FACTA UNIVERSITATIS
Series: **Mechanical Engineering** Vol. 21, Nº 3, Special Issue, 2023, pp. 479 - 500
https://doi.org/10.22190/FUME230901038K

Original scientific paper

A HYBRID DEEP LEARNING APPROACH FOR SENTIMENT ANALYSIS IN PRODUCT REVIEWS

Minghui Kuang[1], Ramin Safa[2], Seyyed Ahmad Edalatpanah[3], Robert S. Keyser[4]

[1]Guangzhou City University of Technology, Guangzhou, Guangdong Province, China
[2]Department of Computer Engineering, Ayandegan Institute of Higher Education, Tonekabon, Iran
[3]Department of Applied Mathematics, Ayandegan Institute of Higher Education, Tonekabon, Iran
[4]Southern Polytechnic College of Engineering and Engineering Technology, Kennesaw State University, Marietta, Georgia, USA

Abstract. *Product reviews play a crucial role in providing valuable insights to consumers and producers. Analyzing the vast amount of data generated around a product, such as posts, comments, and views, can be challenging for business intelligence purposes. Sentiment analysis of this content helps both consumers and producers gain a better understanding of the market status, enabling them to make informed decisions. In this study, we propose a novel hybrid approach based on deep neural networks (DNNs) for sentiment analysis in product reviews, focusing on the classification of sentiments expressed. Our approach utilizes the recursive neural network (RNN) algorithm for sentiment classification. To address the imbalanced distribution of positive and negative samples in social network data, we employ a resampling technique that balances the dataset by increasing samples from the minority class and decreasing samples from the majority class. We evaluate our approach using Amazon data, comprising four product categories: clothing, cars, luxury goods, and household appliances. Experimental results demonstrate that our proposed approach performs well in sentiment analysis for product reviews, particularly in the context of digital marketing. Furthermore, the attention-based RNN algorithm outperforms the baseline RNN by approximately 5%. Notably, the study reveals consumer sentiment variations across different products, particularly in relation to appearance and price aspects.*

Key words: *Deep learning, Recursive neural network (RNN), Resampling technique, Social media marketing*

Received: September 01, 2023 / Accepted October 20, 2023
Corresponding authors: Minghui Kuang, S. A. Edalatpanah
Affiliation, postal address: Guangzhou City University of Technology, Guangzhou, Guangdong Province, China & Department of Applied Mathematics, Ayandegan Institute of Higher Education, Tonekabon, Iran.
E-mail: kuang_minghui@outlook.com & s.a.edalatpanah@aihe.ac.ir

1. INTRODUCTION

Sentiment analysis (SA), also known as opinion mining (OM), is an essential application of data mining that aims to collect and extract sentiments and opinions from various documents. With the ever-increasing amount of unstructured data available on the World Wide Web, processing and interpreting this data pose significant challenges. However, neuro-linguistic programming (NLP) methods can facilitate the analysis [1].

The influence of social media on consumer preferences, behaviors, and attitudes is profound. Monitoring social media provides an excellent opportunity to measure customer loyalty and track their sentiments toward different brands or products. However, extracting relevant content from online social networks (OSNs) is a complex task that requires a comprehensive understanding of dialectal and linguistic semantic rules. SA involves a multi-step process, including feature selection, sentiment data collection, sentiment classification, and sentiment polarity detection. Effective feature selection is crucial for understanding the main aspects and enhancing the accuracy of the system. Detecting and classifying sentiment has become particularly challenging due to the vast volume, variety, and unstructured nature of web-based social networks. Therefore, feature selection plays a critical role in SA feature analysis based on big data.

In general, there are three SA levels: (1) document level, (2) sentence level, and (3) aspect level. Document- and sentence-level SA assume that only one theme is described in a document or sentence, which is often not the case. Therefore, more detailed SA requires further research to identify related entities and aspects and classify sentiment associated with these entities and aspects [2]. Instances of entities include products, services, themes, people, organizations, or events, often with dissimilar aspects. This type of SA generally relies on machine learning (ML) techniques. An increased level of abstraction in ML techniques also improves the efficiency of natural language processing (NLP).

There are three main tasks for aspect-based sentiment analysis (ABSA): (1) opinion target extraction, (2) aspect category detection (ACD), and (3) sentiment polarity. Opinion target extraction involves extracting opinion words, i.e., entity or feature. ACD is related to identifying entities and attributes [3]. Sentiment polarity is about clarifying the sentiment polarity for aspect words.

