Guidelines for Google Classroom usage as an e-learning tool during Covid-19 pandemic based on similarity search

Danielli A. Lima¹,², Seiji Isotani¹,³

¹Instituto Federal do Triângulo Mineiro (IFTM) Campus Patrocínio, MG, Brasil
Laboratório de Inteligência Computacional e Robótica (LICRO)

²Universidade de São Paulo (USP) Campus São Carlos, SP, Brasil
Instituto de Ciências da Computação e Matemática Computacional (ICMC)

³Harvard University (HU) Cambridge, MA, United States
Harvard Graduate School of Education (HGSE)

danielli@iftm.edu.br, sisotani@icmc.usp.br

Abstract. The coronavirus pandemic has changed the routine of almost every school around the world. To solve the problems caused by this disease, many schools adapted and computerized their activities so that the teaching and learning process would not be harmed. Thus, learning management systems were essential, and an example of this type of tool is the Google Classroom, which was widely used in several countries around the world. This article aims to make a similarity search of scientific studies mined by a systematic review to discover guidelines for the implementation of Google Classroom during the pandemic. This study can contribute to the e-learnings implementation not only in the pandemic, but as a support for hybrid and distance classes.

1. Introduction

The Covid-19 pandemic brought a lot of insecurity to the education sector. However, over time, schools sought to find alternatives to overcome the obstacles caused by the absence of in-person classes [Ivenicki 2021]. To get around this problem, some schools worked with the methodology of tutored study plans, asynchronously, while other schools sought to implement some Information and Communications Technology (ICT) tools [Fuady et al. 2021], linked to Learning Management Systems (LMS) in an attempt to keep, whenever possible, synchronous classes, such as, Moodle, Edulogy, Microsoft Teams or Google Classroom [Alves and Lima 2018, Alves et al. 2021].

In this work, we will focus on Google Classroom (GC) to support distance education [Lima et al. 2017b] in times of pandemic, because it is a widely used tool not only in Brazil, but also around the world [A’yun et al. 2021, Baskoro et al. 2020, Espirito Santo and Lima 2020, Yen and Mohamad 2021]. In addition, it has the advantages of being a tool integrated with several Google services [Lima et al. 2017b], and also has a good usability, simplified navigation by hyperlinks, security and privacy, portable file posting and is also considered a m-learning [Lima et al. 2021a, Lima et al. 2021b].

Based on this context, we used a similarity search (SS) to find, among the twenty-five articles presented in [Lima and Isotani 2022] SLR (Systematic Literature Review), the nearest neighbors among the articles, to find the most similar papers to
each other. From the results of this a graph can be created, and then it is possible to group the connected components into clusters. These connected components represent groups with the most similar papers, one nearest neighbor. The main objective of this article is find guidelines, clusters-based, that will be able to help teachers, stakeholders and even government officials to propose a better way of implementing GC as a support for distance or hybrid learning in several institutions [Alves and Lima 2018, Lima et al. 2017b]. In addition, the guidelines presented herein can be expanded to a universal application for the implementation of other e-learning and m-learning tools, considering that the concepts used here are multipurpose to provide quality to distance or hybrid learning.

2. Background

In [Lima and Isotani 2021] work an SLR was carried out in order to find papers from primary studies on the use of Google Classroom during the Covid-19 pandemic. Data were mined from four databases: Web of Science, Scopus, IEEE and Google Scholar and organized on the StArt platform for SLR. From this scientific mining database, 93 studies were found, and one of them was a duplicate study. We kept, for each paper, its original ID defined in the StArt tool. After reading the abstracts of the articles and after applying the inclusion-exclusion criteria, 63 articles were excluded, leaving 30 selected articles.

After the complete reading of each of these articles, 25 articles were extracted for data summary and analysis. The GC tool was chosen precisely because it was used in several countries and also studied in many of them, as shown in Figure 1 in which it is possible to observe the frequency by country of the 25 papers extracted in the work of [Lima and Isotani 2022]. For the presentation of the results a normalization was performed following the formula $2^{\log n}$, where $n$ is the frequency of the countries found in the publication for the first author.

