

A Data Collection Experience with Canvas LMS as a Learning Platform

Aarón Rubio Fernández¹, Fernando Santamaría González², Pedro J. Muñoz Merino¹,
Carlos Delgado Kloos¹

¹ Universidad Carlos III de Madrid, Getafe, Spain
{aarubiof, pedmume, cdk}@it.uc3m.es

² Universidad Isabel I, Burgos, Spain
fernando.santamaria@ui1.es

Abstract. This paper studies the academic information that can be obtained from the “Canvas LMS”, an e-learning platform. The main objective is to obtain academic information that allows teachers to improve the learning process of the students. To obtain academic information, two types of low level data, raw data, are analysed: the messages written by the students and the questionnaires filled by the students. In this way, through this analysis different indicators are obtained, which provide academic information to the lecturers. The same indicators can be obtained in other scenarios or e-learning platforms, since the low level data used can be obtained on most of these platforms. Besides this, different applications of the indicators will be analysed, with the aim of specifying the academic information that a lecturer can obtain by these indicators.

Keywords: Learning Analytics, Academic Indicators, Canvas LMS Platform, E-Learning Platforms, Text Mining, Characterization of Messages, Data Extraction, Virtual Learning Environments.

1 Introduction

In recent times, the use of e-learning platforms has increased due to the development of the information and communication technologies. The e-learning platforms facilitate student learning, consequently they are fundamental in any education system [1]. For this reason, most of the universities use one, or several, e-learning platforms (for example, systems based on Moodle or Open edX).

Nevertheless, the learning process can be improved. The e-learning platforms generate a large amount of data which, correctly analysed, can improve the process of learning. Precisely, this paper analyses the data generated by the “Canvas LMS” e-learning platform [2]. Through text mining techniques, different indicators have been obtained, and they can be used to improve the learning process. In particular, the messages written in the forums, and the questionnaires made by the students, will be analysed. These indicators have been obtained through a software tool, which extract and

analyse the data stored in the platform. In this way, the teachers can obtain the academic information without technical knowledge.

“Canvas LMS” is a cloud-based Learning Management system created by the US Instructure company. This platform for learning was created in 2011. It has been chosen for this study because it is open-source and flexible when extracting the data. The platform has several analytical functionalities, like visualizing the number of views of an information page, but none of them provide detailed information about the messages written in the forums (characteristics of these messages, amount of information, type, sentiment, etc.), at least until now. In any case, the objective of this paper is not to develop analytical tools for the platform, the objective is to define several indicators with which to obtain academic information. These indicators are not exclusive of the Canvas platform, the indicators are obtained through low level data, that are present on most of the e-learning platforms.

As stated previously, the e-learning platforms are fully integrated into the education system. For this reason, it is necessary to improve its performance through the analysis of the academic data that they generate. Besides this, the academic data obtained can be used to perform more complex analysis that provide additional academic information. For example, the emotion of the message can be combined with the theme commented on the message, this allows know if the student is interested in the theme (positive emotion) or not (negative emotion).

Therefore, the main objective of this paper is to obtain relevant academic information, with which improves the learning process of the students. We have obtained the academic information through two different types of indicators. On one hand, we have used indicators related to the messages written in the forums. In the other hand, we have used indicators related to questionnaires made by the students.

The data used to obtain these indicators have been obtained from a real course, with real students that use the “Canvas LMS” platform to learn. Nevertheless, as commented before, the indicators and the processes described in the present paper can be exported to other e-learning platforms, since the low level data used can be obtained on most of these platforms (for example, the messages written in the discussion forums).

Therefore, in this paper, we focus on the academic information that the teacher can obtain through these indicators. But, in the future, we will analyse the actions that the teachers can perform through this academic information in real courses, and the impact of these actions on the learning of the students.

2 Related Works

The studies related to the e-learning platforms have been increasing in recent times, due to the progressive implantation of these platforms in the different educational systems. In the year 2000, about 2970 studies related to “e-learning” were published [3]; however, in the year 2016, about 42300 studies related to “e-learning” were published [4]. These studies can refer to different aspects of the e-learning platforms. We have classified the studies into three types to see their characteristics.

