

Article

Semi-Automatic Approaches for Exploiting Shifter Patterns in Domain-Specific Sentiment Analysis

Pavel Brazdil ^{1,2,*}, Shamsuddeen H. Muhammad ^{2,3,*} , Fátima Oliveira ⁴, João Cordeiro ^{2,5} , Fátima Silva ⁴, Purificação Silvano ⁴ and António Leal ⁴

¹ FEP, University of Porto, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal

² INESC TEC—Institute for Systems and Computer Engineering, Technology and Science, Campus da FEUP, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

³ FCUP—Faculty of Science, University of Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal

⁴ FLUP/CLUP, University of Porto, Via Panorâmica, s/n, 4150-564 Porto, Portugal

⁵ FE/HULTIG, University of Beira Interior, Rua Marquês d'Ávila e Bolama, 6201-001 Covilhã, Portugal

* Correspondence: pbrazdil@fep.up.pt (P.B.); shmuhammad.csc@buk.edu.ng (S.H.M.)

Abstract: This paper describes two different approaches to sentiment analysis. The first is a form of symbolic approach that exploits a sentiment lexicon together with a set of shifter patterns and rules. The sentiment lexicon includes single words (unigrams) and is developed automatically by exploiting labeled examples. The shifter patterns include intensification, attenuation/downtoning and inversion/reversal and are developed manually. The second approach exploits a deep neural network, which uses a pre-trained language model. Both approaches were applied to texts on economics and finance domains from newspapers in European Portuguese. We show that the symbolic approach achieves virtually the same performance as the deep neural network. In addition, the symbolic approach provides understandable explanations, and the acquired knowledge can be communicated to others. We release the shifter patterns to motivate future research in this direction.

Keywords: sentiment analysis; automatic lexicon generation; domain-specific lexicon; contextual shifters; intensification, attenuation and negation; deep learning in sentiment analysis

MSC: 68T50



Citation: Brazdil, P.; Muhammad, S.H.; Oliveira, F.; Cordeiro, J.; Silva, F.; Silvano, P.; Leal, A. Semi-Automatic Approaches for Exploiting Shifter Patterns in Domain-Specific Sentiment Analysis. *Mathematics* **2022**, *10*, 3232. <https://doi.org/10.3390/math10183232>

Academic Editor: José Antonio Sanz

Received: 15 July 2022

Accepted: 29 August 2022

Published: 6 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

We describe two different approaches to domain-sensitive sentiment analysis. The first is a symbolic approach, which exploits a sentiment lexicon together with a set of shifter patterns and rules, sometimes referred to as *contextual shifters*.

These are used to derive sentiment values of sentences in new texts [1]. The second approach exploits deep neural networks (NN) that perform well on many tasks, including sentiment analysis [2].

As often users want to obtain explanations regarding why a certain polarity was attributed to the given text, it is useful to develop solutions that provide such interpretability. Symbolic approaches are interpretable and could be preferred in domains that require an explainable solution. However, a question arises as to whether their performance is not far behind the deep NNs. Our aim is to investigate this issue and provide an answer to this question by applying both approaches to the task of predicting sentiment values of given sentences on economics and finance in European Portuguese (EP).

For the symbolic approach, the sentiment lexicon is developed using an automatic approach by exploiting labeled examples. The examples are in the form of relatively short fractions of given sentences that are relevant for a particular domain. The quality of the induced lexicon depends on the number of available labeled texts. However, since the labeling is conducted by human experts, the amount of labeled data is usually limited.

Therefore, we restrict the automatic approach to single words (unigrams). Phrases, such as *crescimento alto* (high growth), rely on *shifter patterns* that represent *contextual shifters* of the sentiment values of particular words or expressions. Schulder et al. [1,3] stated that content words, such as verbs, nouns and adjectives, can act as polarity shifters. In addition, adverbs, negative function words and modal verbs also act as polarity shifters.

These contextual shifters modify the polarity of words in their scope. We adopt a similar stand; however, as we work with sentiment values, we prefer to use the term “value shifters”, instead of “polarity shifters”. The shifters can be of different types, depending on whether the operation involved is intensification, downtoning/attenuation or reversal [4]. All of these patterns are manually identified from the labeled texts. The sentiment value of the pattern is obtained automatically by applying the rules described in Section 3.3.

The state-of-the-art approach for sentiment analysis uses deep-learning approaches and provides better performance compared with other approaches [2]. The drawback of these approaches is that they do not offer explanations directly on how the prediction was achieved and, hence, are often referred to as black-box methods. This problem can be mitigated by using systems, such as LIME, which provide an explanation of the decision associated with a particular example.

This is performed by identifying the most important features that affected the decision in each example. Our symbolic system provides explanations directly without recourse to additional software. Recently, massively pre-trained language models (PLMs), including BERT-based language models, such as Bert [5] and Roberta [6], have been used for sentiment analysis. In this work, we use a multi-lingual variant of Bert and fine-tune it on Portuguese texts.

The remainder of the paper is organized as follows. The next section discusses related work. Section 3 discusses various aspects of the symbolic approach. Section 3.1 describes the automatic generation method of the sentiment lexicon. Section 3.2 is dedicated to shifter patterns and the rules that are used to derive the sentiment value of specific patterns. Section 3.3 describes how the shifter patterns are processed by the system. Section 3.4 discusses the approach that includes deep NNs—namely, mBERT. Section 4 discusses the experimental setup and the results obtained. Section 5 presents some directions for future work and our conclusions.

2. Related Work

Sentiment analysis has attracted a great deal of attention, and consequently there are many different approaches. These include a lexicon-based approach, classical machine learning-based approach and a deep learning-based approach. Hybrid approaches also exist that combine different aspects of the above approaches. In this section, we discuss these approaches and some comparative studies.

2.1. Lexicon-Based Approaches

Lexicon-based approaches can be divided into different groups, depending on whether the lexicon construction is manual, semi-automatic or automatic. Forte et al. [7] analyzed users’ comments in Portuguese on certain products and services. They used an existing sentiment lexicon as the basis for further domain-specific extensions generated manually, which led to a marked improvement in predictive performance. Similarly, Silva et al. [8] used a similar approach and extended an existing lexicon to the economic and financial domain.

Moreno-Ortiz et al. [9] used a domain-specific lexicon for sentiment analysis applications in English. Their lexicon (SentiEcon) contains 6470 entries (single words and multiwords) annotated with semantic orientation and intensity. This sentiment lexicon is intended for use in the financial/economic domain in conjunction with a general sentiment lexicon.

The disadvantage of manual approaches is that the domain-specific extensions require manual effort and this is why various authors investigated automatic approaches for the construction of lexicons. In the work of Almatarneh et al. [10], words were classified into

three categories—negative, neutral and positive—depending on the rating. Sentiment values are inferred from frequencies in the negative and positive categories. This approach has the disadvantage of considering all ratings in the negative (or positive) category as equivalent. In other words, it does not exploit the valence within the positive/negative category.

The approach described in our paper improves this shortcoming. Muhammad et al. (2020) [11] and Brazdil et al. (2021) [12] improved this approach by processing all tokens in the labeled texts and, for each one, constructing a rating distribution. This distribution was used to infer the sentiment value of the given token. Tokens with values different from zero are added to the sentiment lexicon. This approach is used as the basis for the method adopted in our work.

The automatic approaches can be exploited to generate domain-specific lexicons. As is well known, the same word may convey different sentiment values in different domains. Hence, some words in a general-purpose lexicon may have a different semantic orientation from a given domain-specific lexicon. For example, the word *grown* in *the share has grown* has a positive orientation, as is the case in the financial domain. On the other hand, the same word that appears in a sentence, such as *the tumor has grown*, from the medical domain, has a negative semantic orientation. Therefore, there is a need to adapt or enhance the general-purpose lexicon to use in specific domains [13]. Automated approaches are of great help in this regard.

Semi-automatic approaches to lexicon enhancement start with a relatively small lexicon that contains certain important *seed words* provided manually. The process of label propagation is used to transfer the values from the seed words to other terms with similar meaning (e.g., synonyms). There are several methods for this, including, for example, Wordnet synsets or embeddings [14,15].

2.2. Employing Shifters in Sentiment Analysis

Lexicon-based approaches to sentiment analysis discussed in the previous section have one major shortcoming: lexicons that include only unigrams do not capture contextual information. For example, the sentiment conveyed by *inverta o crescimento de economia* (inverts the growth of economy) is negative, even if the sentiment value of some of the words is positive. One way to resolve this is employing the so-called *sentiment shifters*, which are discussed next. A shifter is a word or expression that changes the polarity of another word or expression [4]. In our example above, the word *inverta* (inverts) is a shifter that changes the polarity of the sentence *o crescimento de economia* (growth of the economy).

In an earlier work, *sentiment shifters* were normally identified manually [4]. Shulder et al. [1] explored a semi-automatic method to extend a given set of *seed shifters*. The work focused on shifters in English and was referred to as a *bootstrapping approach*. The authors employed a set of different resources and a machine learning (ML) system. The process starts with a selection of a certain number (2000) of verbs chosen at random from WordNet. Then, the process was repeated for nouns and adjectives. The selected cases were then labeled manually as *shifters* or *non-shifters*.

For instance, the set of initial seed verbs resulted in 304 verbal shifters, i.e., 15.2% of the initial set. The proportion of shifters in the initial set of nouns and adjectives was much smaller. The resulting shifters represent the so-called *gold set*. Words capturing negation, such as *no*, *not*, *neither* and *without* were handled separately. The authors then defined various features that were used to characterize different cases.

One group of features was called *task-specific features*. These included, for instance, *distributional similarity*, which explores the fact that negation words and some shifters occur in similar contexts. Another group of features was referred to as *generic features*. These captured certain properties derived from, for instance, WordNet, FrameNet and the nominalization lexicon NOMLEX for English. The resulting dataset was used to train a classifier (SVM). This classifier was then applied to new cases characterized in a similar way by the features discussed above. As this process is not entirely reliable, the resulting set of shifters was edited manually, and the final set was added to the initial gold set.

Thus, for instance, the bootstrap process was able to identify additional 676 verbs that were added to the initial gold set. The authors showed that it was possible to increase the generalization performance on a given test data using this method.

Although our approach identified shifters manually, it has other advantages over the automatic approach discussed. First, it uses numeric sentiment values and can thus deal with intensification and attenuation that can affect the final polarities of phrases in which the corresponding shifter appears. Therefore, the concept of *shifter* in our work is more general. Negation words are also regarded as shifters. The second advantage is that we use various shifter patterns, representing rules that determine how the resulting value of a given phrase should be calculated.