To understand the results showed in Figure 2 we created a vector including: Paper ID, publication Year and 12 characteristics: Selection Score, Paper Type, Analysis, Tools, Level, Data, Results, Audience, Overview, Priority Selection, Extraction Score, Priority Extraction. First

Figure 1. Map with normalized data that represents the country of origin of the first author of each of the 25 papers extracted for the study of this work.
we present the automatic analysis performed by StArt, in this case we have 4 attributes: SelectionScore for the selection step in which the articles were classified by the occurrence of the search string words in the title (10-30) score, PrioritySelection ranging (may be (0) low score 10-15, (1) high score 20, and (2) very high score 25-30), ExtractionScore for the extraction in which the articles were classified by the occurrence of the search string words in the title, keywords and abstract (36-104) and the PriorityExtraction varying according to the score, which can be (0) low score \( \leq 50 \), (1) high score 51-60, and (2) very high score > 60, these 4 attributes are represented only in the Figure 2.
Continuing the results of Figure 2, each article was also manually classified according to some criteria, among them we have 8 manually configured features after the articles read: overview (represents the general evaluation of the use of Google Classroom in the classroom, with (−1) negative, (0) neutral and (1) positive), audience (with (0) student, (1) teacher or (2) others), results (challenges, solutions or others, in this case, the sum was used for each parameter (1) only one type of conclusion, (2) two conclusions or (3) three or more types of conclusion), source of data collected in the variable data (survey, interview, observation, documentation or other, where (1) source, (2) data sources, (3) data sources), level for the group’s educational level (where (0) elementary, (1) high, (2) college, (3) undergraduated, (4) postgraduated, (5) others), tools investigated (label (0) only the GC or (1) compares the GC with at least one other LMS tool), analysis (label (0) quantitative, (1) qualitative, (2) mixed and (3) others) and finally, the attribute paperType (with (0) for report, (1) evaluation study, (2) trial study, (3) comparative study, and (4) others), as can be seen in Figure 3, which shows the surface for understanding each of the data for the 25 extracted papers.

3. Methodology

In this work we proposed a workflow, shown in Figure 4, performed in the tool called KNIME Analytics Platform in order to extract some characteristics of the 25 articles in the scientific database that were not included in the precursor work of [Lima and Isotani 2022]. Thus, we performed a similarity search (SS) to group the most similar articles to each other and help us with analyzes to create guidelines that help
teachers and stakeholders in decision making for the use of the GC platform during times of pandemic [Ferreira et al. 2019, Soares et al. 2021].

Initially, we read the .csv data, which are the 25 papers extracted from [Lima and Isotani 2022] work. Later, we made a filtering of the data, in this case, we excluded the 4 attributes automatically generated by StArt, and we kept the 8 criteria performed manually by the authors of this work. Then, we perform a min-max data normalization, in which the lowest values in a column receive 0, the highest 1, and the others values are calculated proportionally. Then the Euclidean numerical distances are calculated in order to serve as input to the SS. The SS algorithm implemented in KNIME performs a search based on row’s comparisons [Lima et al. 2021b, Lima et al. 2017a]. The similarity criteria must be established and the selected attributes were the same as in Figure 3. In this case, we use a similarity filter by distance between the most similar, with a filter between (0.6 − 1.0). As output, SS returns a column with the index with the most similar neighbor (one neighbor was used) and a column with the value of the distance between these neighbors. Denormalized results are stored in a .csv database, with column filtering. Colors and shapes are configured for Scatter Plot, Parallel Coordinates and Histogram presentation. Each of these graphs will be presented in the Results Section. Spark Line was introduced in the previous section for the purpose of presenting all paper features data through an image.

4. Results and discussion

The SS algorithm used herein aims to present the articles with the greatest similarity to each other based on the characteristics of Figure 3. Based on the results found, Figure 5 presents the result of the search for the nearest neighbor, which in this case x-axis represents the paper ID and y-axis represents the row number for the nearest neighbor, the article line number can be seen in the Table 1. Each line that is the set center of SLR papers has been defined with a different color in Figure 5. Furthermore, the results can be better seen in the “SS” column of the Table 1. For a better results understanding, a graph was made manually herein. from the creation of the graph, it was possible to see that it has a set of connected components. Each of these sets can be seen in “Grp” column of Table 1 and Figure 6 in which the vertices represent the IDs of the extracted papers and the colored edges followed the color pattern in Figure 5. At the end, it is possible to
### Table 1. Paper ID, authors (year), tool used, Extraction Score, Similarity Search (SS), Connected Groups (Grp), Overall Evaluation (OvE) and Row.

