritiva contendo a média e o desvio padrão de cada atributo.

Tabela 2. Dados estatísticos de cada atributo por Cluster

Atributos	Cluster 0 (N=2760)	Cluster 1 (N=427)
	$\bar{x} \pm dp$	$\bar{x} \pm dp$
bandwidth	611.47 ± 484.61	582.44 ± 480.89
acesstime	756.54 ± 1104.02	8042.12 ± 3429.54
lectureopen	1.66 ± 1.92	3.59 ± 3.66
slidevisualization	41.96 ± 81.07	144.87 ± 168.03
changechart	0.1 ± 0.67	0.27 ± 1.1
collaboration	0.94 ± 15.59	1.33 ± 11.76
quizanswer	0.33 ± 2.54	1.32 ± 4.18
cquizanswer	0.22 ± 1.50	0.94 ± 2.99

\bar{x} : Média dos valores; dp : Desvio padrão.

A fim de comprovar se existem diferenças estatisticamente significantes entre os valores médios dos atributos dos *clusters*, foi executado o teste U de Mann Whitney, não-paramétrico, uma vez que os dados não são provenientes de uma distribuição normal, confirmado pelo teste Kolmogorov-Smirnov executado em cada atributo de cada *cluster* (com $p < .001$, rejeitando a hipótese nula). A Figura 6 ilustra um comparativo dessas médias e pode-se perceber que o Cluster 1 possui a maior média em todos os atributos, exceto a largura de banda (*bandwidth*), com $\chi^2 = 564935.5$, $p = .08$.

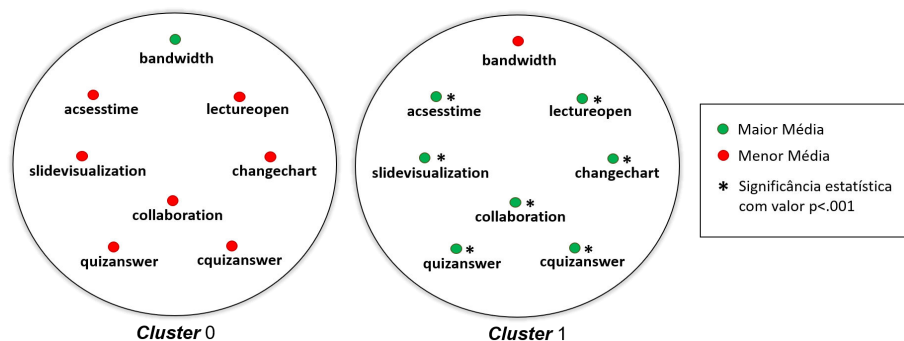


Figura 6. Comparação das médias dos atributos em cada Cluster com subsídio do teste estatístico U de Mann Whitney.

Os *clusters* ilustrados na Figura 6 destacam as diferenças entre os atributos, mostrando que as sessões de aprendizagem do **Cluster 1** apresentaram maiores médias de interações em diferentes recursos do sistema. As sessões com maior tempo de acesso e mais atividades de colaboração também apresentaram evidência de mais respostas corretas aos *quizzes*, o que pode ser um indício de que sessões dessa natureza representam um comportamento autorregulado dentro deste AEU. Segundo [Zimmerman 2008], a aprendizagem autorregulada (SRL, do inglês *Self-Regulated Learning*) consiste no quão os estudantes são proativos e responsáveis pelo próprio processo de aprendizagem, ou seja, os estudantes que participam mais do processo, com mais colaboração e interação, tendem a melhores resultados de desempenho, pois utilizam técnicas de autocontrole, definição de metas e estratégias para o sucesso da aprendizagem.

Neste contexto, é importante destacar que a atividade de *quiz* não é uma avaliação obrigatória para o estudante. [Kitsantas 2013] enfatiza que “estudantes autorregulados tendem a se autoavaliar com frequência e objetivamente usando dados auto-monitorados”, o que indica novamente evidências de comportamentos de estudantes autorregulados. [Kitsantas 2013] mostra, em seu trabalho, várias pesquisas realizadas com o intuito de fornecer diretrizes sobre como AVAs podem ser usados para apoiar a autorregulação dos estudantes, especificamente no ensino superior. Além disso, a autora destaca resultados de algumas pesquisas empíricas voltadas para adequação dos AVAs nesse processo, sugerindo que quando tais ambientes são projetados adequadamente, podem melhorar os processos de autorregulação dos estudantes. Sugerem, ainda, que estudantes tem maior probabilidade de se envolver com a aprendizagem autorregulada quando as tecnologias educacionais possuem ferramentas que apoiam esse processo, o que contribui para melhoria do processo de aprendizagem, conforme atesta [Zimmerman 1986]. Diante disso, pode-se perceber que este AEU possui evidências de processos de autorregulação dos estudantes comprovadas por técnicas de mineração de dados e testes de significância estatística aplicados nos *logs* de acesso e interações do sistema.

