Mining Changes of Opinions Expressed by Students to Improve Course Evaluation

Alaa El-Halees
Faculty of Information Technology, Isalmic Univerity of Gaza, Gaza, Palestine

Abstract: *Opinion mining can be used in many applications. In universities, students' opinions about courses can be considered as a significant informative resource to improve the effectiveness of education. Past works in this area focused on direct mining of students' opinions in regard to the courses. The aim of this paper is to develop a system which detects changes of students' opinions. Understanding such changes can help the management improve course evaluation in academic institutions. For course evaluation, knowing what is changing and how it has changed is crucial as they allow the management to provide the right course features such as teachers, contents, teaching materials and exams which need to satisfy the students' needs. In our work, we present a strategy for mining opinion changes based on the associative classification approach. Firstly, we collect opinions from students in two different semesters in regard to a specific course. Then, we extract rules using association rules. For this purpose, we detect and measure students' change of opinion from one semester to another. We describe types of opinions which can be detected by the students. Finally, we shed light on some of the examples which we have spotted from each type of opinion change.*

Keywords: *Educational Opinion Mining, Change Opinion Mining, Higher Education Evaluation.*

1. Introduction

Opinion mining is a research subtopic of data mining which aims to automatically obtain useful knowledge in subjective texts. It has been widely used in real-world applications such as e-commerce, business-intelligence, marketing, tracking political topics, education and others [1][2][3].

One important application of opinion mining is about using it to improve education in the academic institutions. In previous works, we proposed a model to extract knowledge from students' opinions to improve the teaching effectiveness in academic institutions [3][4]. One of the major academic goals for any university is to improve teaching quality, that is because many people believe that the university is a business-oriented entity and that the responsibility of any business is to satisfy its customers' needs. In this case, the university customers are the students. Therefore, it is important to reflect on students' attitudes to improve teaching quality [3][4].

Much of the opinion mining research has been focused to discover rules from subjective data, and little attention has been paid to mining changes in these data over time. Opinion change mining is the process of monitoring and detecting possible changes to specific opinions over any given period of time. For businesses, knowing what is changing and how it has changed is of importance because they allow businesses to provide the right products and services to go with the changing market needs. If negative changes are detected, corrective measures need to be taken to stop or to delay such changes [5][6].

In this work we continue to use opinion change mining to improve quality of teaching. An improvement in quality should be followed by an action. Opinion change mining can be used to allow the universities to compare results of their course evaluation in regard to some key features, (e.g. teachers, courses, exams, etc.) during two intervals before and after the quality action takes effect, to see whether there was any improvement in these key course features.

Given our keenness to achieve course improvement, we developed a system structure for opinion change mining. The system can be used to detect and analyze opinion change of students. In order to do that, we need first to collect comments of students in relation to the course features to create user-generated information in the different semesters. A course feature is one key aspect of a specific course such as contents, teacher, ..etc. We will then determine the attitude of students towards the feature (i.e. if the student likes or dislikes the feature). In this step, associative classification of sentiment analysis will be used. After generating association rules from the two datasets, we generate opinion rule sets, and then identify change rules from the datasets.

The rest of the paper is structured as follows: the second section discusses all the related works, the third section addresses opinion change mining background, the fourth section describes the proposed system, the fifth section contains a description of the experiments by giving sample results of experiments, while the sixth section implies the conclusion of the paper.

2. Related Works

There has not been any previous work in relation to using opinion change mining for educational purposes, however some similar works have been used in other domains. For example, Cheng et. al in [5] proposed an opinion change mining approach for reviewing customer mining and analysis. The objective of their work was to automatically discover the trends of online reviews. They provided product decision makers with explicit information which enable them to rapidly understand changes in consumer opinions as demonstrated by the discovered change patterns. The proposed method combined the associative classification methodology with the overall rating to discover the correlation between the different features.

Akcora et. al. in [7] presented a way to monitor public opinion in terms of the temporal dimension. Their methods can identify break points and find related events that caused these opinion changes. They proposed to use micro-blogging sites to capture developments in public opinions. Their experiments showed that the proposed product was able to find out changes in public opinion and extract the key views about the events.

