Article

A Hybrid Hierarchical Mathematical Heuristic Solution of Sparse Algebraic Equations in Sentiment Analysis

Maryam Jalali [1], **Morteza Zahedi** [1], **Abdorreza Alavi Gharahbagh** [2], **Vahid Hajihashemi** [2],
José J. M. Machado [3] and **João Manuel R. S. Tavares** [3,*]

1   Faculty of Computer and IT Engineering, Shahrood University of Technology, Shahrood 3619995161, Iran; maryamjalali@shahroodut.ac.ir (M.J.); zahedi@shahroodut.ac.ir (M.Z.)
2   Faculdade de Engenharia, Universidade do Porto (FEUP), Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal; up202003516@edu.fe.up.pt (A.A.G.); up201912327@edu.fe.up.pt (V.H.)
3   Instituto de Ciência e Inovação em Engenharia Mecânica e Engenharia Industrial, Departamento de Engenharia Mecânica, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal; jjmm@fe.up.pt
*   Correspondence: tavares@fe.up.pt; Tel.: +351-22-041-3472

Abstract: Many text mining methods use statistical information as a text- and language-independent approach for sentiment analysis. However, text mining methods based on stochastic patterns and rules require many samples for training. On the other hand, deterministic and non-probabilistic methods are easier and faster to solve than other methods, but they are inefficient when dealing with Natural Language Processing (NLP) data. This research presents a novel hybrid solution based on two mathematical approaches combined with a heuristic approach to solve unbalanced pseudo-linear algebraic equation systems that can be used as a sentiment word scoring system. In its first step, the proposed solution uses two mathematical approaches to find two initial populations for a heuristic method. The heuristic solution solves a pseudo-linear NLP scoring scheme in a polarity detection method and determines the final scores. The proposed solution was validated using three scenarios on the SemEval-2013 competition, the ESWC dataset, and the Taboada dataset. The simulation results revealed that the proposed solution is comparable to the best state-of-the-art methods in polarity detection.

1. Introduction

Due to the expansion of the Internet and social media, sentiment analysis, opinion mining, and polarity detection have become very popular in recent years [1–3]. Sentiment analysis is a subfield of text mining, computational linguistics, and Natural Language Processing (NLP) [1]. NLP and opinion mining on the Internet are widely used in various domains, including e-commerce [4], economics [5], politics [6], social service [7], culture [8], and global health [9]. In addition to identifying abnormal words and speech, opinion mining distinguishes between normal and abnormal phrases. Considering the large amount of information exchanged daily in the form of texts in the virtual space and social media, along with the diverse linguistic complexities involved, opinion mining faces many challenges, making it a highly suitable research area. Social media platforms like X (formerly Twitter), LinkedIn, Amazon, and Facebook [2] are among the possible applications of these algorithms. Depending on the media type and the analyzed text, sentiment analysis is usually categorized into short sentences, such as tweets, and long sentences, such as reviews. Typically, users employ short sentences on social media platforms, whereas long sentences are used, for example, in movie reviews, bookings, or e-commerce sites. In both cases, it is crucial to identify sentiment words as significant parts of the sentence, extract features based on them, and subsequently analyze the sentiment or polarity. Several sentiment

dictionaries have been made for general sentiment words or a specific domain. Some are created manually, while others are machine- or semi-manual-generated. Considering the large number of datasets on the Internet and the daily changing sentimental scores of words, it is evident that regular updates are necessary for all sentiment dictionaries [10]. Since building general and domain-specific sentiment dictionaries is time-consuming and expensive, and dictionaries based on machine learning miss some linguistic details of natural language, this article proposes a hybrid mathematical-heuristic solution to update sentimental scores of existing dictionaries. In the proposed solution, the sentiment words from the database and their tags are first mapped to a sparse space, and then their scores are updated based on solving an unbalanced pseudo-linear NLP equation system. Moreover, the solution provides an opportunity to extract and suggest new words repeated in sentimental phrases and use sentiment words assigned in existing dictionaries. The innovative features of this research are the following:

- A novel approach that converts text and labels into sparse equations;
- Two mathematical methods used to determine two initial populations for a genetic algorithm (GA);
- A novel hybrid hierarchical mathematical heuristic approach to address a polarity detection problem.

2. Literature Review

Several studies on sentiment analysis/mining and opinion mining/analysis have been conducted on X data. In general, sentiment analysis is divided into three categories [11]:

- Document level: The document is labeled as positive, negative, or neutral.
- Sentence level: Each sentence is assigned one of the three categories: positive, negative, or neutral.
- Aspect and feature level: In this case, the document is categorized as positive, negative, or neutral, considering its entire structure. This model is also referred to as the perspective level.

The present study belongs to the first two categories: Document and Sentence. Regarding the used approach, sentiment analysis methods are often divided into four categories: supervised, unsupervised, lexical, or hybrid. The proposed solution is a polarity detection method that belongs to the hybrid group because it combines a supervised approach with lexicons [12]. Zhang et al. [13] used word embedding [14,15] for feature extraction and a Convolutional Neural Network (CNN) for polarity detection. Xia et al. [16] proposed a three-step method for polarity detection and made a dictionary of sentiment words through joint learning. Wu et al. [17] used a semi-supervised method based on a variational autoencoder for polarity detection and sentiment word splitting. Samb et al. [18] increased the accuracy of dictionary-based sentiment analysis methods using trigrams. Raju et al. [19] analyzed the performance of the Support Vector Machine (SVM) classifier in sentiment analysis based on various features.

Ito et al. [20] employed a contextual neural network to detect the polarity of a sentence. Kande et al. [21] used a lexicon-based method and scored sentiment words. Their approach determines the polarity of a sentence using a multiplicative combination of sentiment word scores. Alharbi and Doncker [22] combined user posting history on X with sentiment analysis using a deep neural network to detect sentiment in the data under study. Kraus and Feuerriegel [23] developed a tree-structured tensor-based deep neural network method for sentiment analysis, using linguistic dependencies to create a tree structure. Shuang et al. [24] proposed an aspect-level sentiment classification method using an attention-enabled dual long short-term memory (LSTM) network that relies on sentiment word sequences. Zhao et al. [25] used convolutional graph networks to extract relationships between sentiment words. Their work belonged to the category of aspect-level sentiment classification. Kumar and Jaiswal [26] compared different sentiment analysis methods on X data that used soft computing techniques. According to the authors, hybrid methods

yield better results. For contextual sentiment analysis at the word level, Ito et al. [27] used sentiment word scores in a three-step process. Dashtipor et al. [28] proposed a method for sentiment analysis in the Farsi language, considering grammatical rules using deep neural networks. Wei et al. [29] used a bidirectional LSTM for sentiment analysis. The main innovation of their work is the use of orthogonal multipolarity attention.

