

THE INTERDISCIPLINARITY OF SENTIMENT ANALYSIS AND ITS USE IN TEACHING

<https://doi.org/10.47743/jopafl-2023-30-04>

BRANZILA Carina Ionela

Alexandru Ioan Cuza University of Iași, Faculty of Economics and Business Administration
Iași, Romania
carina.branzila@gmail.com

Abstract: Sentiment analysis is no longer considered an emerging field in the context of today's language analysis tools. Widely used in the business area, it is a very useful tool in analyzing reviews and feedback provided by customers for different types of products and services. Beyond its wide use in marketing and tourism, for instance, sentiment analysis has proven its interdisciplinary character, going far beyond its original purpose. One innovative and creative way is its use in education. There are already quite a few studies in this direction, highlighting the importance and usefulness of sentiment analysis in teaching. This article adds to this research, focusing on its use in an English as a second language class, in the setting of an Economics faculty, with first-year Bachelor students. The general context of this teaching approach, its details, results, and limitations will be described in detail.

Keywords: sentiment analysis, teaching, English as a second language

This Article was presented as a paper at the 15th edition of the Annual International Conference Globalization and Higher Education in Economics and Business Administration (GEBA 2023), which was held at the Alexandru Ioan Cuza University, Faculty of Economics and Business Administration in Iasi, Romania from the 19-21 October 2023.

Introduction

‘Sentiment analysis, also called opinion mining, is the field of study that analyzes people’s opinions, sentiments, appraisals, attitudes, and emotions toward entities and their attributes expressed in written text. The entities can be products, services, organizations, individuals, events, issues, or topics’ (Liu, 2015). The term ‘opinion mining’ appeared first when researchers attempted to classify a larger corpus of reviews retrieved from important sites available online like Amazon and C/net; the classifier used was gradually refined by mining at sentence level and large number of sentences were ‘parsed and thresholded by the mining tool employed’ (Dave, Lawrence and Pennock, 2003). The results of these authors, published more than 20 years ago were yet unclear and uncertain at the time and more work in the field was obviously recommended. However, the two tests engaged in this particular ground-breaking research offered fairly good results for the classification of online

reviews, as long as the ‘appropriate features and metrics’ were used (Dave, Lawrence and Pennock, 2003). On the other hand, the sentiment analysis as a term was coined by Nasukawa and Yi, in the same year of 2003. Their research highlighted the extraction of opinions using reviews and texts from various websites, as well. However, it did not focus on the sentiment from a large text such as a full document, but it was narrowed to references to the particular object reviewed. The use of natural language processing (NLP) as a tool proved quite precise, as it offered consistent results of over 86% accuracy in all tests performed.

The main use of sentiment analysis is on analyzing information using a special algorithm in order to discern if that information expresses a positive, neutral or negative sentiment – hence the name of the tool itself. With the widespread of internet shopping, a larger and larger volume of reviews and opinions on various products, online shops, services etc. are available today. According to Bing, the creation and development of sentiment analysis concur with the one of the social media: ‘sentiment analysis is now right at the center of the social media research’ (Bing, 2015). It is therefore widely used in marketing, tourism and sales, but the research on sentiment analysis has impacted the society in general, too, because people’s opinions influence everything. Furthermore, the overwhelmingly rich resource of opinions, reviews and feedback available online, companies and researchers can gather insights into the perspective of all categories ‘of people or communities towards any service or any product’. This wealth of feedback and reviews are precious to various companies worldwide, as this ‘active feedback actually may help the development of services and products by incorporating the opinions expressed by users and clients. (Mehto & Indras, 2016)

There are various types of sentiment analysis and different software tools using it. This paper focuses mainly on the model called Lexicon based approach. A ‘lexicon sentiment model’ approaches sentiment analysis using a dictionary (lexicon) of words and their associated sentiment scores. More precisely, this refers to a database of words and their positive, neutral or negative connotation. In order to analyze a text from the sentiment point of view, this approach looks into each word of the text and its lexicon score, aggregating them to determine the general sentiment of the text.

According to a pre-set threshold, the overall sentiment score of the text may prove to be positive, negative, or neutral, depending on the number of words with that score found in the text. This lexicon-based sentiment model is easy to use, but they are limited in analyzing irony or other types of nuances. Human intervention is therefore needed or at least the use of more accurate machine learning techniques, in order to deal with more complex texts. To offer a very basic and simple example of how this approach work, we may have a lexicon/dictionary with the following words and scores: Happy: +1; Sad: -1; Good: +1; Bad: -1; Neutral: 0.

