Article

Comparison of Sentiment Analysis Methods Used to Investigate the Quality of Teaching Aids Based on Virtual Simulators of Embedded Systems

Andrzej Radecki [ID] and Tomasz Rybicki *[ID]

Institute of Automatic Control, Lodz University of Technology, 90-537 Lodz, Poland; andrzej.radecki@p.lodz.pl
* Correspondence: tomasz.rybicki@p.lodz.pl

Abstract: Virtual simulators of embedded systems and analyses of student surveys regarding their use at the early stage of the process of learning embedded systems, are presented in this article. The questionnaires were prepared in the Polish language, and the answers were automatically translated into English using two publicly available translators. The results of users' experiences and feelings related to the use of virtual simulators are shown on the basis of detected sentiment using three chosen analysis methods: the Flair NLP library, the Pattern library, and the BERT NLP model. The results of the selected sentiment detection methods were compared and related to users reference answers, which gives information about the methods quality of the methods and their possible use in the automated review analysis process. This paper comprises detailed sentiment analysis results with a broader statistical approach for each question. Based on the students feedback and sentiment analysis, a new version of the TMSLAB v.2 virtual simulator was created.

Keywords: embedded systems simulators; user feedback; sentiment analysis; natural language processing

1. Introduction

Sentiment analysis of student evaluation is a crucial area of research in understanding the opinions and attitudes of students towards various aspects of education [1–6]. It comprises students' attitudes towards courses, teachers, laboratories, and different educational experiences. There is a growing trend in sentiment analysis in the context of student feedback, utilizing Natural Language Processing (NLP) [7–10], Deep Learning (DL) [11,12], and Machine Learning (ML) [13,14] techniques. The review papers [15–17] provide a systematic study on sentiment analysis of students' feedback and the main research topics, venues, and top papers. The authors of [15] have analyzed 92 relevant papers showing the main trends in NLP, ML, and DL techniques. It was shown that the application of DL is rapidly growing.

The interesting proposal for sentiment analysis provides an article [18] that discusses the real-time identification and analysis of student sentiment in classroom teaching. It has many advantages compared to standard post-processing text analysis, as it can assist teachers in understanding student learning stages in time so that they can take appropriate actions. The literature review on the subject encompasses various methodologies and applications. Another proposal for real-time student feedback analysis can be found in [19]. The highest results were found for four aspects (preprocessing, features, machine learning, and use of the natural class) with Support Vector Machines (SVMs). Another application of SVM with a radial kernel and CNB (Complement Naïve Bayes) to analyze students' feedback for real-time interventions in the classroom can be found in [20]. Some other proposals for real-time feedback can be found in [21,22].

Several authors were using different methods of sentiment analysis connected with remote and distance learning that were especially used during the COVID-19 pandemic [23–26].

The authors of [24] deployed a questionnaire to 660 postgraduate students who had access to lecture recordings. Key findings showed that students who accessed lecture recordings reported an enhanced learning experience.

The paper [27] concentrates on the preprocessing method for closed-ended questionnaires. It uses sentiment analysis through polarity. The main goal was to develop a mechanism for analyzing questions and students' emotions in closed-ended responses. The authors proposed a Quest_SA software tool for this purpose. Another type of sentence whose sentiment can be analyzed are conditional sentences.

In turn, the paper [28] proposes a system for automatically extracting aspects from the available text and their corresponding orientation. The study proposes a supervised aspect-based opinion mining system based on a two-layered LSTM model. Another new method of sentiment analysis of student feedback is proposed in [29]. It proposes an innovative model to find the targets of the given sentence using Bi-Integrated Conditional Random Fields (CRFs).

The natural language processing of student online sentiment was analyzed in [30]. The authors were looking for new and better ways to support and understand the learning experience of their students. The idea was to investigate the attitudes and emotions of students when they were interacting on social media about their course experience. The short, informal texts were also analyzed in [31]. The sentiment features were primarily derived from novel, high-coverage tweet-specific sentiment lexicons.

The fuzzy logic rules are also used to analyze the sentiment and understand students' opinions and satisfaction about the university and its services. Some original proposals can be found in [32,33]. Also, a neural network approach can be used in sentiment and behavior analysis [34].

Additionally, the sentiment analysis of student evaluations can benefit from the application of sentiment analysis in different languages, as demonstrated by [35–37] in the context of Korean, Vietnamese, and Kurdish, respectively. There is also a valuable, from the author's point of view, example of sentiment analysis methods in Polish [38].

In the article, the authors focused on a comparative study of sentiment analysis methods [39]. The students' opinions in Polish regarding embedded systems and their simulators used at the Institute of Automation of the Lodz University of Technology [40,41] were analyzed. These opinions contain various emotions and impressions that accompanied the students' work both during laboratory classes and on their own. Based on the analysis of the described student opinions, it was proposed to implement several improvements to the simulators.

For the comparison of sentiment analysis methods, available tools with pretrained NLP models were picked. There are a large number of these types of tools [42], especially with their availability in the Python language [43]. The following requirements were taken into account when choosing analysis methods:

- Ability to analyze texts composed of many sentences [44];
- At least a five-level sentiment analysis scale returned;
- Availability of ready-to-use, pre-trained NLP models [45].

Based on these criteria, Pattern [46], BERT [36], and Flair [47] were used for sentiment analysis.

Since the survey was conducted in Polish and the NLP models for sentiment analysis are trained mainly based on English [48], automatic translations [49] using two translators were used. The two publicly available translators were compared: Google Translator [50] and MyMemory Translator [51].

As a final result of the statistical processing analysis of detected sentiments in relation to the reference answers provided by the participants of the proposed survey, the best method for determining sentiment in Polish surveys with automated language translation was selected.

2. Virtual Simulators of Didactical Embedded Systems

2.1. Embedded Systems at the Institute of Automatic Control at Lodz University of Technology

The Institute of Automatic Control at the Lodz University of Technology (LUT) [52] has a long tradition of teaching embedded systems [53]. It has an extensive hardware base and covers typical issues of using these systems in real applications. The set of didactic microprocessor modules shown in Figure 1 is a subset of the hardware base and is used in two subjects: Microprocessor Techniques [54] and Microprocessor Systems Software [53].
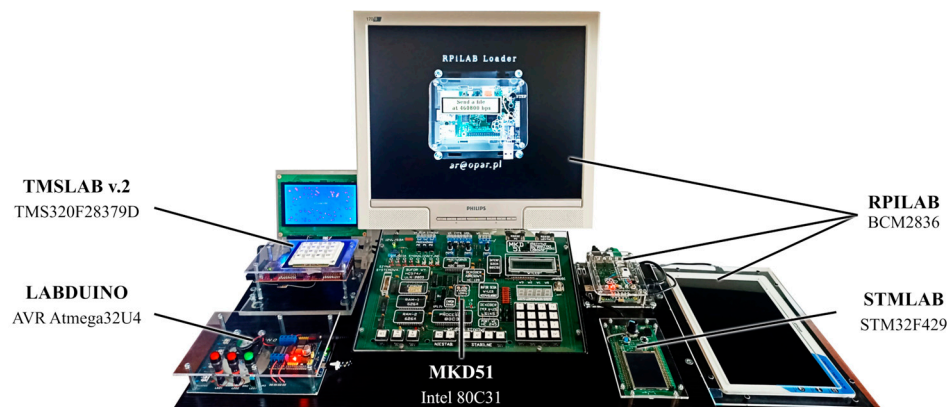

Figure 1. A didactical laboratory stand for microprocessor system teaching.

This didactic set is, in some cases, special because each hardware module has its own virtual simulator [40,41]. These simulators are increasing the flexibility of students' work and thus improving the quality of teaching [55].

2.2. The TMSLAB Didactical Module Simulator

Virtual simulators of embedded systems are very important for students. The authors of the paper draw that conclusion based upon nearly 30 years of teaching experience at the LUT and recent experiences connected with the COVID-19 pandemic. The teaching goals are clear—providing access to high-quality material to study outside the university in order to improve skills and knowledge. Specialized laboratories limit the opportunity for education without access to hardware resources. Embedded systems programming classes provide a unique opportunity to conduct education outside the university, providing students with appropriate tools, in this case simulators of embedded systems used in the classes. However, these systems have a unique architecture, and developing simulators for them is not an easy task. Therefore, before upgrading or modifying a given tool, knowledge about the opinions and feelings of the students on this subject is needed. At the same time, students constantly emphasize how much simulators make it easier for them to work on laboratory tasks, whether it is in terms of preparation for task problems or in connection with the need to complete work started during classes.

One of the microprocessor modules used in teaching classes is the TMSLAB system [53]. Its current, second version, is based on the Texas Instruments [56] TMS320F28379D MCU and is equipped with intermediate circuits enabling the use of a graphic-text LCD, a 4×4 matrix keyboard, and an LED line. The laboratory module evolved from the first version, equipped with the TMS320F2812 MCU system and a similar set of peripherals [41]. With a lot of effort, a system-level simulator [41] was introduced a few years ago for TMSLAB v.1 (Figure 1) and has been used since then in laboratory classes.

The introduction of a new version of the TMSLAB v.2 laboratory hardware module during classes resulted in several differences from the first version. The developed simulator, TMSLAB v.1, was still useful, but students had the impression that it was not suitable for the equipment, which resulted in their different reactions.

Feedback from students was important about the need to further develop the simulator, and therefore it was necessary to conduct an appropriate survey. Finally, at the beginning

of 2024, a decision was made to update the simulator, which resulted in the development of its newer version, shown in Figure 2, in the form of the TMSLAB v.2 simulator.
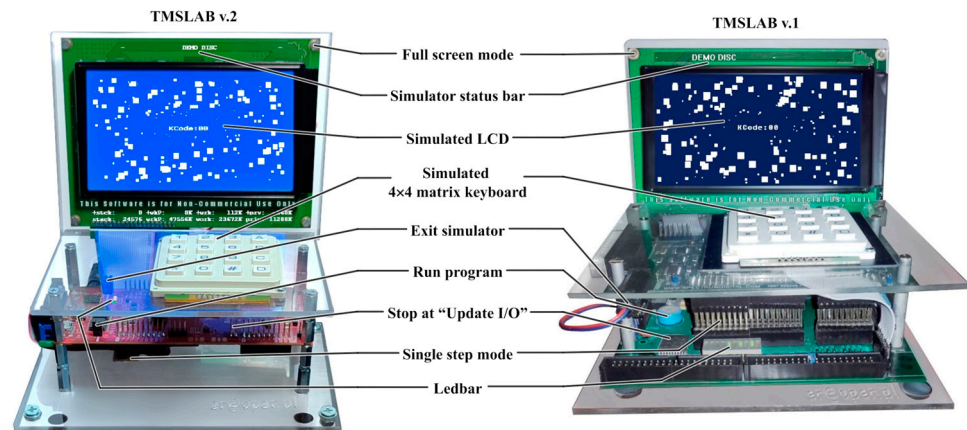

Figure 2. Virtual system-level simulators of TMSLAB didactic modules in two versions.

3. Computing Methods and Tools for Sentiment Analysis

3.1. User Feedback as a Guideline for Product Improvements

In addition to the desired didactic effect related to the use of virtual simulators of embedded systems, the authors also aimed to encourage students to use these tools. This required the implementation of some simulator functionality that is not only about elements that lecturers care about but also taking into account users' opinions. Therefore, simulator developers need to consider components that are needed from a purposeful perspective but also which functionalities are redundant and which are missing.

A good method to obtain information about the last two aspects is through feedback from surveys. The questions they contain may be open or closed; they may concern very narrow and precisely formulated issues; or they may allow greater freedom of expression [27]. As a final result, the goal of the survey is to obtain feedback related to the purpose of conducting it.

3.2. Text Sentiment Analysis

The opinion and sentiment expressed in the respondent statement are very important pieces of feedback. This sentiment guides the interpretation of the user's answers [33] and allows for a better understanding of his needs and emotions [15]. With a large number of surveys, manual sentiment analysis takes a lot of time and requires the operator's work. The results of manual analyses may not be repeatable and may depend on the emotional state of the person interpreting the answers [42]. Automated sentiment detection systems are free of these drawbacks, the results of which can be used to automate processes that meet the needs of respondents [17].

Basically, the result of the sentiment analysis should be the classification of the text into one of the following categories: positive or negative. More precise analysis techniques add a neutral (mixed) category, or additionally, positive mixed and negative mixed. Still, other sentiment analysis methods use indicators that are numbers in the range usually [0,1] or [−1,1], specifying a scale from negative to positive for sentiment description and additionally supplementing the sentiment estimation with a parameter of its detection quality [30].

In accordance with the assumptions presented in the introduction, three tools available for research and comparison were chosen. They were accordingly:

- The Flair NLP library [47];
- The Pattern library [46];
- The BERT NLP model [36].

All three tools were used with default parameters and were in versions: Flair (BLSTM—Bidirectional Long Short-Term Memory) v. 0.13.0, Pattern v.3.6 (DbA—Dictionary-based Approach, SLP classifier based on WSJ trained on the John = "NNP-PERS" dictionary), and Multilingual BERT (mBERT) used in Transformers library v. 4.35.2 (BERT was used with a pretrained model based on English datasets for a multilingual model).

For the purposes of this article, results were unified to give results in a range from one to five, with subsequent sentiment levels assigned to: positive, mixed positive, mixed, mixed negative, and negative sentiment. All the sentiment analysis methods used in the article were scaled to those five levels to give comparable results. Typically, the literature considers a two-point, three-point, and five-point sentiment. The authors decided to use the last variant, which allows for sentiment assessment with greater diversity, providing better comparison possibilities for analysis methods. The libraries used for analysis return not only the probability of sentiment classification but also a five-point sentiment assessment scale. The probability of the determined sentiment result was used as an element of filtering the results, introducing two sub-variants of the comparison of research results.

4. Research Methodology

4.1. The Quality of Teaching Aids Based on Virtual Simulators of the Embedded System Survey

The results presented in this article are based on a survey prepared to verify the suitability of methods and tools used in the subject of Microprocessor Systems Software [53] lectured at the Institute of Automatic Control, Lodz University of Technology [52]. The survey was prepared in such a way as to obtain answers to questions about the quality of tools and materials supporting the teaching process and not discourage students from completing it. This forced us to limit the questions to the minimum necessary set, giving a picture of opinions, thoughts, and emotions related to the use of embedded system simulators. The survey consisted of five open questions and five corresponding closed questions with answers on a point scale. The survey itself was conducted on the university teaching platform WIKAMP [57], which is based on Moodle [58]. Essentially, the students were asked about:

- The need for using software simulators in question 1;
- Thoughts on the TMSLAB module simulator in question 2;
- The significance of creating a simulator of the STMLAB system supporting FreeRTOS in question 3;
- The impressions, feelings, and emotions they experienced when using microprocessor system simulators during laboratory classes in question 4;
- Opinions regarding the plan to introduce an additional microprocessor system with a simulator and an additional programming task within unchanged teaching hours in question 5.

The survey was not obligatory. It was completed by students interested in providing feedback, which increases its reliability. A total of 14 questionnaires were collected from a group of 60 students. This is a satisfactory result, considering the fact that the semester surveys completed by students at the Lodz University of Technology as part of the Microprocessor Systems Software course, regarding the quality of teaching, are usually completed by a much smaller group. Some of the students aversion to completing surveys may be related to the requirement of many forms of surveys at the university level, which results in overloading the students in this area.

4.2. Sentiment Analysis of Open Questions

The open answers to the survey were posted in Polish. The intention at this point was to avoid errors in the quality of the respondents' answers resulting from the possible insufficiently precise expression of emotions in a foreign language. This forced the use of language translators to translate the answers into English. Two publicly available translators were used. Google Translator [50] and MyMemory [51] translators. The results of using both translation tools are included in the article.

Due to the exhaustive answers of some respondents, the article also presents two variants of sentiment analysis. The first one, in which entire answers were analyzed, and the second one, where the answers were divided into sentences that were subject to separate analysis. The second approach posed a serious challenge to the compared sentiment analysis methods because some sentences, taken from a broader opinion, may not correctly represent its sentiment.

4.3. Survey Reference Questions

The sentiment detected in the survey responses was verified based on the answers provided by the respondents themselves. For each open question, there was one closed question with an answer on a five-point scale. This scale corresponded to the five levels of detected sentiment. On this basis, the sentiment detection error was determined, which was used to compare sentiment analysis methods. This error was calculated according to the following formula:

$$\Delta S = \left| S_{ref} - \hat{S} \right| \tag{1}$$

where ΔS—sentiment error, S_{ref}—sentiment reference level, and $\hat{S}$—sentiment detected level.

Only those survey answers that led to the determination of Equation (1) were used for result interpretation. This means that the statistical analysis conducted in the article was performed only in cases where both the open question and the corresponding closed question were answered.