In principle, each pattern has a set of shifters associated with it. Joining all these into one large set does not seem to be a good idea. For instance, we can *esbanjar* (squander) something that is positive (e.g., money, opportunity) but not negative (e.g., *crise* (crisis)). Thus, this knowledge can be incorporated into the design of shifter patterns (see Section 3.2). Another advantage is that our system can deal with iterative application of shifter patterns (e.g., *não funciona mal* (does not work badly) including two different shifters *não* and *mal* (see Section 3.3.1).

2.3. Deep-Learning-Based Approaches

Classical machine learning and lexicon-based approaches use manually crafted features, such as the n-gram or lexicon for sentiment-classification tasks [16]. In contrast, deep-learning approaches use new architectures of neural networks to learn features automatically and provide state-of-the-art performance [2,17,18]. However, they offer no explanation on how the prediction is achieved and are often referred to as black-box methods. Furthermore, these models have a difficulty in incorporating prior human knowledge into the process.

Deep-learning architectures, such as, *Recurrent Neural Networks (RNNs)*, *Long short-term memory (LSTM)*, *Convolutional Neural Networks (CNNs)* and, recently, *transformer-based models* have been widely used for sentiment analysis [2]. Sharfuddin et al. [19] used RNN-based model for binary sentiment classification in the Bengali language and achieved 85.67% accuracy. Similarly, Mudammad et al. [20] used word2vec and LSTM to classify hotel reviews in Indonesia and achieved a performance of 85.9% accuracy. Although originally developed for computer vision tasks, such as image classification [21], CNN can be used for natural language processing tasks. Ouyang et al. [22] used *word2vec* with CNN and provided competitive performance for sentiment classification.

State-of-the-art deep-learning approaches use transformer architecture and exploit the idea of self-attention and positional encoding. Transformers have the ability to deal with longer sequences of words, enabling them to learn from a larger context [23]. Transformer-based pre-trained models, such as *Bert* [5] and *Roberta* [6], are used in many NLP tasks, such as text classification and machine translation.

Multilingual variants of these models, including *mBert* [5], enable the user to adapt the system to other languages that have not been included during training [24]. Muhammad et al. [25] used transformer-based pre-trained multilingual models on a benchmark dataset for sentiment classification in four low-resourced Nigerian languages (Hausa, Yoruba, Igbo and Pidgin).

As we stated above, deep-learning approaches have the disadvantage that they do not provide an explanation regarding how the prediction was achieved. This has motivated us to revisit some of the 'older' approaches that exploit sentiment lexicons and shifter patterns, as these can provide understandable explanations to users. As our approach with shifter patterns captures some contextual information, we hypothesize that they might still provide reasonable performance. The objective of this work is to determine whether the loss of performance is acceptable.

2.4. Comparative Studies

Various researchers have conducted comparative studies among different approaches. Atteweldt et al. [26] compared manual annotation, crowd-coding, dictionary approaches and machine learning (ML). This work is relevant to the research discussed here, as it focuses on a similar domain—economics (more precisely economic headlines) in a language with limited language resources (Dutch), which, in this respect, can be compared to Portuguese. The conclusion of the authors is that off-the-shelf dictionaries do not perform well on new tasks when compared to machine learning (ML) methods, particularly deep-learning methods, such as Convolutional Neural Networks (CNN).

Although deep learning techniques often provide well-performing solutions, it is not easy to see why a certain prediction was made and what it is based on. This motivated us to explore automated techniques for lexicon construction. The work of Tavares et al. [27] is oriented towards sentiment classification of economic texts in Portuguese. The data used involve 400 manually annotated sentences in the economics domain extracted from Portuguese newspapers. Although they performed experiments with different ML approaches, the best performance was obtained with manually generated rules.

2.5. Combined Approaches

Various researchers combined different approaches to sentiment analysis into a single hybrid system. Mudinas et al. [28] and Almatteh et al. [10] combined lexicon-based and learning-based approaches and experimental result shows an increase performance. Zou et al. [29] observed that not all words have the same importance. They proposed a lexicon-based supervised attention model (LBSA), which allows a recurrent neural network to focus on the sentiment content and obtain sentiment-informative representations as a result. The area of neuro-symbolic AI [30,31] raised a great deal of interest in the research community recently, as it can exploit the best of neural and symbolic approaches. To the best of our knowledge, the area of sentiment analysis has not yet explored this venue. Section 5 discusses our plans in this respect.

2.6. Related Work on Negation Detection

Many researchers from the natural language processing (NLP) community have studied *negation detection*, as it is relevant in many real-world applications, in virtually all areas. Consider, for instance, the area of health. It is clearly important to know whether the patient was diagnosed to have (or not to have) a certain disease [32]. This work is relevant to the area of sentiment analysis, as negation affects the sentiment values. Many authors working in the area divide the process into two major sub-tasks - *cue identification* and *scope recognition*. The term *cue* refers to certain words (e.g., *not*, *despite*, *inverts*) that we call *shifters* in our article. The term *scope* refers to a sentence fraction affected by the corresponding cue (shifter).

The approaches to negation detection can be divided into several different types. One involves the so-called rule-based approaches, which involve sentiment lexicons, shifters and rules, that are used to derive the final sentiment value of a given sentence or another piece of text. In our article, we use the term *symbolic approaches* to refer to this kind of approach. The second group involves ML-based approaches to identify cues and the corresponding scope. In principle, any ML method could be used; however, the current trend is to exploit deep neural networks for this aim. Various authors used this in the area of negation detection (e.g., [32]) and also in the area of sentiment analysis (e.g., [1]).

Let us analyze the different types of negation considered in the work of Solarte et al. (2022) [32]. One category was referred to as *syntactic negation*, where the cues are represented by certain function words or adverbs. These can be compared to the reversal/inversion shifters that are adverbs discussed further on.

Another category involved the so-called *lexical negation*, which include verbs, adjectives or noun phrases, such as, *denies*, *negative* and *reduction*. These can again be compared with some of the reversal/inversion shifters used in our work.

The advantage of automatic approaches over manual approaches is that they can detect new cues (reversal/inversion shifters) using a trained deep-learning model [1,32]. We are planning to investigate how we could exploit the methods of this kind to suggest new shifters. Regarding scope detection, we note that we do not need any special detection method, as we exploit the existing knowledge encoded by the dependency parser. The scope is simply computed from its output.

3. Exploiting Symbolic and Deep NN Approaches in Sentiment Analysis

This section provides more details about the proposed symbolic approach to sentiment analysis and covers the following topics:

- Automatic generation of sentiment lexicon (Section 3.1).
- Shifter patterns (Section 3.2).
- Incorporating shifter patterns in the SA system (Section 3.3).

The last topic covers the usage of deep NN approach in sentiment analysis (Section 3.4).

3.1. Automatic Generation of Sentiment Lexicon

In this section, we describe the automatic generation of the sentiment lexicon, which is based on the work of [10] but extends it in various aspects. The methodology adopted is shown in Figure 1, and it involves the following steps:

1. Corpus used and its annotation.
2. Preprocessing.
3. Separating long sentences into short labeled fractions.
4. Generation of the sentiment values.

More details regarding each step are given in the respective subsections.

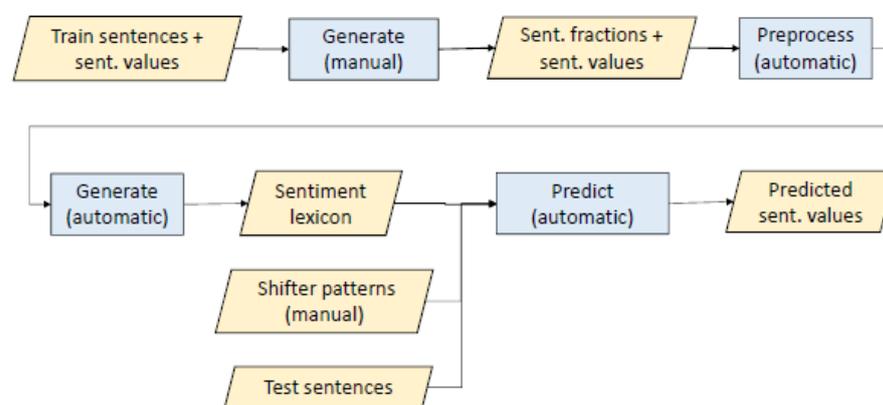


Figure 1. Basic architecture of the SA system.

3.1.1. Corpus Used and Its Annotation

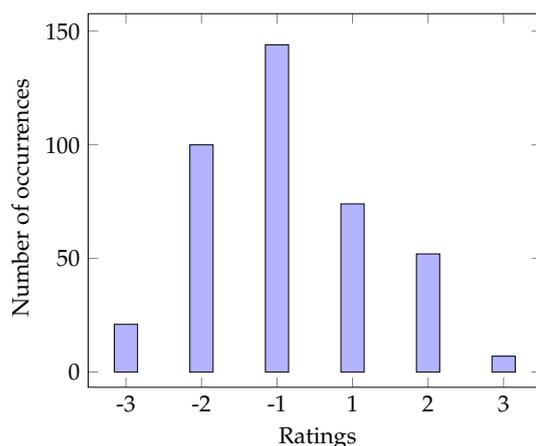
The corpus used in this study involved 23 selected texts from different online Portuguese newspapers that publish articles on finance and economics. Each text contained varying numbers of sentences, ranging from 2 to 81 sentences with a median of 12. The total number of sentences is 398, and each sentence has a different length ranging from 2 to 30 tokens. Table 1 shows examples of some sentences in this corpus.

For the annotation task, we developed annotation guidelines, which consisted of typical examples of sentences (or phrases) with appropriate sentiment values that were agreed on by all linguists. Two linguists (The linguists referred to here are co-authors of this paper) annotated each sentence (or phrase) with a sentiment value on a scale of -3 to 3 . In such instances, where there is disagreement, the other two linguists adjudicated such cases and produced a single sentiment value. Table 1 shows examples of some sentences/phrases and their ratings.

Table 1. Examples of some sentences and ratings.

| Sentence | Translation | Val. |
|---|--|------|
| Portugal não pode estar a governar só para os mercados , ou seja, para tentar demonstrar que o défice está melhor. | Portugal cannot govern only for the markets , that is, try to demonstrate that the deficit is better. | −1 |
| Governo mostra-se mais inseguro e débil, e sem um rumo definido . | Government appears more insecure and weak and without a defined direction . | −2 |
| Quem lida com a exportação de serviços sabe que a falta de qualificação dos portugueses é uma falsa questão. | Anyone who deals with the export of services knows that the lack of qualifications of the Portuguese is a false issue. | 1 |
| O saldo positivo das nossas trocas compensa largamente o financiamento das atividades do país. | The positive balance of our exchanges largely compensates for the financing of the country's activities. | 2 |

The distribution of the ratings, as shown in Figure 2, is skewed towards negative values. This probably reflects the general tendency that was observed in some studies suggesting that people tend to report on various negative aspects more frequently than on the positive ones [33].

