

Interaction through Online Forum and Its Analytical Study Using Text Mining

C. Pooranachandran¹, E. Sennavazhagan² and T. Perumal³

¹ Associate Professor
Government Arts College for Men(Autonomous), Nandanam, Chennai-35
cpooranachandran@hotmail.com

² Research Scholar
Government Arts College for Men(Autonomous), Nandanam, Chennai-35
sennavazhagan@gmail.com

³Research Scholar
Government Arts College for Men(Autonomous), Nandanam, Chennai-35
tperumal24@gmail.com

ABSTRACT— *This research study proposes a cluster analysis of discussion forums using a text mining technique which uses a graph formalism to represent relevant terms found in a text, as well as their relationships. In proposed approach is based on the comparison of graphs. Students centered methods, such discussion are more effective in developing higher order of intellectual skills and problem solving skills. This research process the intervention in which teachers try to mediate the debate by motivating and challenging those who posted the relevant contributions and to participate more. This research results using text mining approaches where to concentrate. The discussion forum students follow general discussions, contributions and to had better understand how to mediate forums discussions as to support students learning.*

Keywords— Cluster analysis, e-learning, forum conversation, Online discussion

1. INTRODUCTION

Technology is increasingly being embraced as an efficient tool which provides increased flexibility for learners in higher education. Computer-mediated communication is a key element of Web Based Learning systems and strategies. Online discussions are one of the most important applications of computer-mediated communication in e-learning environments. This is because they provide an asynchronous collaborative learning environment where interaction takes place between group members, and they have been included in many learning management systems. Assessing students' contributions hence becomes difficult and time-consuming for the course instructors. Most methods of assessing students' participation in online discussions follow a strictly quantitative approach that does not necessarily correlate with learning or students' effort. Indeed, this may encourage students simply to post frequent messages, without making a serious attempt to address the problem under discussion.

In online forums, teachers suggest a discussion topic to students who, in turn, present their reports. Depending on the topic, on the learner's profiles, on the interest raised by the theme, forums may contain hundreds or only a few contributions. Some of these messages may present favorable arguments or opposing ideas concerning the topic of discussion. While some of the texts do not present relevant ideas concerning the topic, other contributions are meaningful as they deal with important concepts by either citing or relating them.

2. RESEARCH MOTIVATION

Broadly speaking there are several reasons to embark on text mining: to enriching the content in some way; to enable systematic review of literature; for discovery or computational linguistics research.

Text mining can be used to improve the indexing of text. This is especially useful for publishers since they can create indexes more efficiently. They can also add value in a digital environment by creating links from articles to relevant further reading

Text mining can help a researcher systematically review a much large body of content and do it faster. There is considerable demand for this kind of text mining in the online learning environment: Furthermore, text mining can help researchers keep up with their field and reduce the risk that they miss something relevant.

3. TEXT MINING

Text mining is the raw material for this research. This area appears to be driving a very large part of the current activity in text mining. Around half of the publishers recently surveyed have been approached by researchers in this field requesting permission to mine their content. This is also the area where the most progress is being made in developing new tools and techniques.

This study aims at verifying a viable way to help teachers check relevant contributions posted in discussion forums. In this approach, teachers can observe how learners interact in the discussion and, by doing so, they may identify less significant texts.

The benefits of using online discussion forums and peer to peer learning for enhancing student learning are well known. Other than in full online courses, their adoption in traditional learning environments complementing face-to-face teaching is increasing and has now become a common educational strategy in higher education. In fact, interest in using online discussion forum as an indicator of students' performance is increasing. online discussion forum and case study analysis in teaching a postgraduate elective unit to accounting and other business students. The objective is to combine the benefits of case study method of teaching with online discussion forum to enhance the quality of learning while making this an assessment component in order to ensure active students' participation.