5. Sentiment Analysis Results

5.1. Sentiment of Survey Answers

All fourteen collected surveys were subjected to sentiment analysis, considering two different language translators, two sets of responses (complete answers and their split into single sentences), and two groups of results: all results and results rejecting low reliability ones. The assumption made for reliable sentiment detection results is that the method returning the sentiment estimation quality indicator returns its value above half of its total range. For all possible combinations, the results of sentiment are shown in Figure 3 for using Google translator and in Figure 4 for using MyMemory translator.
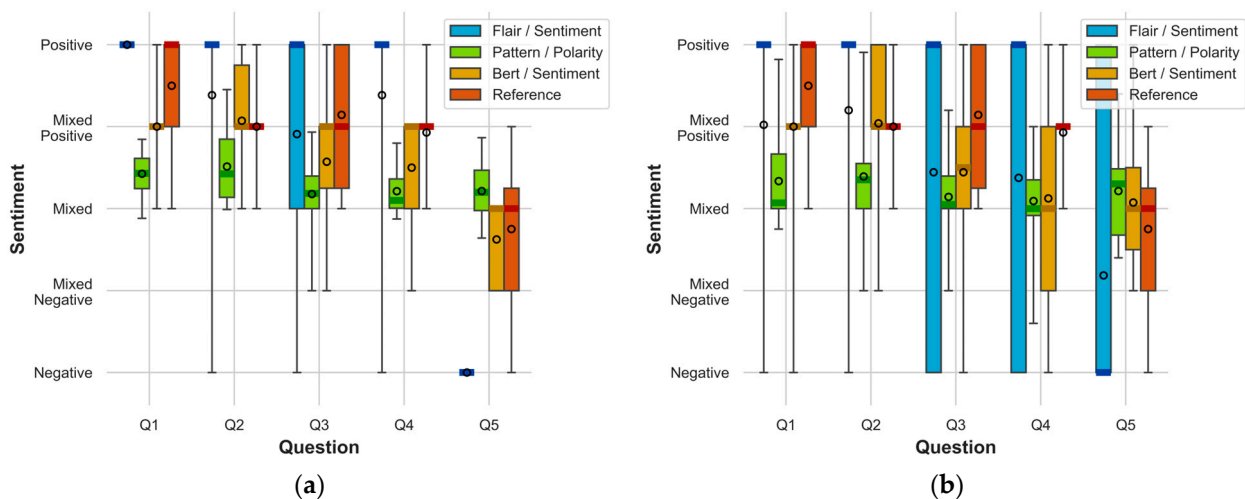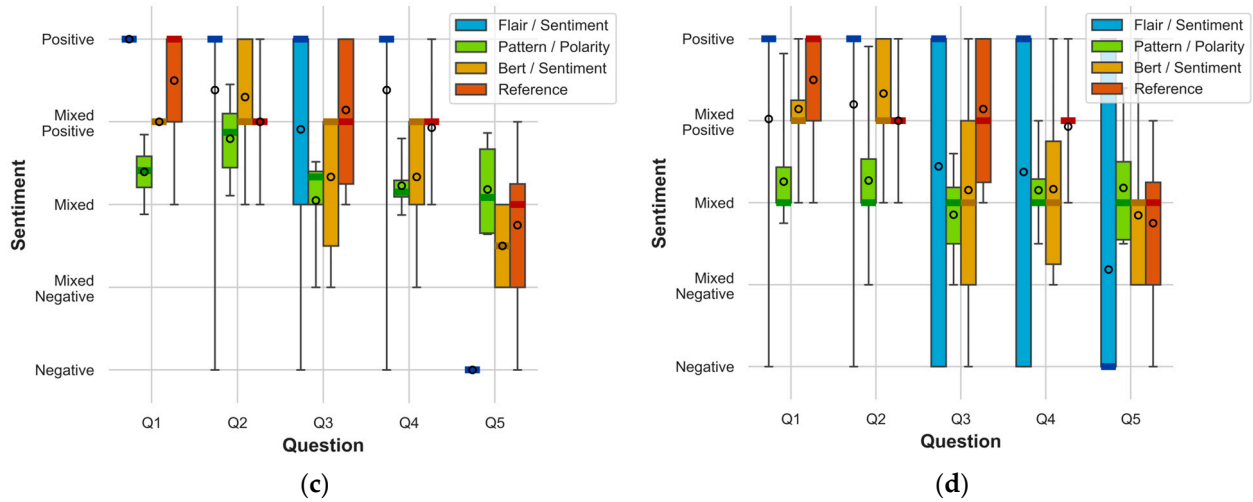

(a)

(b)

Figure 3. *Cont.*

Figure 3. Detected sentiment in survey responses using different methods using Google translator: (**a**) all results obtained for complete answers; (**b**) all results obtained for answers split into sentences; (**c**) results rejecting low reliability ones for complete answers; (**d**) results rejecting low reliability ones for answers split into sentences.
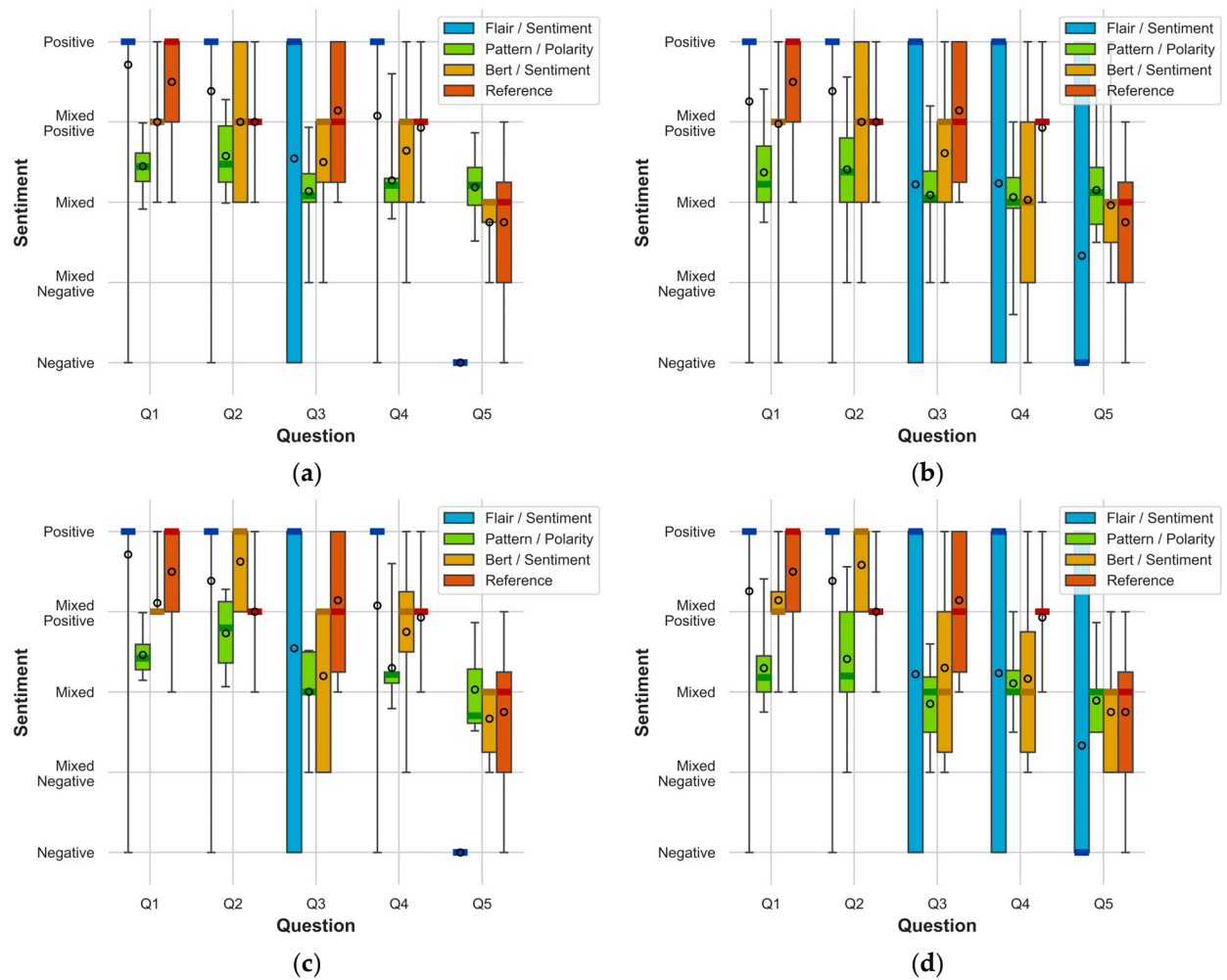

Figure 4. Detected sentiment in survey responses using different methods using MyMemory translator: (**a**) all results obtained for complete answers; (**b**) all results obtained for answers split into sentences; (**c**) results rejecting low reliability ones for complete answers; (**d**) results rejecting low reliability ones for answers split into sentences.

The presented results of the sentiment analysis have been divided according to the question number and presented in statistical terms with the median and average values separately for each of the analysis methods and reference responses that play a verification role. To better illustrate the collected data and their analysis, a broader statistical approach is presented individually for each question in the attachments in Figures A1–A16 for all study cases. There are both probability density distributions obtained for sentiment analysis methods and reference responses, as well as histograms that cover this issue as a supplement to the documentation of statistics regarding the obtained research results. The statistical results included in Appendix A to the article act as complete documentation of the survey results without the need to publish confidential results—another form of sharing survey results. They are also an intermediate element in determining the sentiment estimation error (1), the results of which are presented in Section 5.2.

5.2. Sentiment Analysis Errors

The obtained sentiment analysis results were related to the reference responses in accordance with Equation (1). On this basis, sentiment detection errors were determined and presented in Figures 5 and 6 for all study variants.
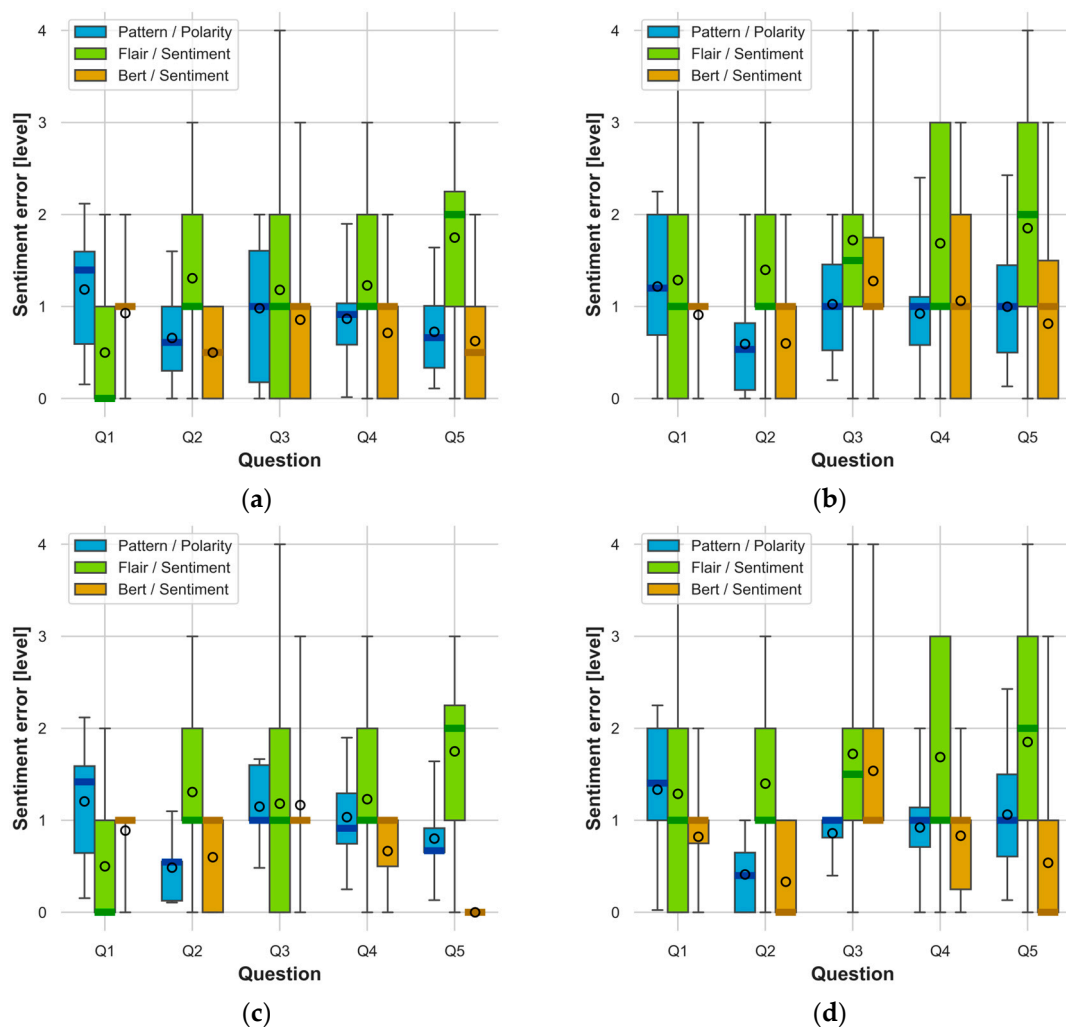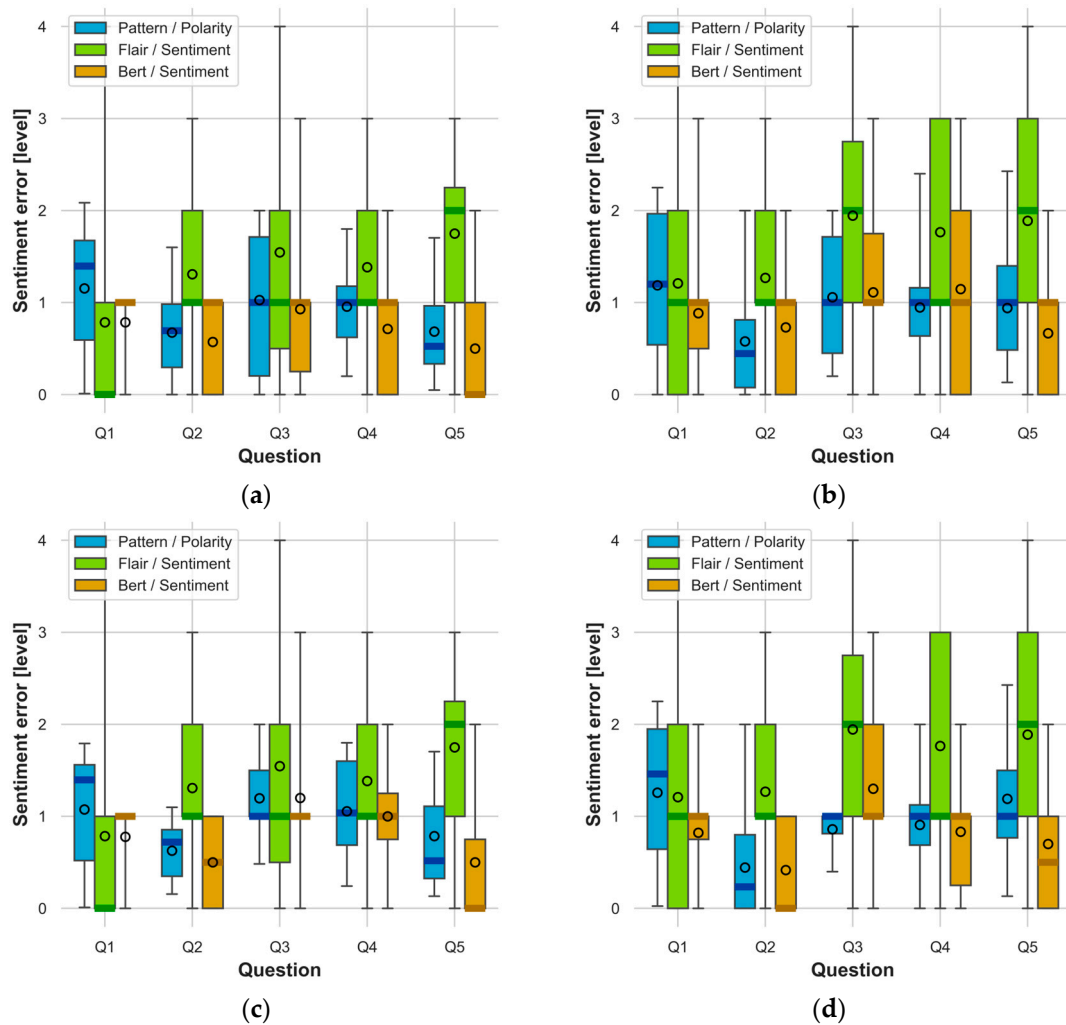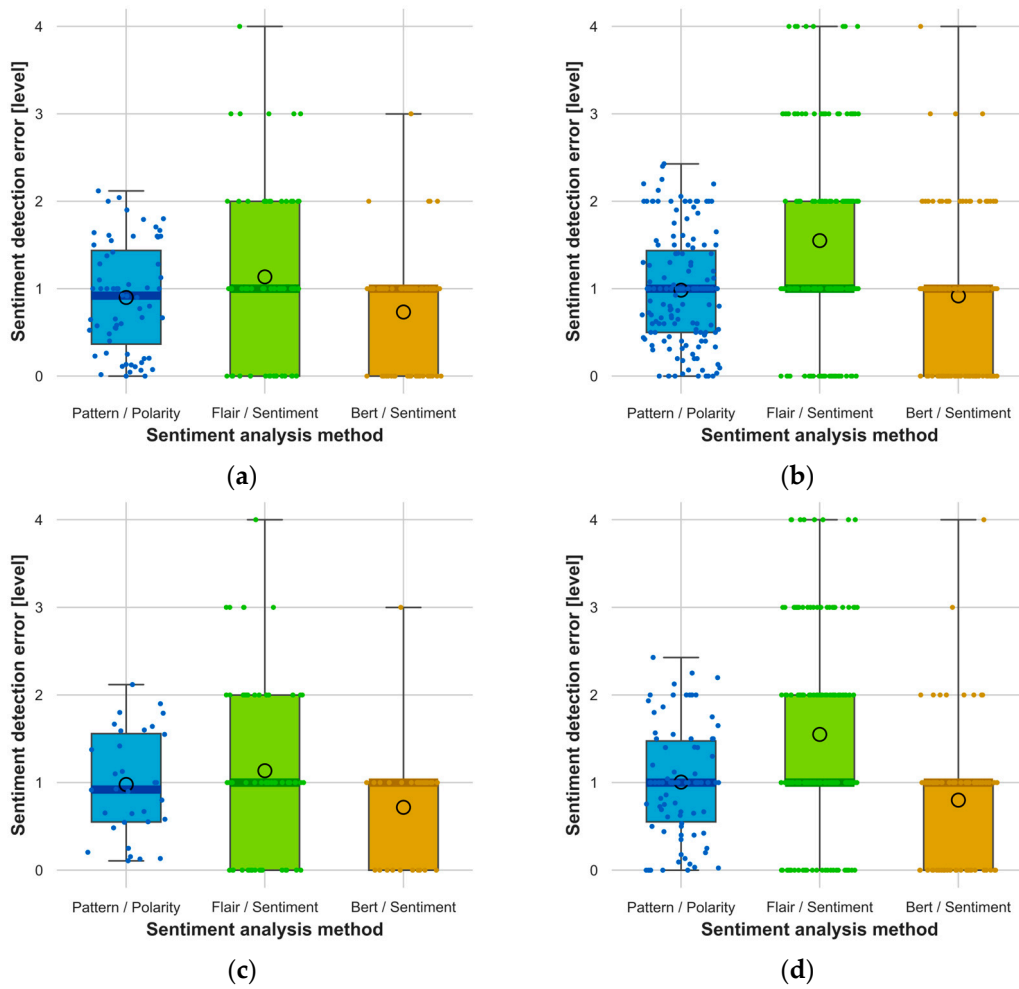