Learning-based SA techniques are generally divided into two categories: lexicon-based and ML-based. Lexicon-based techniques require a predefined word source, such as a dictionary containing a list of words used to express emotion and predetermined emotional values. Then, based on the sentiment lexicon and the features extracted from the text, the sentences are classified according to their polarity. There is, however, a cost and time involved in creating a valid lexicon for these methods, as they don't need any pre-training.

Meanwhile, ML-based techniques, such as naive Bayesian and support vector machine (SVM), have demonstrated significant success in sentence classification and sentiment analysis [4]. ML methods are based on classifying words into sentiment-related labels. The main advantage of ML approaches is their representation of learning ability. ML algorithms require a training dataset that helps automate the classifier, and a test dataset is also used to verify the classifier's performance. Therefore, ML-based approaches are better than lexicon-based approaches for SA because of their capacity to deal with large amounts of data. In addition, extensive research has been conducted on SA using ML-based approaches [5].

Despite their somewhat acceptable performance, ML-based techniques cannot thoroughly investigate texts' structure and meaning, and disregarding a positive phrase can change the text polarity [6]. Alternatively, the extraction of features is a vital part of ML-based techniques, and the extracted features play a crucial role in classifying data. Hence, SA and deep learning (DL) techniques are integrated due to the effectiveness of DL in automated machine learning (AutoML) and the lack of dependence on manual feature extraction. DL includes several networks: convolutional neural network (CNN), deep belief network (DBN), recursive and recurrent neural network (RNN), and many others. Albeit DL-based networks outperform other SA algorithms, they still face some limitations [6].

RNN is a SA algorithm. The superiority of RNN over other neural networks (NNs) is that it outperforms in predicting structured data for variable input. In the RNN architecture, the input layer includes a word feature vector at time t. The input layer is attached to the hidden layer, containing the information history. In addition to recursive connections to the output layer, the hidden layer has recursive connections to itself. The hidden and output layers also contain neurons for storing values at time t. Recurrence permits the network to be deeper than a conventional NN [7].

This study proposes a novel NN-based approach to SA with an emphasis on digital marketing. Although simple RNN is considered a very robust tool, it suffers from some fundamental challenges, including vanishing gradient and the inability to learn in deeper layers. Therefore, this study seeks to overcome these challenges by proposing a modified RNN approach that uses DL-based techniques. The proposed approach is presented so that it can be compared with the evaluation metrics of other existing approaches. Facebook and Amazon datasets are also used to evaluate the proposed approach compared to other approaches.

Although the terms online marketing and digital marketing are prevalent, SMM is becoming more popular with the increasing use of social media platforms such as Instagram, Twitter, Facebook, and LinkedIn [8]. Several social media platforms provide companies with data analytics tools for tracking the progress, success, and engagement of ad campaigns. Companies can address a wide range of stakeholders through SMM. At the strategic level, SMM involves handling a marketing campaign and creating a company's intended social media platforms.

In data analytics, SA identifies and extracts expressed opinions or feelings. As social networks evolve, Research in computer science has become increasingly focused on SA. In particular, many scholars have focused on user behavior toward specific topics based on individuals' interaction and curiosity toward social media platforms [9]. It helps decision-makers, policy-makers, and managers judge the influence of their products, services, or policies on society. SA is a helpful technique for online marketing, especially on social networks.

Using SA in commercial products is essential from three perspectives: customer, organization, and ad agencies [10]:

1) Customer's perspective: When a person intends to purchase a product, a summary of other people's opinions can be more helpful than extensive information about this product. The customer can also conveniently compare products with summarized reviews or opinions.

2) Business organizations and producers' perspective: Organizations must improve their products. A product's design and development depend on this information as well as its marketing and evaluation. Based on customer feedback and opinions, manufacturing companies may decrease, increase, or replace their products. Therefore, you can compare the products and analyze the causes. Business organizations and producers can also utilize sentiment summarization. When the number of sentiments increases, it becomes more challenging for producers or consumers to distinguish them. By summarizing sentiments, it becomes easier for a customer to understand other people's feelings about the product, and it becomes easier for producers to know the customer's feelings about the products [11].

3) Ad agencies' perspective: Feedback and opinions are crucial for advertising agencies as they can elicit market demand ideas. People's general perspective and their favorite product categories can be extracted using opinion mining.