**Figure 2.** Distribution of ratings for the data used.

3.1.2. Preprocessing

The data is read-in and processed by a dependency parser (package *udpipe* of R was used) to extract the tokens, the corresponding lemmas, part of speech (POS) tags and other elements discussed further on. Previous studies have shown that not all lexical categories include sentiment bearing words (e.g., [34–36]). Our experiments have indicated that the most useful categories to consider are nouns, verbs, adjectives and adverbs. Therefore, we focused only on these four categories when considering unigrams. However, we also consider a few rather exceptional cases, such as the word *nada*/PRON (nothing). They have been introduced, as they appear in some shifter patterns.

In this work, we have not eliminated *stopwords*, as is often done in many text mining applications. This is because the stopwords in Portuguese usually include words, such as “*não*” (no), “*mais*” (more) and “*muito*” (many), which play an important role in the shifter patterns discussed in Section 3.2.

3.1.3. Separating Long Sentences into Shorter Elements

We observed that long sentences make the task of inducing the sentiment lexicon difficult. For instance, if a sentence with 20 words has been rated −1, a question arises as to which word(s) contributed to this negative rating. Therefore, to mitigate this difficulty,

we use two different strategies. The first reformulates long sentences into shorter sentence fractions.

Each typically includes words that appear to affect the overall sentiment value. Various irrelevant words can be dropped. The second strategy splits the given sentence into several sentence fractions, some of which may be phrases or even words (e.g., *receitas* (income)). Each of these elements is assigned a rating. To illustrate this process in an example, let us consider the first sentence in Table 1. This sentence was divided into the following fractions/phrases/words and each part is attributed a rating:

- *governar só para os mercados* (govern only for the markets)/−1,
- *défi ce* (deficit)/−1,
- *demonstrar que o défi ce está melhor* (demonstrate that the deficit is better)/1.

The total number of new sentence fractions/phrases obtained from the original sentences was 1754 items (This set belongs to what we call *setup 1*. More information on the setups is given in Section 4). On average, each original sentence gave origin to 4.4 new items. The reformulated sentences or sentence fractions/phrases are used as a source for the generation of sentiment lexicon. The approach is an extension of Brazdil et al. (2021) [12] and is described in the next section.

3.1.4. Generation of the Sentiment Values

The approach presented here considers the length of fractions, as these can vary in their length (when considering sentiment-bearing words). Consider the following two fractions, both of which include the word *impostos* (taxes):

- *os impostos baixos* (low taxes)/1
- *impostos* (taxes)/−1

Suppose the aim is to calculate its final sentiment value from these two phrases. The mean of the two values (1,−1) is 0; however, this does not seem right. The second phrase includes only one word (*impostos*) and, hence, should influence more the final sentiment value. This problem can be overcome by attributing different weights to different phrases/fractions, while taking into account the number of sentiment-bearing words in each phrase. The larger the number of sentiment-bearing words, the smaller the weight.

Let us come back to our example. As the first phrase/fraction includes two sentiment-bearing words (*impostos, baixos*) its weight is 1/2. The second phrase is just one word and so its weight is 1. The final sentiment value SV_{ti} of word/term ti is calculated as the weighted mean of all values, as expressed by the following formula:

$$SV_{ti} = \frac{\sum R_{pj,ti} * W_{pj}}{W_0 + \sum W_{pj}} \tag{1}$$

where $R_{pj,ti}$ represents the rating of term ti in phrase/fraction pj and W_{pj} the respective weight of the phrase/fraction pj . Symbol W_0 represents a small constant (e.g., 0.5) that provides a kind of smoothing. When the term ti is relatively rare, the value of $\sum W_{pj}$ will be rather small, and thus W_0 will pull the sentiment value down. Hardly any effect can be observed for terms that are relatively frequent. Let us consider our previous example again. With $W_0 = 0.5$, we get :

$$SV_{imposto} = \frac{(1 * 0.5) + (-1 * 1)}{0.5 + (0.5 + 1)} = \frac{-0.5}{2} = 0.25 \tag{2}$$

As we can see, the weight of the shorter phrase (*impostos*) influence the final outcome, and hence the final sentiment value is negative (−0.25).

3.1.5. Induced Lexicon: Ecolex

The induced sentiment lexicon dubbed Ecolex obtained from the training data in one of the runs included 1246 entries (In this run we used the training data of setup 1 discussed

in Section 4). Table 2 shows some entries from this lexicon, which are ordered by sentiment value starting with the highest value. Some of the words are related to economics and finance (e.g., *resolver* (resolve) and *corrigir* (correct)).

Table 2. Examples of some positive lexicon entries in Ecolex and Sentilex’.

| Word | Ecolex | | Sentilex’ | |
|-----------------------------------|--------|-------|------------------|----|
| | SV | Class | Word/Idiom | SV |
| resolver (resolve) | 0.909 | V | - | - |
| bom (good) | 0.897 | ADJ | bom | 1 |
| felizmente (luckily) | 0.889 | ADV | - | - |
| corrigir (correct) | 0.880 | V | corrigir | 1 |
| qualidade (quality) | 0.857 | N | - | - |
| amenizar (soften) | 0.854 | V | - | - |
| responsabilidade (responsability) | 0.851 | N | responsabilidade | −1 |
| importante (important) | 0.845 | ADJ | - | - |
| ganhar (win) | 0.835 | V | ganhar | 1 |
| capacidade (capacity) | 0.824 | N | - | - |

The columns on the right show the sentiment values for the same words in Sentilex’, which is a version of Sentilex-PT [37,38], an off-the-shelf general purpose lexicon for Portuguese (Lexicon Sentilex’ was obtained from Sentilex-PT by eliminating all entries with neutral class and idiomatic expressions). We note that Sentilex-PT does not include various words (marked by - in our table) that are included in Ecolex and clearly seem useful. Some of these words are specific to economics and finance, which is the domain used here.

Sentilex-PT lexicon includes not only words but also various idiomatic expressions and phrases (e.g., *honrar os compromissos* (honor the commitments)). This seems to be an advantage. However, our solution based on shifter patterns, discussed in Section 3.2, can generate many different phrases, such as *reforçar crescimento*. Sentilex-PT does not include this facet. On the other hand, the addition of some idiomatic expressions that are not captured by our patterns could still be useful.

Table 3 shows the distribution of four different lexical classes (N, V, ADJ, ADV) in Ecolex. It also contains a few other entries that represent exceptional cases that have been introduced, including, for instance, words, such as *nada* (nothing) (class PRON), and words, such as *nem* (neither) (class Other), which form parts of some shifter patterns.

Table 3. Counts and proportions of lexicon entries for different lexical classes.

| Class | Count | % |
|--------------|-------------|------------|
| N (noun) | 595 | 47.8 |
| V (verb) | 329 | 26.5 |
| ADJ | 251 | 20.1 |
| ADV | 64 | 5.1 |
| PRON | 3 | 0.2 |
| Other | 4 | 0.3 |
| Total | 1246 | 100 |

3.1.6. Combining Domain-Specific and General Purpose Lexicons

Due to the fact that Ecolex was induced from a relatively small domain-specific dataset, it has a relatively small number of entries. Many general-purpose entries are not included,

which is a disadvantage. To overcome this limitation, we combine both the domain-specific lexicon (Ecolex) and the general purpose lexicon (Sentilex') [7,8]. The resulting lexicon is referred to as Eco-Senti-lex. If the same word is used in both lexicons, preference is given to the sentiment value in the domain-specific lexicon (i.e., Ecolex).

3.2. Shifter Patterns

Sentiment lexicons that include only unigrams do not capture contextual information and the retrieved sentiment value is fixed in all contexts. This is clearly a shortcoming. To mitigate this, we exploit shifter patterns that capture some of the information provided by the context. As various authors have shown, certain words or phrases can act as *shifters* or *contextual valence shifters* of the sentiment value of words (or expressions) that appear in the same syntactic structure [39].

A shifter is a word or expression that changes the polarity of another word or expression. This change can be made by the operations of *intensification*, *downtoning/attenuation* or *reversal*. In this work, the shifters are represented by the symbol *S*. The words (or expressions) whose sentiment value is affected by the shifter are referred to as focal elements and are represented by the symbol *F*.

Whenever we need to indicate that the focal element has a positive (or negative) sentiment value, we will use F^+ (or F^-). The combination of the shifter and the corresponding focal element is referred to as *shifter pattern*. Our previous work [12] uses a representation of shifter patterns based on proximity. However, this representation has the disadvantage that it is necessary to set the distance value between *S* and *F*. The representation of patterns adopted here avoids this problem, by incorporating the information of *dependency links* provided by the dependency parser.

To explain this, we introduce two concepts—namely, the concept of *head token* and the concept of *dependent token*. The phrase *crecer muito* (grow very much), for instance, includes the word *crecer* (grow), which is the head token. The word *muito* (very much) is the dependent token. This information is provided by the dependency parser, as shown in Table 4 (see the first two lines).

Table 4. Dependencies between tokens provided by the dependency parser.

| sno | token_id | token | upos | head_token_id | dep_rel |
|-----|----------|-------------|------|---------------|---------|
| 1 | 1 | crecer | V | 0 | root |
| 1 | 2 | muito | ADV | 1 | advmod |
| 2 | 1 | crescimento | N | 2 | nsubj |
| 2 | 2 | parece | V | 0 | root |
| 2 | 3 | bom | ADJ | 2 | xcomp |

The information in Table 4 is represented graphically in Figure 3 (see the left figure), where the head token (*crecer*) is drawn in a box positioned above the dependent token (*muito*) and the link from the dependent token to the head token is visible. The graphic representation is easier to analyze than the structures shown earlier in Table 4.

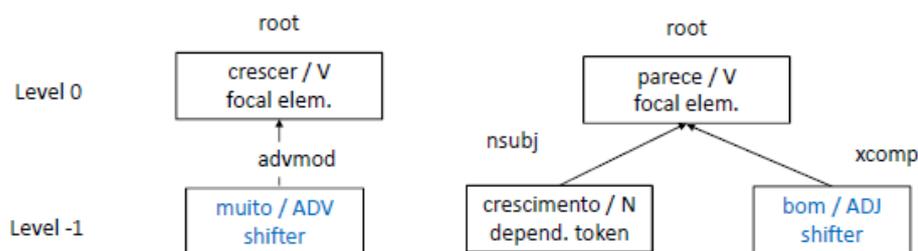


Figure 3. Two examples with dependency links between some tokens.

Apart from these two representations, we also present an algebraic one, which has many advantages of the graphical representation but is more compact. Let us consider our example *crescer muito* (grow very much). The pattern $F_V^+ \leftarrow S^I$ identified as IAV+ in Table 5 captures much of the information shown in Figure 3.

Notice that Table 5 (lines 3 to 5) and Figure 3 (right sub-figure) include the dependency structure of another example—*crescimento parece bom* (growth seems good). This example illustrates that, in general, the graphical representation of the dependency structure has the form of a tree.