Figure 5. Error statistics of sentiment analysis methods for each detected sentiment level using Google translator: (**a**) all results obtained for complete answers; (**b**) all results obtained for answers split into sentences; (**c**) results rejecting low reliability ones for complete answers; (**d**) results rejecting low reliability ones for answers split into sentences.

Figure 6. Error statistics of sentiment analysis methods for each detected sentiment level using MyMemory translator: (**a**) all results obtained for complete answers; (**b**) all results obtained for answers split into sentences; (**c**) results rejecting low reliability ones for complete answers; (**d**) results rejecting low reliability ones for answers split into sentences.

Similarly to the collected research analysis results, the obtained statistics of sentiment detection errors are presented in more detail in Appendix A in Figures A17–A24. Those figures cover not only errors for each question, like in Figures 5 and 6, but also for each sentiment level, which gives a broader look at the issue.

6. Discussion

The conducted review study showed general consistency between emotions expressed descriptively and the respondents ratings from reference answers. Every method used tended to correctly identify the sentiment, but each of them was characterized by different accuracy and statistical properties. Summary error statistics obtained as a result of the study conducted for each analysis method and all study criteria are shown in Figures 7 and 8.

Figure 7. Total analysis method error of detected sentiment using Google translator: (**a**) all results obtained for complete answers; (**b**) all results obtained for answers split into sentences; (**c**) results rejecting low reliability ones for complete answers; (**d**) results rejecting low reliability ones for answers split into sentences.

Figure 8. *Cont.*

(**c**)   (**d**)
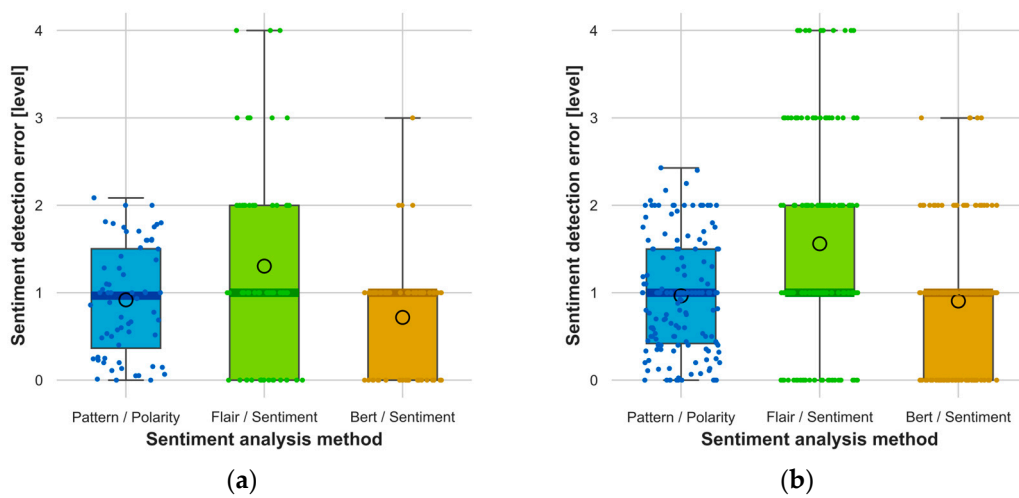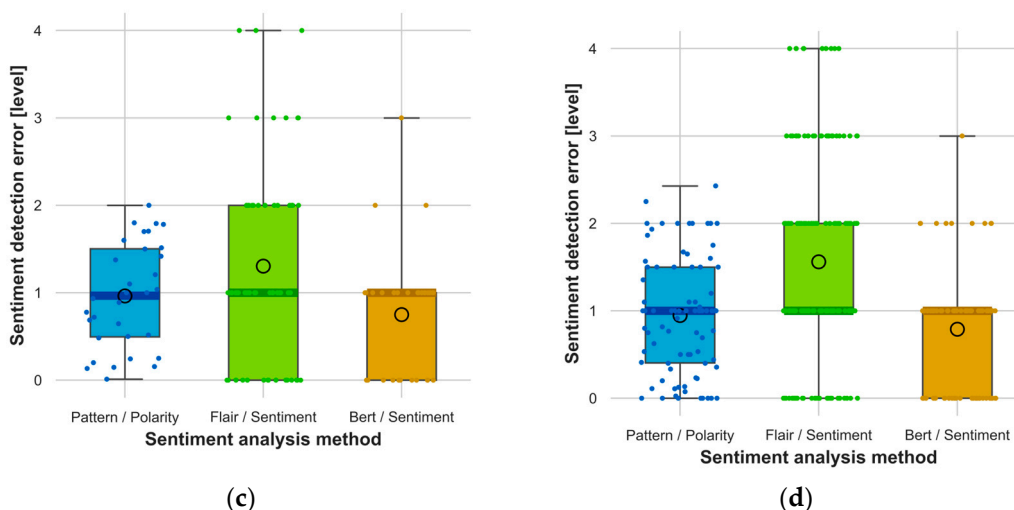
Figure 8. Total analysis method error of detected sentiment using MyMemory translator: (**a**) all results obtained for complete answers; (**b**) all results obtained for answers split into sentences; (**c**) results rejecting low reliability ones for complete answers; (**d**) results rejecting low reliability ones for answers split into sentences.

Based on the review answers detected sentiment levels (Figures 3 and 4), their accuracy expressed as detection error (Figures 5 and 6), and the summarized error for all obtained results per method (Figures 7 and 8), several interesting conclusions can be drawn:

- Each method produced results that make it useful for sentiment analysis. In no case did the average error value of the results exceed two sentiment levels;
- As was to be expected, the BERT model is characterized by the highest precision in sentiment detection. Regardless of the analysis used, the average error made by this method was the smallest;
- The Flair library produced very highly polarized results and occasionally balanced ones. This resulted in a large impact on the error statistics made by this method with every detection mismatch. The method can be useful to distinguish the emotional state of opinions very clearly;
- The Pattern library gave very balanced results in terms of sentiment level. This leads to a reduction in the sentiment detection error but does not allow for a clear distinction between the emotional potential of the review responses;
- In the case of very short statements (splitting respondents' answers into single sentences), all methods returned results consistent with expected ones. The BERT model, which was generally the best, lost the most in terms of analysis quality while at the same time signaling that the obtained results are unreliable;
- The method based on the Pattern library gave surprisingly good results, considering that it is very simple [46] compared to the other two;
- The change of the translator did not affect the nature or general interpretation of the results obtained. This leads to the indirect conclusion that the above methods can be successfully used to analyze the sentiment of texts in Polish using language translators;
- From a statistical point of view, the results obtained for complete answers were similar to the results of analyzing responses split into sentences. This gave a surprisingly high accuracy of analyses of very short texts, which may not necessarily contain emotional potential;
- The results of the error analysis indicate that in both cases of automatic translation, the potential impact of losing the nuances by the translation, discussed in literature [59,60], did not affect the accuracy of sentiment determination.

7. Conclusions

This article compares three sentiment analysis methods for short texts and presents their results in the case of analyzing the quality of teaching aids based on virtual simulators of embedded systems. The comparison took into account various criteria of text taken into account in the analysis, as well as two different language translator tools. It was shown that every sentiment analysis tool gave correct results when analyzing texts in Polish with automated language translations. Even limiting the length of the analyzed text to a single sentence, which may not be emotionally representative in statistical terms, gave satisfactory sentiment detection results.

The issue of potential loss of text nuances in automatic translation was also one of the matters investigated in the article. Two different translators and a given methodology for determining sentiment errors were used for this purpose. The effects of the error analysis show that in all the cases of automatic translation, the potential impact of nuances did not affect the accuracy of sentiment determination. This was observed for both full responses and when questionnaire answers were split into sentences. It reflects the high usability of sentiment analysis through automatic translation and opens the way for automatic sentiment analysis of short texts.

The conclusions of the sentiment analysis method are presented to highlight the strengths and weaknesses of each comparison method. The method based on BERT NLP gave the most precise results, while Pattern was more balanced and Flair strengthened the emotional potential of the statement, often giving mostly negative or positive sentiment.

The article describes what factors may influence the results and what aspects were not taken into account in the study. The results obtained from the comparative analysis are largely influenced by the size of the result set and the expected output sentiment. Due to the limited scope of the survey, the total number of respondents was 14. For this reason, some of the research was carried out by breaking down the survey into individual sentences to enhance their sentimental potential. The main limitation in assessing the obtained results was the level of reference sentiment, which gave less reliable comparative results on mixed sentiment. This was due to the fact that for the mixed reference sentiment, the maximum error that could be made in the analysis was two sentiment levels, while for borderline sentiments it was four levels, i.e., twice as much. For this reason, the questions that were prepared for the survey with the intention of the authors were intended to lead to obtaining borderline reference answers.

The paper also compared not only the currently leading methods but also the simplest ones, based on the DbA (Dictionary-based Approach). This method, considering the mechanisms it uses, may be several times faster than BERT and consume fewer resources, leading to its effective use in small embedded systems, but it requires further research.

Based on the students feedback and sentiment analysis, a new version of the TMSLAB v.2 simulator with more realistic graphics, greater control over the output code and data size in the simulated IDE, and a more interactive user interface was prepared. The authors will also provide our students in the future with a newly developed starter project for the STMLAB simulator (based on the STM32F429 MCU QEMU emulator and LCD system-level simulation) with new features in the form of STM LL (low-layer) drivers and FreeRTOS support in the STMLAB simulator. The ability to run this type of software in a simulated environment is desirable and was assessed mainly positively in the survey. As a result of respondents answers sentiment analysis, the intention to introduce an additional microprocessor module into the teaching program was abandoned, which, according to the survey, would have been perceived relatively negatively.
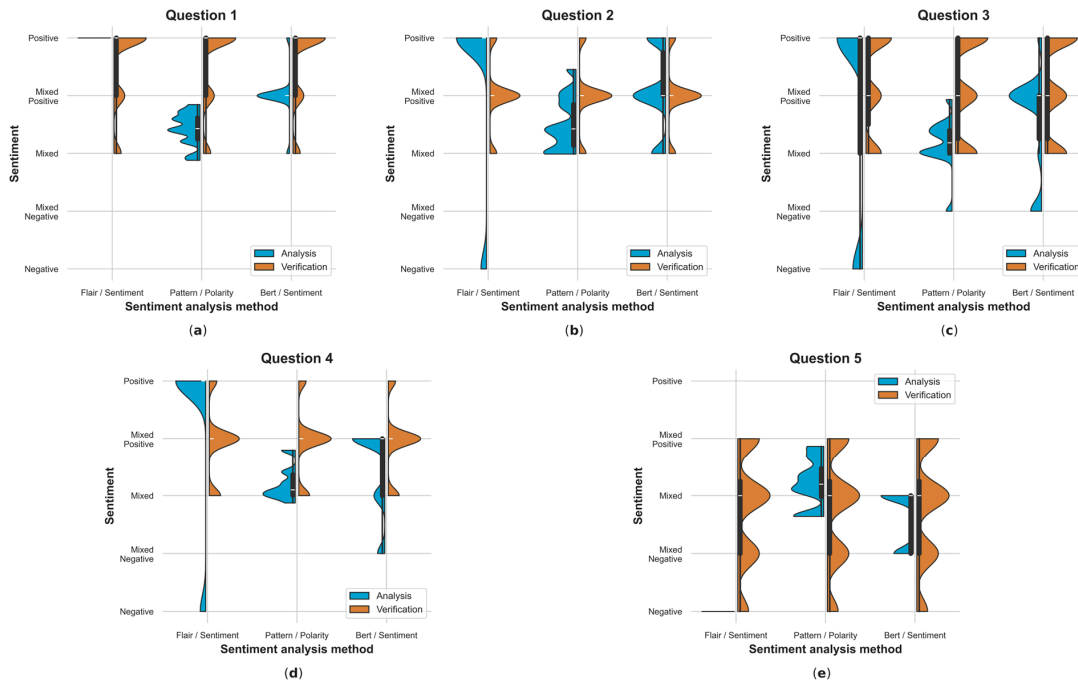
Appendix A