Studies Related to the Efficiency of the Platforms. These studies are aimed to optimize the efficiency of the e-learning platforms. From the optimization of the source code, to the optimization of the global architecture of the platform (for example, the study of the efficiency of the system in a cloud computing environment [5], or the analysis about the optimal architecture for the e-learning platform [6]).

Studies Related to the Learning Analytics Policies. The learning analytics is a field of study quite young, therefore studies that allow to establish the most appropriate institutional policies are needed (especially, studies related to the privacy and confidentiality of all the members of the education system). An example of this type of studies is the European project SHEILA [7], that aims to build a policy development framework for the policies related to the learning analytics.

Studies Related to the Improvement of the Learning Process. In this case, the studies aim to improve the learning process of the students. For this reason, in this type of studies (in which this paper would be included), is important to analyse correctly the learning process of the students [8], as well as to define high level indicators that allow to improve the learning process of the students [9]. Through studies of this type, several e-learning platforms have been able to integrate analytic modules (for example, Moodlerooms, Blackboard, Brightspace Insights, Cognos Analytics for Jenzabar) or specific plugins like KlassData, and these modules or plugins have allowed to improve the learning process of the students.

3 Characterization of the Messages Written in the Forums

To obtain relevant academic information, it is necessary to characterize the messages written in the forums through different indicators. These messages contain different types of academic information, for example, the messages can show the relations established between the students, or the emotion of the students. In this case, the indicators obtained are the ones that follow.

Table 1. Message indicators.

Indicator	Unit of measure
Longitude	Number of words
Mean longitude of the sentences	Number of words
Amount of information	Value included in the interval [0,1]
The message includes a theme specified by the teacher	Yes / No
Type of message	Element of the set [intervention, comment, question, answer]
Emotion of the message	Element of the set [positive, negative]

The longitude of the message and the mean longitude of its sentences are easy to understand, but the rest of indicators need an additional explanation. These “complex” indicators are detailed below.

3.1 Amount of Information

To measure the amount of information contained in the message, metrics based on the longitude of the message cannot be used. The students can write very long messages, with many words, that contain little information (for example, if the students talk about the same concept over and over again). For this reason, it is necessary a specific metric that measure the amount of information contained in the message.

To obtain this metric, the following reasoning has been used. The amount of information contained in the message is determined by the number of different concepts included in the message. In this context, a concept is a group of words that have the same core meaning, for this reason, the concepts will be grouped around words that have the same root (the root of a word is the minimum and irreducible base that share all the words of a word family). For example, if the message has many words with the same root, then the student is talking about the same concept over and over again. Therefore, if the message has many words with different root, it is exposing many concepts and, consequently, it has a large amount of information.

We can see that the concepts and the information of a message are closely related; in fact, in this context, the information of a message can be seen as the set of concepts included in the message.

The value of this metric is the quotient between the number of words with different roots of the message and the total number of words of the message. For this reason, the possible values are included in the interval $[0, 1]$, where the value 0 means that the message has no information, and the value 1 means that the message has the maximum information possible.

To obtain this indicator the “stop words” (words without meaning, for example, the prepositions) have to be filtered. Once the “stop words” have been filtered, we have to obtain the roots of the words and divide the number of different roots by the total number of words of the message (total number of words without considering the “stop words”, i.e., the number of words of the message after having filtered the “stop words”).

3.2 The Message Includes a Theme Specified by the Teacher

As its name suggests, this indicator allows to know if the message includes a theme specified by the teacher. A theme is any set of knowledge related to each other, that the teacher explains to its students. To define a theme, the teacher writes a list of key words in a text file. This indicator can be used to perform more complex analyses, for example, if the message includes a theme and the emotion of the message is positive, then the teacher knows that the student is interested in that theme.

This indicator also uses the roots of the words to know if the message includes a theme specified by the teacher. The idea is that if the message includes a theme, then

it will use a word that share the root with some of the words that define that theme. Therefore, to know if the message includes a theme a single search is required (the roots of the words that define the theme are searched in the message).

The indicator provides quite acceptable results, as long as the teacher defines the themes correctly. For a set of 151 messages, and three themes defined by the teacher, the following results have been obtained.

Table 2. Results obtained.