Table 5. Some shifter patterns that include intensification.

| Id | Shifter Pat. | Token 1 | Token 2 | S | Use |
|-------|-----------------------------|-----------------------------|-----------------------------------|----|-----|
| IAV+ | $F_V^+ \leftarrow S^I$ | (F = crescer (grow)) | S = muito (very much) | 28 | 2.7 |
| IAV- | $F_V^- \leftarrow S^I$ | (F = prejudicar (harm)) | S = muito (very much) | 10 | 1.5 |
| IAJ+ | $S^I \rightarrow F_{ADJ}^+$ | S = muito (very) | (F = satisfatório (satisfactory)) | 29 | 6.8 |
| IAJ- | $S^I \rightarrow F_{ADJ}^-$ | S = muito (very) | (F = negativo (negative)) | 11 | 4.3 |
| IJN+ | $S^I \rightarrow F_N^+$ | S = bom (good) | (F = crescimento (growth)) | 23 | 9.5 |
| IJN- | $S^I \rightarrow F_N^-$ | S = mau (bad) | (F = crise (crisis)) | 10 | 3.3 |
| IV'N+ | $S^I \leftarrow F_N^+$ | S = reforçar (reinforce) | (F = crescimento (growth)) | 20 | 2.3 |
| IV'N- | $S^I \leftarrow F_N^-$ | S = ampliar (amplify) | (F = crise (crisis)) | 15 | 1.8 |
| IVN+ | $F_N^+ \leftarrow S^I$ | (F = crescimento (growth)) | S = reforçado (reinforced) | 20 | 0.8 |
| IVN- | $F_N^- \leftarrow S^I$ | (F = crise (crisis)) | S = ampliado (amplified) | 15 | 0.3 |
| IN'N+ | $S^I \leftarrow F_N^+$ | S = aumento (increase of) | (F = crescimento (growth)) | 11 | 0.7 |
| IN'N- | $S^I \leftarrow F_N^-$ | S = aumento (increase of) | (F = despesa (expenditure)) | 8 | 0.8 |
| INN+ | $F_N^+ \leftarrow S^I$ | (S = crescimento (growth)) | F = aumentado (increased) | 11 | 0.8 |
| INN- | $F_N^- \leftarrow S^I$ | (S = despesa (expenditure)) | F = aumentado (increased) | 8 | 1.0 |

3.2.1. Different Types of Shifter Patterns

Shifters can perform one of the following operations: intensification, downtoning and reversal, following the terminology of Trnavac et al. [4]. All three types of operations are discussed below.

Intensification

As the name of the operation suggests, intensification will increase the magnitude of the sentiment value of the focal element. That is, the sentiment value (SV) of the shifter pattern that includes an intensifier will normally be larger and more positive than the SV of the focal element if its sentiment value is positive, as in *crescer muito* (grow very much). This can be captured by the following rule:

$$SV(F, S^I) = C_I * SV(F) \tag{3}$$

where S^I represents the given intensifier (e.g., *muito*), and C_I a constant greater than 1 (e.g., 2). If the sentiment value of the focal word is negative, the sentiment value of this pattern will be further down on the negative scale, as in *piorar muito* (deteriorate greatly).

Many different shifter patterns appear in texts. Therefore, our goal was to analyze the given texts and to define a set of patterns considered useful. These patterns were organized into three different tables, depending on the type of shifter involved. Some patterns involving intensifiers are shown in Table 5.

All patterns show the class of the focal element. Thus, for instance, F_V (or F_{ADJ} , etc.) represents a focal element that is a verb (or adjective, etc.). The arrow \leftarrow (or \rightarrow) points in the direction of *head token*. The shifter is shown in column *Token 1* or in *Token 2*. This arrangement is used simply to aid in legibility. Consider again *crescer muito* (grow much).

As *crescer* appears before *muito*, it is shown in the column *Token 1*, while *muito* appears in *Token 2*.

The focal element is shown in parentheses to indicate that the focal element is not really a part of the pattern. All patterns require that the focal element is of some specific class (e.g., F_V requires that it is a verb). It is shown in our tables to aid in legibility. The column identified as $|S|$ shows the size of the list of words that represent shifters for each shifter pattern. For instance, let us consider the shifter pattern no. 1 in Table 5. We see that the list of shifters that has been assembled manually includes 28 elements (the tokens *muito* and *tanto* form part of this list).

The column identified as *Use* shows how many times this pattern was invoked on test set consisting of sentences discussed in Section 4.1. As we use six different setups, the numbers shown represent average values of the test sets across different setups. Thus, for example, the pattern *IAV+* was invoked 2.7 times on average in different setups.

We note that, in general, the intensifier patterns got invoked quite frequently. For instance, intensifier pattern *IJN+* that covers *bom crescimento* (good growth), among other phrases, was invoked 9.5 times on average on a test set with average length of 66.3 sentences.

Attenuation/Downtoning

This operation works opposite to intensification. It decreases the absolute value of the sentiment value of the focal element, as in *crescer pouco* (*grow slightly*). This can be captured by the following rule:

$$SV(F, S^A) = C_A * SV(F) \tag{4}$$

where S^A represents the given attenuator (e.g., *pouco*), and C_A is a constant less than 1 (e.g., 0.5). Table 6 shows some shifter patterns with downtoners/attenuators that we have identified.

Table 6. Some shifter patterns that include attenuation/downtoning.

| Id | Shifter Pat. | Token 1 | Token 2 | S | Use |
|-------|-----------------------------|------------------------------|-----------------------------------|----|-----|
| AAV+ | $F_V^+ \leftarrow S^A$ | (F = crescer (grow)) | S = pouco (not much) | 5 | 0.8 |
| AAV- | $F_V^- \leftarrow S^A$ | (F = prejudicar (harm)) | S = pouco (not much) | 4 | 0.7 |
| AAJ+ | $S^A \rightarrow F_{ADJ}^+$ | S = pouco (not much) | (F = satisfatório (satisfactory)) | 4 | 1.8 |
| AAJ- | $S^A \rightarrow F_{ADJ}^-$ | S = pouco (not much) | (F = negativo (negative)) | 4 | 0.8 |
| AJN+ | $S^A \rightarrow F_N^+$ | S = fraco (weak) | (F = crescimento (growth)) | 12 | 1.2 |
| AJN- | $S^A \rightarrow F_N^-$ | S = fraca (weak) | (F = crise (crisis)) | 7 | 1.3 |
| AV'N+ | $S^A \leftarrow F_N^+$ | S = diminuir (decrease) | (F = crescimento (growth)) | 8 | 0.7 |
| AV'N- | $S^A \leftarrow F_N^-$ | S = controlar (control) | (F = despesa (expenditure)) | 23 | 1.5 |
| AVN+ | $F_N^+ \leftarrow S^A$ | (F = crescimento (growth)) | S = diminuído (decreased) | 8 | 0.7 |
| AVN- | $F_N^- \leftarrow S^A$ | (F = crise (crisis)) | S = controlado (controled) | 23 | 1.5 |
| AN'N+ | $S^A \leftarrow F_N^+$ | S = diminuição (decrease of) | (F = crescimento (growth)) | 5 | 0.0 |
| AN'N- | $S^A \leftarrow F_N^-$ | S = diminuição (decrease of) | (F = despesa (expenditure)) | 5 | 0.2 |
| ANN+ | $F_N^+ \leftarrow S^I$ | (S = crescimento (growth)) | F = diminuído (decreased) | 5 | 0.8 |
| ANN- | $F_N^- \leftarrow S^I$ | (S = despesa (expenditure)) | F = diminuído (decreased) | 5 | 1.0 |

Reversal/Inversion

The result of this operation depends on whether the focal element has a positive or negative sentiment polarity. In the case of F^+ , reversal inverts the polarity of the sentiment value of the focal element, such as, for instance, in *não é bom* (is not good). Most people would agree that the sentiment value of this phrase is rather negative. Thus, this can be captured by the following rule:

$$SV(F^+, S^R) = C_{R+} * SV(F^+) \tag{5}$$

where S^R represents the given inversion shifter (e.g., *não*, and C_{R+} a constant that is negative (e.g., -0.8). The use of inversion operation applied to a focal element with negative sentiment value is more complicated. Consider, for instance, the expression *não é mau* (is not bad). Most people would not consider this to be equivalent to *bom* (good) but rather slightly positive and near the neutral value. This can be captured by the following rule:

$$SV(F^-, S^R) = C_{R-} * SV(F^-) \tag{6}$$

where and C_{R-} a negative constant near 0 (e.g., -0.2). Table 7 shows some shifter patterns with reversal/inversion that we identified when analyzing our corpus.

Table 7. Some shifter patterns that include reversal/inversion.

| Id | Shifter Pat. | Token 1 | Token 2 | S | Use |
|-------|-----------------------------|----------------------------|-----------------------------------|----|-----|
| RAV+ | $S^R \rightarrow F_V^+$ | S = não (not) | (F = crescer (grow)) | 12 | 0.7 |
| RAV- | $S^R \rightarrow F_V^-$ | S = não (not) | (F = prejudicar (harm)) | 12 | 6.2 |
| RAJ+ | $S^R \rightarrow F_{ADJ}^+$ | S = não (not) | (F = satisfatório (satisfactory)) | 6 | 1.0 |
| RAJ- | $S^R \rightarrow F_{ADJ}^-$ | S = não (not) | (F = mau (bad)) | 6 | 1.0 |
| RAN+ | $S^R \rightarrow F_N^+$ | S = não (no) | (F = crescimento (growth)) | 7 | 3.7 |
| RAN- | $S^R \rightarrow F_N^-$ | S = não (no) | (F = crise (crisis)) | 7 | 3.3 |
| RV'N+ | $S^R \leftarrow F_N^+$ | S = inverter (reverse) | (F = crescimento (growth)) | 22 | 1.7 |
| RV'N- | $S^R \leftarrow F_N^-$ | S = inverter (reverse) | (F = crise (crisis)) | 13 | 0.7 |
| RVN+ | $F_N^+ \leftarrow S^R$ | (F = crescimento (growth)) | S = invertido (reversed) | 22 | 0.3 |
| RVN- | $F_N^- \leftarrow S^R$ | (F = crise (crisis)) | S = reduzida (reduced) | 13 | 0.0 |
| RN'N+ | $S^R \leftarrow F_N^+$ | S = inversão (reversal) | (F = crescimento (growth)) | 21 | 0.3 |
| RN'N- | $S^R \leftarrow F_N^-$ | S = redução (reduction) | (F = despesa (expenditure)) | 12 | 0.2 |
| RNN+ | $F_N^+ \leftarrow S^I$ | (S = crescimento (growth)) | F = reduzido (reduced) | 5 | 0.8 |
| RNN- | $F_N^- \leftarrow S^I$ | (S = dívida (debt)) | F = reduzido (reduced) | 5 | 1.0 |
| RVV+ | $S^R \leftarrow F_V^-$ | S = faltar (fail to) | (F = crescer (grow)) | 1 | 0.2 |

3.2.2. Methodology Used When Defining Shifter Patterns

As we stated earlier, different shifter patterns were generated manually. This was not done in a haphazard way but rather by following certain methodology discussed below. This work started by examining the concrete sentences/sentence fractions available. We accepted the approach in previous work (e.g., [35]) that separated the shifters into three groups—namely, intensifiers, attenuators and reversal/inversion. Having observed various regularities in each of these groups enabled us to come up with different types of shifter patterns, such as those shown in Table 7. Initially, each group included only a few shifter patterns and the associated sets of shifters.