Hence, the practical implications are: 1) Refine marketing campaigns by analyzing sentiment patterns and adjusting messaging, visuals, and overall strategy. 2) Identify key customer segments by analyzing sentiment scores, enabling tailored targeting and promotions. 3) Improve customer targeting by identifying potential customers with positive sentiments for personalized advertising and partnerships. 4) Enhance product development by analyzing sentiment feedback to identify areas of improvement and innovation. 5) Monitor brand reputation through real-time sentiment analysis to address negative sentiments and manage PR crises. 6) Benchmark competitor performance by comparing sentiment scores and trends to capitalize on market opportunities.

A look at the relevant studies shows that DL is one of the new approaches to social network analysis (SNA) with different purposes [12]. No need for a domain expert and automated feature detection are essential features that make DL a desirable approach to SNA, especially when new classes are constantly emerging, and previous class patterns evolve. Another critical feature of DL is that it has a significantly higher learning capacity than traditional ML-based approaches, and thus, it can learn very complicated patterns. By incorporating these two features as an end-to-end (E2E) approach, DL can learn the nonlinear relationship between the input and the corresponding output without splitting the problem into smaller feature selection problems. Therefore, the following section reviews related work.

2. Related Work

DL-based approaches have recently become more popularized for sentiment classification. Artificial neural networks (ANNs), initially proposed in 1998, have been used for data recognition to classify short-term sentiment [13]. Research demonstrated that a learner's combination of different architectures, including CNN and gated recurrent units (GRUs), improves sentiment classification performance [14].

Due to its significance in business and society, SA is one of NLP's most active research areas, extending beyond computer science, such as management and social sciences. The increasing importance of SA coincides with social media growth. Most businesses and social organizations use SA systems since opinions are a fundamental component of human activity and play a critical role in influencing our actions [15]. Research in the field of marketing with social media has explored various aspects of social media marketing strategy, including content creation, platform selection, audience targeting, and engagement tactics. These studies have provided valuable insights into effective marketing practices on social media platforms [21] [22].

The book "Affective Computing and Sentiment Analysis: Emotion, metaphor, and Terminology" was published by Ahmad [15] in 2011 with 148 pages. This book relies on various research fields, including artificial intelligence (AI), especially reasoning and ML, data mining, linguistics, and psychology. Liu's book, "Sentiment Analysis and Opinion Mining"[16], was published in 2012. This book includes the results of lectures and extensive studies of data mining, web mining, and text mining. It is a comprehensive introductory book that covers all main topics from the latest relevant developments with over 400 references.

2.1. The unbalanced dataset in sentiment analysis

The issue of class imbalance is a common challenge in sentiment analysis (SA) and is prevalent in many real-world applications. Class imbalance refers to situations where the number of samples in some sentiment classes is significantly higher than in others. This poses a problem for machine learning (ML) algorithms as they tend to prioritize the majority class, leading to a bias and reduced sensitivity in detecting minority-class samples during classification training [17]. Consequently, handling classification tasks with imbalanced datasets has become one of the most complex problems in ML.

While various algorithms have been proposed to address imbalanced data distribution in structured datasets, the issue of class imbalance has not been extensively explored in the domain of text classification. However, a seminal study by Li et al. [18], among the first to investigate imbalanced sentiment classification, demonstrated the adverse effect of imbalanced sample distributions on classifier performance. Resampling techniques have been widely adopted to mitigate this effect. These techniques involve manipulating the class distribution by either under-sampling the majority class or oversampling the minority class. Previous research has suggested that oversampling approaches tend to outperform under-sampling methods [7]. One popular oversampling technique is the synthetic minority over-sampling technique (SMOTE) [8], which generates synthetic data points based on the similarity between minority samples in the numerical feature space. However, when it comes to text data, oversampling may not yield semantically comparable terms as texts are composed of words or characters. Therefore, there is a need to develop an oversampling approach that can generate synthetic texts directly from the contextual space.

Research conducted in the field of marketing with social media has provided valuable insights into effective marketing practices on social media platforms. Leveraging social media platforms allows marketers to establish connections with their target audience, enhance brand awareness, and drive customer engagement [21] [22]. Gaining a deep understanding of the dynamics of social media marketing can contribute to the development of more effective sentiment classification models and techniques, enabling marketers to better analyze and interpret customer sentiments and opinions.

