

## Identificação e caracterização de níveis de interação no ensino remoto de emergência na Educação Básica

Aluisio José Pereira<sup>1</sup>, Alex Sandro Gomes<sup>1</sup>, Tiago Thompsen Primo<sup>2</sup>, Rosane Maria Alencar da Silva<sup>1</sup>, Rodrigo Lins Rodrigues<sup>3</sup>, Amadeu Sá de Campos Filho<sup>1</sup>, Ricardo Massa Ferreira Lima<sup>1</sup>, Ronaldo Pereira de Melo Júnior<sup>4</sup>

<sup>1</sup>Universidade Federal de Pernambuco (UFPE) – Recife-PE, Brasil.

<sup>2</sup>Universidade Federal de Pelotas (UFPEL) – Pelotas-RS, Brasil.

<sup>3</sup>Universidade Federal Rural de Pernambuco (UFRPE) – Recife-PE, Brasil.

<sup>4</sup>Colégio Militar do Recife (CMR) – Recife-PE, Brasil.

ajp3@cin.ufpe.br, asg@cin.ufpe.br, tiago.primo@inf.ufpel.edu.br,  
rmas3@cin.ufpe.br, rodrigo.linsrodrigues@ufrpe.br,  
amadeu.campos@nutes.ufpe.br, rmfl@cin.ufpe.br, ronaldo.melo@eb.mil.br

**Abstract.** *This article aims to capture evidence of the effectiveness of educational technology-mediated emergency remote learning about the interaction levels of elementary and high school students. Data from 963 students about the domain and use of the virtual environment Redu were collected and analyzed in the context of a public institution that adopted remote learning during the COVID-19 pandemic. An unsupervised approach was adopted, using the k-means algorithm for clustering students based on values of eleven variables voted to portray the level of student interaction. The results pointed to interaction patterns characterized by students who interact the most, interact sporadically or rarely.*

**Resumo.** *Este artigo objetiva capturar evidências da eficácia da aprendizagem remota de emergência mediada por tecnologia educacional em relação aos níveis de interação de estudantes do ensino fundamental e médio. Foram coletados e analisados dados de 963 estudantes a cerca do domínio e uso do ambiente virtual Redu no contexto de uma instituição pública que adotou aprendizagem remota durante a pandemia do COVID-19. Uma abordagem não-supervisionada foi adotada através do algoritmo k-means para clusterização dos estudantes com base em valores de onze variáveis votadas para retratar o nível de interação dos estudantes. Os resultados apontaram para padrões de interação caracterizados a partir dos estudantes que mais interagem, interagem esporadicamente ou raramente.*

### 1. Introdução

Com as restrições e medidas de distanciamento social, impostas pela pandemia do COVID-19, instituições adotaram o ensino remoto como estratégia emergencial de ensino (Bozkurt, 2020) (Yan *et al.*, 2021). Neste contexto de ensino remoto, a utilização de tecnologias digitais foi fundamental durante o período de fechamento de universidades e escolas (Subedi *et al.*, 2020). Usar tecnologias para aprimorar o ensino-aprendizagem é comumente conhecido como *e-learning* (Al-Arimi, 2014). Entretanto, o ensino remoto adotado de forma não planejada, emergencial e mediante transição súbita de aulas presenciais para remotas criou uma situação diferente do *e-learning* em circunstâncias normais (Affouneh; Salha; Khlaif, 2020) (Hodges; Moore; Lockee; Trust;

Bond, 2020). Dadas as condições de transição do ensino presencial para o remoto as interações entre estudantes, professores e conteúdos escolares podem ter sido afetadas, pois em uma estrutura de diálogo sem o fator distanciamento (presencial) as interações são diretas com grau de proximidade entre os sujeitos (Rumble, 2019).

A interação entre indivíduos, ambiente (natural, artificial ou imaginário) é fundamental no processo de ensino-aprendizagem. A privação social imposta pelo distanciamento físico, à redução das interações sociais face a face e o aumento do uso de meios digitais pode prejudicar o desenvolvimento e a saúde mental de adolescentes (Orben; Tomova; Blakemore, 2020). Entretanto, quando às interações ocorrem mediadas por tecnologias, surgem outras possibilidades a serem exploradas, por exemplo, a análise de dados educacionais (Laisa; Nunes, 2015) (Rabelo *et al.*, 2017), possibilidade de envolver muitos estudantes em torno de uma discussão, coletar dados que podem ser analisados para descoberta de padrões que forneça uma compreensão do engajamento (Macêdo; Santos; Maciel, 2020) (Moubayed *et al.*, 2020), e de como se constituem e quais as características formativas das interações ocorridas para construir modelos (Sari *et al.*, 2018), capturar padrões de aprendizagem e fornecer suporte adaptativo aos estudantes (Chanel *et al.*, 2016) (Strauss; Rummel, 2020).

Neste sentido, exploramos o uso do ambiente virtual Redu durante a sua adoção na pandemia provocada pelo COVID-19. E pressupomos que a compreensão dos níveis de interações que envolvem estudantes em contexto de ensino remoto de emergência possibilita ampliar o conhecimento sobre a experiência de ensino e aprendizado e inspirar a evolução de ambiente virtual adotados nessa modalidade de educação.

O artigo encontra-se estruturado em três outras seções. A Seção 2 trata-se do método com os processos e as abordagens quantitativas utilizadas. A Seção 3 os resultados encontrados. E na Seção 4 apresentamos as considerações finais deste estudo.