Suppose that we encountered *esta medida anulou o crescimento* (this measure canceled the growth) among our sentence fractions and found that the relevant pattern “RV'N+” did not include the shifter *anular* (cancel). Thus, this would prompt us to extend the list of shifters associated with this rule to include also this word. However, we do not stop here. The next aim is to exploit this new piece of knowledge in the attempt to further improve existing shifter patterns using various strategies detailed below.

- Examine the pattern of opposite polarity.
The aim is to determine whether the same shifter can be added to the shifter pattern with opposite polarity. If we apply this strategy to our example, the question is whether the shifter *anular* (cancel) can also be associated with “RVN-”, as in *anular a crise* (cancel the crisis). As this seems admissible, we would add this shifter to this shifter pattern.
- Explore the knowledge relative to nominalization of verbs.

There is a close relationship between verbal shifters (e.g., *anular* (cancel)) and other related counterparts, such as, nominalizations (e.g., *anulação* (cancelation)) and adjectives (e.g., *nulo* (null)). It would be useful to have a lexicon that lists these related forms; however, to the best of our knowledge, this does not yet exist, and hence this issue needs to be resolved manually. Thus, if one form (e.g., the verbal shifter) was introduced in one pattern, we need to make sure that other patterns include the other related forms.

- Extending shifters by introducing similar terms. This strategy involves searching for words with meaning similar to those introduced. If we consider our example with *anular* (cancel), we can use a dictionary to suggest other words (shifters) with a similar meaning. The on-line dictionary *Priberam* lists, for instance, *aniquilar* (aniquilate), *destruir* (destroy), *exterminar* (exterminate). In addition, *Infopédia* suggests *inutilizar* (turn unusable) and *invalidar* (invalidate). Thus, the steps discussed above (in Section 3.2.2) can be repeated with these new candidate shifters. This way, the shifter patterns and the associated shifters can grow rather rapidly, even though the process is manual.

In the future, we plan to automate the above processes. However, as Portuguese is a relatively low-resource language, the range of opportunities is more limited than if we were dealing with English.

3.3. Incorporating Shifter Patterns in the SA System

In this section, we discuss the process of using the shifter patterns to obtain sentiment values of a given text (e.g., a sentence). This process requires the sentiment lexicon to be used first to retrieve the values of single words (unigrams) and then we try to apply the appropriate shifter pattern(s) and apply it to the focal element(s) obtaining thus the modified value(s). At a more detailed level, the process involves the following steps:

1. Process all portions of the given text (lines) one by one.
2. For each line of text, process all sentences in it one by one. Suppose, for example, the sentence, for the sake of simplicity, is just *crescer muito*.
3. For each sentence, generate a dependency tree. The dependency tree for our example is shown in Figure 3.
4. For each dependency tree, identify all shifter words that appear in it. Then, it is necessary to determine how the order in which these should be processed. Section 3.3.3 provides more details regarding how this was conducted. Considering our example, the system identifies only one shifter—namely, the word *muito*.
5. For each identified shifter word, determine all definitions of shifter patterns associated with the respective shifter. Process all the pattern definitions one by one. Considering our example, the system would identify the definition $I_{AV+} : F_V^+ \leftarrow S^I$ shown in Table 5.
6. Determine whether the shifter is the *head token* or the *dependent token*. Take this information into account when searching for the focal element. Examine whether the focal element satisfies all the requirements stipulated in the definition of the shifter pattern. Check, for instance, whether the focal element is of the required class. Considering the pattern I_{AV+} , it should be a verb (V). Our previous example with *crescer/V* satisfies this condition.
7. For the focal element found, retrieve its sentiment value from the sentiment lexicon and apply the respective rule to determine the sentiment value of the pattern. Suppose that the sentiment value of the token *crescer* is 0.64. Thus, the sentiment value of the phrase *crescer muito* is $2 \times 0.642 = 1.284$. This value is used as the modified value of this phrase.
8. Generate an explanation showing how the final sentiment value was calculated and output it on demand. Section 3.3.1 discusses the format of the explanations adopted here.

More complex structures require special care for two reasons. First, a given shifter may affect more than one word and so it is necessary to determine how to deal with such situations. Section 3.3.2 describes the method adopted. Second, sentences may include more than one shifter and so it is necessary to determine the order in which these should be processed. This issue is discussed in Section 3.3.3.

3.3.1. Explanations Provided by the System

The format of the explanations adopted here lists all sentiment-bearing words in the given text. Thus, for instance, the explanation for the phrase *crescer muito* is

crescer/*0.642/1.285/D *muito*/S/IAV+

The explanation above shows both *crescer* and *muito* and is accompanied by additional information, including:

- Sentiment values retrieved from the sentiment lexicon (here “*0.642”). The symbol “*” is used to identify the values retrieved from the sentiment lexicon.
- Sentiment values that have been derived from the values in the sentiment lexicon (here “1.285”).
- symbol “S” that accompanies each shifter identified (here it is associated with the token “muito”).
- Identifier of the shifter pattern(s) identified (here “IAV+”).
- Identification with symbol “D” all tokens that are in the scope of the shifter (here the token “crescer” is accompanied by “D”, as it is in the scope of the shifter “muito”).

3.3.2. Dealing with Shifters That Affect Multiple Words in Their Scope

As we mentioned earlier, a given shifter may affect more than one word. For instance, consider the phrase ,

não gera receitas fiscais adequadas
(does not generate adequate fiscal income).

The shifter *não* (not) affects the phrase *gera receitas fiscais adequadas* (generates adequate fiscal income) that is in its scope. The adopted solution exploits the information provided by the dependency parser (see Figure 4, sub-figure on the left).

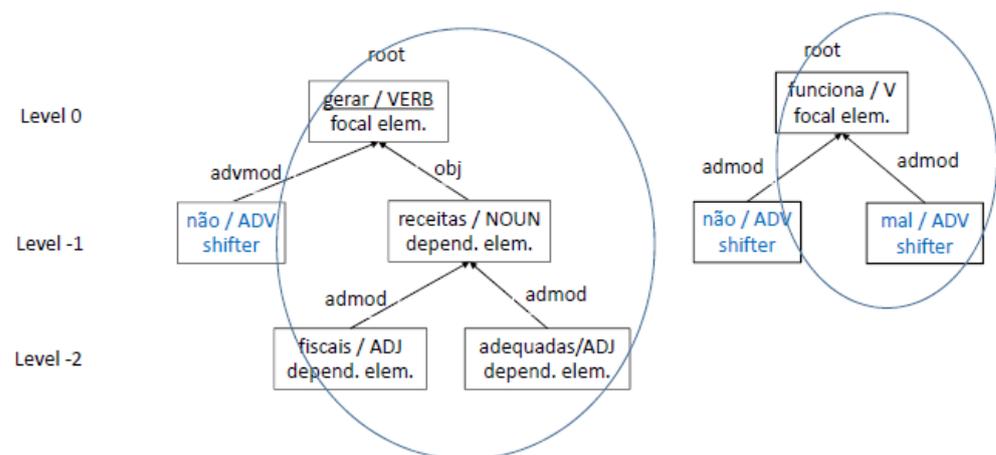


Figure 4. Example of tokens linked by dependencies.

The token *não* points to the head token *gerar* and its dependent substructure shown inside an ellipse in our figure. Thus, the inversion rule is applied to this substructure, not just the token *gerar*. Let us now analyze the explanation generated by the system.

não/*−0.401/−0.627/S/RAV+ *gera*/*0.333/D *receitas*/D *fiscais*/D *adequadas*/*0.4/D

We note that all dependent tokens in this substructure are identified with the symbol “D”. Consequently, the values of “*gera*/*0.333/D” and “*adequadas*/*0.4/D” retrieved from

the sentiment lexicon are used by the inversion rule (RAV+) to obtain the final value of this phrase (-0.627) (The sum of $0.4 + 0.333$ is multiplied by -0.8 giving -0.5864 . In addition, all values of inversion shifters are reduced to 10% of its original value and added to the previous result. Thus, in our case we get $-0.5864 + (0.1 * -0.401) = -0.627$).

3.3.3. Processing Phrases with Multiple Shifters

Sentences may include two (or more) shifters, where one shifter may be in the scope of another one. Whenever this occurs, it is necessary to determine the order in which these shifters should be processed. For instance, the phrase *não funciona mal* (does not work badly) includes two shifters *não* and *mal* (see Figure 4, sub-figure on the right). The system needs to determine what should be processed first.

Intuitive analysis tells us that the shifter *não* should act on the phrase *funciona mal*. Thus, how can we devise a method that works not only for this example but also for other ones? The method adopted conducts the search level by level, starting with the lowest level (here -1). The tokens on each level are processed in the order from right to left. This method identifies the shifter *mal* as the first to be processed, as it is the rightmost token on lowest level (-1). The shifter *não* is then applied to the result of processing *funciona mal*, in accordance with our intuitive analysis. Thus, as a result, our system generates the following explanation:

não/* $-0.401/0.001/S/RAV-$ funciona/* $0.625/D$ mal/* $-0.286/D/RAV+$

As we can see, two shifters have been applied to this text (first RAV+ and then RAV-). The final sentiment value of this phrase is slightly positive (0.001), which seems correct.

Let us consider another phrase with two shifters—*inverte o crescimento de economia* (inverts the growth of economy). In this example, the system SAP predicts the value -1.151 as the final sentiment value. The following explanation generated by the system shows how this value was obtained:

inverter/* $0.4/-1.151/M/RVN+$ crescimento/* $0.598/D/INN+$ economia/* $0.445/D$

First, shifter pattern INN+ is applied. The shifter *crescimento* is applied to the focal word *economia*. This increases the positive value of the focal word retrieved from the sentiment lexicon (0.445) to a larger positive value. Subsequently, the shifter *inverter* (lemma of *inverte*) is applied to the phrase *crescimento economia* and so the value -1.151 is obtained as a result.

