International Journal of Scientific Research in Computer Science and Engineering

Research Paper

ISSN 2320-7639

Volume-1, Issue-1, Jan- Feb-2013

Available online at www.isroset.org

Evaluating and Summarizing Student's Feedback Using Opinion Mining

Vaibhav Jain

Institute of Engineering & Technology, Devi Ahilya Vishwavidyalaya Indore, India

Received: 13 Jan 2013 Revised: 20 Jan 2013 Accepted: 07 Feb 2013 Published: 28 Feb 2013

Abstract- Collecting Students feedbacks for the subjects taught is a regular activity of an academic institution. Automating the process of collecting these feedbacks becomes an important requirement. This provides an opportunity to analyze these feedbacks efficiently and summarize the performance of a teacher in the subjects he taught. Opinion mining (or Sentiment Analysis) which is generally used for classifying customer reviews in terms of positive or negative sentiments can be used effectively in evaluating and summarizing the student's feedback.

Keywords: Opining Mining, Information Retrieval, Text Summarization, Text Mining etc.

I. INTRODUCTION

"What student thinks about a subject, a teacher's teaching ability and the curriculum?" is what a school or a college seeks to get from students in an academic calendar on regular basis. Manually collecting these feedbacks from a large group of students could consume a lot of resources. Generating a report from the students feedbacks collected could be equally challenging. Developing a system that can process subjective information effectively requires overcoming number of challenges such as categorizing and summarizing reviews/feedbacks given by a customer/student.

In recent time, Opinion mining (or Sentiment Analysis) attracted lot of attention of researchers because of its potential applications. Opinion mining seeks to identify the viewpoint underlying a text span; an example could be a student's response to a subject as whether he or she likes it or not. Everybody knows about the popularity of recommendation systems used by E-commerce companies which helps them and customers to get feedback about the products they are interested in. These systems us opinion mining to get summarized reviews of a product. However, it is a challenging task as opinions in natural language are often expressed in subtle and complex ways. Using simple text categorization techniques like n-gram or keyword identification technique may not be enough. Recognizing the semantic interpretation of the opinions may involve multiple levels. For example, a negative review may contain many positive phrases and vice-versa.

Opinion mining often uses support vector machines (SVM) or naïve bays (NB) models or their variants for text categorization. I am proposing to use SVMs which are well-known and powerful tool for classification of vectors of real-values features. On other hand, generating multiple-document opinion-oriented summaries poses a challenge. Determining which documents or portions of documents express the same opinion is not always an easy task. A bounded summary of opinions based on average or relative frequencies could be trivial but for unbounded summaries of opinions, visualization methods can be used.

This paper proposes an online feedback system which is about automating the process of recording student's feedback. The proposed system collects the feedback submitted by students and then classifies them as positive feedbacks or negative feedbacks using SVM classifier. Then, it generates a performance summary of a teacher for the subjects he or she taught in that academic session using a bounded summarization algorithm. In my view, the proposed system could be useful for the academic institutions and can also help teachers to understand their teaching performance in a summarized manner.

II. RELATED WORK

The term *opinion mining* coined by Dave et al. [1] and his idea was about generating a list of product attributes like quality, features etc. and aggregating opinions in form of poor, mixed and good opinions. According to him, the ideal opinion-mining tool would "process a set of search results for a given item, generating a list of product attributes, and aggregating opinions about each of them as poor, mixed or good". Pang and Lee [2] suggested an efficient technique for subjectivity summarization using minimum-cut framework. Hu and Liu [3] addressed how feature-based summaries can be generated from the customer reviews. Shein and Nyunt [4] further established the role of SVMs in sentiment

analysis by introducing an ontological based approach. Further, Mullen and Collier [5] suggested how SVMs can be used for text categorization where we have diverse information sources. Lerman et al. [6] evaluated different sentiment summarization algorithms and showed through their study that how sentiment informed summarized models could be useful over a baseline model.