<table>
<thead>
<tr>
<th>ID</th>
<th>Authors</th>
<th>Tool</th>
<th>Score</th>
<th>SS</th>
<th>Grp</th>
<th>OvE</th>
<th>Row</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[Octoberlina and Muslimin 2020]</td>
<td>CP</td>
<td>42</td>
<td>$R_{20}$</td>
<td>$G_A$</td>
<td>−1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>[Tinungki and Nurwahyu 2020]</td>
<td>GC</td>
<td>66</td>
<td>$R_{13}$</td>
<td>$G_C$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>[Wan Hassan et al. 2020]</td>
<td>GC</td>
<td>59</td>
<td>$R_6$</td>
<td>$G_E$</td>
<td>−1</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>[Baharun et al. 2021]</td>
<td>GC</td>
<td>36</td>
<td>$R_{22}$</td>
<td>$G_B$</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>[Oyarinde and Komolafe 2020]</td>
<td>GC</td>
<td>77</td>
<td>$R_0$</td>
<td>$G_A$</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>[Oktaria and Rahmayadevi 2021]</td>
<td>GC</td>
<td>76</td>
<td>$R_0$</td>
<td>$G_A$</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>19</td>
<td>[Rahmawati et al. 2021]</td>
<td>CP</td>
<td>50</td>
<td>$R_7$</td>
<td>$G_E$</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>23</td>
<td>[Hasanah et al. 2020]</td>
<td>GC</td>
<td>49</td>
<td>$R_0$</td>
<td>$G_A$</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>25</td>
<td>[Fuady et al. 2021]</td>
<td>CP</td>
<td>51</td>
<td>$R_{24}$</td>
<td>$G_D$</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>26</td>
<td>[Fatmahamin 2021]</td>
<td>GC</td>
<td>104</td>
<td>$R_6$</td>
<td>$G_E$</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>27</td>
<td>[Sultanova et al. 2020]</td>
<td>GC</td>
<td>63</td>
<td>$R_6$</td>
<td>$G_E$</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>29</td>
<td>[Taufan et al. 2021]</td>
<td>CP</td>
<td>55</td>
<td>$R_1$</td>
<td>$G_C$</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>30</td>
<td>[Susanto et al. 2021]</td>
<td>GC</td>
<td>66</td>
<td>$R_1$</td>
<td>$G_C$</td>
<td>−1</td>
<td>13</td>
</tr>
<tr>
<td>45</td>
<td>[Sugiarto et al. 2020]</td>
<td>GC</td>
<td>74</td>
<td>$R_0$</td>
<td>$G_A$</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>50</td>
<td>[Isda et al. 2021]</td>
<td>GC</td>
<td>52</td>
<td>$R_{18}$</td>
<td>$G_E$</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>52</td>
<td>[Baskoro et al. 2020]</td>
<td>GC</td>
<td>46</td>
<td>$R_6$</td>
<td>$G_E$</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>54</td>
<td>[Purnama 2020]</td>
<td>CP</td>
<td>62</td>
<td>$R_7$</td>
<td>$G_E$</td>
<td>−1</td>
<td>18</td>
</tr>
<tr>
<td>61</td>
<td>[Mokhtar and Karim 2021]</td>
<td>GC</td>
<td>67</td>
<td>$R_3$</td>
<td>$G_B$</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>66</td>
<td>[Ahmad et al. 2020]</td>
<td>GC</td>
<td>68</td>
<td>$R_0$</td>
<td>$G_A$</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>70</td>
<td>[Zulkefli et al. 2020]</td>
<td>CP</td>
<td>59</td>
<td>$R_1$</td>
<td>$G_C$</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>72</td>
<td>[Yen and Mohamad 2021]</td>
<td>GC</td>
<td>46</td>
<td>$R_3$</td>
<td>$G_B$</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>91</td>
<td>[Espirito Santo and Lima 2020]</td>
<td>GC</td>
<td>46</td>
<td>$R_0$</td>
<td>$G_D$</td>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>92</td>
<td>[Satenko and Chugai 2020]</td>
<td>CP</td>
<td>48</td>
<td>$R_9$</td>
<td>$G_D$</td>
<td>1</td>
<td>24</td>
</tr>
</tbody>
</table>

See that there are 5 well-defined groups, with the data distributed as Figure 7, following the order: paper ID, year of publication of the paper, the 8 criteria defined in Figure 3 and finally the Group, that represents the graph connected component. Thus, for a better understanding of the data for each of the groups, we present a histogram in Figure 8 with the separate arithmetic mean for each of the 8 attributes used in this analysis. In this way, it is possible to see that each group has characteristics that make it different from the others, for example, the $G_D$ is the group with the highest average for the respondents’ education studies level, the $G_B$ is the only group that presents only works that deal exclusively with the GC tool as e-learning during the pandemic, $G_C$ with the smallest overall among all other clusters, $G_A$ with the highest average for analysis, which means that they were mostly mixed or qualitative in addition to the $G_E$ lower value assumed for the feature audience, meaning that more students were interviewed on average.

