Text Modeling in Adaptive Educational Chat Room Based on Data Mining

Nidal A.M Jabari

Palestine Technical University, Kadoorie

Abstract: *Through the previous studies I have developed a new methodology for adaptive virtual education management. A basic part of virtual learning is that the students should have a discussion on a particular subject with each others. As this process is a part of the educational process, there is a need to evaluate this discussion and provide a numeric assessment. The problem with the methodology is the size of the words used in the text of the dialogue, which create difficulty in calculating the evaluation. This paper developed a solution to reduce the text to a minimum, which can enable to extract the appropriate assessment of the text. The paper also suggested the possibility of using machine learning to deal with the unknown words in the text .*

Keywords: *Adaptive, chat, learning, AVCM, text, modelling.*

1. Introduction

The virtual classroom model based on student modeling and course sequencing was an intelligent model for adaptive virtual classroom [4]. This model offers online presentation, online testing and online peer to peer chat . Student modeling is a methodology to extract the student characteristics and evaluation process within the virtual classroom . In [5] also, it was discussed the student modeling in the adaptive educational chat room by itself including text modeling , time modeling and peer evaluation. This paper will discuss the text modeling in the educational adaptive chat room as one component of the student modeling in educational chat room based on data mining. Text model means extracting the evaluation criteria which we may use to evaluate the text and expressions that the student used while chatting with his peer in the virtual classroom. The researchers in [6] discussed this approach and extract a mathematical equation to evaluate the text used by two peers in an educational chat room . But in our research we are looking for minimizing the degree of complexity of the used text by mining the text into the useful keywords which are discussed in [6] .

2. Background and Theoritical Framework

2.1. AVCM

AVCM [4] is a virtual classroom model which applies adaptive e-learning, It contains of three main facilities: 1. Adaptive presentation 2. Adaptive testing and 3. Adaptive chat Student modeling is crucial for an intelligent learning environment to be able to adapt to the needs and knowledge of individual student [7]. AVCM student modeling process relies mainly on two methods: Static Student Modeling (Stereotype) and Dynamic modeling (Overlay Model). The AVCM overall architecture is shown in figure 1
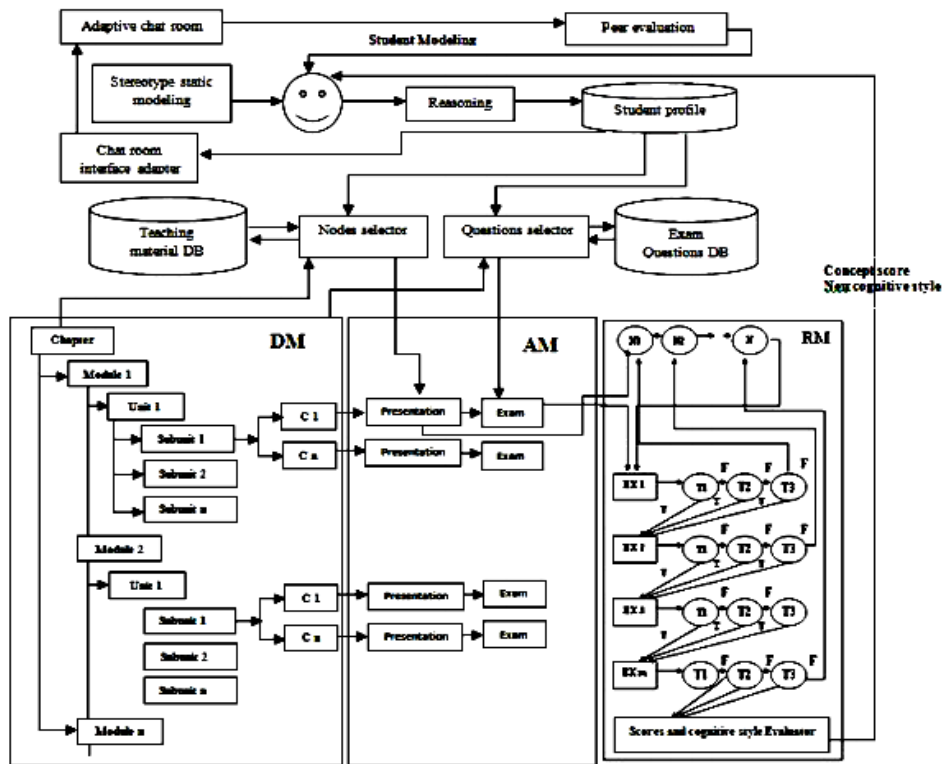
Fig. 1: AVCM Overall Architecture