Figure A1. Probability density distributions obtained for sentiment analysis methods and reference responses using Google translator for complete answers and statistics calculation for all results obtained: (**a**–**e**) results obtained for questions 1–5.
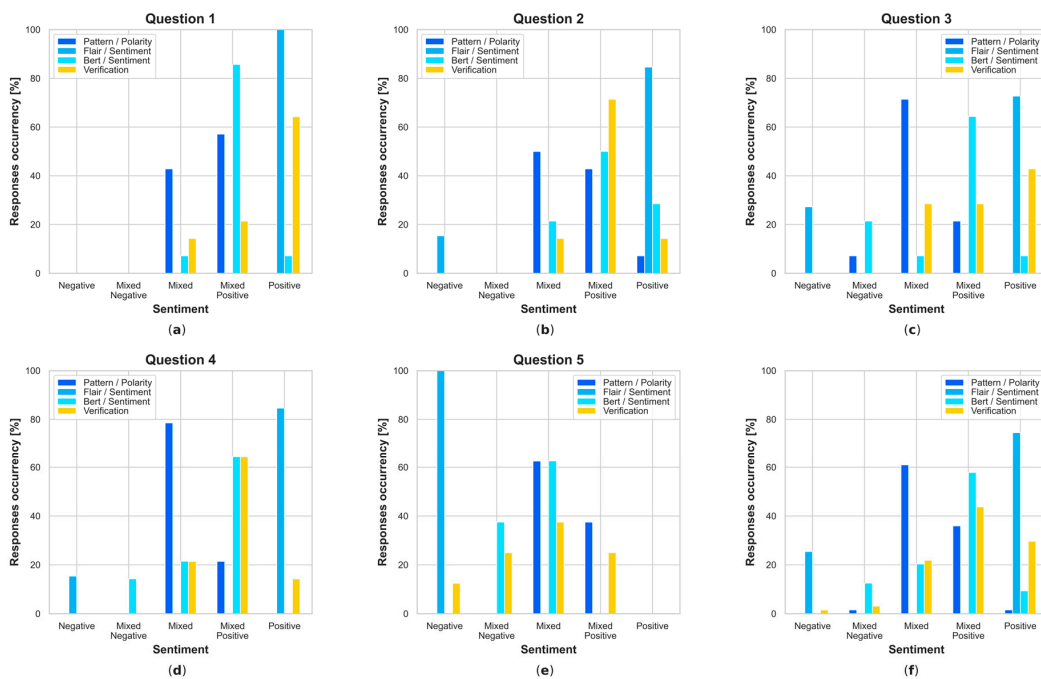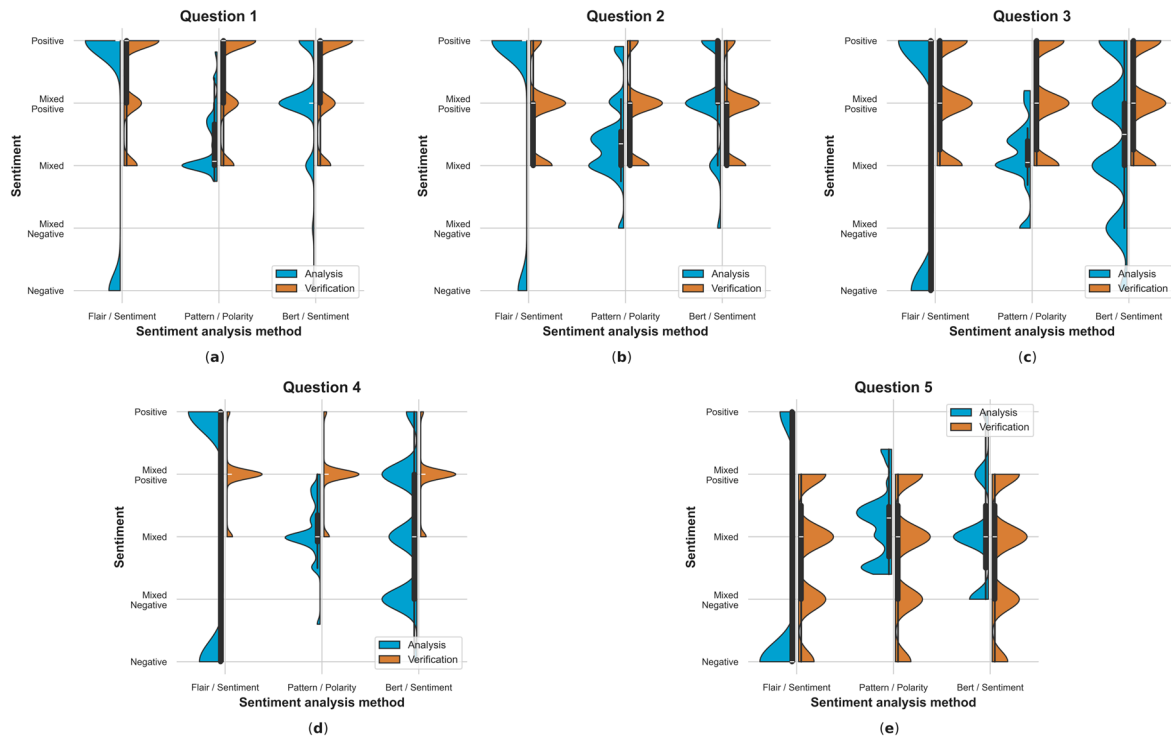

Figure A2. Histograms obtained for sentiment analysis methods and reference responses using Google translator for complete answers and statistics calculation for all results obtained: (**a**–**e**) results obtained for questions 1–5; (**f**) sentiment level occurrence for all answers.

Figure A3. Probability density distributions obtained for sentiment analysis methods and reference responses using Google translator for answers split into sentences and statistics calculation for all results obtained: (**a**–**e**) results obtained for questions 1–5.
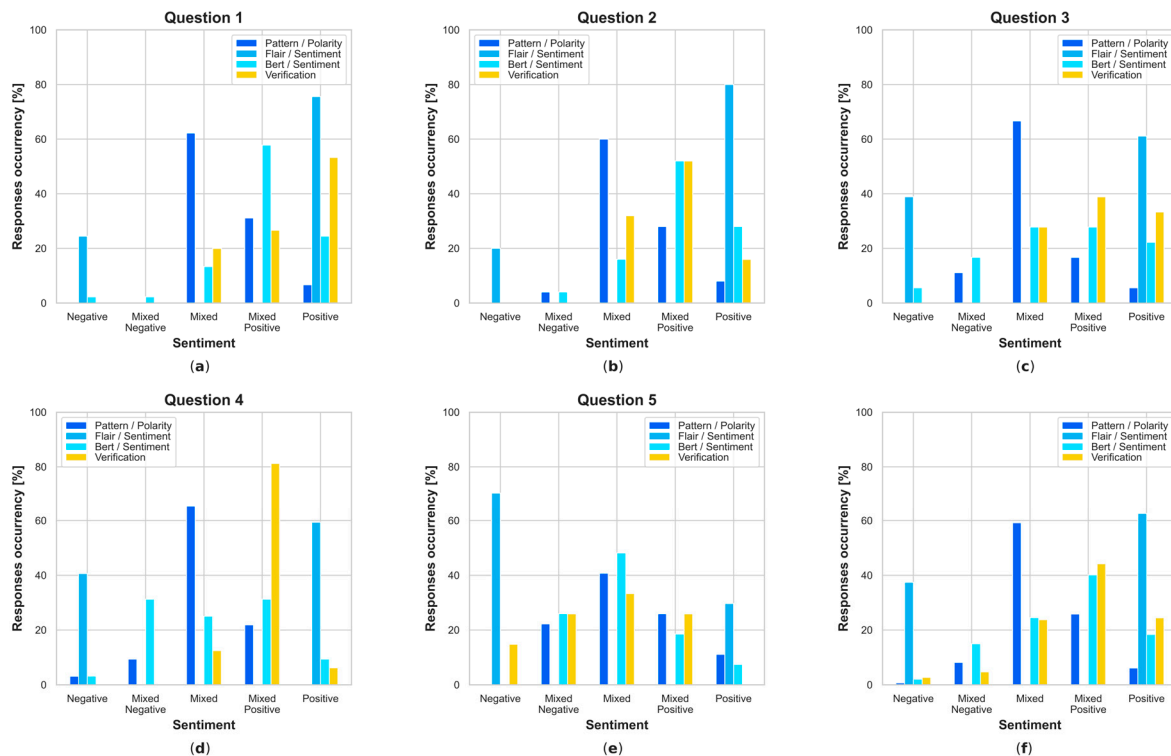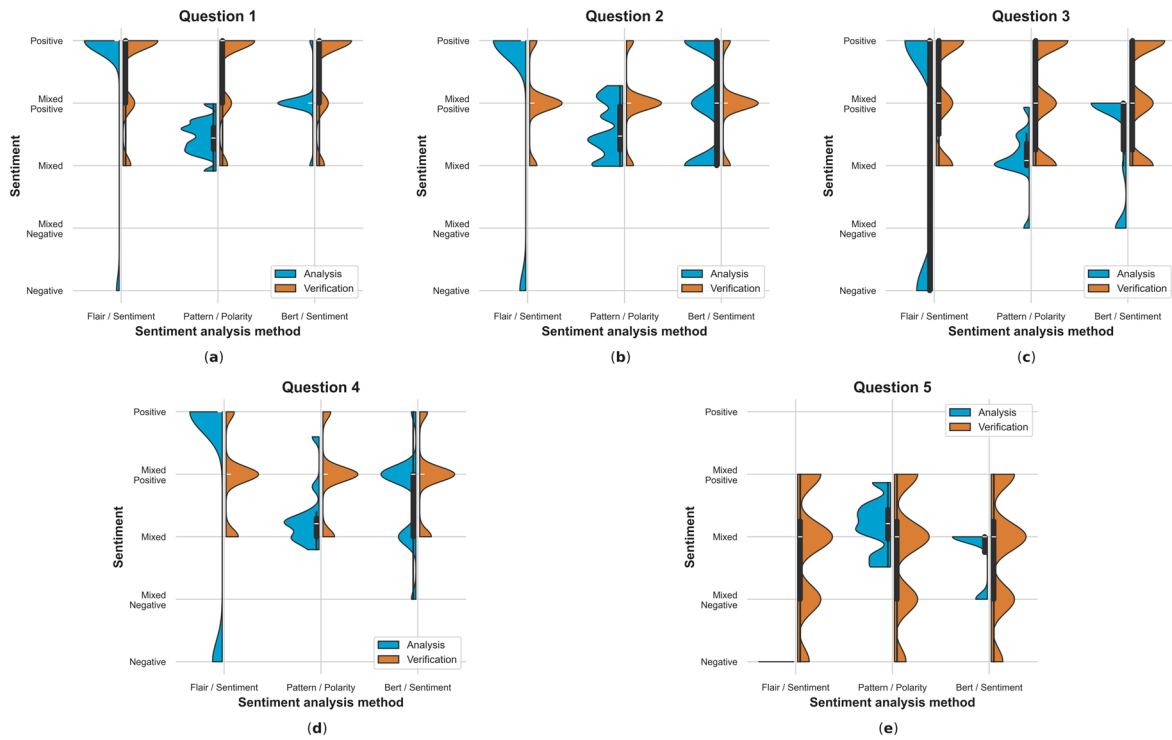

Figure A4. Histograms obtained for sentiment analysis methods and reference responses using Google translator for answers split into sentences and statistics calculation for all results obtained: (**a**–**e**) results obtained for questions 1–5; (**f**) sentiment level occurrence for all answers.

Figure A5. Probability density distributions obtained for sentiment analysis methods and reference responses using the MyMemory translator for complete answers and statistics calculation for all results obtained: (**a**–**e**) results obtained for questions 1–5.
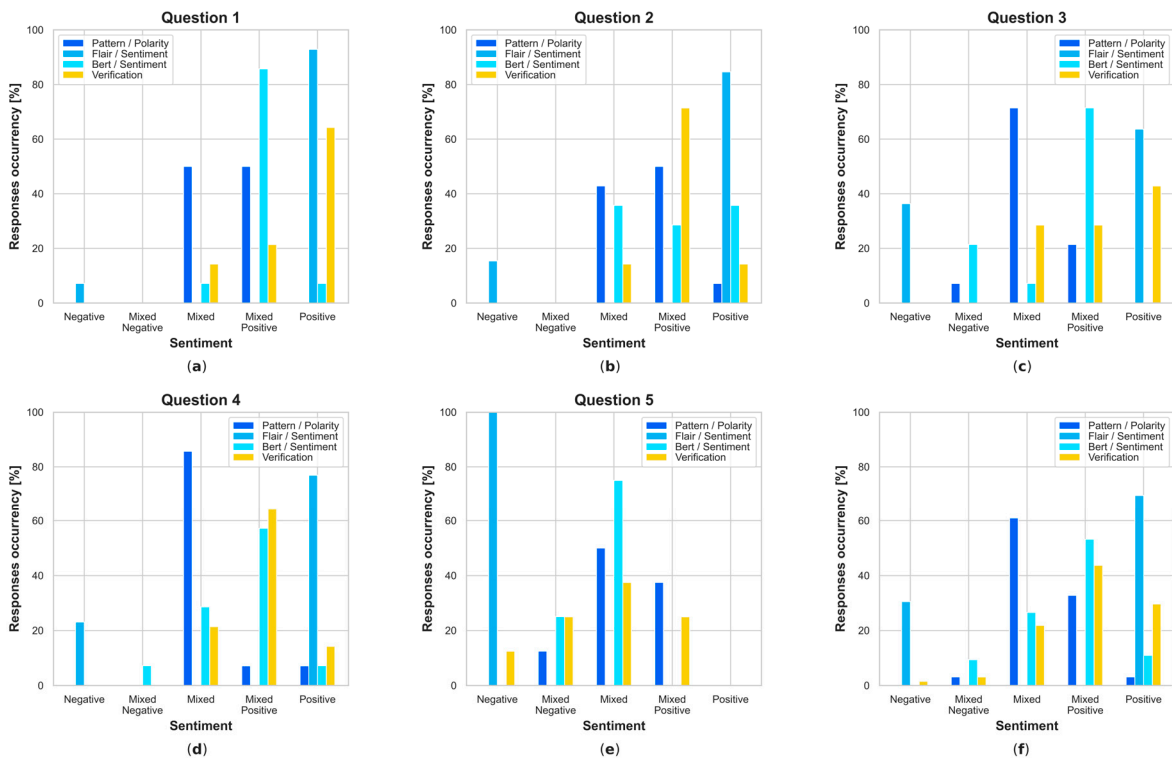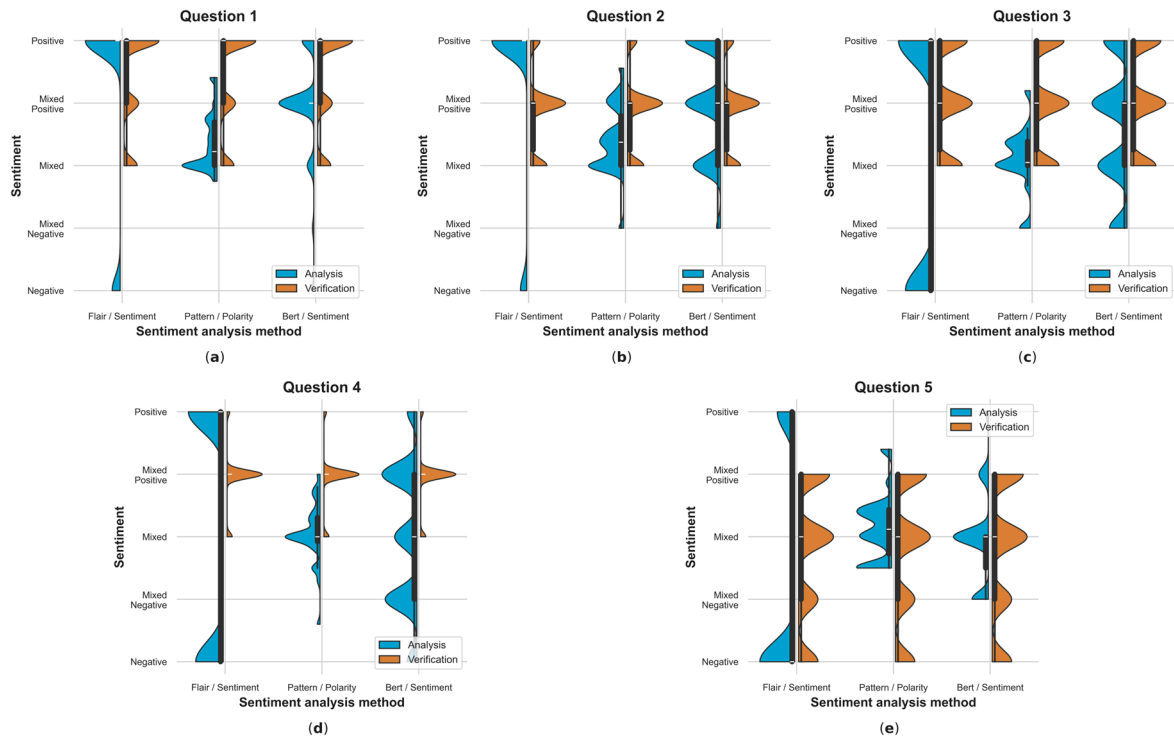

Figure A6. Histograms obtained for sentiment analysis methods and reference responses using the MyMemory translator for complete answers and statistics calculation for all results obtained: (**a**–**e**) results obtained for questions 1–5; (**f**) sentiment level occurrence for all answers.