5. Guidelines obtained by summarizing results

From the reading of the 25 articles and based on the 5 groups of the graph in the Figure 5 and in the characteristics of parallel coordinates, it was possible to find a set with 5 guidelines for the implementation of GC as tool for e-learning during Covid-19 disease.
pandemic. Each guideline represents central theme extracted by each group, which represents a feature in common to the papers of the groups. This stage of the central choice of the theme of each group (for guideline topic) was herein performed manually.

After reading the articles, the following solutions were pointed out for the implementation of GC as an LMS tool during the pandemic: (1) preparation of all stakeholders involved in the teaching-learning (T&L) process, focusing not only on technical aspects of using the tool, as well as related aspects of reflection-action during distance learning. After reading the articles, it was also possible to understand that (2) it is necessary to verify that everyone involved in process has adequate internet, avoiding problems with data quotas, connection instability, or internet slowness. In this way, it is possible for everyone to have access to the knowledge passed on by the teachers and also be able to perform the tasks (or homework) completely, as well as sign the attendance. Another important aspect is the (3) certification of posted digital materials, whether they
are: videos, texts, audios and images. In order that they have good quality, but reducing these files to smaller sizes, so that students do not have problems downloading them, watch them online or still face physical problems, such as eye pain or anxiety. It is also important to ensure that the contents are in an appropriate language and are summarized so that the student remains engaged, and the teacher must provide adequate time for students to complete assignments within the stipulated time.

Another guideline that becomes essential is the question of whenever possible (4) vary in the tools used, for example, the Google Forms tool is suggested for the application of questionnaires, and the Google Drive and Google Agenda are tools recommended to store and organize the T&L process to increase the students’ motivation to face online learning, always giving students time to assimilate the tools. Finally, whenever possible, (5) use other tools to support face-to-face communication with students whenever possible. Face-to-face communication is extremely important for students and teachers to establish a contact relationship for feedback and even for exchanging experiences, reducing students’ anxiety and give them more autonomy. Thus, a series of tools are presented by the authors, such as Zoom, Google Meet, and tools such as Whatsapp facilitate communication with students and their parents, especially those at levels of Elementary or High Education. Another important aspect is creating a break between activities or themes so that students are less anxious. Figure 9 represents the compilation of the results presented here in the form of 5 guidelines, which will be useful for the implementation of GC in the classroom, according to the reading of the 25 articles extracted by the SLR. Furthermore, for some of these actions to take place effectively, it is necessary for governments to be able to contribute more financial budget to educational institutions, especially in underdeveloped or developing countries.

6. Conclusions

The pandemic changed people’s lives and this affected their daily activities in various sectors, one of them was the education sector. So that school activities were not harmed, educational stakeholders had to overcome this problem with the use of LMS for educational support. One of the main LMS tools used around the world was the GC tool [Espirito Santo and Lima 2020] [Tinungki and Nurwahyu 2020] [Baharun et al. 2021]. To show a set of guidelines for the implementation of this tool in schools to be effective, we...
Figure 9. Guidelines created from SLR for GC usage as LMS.

Guidelines for using Google Classroom as an e-learning tool

1. Promote teacher and student training
   Training students and teachers is key to a successful GC implementation, because the lack of technical literacy to use the tool hinders T&L.

2. Make sure all students have adequate internet
   When students do not have adequate internet quota or when the signal is poor, it can hinder the teaching-learning process.

3. Certify about size and quality of digital materials
   Build digital materials that are short, take up little disk space, are of good quality, and can be quickly downloaded by the students.

4. Use other platforms to support education
   The use of other platforms such as Edulogy, Google Meet, Zoom, Google Forms and Moodle can diversify e-learning tools to improve teaching-learning process.

5. Establish communication with students
   Encourage students’ independence, but establish communication through the GC message board and discussion forums with the class to improve teaching-learning.

carried out a study based on the SLR authors [Lima and Isotani 2022], which presented a set of 25 works. In this work, we used a SS, and initially 11 different groups were found, using the KNIME tool. After assembling a graph, we realized that 5 connected components were found. From these 5 groups and after reading the articles, we found a set of 5 guidelines that can be used by teachers to implement the GC as an LMS tool. In addition, we hope that with future work we will be able to conduct courses for teacher training in the GC tool and also in aspects of reflection-action to guide e-learning, not only in the pandemic, but also as an alternative to higher quality hybrid classes.

References