## **2. Método**

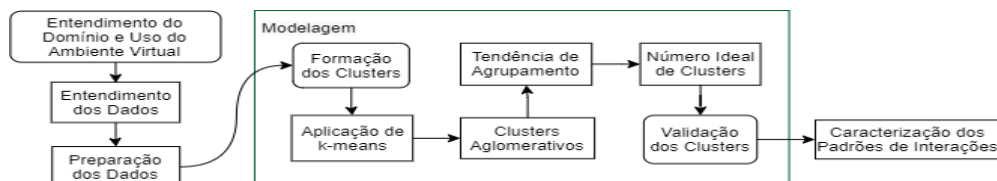
A partir de uma abordagem quantitativa, o estudo foi realizado em um contexto atípico provocada pela pandemia do COVID-19, no decorrer da qual, estudantes praticavam o ensino remoto emergencial que circunstancialmente permitiu coletar dados de uso do ambiente virtual durante oito meses em turmas do Ensino Fundamental (EF) e do Ensino Médio (EM), de uma Instituição Pública Federal de Ensino Básico, localizada na região metropolitana do Recife, Pernambuco, Brasil. A instituição foi adotada como local do estudo de caso por integralizar as atividades escolares mediadas por tecnologias digitais para atender o público de 963 estudantes matriculados no ano letivo de 2020.

### **2.1. Coleta e análise dos dados digitais: modelagem e clusterização**

Solicitou-se autorização para conduzir a pesquisa e ter acesso aos dados de uso do ambiente virtual. As coletas, tratamento e análise dos dados seguiram etapas do processo de *CRoss Industry Standard Process for Data Mining* (CRISP-DM) (Wirth e Hipp, 2000), conforme fluxo de atividades apresentado na Figura 1.

Para entendermos o domínio, analisamos o uso do ambiente virtual a partir da instância implantada na instituição de ensino e coletamos os dados que correspondem ao período de março a outubro de 2020. Foram tomadas medidas respeitando a Lei Geral de Proteção de Dados Pessoais (LGPD) (Brasil, 2019), para anonimidade dos dados, e nos casos em que se tratava de informações pessoais (nome, sobrenome, e-

*mail*, telefone) foram substituídos por “anônimo”, e quando sensíveis (*Login* e senha de acesso à plataforma) foram suprimidas totalmente da cópia da base de dados fornecida para condução da pesquisa.



**Figura 1. Processo de análise dos dados de uso do ambiente virtual**

No entendimento dos dados, um grupo de especialistas em Educação, Sociologia, Análise de Desempenho e Mineração de Dados Educacionais foi mobilizado para atuarem em consenso na escolha, dentre um conjunto de 36 variáveis coletadas do ambiente virtual Redu, as que melhor poderiam gerar evidências sobre o perfil de interação dos estudantes. O resultado do trabalho desta comissão permitiu definir (Tabela 1) as variáveis (VAR01 a VAR11) utilizadas como base deste estudo. Os valores relativos a elas foram consultados na base de dados do Redu por *scripts* SQL.

**Tabela 1. Variáveis selecionadas para condução das análises**

| Variáveis | Descrição sobre a variável  |
|-----------|---|
| VAR01     | Quantidade geral de amigos do estudante                                 |
| VAR02     | Quantidade de amigos do estudante que são (estudante)                   |
| VAR03     | Quantidade de colegas diferentes para quem o estudante enviou mensagens |
| VAR04     | Quantidade geral de mensagens enviadas pelo estudante                   |
| VAR05     | Quantidade de mensagem enviada (estudante-estudante)                    |
| VAR06     | Quantidade de mensagem enviada (estudante-professor)                    |
| VAR07     | Quantidade de exercícios realizados pelo estudante                      |
| VAR08     | Quantidade de pedidos de ajuda do estudante                             |
| VAR09     | Quantidade de comentários realizados pelo estudante                     |
| VAR10     | Quantidade geral de respostas às postagens recebidas pelo estudante     |
| VAR11     | Turno do dia que o estudante mais comentou                              |

Antes da aplicação do algoritmo de *Clustering* desagregaram-se os registros do conjunto de dados relativos aos estudantes do EF (504 registros) dos estudantes do EM (459 registros) para conduzir as análises separadamente. Foram avaliadas as tendências de agrupamento, conforme etapas definidas por Kassambara (2017), para confirmar se a aplicação de *Clustering* seria adequada para o conjunto de dados. A função *fviz\_nbclust()* no R permitiu estimar o número de *Clusters* ( $k = 4$ ) como sendo o número ideal de *Clusters* a partir dos dados do EF e EM. A formação dos *Clusters* ocorreu através da aplicação do algoritmo *k-means* (Hartigan-Wong, 1979). O *k-means* define a variação total dentro do *Cluster* como a soma das distâncias quadradas com as distâncias euclidianas entre itens e correspondentes centroides (Amatriain *et al.*, 2011).

Conduziu-se a normalização das variáveis através da função *preProcess()* opção *range* para escala entre 0 (zero) e 1 (um) para atribuir os dados normalizados (somente a variáveis numéricas VAR01 a VAR10) a uma variável *df*. O código no R descreve a associação realizada: *kmeans(df, x, iter.max = 10, nstart = 25)*, onde: *df*: representa o *data frame* das variáveis numéricas VAR01 a VAR10; *x*: o número de *Clusters* (*k*), (foi

usado  $x = 4$ ); *iter.max*: número máximo de iterações, (usou o valor padrão de *iter.max* = 10); *nstart*: números de partições aleatórias, (foi usado *nstart* = 25 para permitir buscar aleatoriamente 25 diferentes maneiras para iniciar as tarefas de selecionar os melhores resultados correspondentes).