Amiri and Chua in [8] mined the sentiment terminology which includes the detection of new opinion words as well as understanding their polarities. In their paper, they proposed an approach based on the interchangeability characteristic of words to detect new opinion words through time. Thereafter, they discovered that the current non-time-based polarity inference approaches may assign an opposite polarity to the same opinion word at different times. They proposed the method to find new opinion words through time and utilized Dempster-Shafer theory to obtain a time accumulated polarity for each new opinion word.

Fukuhara et. al. in [9] proposed a method for analyzing temporal trends of sentiments and topics from texts with timestamps. The method accepts text with timestamp such as weblogs and news articles and produces two kinds of graphs; topic graph that shows temporal change of topics associated with a sentiment, and a sentiment graph that shows temporal change of sentiments associated with a topic.

3. Background

Opinion change mining is the process of detecting possible changes to specific opinions over any given period of time. It can be seen as an extension of opinion mining [5]. It takes its basic concepts from change data mining. This section describes the basic concepts of opinion mining and its change aspects. Also, we will discuss association rules and association rule based classification which is the method that will be used to detect changes in opinions. Finally, we talk about types of changes in opinions and how to measure the relevant changes.

3.1 Opinion mining

Opinion mining is the process which enables the system to identify opinions from language texts written manually by humans [10]. Opinion mining is a subtopic of text mining that classifies an opinionated document as expressing positive or negative indication. It is also commonly known as sentiment analysis. It aims to find the general sentiment of the author in an opinionated text [11]. For example, in educational data, a student may express his/her opinion about a course using discussion forums. Opinion classification determines whether the student attitude is positive or negative about that course [3].

In opinion mining, a set of user-generated content review is formally referred to with the symbol R by students expressing opinions about a course, opinion classification aims to classify each review $r \in R$ to determine whether the review is positive or negative [3].

3.2 Change mining

Change mining can be defined as the activity of looking for interesting correlations or patterns in large sets during different time periods. Cheng et. al in [5] proposed the following change mining steps: First of all, calculating the degree of similarity in the conditional parts of two rules for different time periods. Secondly, for each rule at any given time t, it can find the most similar rule for time $t + k$, and finally, classify the type of change for the rules using the maximum similarity value and the difference measures for the conditional part.

There are many ways to detect change mining, one common method is use to association based on classification.

3.3 Association Based on Classification

Association rule mining was first introduced by [12]. It aims to extract a set of association rules from a given dataset. Association rule describes an implicative co-occurring relationship between two binary-valued sets in transactional dataset items, expressed in the form of an "antecedent $\Rightarrow$ consequent" rule. Where antecedent is the frequent item set driving the rule and the consequent is the frequent items included in a transaction due to the antecedent.

Classification problems can be solved by the association rule mining in the form of "antecedent $\Rightarrow$Y", where the antecedent includes a set of terms and Y is a class label. Several methods have been proposed and applied in classification problems such as Classification Based on Associations (CBA) [13]. CBA rule is called *ruleitem* and has received a much wider support than *minsup*. The support count of the antecedent part (called *condsupCount*) is the number of cases in dataset that contain the condition. The support count of *ruleitm* (called *rulesupCount*) is the number of cases in a dataset that contain the condition and are labeled with class Y.

3.5 Association based on classification opinion mining

We proposed to use association based on classification as basic components of the classification model. These rules have the form $X \Rightarrow c$, where X is a combination of *[feature, opinion]* pairs where features of the course (*e.g.* Teacher, Grades, ..etc). and opinion about this feature (e.g. "Good", "Easy",….). *c* is the overall semantic orientation (i.e., positive or negative) of the generated rule.

4.0 SYSTEM ARCHITECTURE

To mine changes of opinions from students to be able to improve the course, we propose a system as shown in figure 4.1.