3.4. Applying a Deep Learning Approach to Sentiment Analysis

As we stated earlier in Section 2, deep-learning approaches can achieve good performance in many tasks, including sentiment analysis. This fact has motivated us to apply these models to the task discussed in this paper. We use pre-trained multilingual variant of Bert (Bidirectional Encoder Representations from Transformers)—mBERT. This choice was motivated by the fact that mBERT achieved state-of-the-art performance on eleven natural language processing tasks, as Devlin et al. [5] have shown. The mBERT language model was recently used for various sentiment-classification tasks [25].

The language model was pre-trained on large Wikipedias in 100 different languages, including Portuguese. This system pre-trains deep bidirectional representations from unlabeled text using both left and right context. We downloaded the pre-trained system and performed fine-tuning of a single layer. In the problem studied in this paper, fine-tuning was performed using the sentence fractions discussed in Section 3.1.3.

4. Experimental Setup, Evaluation and Results

We conducted two groups of experiments. One involved an experiment with the symbolic approach, which uses a lexicon and shifter patterns. The second involved an experiment with the deep learning-based approach. Both groups of experiments were performed on exactly the same setup as described in the next subsection.

4.1. Experimental Setups

Each group of experiments used the following three types of components:

- Sentence fractions used for training.
- Sentence fractions used for testing.
- Sentences used for testing.

Let us see better what we mean by “training” in the first item above. In the symbolic approach this involves generation of the sentiment lexicon. In the deep learning-based approach this involves the so-called *fine-tuning*. We performed two types of tests. The first type used certain sentence fractions for testing, which were different from those used in training. The second type used sentences for testing.

4.2. Method Used to Construct the Train and Test Sentence Fractions

The division of the given fractions into train and test portions is similar to the scheme of N-fold (N = 6) cross-validation. The existing data is divided into N partitions (folds) and the data of one fold is used as test data, while the data of the other folds are used as *training data*. As there are N different ways we can select a fold for testing, we obtain N different pairs of training and test datasets and, hence, also N separate results.

The method used here is similar but not the same. The main difference is that we do not start with a single pool of data but rather with six different sets of sentences and the corresponding sentence fractions. Consequently, we cannot guarantee that the distribution of positive and negative cases is the same in each set. For this reason, we prefer to avoid the terms *fold j* and prefer to use the term *setup j*. As there are N (here, N = 6) different ways we can select data for testing, we obtain N different pairs of training and test datasets (and, hence, also N separate results).

Table 8 provides more details about the six setups. The table shows, for each setup, the number of items included and the proportion of items with the negative label. For instance, setup 1 includes 1422 sentence fractions in the training set. Of these, 51.0% are negative. We do not show the proportion of positive items, as these can be easily calculated, since both groups total 100%.

Table 8. Overview of train and test datasets in six setups.

| Setup | 1 | 2 | 3 | 4 | 5 | 6 | Mean |
|------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|---------------|
| Train set (fractions) | 1422 | 1452 | 1577 | 1570 | 1475 | 1246 | 1457.0 |
| % of negative | 51.0 | 49.9 | 48.3 | 48.1 | 48.7 | 53.4 | 49.9 |
| Test set (fractions) | 332 | 291 | 183 | 188 | 269 | 497 | 292.3 |
| % of negative | 44.6 | 49.1 | 62.3 | 63.8 | 55.8 | 40.4 | 52.7 |
| Test set (sentences) | 59 | 68 | 66 | 70 | 68 | 67 | 66.3 |
| % of negative | 66.1 | 75.0 | 69.7 | 67.1 | 85.3 | 35.8 | 66.5 |

4.3. Evaluation Measures

In general, different types of measures can be used to evaluate the quality of predictions. The choice depends on whether the system is used to generate numerical or categorical predictions. Here, we focus on categorical predictions in terms of polarity. The measures commonly used for this task are *accuracy*, *precision*, *recall*, *F1*. As *F1* is calculated for each class separately, the measures *microF1*, *macroF1* and *wF1* (weighted *F1*) combine the individual *F1* measures [40].

In the experiments, we used both *accuracy* and *wF1*. As the two sets of results were rather consistent, we report the results only in terms of *wF1*, which is calculated as follows:

$$wF1 = prop.pos * F1.pos + prop.neg * F1.neg \quad (7)$$

where $F1.pos$ ($F1.neg$) represents the $F1$ measure for the positive (negative) class and $pop.pos$ ($prop.neg$) the proportion of positive (negative) examples in the test set.

4.4. Experiments with the Symbolic Approach

In this study, we compare the results of our sentiment analysis systems that use the following lexicons:

- Ecolex: The induced sentiment lexicon (see Sections 3.1.4 and 3.1.5).
- Sentilex' (see Section 3.1.6).
- Eco-Senti-lex: Combination of Ecolex and Sentilex' (see Section 3.1.6).

The lexicon sizes used in different setups are shown in Table 9.

Table 9. Overview of lexicons used.

| Setup | 1 | 2 | 3 | 4 | 5 | 6 | Mean |
|---------------|------|------|------|------|------|------|--------|
| Ecolex | 1246 | 1230 | 1280 | 1220 | 1227 | 1162 | 1227.5 |
| Sentilex' | 5496 | 5496 | 5496 | 5496 | 5496 | 5496 | 5496 |
| Eco-Senti-lex | 6561 | 6561 | 6583 | 6541 | 6551 | 6480 | 6546.2 |

4.4.1. Experiments Involving Sentence Fractions

The first set of experiments involved sentence fractions as test data. We used two approaches to sentiment analysis in conjunction with sentiment lexicons in Table 9. In the first, referred to as *SA*, we used only sentiment lexicons. In the second, referred to as *SAP*, we used both lexicon and shifter patterns. The results are shown in Table 10. Columns 1 to 6 show the performance for different setups. The mean value across different setups is shown in the last column.

Table 10. Weighted F1-score results on sentence fractions.

| Setup | 1 | 2 | 3 | 4 | 5 | 6 | Mean |
|-------------------|-------------|-------------|------|-------------|-------------|-------------|-------------|
| SA-Ecolex | 68.6 | 68.8 | 67.4 | 70.5 | 62.9 | 66.9 | 67.5 |
| SAP-Ecolex | 66.7 | 69.9 | 68.9 | 69.5 | 65.0 | 64.6 | 67.6 |
| SA-Sentilex' | 51.1 | 47.0 | 66.4 | 60.5 | 57.6 | 46.3 | 55.0 |
| SAP-Sentilex' | 46.6 | 47.8 | 68.1 | 60.5 | 57.6 | 43.9 | 54.1 |
| SA-Eco-Senti-lex | 71.8 | 68.5 | 68.6 | 69.1 | 65.9 | 68.6 | 68.8 |
| SAP-Eco-Senti-lex | 71.3 | 70.8 | 68.6 | 68.6 | 68.5 | 67.0 | 69.5 |

The best result is obtained with SAP-Eco-Senti-lex with a mean weighted F1 of 69.5% across different setups. This value is somewhat better (0.7% higher on average) than the performance of a similar system SA-Eco-Senti-lex that uses only the sentiment lexicons, suggesting that it is useful to use shifter patterns. The SAP-Eco-Senti-lex result is also better (1.9% higher on average) than the performance of a similar system SA-Ecolex, which uses only Ecolex. This suggests that it is indeed useful to use the combination of the domain-specific lexicon with the general-purpose one (Sentilex).

4.4.2. Experiments Involving Sentences

The second group of tests involved sentences as test data. The results are shown in Table 11.

Table 11. Weighted F1-score results on long sentences.

| Setup | 1 | 2 | 3 | 4 | 5 | 6 | Mean |
|-------------------|-------------|------|-------------|-------------|-------------|-------------|-------------|
| SA-Ecolex | 54.7 | 65.6 | 62.1 | 60.9 | 61.7 | 63.4 | 61.4 |
| SAP-Ecolex | 68.5 | 65.3 | 69.0 | 60.0 | 72.4 | 61.7 | 66.2 |
| SA-Sentilex' | 71.0 | 70.3 | 62.8 | 63.0 | 70.9 | 61.6 | 66.2 |
| SAP-Sentilex' | 69.8 | 72.8 | 66.1 | 55.5 | 32.1 | 63.2 | 59.9 |
| SA-Eco-Senti-lex | 56.2 | 68.0 | 65.1 | 63.7 | 59.4 | 62.0 | 62.4 |
| SAP-Eco-Senti-lex | 61.7 | 60.2 | 76.5 | 68.1 | 67.3 | 64.7 | 66.4 |

The best result is obtained again with SAP-Eco-Senti-lex with a mean weighted F1 of 66.4% in different setups. This value is substantially better (4.0% higher on average) than the performance of a similar SA-Eco-Senti-lex system that uses only the sentiment lexicons. Thus, the positive effect of using shifter patterns is also confirmed in sentences. This result is comparable to that obtained with SAP-Ecolex. The use of the combined sentiment lexicon did not provide an advantage in this setting.

4.5. Deep Learning Approach Experiment

For the deep learning experiment, we follow the approach described in Section 3.4. First, we reuse a pre-trained Bert model with the HuggingFace transformer tool [23]. This model is then fine-tuned with the training sentence fractions. We used a batch size of 32, a maximum sequence length of 128, the number of epochs 20 and the learning rate 5×10^{-5} .

As the deep-learning approach has higher variability, we experimented with five different runs. Tables 12 and 13 show the results of the tests performed on sentence fractions. As expected, the deep-learning approach achieves better performance compared to the symbolic approach. The mean performance in terms of the weighted F1 of the deep-learning approach is 5.4% higher than the corresponding performance of the best variant of the symbolic approach (SAP-Eco-Senti-lex).

Table 12. Weighted F1-score results on short sentences for six different dataset setup and five runs.

| Run Setup | 1 | 2 | 3 | 4 | 5 | 6 | Mean |
|-------------|------|------|------|------|------|------|-------------|
| 1 | 82.6 | 78.4 | 74.1 | 73.9 | 78.1 | 73.9 | 76.8 |
| 2 | 77.8 | 77.5 | 76.2 | 75.7 | 79.9 | 75.7 | 77.1 |
| 3 | 81.9 | 69.6 | 66.5 | 74.0 | 78.1 | 74.0 | 74.0 |
| 4 | 81.6 | 74.9 | 71.7 | 74.8 | 75.8 | 76.3 | 75.9 |
| 5 | 80.2 | 81.9 | 74.2 | 71.2 | 76.3 | 71.2 | 75.3 |
| Mean | 80.4 | 76.5 | 72.5 | 73.9 | 77.6 | 74.2 | 75.9 |

Tables 12 and 13 show the results of the tests performed on sentences. In this setting, the deep-learning approach achieved a mean performance of 66.8%. The performance of the symbolic approach was 66.4%, that is only 0.4% lower. Presumably, the usage of shifter patterns captures a great deal of contextual information and, hence, could compete with the deep-learning approach. This shows that it is worth pursuing this further, as the symbolic approach can, in addition, provide useful explanations.