Naseem et al. [30] used a Deep Intelligent Contextual Embedding method for sentiment analysis of X data. Cambria et al. [31] proposed SenticNet 6, an ensemble sentiment analysis application. Ito et al. [32] presented a contextual sentiment neural network for document-level sentiment analysis. Gupta and Joshi [33] suggested a hybrid approach that includes feature extraction, SVM, and a local contextual semantic dictionary to increase the efficiency of X sentiment analysis. Santhiya et al. [34] compared polarity detection methods in social media and concluded that hybrid methods generally have higher accuracy than others. Carvalho and Plastino [35] performed a comparative study on the features used in X sentiment analysis and highlighted the strengths and weaknesses of different parts. Bandhakavi et al. [36] improved the efficiency of X sentiment analysis methods using emotion-sensitive polarity lexicons.

Kochari et al. [37] presented a Persian part of a speech tagging system using an LSTM model, which can be used as a sentiment analysis method. Zargari et al. [38] suggested an order-sensitivity sentiment dictionary of word sequences containing intensifiers. Žunić et al. [39] proposed an aspect-based method for sentiment analysis based on a dependency analyzer. Cambria et al. [40] suggested a sentiment analysis method based on concept extraction. Junior et al. [41] used a set of sentiment analysis features and employed genetic programming to combine them to achieve the best accuracy. Li et al. [42] suggested a domain-dependent X sentiment analysis method that uses two types of word embedding for more precise sentiment expression, attention-based bidirectional LSTM and multi-contextual CNN. Polignano et al. [43] proposed a polarity detection model that simultaneously used transformer learning and data obtained from a polarized lexicon, the Bidirectional Encoder Representations from Transformers (BERT) model. Table 1 presents details of state-of-the-art studies in sentiment analysis, which are dictionary-based and use deep or machine learning-based methods for polarity detection.

Table 1. Recent sentiment analysis methods.

Reference	Year	Classifier	Feature Extraction
Zhang et al. [13]	2018	CNN	Word embedding
Xia et al. [16]	2016	Joint learning	–
Wu et al. [17]	2019	Variational autoencoder	Sentiment word splitting
Samb et al. [18]	2019	Dictionary-based	Trigrams
Alharbi and Doncker [22]	2019	Deep neural network	User posting history
Kraus and Feuerriegel [23]	2019	Deep neural network	Tree-structured tensor-based
Shuang et al. [24]	2019	Dual LSTM network	Sentiment word sequences
Zhao et al. [25]	2020	Convolutional graph network	–
Wei et al. [29]	2020	Bidirectional LSTM	–
Naseem et al. [30]	2020	Deep Intelligent Contextual Embedding	–
Zargari et al. [38]	2023	Fuzzy	Order-sensitivity sentiment dictionary
Polignano et al. [43]	2022	BERT	–
Kim and Jang [44]	2023	Deep Learning	Sentiment Dictionary
Gupta et al. [45]	2023	Machine learning	Senticnet 5
Gopi et al. [46]	2023	SVM	TF-IDF
Tong et al. [47]	2024	Bidirectional LSTM	BERT as word embedding model

Table 1. *Cont.*

Reference	Year	Classifier	Feature Extraction
Aslan [47]	2024	CNNs and Bidirectional LSTM	Multistage Feature Extraction
Ben et al. [48]	2023	Machine learning and deep learning	TF-IDF
Raza et al. [49]	2023	SentiWordNet	PoS tagging
Ramos et al. [50]	2024	Machine learning and deep learning	Multiple sentilexicons and dense word representations
Bashiri and Naderi [51]	2024	Gaussian Mixture Model	Emotional expression
Young et al. [52]	2024	SocialSent algorithm	–
Shahade et al. [53]	2023	Deep Learning	Text vectorization

All of the methods above use sentiment dictionaries or build their dictionaries. These dictionaries, such as WordNet [54], include sentiment words and their scores as a form of polarity or sentiment. Considering the advantages and disadvantages of previous studies and based on our best search, no hybrid algebraic-heuristic solution has been used to assign scores to sentiment words. The proposed solution gives scores to the sentiment words using a system of equations and mathematically solves the polarity detection problem with maximum accuracy. This research focuses on assigning sentiment word scores using a heuristic method, deterministic mathematical analysis, equations solved in sparse space, and linear algebra.

3. Proposed Solution

The proposed solution consists of training and testing steps. Figure 1 shows the block diagrams of the proposed solution training step. The sentimental words are selected in this step, and their scores are calculated. The training consists of seven blocks, including reading the tagged dataset, removing unnecessary words, creating summarized phrases, establishing equations, finding two initial answers using two different approaches, and finding the final answer using a GA (Figure 1). In the testing step, the output of the training step is used to find the polarity of a new sentimental phrase, and the accuracy of the proposed dictionary is assessed. The details of these two steps are explained below.
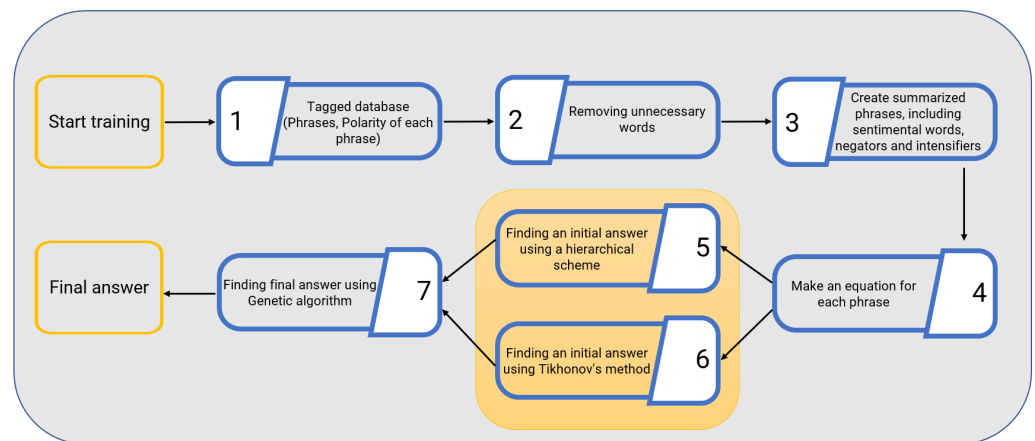

Figure 1. Training step of the proposed solution.

3.1. Training Step

The main blocks of the training phase are blocks 5, 6, and 7 (Figure 1). The solution begins by reading the labeled sentences in block 1. After removing unnecessary words from an input sentence, the sentence is summarized into phrases containing sentiment words, negators, and intensifiers (blocks 2 and 3).

Each unigram, including sentiment words, is assumed to be a variable. Intensifiers and negators are considered as multiplicative variables. For instance, "It's so good and

tasty" is a positive sentence consisting of three essential words: "so", "good" and "tasty". The corresponding equation can be represented as:

$$x_1 \times x_2 + x_3, \tag{1}$$

where x_1 is the score of "so" as an intensifier, while x_2 and x_3 are the scores of "good" and "tasty" as sentiment words, respectively. The answer to the equation is considered 1 (one) for positive sentences, -1 for negative sentences, and 0 (zero) for neutral sentences, so Equation (1) can be rewritten as:

$$x_1 \times x_2 + x_3 = 1. \tag{2}$$

In block 4, for each sentence, a combination of some variables is considered that leads to a result: positive/negative (± 1) as sentimental intensity. At the end of this stage, an algebraic system of pseudo-linear equations with N equations and M unknowns is established (block 4 of Figure 1). Since the number of equations and unknowns is unequal, the proposed solution finds two initial answers for this sparse system: polarity detection, blocks 5 and 6. These two initial solutions and a set of random keys serve as initial populations for a heuristic method to identify sentimental scores associated with sentiment words in block 7. The main innovation of this work is the generation of equations and solving them with a combination of mathematical and heuristic approaches. The following sections discuss the proposed mathematical methods for finding two initial solutions and the heuristic approach to finding final scores.