III. Support Vector Machines For Classification

SVMs are a machine learning classification technique which uses a function called a kernel to map a space of data points in which the data is not linearly separable onto a new space in which it is. One important property of SVMSs is that their ability to learn can be independent of the dimensionality of the feature space. SVMs can handle high-dimensional input space since they use overfitting protection, which does not necessarily depend on the number of features; they have the potential to handle these large feature spaces.

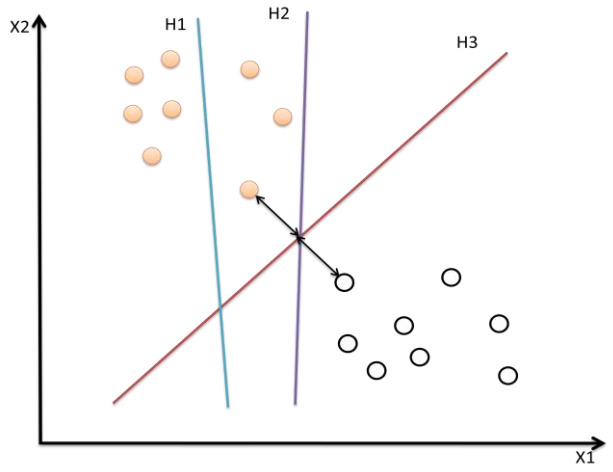

Fig. 1: SVM separating hyper planes

In SVM, a data point is viewed as a *p-dimensional* vector, and we want to know whether we can separate such points with a *(p-1)-dimensional* vector. There are many hyper planes that can classify the data. One reasonable choice as the best hyper plane is the one that represents the largest separation or margin between the two classes. Fig. 1 shows an example where we have two sets of classes. Here, H1 does not separate the classes. H2 does, but only with a small margin. H3 separates them with maximum margin.

SVMs can be used to solve various real world problems of uncertainty in knowledge-based systems. They can be used for hand-written character recognition and image classification. They can also be useful in text and hypertext categorization as their application can significantly reduce the need for labeled training instances in both the standard inductive and transductive settings. Most text categorization problems are linearly separable and SVMs are capable of identifying such linear separators. More details about SVMs to be used for text categorization are well addressed in [7].

IV. Student Feedback Format

Fig. 2 shows a sample feedback format. The feedback shown in the figure consists of two sections. In the first section, students give ratings to the teacher's performance in the course based on 10 points. The lowest rating i.e. *0* indicates the poor performance while the highest rating i.e. *5* indicates an excellent performance. The next section is for recording the student's comments on teacher's strengths, weaknesses and any general suggestions.

Points \ Ratings	1 Poor	2 ...	3 ...	4 ...	5 Excellent
Conceptual Clarity					
Motivation Provided					
Communication Skill					
Regularity & Punctuality					
Subject Knowledge					
Course Coverage					
Compliments Theory with Practical Skills					
Interaction & Guidance					
Computer/ IT Skills					
Overall Performance					

What are the Strengths of Teacher _____

What are the weaknesses of Teacher _____

Any Other Suggestion (syllabus, faculty, etc.)_____

Fig. 2: A Sample Feedback Format

V. Proposed Framework

I am proposing an online feedback system which would record the student's feedback and analyze the teacher's performance based on opinion mining using SVM classifier and then summarizes the teacher performance based on bounded summary algorithm i.e. counting frequency of positive and negative reviews.

Fig. 3 shows our proposed online feedback system based on opinion mining. First, a feedback scheduler would be responsible for scheduling the student's feedback for each course. In order to maintain the anonymity of a student, the scheduler would ensure that the feedback submission link would be up for may be 15-30 minutes based on the number of students registered in the course. On a given schedule, a link will be activated and the students of that course would record the time table of that class. Then, students would enter feedbacks of the subjects he or she had enrolled for

based on the schedule fixed using the feedback scheduler. A course manager would be maintaining the course information including list of courses taught by a teacher. The SVM classifier module would extract the opinions from the feedbacks collected and would classify them and pass it to the summary generator module. The summary generator module would take classified opinions and generate the summarized feedback report.

