An Review of Student Sentimental Analysis For Educational Database Using Unsupervised Machine Learning Approaches

Laishram Kirtibas Singh
Research Scholar, Department of Computer Science, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai.
kirtibas9@gmail.com

R.Renuga Devi
Assistant Professor, Department of Computer Science, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai.
nicrdevi@gmail.com

Abstract

With the World Wide Web's growth at a relatively high rate, a huge increase in network communications has become inevitable. Online communication data includes feedback posted by students. Sentiment analysis, natural language processing, computational linguistics and text analysis, application recognition and text retrieval of information. Social media platforms are considered to be the most popular form of online communication. Through platforms such as the e-learning education system, multiple pieces of information that reflect student opinions and attitudes are published and shared among users every day. These platforms have recently focused on tracking and monitoring the reputation of their study time and student review and helping decision-makers and politicians assess their public opinion on policy and political issues. The opinions and emojis of understanding its important role in influencing student study opinion decisions are becoming more important for the educational field. It is possible to consider how emotions are used to advance e-learning intelligence tasks so that organizations can gain access to analyze education and detect unfavorable rumor risk management capabilities. Machine Learning is widely used for reputation analysis, and many sentiment analysis systems use unsupervised learning methods. This existing sentimental analysis for dictionary-based approaches is to provide a higher time complexity. The words are utilizing unsupervised learning based lexicon pattern methodology. The educational, sentimental pattern dependency factor between aspects and the sentiment word is considered. This approach evaluates the sentimental information to find the aspect and provides less complexity and a higher classification rate than existing methods.

Keywords: lexicon pattern methodology, higher time complexity, Machine Learning, unsupervised learning methods, sentiment analysis.

1. Introduction

Student feedback allows teachers to understand student learning behavior and improve education. Feedback allows students to highlight issues that may differ from the lecture. This happens when students do not understand some of the lectures or instances. Another example is the teacher's instruction is too fast or too slow. Feedback is usually taken in [24], and the end unit is more advantageous to its real-time.
Students discussed whether access to record lectures would be mandatory. Research on the effectiveness of using the traditional lecture records method (such as surveys and interviews) is very structured, but personal information is posted on checking student Facebook page. Therefore, social networks and sentiment analysis technology are to be used. Choose to use social networking and market sentiment analysis. These methods can be used to semi-structured and unstructured data useful to study social media. The overall survey results do not show that students generally agreed that, as a resource to supplement their speech recordings, they are using on-site seminars but exchange them [2].

Feedback is sent to the entity about past behavior from the statement to analyze future and current behavior to achieve the desired result. Feedback helps us to follow new knowledge and prevent repetitive mistakes. Education and learning play an important role. Many people want to know whether their opinions must be serious or the most important in the decision-making process in higher education quality. Feedback is the process of helping and assessing an organization on a monitor and standardizing the overall working environment [23].

Feedback analysis systems are an online and offline method available for student feedback in any institution. This is a comfortable way to find a request for comment given through various forms of documents and people in the feedback step and is a green place to get feedback analysis. This can be a powerful way to decorate and enhance scholars with acquired knowledge. Feedback is one of the most powerful impacts of learning and achievement, but this impact can also be [33] hugely negative.

2. Students Feedback Analysis

Feedback from students is very important as it helps the instructor understand the learning behavior of the students. Sometimes the instructor does not understand what the instructor is trying to explain, as the student can display this to the instructor by providing such feedback. Student feedback can understand another issue, and the student feedback also includes [37] not understanding the lecture's help.

Take the need for feedback to improve the lessons. If the student is not involved in providing feedback, there is no way to see if the lesson needs improvement. Students often act as classroom observers and expect instructor information to feed them. This is a matter of design for different teaching methods and international students with different backgrounds and experiences [38].

Student participation is important to education, and one way to measure it is through participation. The traditional way for students to ask questions is by raising their hands, but this method is not suitable for everyone, especially for a shy person. Studies on student participation showed that participation was higher than the method of raising hands and answering and that students voting were more popular than raising hands for learning. Another drawback of raising the hand is that the student can see other answers before copying with another student [39-40].