4.6. Limitations of the Symbolic and Deep Learning Approaches

The task of constructing a domain-specific sentiment analysis system is not easy for various reasons. We analyze the errors of the symbolic and deep-learning approaches and their causes. This is useful, as it points to directions that could be followed to come up with an improved system.

Table 13. Weighted F1-score results on long sentences for six different dataset setup and five runs.

| Run Setup | 1 | 2 | 3 | 4 | 5 | 6 | Mean |
|-------------|------|------|------|------|------|------|-------------|
| 1 | 78.4 | 77.8 | 74.0 | 52.8 | 79.3 | 52.8 | 69.2 |
| 2 | 77.5 | 76.7 | 76.2 | 40.9 | 75.2 | 40.9 | 60.3 |
| 3 | 69.6 | 72.0 | 66.5 | 52.8 | 79.7 | 54.8 | 65.9 |
| 4 | 74.9 | 74.7 | 75.3 | 54.6 | 81.6 | 54.6 | 69.3 |
| 5 | 81.9 | 77.9 | 72.1 | 37.8 | 84.7 | 37.8 | 51.2 |
| Mean | 76.5 | 75.8 | 72.8 | 47.8 | 80.1 | 47.8 | 66.8 |

4.6.1. Analysis of Some Errors of the Symbolic Approach

We present three different errors for the symbolic approach that include: missing entries in the sentiment lexicon, incoherent labeling of some sentence fractions and incorrect output provided by the dependency parser.

First, some errors occurred because our sentiment lexicon was missing some words and this affected the functioning of the shifter patterns. For instance, applying an intensifier pattern rule (IAJ+) to the phrase *mais desfavorecido* (more disadvantaged) does not have any effect since *desfavorecido* (disadvantaged) does not appear in our sentiment lexicon and its sentiment is 0. Therefore, when the pattern IAJ+ is invoked, multiplying the sentiment value of 0 by a given constant (e.g., 2) does not alter the sentiment value (i.e., returns 0). This problem could be overcome by reusing, for instance, the semi-automatic approaches discussed in Section 2.1.

Another example is *explosivo crescimento* (explosive growth). The word *explosivo* (explosive) act as an intensifier of the focal word *crescimento* (growth). However, as the shifter pattern IJN+ does not include the word *explosivo* (explosive) in its list of intensifier words. This problem could be corrected manually, by extending the list of words associated with the shifter pattern mentioned above. Another possibility would be to use automatic techniques, such as the one used by Shulder et al. [1] (see Section 2.2) to determine whether this word is similar to others that already exist in the given shifter rule.

Second, some errors occur because some sentence fractions are labeled in a way that is incoherent with the way the given shifter patterns are interpreted. For instance, the sentence *o objetivo é reduzir a dívida* (the aim is to reduce the debt). Some people would consider that this sentence has a positive sentiment value, as reducing a debt seems a “noble aim”. This implies that some future state is considered when attributing the sentiment value to this sentence.

However, others could argue that the sentence implies that the country still has a debt and, hence, conveys a negative sentiment value. This implies that the current state is considered when attributing the sentiment value. Therefore, a problem arises if one interpretation is used when labeling the sentence fractions and another in the application of shifter patterns. This decreases the performance of the system. There are numerous examples similar to the one discussed above, some of which include modal verbs, others are paraphrases conveying more or less the same meaning:

1. *devemos reduzir a dívida* (we should reduce the debt) (+),
2. *podemos reduzir a dívida* (we can reduce the debt) (+),
3. *é imprescindível reduzir a dívida* (it is necessary to reduce the debt) (+) and
4. *quem nos dera a dívida menor* (who could arrange us a smaller debt) (−).

In all these examples, the dilemma arises regarding what should be the sentiment value. The first three phrases seem to have a positive sentiment value (+), as there seems to be a solution to the problem with the debt. This does not seem to be the case in the last example and so it is annotated with (−).

Another problem that we encountered involves a decision whether a certain shifter should be included in attenuation/downtoning pattern or in the reversal/inversion pattern.

For instance, the phrase *reduzir a crise* (reduce the crisis). It is possible to argue that the term *reduzir* (reduce) just attenuates the negative sentiment value associated with the term *crise* (crisis), implying that the crises still persists, albeit in a reduced form.

However, one could also argue that the phrase indicates that the problem with the crisis was in fact resolved. If this interpretation is used, the shifter *crise* (crisis) forms part of a reversal/inversion pattern. The fact that some errors occurred with some sentence fractions, such as the ones shown, indicates that care is needed when designing the system. It is important that the labeling is performed in a coherent way with the method used to derive the sentiment value from the given shifter patterns.

Lastly, some errors are due to incorrect output provided by the dependency parser. We believe that the decision of using a dependency parser in this work was a correct one, as it builds on existing knowledge and simplifies the task of developing the sentiment analysis system. However, this has also a downside, as the output of the dependency parser is sometimes not perfect. We observed that some of its errors affect the functioning of our system.

For instance, the phrase *confunde X com Y* (confuses X for Y), the verb *confunde* was incorrectly lemmatized into *confunder*. For this reason, if we had *confundir* in the sentiment lexicon (e.g., Sentilex'), it would not have been recognized. However, it could happen that if a similar phrase to *confunde X com Y* appeared in the training data, a wrong entry would have been introduced also in the induced lexicon. Consequently, if the test data included the same word (albeit with wrong spelling), the system would recognize this word and the error would not be revealed.

4.6.2. Analysis of Some Errors of the Deep-Learning Approach

The deep-learning model commits different types of errors. Some errors occur when the deep-learning model generates predictions of sentiment values for individual words. Some examples of this type of errors are shown in Table 14 (cases 1 to 4). This kind of errors can affect predictions of longer phrases that include these words. For instance, as the prediction for case 4 *rentável* (profitable) is wrong, so is the prediction for *tornar rentável* (make profitable).

The second type of errors occur with phrases that include shifters. These are the cases 5 and 6 in Table 14. Although we would attribute a negative sentiment value to the term *problema* (problem), the shifter *resolver* (to resolve) inverts this. This is apparently not recognized by the deep-learning model in this case. Case 6, which includes the shifter *não* (not) is similar.

Table 14. Examples of some errors of the deep-learning model.

| N° | Term/Phrase | True | Pred. |
|----|---|------|-------|
| 1 | custo (cost) | – | + |
| 2 | desajustada (maladjusted) | – | + |
| 3 | limpar (clean) | + | – |
| 4 | rentável (profitable) | + | – |
| 5 | resolver problema (resolve the problem) | + | – |
| 6 | não é um problema económico (it is not an economic problem) | + | – |
| 7 | têm de ser resolvidos (have to be resolved) | + | – |

The problem of incoherent labeling of some sentence fractions that was discussed in Section 4.6.1 also affects the deep-learning approach. Case 7 in Table 14 illustrates this. As we stated before, it is important that the same stance is taken (i.e., considering the situation in the future) when labeling phrases with a similar meaning

5. Future Work and Conclusions

5.1. Future Work

The presented approach could be improved in various ways. Given that the training data was small (data consisting of labeled sentence fractions), one cannot expect the predictive performance to be very high. It would be interesting to see how this performance improves with more training data. More training data would help the symbolic approach to construct a more comprehensive sentiment lexicon and thus lead to improved performance. As for the deep-learning approach, a question arises of whether the dataset used is too small for the model to manifest its strength. In addition, it would be useful to carry out an experiment with BERTimbau [41], which was trained on texts in Brazilian Portuguese, which could lead to better results. Future work could provide answers to these issues.

Further improvements can be obtained by enhancing the set of shifter patterns and, in particular, the lists of words acting as shifters that are associated with each pattern. Some of these patterns are general and not specific to the domain of economics and finance. Therefore, it would be interesting to verify how many of the existing patterns could be reused in other domains (e.g., health). Another issue is how easy (or difficult) it is to translate and adapt existing patterns to another language. As we stated in Section 2.5, the area of neuro-symbolic AI [30,31] raised a great deal of interest in the research community. We plan to explore this line in our future work.

5.2. Conclusions

We presented two approaches to the construction of a sentiment analysis system in a specific domain (economics and finance in European Portuguese). The first was a symbolic approach, which exploited a sentiment lexicon and shifter patterns. The second exploited a deep-learning model. This included a pre-trained Bert model, which was subsequently fine-tuned on the domain-specific training data. The performance of both approaches was compared on the given common data.

In one group of experiments, the tests were performed on relatively short sentence fractions. As expected, the deep-learning model achieved better performance (wF1 = 75.9%) compared with the symbolic approach (wF1 = 69.5%). In the second group of experiments, the tests were performed on sentences. We demonstrated that the performance of the symbolic system (wF1 = 66.4%) was not much lower than the performance of the deep-learning model (wF1 = 66.8%). We believe that this is due to the several proposed innovative features as highlighted below.

Proposed automatic method to induce the sentiment lexicon. The lexicon construction is performed automatically, based on the distribution of ratings for different words that occur in the training data. The proposed method is an improved variant of the previous version [12] and considers the length of sentence fractions used in training, which affects the quality of the lexicon entries.

Proposed enhanced set of shifter patterns. The shifter patterns proposed here represent an improved version compared to the previous variant [12] and capture contextual information. We believe that this is the main reason why the symbolic approach competes with the deep-learning approach in sentiment prediction. The shifter also captures linguistic knowledge that is not only interpretable to us but can also be used in sentiment analysis systems. We release the shifter patterns to motivate future research in this area at <https://github.com/shmuhammad2004/shifter-patterns>, accessed on 10 July 2022.

An improved method to apply shifter patterns. The presented method exploits the information provided by the dependency parser and, in this way, builds on existing knowledge. The method presented in the earlier variant [12] was improved so that shifters (e.g., inversion) would apply to entire phrases (or sentences) in its scope (i.e., not just to the first word in the phrase (sentence)).