If we were to analyze the sentiment of a sentence like: ‘This particular product is good, but its second feature is bad’, we would break the sentence into separate words and search their scores in the predefined lexicon: "This product": Neutral (not in the lexicon); "is": Neutral "good": +1; "but": Neutral; "its second": Neutral; "feature": Neutral; "is": Neutral; "bad": -1

Summing up the sentiment scores of this simple sentence we end up with:

$$1(\text{Good})+(-1) (\text{Bad})=01(\text{Good})+(-1)(\text{Sad})=0$$

The overall sentiment score in this case is therefore 0; in other words, the score is neutral and the sentiment analysis of the sentence classifies the sentence as neutral.

This is a first, basic step in lexicon based analyses. In complex lexicon sentiment analysis, which uses machine learning techniques, sentiment scores are weighted, the topic and order of words are also looked into and the result is more nuanced and accurate. For example, a word like 'OK' could be assigned in such a lexicon a score of 1, 'good' a score of 2, and 'awesome' a score of 3; on the negative side, 'terrible' could be -2 and 'bad' -1. The actual score values would always depend on the lexicon and the assigned values. The sentiment score of a sentence is calculated as a sum of individual word scores, just like in the simple example above.

Some popular and widely used lexicons for sentiment analysis are the AFINN Lexicon, with almost 4000 listed words and their scores, SentiWordNet - an opinion lexicon, its origin database being the WordNet database, or VADER (Valence Aware Dictionary and Sentiment Reasoner) which is mainly designed for analyzing social media texts.

Literature review

An interesting approach to sentiment analysis uses fuzzy logic, as explained by authors like Subasic and Huettner. They developed a lexicon associating words with affect categories, as described above, and specified the intensity of the given word and its relation to the category it belonged to, also called centrality (Subasic and Huettner, 2001). A possible example would be the word 'dislike' from the category of 'sentiments', displaying a certain level of intensity and centrality. Words are then grouped in sets and these sets help analyze the sentiment of a document (Subasic and Huettner, 2001). A separate method would be to make use of a manually-constructed lexicon to analyze the text in order to discern if its general sentiment is positive, neutral or negative.

In their research on fuzzy logic, Subasic and Huettner propose an innovative approach that combines natural language processing and fuzzy logic techniques for analyzing affect content in a text. Their objectives include rapid analysis and visualization of affect content to aid decision-making. The key resource is the already mentioned fuzzy-affect lexicon, from which the fuzzy thesaurus and affect category groups are derived. The text is tagged with affect categories, and centralities/intensities are combined using fuzzy logic to generate affect sets, representing the document's general sentiment. Analysis on social media or news content or movie reviews show a correspondence between affect sets and human judgments. This approach allows for a certain ambiguity, leveraging fuzzy logic's ability to handle word vagueness in natural language. The authors highlight the importance of affect-related information in various electronic documents, emphasizing its role in decision-making and human communication. They introduce a fuzzy semantic typing approach that employs an affect lexicon, addresses ambiguity, and integrates with general-purpose text management systems. Fuzzy logic is chosen for its qualitative emphasis and effectiveness in managing ambiguity and imprecision. The system generates a fuzzy thesaurus and affect category groups for enhanced functionality, as well (Subasic and Huettner, 2001).

A different type of sentiment analysis model uses the so-called Aspect Catalogue, which is able to analyze a text considering its 'intended meaning' and classify it as positive, negative or neutral. The authors analyzed a word, for instance, on a two dimensional grid which can discern this intended meaning and, according to the context, can 'detect

underlying sentiments such as sarcasm satire praise or slander' (Mehto and Indras, 2016). This is an important factor in sentiment analysis because the software often fails in this respect, this lack of refinement being one of its major shortcomings, according to researchers.

Generally speaking, one of the most common areas of application of sentiment analysis is analyzing customers' reviews of products and services from the feedback section of websites or from social media. The reviews and comments from customers help future clients decide what to purchase, if the case, and at the same time foster customer engagement. They also enhance brand visibility and very importantly, they help the company improve their products and services. In fact, both positive and negative reviews contribute to the overall reputation and performance of a business, as there is no bad publicity, it is generally considered. A specific application of such a type of analysis is the so-called 'flame detection, where highly heated or antagonist language can be identified' (Mehto and Indras, 2016). Flame detection in sentiment analysis contributes to creating a more positive and respectful online environment by identifying and managing content that could incite negativity. It requires a nuanced understanding of language and context, often involving a combination of linguistic analysis, machine learning, and user moderation mechanisms. In this way, offensive comments can be filtered and moderated, as needed. Furthermore, customer opinions play a crucial role in guiding business decisions. Monitoring the public's perspectives is obviously a valuable tool for discovering potential customers and understanding their preferences. This insight aids in fostering positive public relations and can also be instrumental in predicting future business trends. Also, companies can adjust their services and products to correspond to the demands of the market. Perhaps this is one of the most useful uses of sentiment analysis in the case of companies.