Theme	Number of messages that include the theme	Number of messages that have been correctly classified	Number of messages that have been incorrectly classified
Learning	99	97	2
Ecosystem	20	20	0
Connectivism	71	71	0
Without theme	44	44	0

The success rate obtained is quite high, since only two of the 151 messages have been incorrectly classified (there are messages that included more than one theme, and messages that do not include any theme, for this reason the total sum of messages is not 151). As stated previously, it is necessary that the teacher defines the theme appropriately to obtain high success rates.

3.3 Type of Message

This indicator allows to know the relation established between the student that wrote the message, and the rest of the students that participate in the forum. The objective is to know if the student that wrote the message has some type of relation with the rest of students that wrote messages in the forum. For this purpose, four types of messages are defined.

- **Intervention.** The student doesn't establish any type of relation with the rest of the students. The message exposes a series of concepts, but the student doesn't interact with the rest of the students.
- **Comment.** In this case, the student analyses a message written by another student. For this reason, this type of message indicates that the student has established a relation with another student.
- **Question.** In this type of message, the student asks for information to the rest of the students. Besides, this indicator can be used to know the doubts of the students (the concepts more complex to understand).
- **Answer.** In this case, the student answers the question posed by another student. Therefore, a relation with the user that asked the question is established. Also, this indicator can be used to know if the students help each other.

To obtain this indicator, supervised machine learning techniques have been used, in particular, a multi-class classifier has been implemented. The classifier is based on three binary classifiers, for this reason, the classification is done in three steps. First of all, the classifier distinguishes between messages of type “intervention” and messages of type “non-intervention” (one of the three remaining types). If the message is not an intervention, the second binary classifier distinguishes between messages of type “comment” and messages of type “non-comment” (a question or an answer). Finally, the third binary classifier allows to know if the message is a question or an answer. To train the classifier a set of 642 messages, tagged manually, has been used.

A comparative study of different classifiers has been done through the data mining tool “Weka” [10]. The results of this study are the following.

Table 3. Results obtained.

Weka’s classifier	First classification [Success rate %]	Second classification [Success rate %]	Third classification [Success rate %]
KStar	90.9502	91.9811	93.9394
Classification Via Regression	90.7994	84.434	93.9394
Dagging	90.0452	87.7358	84.8485
Decorate	90.7994	91.9811	93.9394
LogitBoost	89.8944	91.9811	93.9394
Rotation Forest	91.7044	91.5	93.9394
Random Forest	92.1569	92.9245	93.9394

The classifier with the highest success rate is the “Random Forest” classifier [11]. Therefore, the classifier is based on three binary “Random Forest” classifiers. This classifier provides a success rate higher than 80% (all these success rates have been obtained by cross validation).

The messages written by the students are usually similar, for this reason, with another set of messages the success rate should be close to 80%.

Finally, we would like to highlight that we can see similar analyses in other works (for example, in [12]), therefore, although our indicator “type of message” is different than the indicators used in these analyses, it is possible to glimpse the utility of our indicator through these works.

3.4 Emotion of the Message

This indicator allows to know the emotion expressed in the message and, therefore, the emotion of the student that wrote the message. In this way, it is possible to know if the student is interested in the forum.

Two types of emotions have been defined [13], on one hand a positive emotion that indicates that the student is interested in the forum and, in the other hand, a negative emotion that indicates that the student is not interested in the forum.

The positive and negative emotions are detected through positive and negative opinions expressed in the messages of the students. For example, if the message is “I don’t understand the theme, is not well explained”, the message expresses a negative opinion that can be used as an indicator of a negative emotion. It is true that a negative opinion is not a “direct indicator” of interest, but the opinions of the students can be used as warning indicators, because if the students have several negative opinions, they will lose the interest in the theme. On the other hand, if the students have several positive opinions, they will have more interest in the theme. In future works, we will verify the validity of this hypothesis.

As in the previous case, we have performed a comparative study of different classifiers (using the tool “Weka”) and, again, the “Random Forest” classifier has obtained the best results; for this reason, we have used the “Random Forest” classifier to know the emotion expressed in the messages. In this case, only a binary “Random Forest” classifier is required (the classifier only has to distinguish between positive and negative messages). The classifier provides a success rate of 85 % (this rate has been obtained by cross validation). As in the previous indicator, with another set of messages, the success rate should be close to 85%, since the messages written by the students are usually similar.