Figure A7. Probability density distributions obtained for sentiment analysis methods and reference responses using the MyMemory translator for answers split into sentences and statistics calculation for all results obtained: (**a**–**e**) results obtained for questions 1–5.
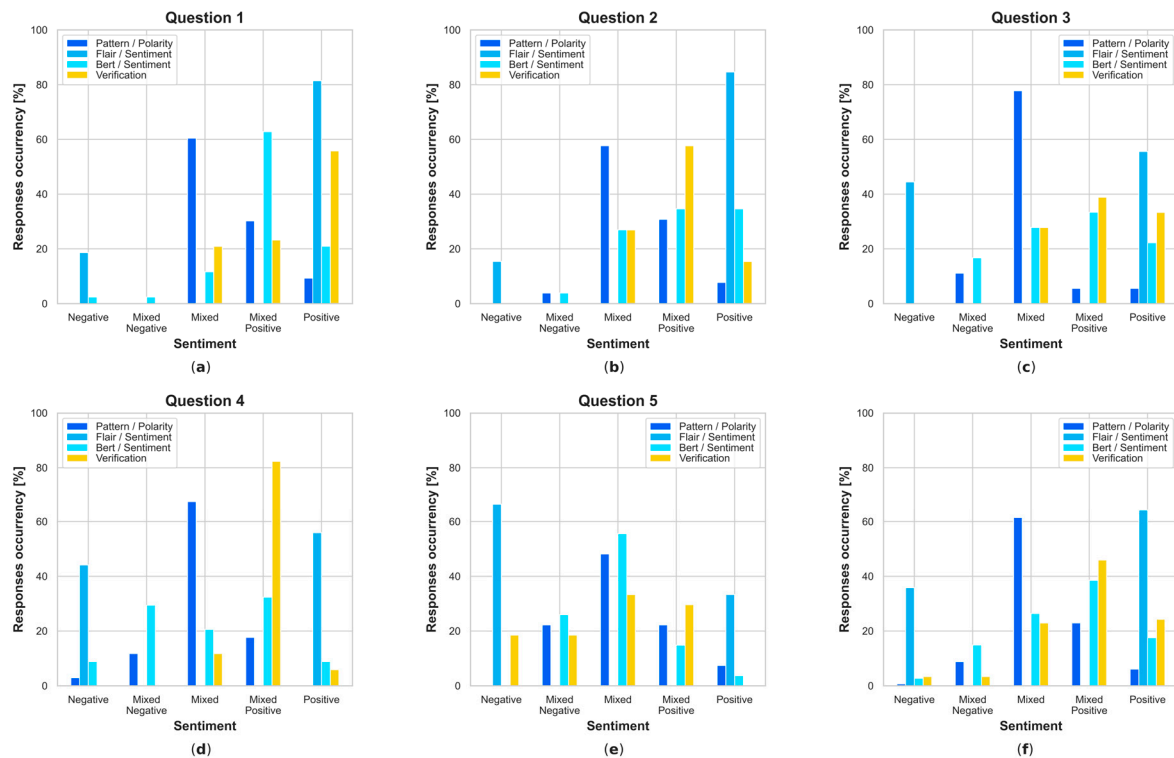

Figure A8. Histograms obtained for sentiment analysis methods and reference responses using the MyMemory translator for answers split into sentences and statistics calculation for all results obtained: (**a**–**e**) results obtained for questions 1–5; (**f**) sentiment level occurrence for all answers.
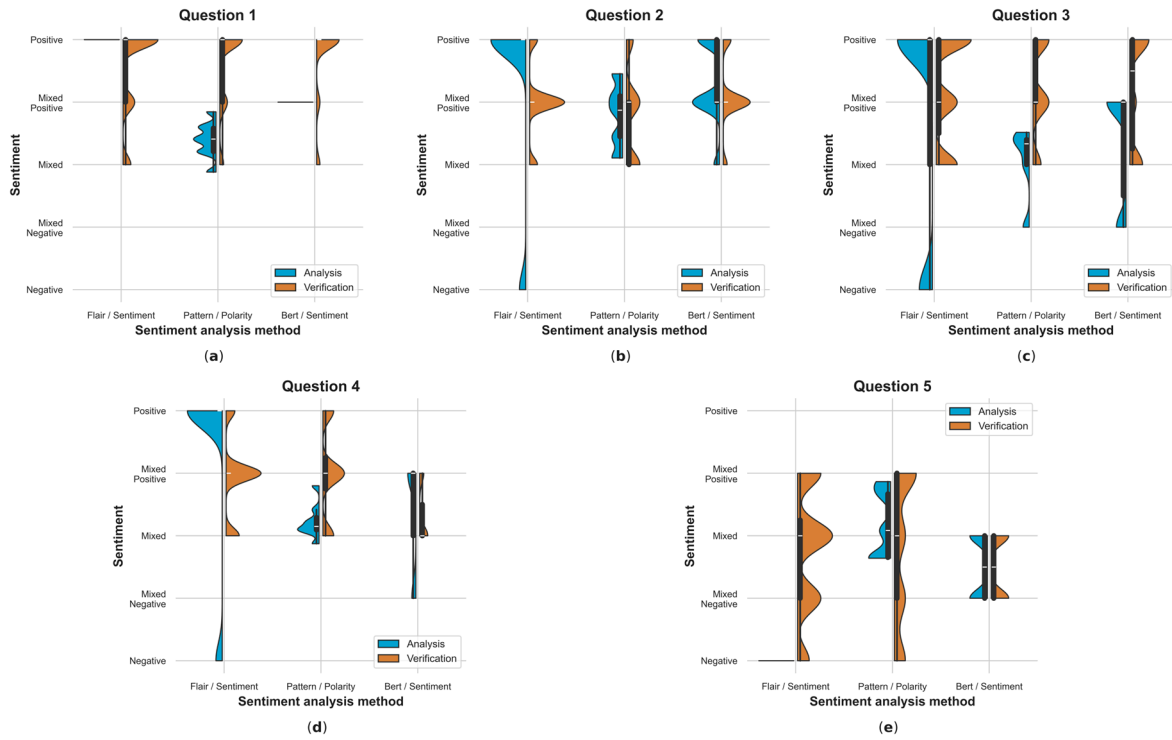
Figure A9. Probability density distributions obtained for sentiment analysis methods and reference responses using Google translator for complete answers and statistics calculation for results rejecting low reliability ones: (**a**–**e**) results obtained for questions 1–5.
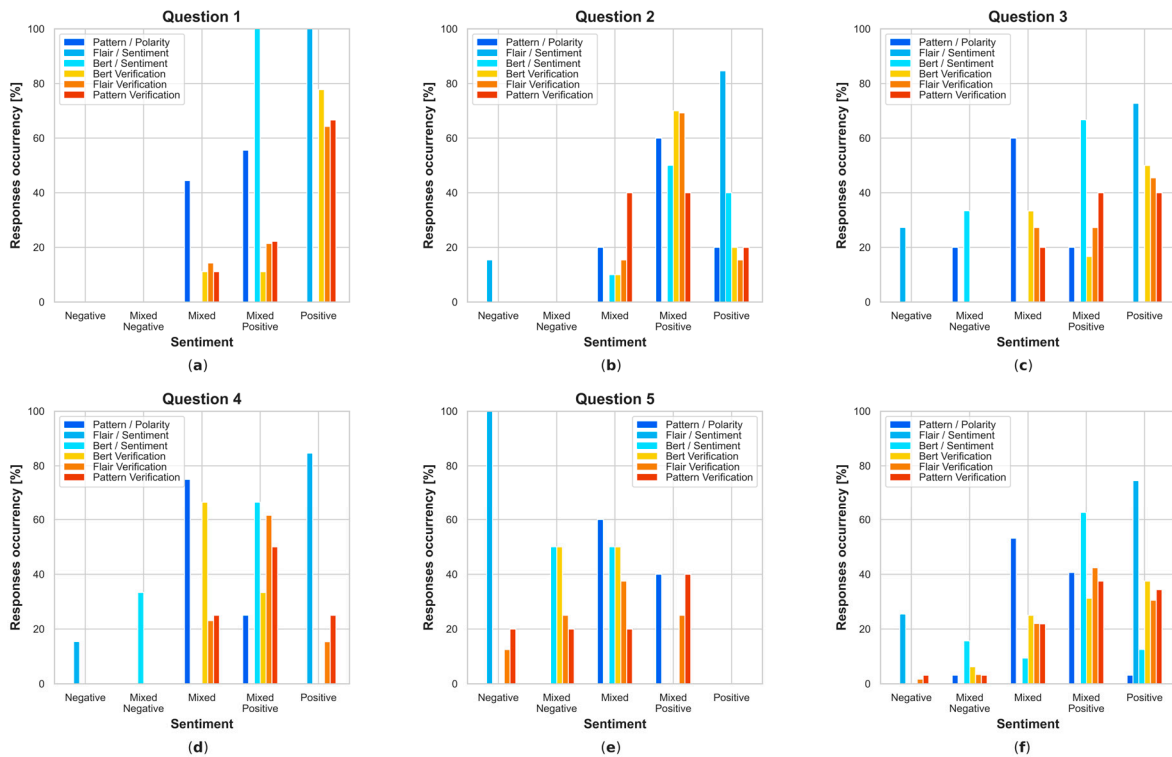

Figure A10. Histograms obtained for sentiment analysis methods and reference responses using Google translator for complete answers and statistics calculation for results rejecting low reliability ones: (**a**–**e**) results obtained for questions 1–5; (**f**) sentiment level occurrence for all answers.
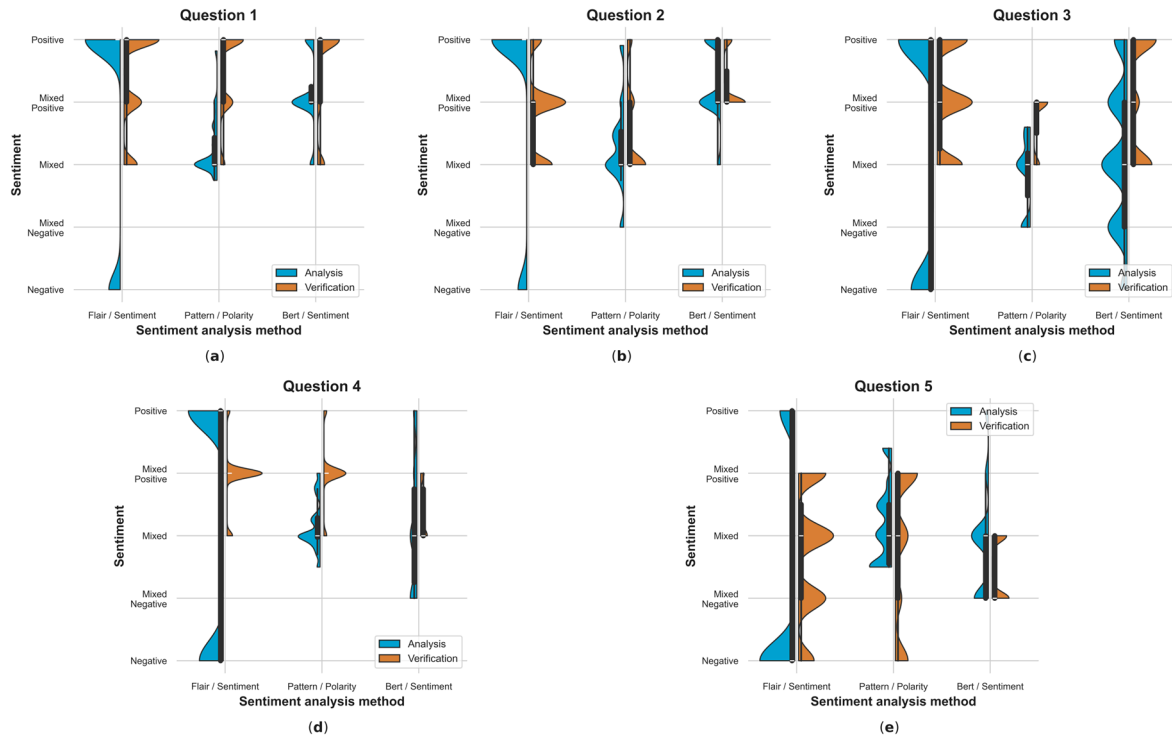
Figure A11. Probability density distributions obtained for sentiment analysis methods and reference responses using Google translator for answers split into sentences and statistics calculation for results rejecting low reliability ones: (**a**–**e**) results obtained for questions 1–5.
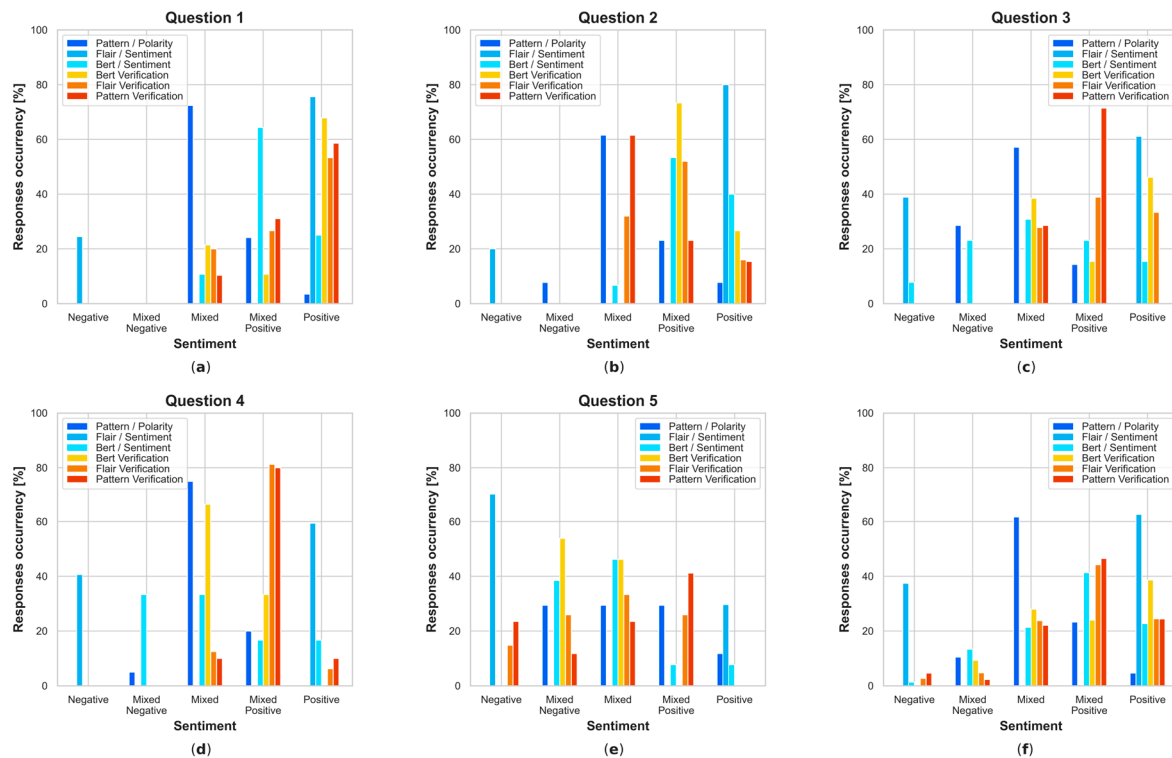