3. PROPOSED APPROACH

This section proposes a DL-based approach to SA in digital marketing. This approach analyzes users' opinions about different product categories using an appropriate technique to detect user sentiment. Their views and ratings also identify trends for the product categories we want to market. Figures 1 and 2 illustrate a flowchart for the proposed approach, which comprises two main stages: (a) model generation and (b) product marketing. The first stage generates an RNN-based model using Amazon data. This stage consists of four sub-stages: (1) preprocessing, (2) resampling, (3) input vector generation using Word2Vec, and (4) RNN-based sentiment model generation. The second stage applies Facebook data to the generated model and performs product marketing based on different aspects. This stage also includes four sub-stages: (1) preprocessing using Stanford CoreNLP, (2) input vector generation using Word2Vec, (3) applying Facebook data to the model, and (4) marketing the intended product using rating and cosine similarity.

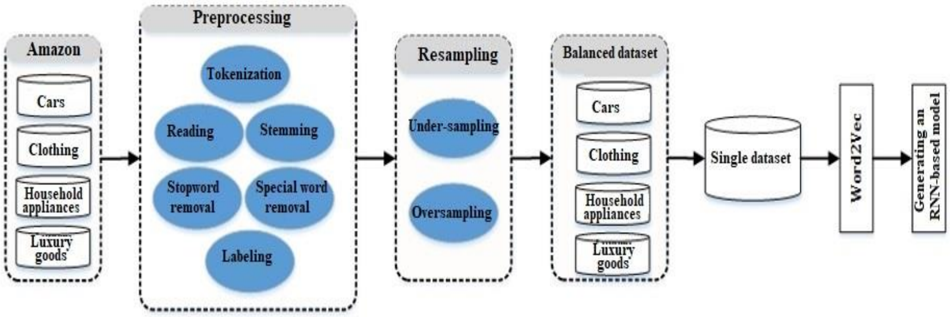


Fig. 1 Model generation

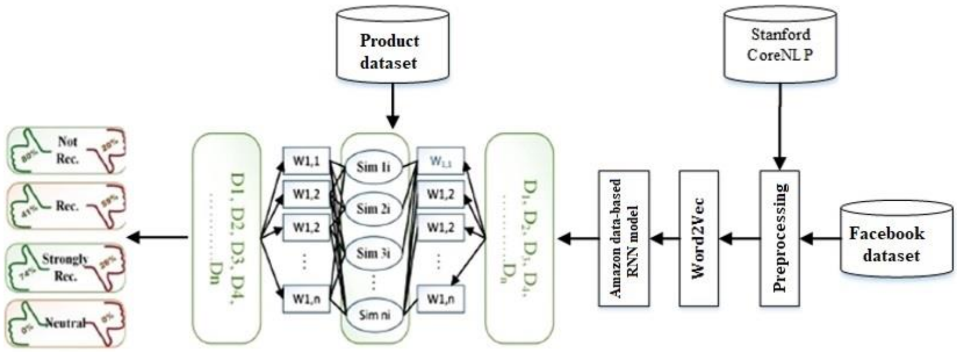


Fig. 2 Product marketing

3.1. Datasets: Amazon data

Amazon is one of the largest e-commerce websites with countless reviews. In the research analysis section, we use Amazon data prepared by Amazon Web Services. Since the dataset is unlabeled, we need to label the data to use it in a supervised learning model. We used four JSON files whose data structure is presented in Table 1.

Table 1 Review text features

Description	Feature Title
Reviewer identifier	reviewerID
Product identifier	asin
Name of the reviewer	reviewerName
Review helpful vote	vote
Text of the review	reviewText
Rating of the product	overall
Summary of the review	summary
Time of the review (UNIX time)	unixReviewTime
Time of the review (raw)	ReviewTime

"Text of the review" (reviewText) and the "rating of the product" (overall) are the primary data features used for sentiment classification. In total, four files include about fifteen million reviews. The distribution of Amazon data collected for four product categories, including cars, household appliances, jewelry and clothing, and luxury goods, is 1,711,519, 2,070,831, 11,285,464, and 34,278, respectively. As can be seen, the number of samples for jewelry and clothing is much higher than in other categories (75%). On the other hand, the distribution rates for household appliances, cars, and luxury goods are 11.4%, 13.4%, and 0.002, respectively.