4 Characterization of the Questionnaires

The questionnaires filled by the students are a valuable source of academic information (a questionnaire is a set of questions with several answers to choose, the questionnaire is completed in the own platform). Through the information associated with the questionnaires, the student’s progress in the acquisition of knowledge can be estimated. In particular, the following academic information has been extracted from the platform.

- Score obtained in the questionnaire.
- Maximum score of the questionnaire.
- Number of times that the questionnaire has been done (number of attempts).
- Time used to complete the questionnaire.
- Identifier of the questionnaire.
- Identifier of the student that has completed the questionnaire.

With all this information, relevant academic information can be obtained. On one hand, it is possible to know if the students understand the concepts asked in the corresponding questionnaire. Using the scores obtained by the students, it can be seen if the students understand the concepts (if the students obtain, on average, high scores) or not (if the students obtain, on average, low scores). In the other hand, it can be seen the individual progress of each student. As is known the score obtained by the student in every one of his attempts, it can be seen if the student progresses (if he obtains an increasing score), or if the student stagnates (if he obtains a decreasing score, or similar scores in the same questionnaire). Besides, as it is known the score obtained by the

student, the number of attempts done and the time used, it is possible to differentiate different types of students [14]. For example, it is possible to differentiate between a student that has obtained the score X in a time Y , and a student that has obtained the same score (X) but in a lesser time (for instance $Y - 10$ minutes); the last student has the concepts more clear, because he has obtained the same score in less time.

Although most of these indicators are basic and well known, it is important to remember the educative potential that these indicators have. Often, the results obtained in a questionnaire, exam, etc., are not analysed properly by the teachers, for this reason it is necessary to encourage them to use these data. Providing to the teachers the information related to the progress of the students, the themes more complicated to understand, etc., it is more likely that the teachers use this information (although, ultimately, it is the teacher who decides if this information will be used). For this reason, this section is important to improve the learning process of the students.

In this way, with the information associated with the questionnaires, it is possible to obtain relevant academic information that allows to improve the learning process of the students.

5 Applications

As commented before, the main objective of this paper is to obtain relevant academic information, with which improves the learning process of the students. Therefore, the applications considered in the present paper focus on the improvement of the learning process of the students.

First of all, the platform provides data that can be used as academic information source almost directly. In particular, these data are the data associated with the questionnaires made by the students. Simply combining the information of the questionnaires (scores obtained, time used, number of attempts, etc.) with the concepts associated with these questionnaires (the concepts asked in the questions), it is possible to obtain a general panoramic of the student situation. Therefore, the teacher can know what concepts are more complex to learn, what students have more difficulties to learn certain concepts, what evolution have the students, etc. All this academic information allows the teacher to improve the learning process of the students.

The indicators associated with the messages written in the forums can also be used to obtain relevant academic information. Although some uses of these indicators have already been commented previously, it is necessary to specify the academic applications that could be developed with these indicators. We will use these applications in future courses to get a better feedback of the learning process of the students. Besides, these applications provide academic information that will allow us to perform more in-depth analyses.

5.1 Applications of the Indicators Associated with the Messages

To obtain relevant academic information it can be used the indicators (associated with the messages) separately, or it can be combined several of these indicators to obtain

more detailed academic information. The applications of the indicators associated with the messages are, principally, the following.

Applications that Use a Single Indicator. As its name suggests, these applications only use one of the indicators commented previously. These applications are shown below.

- The indicator “The message includes a theme specified by the teacher” can be used to know the degree of involvement of the students in the dynamic of the forum. For example, if in the forum the students can write about three themes (the themes A, B and C), it would be possible to identify the students that only have written about one theme (for example, the theme A), the students that have written about two themes, and the students that have written about the three themes. In this way, the teacher can know the students that are most implicated in the forum. The teacher can use this academic information to value the involvement of the students in the dynamic of the class.
- The indicator “Type of message” can be used to know the relationships established among the students of a certain forum. In addition to this application, this indicator has a great potential when it is combined with other indicators, as will be shown in the next section.
- As for the indicator “Emotion of the Message”, it can be said that the indicator allows to know the degree of interest of the students in the forum. As in the previous case, this indicator has a great potential when it is combined with other indicators, since it is possible to obtain more academic information than in the case of using the indicators separately.