Traditional forums do not provide functionality to re-order responses based on correctness, they merely list responses chronologically. Moreover, users are free to respond with non-answers such as: "I have this same problem!", "Can you provide more information?", "No idea, good luck.", and "Thanks, that solved it." None of these responses constitute answers to the original poster's question. New visitors with the same question must invariably sift through non-answers to find answers. This often leads to the creation of a new thread with the same question and the ubiquitous non-answer response.

4. FORUM DEVELOPMENT

Online forum is the discussion area now a day the most of the conversation area is the forum the many fields are in the forum discussion each area they have one forum to discuss about their product and the clarification. This chapter we discuss how we developed the forum for the education area. The heading wise we discuss the development of the forum.

Here researcher developed that online forum for the education side. All the students are having many questions to ask to professor so the time and the professor availability not there in the college hours so the students not getting their answer with time and they not able to fulfil their answers. Here all the students can able to register in the forum and they can rise their questions in the forum. Some students have some ideas they will reply for the posts. Our staff also registered in the forum so they can give some idea in the forum so the student gets answer in both the side (student or staff) they will get their posts reply

This study is relevant due to the importance of doing a qualitative evaluation of the messages registered by students in online discussion forums, that is, meaningful contributions concerning the theme being debated. Quantitative analysis is simpler to be implemented, for instance, when verifying the participation rate of learners and the extent of the discussions. However, according to Mazzolini, it does not represent an adequate way of portraying contributions in a forum. For teachers, qualitative analysis is essential so that they carry out a close assistance of how learners are building their knowledge. Moreover, it allows teachers to verify the real contribution of each student, observe and interfere in the discussions.

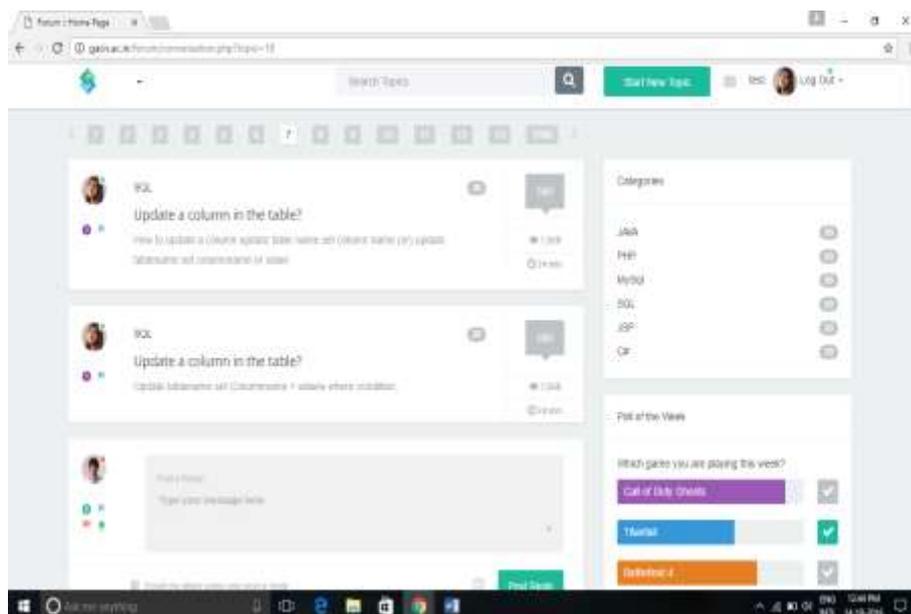


Fig 4.1

5. FORUM DISCUSSION

In online forums, teachers suggest a discussion topic to students who, in turn, present their reports. Depending on the topic, on the learners profiles, on the interest raised by the theme, forums may contain hundreds or only a few contributions. Some of these messages may present favorable arguments or opposing ideas concerning the topic of discussion. While some of the texts do not present relevant ideas concerning the topic, other contributions are meaningful as they deal with important concepts by either citing or relating them.