A validação da formação dos *Clusters* permitiu avaliar os resultados da *Clustering*. A função *eclust()* no R: *eclust(df, FUNcluster = c("kmeans"), k = NULL, k.max = 10, nstart = 25, graph = TRUE, hc\_metric = "euclidean", hc\_method = "ward.D2")*, onde: *df*: representa o *data frame* das variáveis numéricas VAR01 a VAR10; *FUNcluster*: a função de agrupamento, no caso o "kmeans"; *k*: o número de *Clusters*, ao atribuir o valor *NULL* para *k*, o número de *Clusters* é estatisticamente estimado pela função para fornecer informações de silhueta para o particionamento *k-means*; *k.max*: número máximo de *Clusters* que por padrão é 10; *nstart*: números de partições aleatórias, especificou-se *nstart* = 25; *graph*: se *TRUE* exibe-se o gráfico; *hc\_metric*: métrica usada, no caso a distância euclidiana; *hc\_method*: método de aglomeração, no caso o *ward.D2*.

A variação dos valores e a correlação das variáveis foram calculadas. Matrizes de correlação foram usadas para representar o coeficiente de correlação de Pearson entre as variáveis calculados através da função *cor()* do pacote *ggcorrplot*, com escala de valores entre -1 até +1, possibilitou identificar quais variáveis se correlacionavam (positivo ou negativamente) ou inexistia correlação.

### 3. Resultados

A análise dos dados quantitativos oriundos da interação dos estudantes no ambiente virtual permitiu agrupar os estudantes por níveis de interação. Na Seção 3.1 temos os resultados da clusterização conduzidas a partir das análises dos valores coletados para as variáveis relativas aos 504 estudantes do EF e na Seção 3.2 para os 459 estudantes do EM. Os resultados da caracterização dos padrões (Seção 3.3) identificados ao unificar os resultados obtidos entre os níveis de interação dos estudantes do EF e EM.

#### 3.1. Níveis de interação dos estudantes do ensino fundamental

A formação dos quatro *Clusters* para os estudantes do Ensino Fundamental (EF) ocorreu de modo não-supervisionado e permitiu a distribuição os estudantes da seguinte maneira: 5 (1%) dos estudantes no *Cluster 1*, 164 (32,5%) dos estudantes no *Cluster 2*, 298 (59,1%) dos estudantes no *Cluster 3*, e 37 (7,3%) dos estudantes no *Cluster 4*. A Figura 2 apresenta os *boxplots* das variáveis (VAR01 a VAR10) com valores normalizados (escala de 0 a 1) para os quatro *Clusters* formados.

É possível observar nos gráficos (Figura 2) que a dispersão dos valores do conjunto de dados para os estudantes agrupados nos *Clusters* 1 e 4 apresentam amplitude interquartilica maiores que os dos *Clusters* 2 e 3. A Tabela 2 apresenta os valores de média e percentual das variáveis e para os quatro *Clusters*. É possível evidenciar que no *Cluster 1* a quantidade geral de mensagens enviadas pelo estudante (VAR04) possui média 524 (91,4%), as mensagens enviadas estudante-estudante (VAR05) possui média 516,6 (94,8%), e são os maiores valores em relação às demais variáveis observadas. Constatam-se comportamentos semelhantes em outros casos, no *Cluster 2* apenas o quantitativo de exercícios realizados pelo estudante (VAR07) apresenta média de 3,28 (61,9%) com o 2º quartil (mediana) acima de 0,50.

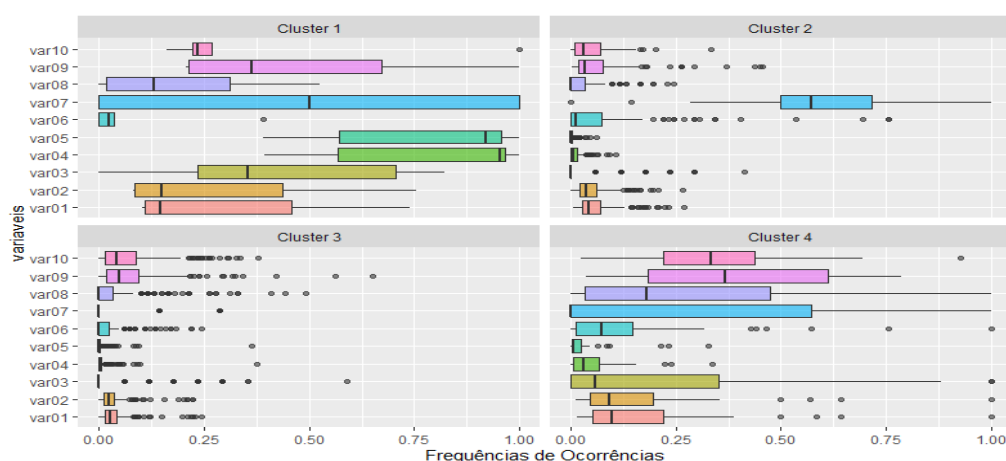


Figura 2. Frequência de ocorrência por variáveis para os Clusters EF

Todas as variáveis do *Cluster 3* apresentaram amplitude interquartílica abaixo de 0,25 e valores de média 10,22 e representa 4,4% do nível de interação. No *Cluster 4* a quantidade de comentários realizados pelos estudantes (VAR09) e a quantidade geral de respostas recebidas pelo estudante nas postagens realizadas (VAR10) são as que apresentam o 2º quartil (mediana) maiores para o grupo.