3. Machine Learning-based Sentiment Analysis

The machine learning method of reputation analysis uses a training data set that provides a training forecast model and evaluates the training model's performance in the test data set. It can
be further divided into supervised learning and unsupervised learning methods that do not require data sets and actual emotional labels. The polarization of text content is determined by combining the polarization of phrases with adjectives and adverbs.

The phrasal verb is used to identify the Pointwise Mutual Information (PMI) system's statistical information. Machine learning calculates the semantic approach of text that is not structured based on pre-analysis technology. A dictionary of emotions is generated using a semi-automatic polarization expansion algorithm. Reputation Analysis for unsupervised machine learning methods and training classification involves extracting feature words from the text. Text document tag set, classification requiring training.

Language features widely used in reputation analysis include Part Of Speech (POS) tags, punctuation, and emoticon representations. Machine learning methods predict emotions from student feedback. Extracted from feedback text [15], the model's use reported in training was n-grams.

4. Related Work

Valuable insight into the teaching and learning process of student feedback also provides an excellent mechanism. Implement and review student feedback as manual intervention is usually the means of laborious and difficult tasks. This task may be feasible, including short-term student feedback in small courses, but it is large, as it applies specifically to the general online course, MOOCs (Massive Open Online Courses). It is unrealistic for large cases. Therefore, CNN (Conventional Neural Network) can solve this problem and automatically analyze the framework for expressing student opinions in comments. In particular, structure-related MOOCs [1] rely on feature-level conceptual analysis and automated target recognition to express certain aspects of emotion and polarization.

The automated method, of course, provides visualization of student free-text annotations from satisfaction surveys. Focusing on emotions, the visual learning and education of these spirit beasts require improvement in either visualization or performed in all aspects of the curriculum. They offer simple and systematic education to monitor their curriculum and educational strategies to make the right decisions. [3] This can be achieved in several ways, but it is often recognized by closed-loop or standardized answers and questions, collecting additional free-text response fields and complimentary feedback on student satisfaction.

5. Contextual rankings for synonyms data analysis

Pattern mining has been extensively researched and has brought many applications into the real world over the past decade. High data requirements highlight this article and place the age of non-question systems in the definition of broadcast question and non-question systems. For this purpose, a system was proposed based on relevant rules, general rules, market sentiment analysis and finding public opinion directories on Twitter and the patterns of social networks in the non-query system. Association Rules were previously used for sentiment analysis. Still, once the sentiment analysis process is complete, they often appear to be associated with a particular sense,
indicating the token used in it. Meanwhile, they are also used to find patterns between emotions [4].

Most companies handle effective, quantitative feedback in practice, while quality feedback is manually manipulated or completely ignored. In this research, based on the oversight, the feedback processing system was based on the two-tier LSDM (long Short-Term Memory Network) model. The first-level predictive concept describes methods in a particular direction and its predictive methods [5] (positive, negative, and neutral).

Limited vocabulary knowledge of language learners leads to inaccurate wording. This is especially true when they try to express their emoji. Automated Composing Scoring System relies heavily on traditional thesauruses. Unfortunately, this does not provide proper advice for vocabulary selection. To assist English as a better second language, learners' word choice becomes a solution that provides contextual rankings for synonyms for emotional words. RESOLVE presents accurate emotional language about events in related areas. As the pattern is learned to capture emotional events, various factors are considered for the emotional word's scoring function [6] ranking.

Table 1 Previous Technique Advantages And Disadvantages

<table>
<thead>
<tr>
<th>Technique</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flipping Model [7]</td>
<td>A two-level marking strategy can provide more supervision information.</td>
<td>However, the polarity label in the polarity dictionary is not clear and correct. The dictionary may contain tag noise, and the polarity of the word varies according to its context.</td>
</tr>
<tr>
<td>Sentiment Convolutional Neural Network [8]</td>
<td>It is used to predict each word's emotional weight together with the score of sentence-level emotional bias.</td>
<td>This method does not take into consideration emotional information embedded in emotional words, which can introduce noise features as predictions</td>
</tr>
<tr>
<td>Machine Learning Algorithms [9]</td>
<td>Construct multiple emotion dictionaries, including the original emotion dictionaries, dictionary emoticons, and other related dictionaries</td>
<td>It takes more time to process the data.</td>
</tr>
</tbody>
</table>
The Support Vector Machine (SVM) of the attention model is used for classification tasks, such as sentiment analysis. Some words are more important than others in training the attention model. In this review work, a new functional model is trained in eye-tracking data on cognitive earth. First, the model builds the reading prediction in the context of another aspect of independent data using eye-tracking data. Predictive reading time Neuro-emotional analysis [10] is used to establish a cognitive focus ground layer.