Figure A12. Histograms obtained for sentiment analysis methods and reference responses using Google translator for answers split into sentences and statistics calculation for results rejecting low reliability ones: (**a**–**e**) results obtained for questions 1–5; (**f**) sentiment level occurrence for all answers.
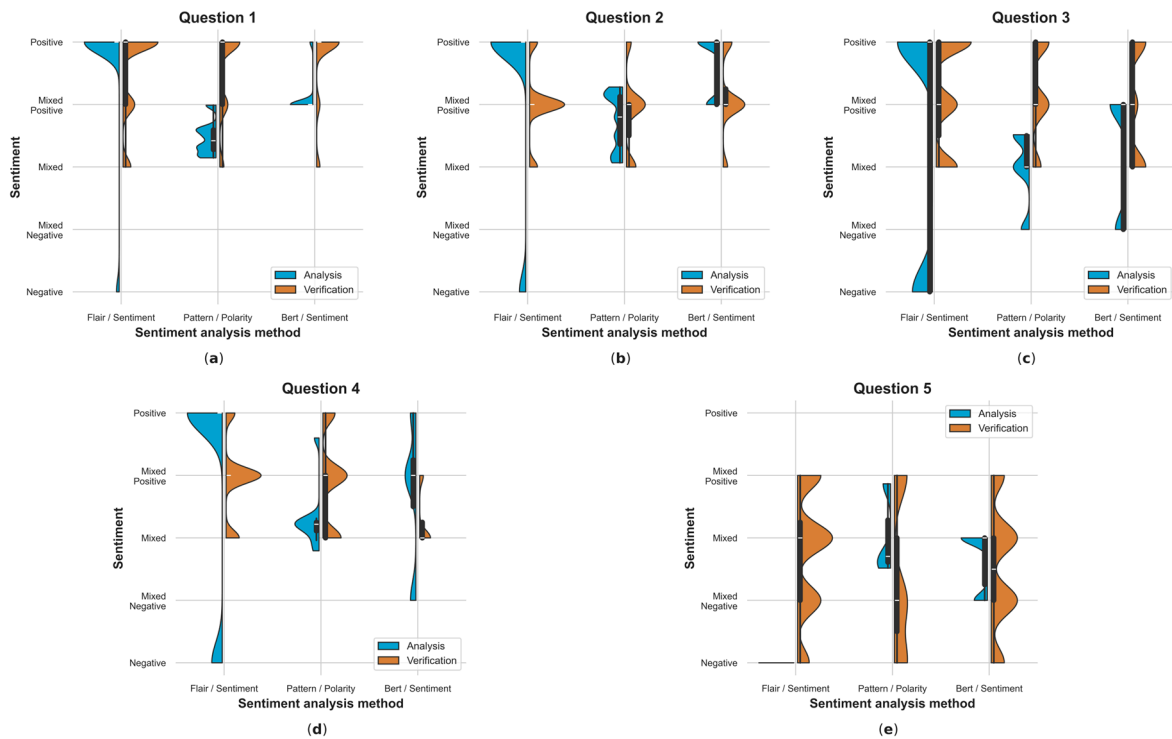
Figure A13. Probability density distributions obtained for sentiment analysis methods and reference responses using the MyMemory translator for complete answers and statistics calculation for results rejecting low reliability ones: (**a**–**e**) results obtained for questions 1–5.
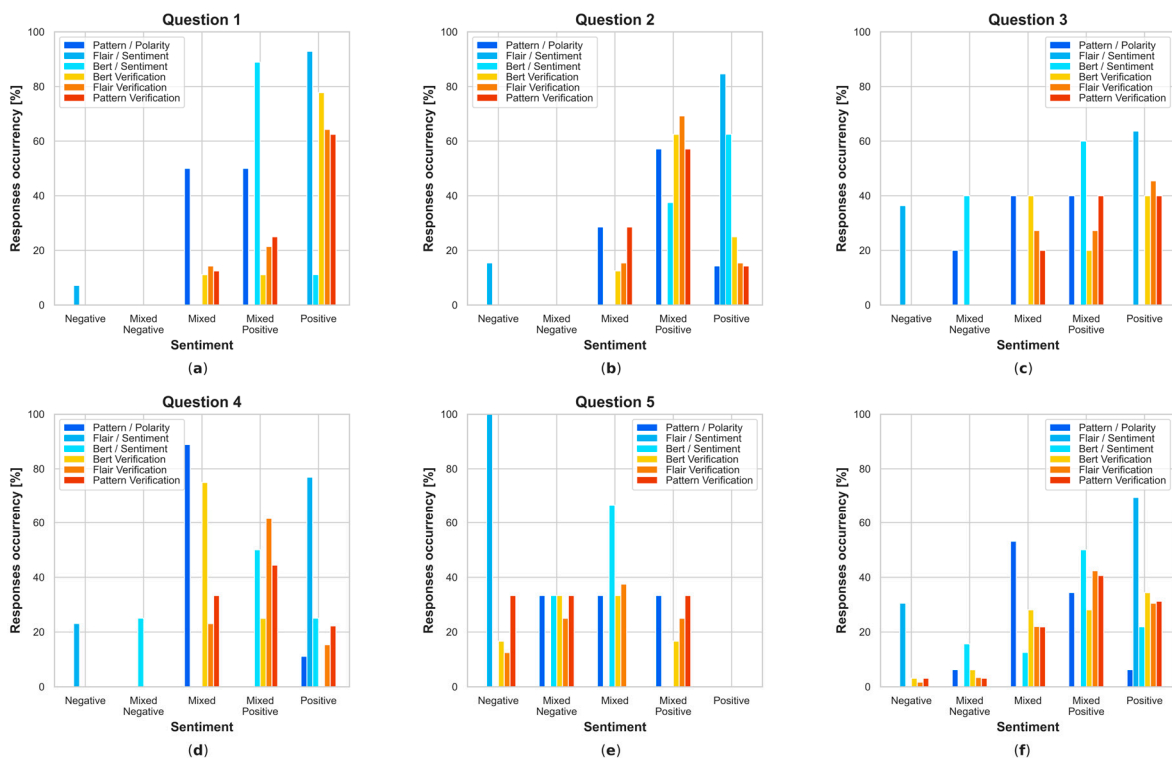

Figure A14. Histograms obtained for sentiment analysis methods and reference responses using the MyMemory translator for complete answers and statistics calculation for results rejecting low reliability ones: (**a**–**e**) results obtained for questions 1–5; (**f**) sentiment level occurrence for all answers.
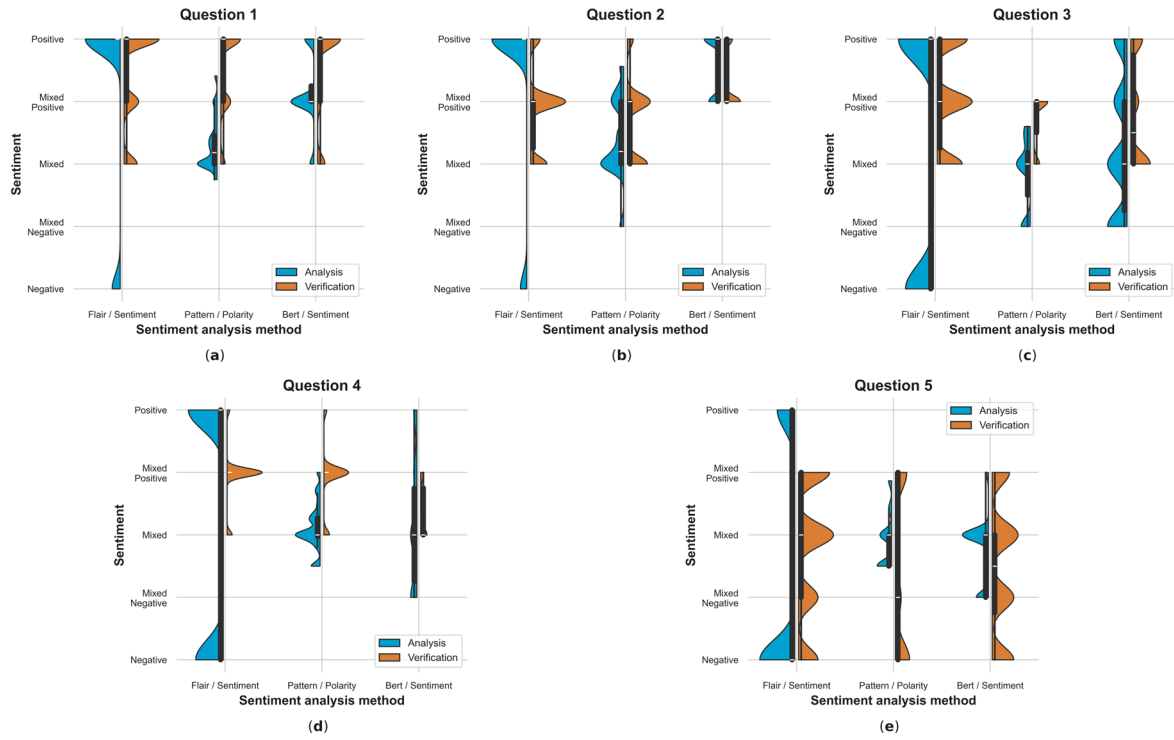
Figure A15. Probability density distributions obtained for sentiment analysis methods and reference responses using the MyMemory translator for answers split into sentences and statistics calculation for results rejecting low reliability ones: (**a**–**e**) results obtained for questions 1–5.
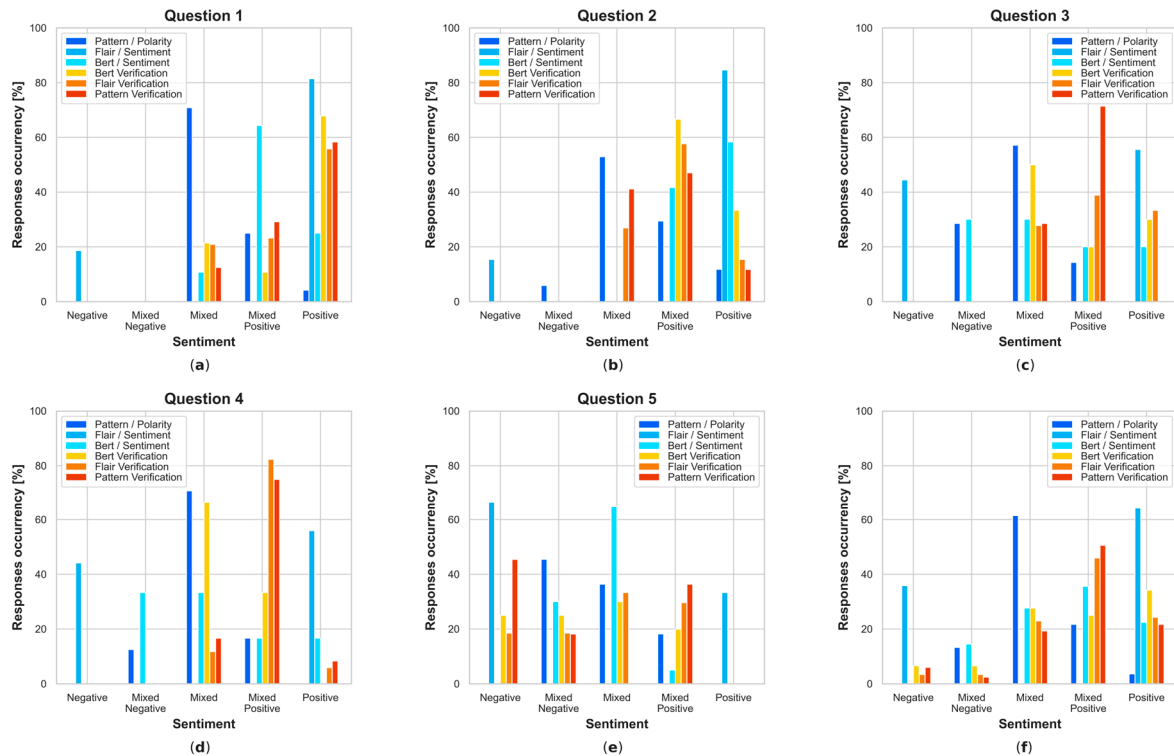

Figure A16. Histograms obtained for sentiment analysis methods and reference responses using the MyMemory translator for answers split into sentences and statistics calculation for results rejecting low reliability ones: (**a**–**e**) results obtained for questions 1–5; (**f**) sentiment level occurrence for all answers.

Figure A17. Error statistics of sentiment analysis methods for each detected sentiment level using Google translator for all results obtained considering complete answers: (**a**–**e**) results obtained for questions 1–5.
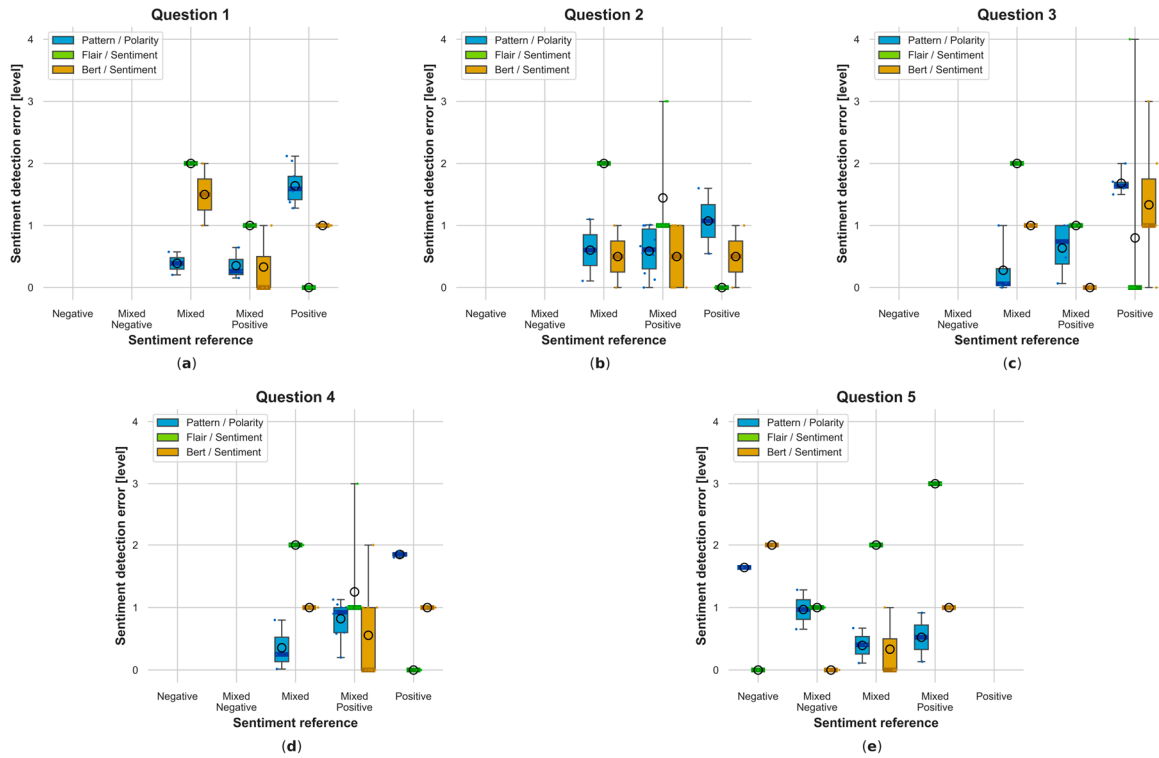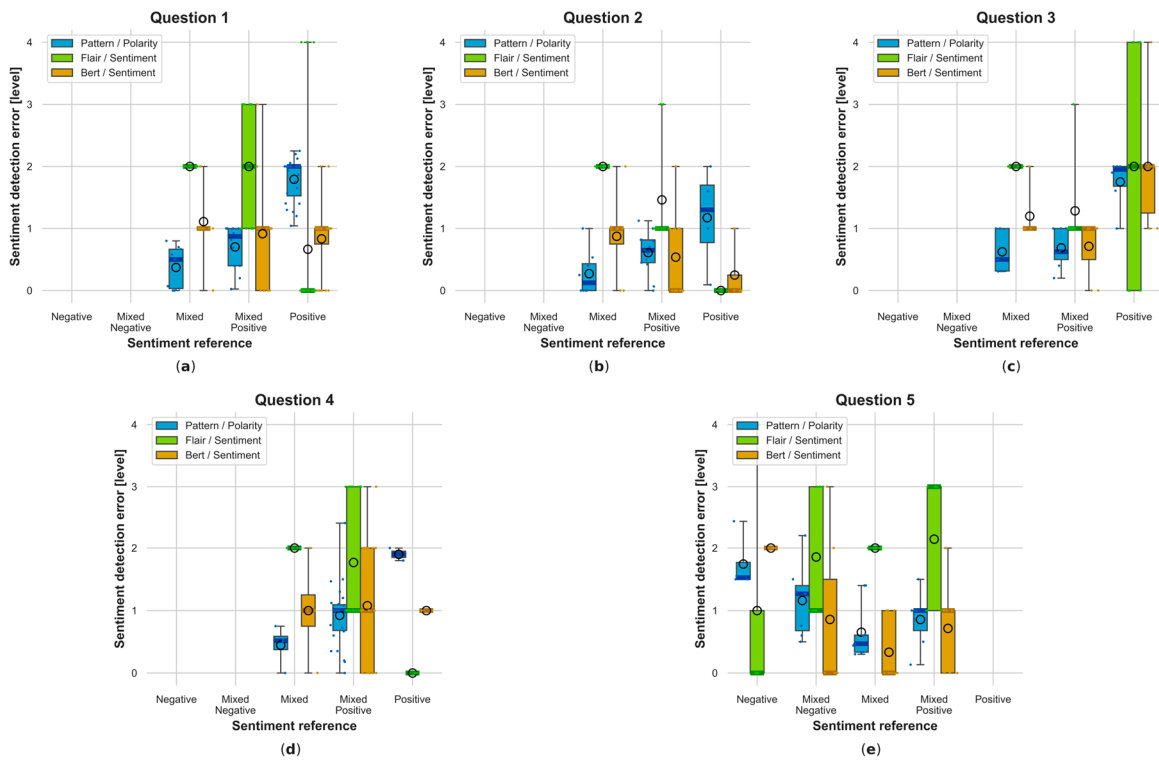

Figure A18. Error statistics of sentiment analysis methods for each detected sentiment level using Google translator for all results obtained considering answers split into sentences: (**a**–**e**) results obtained for questions 1–5.