All methods described have their merits and their drawbacks. Since this paper focuses mainly on the lexicon based method, it is worth mentioning that its advantages lie in its generality, as the words hold consistent meaning regardless of the topic. As previously mentioned, the lexicon-based approach involves using dictionaries with words annotated by their semantic orientation, polarity and centrality. However, a drawback is that the polarity of a word may change with the context. For instance, consider the sentences: 'This university offers a high standard of education' and 'The criminality is high in that area'. In both sentences, the word "high" retains its meaning, but in the first case, "high level" conveys a positive sentiment, whereas in the second case, "high level of criminality" conveys a negative sentiment. Obviously, this shows once more that the complexity of natural language cannot be analyzed accurately by machines, at least not yet. Human intervention is needed in this case, in order to observe the context and decide if the sense is indeed positive or perhaps neutral or negative. For instance, domains like tourism or websites selling electronic equipment can be quite straightforward, but this is not always the case with blogs, vlogs or social media in general, where feedback and opinions can be nuanced and subtler.

Lexicons can be created manually or automatically. In the following teaching approach, the tool called SentiWordNet, available on the internet as a free resource for opinion mining, was used in class. As most such tools, it can analyze text and categorize it as positive, negative or neutral.

Methodology

In the following we will detail a teaching approach, in other words, a practical lesson plan that attempted to cover the key components of a lexicon-based sentiment analysis model. This lesson plan was designed for an intermediate level of students at the Faculty of Economics, studying English as part of their curricula. The lesson was intended as an exercise in sentiment analysis, and its main purpose was not to teach opinion mining, in any way, but rather to explore a new area using English as a second language. The class itself was a class of English for special purposes, and in one of the seminars, the theme of sentiment analysis was suggested to students, and accepted by them, as well, as a hands-on experiment. The main idea of the class was to only speak English, to work in pairs and groups and to use language in a practical way, by employing a free tool of natural language processing and observe the outcome. No technical background was required, just the possibility of using this type of analysis on an electronic device, be it a phone or a tablet. The entire activity focused on understanding lexicon-based sentiment analysis without any need of programming skills. In fact, the whole point of the lesson was to explain and enforce the concept of sentiment analysis, more precisely that of lexicon-based sentiment analysis in a unpretentious, non-technical manner. The length of the lesson being of maximum two hours, time was also an issue, so things were kept simple and straightforward. The lesson outline proceeded as follows:

- Introduction (15 minutes): Explain briefly some basics of sentiment analysis as a tool and of lexicon-based sentiment analysis using simple language and a short YouTube presentation film.
- Text Highlighting (15 minutes): Provide printed copies of short texts or product reviews (reviewing clearly aspects like performance, design, usability). Ask students to read through the text and use colored pens to highlight words or phrases that convey positive, negative, or neutral sentiments. Encourage them to refer to a simple sentiment lexicon provided by the instructor.
- Identify aspects (10 minutes): Discuss the aspects of the text (e.g., product features like performance, design, usability). Have students underline or circle words related to specific aspects using a different color.
- Group Discussion (15 minutes): Divide the class into pairs or small groups and ask them to share their highlighted texts and identified aspects. Facilitate a discussion on how different aspects contribute to the overall sentiment of the text.
- Break (10 minutes): A break was offered at this time, as the class itself was meant to last for 100 minutes, with a 10 minutes break in the middle: 50 minutes of lesson, 10 minutes break and another 50 minutes teaching time. After the break, students continued working on the topic of the lesson, following with the next stage, the fifth one.
- Create Aspect-Specific Sentiment Summary (10 minutes): On the whiteboard, flip chart, laptop, create a table with aspects and sentiment scores (positive, negative, neutral). In open class, discuss and decide on the sentiment scores for each aspect based on the highlighted words.
- Presentation (10 minutes): Ask each group to present their aspect-specific sentiment summary to the class. Encourage students to explain how they identified sentiments and aspects in their respective texts and if they consider these relevant.