Aspect-Based Sentiment Analysis (ABSA) User feedback and quest for preferences are essential to production and service. Many classical methods have been deep learning, as figure 1 shown in the previous literature, but some useful clues (e.g., context, vocabulary, grammar) has been sufficiently considered, instead. In this study, the ABSA method is to guide background, form, grammar and clues. First, a new subnet has been introduced to consider the overall context in which ABSA expresses its stated objectives. Second, vocabulary embedding is used to add additional vocabulary clues. Third, the interest-related focus is a new module that provides clues to the grammatical structure of comprehensive focus inference [11].

6. Lexicon based word polarity annotations
Several lexicons have been developed for reputation analysis, mostly related to word polarity annotations (e.g., positive/negative), but to build fine-grained sentiment analysis (e.g., happiness, sadness). Vocabulary attempts have become an important cause these days. They frequently learn models using CNNs (Convention Neural Networks) and RNNs (Recurrent Neural Networks) as building blocks for developing emotional perceptions and using them as baselines for comparing model performance. In both cases, embedding usually contains similar concepts. Nevertheless, it makes unmolded emotions and emotions, and the "good" and "bad" operators will be similar, despite their opposite polarities [12]. To determine the need to get the sentence context (soccer), bias word influences a particular domain's corpus. Then check if the context word price can accurately indicate the sentence price. [13] In the second part, four different existing emotional
dictionaries are compared from the first part with word scores and determine the price of sentences at different sensitivity levels in semantics and context.

The Viterbi Algorithm approaches text-independent phoneme segmentation at the sampling point level. The algorithm consists of two stages: First, the vocal cord vibration information in Electroglottograph (EGG) is used to detect the voice data's voiced part. The Hilbert envelope function is used to achieve a level of sampling accuracy for point detection. In guiding the decoder, telephone transcription is obtained by the voice recognition system, [14], where the voice audio stream is first decoded.

Sentiment Analysis (SA) focuses on unstructured textual data, such as product reviews and Weibo mining opinions (identification/classification). It is widely used in product reviews, political campaigns and market analysis and customer feedback. Using Unsupervised Machine Learning (SML) is an important method based on the database methods to learn the defined class labels from the training database's mathematics. The results are particularly promising in emotion, but there is no guarantee that they will provide real-time data with the same performance as the new data set in the model. New features have also been applied to [15-16] cross-domain datasets that appear in different domains, and previous studies have shown that SML is the result of a decline.

Online reviews and posts on social media are useful because they provide valuable feedback for businesses to shift their hobbies and preferences to their marketing strategies to make more informed decisions. Because analysis is very important to determine the overall view or service on a particular subject. Traditionally, sentiment analysis, such as online reviews and Weibo, has been used from a single data source. However, more accurate and comprehensive results need to be developed to encourage emotional analysis from multiple data sources. It is possible to increase the size of emotionally categorized data blocks by specific interest training to use multiple data sources. Until now, the problem of adequate datasets for training classification has been solved by multiple domain reputation analysis [17].

### Table 2 Comparison of Existing Technique Advantages and Disadvantages

<table>
<thead>
<tr>
<th>Technique</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>eXtreme Gradient Boosting and a genetic algorithm [18]</td>
<td>Review methods showed the online ranking mood when shifting; some customers see, for example, both the information and pictures.</td>
<td>Analysis of these comments is manually taken, more time-consuming, and requirements are less stringent.</td>
</tr>
<tr>
<td>Machine Learning Classifier Algorithm [19]</td>
<td>The system supports a variety of applications and generates useful active mode.</td>
<td>It does not give better performance, and many processed data failed.</td>
</tr>
</tbody>
</table>
With the rapid burst of user-generated data, there are other options to enter personal points of view, thinking and behavior. Random Forest Algorithm analysis of such data always requires different marketing, political and scientific application events, resulting in the realization of any individual or public truth thoughts. In addition to the statistical data collected, viewers' views, interests, and the probability of channel changes per clip [21] are also explored.