Applications that Use Two or More Indicators. In this case, the applications use several indicators to obtain academic information. It is logical that if the application uses several indicators, it will obtain more complete academic information than in the case of using a single indicator. Therefore, the applications shown below, will provide very relevant academic information.

- Combining the indicators “Longitude”, “Mean longitude of the sentences” and “Amount of information”, it is possible to know the way in which the student expresses himself. Therefore, it can be differentiated different type of students, for example, there may be students that write very long messages, with complex sentences (long sentences), but that contain a very redundant information. At the other extreme, there may be students that write very short messages, with simple sentences (short sentences), but that contain more information. Through an application that combines these indicators, and shows the results, the teacher will be able to obtain a first information of the messages written in a forum, which can have repercussions in a better feedback for the students. For example, the teacher can tell the student that he should write shorter messages with less redundant in-

formation, or that he should write more elaborate messages so that they don't look like a list of unrelated concepts.

- Combining the indicators "The Message Includes a Theme Specified by the Teacher", "Type of Message" and "Emotion of the Message", it is possible to know the way in which the students approach the themes. Therefore, the teacher can know if the student is exposing a theme, if the student is asking about a theme, or if the student has such a level of knowledge that he can answer the questions of his classmates. Besides, thanks to the indicator "Emotion of the Message", the teacher can know if the student is interested in the theme. In this way, a very complete academic information is obtained, with which the teacher can better understand the learning process of the students.
- The teachers have to be able to compare all these indicators graphically, for this reason, visual analytics applications are necessary. With this type of applications, the teachers can extract specific academic information for each group of students which will allow them to follow correctly the learning process of the students. The teachers need to visualize and compare the indicators commented previously, since they can't be based, exclusively, on academic information coming from other groups of students. The experiences with other groups of students are important, the teacher has to lean on them, but each student learns in a unique way, different from the rest of students, therefore, it is necessary to be able to access all the academic indicators that the student generates.

Therefore, through the applications described above, the teachers can obtain relevant academic information, with which to improve the learning process of the students.

5.2 Some Results

Once commented the most relevant applications, it is important to show some results that these applications would provide. The applications use a set of 642 messages, generated by approximately 30 students. For example, an application that could compare the amount of information contained in the message with the longitude of the message, would provide results like the following.

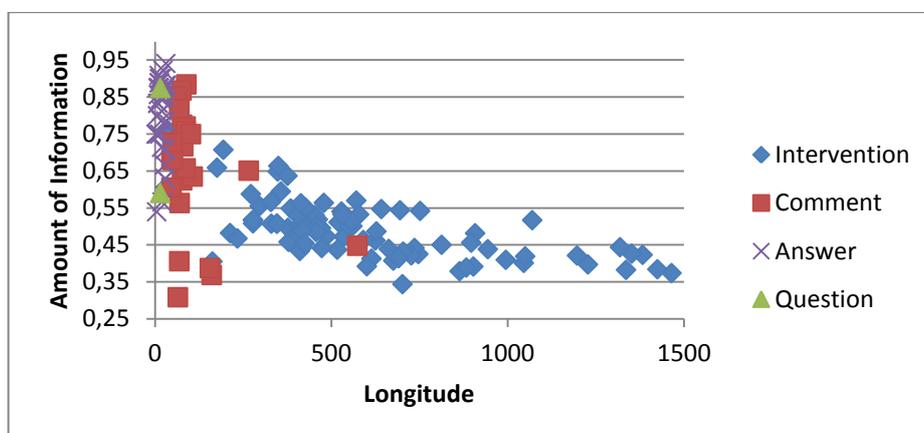


Fig. 1. Amount of information contained in the messages compared to the longitude of these messages.

The figure shows that a non-linear relation between these two indicators exists. The students that write long messages, with many words, add a lot of redundancy, for this reason the amount of information decreases. The value associated with the amount of information has been normalized with regard to the total number of words of the message, for example, if a message includes three concepts explained in 30 words, that message has more information than another message that includes these three concepts but explained in 100 words.