Author Contributions: Conceptualization, P.B. and S.H.M.; methodology, P.B., S.H.M. and J.C.; software, P.B. and S.H.M. and J.C.; validation, P.B.; formal analysis, P.B. and S.H.M.; investigation, P.B., S.H.M., J.C., F.O., F.S., P.S. and A.L.; resources, P.B. and S.H.M.; data curation, F.O., F.S., P.S. and A.L.; writing—original draft preparation, P.B., S.H.M. and J.C.; writing—review and editing, P.B., S.H.M., J.C., F.O., F.S., P.S. and A.L.; visualization, P.B., S.H.M. and J.C.; supervision, P.B.; project administration, P.B. and S.H.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work is financed by National Funds through the Portuguese funding agency, FCT—Fundação para a Ciência e a Tecnologia, within project LA/P/0063/2020 of INESC TEC and project UIDB/00022/2020 of CLUP.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We wish to thank Alípio Jorge from LIAAD INESC TEC for their support in relation to this work. We also thank the anonymous reviewers for their useful comments that allowed us to improve the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Schulder, M.; Wiegand, M.; Ruppenhofer, J. Automatic generation of lexica for sentiment polarity shifters. *Nat. Lang. Eng.* **2021**, *27*, 153–179. [[CrossRef](#)]
- Yadav, A.; Vishwakarma, D.K. Sentiment analysis using deep learning architectures: A review. *Artif. Intell. Rev.* **2020**, *53*, 4335–4385. [[CrossRef](#)]
- Schulder, M.; Wiegand, M.; Ruppenhofer, J.; Roth, B. Towards bootstrapping a polarity shifter lexicon using linguistic features. In Proceedings of the International Joint Conference on Natural Language Processing (IJCNLP), Taipei, Taiwan, 27 November–1 December 2017; Asian Federation of Natural Language Processing: Taipei, Taiwan, 2017; pp. 624–633.
- Trnavac, R.; Das, D.; Taboada, M. Discourse relations and evaluation. *Corpora* **2016**, *11*, 169–190. [[CrossRef](#)]
- Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
- Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; Stoyanov, V. Roberta: A robustly optimized bert pretraining approach. *arXiv* **2019**, arXiv:1907.11692.
- Forte, A.C.; Brazdil, P.B. Determining the level of clients' dissatisfaction from their commentaries. In Proceedings of the International Conference on Computational Processing of the Portuguese Language, Tomar, Portugal, 13–15 July 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 74–85.
- Silva, F.; Silvano, P.; Leal, A.; Oliveira, F.; Brazdil, P.; Cordeiro, J.; Oliveira, D. Análise de sentimento em artigos de opinião. *Linguíst. Rev. Estud. Linguíst. Univ. Porto* **2018**, *13*, 79–114.
- Moreno-Ortiz, A.; Fernández-Cruz, J.; Hernández, C.P.C. Design and evaluation of SentiEcon: A fine-grained economic/financial sentiment lexicon from a corpus of business news. In Proceedings of the 12th Language Resources and Evaluation Conference, Marseille, France, 11–16 May 2020; pp. 5065–5072.
- Almatarneh, S.; Gamallo, P. Automatic construction of domain-specific sentiment lexicons for polarity classification. In Proceedings of the International Conference on Practical Applications of Agents and Multi-Agent Systems, Porto, Portugal, 21–23 June 2017; Springer: Berlin/Heidelberg, Germany, 2017; pp. 175–182.
- Muhammad, S.H.; Brazdil, P.; Jorge, A. Incremental Approach for Automatic Generation of Domain-Specific Sentiment Lexicon. In Proceedings of the Advances in Information Retrieval, LNCS, Lisbon, Portugal, 14–17 April 2020; Springer: Berlin/Heidelberg, Germany, 2020; Volume 12036, pp. 619–623.
- Brazdil, P.; Silvano, P.; Silva, F.; Muhammad, S.; Oliveira, F.; Cordeiro, J.; Leal, A. Extending General Sentiment Lexicon to Specific Domains in (Semi-)Automatic Manner. In Proceedings of the SALLD-1: Proceedings of the Workshop on Sentiment Analysis & Linguistic Linked Data, Zaragoza, Spain, 1 September 2021; Volume 3064.
- Wang, Y.; Zhang, Y.; Liu, B. Sentiment lexicon expansion based on neural PU learning, double dictionary lookup, and polarity association. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 7–11 September 2017.
- Muhammad, A.; Wiratunga, N.; Lothian, R.; Glassey, R. Domain-Based Lexicon Enhancement for Sentiment Analysis. In Proceedings of the BCS SGAI Workshop on Social Media Analysis, Cambridge, UK, 10 December 2013; Citeseer: Gaithersburg, MD, USA, 2013; pp. 7–18.

15. Hamilton, W.L.; Clark, K.; Leskovec, J.; Jurafsky, D. Inducing domain-specific sentiment lexicons from unlabeled corpora. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Austin, TX, USA, 1–5 November 2016; pp. 595–605.
16. Mukhtar, N.; Khan, M.A.; Chiragh, N. Lexicon-based approach outperforms Supervised Machine Learning approach for Urdu Sentiment Analysis in multiple domains. *Telemat. Inform.* **2018**, *35*, 2173–2183. [[CrossRef](#)]
17. Minaee, S.; Kalchbrenner, N.; Cambria, E.; Nikzad, N.; Chenaghlu, M.; Gao, J. Deep Learning Based Text Classification: A Comprehensive Review. *arXiv* **2021**, arXiv:2004.03705.
18. Dogra, V.; Verma, S.; Kavita; Chatterjee, P.; Shafi, J.; Choi, J.; Ijaz, M. A Complete Process of Text Classification System Using State-of-the-Art NLP Models. *Comput. Intell. Neurosci.* **2022**, 2022. [[CrossRef](#)] [[PubMed](#)]
19. Sharfuddin, A.A.; Tihami, M.N.; Islam, M.S. A deep recurrent neural network with BiLSTM model for sentiment classification. In Proceedings of the 2018 International Conference on Bangla Speech and Language Processing (ICBSLP), Sylhet, Bangladesh, 21–22 September 2018; pp. 1–4.
20. Muhammad, P.F.; Kusumaningrum, R.; Wibowo, A. Sentiment analysis using Word2vec and long short-term memory (LSTM) for Indonesian hotel reviews. *Procedia Comput. Sci.* **2021**, *179*, 728–735. [[CrossRef](#)]
21. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
22. Ouyang, X.; Zhou, P.; Li, C.H.; Liu, L. Sentiment analysis using convolutional neural network. In Proceedings of the 2015 IEEE International Conference on Computer and Information Technology, Ubiquitous Computing and Communications, Dependable, Autonomic and Secure Computing, Pervasive Intelligence and Computing, Liverpool, UK, 26–28 October 2015; pp. 2359–2364.
23. Wolf, T.; Debut, L.; Sanh, V.; Chaumond, J.; Delangue, C.; Moi, A.; Cistac, P.; Rault, T.; Louf, R.; Funtowicz, M.; et al. Transformers: State-of-the-art natural language processing. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Online, 16–20 November 2020; pp. 38–45.
24. Grote, H.; Schmidt, F. MAD-X-an upgrade from MAD8. In Proceedings of the 2003 Particle Accelerator Conference, Portland, OR, USA, 12–16 May 2003; Volume 5, pp. 3497–3499.
25. Muhammad, S.H.; Adelani, D.I.; Ahmad, I.S.; Abdulmumin, I.; Bello, B.S.; Choudhury, M.; Emezue, C.C.; Aremu, A.; Abdul, S.; Brazdil, P. NaijaSenti: A Nigerian Twitter Sentiment Corpus for Multilingual Sentiment Analysis. *arXiv* **2022**, arXiv:2201.08277.
26. van Atteveldt, W.; van der Velden, M.A.; Boukes, M. The Validity of Sentiment Analysis: Comparing Manual Annotation, Crowd-Coding, Dictionary Approaches, and Machine Learning Algorithms. *Commun. Methods Meas.* **2021**, *15*, 121–140. [[CrossRef](#)]
27. Tavares, C.; Ribeiro, R.; Batista, F. Sentiment Analysis of Portuguese Economic News. In Proceedings of the tenth Symposium on Languages, Applications and Technologies (SLATE 2021), Vila do Conde/Póvoa de Varzim, Portugal, 1–2 July 2021; Schloss Dagstuhl-Leibniz-Zentrum für Informatik: Wadern, Germany, 2021; Volume 94; pp. 17:1–17:13.
28. Mudinas, A.; Zhang, D.; Levene, M. Combining lexicon and learning based approaches for concept-level sentiment analysis. In Proceedings of the first International Workshop on Issues of Sentiment Discovery and Opinion Mining, Beijing, China, 12–16 August 2012; pp. 1–8.
29. Zou, Y.; Gui, T.; Zhang, Q.; Huang, X.J. A lexicon-based supervised attention model for neural sentiment analysis. In Proceedings of the 27th International Conference on Computational Linguistics, Santa Fe, NM, USA, 20–26 August 2018; pp. 868–877.
30. Hitzler, P.; Sarker, M.K. *Neuro-Symbolic Artificial Intelligence: State of the Art*; IOS Press: Dordrecht, The Netherlands, 2022.
31. Besold, T.R.; d’Avila Garcez, A.; Bader, S.; Bowman, H.; Domingos, P.; Hitzler, P.; Kühnberger, K.U.; Lamb, L.C.; Lima, P.M.V.; de Penning, L.; et al. Neural-Symbolic Learning and Reasoning: A Survey and Interpretation. In *Neuro-symbolic Artificial intelligence: State of the Art*; Hitzler, P., Sarker, M.K., Eds.; IOS Press: Dordrecht, The Netherlands, 2022; pp. 1–51.
32. Solarte, O.; Montenegro, O.; Torrente, M.; González, A.R.; Provencio, M.; Menasalvas, E. Negation and uncertainty detection in clinical texts written in Spanish: A deep learning-based approach. *PeerJ Comput. Sci.* **2022**, *8*, e913.
33. Schöne, J.P.; Parkinson, B.; Goldenberg, A. Negativity spreads more than positivity on Twitter after both positive and negative political situations. *Affect. Sci.* **2021**, *2*, 379–390. [[PubMed](#)]
34. Martin, J.R.; White, P.R. *The Language of Evaluation: Appraisal in English*; Palgrave Macmillan: London, UK, 2005.
35. Taboada, M.; Brooke, J.; Tofiloski, M.; Voll, K.; Stede, M. Lexicon-based methods for sentiment analysis. *Comput. Linguist.* **2011**, *37*, 267–307. [[CrossRef](#)]
36. Liu, B. Sentiment analysis and opinion mining. *Synth. Lect. Hum. Lang. Technol.* **2012**, *5*, 1–167.
37. Silva, M.J.; Carvalho, P.; Sarmiento, L. Building a sentiment lexicon for social judgement mining. In Proceedings of the International Conference on Computational Processing of the Portuguese Language (PROPOR), Coimbra, Portugal, 17–20 April 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 218–228.
38. Carvalho, P.; Silva, M.J. SentiLex-PT: Principais características e potencialidades. *Oslo Stud. Lang.* **2015**, *7*, 425–438. [[CrossRef](#)]
39. Polanyi, L.; Zaenen, A. Contextual valence shifters. In *Computing Attitude and Affect in Text: Theory and Applications*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 1–10.
40. van Rijsbergen, C. *Information Retrieval*, 2nd ed.; Butterworth-Heinemann: Woburn, MA, USA, 1979.
41. Souza, F.; Nogueira, R.; Lotufo, R. BERTimbau: Pretrained BERT Models for Brazilian Portuguese. In Proceedings of the Brazilian Conference on Intelligent Systems, Rio Grande, Brazil, 20–23 October 2010; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2020; Volume 12319. [[CrossRef](#)]