2.2. AVCM Chat Model

Chat is a peer-to-peer service for communication with others [4]. Chat room in AVCM is an additional separated system . AVCM Chat tool is an adaptive chat room which can assist the student to select the best peer to chat with in an educational chat session. The tool can give evidence about the student knowledge level about one concept which is matter of discussion. At the same time it is a collaborative tool which can be used to model the students during the chat session . AVCM chat room is shown in figure 2.

Fig. 2: AVCM chat room

Hanini etl in [6] discussed the text modeling process in an educational chat room . They introduced a model for monitoring the text and the expressions used by the student with his colleagues during the educational chat session about a particular topic assigned to be discussed during a particular session. Monitoring is done through text modeling process by evaluating the expressions used by the student within the session time. This method is used to improve the general idea of the educational Chat among students and make it formalized which will lead to make the students behave more seriously during the session chat. Experiments performed as well as previous

studies have produced mathematical equations based on the parameters extracted by analyzing and mining the messages submitted by the peers. Text model will be added to the time and peer model to construct the final student model in the educational chat room. Hanini etl , extracted nine levels for student modeling as seen in table 1 .

TABLE I: levels of text modeling levels in chat room

Category	Level	Target Tag
Main concept	Level 1	Self
Related directly to the main concept	Level 2	Self
Useful Words	Level 3	Self
Positive Expressions	Level 4	Both
Agreement Expressions	Level 5	Both
Enquiry Expressions	Level 6	Both
Respect Expressions	Level 7	Self
Negative Expressions	Level 8	Peer
Unused Expressions	Level 9	Both

The text model described and ensures that the student used his chat expressions efficiently in the chat session. Chat expressions are from several levels and each level should have rate of occurrences. Two of these levels are related directly to the concept of chat while the others are general words affect the chat session quality.Using these levels and applying two experiments they got some parameters.

Using these levels Hanini etl extracted the following equation (Eq.1) for text modelling in the chat room .

Text Model Score = -6.91+(0.9*level 1)+(0.14*level 2)+(0.11*level 3)+(0.06*level 4)+(0.06*level 5)+ (0.18*level 6)+(0.02*level 7) .... (1)

Eq 1: Text model mathematical calculation

2.3. Text Mining Introduction

Text mining is the process of mining the useful information from the text documents. It is also called knowledge discovery in text (KDT) or knowledge of intelligent text analysis. Text mining is a technique which extracts information from both structured and unstructured data and also finding patterns. Text mining techniques are used in various types of research domains like natural language processing, information retrieval, text classification and text clustering[2]. Text Mining process is shown in Figure 3.
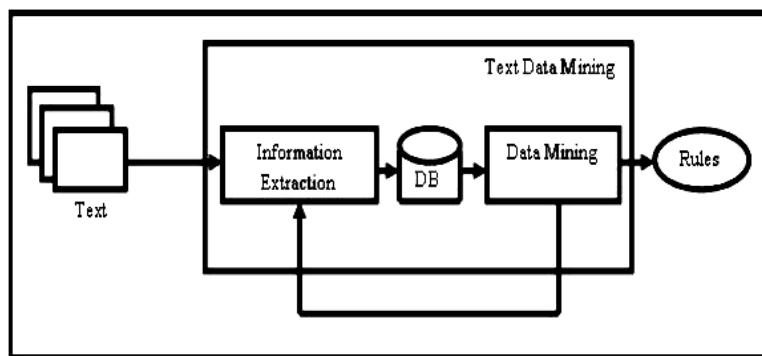

Fig. 3 : Text Mining process

In the most general terms, text mining will "turn text into numbers" (meaningful indices), which can then be incorporated in other analyses such as predictive data mining projects, the application of unsupervised learning methods (clustering), etc [3].