Besides this, we can see several characteristics of the different type of messages. For example, the figure shows that the messages of type “intervention” contain more words than the other types of message, but they have less information (as commented before, the amount of information has been normalized with regard to the total number of words of the message). Therefore, the “comments”, the “questions”, and the “answers”, contain less words than the “interventions” but they have more information. We can see also that the number of “answers” is higher than the number of “questions”. This can be explained if several students answer the same question, which is usual in the discussion forums.

Therefore, through the figure shown, the teacher can obtain academic information about the messages written by the students.

In this other example, the application allows to know if the message includes the theme specified by the teacher and also the emotion associated with the message. In this way, the messages can be classified into four classes.

- Class 1: the message includes the theme and expresses a positive emotion.
- Class 2: the message includes the theme and expresses a negative emotion.
- Class 3: the message doesn't include the theme and expresses a positive emotion.
- Class 4: the message doesn't include the theme and expresses a negative emotion.

The next figure shows the distribution of the messages written in a certain forum, according to these four classes.

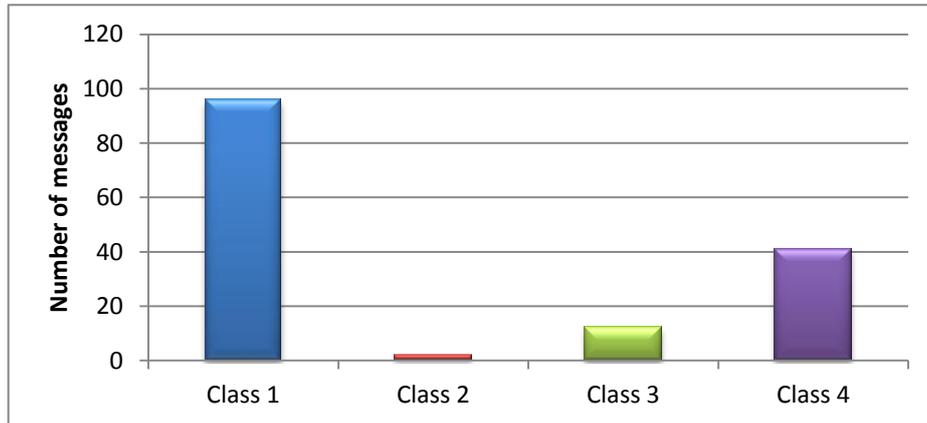


Fig. 2. Number of messages of each class.

In this case, the figure shows that the theme defined by the teacher generates a positive emotion to the students (in other words, the theme generates interest to the students), since the majority of the messages that include the theme have a positive emotion associated with them. In this way, the teacher obtains a relevant academic information, since he can know if the students are interested in the different themes.

Also, we can see the distribution of the messages in different forums, i.e., the number of “interventions”, “comments”, “questions” and “answers”, in the forums. In the following example, we have analysed three forums with similar results (on average, the students have obtained 75.56 out of 100 points in the forum #1, 69.78 out of 100 points in the forum #2, and 72.59 out of 100 points in the forum #3).

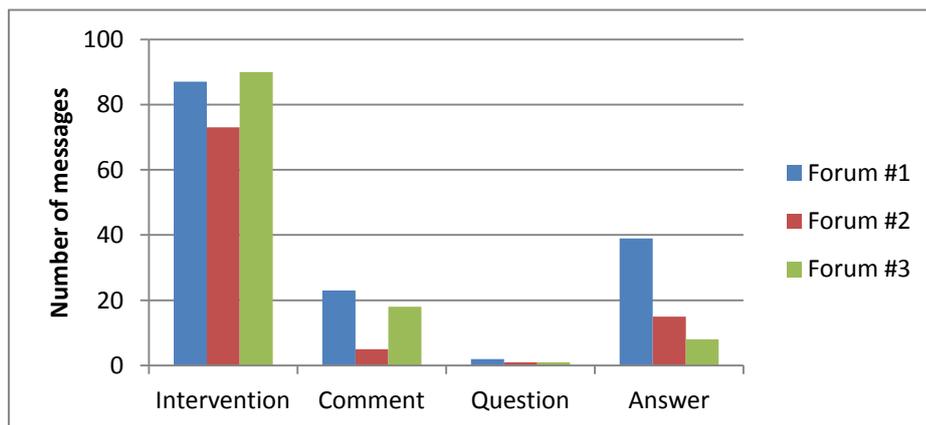


Fig. 3. Number of messages of each type.