Tabela 2. Média das variáveis para cada grupo de estudantes do EF

| Variáveis                 | Cluster 1             | Cluster 2            | Cluster 3           | Cluster 4            | Média das variáveis |
|---------------------------|-----------------------|----------------------|---------------------|----------------------|---------------------|
| VAR01                     | 195,4 (53,2%)         | 36,40 (9,9%)         | 21,31 (5,8%)        | 113,86 (31,0%)       | <b>91,74</b>        |
| VAR02                     | 162,8 (54,7%)         | 27,29 (9,2%)         | 16,24 (5,5%)        | 91,54 (30,7%)        | <b>74,47</b>        |
| VAR03                     | 7,2 (60,2%)           | 0,52 (4,3%)          | 0,47 (3,9%)         | 3,78 (31,6%)         | <b>2,99</b>         |
| VAR04                     | 524 (91,4%)           | 7,66 (1,3%)          | 4,28 (0,7%)         | 37,43 (6,5%)         | <b>143,34</b>       |
| VAR05                     | 516,6 (94,8%)         | 2,04 (0,4%)          | 2,74 (0,5%)         | 23,73 (4,4%)         | <b>136,28</b>       |
| VAR06                     | 7,4 (27,2%)           | 5,46 (20,1%)         | 1,49 (5,5%)         | 12,86 (47,3%)        | <b>6,80</b>         |
| VAR07                     | 0,6 (11,3%)           | 3,28 (61,9%)         | 0,07 (1,3%)         | 1,35 (25,5%)         | <b>1,33</b>         |
| VAR08                     | 12 (37,0%)            | 1,48 (4,6%)          | 2,12 (6,5%)         | 16,81 (51,9%)        | <b>8,10</b>         |
| VAR09                     | 284,4 (48,9%)         | 38,69 (6,6%)         | 39,58 (6,8%)        | 219,24 (37,7%)       | <b>145,48</b>       |
| VAR10                     | 81,8 (45,2%)          | 10,10 (5,6%)         | 13,89 (7,7%)        | 75,03 (41,5%)        | <b>45,20</b>        |
| <b>Média dos Clusters</b> | <b>179,22 (52,4%)</b> | <b>13,29 (12,4%)</b> | <b>10,22 (4,4%)</b> | <b>59,56 (30,8%)</b> | <b>65,57</b>        |

No cálculo das correlações entre as variáveis (VAR01 a VAR10) para cada *Cluster*, conforme matrizes de correlação (Figura 3) resultante é possível observar que ocorrem correlações fortemente negativas, por exemplo, no *Cluster 1* entre a quantidade geral de mensagens enviadas (VAR04), a quantidade de mensagens enviada estudante-estudante (VAR05) correlacionadas com a quantidade de pedidos de ajuda (VAR08) e de comentários realizados pelo estudante (VAR09). Ou seja, para os estudantes deste *Cluster* ao postarem seus pedidos de ajuda ou realizarem comentários diretamente no ambiente virtual enviam menos mensagens diretas. Correlações equivalentes ocorrem também nos *Clusters 2* e *4*, ao correlacionar a realização dos exercícios (VAR07) os estudantes parecem enviar menos comentários (VAR09).

Por outro lado, é possível perceber correlação fortemente positiva nos quatro *Clusters* quando a quantidade de amigos do estudante (VAR01) é correlacionada à

quantidade de amigos do estudante que são estudantes (VAR02). Assim como, nos *Clusters* 1, 3 e 4, ao correlacionar a quantidade geral de mensagens enviadas pelo estudante (VAR04) com a quantidade de mensagens entre estudantes (VAR05). Para os estudantes do *Cluster* 1, quanto mais formulam pedidos de ajuda (VAR08), mais comentários são realizados (VAR09).

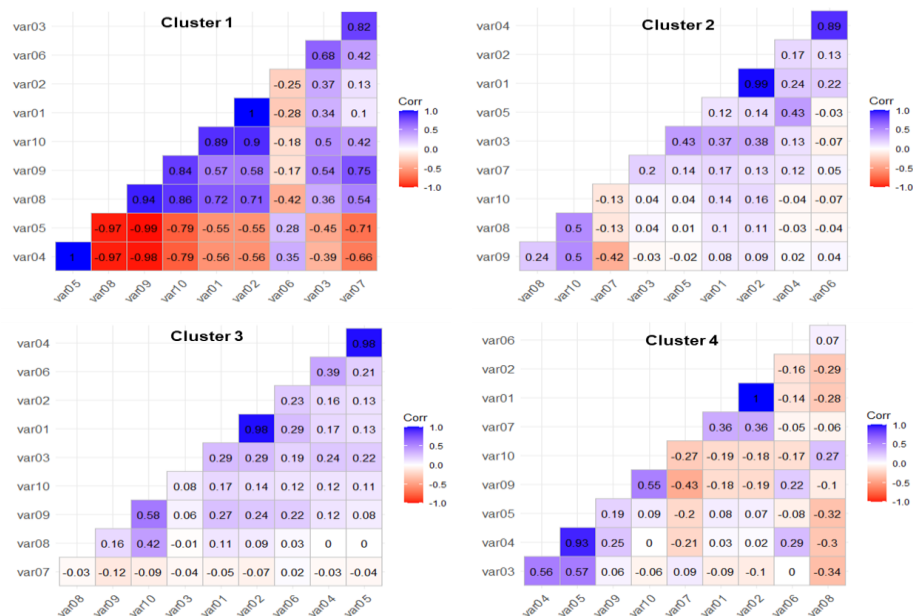


Figura 3. Matrizes de correlação para os *Clusters* do EF

Quanto maior a quantidade de amigos no ambiente virtual que têm os estudantes (VAR02), mais respostas ele recebe às suas postagens (VAR10). Já para os estudantes do *Cluster* 2, quanto mais enviam mensagens diretas (VAR04), mais aumenta a frequência de envio de mensagens estudante-professor (VAR06).