Figure A19. Error statistics of sentiment analysis methods for each detected sentiment level using the MyMemory translator for all results obtained considering complete answers: (**a**–**e**) results obtained for questions 1–5.
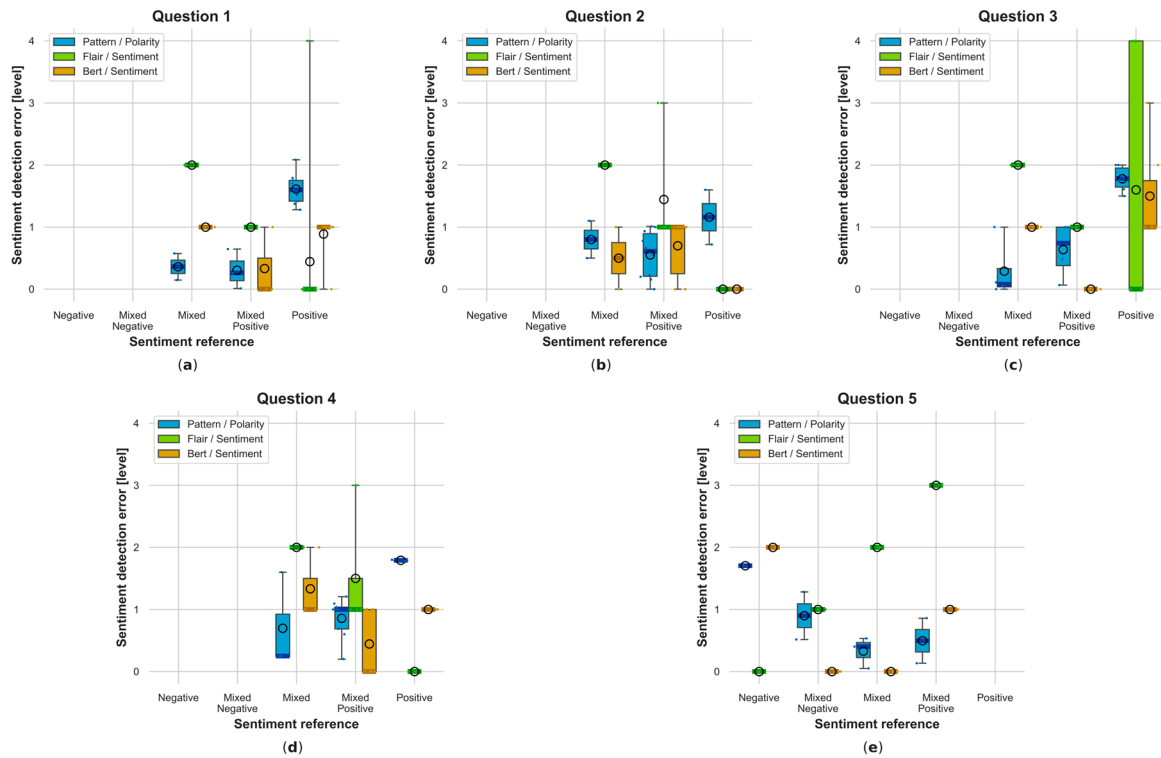

Figure A20. Error statistics of sentiment analysis methods for each detected sentiment level using the MyMemory translator for all results obtained considering answers split into sentences: (**a**–**e**) results obtained for questions 1–5.

Figure A21. Error statistics of sentiment analysis methods for each detected sentiment level using Google translator for results rejecting low reliability ones considering complete answers: (**a**–**e**) results obtained for questions 1–5.

Figure A22. Error statistics of sentiment analysis methods for each detected sentiment level using Google translator for results rejecting low reliability ones considering answers split into sentences: (**a**–**e**) results obtained for questions 1–5.

Figure A23. Error statistics of sentiment analysis methods for each detected sentiment level using the MyMemory translator for results rejecting low reliability ones considering complete answers: (**a**–**e**) results obtained for questions 1–5.
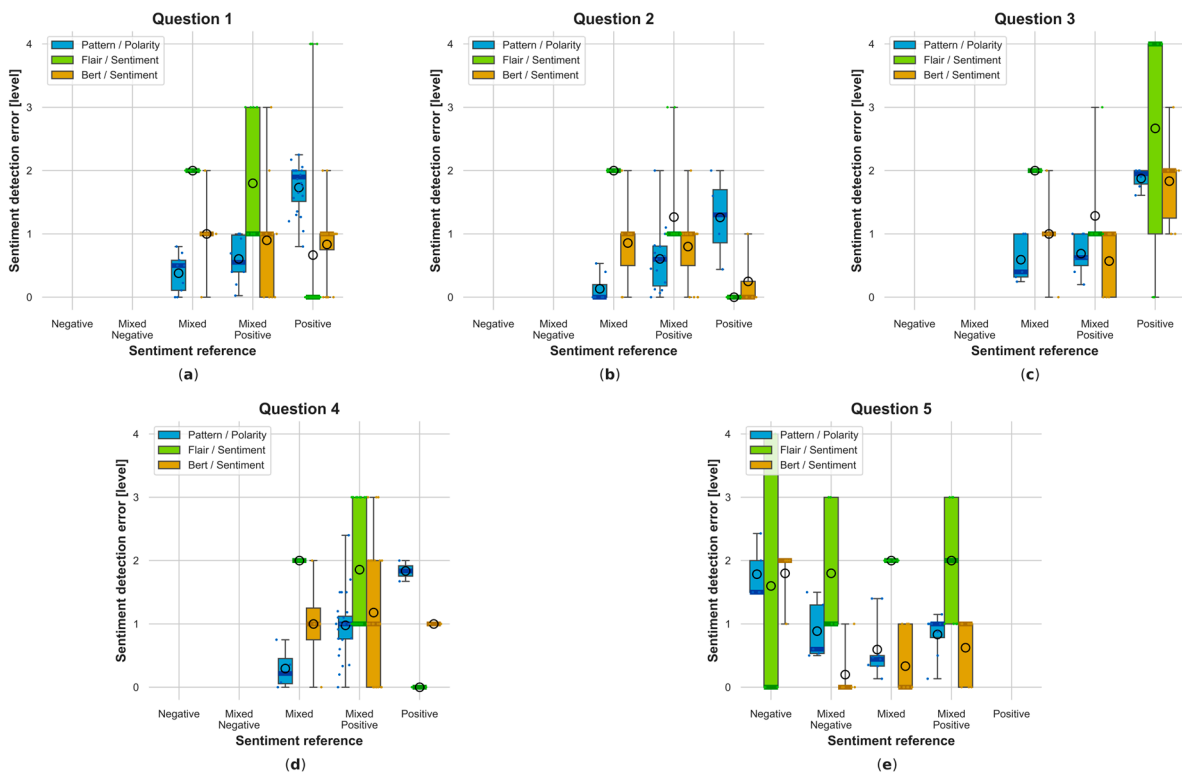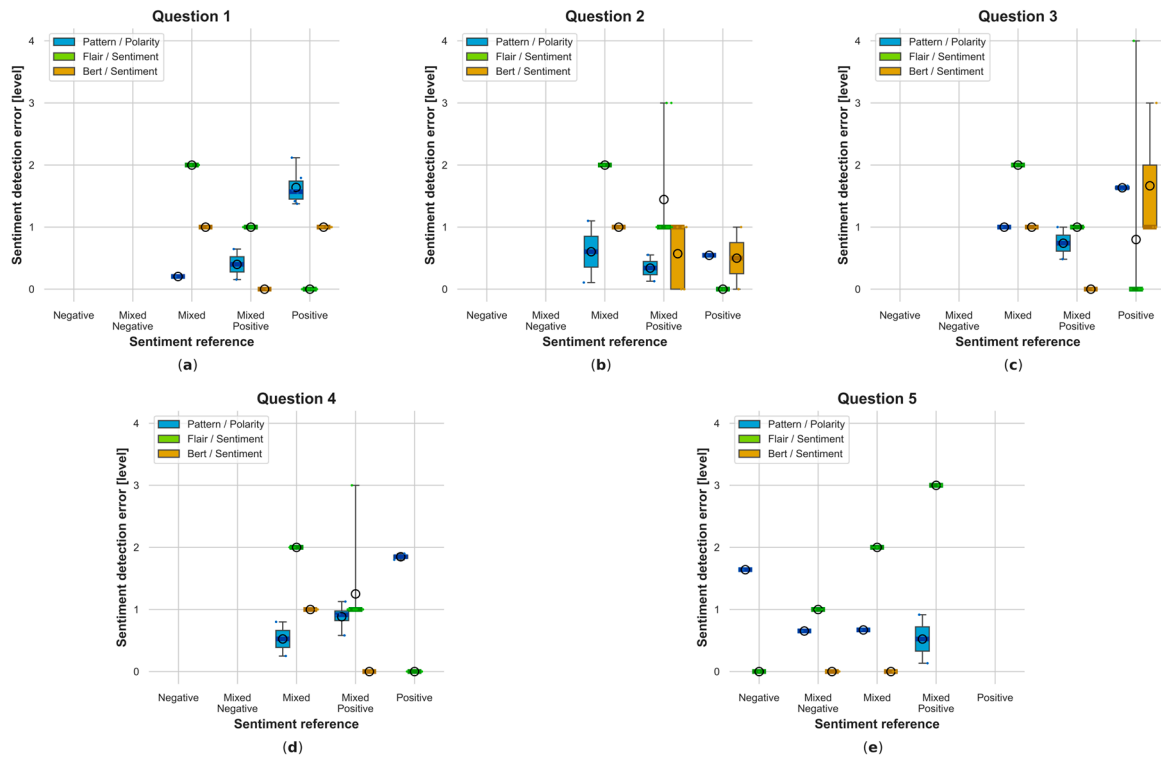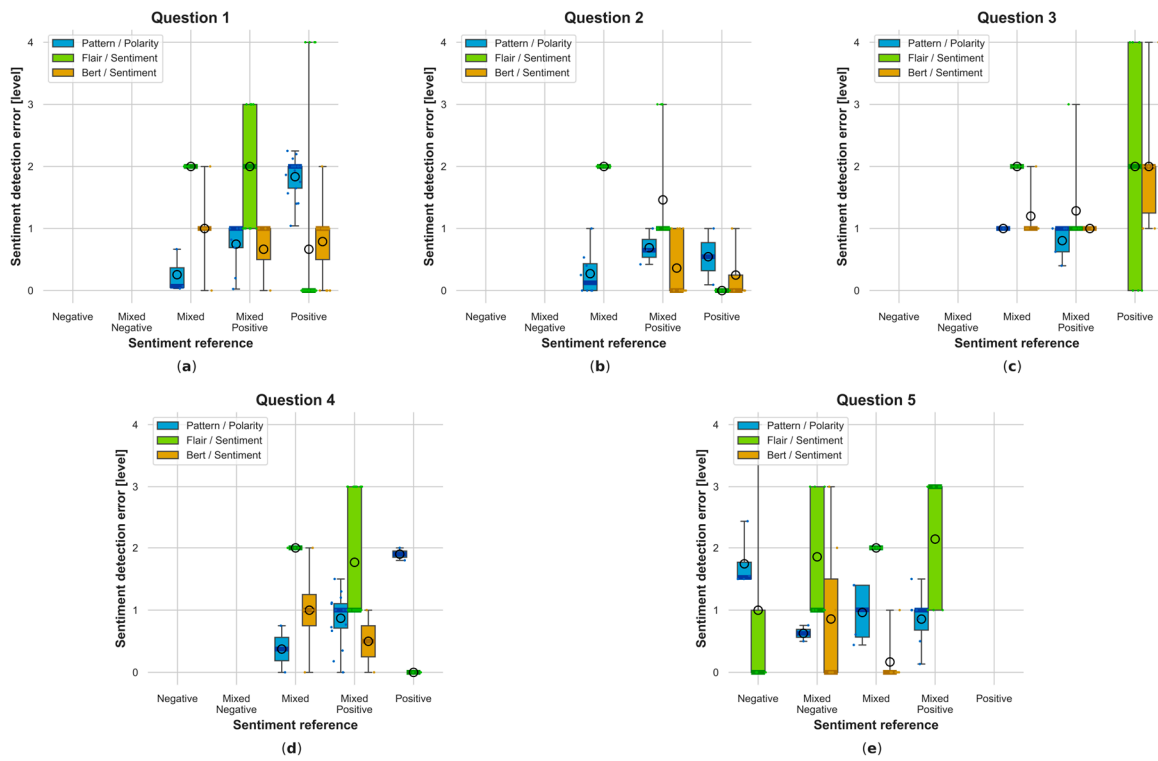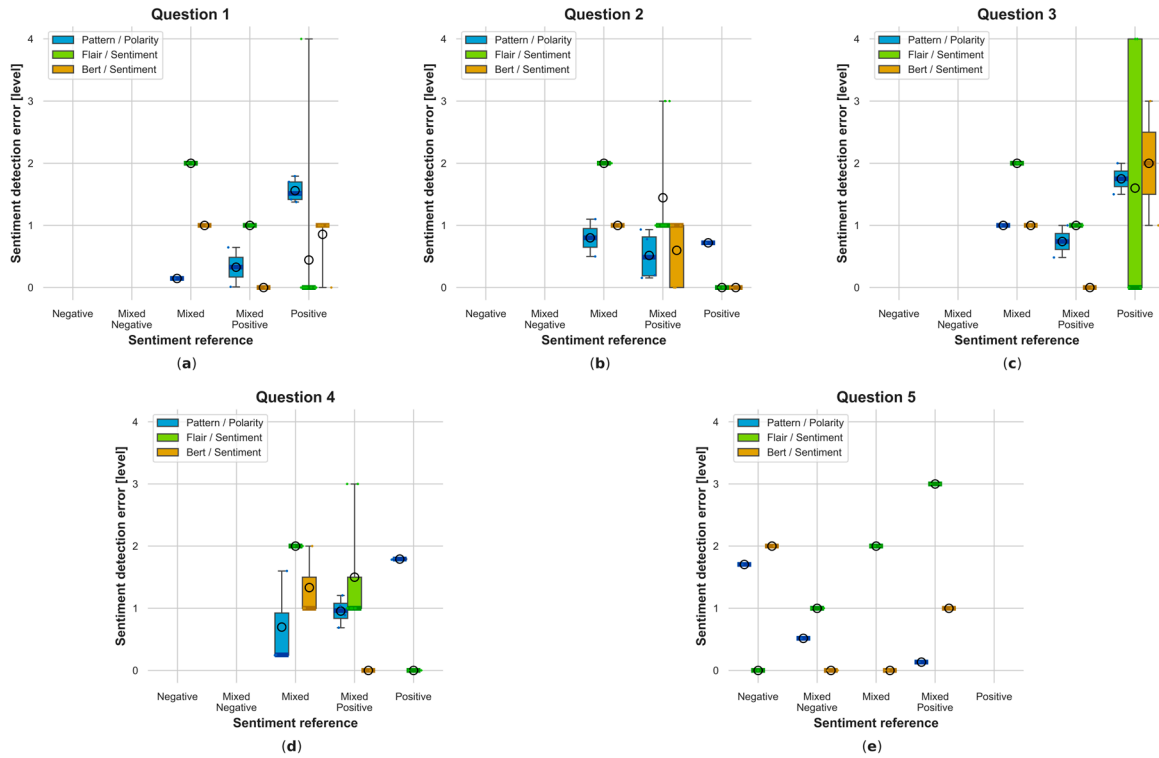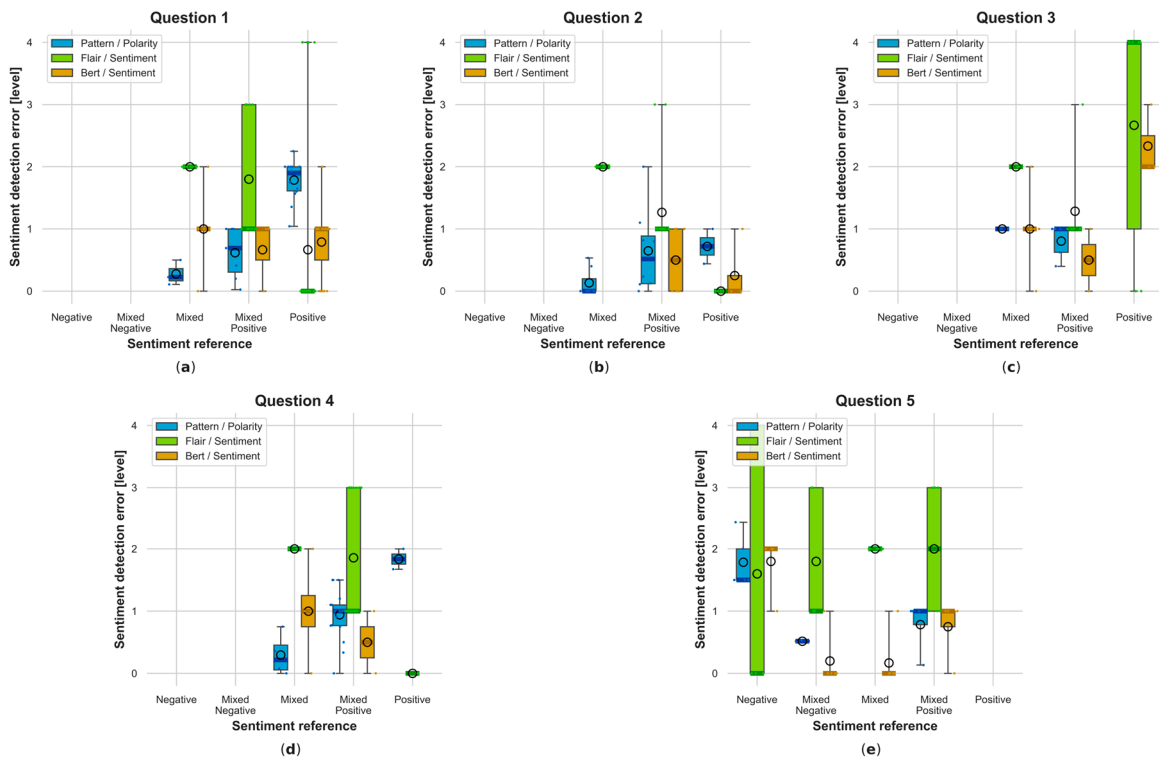

Figure A24. Error statistics of sentiment analysis methods for each detected sentiment level using the MyMemory translator for results rejecting low reliability ones considering answers split into sentences: (**a**–**e**) results obtained for questions 1–5.