6. Considerações Finais

Neste trabalho foi possível observar a importância das técnicas de mineração de dados educacionais para análise do comportamento do estudante em sessões de aprendizagem de um AEU. Foi utilizado o algoritmo de agrupamento K-means e obteve-se dois grupos com diferenças estatisticamente significantes, comprovadas por testes estatísticos. Observou-se, ainda, a existência de indícios de Aprendizado Autorregulado nas sessões de aprendizagem em decorrência das interações ocorridas no AEU utilizado, o que é uma característica bastante positiva neste tipo de ambiente.

Como trabalhos futuros, pretende-se aplicar outros algoritmos de agrupamento, bem como outras técnicas de mineração de dados e testes estatísticos a fim de analisar o comportamento de estudantes, tanto em sessões de aprendizagem quanto em nível de curso. Ademais, deseja-se realizar um estudo mais aprofundado sobre características comportamentais nas sessões de aprendizagem que fomentam ou dificultam o aprendizado autorregulado nesse tipo de ambiente.

Agradecimentos

Os autores agradecem o apoio da Universidade Federal de Uberlândia (PPGCO/FACOM/UFU e PROPP/UFU) e do IFSULDEMINAS.

Referências

- Abowd, G. D., Atkeson, C. G., Feinstein, A., Hmelo, C., Kooper, R., Long, S., Sawhney, N., e Tani, M. (1997). Teaching and learning as multimedia authoring: The classroom 2000 project. In *Proceedings of the Fourth ACM International Conference on Multimedia*, página 187–198, New York, NY, USA. ACM.
- Aggarwal, C. C. (2015). *Data mining: the textbook*. Springer, 1 edition.
- Araújo, R. D., Brant-Ribeiro, T., Ferreira, H., Dorça, F., e Cattelan, R. (2016). Segmentação colaborativa de objetos de aprendizagem utilizando bookmarks em ambientes educacionais ubíquos. In *Anais do XXVII do Simpósio Brasileiro de Informática na Educação*, páginas 1205–1214. SBC.

- Araújo, R. D., Dorça, F. A., e Cattelan, R. G. (2018). A Computational Architecture for Learning Objects Authoring and Personalization in Ubiquitous Learning Environments. In *Anais dos Workshops do VII Congresso Brasileiro de Informática na Educação*, páginas 22–31. SBC.
- Baker, R., Isotani, S., e Carvalho, A. (2011). Mineração de dados educacionais: Oportunidades para o Brasil. *Revista Brasileira de Informática na Educação*, 19(02):03.
- Chrysafiadi, K. e Virvou, M. (2013). Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications*, 40(11):4715 – 4729.
- do Carmo, Ê. P., Gasparini, I., e Oliveira, E. (2019). Captura e visualização das trajetórias de aprendizagem como ferramentas para a análise do comportamento dos estudantes em um ambiente adaptativo educacional. In *Anais do XXX Simpósio Brasileiro de Informática na Educação*, páginas 309–318. SBC.
- El-Halees, A. M. (2009). Mining students data to analyze e-learning behavior: A case study. *Mining students data to analyze e-Learning behavior: A Case Study*, 29.
- García, E., Romero, C., Ventura, S., e De Castro, C. (2011). A collaborative educational association rule mining tool. *The Internet and Higher Education*, 14(2):77–88.
- Jie, W., Hai-yan, L., Biao, C., e Yuan, Z. (2017). Application of educational data mining on analysis of students' online learning behavior. In *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*, páginas 1011–1015. IEEE.
- Kitsantas, A. (2013). Fostering college students' self-regulated learning with learning technologies. *Hellenic Journal of Psychology*, 10(3):235–252.
- Lallé, S. e Conati, C. (2020). A data-driven student model to provide adaptive support during video watching across moocs. In *International Conference on Artificial Intelligence in Education*, páginas 282–295. Springer.
- Pimentel, M. d. G., Ishiguro, Y., Kerimbaev, B., Abowd, G., e Guzdial, M. (2001). Supporting educational activities through dynamic web interfaces. *Interacting with Computers*, 13(3):353–374.
- Razali, N. M., Wah, Y. B., et al. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics*, 2(1):21–33.
- Urdan, T. (2010). *Statistics in Plain English, Third Edition*. Taylor & Francis.
- Weiser, M. (1991). The Computer for the 21st Century. *Scientific American*, 265(3):66–75.
- Zhao, X. e Okamoto, T. (2011). Adaptive multimedia content delivery for context-aware u-learning. *International Journal of Mobile Learning and Organisation*, 5(1):46–63.
- Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology*, 11(4):307–313.
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American educational research journal*, 45(1):166–183.