Participation in discussion forums is an important learning activity in distance education, allowing the teacher to identify important information about their pupils. For teachers, forum analysis is essential as a means to provide a closer view on how learners build their knowledge. Moreover, it allows teachers to verify the real contribution of each student, observe and interfere in the discussions. However, in the case of very large groups, analyzing students' posts can be quite time consuming. Thus, a content analysis technique may be helpful to teachers as to enable them to focus more on finding out why some students have not discussed important concepts related to the main topic.

A potentially useful intellectual tool for researchers is the ability to make connections between seemingly unrelated facts and as a consequence create inspired new ideas, approaches or hypotheses for their current work. This can be achieved through a process known as text mining.

6. ANALYSIS AND INTERPRETATION

Analysis of data involves styling the tabulated material in order to determine inherent facts or meanings. Analytical techniques have contributed greatly in gathering, organizing and interpreting text data. The processing of text data through analytical calls for competence in the use of analytics methods and for understanding of concepts that underline their development and their application.

The data collected have been analysed text mining with the help of R studio. The analysis of the study is divided into two parts namely preliminary analysis and major analysis.

Sno	Category	Replies
1	Post 1	11
2	Post 2	7
3	Post 3	5
4	Post 4	22
5	Post 5	6

Table 6.1

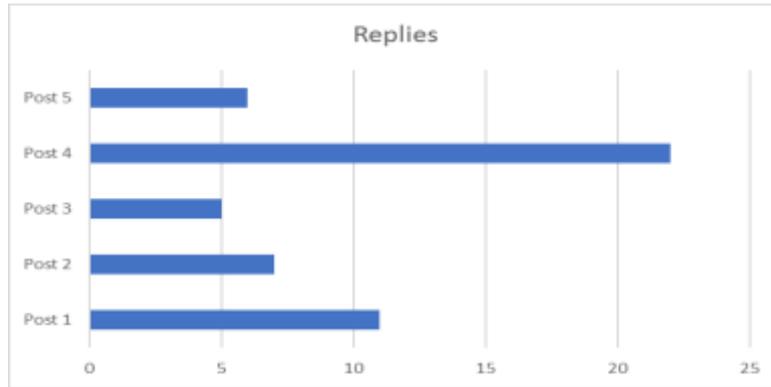


Fig 6.1

First finding the word frequency , word cloud and hierarchical clustering for the number of posts taken by sample in the college forum.

the most of the bold or highlighted words are the maximum possibility words or possible way the questions occur in the practical viva vise time. Use of the most repeated values are the most possible questions so the text mining used to represent the most asking questions and hints of the post.



Fig 6.2

The number of posts taken from the forum are concuded that the given word cloud diagram. The maximum number of users the posts are related to the education oreinted because the students concentrate their studies compared to the extra curricular activites so the teachers should note these from the researach then they concentrate those students and participate in the forum higher in the education side.



Fig 6.3

7. CONCLUSION

Interest in using online discussion forum as an indicator of students' performance is increasing and behavior. The initial experiments made for this paper show that text mining using graph representation can be applied in the creation of relevance indicators for forum messages. With such analysis at hand, students have a tool that may help them follow relevant contributions made by learners in discussion forums. There is further scope for study on Student and faculty in the forum conversation.