References

1. Rani, S.; Kumar, P. A Sentiment Analysis System to Improve Teaching and Learning. *Computer* **2017**, *50*, 36–43. [CrossRef]
2. Karunya, K.; Aarthy, S.; Karthika, R.; Jegatha Deborah, L. Analysis of Student Feedback and Recommenda-tion to Tutors. In Proceedings of the 2020 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India, 28–30 July 2020; IEEE: Piscataway, NJ, USA; pp. 1579–1583.
3. Kastrati, Z.; Imran, A.S.; Kurti, A. Weakly Supervised Framework for Aspect-Based Sentiment Analysis on Students' Reviews of MOOCs. *IEEE Access* **2020**, *8*, 106799–106810. [CrossRef]
4. Hynninen, T.; Knutas, A.; Hujala, M. Sentiment Analysis of Open-Ended Student Feedback. In Proceedings of the 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 28 September 2020; IEEE: Piscataway, NJ, USA; pp. 755–759.
5. Katla, S.N.; Korivi, N.; Manikandan, V.M. A Sentiment Analysis-Based Intelligent System for Summarizing the Feedback of Educational Institutions. In Proceedings of the 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 6 July 2023; IEEE: Piscataway, NJ, USA; pp. 1–5.
6. Kumarasiri, A.D.S.S.; Delwita, C.E.M.S.M.; Haddela, P.S.; Samarasinghe, R.P.; Udishan, R.P.I.; Wick-ramasinghe, L. Student Teaching and Learning System for Academic Institutions. In Proceedings of the 2022 IEEE 7th International conference for Convergence in Technology (I2CT), Pune, India, 7 April 2022; IEEE: Piscataway, NJ, USA; pp. 1–6.
7. Iatrellis, O.; Samaras, N.; Kokkinos, K.; Xenakis, A. Elevating Academic Advising: Natural Language Processing of Student Reviews. *Appl. Syst. Innov.* **2024**, *7*, 12. [CrossRef]
8. Akritidis, L.; Bozanis, P. Low-Dimensional Text Representations for Sentiment Analysis NLP Tasks. *SN Comput. Sci.* **2023**, *4*, 474. [CrossRef]
9. Jelodar, H.; Wang, Y.; Orji, R.; Huang, H. Deep Sentiment Classification and Topic Discovery on Novel Coronavirus or COVID-19 Online Discussions: NLP Using LSTM Recurrent Neural Network Approach. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 2733–2742. [CrossRef]
10. Soong, H.-C.; Jalil, N.B.A.; Kumar Ayyasamy, R.; Akbar, R. The Essential of Sentiment Analysis and Opinion Mining in Social Media: Introduction and Survey of the Recent Approaches and Techniques. In Proceedings of the 2019 IEEE 9th Symposium on Computer Applications & Industrial Electronics (ISCAIE), Kota Kinabalu, Malaysia, 27–28 April 2019; pp. 272–277.
11. Nguyen, P.X.V.; Hong, T.T.T.; Van Nguyen, K.; Nguyen, N.L.-T. Deep Learning versus Traditional Classifiers on Vietnamese Students' Feedback Corpus. In Proceedings of the 2018 5th NAFOSTED Conference on Information and Computer Science (NICS), Ho Chi Minh City, Vietnam, 23–24 November 2018; IEEE: Piscataway, NJ, USA; pp. 75–80.
12. Purushotham, P.; Kiran, A. Sentiment Analysis Using Deep Learning for Students' Feedback: A Survey. In Proceedings of the 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 23 January 2023; IEEE: Piscataway, NJ, USA; pp. 1–4.
13. Jain, K.; Kaushal, S. A Comparative Study of Machine Learning and Deep Learning Techniques for Sentiment Analysis. In Proceedings of the 2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 29–31 August 2018; IEEE: Piscataway, NJ, USA; pp. 483–487.
14. Katragadda, S.; Ravi, V.; Kumar, P.; Lakshmi, G.J. Performance Analysis on Student Feedback Using Machine Learning Algorithms. In Proceedings of the 2020 6th International Conference on Advanced Compu-ting and Communication Systems (ICACCS), Coimbatore, India, 6–7 March 2020; IEEE: Piscataway, NJ, USA; pp. 1161–1163.
15. Kastrati, Z.; Dalipi, F.; Imran, A.S.; Pireva Nuci, K.; Wani, M.A. Sentiment Analysis of Students' Feedback with NLP and Deep Learning: A Systematic Mapping Study. *Appl. Sci.* **2021**, *11*, 3986. [CrossRef]
16. Mäntylä, M.V.; Graziotin, D.; Kuutila, M. The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Comput. Sci. Rev.* **2018**, *27*, 16–32. [CrossRef]
17. Umar, M.; Aliyu, M.; Modi, S. Sentiment Analysis in the Era of Web 2.0: Applications, Implementation Tools and Approaches for the Novice Researcher. *Caliphate J. Sci. Technol.* **2022**, *4*, 1–9. [CrossRef]
18. Tian, X.; Tang, S.; Zhu, H.; Xia, D. Real-time sentiment analysis of students based on mini-Xception architecture for wisdom classroom. *Concurr. Comput. Pract. Exp.* **2022**, *34*. [CrossRef]
19. Altrabsheh, N.; Cocea, M.; Fallahkhair, S. Sentiment Analysis: Towards a Tool for Analysing Real-Time Stu-dents Feedback. In Proceedings of the 2014 IEEE 26th International Conference on Tools with Artificial Intelligence, Limassol, Cyprus, 10–12 November 2014; IEEE: Piscataway, NJ, USA; pp. 419–423.
20. Altrabsheh, N.; Cocea, M.; Fallahkhair, S. Learning Sentiment from Students' Feedback for Real-Time Interventions in Classrooms. In Proceedings of the International Conference on Adaptive and Intelligent Systems, Bournemouth, UK, 8–9 September 2014; pp. 40–49.
21. Altrabsheh, N.; Cocea, M.; Fallahkhair, S.; Dhou, K. Evaluation of the SA-E System for Analysis of Students' Real-Time Feedback. In Proceedings of the 2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT), Timisoara, Romania, 3–7 July 2017; IEEE: Piscataway, NJ, USA; pp. 60–61.
22. Adachi, Y.; Negishi, T. Development and evaluation of a real-time analysis method for free-description questionnaire responses. In Proceedings of the 2020 15th International Conference on Computer Science & Education (ICCSE), Delft, The Netherlands, 18–22 August 2020; pp. 78–82.

23. Aiyanyo, I.D.; Samuel, H.; Lim, H. Effects of the COVID-19 Pandemic on Classrooms: A Case Study on Foreigners in South Korea Using Applied Machine Learning. *Sustainability* **2021**, *13*, 4986. [CrossRef]

24. Nkomo, L.M.; Daniel, B.K. Sentiment Analysis of Student Engagement with Lecture Recording. *TechTrends* **2021**, *65*, 213–224. [CrossRef] [PubMed]

25. Lau, S.L.; Sim, T.Y. Feedback of University Students on Online Delivery Learning During the COVID-19 Pandemic Period. In Proceedings of the 2020 IEEE Conference on e-Learning, e-Management and e-Services (IC3e), Kota Kinabalu, Malaysia, 17–19 November 2020; pp. 13–18.

26. Palit, S.; Nur, S.; Khatun, Z.; Rahman, M.; Ahmed, T. Analysis of Online Education System of Bangladesh during COVID-19 Pandemic Based on NLP and Machine Learning: Problem and Prospect. In Proceedings of the 2021 Emerging Trends in Industry 4.0 (ETI 4.0), Raigarh, India, 19–21 May 2021; pp. 1–6.

27. Jayanthi M, A.; Shanthi I, E. Quest_SA: Preprocessing Method for Closed-Ended Questionnaires Using Sentiment Analysis through Polarity. *Mob. Inf. Syst.* **2022**, *2022*, 4733550. [CrossRef]

28. Sindhu, I.; Daudpota, S.M.; Badar, K.; Bakhtyar, M.; Baber, J.; Nurunnabi, M. Aspect-Based Opinion Mining on Student's Feedback for Faculty Teaching Performance Evaluation. *IEEE Access* **2019**, *7*, 108729–108741. [CrossRef]

29. Sangeetha, K.; Prabha, D. Understand Students Feedback Using Bi-Integrated CRF Model Based Target Extraction. *Comput. Syst. Sci. Eng.* **2022**, *40*, 735–747. [CrossRef]

30. Pham, T.D.; Vo, D.; Li, F.; Baker, K.; Han, B.; Lindsay, L.; Pashna, M.; Rowley, R. Natural language processing for analysis of student online sentiment in a postgraduate program. *Pac. J. Technol. Enhanc. Learn.* **2020**, *2*, 15–30. [CrossRef]

31. Kiritchenko, S.; Zhu, X.; Mohammad, S.M. Sentiment Analysis of Short Informal Texts. *J. Artif. Intell. Res.* **2014**, *50*, 723–762. [CrossRef]

32. Almasani, S.A.M.; Qaid, W.A.A.; Saif, J.A.M.; Alqubati, I.A.A. Fuzzy rule based sentiment analysis for finding University Student Satisfaction in Yemen. *Indian J. Sci. Technol.* **2021**, *14*, 3264–3269. [CrossRef]

33. Alzaid, M.; Fkih, F. Sentiment Analysis of Students' Feedback on E-Learning Using a Hybrid Fuzzy Model. *Appl. Sci.* **2023**, *13*, 12956. [CrossRef]

34. Xiaoning, T. Application of Artificial Neural Network in Teaching Quality Evaluation. In Proceedings of the 2022 International Conference on Computers, Information Processing and Advanced Education (CIPAE), Ottawa, ON, Canada, 26–28 August 2022; pp. 35–38.

35. Lee, K.J. Compositional rules of Korean auxiliary predicates for sentiment analysis. *J. Korean Soc. Mar. Eng.* **2013**, *37*, 291–299. [CrossRef]

36. Truong, T.-L.; Le, H.-L.; Le-Dang, T.-P. Sentiment Analysis Implementing BERT-Based Pre-Trained Language Model for Vietnamese. In Proceedings of the 2020 7th NAFOSTED Conference on Information and Computer Science (NICS), Ho Chi Minh City, Vietnam, 26 November 2020; IEEE: Piscataway, NJ, USA; pp. 362–367.

37. Amin, M.H.S.M.; Al-Rassam, O.; Faeq, Z.S. Kurdish Language Sentiment Analysis: Problems and Challenges. *Math. Stat. Eng. Appl.* **2022**, *71*, 3282–3293. [CrossRef]

38. Wawer, A.; Sobiczewska, J. Predicting Sentiment of Polish Language Short Texts. In Proceedings of the Proceedings—Natural Language Processing in a Deep Learning World, Varna, Bulgaria, 22 October 2019; Incoma Ltd.: Shoumen, Bulgaria; pp. 1321–1327.

39. Rajput, G.K.; Kumar, A.; Kundu, S. A Comparative Study on Sentiment Analysis Approaches and Methods. In Proceedings of the 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART), Moradabad, India, 4 December 2020; IEEE: Piscataway, NJ, USA; pp. 427–431.

40. Radecki, A.; Rybicki, T. An Accurate State Visualization of Multiplexed and PWM Fed Peripherals in the Virtual Simulators of Embedded Systems. *Appl. Sci.* **2022**, *12*, 3137. [CrossRef]

41. Radecki, A.; Rybicki, T. Simulation Oriented Layer of Embedded Software Architecture for Rapid Development of Custom Embedded Systems Virtual Simulators Used in Didactics. *Appl. Sci.* **2022**, *12*, 6322. [CrossRef]

42. Das, N.; Gupta, S.; Das, S.; Yadav, S.; Subramanian, T.; Sarkar, N. A Comparative Study of Sentiment Analy-sis Tools. In Proceedings of the 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Chennai, India, 24 September 2021; IEEE: Piscataway, NJ, USA; pp. 1–7.

43. Kaur, C.; Sharma, A. Social Issues Sentiment Analysis Using Python. In Proceedings of the 2020 5th Interna-tional Conference on Computing, Communication and Security (ICCCS), Patna, India, 14 October 2020; IEEE: Piscataway, NJ, USA; pp. 1–6.

44. Che, W.; Zhao, Y.; Guo, H.; Su, Z.; Liu, T. Sentence Compression for Aspect-Based Sentiment Analysis. *IEEE/ACM Trans. Audio, Speech, Lang. Process.* **2015**, *23*, 2111–2124. [CrossRef]

45. Izsak, P.; Guskin, S.; Wasserblat, M. Training Compact Models for Low Resource Entity Tagging Using Pre-Trained Language Models. In Proceedings of the 2019 Fifth Workshop on Energy Efficient Machine Learning and Cognitive Computing—NeurIPS Edition (EMC2-NIPS), Vancouver, BC, Canada, 13 December 2019; IEEE: Piscataway, NJ, USA; pp. 44–47.

46. De Smedt, T.; Daelemans, W. Pattern for Python. *J. Mach. Learn. Res.* **2012**, *13*, 2063–2067.

47. Arifiyanti, A.A.; Kartika, D.S.Y.; Prawiro, C.J. Using Pre-Trained Models for Sentiment Analysis in Indonesian Tweets. In Proceedings of the 2022 6th International Conference on Informatics and Computational Sciences (ICICoS), Semarang, Indonesia, 28 September 2022; IEEE: Piscataway, NJ, USA; pp. 78–83.

48. Kavitha, M.; Naib, B.B.; Mallikarjuna, B.; Kavitha, R.; Srinivasan, R. Sentiment Analysis Using NLP and Machine Learning Techniques on Social Media Data. In Proceedings of the 2022 2nd International Confer-ence on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 28 April 2022; IEEE: Piscataway, NJ, USA; pp. 112–115.
49. Wang, X. Automatic Scoring of English Online Translation Based on Machine Learning Algorithm. In Proceedings of the 2022 International Conference on Artificial Intelligence of Things and Crowdsensing (AIoTCs), Nicosia, Cyprus, 26–28 October 2022; IEEE: Piscataway, NJ, USA; pp. 111–115.
50. Brkić, M.; Bašic Mikulić, B.; Matetić, M. Can We Beat Google Translate? In Proceedings of the ITI 2012 34th International Conference on Information Technology Interfaces, Cavtat/Dubrovnik, Croatia, 25–28 June 2012; pp. 381–386.
51. Amin, A.; Hossain, I.; Akther, A.; Alam, K.M. Bengali VADER: A Sentiment Analysis Approach Using Modified VADER. In Proceedings of the 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox's Bazar, Bangladesh, 7–9 February 2019; IEEE: Piscataway, NJ, USA; pp. 1–6.
52. Institute of Automatic Control, Lodz University of Technology. Available online: https://www.automatyka.p.lodz.pl/?lang=en (accessed on 21 March 2024).
53. Microprocessor Systems Laboratory. Available online: https://automatyka.p.lodz.pl/en/microprocessor-systems-laboratory (accessed on 21 March 2024).
54. Mroczek, H. *Microprocessor Technique*; Lodz Uniwersity of Technology Press: Łódź, Poland, 2007; ISBN 978-83-7283-238-2. (In Polish)
55. Ghosh, A.; Bershteyn, M.; Casley, R.; Chien, C.; Jain, A.; Lipsie, M.; Tarrodaychik, D.; Yamamoto, O. Hardware-Software Co-Simulator for Embedded System Design and Debugging. In Proceedings of the Asia and South Pacific Design Automation Conference ASP-DAC, Chiba, Japan, 29 August–1 September 1995; pp. 155–164. [CrossRef]
56. Zheng, S. TPA31xxDx Bootstrap Circuit. Application Report No. SLOA259, Texas Instruments, Dallas, Texas, November 2017. Available online: http://www.ti.com (accessed on 24 January 2011).
57. WIKAMP Lodz University of Technology Teaching Platform. Available online: https://edu.p.lodz.pl/?lang=en (accessed on 21 March 2024).
58. Wang, X. Designing Method of Art Education Course Based on Moodle Platform. In Proceedings of the 2019 International Conference on Smart Grid and Electrical Automation (ICSGEA), Xiangtan, China, 10–11 August 2019; IEEE: Piscataway, NJ, USA; pp. 361–365.
59. Taber, K.S. Lost and found in translation: Guidelines for reporting research data in an 'other' language. *Chem. Educ. Res. Pract.* **2018**, *19*, 646–652. [CrossRef]
60. Michael, J.S. Nuance Lost in Translation. *NTM Z. für Gesch. der Wiss. Tech. und Med.* **2017**, *25*, 281–309. [CrossRef]