We can see that the students usually write a lot of “interventions” but very few “questions”, this suggests that they have the concepts clear. If we compare this information with the grades obtained by the students, we can ensure that the students have understood the concepts, since they haven’t had many doubts and they have obtained good grades in the forums. Besides, we can see that the students help each other, since there are more answers than questions, i.e., if a student has a question several of his partners try to help him with their answers.

The fourth and last example shows an application that allows to know the amount of information contained in each type of message. Therefore, the application shows the amount of information contained in the messages of type “intervention”, “comment”, “question” and “answer”. For example, the application could show the amount of information contained in the messages of type “answer” (in the messages that answer the questions posed by the students).

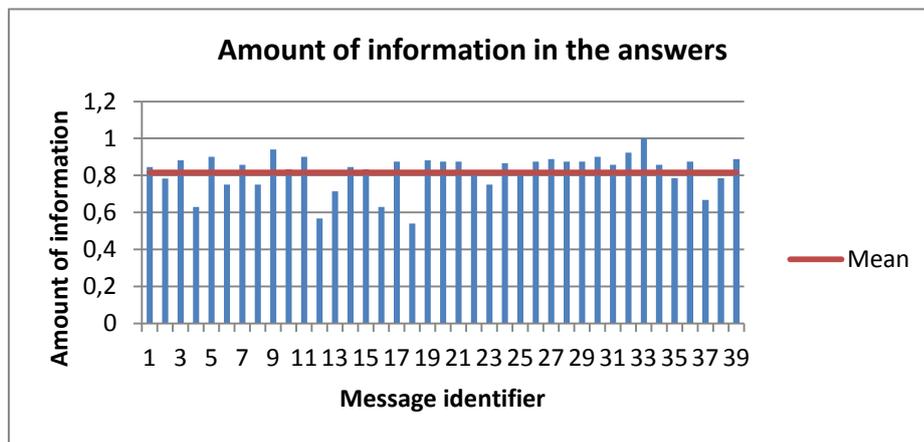


Fig. 4. Amount of information contained in the messages of type “answer”.

In this way, the teacher can obtain a first impression of the “quality” of the messages, in other words, the teacher can know if the messages contain an amount of information reasonable for its class. For example, the messages of type “answer” are short messages that don’t usually have excessive redundancy, since the student is answering a question; therefore, the amount of information will be high, as can be seen in the figure.

5.3 Considerations

Once commented some of the applications of the indicators defined in this paper, it is important to highlight some considerations to take into account for a correct use of the applications.

The first consideration is related with number of needed messages for training the classifiers (the classifier associated with the type of message, and the classifier asso-

ciated with the emotion of the message). The classifiers need a large number of messages, in particular, the classifiers need several hundred of messages. In this paper, a set of 642 messages has been used to train the classifiers.

The teachers that use the applications may not have technical knowledge, for this reason, it is necessary to automate the process of extraction, analysis and visualization of the data. In this way, a teacher without knowledge about machine learning, programming, etc., can use the applications without any type of problem. For example, the ANALYSE tool [15] allows to visualize several academic indicators, and can be used by teachers without technical knowledge.

This paper shows several indicators that allow to obtain academic information, but there are more indicators with academic potential, for example, indicators associated with the time. The time invested by the student to read an article, watch a video, or make a questionnaire is a valuable source of information, since it can be taken as an indirect indicator about the work done by the student. Although, like any other indicator, this indicator can provide wrong results if the student doesn't use the platform as it is expected (for example, if the student open a document, he doesn't read it and, after a while, he closes it). Therefore, the fact that the users can use the platform in other ways has to be taken into consideration, since this fact substantially decreases the quality of the academic indicators obtained. Nevertheless, this doesn't mean that these indicators are not useful, since they have a lot of potential from the academic point of view. Consequently, it is reasonable to say that all the indicators provide academic information, therefore it is necessary to know what information is needed and what indicators can provide it.

6 Conclusions

In this paper, different academic indicators have been defined that can be obtained from the data generated by the "Canvas LMS" e-learning platform. As stated previously, these indicators can be extracted from other e-learning platforms (logically, these platforms have to generate the necessary low level data to extract the indicators); therefore, if the teacher uses other platforms, different from "Canvas", he can obtain the same information that in the case of using the "Canvas" platform.