- Reflection (10 minutes): Conclude the activity with a brief reflection. Ask students to share what they learned about sentiments and aspects in text. Discuss, if applicable real-world applications of understanding sentiments in different aspects.
- Assessment (5 minutes): Assess student participation in text highlighting, sentiment identification, and group discussions through verbal feedback and appraisal. Ponder on student understanding through class participation, discussion contributions, and completion of the hands-on activity
- Evaluate the accuracy of sentiment summaries presented by each group.
- Discussion and Reflection (12 minutes): Facilitate a class discussion on the challenges and potential applications of lexicon-based sentiment analysis with aspect-based analysis.
- Encourage students to reflect on how this approach might be useful in real-world scenarios.
- Optional final assignment (3 minutes):
- Assign, according to needs, an optional task that requires students to research and write a short report on a specific application of lexicon-based sentiment analysis in a chosen industry or domain. This can be used in future lessons with the same topic.
- As a final note, this hands-on activity focused on a more visual and collaborative approach to understanding sentiments and its use. It tried to remove the technical barrier and strongly encouraged active participation through group discussions and presentations.

Results

The proposed lesson plan was designed to achieve several learning results related to lexicon-based sentiment analysis, some of them being perhaps ambitious, such as understanding some fundamentals about Sentiment Analysis and recognizing its purpose and significance in understanding text data. Furthermore, an important objective of the class was to help students understand the role, purpose and even use of sentiment analysis, recognizing how words are associated with sentiment scores. Students were supposed to understand some key concepts of sentiment analysis, understanding that sentiments can vary for different aspects within the same piece of text. The real-world case study was meant to highlight the students' ability to understand and practically apply concepts in class, by analyzing sentiments within a pre-selected text, employing critical thinking skills to assign sentiment scores. Most importantly, given the fact that the main purpose of this lesson was to practice English as a second language, the focus was not on sentiment analysis itself as a goal, but rather a means of achieving the goal of conversation in the target language, pair work, group work, short presentation and class collaboration in general. This is in fact part of a larger teaching approach, known as CLIL, which stands for Content and Language Integrated Learning. Through CLIL, students learn about topics new to them, such as sentiment analysis and how to use it, and also a second language which is basically the means of instruction during the CLIL lessons. This approach has been extensively promoted by the European Union as a means to achieve multilingualism and many European countries have integrated it into their curricula at all levels of education. This particular exercise of a lesson was devised as an experiment in this regard. Regarding the outcomes of the class, it is worth noting that students participated actively in group discussions, sharing insights and collaboratively deciding on sentiment scores for

the text. They also presented their findings to the class, which can help a lot enhancing their communication and presentation skills. Reflecting on the real-world applications of lexicon-based sentiment analysis and connecting theoretical concepts to a practical scenario has proven helpful, according to the feedback received from students afterwards. These learning outcomes covered a range of cognitive, practical, collaborative, reflective, and application-oriented skills, promoting at the same time the understanding of the concepts presented in class. The proposed lesson plan incorporated interactive and collaborative elements, which positively influenced student engagement, from class observation. From the ulterior discussions with students, they found the subject of sentiment analysis intriguing, especially if they can relate it to real-world scenarios like product reviews or social media comments, which was the case here. They appreciated the application-oriented approach, particularly the case study that connects theoretical concepts to practical scenarios. The hands-on activities, such as text highlighting, aspect identification, and group discussions, promoted engagement. Students appreciated the opportunity to actively participate in a practical activity and actually see the results of their work in class. Group discussions and presentations encourage collaboration. Students mentioned the fact that they enjoyed working together, sharing ideas, and collectively determining sentiment scores for different aspects. This enhanced the understanding of the relevance of sentiment analysis. Encouraging reflection on real-world applications and prompting critical thinking during group discussions can normally foster a deeper understanding of the subject matter.

Limitations

The activity was adjusted for lower-level students, as well, so as to make the concepts more accessible, in the hope that students who might have been intimidated by technical aspects could feel more comfortable and engaged during the lesson. And last but not least, the combination of interactive activities, group discussions, and a reflective component can contribute to creating a positive and inclusive learning environment. Important authors in education mention that ‘the lens of reflection’ in teaching, as outlined by John Dewey as early as 1933, is getting more and more prominence, up to the point when American universities ‘teach reflection as a cornerstone’ in their educational programs (Stanley, 1998). It is therefore useful to include reflection in one’s teaching and classes as often as possible, helping students to objectify themselves. It is worth mentioning that individual student reactions may vary in such a class, and it may be beneficial to ask for feedback even during the class or at least immediately after the lesson, in order to understand how well the objectives were met and to make adjustments for possible future sessions. Overall, a well-structured and interactive lesson plan can contribute to positive student engagement and learning experience and it was hoped this class played this exact role. Among the challenges the students face, apart from the ones mentioned above, students may also have encountered other difficulties, as well. The main issue in a foreign language class, is definitely that students with limited technical proficiency might feel overwhelmed, especially during hands-on activities involving CLIL activities. This may lead to class avoidance, which is problematic, because students simply tune-out of the lesson altogether. In order to manage this problem, the teacher should provide an alternative non-technical hands-on activity that focuses on visual and collaborative aspects, ensuring everyone can actively participate. Furthermore, students may struggle if they have little or no prior