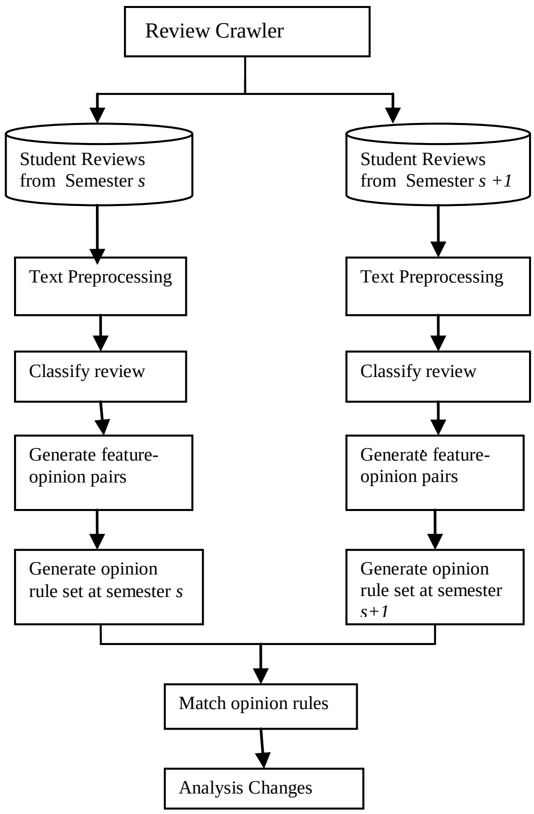

Figure 4.1: System Architecture to Mine Changes of Opinions of Students to Improve Course

The system has the following parts:

1) Given a specific course at a specific time interval (e.g. any given semester) extending from the course start date to the end date, the system first collects the related reviews and places them in the review database. Then, again for the same course and for other semester time intervals (mostly the next one) , it collects the reviews and puts them in another database.

2) A preprocessing step is done for both databases. The preprocessing module contains five steps which are: Firstly, tokenization, which breaks down the text by splitting it to spaces and punctuation marks and forms a bag of words. Secondly, it applies stop words removal by removing common words such as articles. Thirdly, light stemming is applied. Fourthly, vector representations are made for the terms from their textual representations by performing TFIDF weight (term frequency–inverse document frequency). Fifthly, low frequency of occurring term is removed.

3) Classify each review: following the work explained in [3], each review is classified as positive or negative.

4) Generate feature-opinion pairs: The purpose of this step is to extract the candidate feature-opinion from the parsed reviews. Mostly we can find most course features indicating words are nouns or noun phrases. Using a part- of-speech (POS), noun and noun-phrase as a possible features set. If the feature is repeated in the reviews we choose it as a selected feature. Therefore, if a review contains any frequent feature, then we extract the adjacent *adjective*. If such an adjective is found, it is considered an opinion feature.

5) Generate opinion rules: using FP-Growth algorithm to mine student behavior patterns. The minimum support and minimum confidence of

the association rules are set to *minsup* and the frequent item set is assumed to include up to *maxitems*.

6) Match opinion rules by calculating the degree of similarity in the conditional parts of two rules for different time periods. For each rule at time s, we need to find the most similar rule for time *s + 1*.

7) Analysis changes: Classify the type of change for the rules and measure change factor using the formula 4.1 from [5] :

$$\left| \frac{\sup(r_i^{t1}) - \sup(r_j^{t2})}{\min(\sup(r_i^{t1}), \sup(r_j^{t2}))} \right| \quad (4.1)$$

To detect change of opinions in patterns between the two datasets in different time periods

5. Experiments and Results

To evaluate our method, a set of experiments was designed and performed. In this section we describe the experiments design including the corpus, the preprocessing stage, the used data mining method and types of generated rules. Afterwards, we show some examples of the generated rules.

5.1 Corpus

W collected data for our experiments from three discussion forums which are designed to discuss courses in a university. The language of the forums is Arabic. We focused on the content of five courses including all threads and posts about these courses. Table 5.1 shows a summary of the extracted data. Data details for each selected course are given in tables 5.2 and 5.3.