8. REFERENCES

- [1] A. Clauset, M.E.J. Newman, C. Moore: Finding community structure in very large networks, <http://www.arxiv.org/abs/cond-mat/0408187>
- [2] A.M. Kaplan, and M. Haenlein, "Users of the world, unite! The challenges and opportunities of Social Media", *Business Horizons*, 2010, 53, 1, pp. 59–68
- [3] BrenoFabrício Terra Azevedo, Patricia Alejandra Behar and Eliseo BerniReategui, *Qualitative Analysis of Discussion Forums*, Dynamic Publishers Inc., USA, April 2011
- [4] C. E. Cutrona and J. A. Suhr, "Controllability of Stressful Events and Satisfaction With Spouse Support Behaviors," *Communication Research*, vol. 19, pp. 154-174, 1992.
- [5] De Wever, B., Schellens, T., Valcke, M. and Van Keer, H, 2005, "Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review", *Computers & Education* 46 (1), pp. 6–28.
- [6] Dustdar, S.; Gall, H., *Architectural concerns in distributed and mobile collaborative systems, Parallel, Distributed and Network-Based Processing*, 2003. Proceedings of the Eleventh Euromicro Conference, pp.475,483, 5-7 Feb. 2003
- [7] Dracula Graph Library, 2010. Dracula Graph Library, [Online], Available: <http://www.graphdracula.net> [Oct 2011].
- [8] E. Wenger, *Community of practice*. Cambridge university press, 1998.
- [9] F. Antonelli and M. L. Sapino, "A rule based approach to message board topics classification," presented at the the 11th international conference on Advances in Multimedia Information Systems, 2005.
- [10] G. Somprasertsri and P. Lalitrojwong, "Mining Feature-Opinion in Online Customer Reviews for Opinion Summarization," 2010.
- [11] G. Eysenbach, "Medicine 2.0: Social Networking, Collaboration, Participation, Apomediation, and Openness," *J. Med. Internet Res*, vol. 10, p. e22, 2008.
- [12] K. Cai, S. Bao, Z. Yang, J. Tang, R. Ma, L. Zhang, et al., "OOLAM: an opinion oriented link analysis model for influence persona discovery," presented at the Proceedings of the fourth ACM international conference on Web search and data mining, Hong Kong, China, 2011.
- [13] Krebs, V. *Social Network Analysis*, [Online], Available: <http://www.orgnet.com/sna.html> [Oct 2011]

- [14] L. Rourke, T. Anderson, D. R. Garrison, and W. Archer, "Assessing social presence in asynchronous text-based computer conferencing", *Journal of distance education*, 14(2), 1999.
- [15] L. Hong and B. D. Davison, "A Classification-based Approach to Question Answering in Discussion Boards," presented at the the 32nd international ACM SIGIR conference on Research and development in information retrieval (SIGIR'09), Boston, Massachusetts, USA, 2009.
- [16] Luhn, H.P., 1958. The automatic creation of literature abstracts, *IBM Journal of Research and Development*, 2, 159-165.
- [17] M. Lipman, *Thinking in education*, Cambridge University Press, 1991.
- [18] M. Slater, "A note on presence", *Presence connect*, 3(3), 2003
- [19] Naidu, S and Jarvela, S, 2006, "Analyzing CMC content for what?", *Computer & Education* 46, pp. 96-103.
- [20] NetworkRepository.com - The First Interactive Data Repository, <http://networkrepository.com/>
- [21] Scalar Vector Graphics, 2011. SVG 1.1 2nd Edition, [Online], Available: <http://www.w3.org/TR/SVG> [Nov 2011].
- [22] Scott, J. *Social Network Analysis*, *Sociology*. February 1988 22: pp.109-127. doi:10.1177/0038038588022001007
- [23] Spaho, E., Kulla, E., Xhafa, F. and Barolli, L. P2P Solutions to Efficient Mobile Peer Collaboration in MANETs. *3PGCIC 2012*: 379-383
- [24] Spatariu, A., Hartley, K. and Bendixen, L.D. 2004, "Defining and measuring quality in on-line discussion", *Journal of Interactive Online Learning*. Vol 2 (4).
- [25] Stephen P. Borgatti, *Centrality and network flow*, *Social Networks*, Volume 27, Issue 1, January 2005, Pages 55–71, Elsevier, 2005.
- [26] S. Kim, "Content analysis of cancer blog posts," *J Med LibrAssoc*, vol. 97, pp. 260-266, 2009.
- [27] Xhafa, F., Kolicic, V., Potlog, A., Spaho, E., Barolli, L. Takizawa, M. *Data Replication in P2P Collaborative Systems*. *3PGCIC 2012*: 49-57