In this way, the indicators defined in the present paper can be used to obtain relevant academic information that allows to improve the learning process of the students. Besides this, these indicators can be used to perform more complex analyses that provide more detailed information. For example, it could be studied if exist any type of correlation between the score obtained by the student (for instance, the score obtained by the student in a questionnaire), and the types of message that the student has written in the forums where the themes of the course are explained (for example, the messages written in forums related to the themes asked in the questionnaires). In this way, the indicators analysed in this paper allow to obtain academic information directly and, besides, they can be used to perform more complex analyses.

Acknowledgments. Work partially funded by the eMadrid project, funded by the Madrid Regional Government with grant No. S2013/ICE-2715, and the Spanish Ministry of Economy and Competitiveness Project RESET (TIN2014-53199-C3-1-R) and Project SNOLA (TIN2015-71669-REDT).

References

1. Arkorful, V., Abaidoo, N.: The role of e-learning, advantages and disadvantages of its adoption in higher education. *International Journal of Instructional Technology and Distance Learning*, vol 12(1), pp. 29-42. (2015).
2. Canvas LMS Homepage, <https://www.canvaslms.com>, last accessed 2017/04/10.
3. Studies published in the year 2000 related to “e-learning”, https://scholar.google.es/scholar?q=%22e-learning%22&as_ylo=2000&as_yhi=2000, last accessed 2017/06/12.
4. Studies published in the year 2016 related to “e-learning”, https://scholar.google.es/scholar?q=%22e-learning%22&as_ylo=2016&as_yhi=2016, last accessed 2017/06/12.
5. Pocatilu, P., Alecu, F., Vetrici, M.: Measuring the efficiency of cloud computing for e-learning systems. *WSEAS Transactions on Computers*, vol. 9(1), pp. 42-51. (2010).
6. Muñoz-Organero, M., Delgado, C., Muñoz-Merino, P.J.: Personalized service-oriented e-learning environments. *IEEE Internet Computing*, vol. 14(2), pp. 62-67. (2010). doi: 10.1109/MIC.2009.121
7. SHEILA Project (Homepage), <http://sheilaproject.eu/>, last accessed 2017/04/10.
8. Ruipérez, J.A., Muñoz-Merino, P.J., Leony, D., Delgado Kloos, C.: ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior*, vol. 47, pp. 139-148, ISSN 0747-5632 (2015). doi:10.1016/j.chb.2014.07.002
9. Muñoz-Merino, P. J., Ruipérez, J.A., and Delgado Kloos, C.: Inferring higher level learning information from low level data for the Khan Academy platform. In: *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK '13)*, Dan Suthers, Katrien Verbert, Erik Duval, and Xavier Ochoa (eds), pp. 112-116. ACM, New York, NY, USA (2013). doi: 10.1145/2460296.2460318
10. Weka Homepage, <http://www.cs.waikato.ac.nz/ml/weka/>, last accessed 2017/06/12.
11. Breiman, L.: Random forests. *Machine learning*, vol. 45(1), pp. 5-32. (2001). doi: 10.1023/A:1010933404324
12. Wong, J. S., Pursel, B., Divinsky, A., Jansen, B. J.: Analyzing MOOC discussion forum messages to identify cognitive learning information exchanges. *Proceedings of the Association for Information Science and Technology*, vol 52(1), pp. 1-10. (2015). doi: 10.1002/pa2.2015.145052010023
13. Aballay, L., Aciar, S., Reategui, E.: Método para detección de emociones desde foros utilizando Text Mining. *Campus Virtuales*, vol. 6(1), pp. 89-98. (2017)
14. Pérez, P. T.: E-evaluación y e-learning: el tiempo de respuesta como herramienta de evaluación. *RELADA-Revista Electrónica de ADA-Madrid*, vol. 6(4). (2012).
15. Pijeira Díaz, H. J., Santofimia Ruiz, J., Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Delgado Kloos, C.: A Demonstration of ANALYSE: A Learning Analytics Tool for Open edX. In: *Proceedings of the Third ACM Conference on Learning@ Scale*, pp. 329-330. ACM. (2016). doi: 10.1145/2876034.2893402