Table 5.1: A summary of the used corpus

Number of posts	342
Number of Statements	6210
Number of courses	5
Number of positive posts	227
Number of positive posts	115

Table 5.2: Corpus used for the first semester

Course To Review	No. of Posts	No. of Sentences	No. of Words
Course_1	71	1920	13228
Course_2	35	1321	7280
Course_3	24	617	3587
Course_4	21	524	3183
Course_5	20	635	3407

Table 5.3: Corpus used for the second semester

Course To Review	No. of Posts	No. of Sentences	No. of Words
Course_1	57	1802	10122
Course_2	26	812	4212
Course_3	28	705	3240
Course_4	37	1407	8112
Course_5	23	750	2914

5.2 preprocessing

After we collected the data associated with the chosen five courses in two successive semesters, we striped out the HTML tags and non-textual contents. Then, we separated the documents into posts and converted each post into a single file. For Arabic scripts, some alphabets have been normalized (e.g. the letters which have more than one form) and some repeated letters have been cancelled. Thereafter, the sentences are tokenized, stop words removed and Arabic light stemmer applied. We obtained vector representations for the terms from their textual representations by performing TFIDF weight (term frequency–inverse document frequency) which is a well known weight presentation of terms often used in text mining [14]. We also removed some terms with a low frequency of occurrence.

5.3 Methods

In our experiments to classify reviews, we applied a machine learning method, which is Support Vector Machine from [17]. To extract feature-opinion pairs, we used the Arabic part-of-speech system from [15]. In each statement, we selected noun and noun phrases as features and their adjacent adjective as its opinion. We used FP-Growth methods from Rapidminer [16] to generate association rules with minsup equals 5% and maxitems equals *4*.

5.4 Change Rule types

Six types of rules can be generated from the proposed systems which are:

Type I: Rules that increase positive opinions which means that the rule from the old data has positive support while the support increases in the new rules .

Type II: Rules that reduce positive opinions which means that the rule from the old data has positive support while the support decreases in the new rules.

Type III: Rules that increase negative opinions which means that the rule from the old data has negative support while the support increases in the new rules.

Type IV: Rules that decrease negative opinions which means that the rule from the old data has negative support while the support decreases in the new rules.

Type V: Rules indicating frequent positive to negative opinion changes which means that the rule from the old data has positive support while it changed to a negative support in the new rules.

Type VI: Rules that change from negative to positive which means that the rule from the old data has negative support while it changed to positive support in the new rules.

5.5 Results

In this section, we present experimental results of our methods. Table 5.4 gave the number of rules generated for each rule type for course 1 as an example.

Table 5.4: Number of rules generated for course 1 in each rule type.

Type	No. of Rules
Type I (pos → pos $^{+}$)	41
Type II (pos → pos $^{-}$)	24
Type III (neg → neg $^{+}$)	15
Type IV (neg → neg $^{-}$)	33
Type V (pos → neg)	11
Type VI (neg → pos)	25

We can conclude that the results for this course in general have been in the right direction. That is because many negative opinions changed to positive (25) and many positive opinions developed and became more positive (41). However, many negative opinions became more negative (33) as well.

In addition, table 5.5 gives examples for each type of the generated rules. From the table we can notice that in some courses the exams became more positive, but the change pattern was not so evident (about 0.072).

That means that the positive attitude of students toward *exams* is almost the same in both semesters. On the other hand, students' attitude towards *exams* and *teacher* became positive with a change rate of 0.32. Further we can notice that a big change (about 11.9) was observed in terms of moving from a positive to negative support for the course textbooks. The foregoing results show that a wrong decision has been taken in regard to changing the course textbook in this semester in comparison with the other semester.

Table 5.5: Examples for each type of generated rules.

Type	Example of rule at semester s	Support of semester s	Example of rule at semester s+1	Support of semester s+1	Change of rule factor
Type I	exam -> Pos	0.179	exam -> Pos	0.192	0.072
Type II	exam ,teacher -> Pos	0.139	exam ,teacher -> Pos	0.105	0.32
Type III	Subject, teacher-> neg	0.011	Subject, teacher-> neg	0.103	8.36
Type IV	Subject->neg	0.201	Subject->neg	0.111	0.81
Type V	Book->pos	0.120	Book->neg	0.011	11.90
Type VI	Teacher->neg	0.012	Teacher->pos	0.029	3.41

6.0 Conclusion

In this paper, we have described our approach of mining change of opinions from educational data to help the management structures improve course evaluation in academic institutions.