### 3.2. Níveis de interação dos estudantes do ensino médio

Foram formados quatro *Clusters* para os estudantes do Ensino Médio (EM). O *Cluster* 1 possui 11 (2,4%) dos estudantes, no *Cluster* 2 é representado por 20 (4,4%), o *Cluster* 3 são 194 (42,3%), e no *Cluster* 4 é formado por 234 (51%). A Figura 4 apresenta as *boxplots* a partir dos valores normalizados (escala de 0 a 1) para as variáveis (VAR01 a VAR10) dos quatro *Clusters* formados.

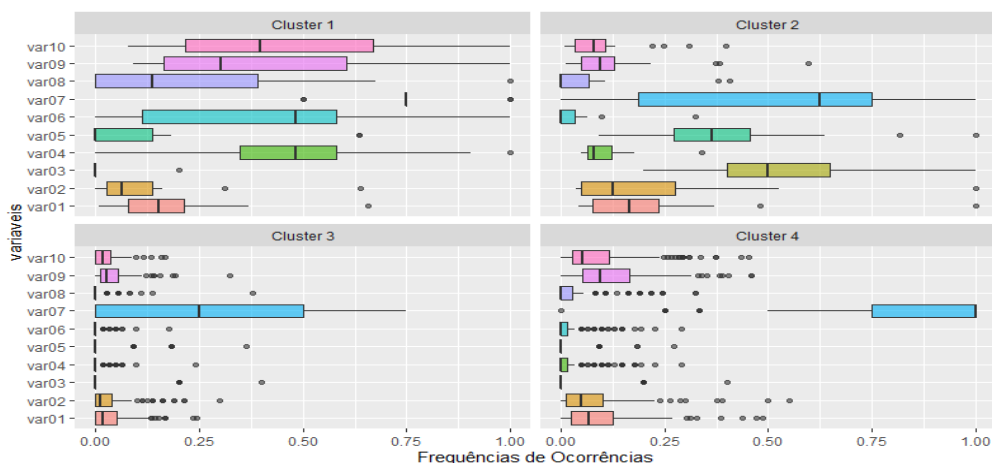


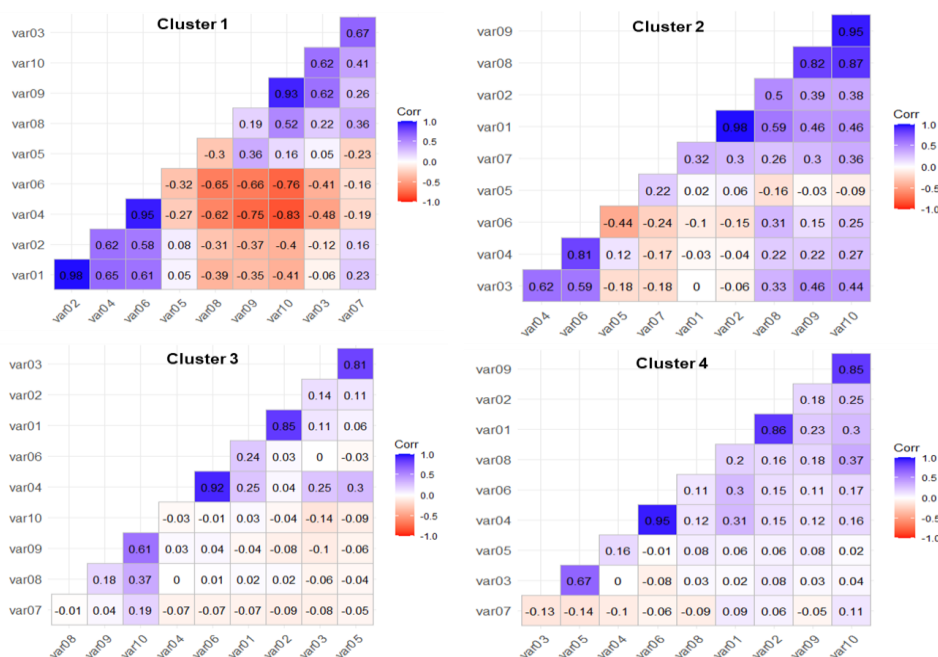
Figura 4. Frequência de ocorrência por variáveis para os *Clusters* do EM

Nota-se (Figura 4) que os valores de extremos e quartis de variáveis presentes nos *Clusters* 1 e 2 são maiores comparados aos valores das mesmas variáveis nos *Clusters* 3 e 4. Complementar a essas visualizações a Tabela 3 apresenta as médias dos valores das variáveis, onde é possível denotar que nos *Clusters* 1 e 2 respectivamente possuem média do *Cluster* de 23,13 (49,8%) e 9,91 (34,9%), sendo, portanto, os *Clusters* que possuem estudantes maiores níveis de interação. Em contrapartida, os *Clusters* 3 e 4 apresenta respectivamente médias de 1,85 (4,1%) e 5,41 (11,1%), isso revela que os estudantes destes *Clusters* possuem menores níveis de interação no ambiente virtual a partir das variáveis observadas.