There are many methods for extract features from text data (Specially for Web and Email Categorization) you can find in the literature. There is also some beginning steps you can follow:

- Trimming Vocabulary: Refers to the process of removing "non-content".
- Stemming: Refers to reducing all variants of a word to a single term.
- Define Classes: Refers to defining how many classes do you have?.

3. Text Modeling in Adaptive Educational Chat Room Based on Text Mining

Text mining is a technique based on data mining which will be used in AVCM chat room to analyze the text used by the two peer students in one session . Referring to Eq. 1 , we need to extract the number of the keywords occurrences in the chat text used by the students . Each level will give positive or negative number to the overall student score as shown in Figure 1. If the extracted word is not from the nine categories it means the system should use machine learning to add it to the suitable level as shown in Figure 4.
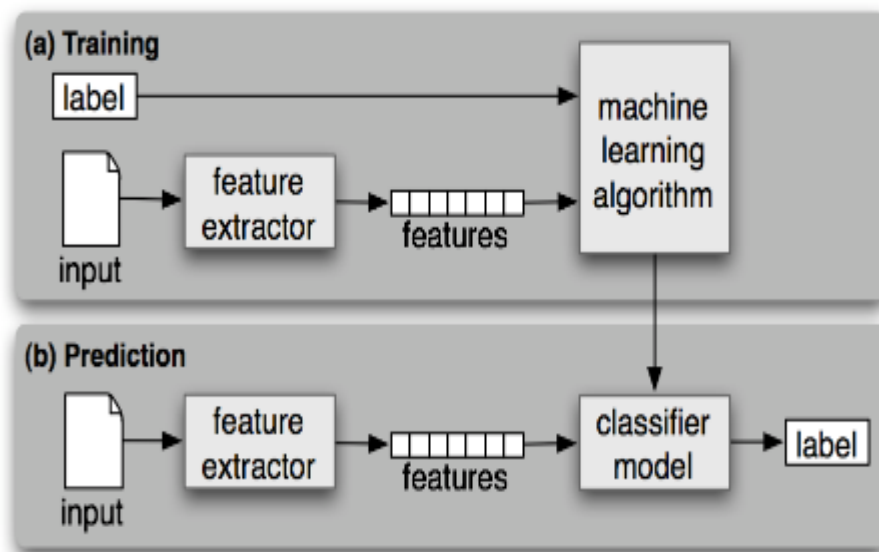

Fig. 4: Text feature extraction

Our approach is based on figure 2 , Table 1 , equation 1 , figure 3 and figure 4 .In order to model the text , it should be mined first to the minimum degree . This way will reduce the amount of words which will be a matter of evaluation by the nine levels .The final result we need is a numeric score calculated by adding the score of the nine levels.