3.1.1. First Approach

This approach solves equations sequentially based on the number of variables. The block diagram of this approach is shown in Figure 2.
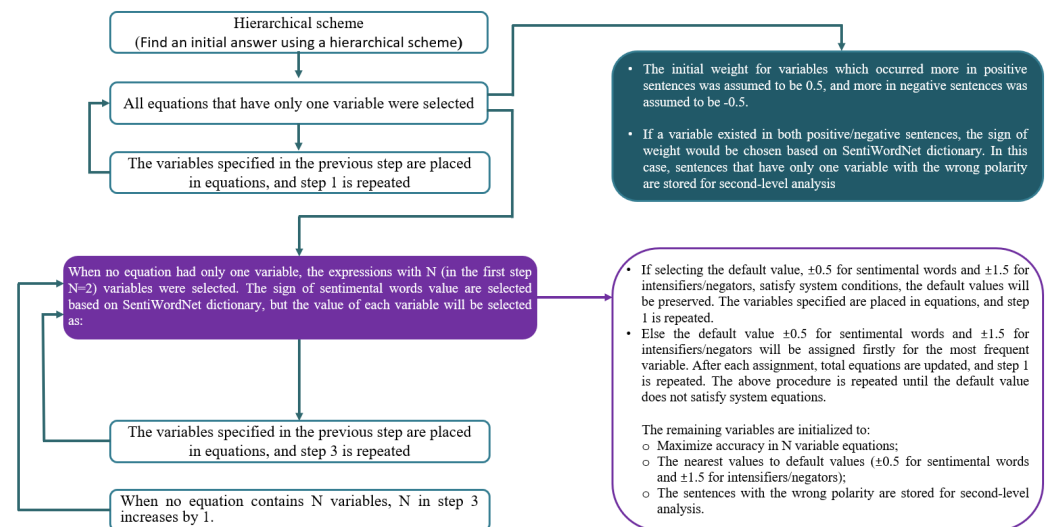

Figure 2. Block diagram of the proposed solution for finding the first initial answer.

The steps of the first approach are as follows:

1. All equations that have only one variable are selected.
    - An initial score of (0.5) is assigned to variables occurring more frequently in positive sentences and (-0.5) in the case of negative sentences.
    - When a variable exists in both positive and negative sentences, the sign of its score would be chosen based on the SentiWordNet dictionary. In this case, some sentences with only one variable show incorrect polarity. These sentences are stored for a second-level analysis.

2. The variables specified in the previous step are substituted into all equations, and step 1 is repeated.
3. If no equation has only one variable, the expressions with N (*initially $N = 2$*) variables are selected. The sign of sentiment word score is determined based on the SentiWordNet dictionary, and the value of each variable is determined as follows:

    *   If selecting the default value (± 0.5) for sentiment words and (± 1.5) for intensifiers/negators) satisfies the system conditions, the default values are retained. The specified variables are substituted into equations, and step 1 is repeated.
    *   Otherwise, the default values, (± 0.5) for sentiment words and (± 1.5) for intensifiers/negators, are initially assigned for the most frequent variable. After each assignment, the overall equations are updated, and step 1 is repeated.

    The above procedure is repeated until the default value no longer satisfies the equations system. The remaining variables are initialized according to the following:

    *   Maximize Accuracy in the N variable equations;
    *   The nearest values are assigned to the default values (± 0.5) for sentiment words and (± 1.5) for intensifiers/negators. In this case, some sentences show incorrect polarity and are stored for a second-level analysis.

4. The variables specified in the previous step are substituted into equations, and step 3 is repeated.
5. When no equation contains N variables, N is increased by 1 (one) in step 3.

    The final scores obtained using the above scheme are used as one of the initial populations in the GA.

3.1.2. Second Initial Approach

The second approach used in this research is based on Tikhonov's method. The sentimental equations formulated in block 4 constitute an unbalanced system of pseudo-linear equations, which can be solved using Tikhonov's method known as ridge regression. The Tikhonov method is a widely employed technique for stabilizing discrete ill-posed problems and was independently proposed by Philips and Tikhonov [55]. Like the least squares method, Tikhonov's method assumes that the observation error is random and that the error probability distribution function (PDF) is a Gaussian (Normal) PDF with zero mean. As it is impossible to obtain an answer using only the least squares condition in ill-posed problems, like the cases being studied, the method adds a regularization term to the error of the least square way to solve the problem. An ill-conditioned linear system has the following form:

$$A_{M \times N} \, x_{N \times N} = b_{M \times N}, \qquad M \geq N, \tag{3}$$

where A is the coefficient matrix, b is derived from labeled data, N is the number of variables and M is the number of labeled data. In ill-posed inverse problems, the answers are highly sensitive to the used labeled data. In ill-posed inverse problems, regularization techniques like Tikhonov's technique are frequently used to provide a decent and stable solution. In Tikhonov's regularization technique, the cost function is a combination of the squared error and a weighted norm2 of the response vector:

$$\min_{x} \left\{ \|Ax - b\|^2 + \mu^2 \|x\|^2 \right\}, \tag{4}$$

where μ is the regularization parameter. Minimizing this cost function is equivalent to solving:

$$(A^T A + \mu^2 I)x = A^T b, \tag{5}$$

where I is the identity matrix. This approach balances answers with the regularization constraint, providing more stable solutions for ill-posed problems. If the Singular Value

Decomposition (SVD) of matrix A is obtained, it can be written as $A = U\Sigma V^T$, where the following are applied:

- $U = (u_1, \ldots, u_m)$ and $V = (v_1, \ldots, v_n)$ are orthonormal matrices;
- $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n) \in \text{R}_{m \times n}$;
- and $\sigma_1, \sigma_2, \ldots, \sigma_n$ are the singular values of A, with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$.

In this case, the Tikhonov solution of Equation (5), i.e., the answer to Equation (3), is

$$x_\mu = \sum_{i=1}^{n} f_i \frac{u_i^T b}{\sigma_i} v_i,\tag{6}$$

where F_i is the Tikhonov filter factor, which is defined as [55]

$$f_i = \frac{\sigma_i^2}{\sigma_i^2 + \mu^2}.\tag{7}$$

In this study, the truncated SVD (TSVD) method [56] is employed to calculate the coefficients X_k. In TSVD, the F_i values are approximated as follows:

$$f_i = \frac{\sigma_i^2}{\sigma_i^2 + \mu^2} \approx \begin{cases} 1, & \sigma_i \geq \mu, \\ 0, & \sigma_i < \mu. \end{cases}\tag{8}$$

TSVD is a regularization method that uses the best low-rank approximation of A. Tikhonov's method has been used in several applications, such as the calculation of the relaxation time distribution function in impedance spectroscopy [57], acoustic tomography [58], and dynamic load identification [59]. In the proposed solution, the scores calculated by the Tikhonov method form the second initial population of the GA.