Manual classification of application standardization is an important, time-consuming development. Automatic categorization of usage reviews allows developers to fix vulnerabilities, especially time. In this perspective, several methods for the automatic classification of reviews have been proposed. However, they did not use the non-textual information reviewed by the application. In that article, an in-depth classification is based on the benefits of learning methods. It evaluates non-textual information applications and does not use Deep Learning Technique that has proven to be highly accurate in text classification in various categories [22].

<table>
<thead>
<tr>
<th>Technique</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Feedback Mining System [24]</td>
<td>Each word and frequency extracts the topic with the highest frequency count.</td>
<td>Evaluate qualitative and quantitative data and provide low sentiment scores</td>
</tr>
<tr>
<td>K Suggests That Agglomeration Algorithmic Rule [25]</td>
<td>To analyze data from the event exploitation sentiment analysis</td>
<td>Each student must submit their feedback in the form of 1-10 as the score, but the concept is similar to the manual form.</td>
</tr>
<tr>
<td>Intelligent Tutoring System [26]</td>
<td>Mindspark provides detailed feedback received from the students and explains the answers.</td>
<td>Lower complexity level in the system</td>
</tr>
<tr>
<td>Markov Chain Monte Carlo Inference Algorithm [27]</td>
<td>Boolean logic function eliminates the problem of identification mentioned.</td>
<td>Estimated response data from the difficulty of grading students' knowledge and questions only from the teacher’s minimal effort.</td>
</tr>
</tbody>
</table>

To evaluate student performance, students need to improve the most constructive learning and educational strategies proposed in previous studies. For example, teachers use traditional questionnaires to evaluate the performance and tactical execution curriculum students' creative
learning. However, most early research of assistant teachers was not automated to detect students' current learning situations, providing immediate feedback. Based on the above problems' interpretation, the Hybrid Biological Data Analysis Approach is used by the teacher study student in the learning process [28] and wearable device biomonitoring study states.

K-NN (K-Nearest Neighbor) is a challenge to identify high-risk students in educational institutions as soon as possible. The time interval between recognition and real-time is dangerous. Early identification can significantly reduce the risk of failure or disconnection. In a small process, identity is easy, but it's unrealistic. Stored online current learning management systems Online can generate predictive models to identify these students, with large amounts of data and blended learning [29].

With the rapid development of the network, Cyber Physical-Social Systems (CPSSs) provides a wealth of information. When a major challenge arises to retrieve valuable information on the Internet, especially with the personalized site that offers large and complex data, the user intends to capture, to precise searches needs of the Common search engines, face the challenge of solving the challenges posed by this information explosion. Provides feedback technology, design and CPSSs [30] personalized websites related to real-time location, efficient configuration, and intelligent search framework.

As one type of operating system, cooperation Dual Rotary Crane Systems (DRCSs) is the most common. The large load-carrying capacity of the case where the work is used to complete the complex dynamics is very low. However, there is a lot of focus on DRCS control at this time. Compared to a single crane, the DRCS has multiple level variables such as geometric controls and link relationships. Therefore, controlling complex dynamics and kinetic properties [31] makes DRCS design/stability analysis very difficult.

### Table 4 Evaluated Lectures And Opinion Analysis

<table>
<thead>
<tr>
<th>Technique</th>
<th>Advantages</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional Tracking Method [32]</td>
<td>Feedback mechanisms, in practice interactive teaching and learning activities, can greatly improve based on this method.</td>
<td>The dual-channel organization only provides the basic functions of the classroom feedback system</td>
</tr>
<tr>
<td>Machine Learning Technique [33]</td>
<td>The system evaluates the talks and comments from students, and lectures are collected through a questionnaire.</td>
<td>Real-time text analysis problem feedback is difficult.</td>
</tr>
<tr>
<td>student feedback Mining System [35]</td>
<td>Qualitative and quantitative feedback can be made about teaching and learning to help the informed decision.</td>
<td>However, qualitative data is not included in the report</td>
</tr>
</tbody>
</table>
Opinion mining has become widespread, providing information on online reviews and feedback systems. Conventional Opinion Mining Technique allows people to see how they feel about a particular subject as positive or negative emotions in such feedback. In the present course, the purpose of sentiment analysis is to provide detailed sensory information about mines in various fields. Therefore, the proposed system has analyzed the students' perception of the conceptual system [34,35] on a feature scale.