3.2. Facebook data

Facebook post reactions dataset is also used for analysis. This dataset comprises Facebook posts on the customer service pages of twelve major US and UK chain stores/retail stores, i.e., Amazon, Tesco, Sainsbury, Walmart, Safeway, Aldi UK, Target, The Home Depot, Walgreens, Best Buy, Macy's, and Publix. Most of the posts are released by the customers of these chain stores. Besides, we obtained the written text of the posts, the Facebook reaction matrix, and posts made by other users with comments attached. Posting a reaction on a reply feature was introduced in 2017, and these reactions belong only to the initial post and not to any subsequent replies. Therefore, it would lead to generating an incomplete or tiny dataset. Reactions encompass "love", "wow", "haha", "sad", and "angry". This dataset contains more than 70,000 posts. Of these, 25,969 posts received at least one reaction. The full dataset is available; Out of 70,000 posts, 45,386 posts are negative (65%), and 24,614 posts are positive (35%).

3.3. Preprocessing

Preprocessing is an essential step in natural language processing (NLP) tasks, including sentiment analysis and text categorization. It involves transforming raw text data into a format that is suitable for analysis and machine learning algorithms. In the preprocessing phase, six operations are performed to prepare the text data for further analysis:

File reading: Each product category has a separate file that is read individually to extract the text data. This ensures that the data is organized and processed systematically.

Text tokenization: Tokenization is the process of extracting individual words from the text. It serves as the foundation for subsequent analysis, as it breaks down the text into its constituent units or tokens. Each word is treated as a separate token.

Stop-word removal: Stop words are commonly occurring words that do not carry significant meaning, such as articles, prepositions, and conjunctions. In addition to these standard stop words, certain punctuation marks and text separator characters are also considered stop words and removed from the text. This step helps eliminate noise and reduce the dimensionality of the data.

Selective word removal: This step involves removing specific types of words that are not relevant to the analysis. Words like links, stickers, and conjunctions such as "and" and "or" are removed, as they do not contribute to the sentiment or categorization tasks.

Stemming: Stemming is a process that reduces the dimensionality of the input features by converting words to their base or root form. It groups together variations of the same word, making it easier to detect common features and their meanings in the text. The Porter stemmer algorithm is used in this case to find word stems. Only sentiment words are considered for stemming, and texts without this category of features are removed. For example, words like "wait," "waiting," or "waited" are transformed into their root form "wait."

Text labeling: In the case of the Amazon data, text samples are labeled as positive or negative based on their star ratings. Texts with ratings of 4 and 5 stars are labeled as positive, indicating favorable sentiment. Texts with ratings lower than three stars are labeled as negative, indicating unfavorable sentiment. Texts with three-star ratings are considered neutral and are excluded from further analysis. This step ensures the data is labeled for supervised learning tasks.

For Facebook data, a different preprocessing approach is used. The Stanford CoreNLP parser is applied, which involves the following steps: 1) Transforming all words to lowercase letters for uniformity in the data.2) Substituting URLs with the generic token "URL" to remove specific URLs and treat them as a single entity. 3) Substituting user links with the generic token "AT_USER" to anonymize specific user information and treat them consistently. 4) Eliminating the hash symbol (#) from hashtag references, treating them as regular words. 5) Substituting three or more occurrences of the same character in a word with a single occurrence. For example, "looooove" becomes "love" to reduce repeated characters. 6) Eliminating sequences containing numbers, as they are not relevant for sentiment analysis.

By applying these preprocessing steps, the text data is transformed into a suitable format for sentiment analysis and text categorization tasks, enabling more accurate and efficient analysis of the data

3.4. Resampling

Resampling is a widely used technique to tackle the challenge of imbalanced data distribution in various domains. While oversampling approaches have shown their effectiveness in mitigating the negative impact of imbalanced data, their applicability to imbalanced contextual data remains uncertain. Traditional oversampling methods often rely on numerical computations to measure similarities between items. However, when dealing with textual data, which consists of words and characters, numerical approaches may fail to generate semantically consistent new items. Therefore, alternative methods are needed to address the oversampling of imbalanced contextual data. Researchers have identified limitations in traditional numerical oversampling methods when applied to textual data. For instance, it has been observed that numerical oversampling may struggle to capture the intricate semantic relationships present in texts. The generation of synthetic samples solely based on numerical computations can result in synthetic data that lacks semantic coherence. These limitations underscore the necessity of exploring alternative approaches that can generate synthetic texts instead of relying solely on numerical vectors, thereby enabling effective oversampling of imbalanced contextual data.