**Tabela 3. Média das variáveis para cada grupo de estudantes do EM**

| Variáveis                 | Cluster 1            | Cluster 2           | Cluster 3          | Cluster 4           | Média das variáveis |
|---------------------------|----------------------|---------------------|--------------------|---------------------|---------------------|
| VAR01                     | 23 (36,2%)           | 25,4 (40,0%)        | 4,43 (7,0%)        | 10,63 (16,8%)       | <b>15,87</b>        |
| VAR02                     | 10,45 (30,0%)        | 16,4 (47,0%)        | 2,39 (6,9%)        | 5,62 (16,1%)        | <b>8,72</b>         |
| VAR03                     | 0,09 (3,0%)          | 2,8 (92,7%)         | 0,08 (2,6%)        | 0,05 (1,7%)         | <b>0,76</b>         |
| VAR04                     | 29,73 (77,9%)        | 6,7 (17,6%)         | 0,43 (101%)        | 1,28 (3,4%)         | <b>9,54</b>         |
| VAR05                     | 1,55 (25,4%)         | 4,35 (71,2%)        | 0,12 (2,0%)        | 0,09 (1,5%)         | <b>1,52</b>         |
| VAR06                     | 26,18 (88,7%)        | 2,0 (6,8%)          | 0,27 (0,9%)        | 1,08 (3,7%)         | <b>7,38</b>         |
| VAR07                     | 2,91 (37,7%)         | 1,7 (22,0%)         | 0,74 (9,6%)        | 2,36 (30,6%)        | <b>1,93</b>         |
| VAR08                     | 9,73 (73,7%)         | 2,2 (16,7%)         | 0,25 (1,9%)        | 1,02 (7,7%)         | <b>3,30</b>         |
| VAR09                     | 77,55 (58,5%)        | 25,65 (79,5%)       | 7,14 (5,4%)        | 22,18 (16,7%)       | <b>33,13</b>        |
| VAR10                     | 50,09 (67,3%)        | 11,85 (15,9%)       | 2,64 (3,5%)        | 9,81 (13,2%)        | <b>18,60</b>        |
| <b>Média dos Clusters</b> | <b>23,13 (49,8%)</b> | <b>9,91 (34,9%)</b> | <b>1,85 (4,1%)</b> | <b>5,41 (11,1%)</b> | <b>10,07</b>        |

As matrizes de correlação entre as variáveis nos quatro *Clusters* (Figura 5) revelam quais variáveis encontram-se correlacionadas fortemente negativamente, a inexistência de correlação, ou as que se correlacionam fortemente positivamente.



**Figura 5. Matrizes de correlação para os Clusters do EM**

Denotam-se correlações fortemente negativas no *Cluster* 1 entre a quantidade geral de mensagens enviadas pelos estudantes (VAR04), a quantidade de mensagens enviadas pelo estudante-professor (VAR06) correlacionadas com a quantidade de respostas obtidas nas postagens no ambiente virtual (VAR10). Revelando que os estudantes deste *Cluster* ao enviarem mensagens diretas aos professores diminui a necessidade de obterem respostas às postagens, pedidos de ajuda e comentários no ambiente virtual. Algo semelhante pode ser percebido no *Cluster* 2 ao correlacionar a quantidade de mensagens estudante-estudante (VAR05) e o envio de mensagens estudante-professor (VAR06). Revelando que ao ser enviado mais mensagens direcionadas aos estudantes reduz o envio direcionado aos professores, e vice-versa.

Percebe-se inexistência de correlação entre a quantidade de colegas diferentes para quem o estudante enviou mensagens (VAR03), correlacionada com: a quantidade de amigos do estudante (VAR01) no *Cluster* 2; a quantidade de mensagem enviada estudante-professor (VAR06) no *Cluster* 3; e com a quantidade geral de mensagens enviadas (VAR04) no *Cluster* 4. Algo parecido ocorre no *Cluster* 3 entre a quantidade geral de mensagens enviadas (VAR04) e os pedidos de ajuda do estudante (VAR08). Isto sugere que o envio de mensagens não interfere nos pedidos de ajuda do estudante para os estudantes presentes no *Cluster* 3.

Para os estudantes nos quatro *Clusters*, a quantidade geral de amigos (VAR01) e a quantidade de amigos que são estudantes (VAR02) estão positivamente correlacionadas. Isso nos permite inferir que essas variáveis são complementares. À medida que alguém tem mais amigos, a quantidade geral de amigos aumenta proporcionalmente. Correlações semelhantes estão presentes nos *Clusters* 1, 3 e 4 em relação à quantidade de mensagens enviadas (VAR04) e as mensagens enviadas estudante-professor (VAR06). Perceba também que quanto mais comentários são postados (VAR09), mais respostas as postagens são recebidas (VAR10) nos *Clusters* 1, 2 e 4. No *Cluster* 2, quanto mais os estudantes realizam pedidos de ajuda (VAR08) maior a quantidade de respostas às suas postagens no ambiente virtual (VAR10). Para os estudantes do *Cluster* 3 quanto mais mensagens são enviadas estudante-estudante (VAR05) proporcionalmente se tem mais colegas diferentes para quem o estudante enviou mensagem (VAR03).

### 3.3. Caracterização das interações dos estudantes do EF e EM

A partir das análises das interações dos estudantes Ensino Fundamental (EF) e Ensino Médio (EM) no ambiente virtual Redu, caracterizaram-se as interações dos estudantes em três níveis: (i) mais interação (representados pelos *Clusters* 1 e 4 do EF e os *Clusters* 1 e 2 do EM), (ii) interação esporadicamente (pelo *Cluster* 2 do EF e o *Cluster* 4 do EM), e, (iii) raramente interação (pelo *Cluster* 3 do EF e o *Cluster* 3 do EM). A Tabela 4 apresenta a caracterização realizada e uma descrição dos perfis do público de estudantes presente em cada nível de interação.