3.1.3. Genetic Algorithm

The last step of the training process is based on a GA, employed to find the final scores as the solution to the proposed system of equations. In the start, the GA needs a set of initial populations. Two of these initial populations are the scores obtained by the previously presented approaches. The other initial populations are randomly selected. The objective function of the GA is to maximize the accuracy of the overall equations. The assumed constraints were as follows:

- For sentiment words, all scores are in the range of ± 1;
- For intensifiers/negators, all scores are limited to the mean of $\pm[1, 2]$ for intensifiers/negators (positive for intensifiers and negative for negators);
- The sign of the scores assigned to sentiment words must be the same as in the SentiWordNet dictionary;
- The sign of negators must be harmful, and the sign of intensifiers must be positive;
- Finally, the best answer of block 6 (Figure 1) is stored as the final score.

3.2. Testing Step

The test step of the proposed solution is shown in Figure 3. The input used to test the system includes test samples, a sentiment dictionary, negators, intensifiers, and the calculated scores in the training step. After removing unnecessary words from the input samples and splitting them into the remaining terms (blocks 1 to 3), the equations are established in block 4 using the scores calculated in the training step. It becomes straightforward to obtain the overall score for each sentence. Positive scores indicate positive polarity and vice versa; negative scores indicate negative polarity.
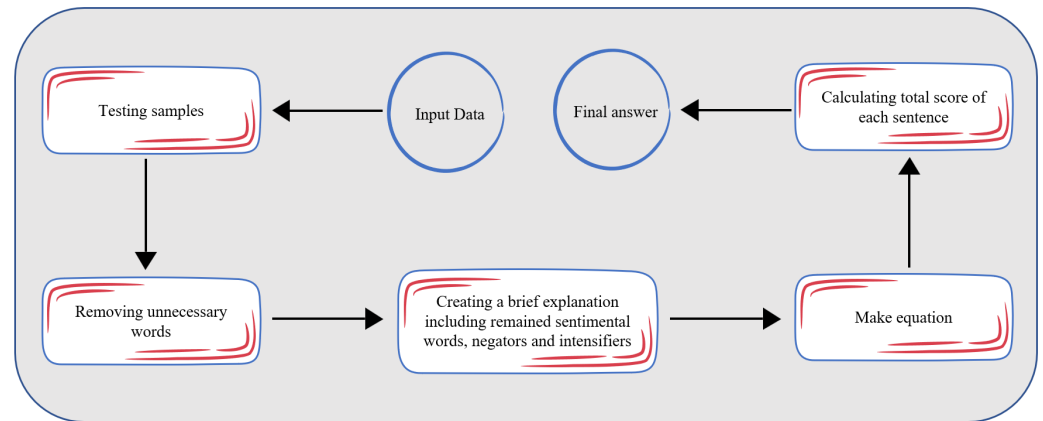
Figure 3. Test step of the proposed solution.

4. Experimental Results

This section describes the datasets and dictionaries used in the experiments and presents details of the training step and the obtained results. Section 4.1 introduces the datasets and dictionaries used for training, testing, and validating the proposed solution. Three scenarios were used in the training and testing step, leading to the experimental results described in Section 4.2.

4.1. Datasets and Dictionaries

To evaluate the efficiency of the proposed solution, the proposed model was trained in scenarios using Amazon and X datasets. Subsequently, the trained model was tested on Amazon, X, and Taboada datasets to assess its performance.

4.1.1. X

X (Formerly Twitter) is a social networking and microblogging service that allows users to send text messages of up to 280 characters, known as tweets. X was created in March 2006 by Jack Dorsey and launched in July 2006. By 2012, more than 100 million users had tweeted 340 million tweets daily, and the service handled an average of 1.6 billion daily search queries. This work used the publicly available X corpus SemEval-2013 (Task 2) for polarity detection, which was published in the SemEval-2013 competition. Task 2 of the competition consisted of term-level and message-level polarity detection. The main objective was message-level polarity detection, i.e., classifying a tweet into positive or negative classes. This dataset includes 2978/1162 positive/negative samples for training and 1306/484 positive/negative samples for testing. Due to the unbalanced nature of the samples, the balanced precision [60] was used as the objective function for the GA during the training step.

4.1.2. Amazon

Amazon's website is one of the first successful examples of an American e-commerce company. This site was first launched in 1994 by Jeff Bezos in Seattle, USA. The company's core business started as an online bookstore in 1995 and followed the legal and regulatory processes in the United States. The ESWC Semantic Sentiment Analysis consisted of several challenges. The refined sentiment analysis, one of these challenges, was divided into five subtasks related to the classification and quantification of sentiment polarity according to a two- or five-point scale in Amazon reviews. The proposed method was tested in the binary polarity detection task.

4.1.3. Taboada Database

The Taboada database contains eight domains containing fifty positive and negative phrases [61]. This database was collected and labeled by Stanford University in 2004 and is often used as a reference for comparing different methods for sentiment analysis.

4.1.4. Dictionary

The proposed solution requires a list of sentiment words and the sign of their scores. In this work, the following commonly used sentiment dictionaries were used: GI Dictionary [62], WordNet dictionary [54], ANEW dictionary [63], and the SentiWordNet dictionary [64]. Among these, the primarily used dictionary is SentiWordNet 3.0 [64], an improved version of SentiWordNet 1.0. This dictionary is a lexical resource created by automatic annotation of WordNet.

4.2. Results

In the training step of the proposed solution, the following assumptions and options were considered:

- In establishing equations, the positive polarity answer was assumed to be 1 (one), and the negative polarity was assumed to be −1;
- The list of sentiment words and their corresponding signs was derived from Senti-WordNet 3.0 [64];
- The list of intensifiers was made using the list of general intensifiers [65];
- The negators were selected based on the research conducted by Kiritchenko and Saif [66];
- The proposed dictionary contains 14,072 negative and 15,023 positive sentiment words (29,095 words). Additionally, it includes 15 negators and 176 general intensifiers.

Three different scenarios were considered for calculating the scores:

- First: The training dataset was an X dataset, and the calculated scores were tested on X dataset test samples;
- Second: The training set was 70% (700,000 samples) randomly selected from the Amazon dataset, and the calculated scores were tested on the remaining samples;
- Third: samples were selected from the Amazon and X datasets for the training dataset, and the calculated scores were tested on the Taboada dataset.