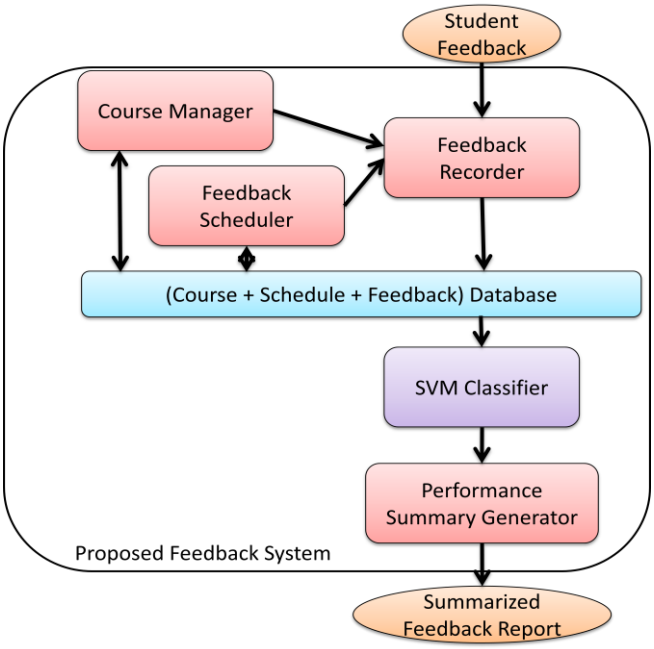

Fig. 3: Proposed Online Feedback System using Opinion Mining

VI. Performance Summary Generator

The final feature-based performance summary of each teacher for all subjects taught can be generated in the following way:

- For each identified feedback parameter, related opinion sentences can be put into positive and negative categories according to the opinion sentences orientation.
- All feedback parameter are ranked according to the frequency of their appearances in the reviews.

The following shows an example summary for the feature "communication skills" provided by a teacher.

Feature: **communication skills**

Positive: 9

- Overall his communication skills are excellent
- His expressions get us engaged throughout lecture
- He can do much better than this.

…

Negative: 2

- He should speak little loudly. Last benchers have listening issues.
- Poor skills.

VII. Discussion and Future Work

Automating the student's feedback could provide many benefits including saving cost, time and making efficient report generation etc. In my view, use of opining mining can help in summarizing the overall feedback report effectively and evaluating faculty performance in form of a summarized view could be useful for academic institutions. Teachers could also effectively review their own performance at this level instead of reviewing each student feedback individually.

Here, I have chosen SVM for opinion classification as it is one of the commonly used classifier in practice and its accuracy is equally good. I know that the feedback form shown in the Fig. 2 provides limited opportunity as opinions can be recorded for only section 2 of the feedback form.

For future work, we are planning to develop the proposed system and that would help us evaluating our proposed system and would definitely benefit the academic institutions. The proposed system can be extended to other areas like collecting feedbacks for workshops and short-term courses.

References

[1] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," in Proceedings of WWW, pp. 519–528, 2003.

[2] B. Pang and L. Lee, "A Sentimental Education: Sentiment Analysis Using Subjectivity", Proc. of ACL, pp. 271-278, 2004.

[3] M. Hu and B. Liu, Mining and summarizing customer reviews. In *Proc. of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (KDD '04). ACM, USA, pp. 168-177, 2004.

[4] K. P. Shein and T. S. Nyunt, "Sentiment Classification Based on Ontology and SVM Classifier," Communication Software and Networks, 2010. ICCSN '10. Second International Conference on, Singapore, pp. 169-172, 2010.

[5] T. Mullen and N. Collier, "Sentiment analysis using support vector machines with diverse information Sources," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 412–418, July, 2004.

[6] K. Lerman, S. Goldensohn, and R. McDonald. Sentiment summarization: evaluating and learning user preferences. In Proc. of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL

'09). Association for Computational Linguistics, USA, pp. 514-522, 2009.

[7] T. Joachims, Learning to Classify Text Using Support Vector Machines, Kluwer Academic Publishers, 2001.