E assim, ao correlacionar de forma unificados os dados do EF e EM por caracterização e padrões de interação (Figura 6) denotamos algumas variáveis (VAR01 e VAR02; VAR03 e VAR05; VAR08, VAR09 e VAR10) estão sempre positivamente correlacionadas nos diferentes níveis de interação (mais, esporadicamente, ou raramente interação). Esses resultados nos permitem perceber que essas variáveis podem ser complementares ou juntas são sempre nas interações dos estudantes.

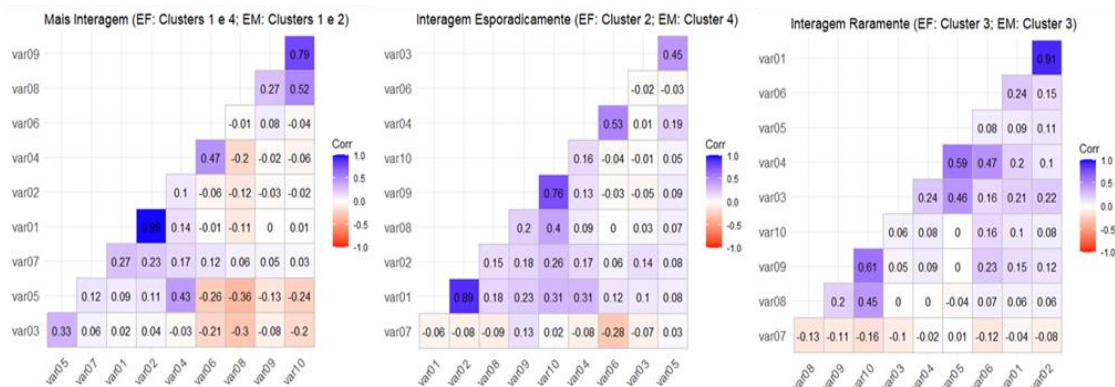


**Tabela 4. Padrões de interação dos estudantes do EF e EM**

| Caracterização das Interações | Descrição dos Perfis de Estudantes  |
|-------------------------------|---|
| Mais Interagem                | Os estudantes presentes neste grupo foram os que mais interagiram dentre os demais. Possuem números maiores de amigos no ambiente virtual, apresentam maiores quantitativos no envio de mensagens para colegas e professores, possuem mais interações via fóruns de discussão ao pedirem ajuda e realizarem comentários. Recebem mais respostas a postagens realizadas. Pode-se destacar que este grupo possui estudantes ativamente mais interativos no ambiente virtual, representam percentual de 83,2% das interações ocorridas entre os estudantes do EF e de 84,7% no EM. Porém, o percentual de estudantes que representam esse grupo é baixo, 42 (8,3%) no EF, e 31 (6,8%) no EM. |
| Interagem Esporadicamente     | Os estudantes deste grupo possuem como características em comum serem eventuais nas interações mediadas pelo ambiente virtual. A partir das correlações das variáveis observadas, evidencia-se pouco envio de mensagens para colegas e professores. No entanto, se descartam pela realização dos exercícios. Concentram percentuais de interações de 12,4% para os estudantes do EF e 11,1% para os do EM. O número de estudantes neste grupo é bem representativo, 113 (32,5%) no EF, e 234 (51,0%) no EM.   |
| Interagem raramente           | Este grupo foi considerado com o pior nível de interação dentre os demais, dado que, os estudantes presentes neste grupo além de apresentarem poucos amigos, enviarem poucas mensagens possuem valores baixos para as variáveis observadas. Possuem níveis de interação baixos: 4,4% no EF e de 4,1% no EM. Entretanto, são bem representativo 298 (59,1%) no EF, e 194 (42,3%) no EM.  |

Nota: EF: Ensino Fundamental; EM: Ensino Médio.

Para os estudantes que interagem esporadicamente, o envio de mensagens (VAR04) apresenta-se mais fortemente correlacionada com a quantidade de mensagens enviadas pelo estudante-professor (VAR06). Diferentemente do que ocorre para os estudantes que mais interagem, ou interagem raramente, visto que neste nível o envio de mensagem (VAR04) aumenta proporcionalmente tanto em relação à quantidade de mensagens enviadas estudante-estudante (VAR05), quanto estudante-professor (VAR06). Isso nos revela que dentre os estudantes que interagem esporadicamente, à medida que envia mensagens, busca contato mais com os professores do que com colegas.



**Figura 6. Correlação das variáveis para os níveis de interação**

Para os estudantes que mais interagem, a quantidade de pedidos de ajuda (VAR08) correlaciona-se negativamente com a quantidade de colegas diferentes para

quem o estudante enviou mensagens (VAR03) e com a quantidade de mensagem enviada estudante-estudante (VAR05). O que nos permite compreender que quanto mais os estudantes que mais interagem pedem ajuda publicamente no ambiente virtual (VAR08), menos enviam mensagens diretas via *chats* aos colegas e professores. Isso reduz consequentemente a quantidade geral de mensagens enviadas. O processo inverso também ocorre, neste caso, ao aumentar a quantidade de mensagens diretas destinadas aos colegas ou professores, os estudantes diminuem ou dispensam interações colaborativas promovidas coletivamente entre os estudantes matriculados em um mesmo curso.

Percebe-se também que, para os estudantes que mais interagem, a quantidade de mensagem enviada à estudante-estudante (VAR05) possui correlação negativa com a quantidade de mensagem enviada estudante-professor (VAR06). Isso nos permite confirmar que os estudantes deste nível de interação se deparam com uma dubiedade. Ao enviarem mais mensagens a seus colegas, enviam menos mensagens aos professores, e vice-versa. Para os estudantes que interagem esporadicamente, quando aumenta a quantidade de mensagens enviadas ao professor (VAR06), diminui a quantidade de exercícios realizados pelo estudante (VAR07). Ao compreender que o processo inverso também pode ser possível, a não realização dos exercícios leva a um número maior de mensagens direcionadas aos professores. Porém, essas análises não nos esclarecem e não nos permitem inferir quais os motivos dessa correlação inversamente proporcional.