These scenarios were selected based on related state-of-the-art studies. As an enhanced scoring system, the proposed solution should increase the accuracy of the existing sentimental word scoring system. The three different NLP scoring systems proposed in [67–73] were considered, and the proposed scheme was validated in all of them. The GA parameters were configured as follows:

- For the first scenario, the initial population size was set to $2^{10} = 1024$, while for the second and third scenarios, it was set to $2^{16} = 65536$, considering the complexity and number of equations involved;
- The GA was implemented in the binary form;
- The resolution of the scores was 0.01 in decimal mode, i.e., 7 bits for each score in binary form;
- The total length of each population was $7 \times NVar$;
- The $NVar$ value for the first scenario was 11,472, meaning there were only 11,472 sentiment words, intensifiers, and negators from the used list in the train samples of the X dataset. For the other two scenarios, the value of $NVar$ was 73,856;
- The type of crossover was multi-point crossover, and the number of crossover points was set to 1% of the binary vector length;
- The probability of mutation was assumed to be 0.01;
- The selection scheme was the roulette wheel;
- The objective function in the first scenario was the balanced accuracy [74] and the standard accuracy for the other two;
- Sentiment words, intensifiers, and negators that did not appear in the train samples were excluded from the list;
- All implementations were accomplished using Matlab 2021a.

All the parameters of the GA were set based on [75]. To compare the proposed solution with other state-of-the-art methods, the True Positives (TP), True Negatives (TN),

False Positives (FP), and False Negatives (FN) were calculated, and the precision, recall, F-measure, and accuracy metrics were derived from them.

4.2.1. First Scenario

In this scenario, the training dataset was an X dataset, and the calculated scores were tested on X dataset test samples. The proposed solution was compared with the methods proposed by Gupta and Joshi [67], Saif et al. [68], MiMuSA [76] and VIKOR Optimization [77]. Table 2 shows the accuracy of the proposed solution on a test dataset of the X dataset. The test dataset comprises 1306 positive samples and 484 negative samples. Based on the simulation results, a weakness of the previous methods is the low number of TN. The proposed solution successfully addressed this weakness and achieved a TN value at least three times higher, although it decreased the TP value. The suggested solution outperforms the previous approaches in terms of overall accuracy, precision, recall, and F-measure, particularly in recall. The last column of Table 2 shows the differences between the accuracy of each state-of-the-art method relative to the proposed solution. As shown in Table 2, only the VIKOR Optimization [77] had a slightly better performance than the proposed solution, which is not dictionary-based and only detects polarity based on a very complex hybrid approach.

Table 2. Comparison of state-of-the-art methods and the proposed solution using the X dataset.

Classifier	TP	FP	TN	FN	Precision	Recall	F1	Accuracy	Accuracy Difference
SVM [67]	958	348	27	457	73.32	67.7	70.4	55.03	2.4
Naive Bayesian [67]	773	533	53	431	59.2	64.2	61.6	46.15	11.28
DT [67]	810	496	54	430	61.99	65.3	63.6	48.27	9.16
SVM [68]	881	425	5	479	67.45	64.8	66.1	49.50	7.93
MiMuSA [76]	849	457	149	335	65.01	71.71	68.19	55.75	1.68
VIKOR Optimization [77]	923	383	166	318	70.67	74.37	72.47	60.84	−3.41
Proposed solution	876	430	152	332	74.72	72.52	69.69	57.43	-

4.2.2. Second Scenario

In this scenario, the training dataset consisted of 70% (700,000 samples) of randomly selected samples from the Amazon dataset, and the calculated scores were tested on the remaining samples. The proposed solution was compared with the methods of Sygkounas et al. [69], Di Rosa and Durante [70], and Petrucci and Dragoni [71]. Table 3 shows the results using the Amazon test dataset. The test dataset consists of 150,000 positive samples and equal negative samples. In the proposed solution, the training procedure was performed ten times to eliminate the effect of a random selection of training and test data, and the reported result is the median of the obtained results.

Table 3. Comparison of state-of-the-art methods and the proposed solution using the Amazon test dataset.

Classifiers	TP	FP	TN	FN	Precision	Recall	F1	Accuracy	Accuracy Difference
Sygkounas et al. [69]	128,512	21,488	136,573	13,427	85.67	90.54	88.04	88.36	2.52
Di Rosa & Durante [70]	125,665	24,335	137,251	12,749	83.78	90.789	87.142	87.64	3.24
Petrucci & Dragoni [71]	122,755	27,245	135,134	14,866	81.84	89.198	85.359	85.96	4.92
Jalali et al. [78]	135,344	14,640	137,289	12,727	90.36	91.40	90.88	90.88	0
Proposed solution	135,354	14,646	137,283	12,717	88.24	91.234	89.71	90.88	-

Based on the simulation results obtained in this scenario, one can conclude that the overall accuracy, precision, recall, and F-measure of the proposed solution are better than those of the other methods. The proposed solution has increased the accuracy by a minimum of 2.52% compared to the two first methods. The closest result to the one obtained by the proposed solution was of Jalali et al. [78], which was expected since both methods are Tikhonov's technique based. However, it is important to note that the TSVD

technique used in the proposed solution is faster than that used in the method proposed by Jalali et al. [78] and requires fewer calculations.

4.2.3. Third Scenario

In this scenario, the training dataset was selected from the Amazon and X datasets, and the methods under study were then tested on the Taboada dataset. The proposed solution was compared with the methods proposed by Zargari et al. [72] and Dey et al. [73]. The evaluation conditions were chosen similarly to those used in the study of Zargari et al. [72]. Considering the two different methods used by Zargari et al., the proposed solution was compared to the one that led to the best result obtained by the authors.

The results in Table 4 demonstrate that the proposed solution enhances scoring efficiency by improving the accuracy for negative sentences. Although the proposed solution may slightly decrease the accuracy of detecting positive polarity, the overall outcome indicates that it outperforms the state-of-the-art methods in polarity detection.

Table 4. Comparison of state-of-the-art methods and the proposed solution using the Taboada dataset.

Classifiers	TP	FP	TN	FN	Precision	Recall	F1	Accuracy	Accuracy Difference
Dey et al. [73]	386	114	124	376	77.2	50.66	61.17	51.00	7
Zargari et al. [72]	390	110	162	338	78	53.57	63.52	55.20	2.8
Proposed solution	356	144	224	276	71.2	56.33	62.9	58.00	-

4.2.4. Sensitivity Analysis

To validate the correctness of the selected GA parameters, a preliminary sensitivity analysis was conducted. Increasing the initial population size to 2^{11} and 2^{17} for the first scenario and the other two scenarios, respectively, increased the execution time but did not affect the accuracy of the proposed solution. Reducing the initial population size to 2^9 for the first scenario and 2^{15} in the other two reduced the execution time, the accuracy of the proposed solution, and the stability of obtained scores in different GA executions. The change in the score obtained in different executions, decreasing the stability of the obtained score, means that the GA is at the local minimum, which is unacceptable.

Concerning the resolution of the score, by increasing the resolution from 0.01 to 0.001, the number of necessary bits increases from 7 to 10, which increases the memory consumption and the number of GA calculations by about 40%. Still, in the end, the final accuracy did not change. Reducing the resolution from 0.01 to 0.1 reduced the final accuracy and was unacceptable.