knowledge of sentiment analysis or lexicons and this should always be taken into consideration, perhaps by preparing a brief explanation in the beginning of the class, thus ensuring that foundational concepts are well-explained. Also, the students should be encouraged to ask questions and ask for clarifications, if the case. Students may also encounter difficulties when working in pairs and in groups, as this may not be their usual way of learning, and this may be overcome by showing them the benefits of collaboration in class. And last but not least, students may struggle with their group presentations, as they are shy or feel in the spotlight when discussing in front of the whole class. However, when they have worked in groups, it may become easier because this task can be assigned voluntarily and more assertive or communicative students can assume this activity with success. By anticipating these potential difficulties and incorporating adjustments into the lesson plan, educators can create a more supportive learning environment, allowing students to overcome challenges and engage effectively with the material and the lesson itself.

Conclusions

In today's highly competitive business world, customer experience has become a critical factor in the success of any organization. Understanding how customers feel about your products, services and brand is invaluable in making informed decisions that can lead to improvements. This is where sentiment analysis or opinion mining comes into play. Sentiment analysis is not just a passive tool for understanding customer sentiment. It can provide valuable insights that enable companies to improve customer service and enhance the overall customer experience. By analyzing the sentiment behind customer reviews and feedback, businesses can better understand the clients' preferences. Sentiment analysis can help recommend products or services that match individual customer preferences, increasing the likelihood of people purchasing more of the services and products available. Analyzing these aspects is important in education, and including them in a class of any type, adjusting the activities and tasks according to the level and interest of students may prove insightful and aligned with the tendencies of today. Using sentiment analysis in class, not as a programming endeavor but as a means to, for instance, to engage in a foreign language, can be of interest to students at Bachelor's level in a faculty of Economics. Considering this, the article has described a practical exercise, meant as an experiment, of employing a natural language processing tool like sentiment analysis in a lesson of English as a second language. The stages of the lesson, its results, and feedback were described in detail, along with the limitations and possible problematic aspects of such a class. Some possible solutions to these problems were also provided. However, the lesson overall proved to be quite dynamic and interesting for students, and their feedback was positive on the whole. The main thing to be mentioned is that such a lesson should be repeated, using sentiment analysis in other creative way, or even included in a larger teaching project of using Content-based lessons in teaching a foreign language.

References

1. Bird, S., (2006), NLTK: the natural language toolkit. Proceedings of the COLING/ACL on Interactive presentation sessions 69–72, Association for Computational Linguistics

2. Dave, K., Lawrence, S., Pennock, D. (2003). Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. *Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews*. <https://doi.org/775152.10.1145/775152.775226>
3. Liu, B. (2015) *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press, Cambridge, pp. 4-11, <https://doi.org/10.1017/CBO9781139084789>
4. Mehto, A., Indras, K. (2016). Data mining through sentiment analysis: Lexicon based sentiment analysis model using aspect catalogue, p. 1-7, <https://doi.org/10.1109/CDAN.2016.7570901>
5. Palanisamy, P., Yadav, V., & Elchuri, H. (2013). Serendio: Simple and Practical lexicon based approach to Sentiment Analysis. *International Workshop on Semantic Evaluation*
6. Pang P., Lee L., Vaithyanathans., (2002) Thumbs up? Sentiment Classification using Machine Learning Techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing*, pp. 79–86
7. Stanley, C. (1998). A Framework for Teacher Reflectivity. *TESOL Quarterly*, 32(3), 584–591. <https://doi.org/10.2307/3588129>
8. Subasic, P. & Huettner, A. (2001). Affect analysis of text using fuzzy semantic typing. *Fuzzy Systems, IEEE Transactions on*, 2, pp. 483-496, <https://doi.org/10.1109/91.940962>



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution - Non Commercial - No Derivatives 4.0 International License.