#### 4. Considerações finais

Este estudo objetivou reunir evidências sobre a eficácia do aprendizado remoto emergencial mediado por tecnologia educacional em relação aos níveis de interação dos estudantes do Ensino Fundamental e Médio de uma instituição pública que adotou o ensino remoto de emergência mediado por ambiente virtual de rede social educativa. Com a análise dos dados de uso do ambiente virtual Redu, agrupamos os estudantes a partir dos níveis de interação e concluímos que os estudantes que mais interagem, recorrem aos diversos recursos disponíveis e vivenciam interações com colegas e professores. Os que interagem espontaneamente centram suas interações na finalidade de realizar as atividades e exercícios. Já os estudantes que interagem raramente apresentam comportamentos difíceis de serem compreendidos a partir dos das análises conduzidas, pois apresentam baixos níveis de interação no ambiente virtual.

Para dimensionar melhor as estratégias que os estudantes conduziram através e as que vão além (externo) dos limites dos dados coletados, se faz necessário conduzir análises qualitativas (entrevistas) para coletar evidências da percepção dos estudantes sobre o ensino remoto mediado pelo ambiente virtual.

#### Referências

- Affounh, S., Salha, S., & Khlaif, Z. N. (2020). Designing quality e-learning environments for emergency remote teaching in coronavirus crisis. *Interdisciplinary Journal of Virtual Learning in Medical Sciences*, 11(2), 135-137.
- Al-Arimi, A. M. A. K. (2014). Distance learning. *Procedia-Social and Behavioral Sciences*, 152, 82-88.

- Amatriain, X., Jaimes, A., Oliver, N., & Pujol, J. M. (2011). Data mining methods for recommender systems. In *Recommender systems handbook* (pp. 39-71). Springer, Boston, MA.
- Brasil (2019). Conselho Nacional. “Lei nº 13.853, de 8 de julho de 2019” [http://www.planalto.gov.br/ccivil\\_03/\\_Ato2019-2022/2019/Lei/L13853.htm](http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2019/Lei/L13853.htm), Outubro.
- Bozkurt, Aras (2020). A global outlook to the interruption of education due to COVID-19 pandemic: Navigating in a time of uncertainty and crisis. *Asian Journal of Distance Education*, 1-126.
- Chanel, G., Lalanne, D., Lavoué, E., Lund, K., Molinari, G., Ringeval, F., & Weinberger, A. (2016). Grand Challenge Problem 2: Adaptive awareness for social regulation of emotions in online collaborative learning environments. In *Grand challenge problems in technology-enhanced learning II: MOOCs and beyond* (pp. 13-16). Springer, Cham.
- Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, 28(1), 100-108.
- Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching online learning. *Educause review*, 27(1), 1-9.
- Kassambara, A. (2017). *Practical guide to cluster analysis in R: Unsupervised machine learning* (Vol. 1). Sthda.
- Laisa, J., & Nunes, I. (2015, October). Mineração de Dados Educacionais como apoio para a classificação de alunos do Ensino Médio. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)* (Vol. 26, No. 1, p. 1112).
- Macêdo, P. H., Santos, W. B., & Maciel, A. M. (2020). Análise de Perfis de Engajamento de Estudantes de Ensino a Distância. *RENOTE*, 18(2), 326-335.
- Maschio, P., Vieira, M. A., Costa, N., de Melo, S., & Júnior, C. P. (2018, October). Um panorama acerca da mineração de dados educacionais no Brasil. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)* (Vol. 29, No. 1, p. 1936).
- Moubayed, A., Injadat, M., Shami, A., & Lutfiyya, H. (2020). Student engagement level in an e-learning environment: Clustering using k-means. *American Journal of Distance Education*, 34(2), 137-156.
- Orben, A., Tomova, L., & Blakemore, S. J. (2020). The effects of social deprivation on adolescent development and mental health. *The Lancet Child & Adolescent Health*.
- Rabelo, H., Burlamaqui, A., Valentim, R., de Souza Rabelo, D. S., & Medeiros, S. (2017, October). Utilização de técnicas de mineração de dados educacionais para predição de desempenho de alunos de EaD em ambientes virtuais de aprendizagem. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)* (Vol. 28, No. 1, p. 1527).
- Rumble, G. (2019). *The planning and management of distance education*. Routledge.

- Sari, I., Maselena, A., Satria, F., & Muslihudin, M. (2018). Application model of k-means clustering: insights into promotion strategy of vocational high school. *International Journal of Engineering & Technology*, 7(2.27), 182-187.
- Strauss, S., & Rummel, N. (2020). Promoting interaction in online distance education: Designing, implementing and supporting collaborative learning. *Information and Learning Sciences*, 121(5), 251–260.
- Subedi, S., Nayaju, S., Subedi, S., Shah, S. K., Shah, J. M. (2020). Impact of e-learning during COVID-19 pandemic among nursing students and teachers of Nepal. *International Journal of Science and Healthcare Research*, 5(3), 9.
- Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining* (Vol. 1). London, UK: Springer-Verlag.
- Yan, L., Whitelock-Wainwright, A., Guan, Q., Wen, G., Gašević, D., & Chen, G. (2021). Students' experience of online learning during the COVID-19 pandemic: A province-wide survey study. *British Journal of Educational Technology*.