Increasing the mutation probability from 0.01 to 0.02 did not affect the final accuracy, but it significantly increased the GA convergence time. Reducing the mutation probability to 0.005 led to a decrease in the stability of the obtained score in different GA executions. The change in score obtained in different executions means that the response is at the local minimum, which is unacceptable.

5. Conclusions

This study proposes a novel hybrid solution for solving unbalanced systems of pseudo-linear algebraic equations. The proposed solution uses two mathematical approaches to find two initial solutions, which are added to a heuristic method randomly assigning initial populations. The heuristic process finds the final solution based on the defined objective function of GA. The proposed hybrid solution solved an NLP scoring scheme in a polarity detection method. The results showed the proposed solution's effectiveness compared to other machine learning, fuzzy, and stochastic methods. Notably, the proposed solution surpasses the state-of-the-art results of the SemEval-2013 competition on X data and also the state-of-the-art results of the ESWC Semantic Sentiment Analysis 2016 competition with comparable accuracy. Finally, the proposed solution was trained on the Amazon and X datasets and tested on the Taboada dataset. In this case, the proposed solution shows

2.8% and 7% better accuracy than the two latest methods tested using this dataset. To the authors' knowledge, the proposed solution is the first NLP approach that combines mathematical and heuristic approaches for sentiment word scoring.

Updating the final sentimental word scores using real-time conditions in social media is an exciting direction for future work. As a suggestion, a modified version of the proposed solution could update some sentiment word scores based on current social, political, or cultural events. This would allow the sentiment scoring system to adapt to changing connotations and associations of words over time based on real-world contexts. For example, certain words may shift to having more positive or negative sentiments following major events reported on social media. The proposed solution can automatically incorporate a mechanism to adjust sentiment values for affected terms. This would keep the scoring dictionary current and prevent sentiment analysis from becoming outdated as language evolves. In addition, one of the most interesting possible future works is to analyze the uncertainty and changes of assigned scores over time and with the change of some concepts in social culture as the robustness of the proposed solution. The very high dependence of the obtained scores on a defined equation system is one of the weaknesses of the proposed method. Obviously, by changing the form of the defined quasi-linear equations, the obtained score will change. This weakness should be addressed in future work.

References

1. Cortis, K.; Davis, B. Over a decade of social opinion mining: A systematic review. *Artif. Intell. Rev.* **2021**, *54*, 4873–4965. [CrossRef] [PubMed]
2. Wang, R.; Zhou, D.; Jiang, M.; Si, J.; Yang, Y. A survey on opinion mining: From stance to product aspect. *IEEE Access* **2019**, *7*, 41101–41124. [CrossRef]
3. Messaoudi, C.; Guessoum, Z.; Ben Romdhane, L. Opinion mining in online social media: A survey. *Soc. Netw. Anal. Min.* **2022**, *12*, 25. [CrossRef]
4. Alnahas, D.; Aşık, F.; Kanturvardar, A.; Ülkgün, A.M. Opinion Mining Using LSTM Networks Ensemble for Multi-class Sentiment Analysis in E-commerce. In Proceedings of the 2022 3rd International Informatics and Software Engineering Conference (IISEC), Ankara, Turkey, 15–16 December 2022; pp. 1–6. [CrossRef]
5. Chen, M.Y.; Chen, T.H. Modeling public mood and emotion: Blog and news sentiment and socio-economic phenomena. *Future Gener. Comput. Syst.* **2019**, *96*, 692–699. [CrossRef]
6. Santos, J.S.; Bernardini, F.; Paes, A. A survey on the use of data and opinion mining in social media to political electoral outcomes prediction. *Soc. Netw. Anal. Min.* **2021**, *11*, 103. [CrossRef]
7. Hajihashemi, V.; Ameri, M.M.A.; Gharahbagh, A.A.; Bastanfard, A. A pattern recognition based Holographic Graph Neuron for Persian alphabet recognition. In Proceedings of the 2020 International conference on machine vision and image processing (MVIP), Qom, Iran, 18–20 February 2020; pp. 1–6. [CrossRef]
8. He, Q. Hot Spot Mining and Analysis Model of Sports Microblog Culture Public Opinion Based on Big Data Environment. In Proceedings of the 2021 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS), Shenyang, China, 29–31 July 2021; pp. 739–743. [CrossRef]
9. Piedrahita-Valdés, H.; Piedrahita-Castillo, D.; Bermejo-Higuera, J.; Guillem-Saiz, P.; Bermejo-Higuera, J.R.; Guillem-Saiz, J.; Sicilia-Montalvo, J.A.; Machío-Regidor, F. Vaccine hesitancy on social media: Sentiment analysis from June 2011 to April 2019. *Vaccines* **2021**, *9*, 28. [CrossRef]

10. Rubtsova, Y. Reducing the deterioration of sentiment analysis results due to the time impact. *Information* **2018**, *9*, 184. [CrossRef]

11. Alsaeedi, A.; Khan, M.Z. A study on sentiment analysis techniques of Twitter data. *Int. J. Adv. Comput. Sci. Appl.* **2019**, *10*, 361–374. [CrossRef]

12. Mittal, A.; Patidar, S. Sentiment analysis on twitter data: A survey. In Proceedings of the 7th International Conference on Computer and Communications Management, Bangkok, Thailand, 27–29 July 2019; pp. 91–95. [CrossRef]

13. Zhang, Z.; Zou, Y.; Gan, C. Textual sentiment analysis via three different attention convolutional neural networks and cross-modality consistent regression. *Neurocomputing* **2018**, *275*, 1407–1415. [CrossRef]

14. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient estimation of word representations in vector space. *arXiv* **2013**, arXiv:1301.3781. [CrossRef]

15. Tang, D.; Wei, F.; Yang, N.; Zhou, M.; Liu, T.; Qin, B. Learning sentiment-specific word embedding for twitter sentiment classification. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Baltimore, MD, USA, 22–27 June 2014; pp. 1555–1565. [CrossRef]

16. Xia, R.; Xu, F.; Yu, J.; Qi, Y.; Cambria, E. Polarity shift detection, elimination and ensemble: A three-stage model for document-level sentiment analysis. *Inf. Process. Manag.* **2016**, *52*, 36–45. [CrossRef]

17. Wu, C.; Wu, F.; Wu, S.; Yuan, Z.; Liu, J.; Huang, Y. Semi-supervised dimensional sentiment analysis with variational autoencoder. *Knowl. Based Syst.* **2019**, *165*, 30–39. [CrossRef]

18. Samb, S.M.K.; Kandé, D.; Camara, F.; Ndiaye, S. Improved bilingual sentiment analysis lexicon using word-level trigram. In Proceedings of the 2019 IEEE 5th International Conference on Computer and Communications (ICCC), Chengdu, China, 6–9 December 2019; pp. 112–119. [CrossRef]

19. Raju, K.D.; Jayasingh, B.B. Influence of Syntactic, Semantic and Stylistic Features for Sentiment Identification of Messages Using Svm Classifier. *Int. J. Sci. Technol. Res.* **2019**, *8*, 2551–2557.