After clarifying the technical level of Python scholars' opinions using achievements in opinion mining and open-source tools, the work is done. See the comments supported by the extraction options such as inspection, guidance, etc.; provides an overall performance comparison. Overall high performance [36] is Compared with Artificial Neural Network (ANN) results to find some metrics related to various design techniques.

In the past decade, emoticons have become a new and popular form of digital communications across different social networks and spoken language. To treat this as a new model, CNN (Conventional Neural Network), ideograms' right are distinct from the two texts' semantic structure. They are typically embedded similar to the image [41].

With the popularity of social platforms, emoticons have become very popular among many users. It expresses more than plain text and enhances the contents. In messages and microblogging posts, use the appropriate emoticons, so lovely and friendly. Recently, emoticons recommendation is becoming a significant task because it is difficult to choose the right one from the thousands of emoticons candidates. Context-Aware Personalized Emoji Recommendation (CAPER) model incorporates context information and personal information [42].

7. Comparative Analysis of Different Control Algorithm

For each randomly selected data set partition. The same training and test sets are used for each category. These three indicators, Precision, Recall, and F-score, are compared for classification using the calculated performance.
Figure 2 Sequence Performance Level Analysis [2]

The above diagram 2 shows the Naïve Bayes classifier precision value is 0.77%, recall value is 0.68%, F-score 0.72%, Rocchio Classifier precision value is 0.75%, and recall value is 0.73%, F-score 0.74%, Perceptron Classifier precision value is 0.65%, and recall value is 0.56%, F-score 0.69%.

Figure 3 Statistical Analysis of positive and negative reviews [20]
As shown in figure 3, these conventional approaches may not have a very high accuracy/precision value compared to the star ratings. Better ratings are based on the recommendation system, ignoring the traditional rating, which inherently functions.

8. Conclusion:

Many emotion analysis methods rely on public opinion dictionaries to assess emotions in texts and paragraphs from education database. Sentimental and emotion dictionaries are dictionaries of words that correspond to their corresponding emotion categories and meaning orientations. Semantic orientation has been a numerical measure used to express a word or phrase's polarity and intensity. The vocabulary can use the opinion word's orientation value in the aggregate that calculates the message's polarity. They have also proved useful for feature extraction in unsupervised classification schemes. The emotional polarity of the sentence is used as a function of the sarcastic detection of product opinions. The data used is a corpus of satirical markers. Although the accuracy has not reached a satisfactory level, results are expected. The F-score of the score obtained has been calculated using the unsupervised machine learning method.

References

2. Troussas, Christos; Virvou, Maria; Espinosa, Kurt Junshean; Llaguno, Kevin; Caro, Jaime (2013). IISA 2013 - Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning. , 1–6. doi:10.1109/IISA.2013.6623713.
5. Sindhu, Irum; Daudpota, Sher Muhammad; Badar, Kamal; Bakhtyar, Maheen; Baber, Junaid; Nurunnabi, Mohammad (2019). Aspect Based Opinion Mining On Student’s Feedback For Faculty Teaching Performance Evaluation. IEEE Access, 1–1. doi:10.1109/ACCESS.2019.2928872.
17. Abdullah, NorAniza; Feizollah, Ali; Sulaiman, Ainin; Anuar, Nor Badrul (2019). Challenges and Recommended Solutions in Multi-Source and Multi-Domain Sentiment Analysis. IEEE Access, 1–1. doi:10.1109/ACCESS.2019.2945340.


40. Sindhu, Irum; Daudpota, Sher Muhammad; Badar, Kamal; Bakhtyar, Maheen; Baber, Junaid; Nurunnabi, Mohammad (2019). Aspect Based Opinion Mining on Student’s
Feedback For Faculty Teaching Performance Evaluation. IEEE Access, 1–1. doi:10.1109/ACCESS.2019.2928872.