The methodology of text modeling is summarized as shown in Figure 5. The algorithm is:
1. convert used text into text file Then the system should remove "non-content" words .
2. remove "non-content" words
3. reduce all variants of a word to a single term
4. create the new reduced file
5. start extracting keywords
6. Feature extraction of keywords
7. Searching keyword in the keywords database
8. If exist Define class of the keyword
    a. If Not exist
    b. Use machine learning to add new keyword to the database assigned to a level .
9. State the level of the keyword
10. Counts Keywords in each level
11. Find the student score by using equation 1 .
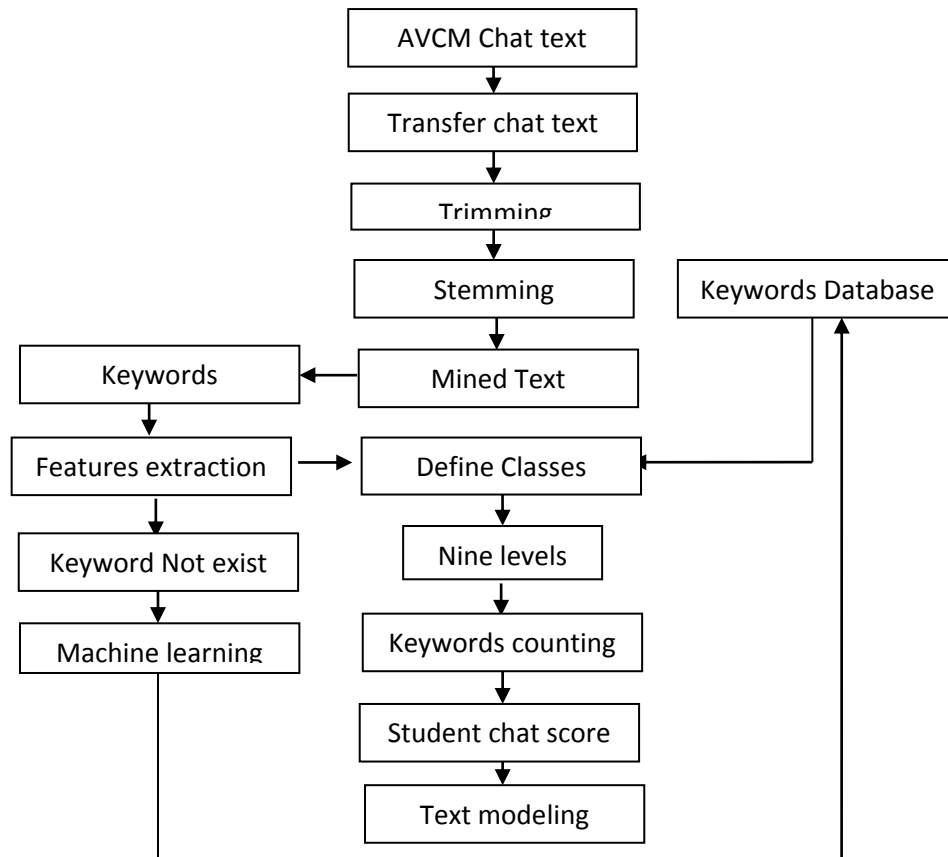
The full procedure is described in figure 5.

Fig. 5: Text Modelling in adaptive educational chat room based on data mining

4. Conclusion

The texts used in the chat between two students are large and many expressions are usually used . Many of them are not useful in the evaluation process. It is very important before the evaluation process starting is to extract the words that have meaning. Therefore, the use of trimming data will help to reduce the amount of words being evaluated. Many of the words used will not be included in the keywords that the teacher collects and inserts, making it necessary for the system to acquire these words and put them in the correct level to be used later in the evaluation process.

5. References

[1] Konstantina Chrysafiadi and Maria Virvou , "Student modeling approaches: A literature review ", *Journal of Expert Systems with Applications* ,Volume 40, Issue 11 , Pages 4715-4729,1 September 2013

https://doi.org/10.1016/j.eswa.2013.02.007

[2] S. Vijayarani1and J. Ilamathi ," Preprocessing Techniques for Text Mining - An Overview" , Ms., Department of Computer Science, School of Computer Science and Engineering, Bharathiar University, Coimbatore, Tamilnadu, India1, 2, 3 (2015)

[3] George Forman and Evan Kirshenbaum , "D. Extremely Fast Text Feature Extraction for Classification and *Indexing Hewlett-Packard Labs Palo Alto*", CA, USA Hewlett-Packard Labs Palo Alto, CA, USA

[4] Nidal A.M Jabari, Mochammad Hariadi and Mauridhi Hery Purnomo "Intelligent Adaptive Presentation and E-testing System based on User Modeling and Course Sequencing in Virtual Classroom" , *International Journal of Computer Applications* (0975 – 8887) Volume 50 – No.9, July 2012.

[5] Muna Hanini , Radwan Tahboub , Nidal Jabari , "Student modeling in adaptive educational chat room" , „*Journal of theoretical applied information technology* ,2014 .

[6] Muna Hanini, Nidal A.M Jabari , Radwan Tahboub "Text Modeling in Adaptive Educational Chat Room" , *International Journal of Computer Applications* (0975 – 8887) Volume 103 – No.5, October 2014 .

[7] Raymund Sison, Masamichi Shimura, "Student Modeling and Machine Learning " , *International Journal of Artificial Intelligence in Education* (1998), 9, 128-158.