20. Ito, T.; Tsubouchi, K.; Sakaji, H.; Izumi, K.; Yamashita, T. Csnn: Contextual sentiment neural network. In Proceedings of the 2019 IEEE International Conference on Data Mining (ICDM), Beijing, China, 8–11 November 2019; pp. 1126–1131. [CrossRef]

21. Kandé, D.; Camara, F.; Ndiaye, S.; Guirassy, F.M. FWLSA-score: French and wolof lexicon-based for sentiment analysis. In Proceedings of the 2019 5th International Conference on Information Management (ICIM), Cambridge, UK, 24–27 March 2019; pp. 215–220. [CrossRef]

22. Alharbi, A.S.M.; de Doncker, E. Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information. *Cogn. Syst. Res.* **2019**, *54*, 50–61. [CrossRef]

23. Kraus, M.; Feuerriegel, S. Sentiment analysis based on rhetorical structure theory: Learning deep neural networks from discourse trees. *Expert Syst. Appl.* **2019**, *118*, 65–79. [CrossRef]

24. Shuang, K.; Ren, X.; Yang, Q.; Li, R.; Loo, J. AELA-DLSTMs: Attention-enabled and location-aware double LSTMs for aspect-level sentiment classification. *Neurocomputing* **2019**, *334*, 25–34. [CrossRef]

25. Zhao, P.; Hou, L.; Wu, O. Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification. *Knowl. Based Syst.* **2020**, *193*, 105443. [CrossRef]

26. Kumar, A.; Jaiswal, A. Systematic literature review of sentiment analysis on Twitter using soft computing techniques. *Concurr. Comput. Pract. Exp.* **2020**, *32*, e5107. [CrossRef]

27. Ito, T.; Tsubouchi, K.; Sakaji, H.; Yamashita, T.; Izumi, K. Word-level contextual sentiment analysis with interpretability. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 4231–4238. [CrossRef]

28. Dashtipour, K.; Gogate, M.; Li, J.; Jiang, F.; Kong, B.; Hussain, A. A hybrid Persian sentiment analysis framework: Integrating dependency grammar based rules and deep neural networks. *Neurocomputing* **2020**, *380*, 1–10. [CrossRef]

29. Wei, J.; Liao, J.; Yang, Z.; Wang, S.; Zhao, Q. BiLSTM with multi-polarity orthogonal attention for implicit sentiment analysis. *Neurocomputing* **2020**, *383*, 165–173. [CrossRef]

30. Naseem, U.; Razzak, I.; Musial, K.; Imran, M. Transformer based deep intelligent contextual embedding for twitter sentiment analysis. *Future Gener. Comput. Syst.* **2020**, *113*, 58–69. [CrossRef]

31. Cambria, E.; Li, Y.; Xing, F.Z.; Poria, S.; Kwok, K. SenticNet 6: Ensemble application of symbolic and subsymbolic AI for sentiment analysis. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management, Virtual Event, 19–23 October 2020; pp. 105–114. [CrossRef]

32. Ito, T.; Tsubouchi, K.; Sakaji, H.; Yamashita, T.; Izumi, K. Contextual sentiment neural network for document sentiment analysis. *Data Sci. Eng.* **2020**, *5*, 180–192. [CrossRef]

33. Gupta, I.; Joshi, N. Enhanced twitter sentiment analysis using hybrid approach and by accounting local contextual semantic. *J. Intell. Syst.* **2019**, *29*, 1611–1625. [CrossRef]

34. Santhiya, P.; Kogilavani, S.; Malliga, S. Sentiment Analysis Classifiers for Polarity Detection in Social Media Text: A Comparative Study. In Proceedings of the 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2–4 December 2021; pp. 1407–1411. [CrossRef]

35. Carvalho, J.; Plastino, A. On the evaluation and combination of state-of-the-art features in Twitter sentiment analysis. *Artif. Intell. Rev.* **2021**, *54*, 1887–1936. [CrossRef]

36. Bandhakavi, A.; Wiratunga, N.; Massie, S.; Deepak, P. Emotion-aware polarity lexicons for Twitter sentiment analysis. *Expert Syst.* **2021**, *38*, e12332. [CrossRef]

37. Koochari, A.; Gharahbagh, A.; Hajihashemi, V. A Persian part of speech tagging system using the long short-term memory neural network. In Proceedings of the 2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), Mashhad, Iran, 23–24 December 2020; Volume 2020. [CrossRef]

38. Zargari, H.; Hosseini, M.M.; Gharahbagh, A.A. Order-Sensitivity Sentiment dictionary of word sequences containing intensifiers. *Multimed. Tools Appl.* **2024**, *83*, 54885–54907. [CrossRef]

39. Žunić, A.; Corcoran, P.; Spasić, I. Aspect-based sentiment analysis with graph convolution over syntactic dependencies. *Artif. Intell. Med.* **2021**, *119*, 102138. [CrossRef]

40. Cambria, E.; Mao, R.; Han, S.; Liu, Q. Sentic parser: A graph-based approach to concept extraction for sentiment analysis. In Proceedings of the 2022 IEEE International Conference on Data Mining Workshops (ICDMW), Orlando, FL, USA, 28 November–1 December 2022; pp. 1–8. [CrossRef]

41. Junior, A.B.; da Silva, N.F.F.; Rosa, T.C.; Junior, C.G. Sentiment analysis with genetic programming. *Inf. Sci.* **2021**, *562*, 116–135. [CrossRef]

42. Li, Q.; Zhang, Q.; Si, L. Tweetsenti: Target-dependent tweet sentiment analysis. In Proceedings of the World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; pp. 3569–3573. [CrossRef]

43. Polignano, M.; Basile, V.; Basile, P.; Gabrieli, G.; Vassallo, M.; Bosco, C. A hybrid lexicon-based and neural approach for explainable polarity detection. *Inf. Process. Manag.* **2022**, *59*, 103058. [CrossRef]

44. Kim, B.K.; Jang, K.B. Development of Sentiment Detection combined with Deep Learning and Sentiment Dictionary. *J. Internet Things Converg.* **2023**, *9*, 21–31. [CrossRef]

45. Gupta, S.; Singh, A.; Kumar, V. Emoji, text, and sentiment polarity detection using natural language processing. *Information* **2023**, *14*, 222. [CrossRef]

46. Gopi, A.P.; Jyothi, R.N.S.; Narayana, V.L.; Sandeep, K.S. Classification of tweets data based on polarity using improved RBF kernel of SVM. *Int. J. Inf. Technol.* **2023**, *15*, 965–980. [CrossRef]

47. Tong, X.; Chen, M.; Feng, G. A Study on the Emotional Tendency of Aquatic Product Quality and Safety Texts Based on Emotional Dictionaries and Deep Learning. *Appl. Sci.* **2024**, *14*, 2119. [CrossRef]

48. Ben, T.L.; Alla, P.C.R.; Komala, G.; Mishra, K. Detecting sentiment polarities with comparative analysis of machine learning and deep learning algorithms. In Proceedings of the 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT), Gharuan, India, 5–6 May 2023; pp. 186–190. [CrossRef]