We proposed a system that takes students opinions during two semesters in regard to a certain course, used text preprocessing methods to classify reviews to positive and negative opinions, generate feature-opinion pairs, generate opinion rule set, and then match opinion rules. At last, we recognized six types of changes that may occur to the opinions of the students. According to these rules the management structure may decide to improve the course by changing teachers, exams , ..etc for instance.

In future work, many improvements may be adopted on the proposed system such as automatically extracting educational features, summarizing the results, graphing the analyzed results, and detecting changes occurring in time intervals exceeding the two semesters period.

References

[1] B. Pang and L. Lee. "Opinion Mining and Sentiment Analysis". In Information Retrieval Vol.2,No.1,21–135, 2008.

[2] M. Hu, and B. Liu. "Mining and Summarizing Customer Reviews". Proceedings of the 10th ACM SIGKDD international conference on Knowledge discovery and data mining, Seattle, WA, USA, 2004

[3] A. El-Halees. "Mining Opinions in User-Generated Contents to Improve Course Evaluation", The Second International Conference on Software Engineering and Computer Systems (ICSECS 2011) , Pahang, Malaysia, 27-29 June 2011-05-28.

[4] A. El-Halees. " Mining Feature-opinion in Educational Data for Course Improvement". International Journal on New Computer Architectures and Their Applications (IJNCAA) 1(3): 1101-1110 The Society of Digital Information and Wireless Communications, 2011 (ISSN: 2220-9085)

[5] L. Cheng; Z. Ke; B. Shiue, "Detecting changes of opinion from customer reviews," In Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2011 vol.3, no., pp.1798,1802, 26-28 July 2011

[6] A. Chiu, and H. Chang. "Mining changes in customer behavior in retail marketing" . An International Journal of Expert Systems with Application. Vol. 28, No. 4, pp. 773-781) May 2005.

[7] C. Akcora, M. Bayir, M. Demirbas, and H. Ferhatosmanoglu. "Identifying breakpoints in public opinion" . In Proceedings of the First Workshop on Social Media Analytics (SOMA '10). ACM, New York, NY, USA, 62-66.

[8] H. Amiri and T. Chua. "Mining sentiment terminology through time". In Proceedings of the 21st ACM international conference on Information and knowledge management (CIKM '12). ACM, New York, NY, USA, 2060-2064.

[9] T. Fukuhara, H. Nakagawa, and T. Nishida "Understanding sentiment of people from news articles: Temporal sentiment analysis of social events" . Proceedings of the International Conference on Weblogs and Social Media ICWSM, (2007).

[10] L. Xia, , A.L. Gentile, J. Munro, and J. Iria" Improving Patient Opinion Mining through Multi-step Classification." In: Matoušek, V., Mautner, P. (eds.) TSD 2009. LNCS, vol. 5729, pp. 70–76. Springer, Heidelberg (2009)

[11] B. Pang and L. Lee. " Opinion Mining and Sentiment Analysis" . Foundations and Trends in Information Retrieval Vol. 2, issue 1-2 (January 2008).

[12] R. Agrawal, T. Imielinski, , and A. N Swami,. 1993. "Mining association rules between sets of items in large databases". In Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, P. Buneman and S. Jajodia, Eds. Washington, D.C., 207-216.

[13] B. Liu, W. Hsu, and Y. Ma. "Integrating Classification and association rule mining" . In KDD '98, New York, NY, Aug. 1998.

[14] G. Salton and C. Buckley. "Term-weighting approaches in automatic text retrieval" . Information Processing & Management 24 (5): 513–523. 1988.

[15] R.Toutanova and C. Manning. "Enriching the Knowledge Sources Used in a Maximum Entropy Part-of-Speech Tagger". In Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora (EMNLP/VLC-2000), pp. 63-70.

[16] http:/www.Rapidi.com

[17] C. Cortes and V. Vapnik. "Support-Vector Networks". Machine Learning, 20, 1995.