49. Raza, A.A.; Habib, A.; Ashraf, J.; Shah, B.; Moreira, F. Semantic orientation of crosslingual sentiments: Employment of lexicon and dictionaries. *IEEE Access* **2023**, *11*, 7617–7629. [CrossRef]

50. Ramos Magna, A.; Zamora, J.; Allende-Cid, H. Senti-Sequence: Learning to Represent Texts for Sentiment Polarity Classification. *Appl. Sci.* **2024**, *14*, 1033. [CrossRef]

51. Bashiri, H.; Naderi, H. LexiSNTAGMM: An unsupervised framework for sentiment classification in data from distinct domains, synergistically integrating dictionary-based and machine learning approaches. *Soc. Netw. Anal. Min.* **2024**, *14*, 102. [CrossRef]

52. Young, J.C.; Arthur, R.; Williams, H.T. CIDER: Context-sensitive polarity measurement for short-form text. *PLoS ONE* **2024**, *19*, e0299490. [CrossRef]

53. Shahade, A.K.; Walse, K.; Thakare, V.M.; Atique, M. Multi-lingual opinion mining for social media discourses: An approach using deep learning based hybrid fine-tuned smith algorithm with adam optimizer. *Int. J. Inf. Manag. Data Insights* **2023**, *3*, 100182. [CrossRef]

54. Miller, G.A. WordNet: A lexical database for English. *Commun. ACM* **1995**, *38*, 39–41. [CrossRef]

55. Hansen, P.C. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*; SIAM: Philadelphia, PA, USA, 1998.

56. Hansen, P.C. The truncated SVD as a method for regularization. *BIT Numer. Math.* **1987**, *27*, 534–553. [CrossRef]

57. Gavrilyuk, A.; Osinkin, D.; Bronin, D. On a variation of the Tikhonov regularization method for calculating the distribution function of relaxation times in impedance spectroscopy. *Electrochim. Acta* **2020**, *354*, 136683. [CrossRef]

58. Zhang, J.; Qi, H.; Jiang, D.; He, M.; Ren, Y.; Su, M.; Cai, X. Acoustic tomography of two dimensional velocity field by using meshless radial basis function and modified Tikhonov regularization method. *Measurement* **2021**, *175*, 109107. [CrossRef]

59. Jiang, J.; Tang, H.; Mohamed, M.S.; Luo, S.; Chen, J. Augmented tikhonov regularization method for dynamic load identification. *Appl. Sci.* **2020**, *10*, 6348. [CrossRef]

60. Wang, L.; Niu, J.; Song, H.; Atiquzzaman, M. SentiRelated: A cross-domain sentiment classification algorithm for short texts through sentiment related index. *J. Netw. Comput. Appl.* **2018**, *101*, 111–119. [CrossRef]

61. Taboada, M.; Anthony, C.; Voll, K.D. Methods for Creating Semantic Orientation Dictionaries. In Proceedings of the LREC, Genoa, Italy, 22–28 May 2006; pp. 427–432.

62. Stone, P.J.; Dunphy, D.C.; Smith, M.S. *The General Inquirer: A Computer Approach to Content Analysis*; MIT Press: Cambridge, MA, USA, 1966.

63. Bradley, M.M.; Lang, P.J. *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings*; Technical Report, Technical Report C-2; University of Florida: Gainesville, FL, USA, 1999.

64. Baccianella, S.; Esuli, A.; Sebastiani, F. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In Proceedings of the Lrec, Valletta, Malta, 17–23 May 2010; Volume 10, pp. 2200–2204.

65. Brooke, J. A Semantic Approach to Automated Text Sentiment Analysis. Master's Thesis, Simon Fraser University, Burnaby, BC, Canada, 2009.

66. Kiritchenko, S.; Mohammad, S.M. The effect of negators, modals, and degree adverbs on sentiment composition. *arXiv* **2017**, arXiv:1712.01794. [CrossRef]

67. Gupta, I.; Joshi, N. Feature-based twitter sentiment analysis with improved negation handling. *IEEE Trans. Comput. Soc. Syst.* **2021**, *8*, 917–927. [CrossRef]

68. Mohammad, S.M.; Kiritchenko, S.; Zhu, X. NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets. *arXiv* **2013**, arXiv:1308.6242. [CrossRef]

69. Sygkounas, E.; Rizzo, G.; Troncy, R. Sentiment polarity detection from amazon reviews: An experimental study. In *Semantic Web Evaluation Challenge*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 108–120. [CrossRef]

70. Di Rosa, E.; Durante, A. App2check extension for sentiment analysis of amazon products reviews. In Proceedings of the Semantic Web Challenges: Third SemWebEval Challenge at ESWC 2016, Heraklion, Crete, Greece, 29 May–2 June 2016; Revised Selected Papers 3, pp. 95–107. [CrossRef]

71. Petrucci, G.; Dragoni, M. The IRMUDOSA system at ESWC-2016 challenge on semantic sentiment analysis. In *Semantic Web Evaluation Challenge*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 126–140. [CrossRef]

72. Zargari, H.; Zahedi, M.; Rahimi, M. GINS: A Global intensifier-based N-Gram sentiment dictionary. *J. Intell. Fuzzy Syst.* **2021**, *40*, 11763–11776. [CrossRef]

73. Dey, A.; Jenamani, M.; Thakkar, J.J. Senti-N-Gram: An n-gram lexicon for sentiment analysis. *Expert Syst. Appl.* **2018**, *103*, 92–105. [CrossRef]

74. Carta, S.; Podda, A.S.; Recupero, D.R.; Saia, R.; Usai, G. Popularity prediction of instagram posts. *Information* **2020**, *11*, 453. [CrossRef]

75. Gharahbagh, A.A.; Abolghasemi, V. A novel accurate genetic algorithm for multivariable systems. *World Appl. Sci. J.* **2008**, *5*, 137–142.

76. Wang, Z.; Hu, Z.; Ho, S.B.; Cambria, E.; Tan, A.H. MiMuSA—Mimicking human language understanding for fine-grained multi-class sentiment analysis. *Neural Comput. Appl.* **2023**, *35*, 15907–15921. [CrossRef]

77. Punetha, N.; Jain, G. Optimizing sentiment analysis: A cognitive approach with negation handling via mathematical modelling. *Cogn. Comput.* **2024**, *16*, 624–640. [CrossRef]

78. Jalali, M.; Zahedi, M.; Basiri, A. Deterministic solution of algebraic equations in sentiment analysis. *Multimed. Tools Appl.* **2023**, *82*, 35457–35474. [CrossRef]

79. Nakov, P.; Rosenthal, S.; Kozareva, Z.; Stoyanov, V.; Ritter, A.; Wilson, T. Semantic sentiment analysis of twitter. In *Proceedings of the Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2: Seventh International Workshop on Semantic Evaluation (SemEval 2013), Atlanta, Georgia, USA, 2013*; Association for Computational Linguistics: Kerrville, TX, USA, 2013; pp. 312–320.