Although oversampling approaches have proven their ability to mitigate the detrimental effect of imbalanced data distribution in multiple domains, their efficacy in dealing with imbalanced contextual data is ambiguous. In conventional sampling methods, similarities between items are measured numerically. However, because texts contain words/characters, numerical computations may not produce semantically consistent new items; hence, we generate synthetic texts instead of numerical vectors to oversample imbalanced contextual data. These studies emphasize the importance of considering the unique characteristics of textual data when designing oversampling methods. By generating synthetic texts instead of relying solely on numerical computations, these approaches aim to create semantically consistent oversampled data. This, in turn, enhances the performance of imbalanced contextual data analysis tasks. In the following section, we will present our novel oversampling method, specifically tailored to address the challenge of imbalanced contextual data. Our method focuses on generating synthetic texts that maintain semantic coherence and context, thereby improving the analysis of imbalanced contextual data.

The proposed method increases minority-class texts with two strategies: inversion and imitation. During the inversion process, sentimental words in majority-class texts are converted into their opposites with sentiment polarity to generate synthetic texts. Despite grammatically incorrect synthetic texts, they must at least be semantically interpretable.

During inversion and imitation, we perform the procedures iteratively. To simplify the description, we assume the negative class is the minority class. Our first step is to train a classifier based on the original training set. Then, a probabilistic prediction is made directly from the training set using the generated classifier. Finally, synthetic texts are constructed from positive and negative samples with high and low predicted confidence, respectively, utilizing inversion and imitation methods.

We use the label propagation approach to find the counterpart lexicon. We first create a weighted undirected graph for a text with n words, where each node represents a word, and each edge is weighted according to semantic similarity. Next, the seed words with robust sentimental polarities are selected, and each word is scored 1. Then, each word should be scored between [0,1], where 0 denotes negative words, and 1 represents positive words. The first step constructs a vector for each word (using the technique described in Section 3.6).

A weighted lexicon graph is formed based on a collection of word vectors, with each node in the graph representing a word. The semantic similarity between two words computes the weight. This study uses cosine similarity. The words w_i and w_j are connected if their cosine similarity exceeds a threshold value δ.

$$E_{ij} = \begin{cases} {w_i}^T w_j & \text{if} \quad \frac{{w_i}^T w_j}{\|w_i\|\|w_j\|} \geq \delta \\ 0 & \text{if} \quad \frac{{w_i}^T w_j}{\|w_i\|\|w_j\|} < \delta \end{cases} \tag{1}$$

A set of seed words is selected, and their sentimental polarity is labeled after the lexicon graph is constructed. Since the seed words function as labels, they must have a positive or negative sentimental polarity. This study uses the words lexicon proposed in [22] (Table 2).

Table 2 Seed words

Positive seed words	Negative seed words
Excellent, Good, Pleasant, Love, Delightful, Perfect, Happy, Fortunate, Lovely, Awesome, Amazing, Fantastic, Nice, Successful, Best	Sad, Bad, Evil, Unfortunate, Horrible, Unhappy, Nasty, Hate, Unpleasant, Disgusting, Terrible, Awful, Poor, Failure, Worst

The proposed oversampling approach is identical to boosting the classifier because the reweighting mechanism in promoting the classifier also highlights the significance of misclassified samples. Since oversampling increases the training samples and model complexity, we remove the majority-class samples using the inversion technique. Samples from the majority class for which a counterpart sentence from the minority class was formed are removed from the dataset. In other words, we use oversampling and undersampling techniques simultaneously.

3.5. Word embedding vector

Suppose we have a sentence $s = \{w_1, w_2, \dots, w_i, \dots, w_n\}$ consisting of n words and the word aspect w_i in sentence s. The aspect-level sentiment classification focuses on the sentiment polarity of the sentence s for the word w_i. For example, the sentiment polarity of the sentence "great food, but the service was dreadful!" is positive for "food," while it is negative for "service." When facing a text, we place words in a continuous, low-dimensional, and real-valued vector called word embedding.

There is a need for vector representation of words to train DNNs aiming to understand the input words deeply. Hence, we use the Word2Vec model [13], a two-layer shallow neural network, in the model's first layer. The Word2Vec model has two versions: a continuous bag of words (CBOW) and a skip-gram (SG). According to previous studies indicating its effectiveness in SA, the proposed approach uses an SG-based model to generate input word vectors. We selected this algorithm in the first layer because of its proven performance in representing semantic words without human involvement. A vector representation of a word is learned based on the context in which it appears.

3.6. RNN-based sentiment classification

DL's data representation learning method is based on neural network algorithms, which use multiple layers of modules to analyze and classify inputs, with output from one layer fed into the next. The gradient is computed using an activation function in a process known as backpropagation. The input data (image, audio, or text) generates numeric vectors. Since each layer corresponds to a higher level of abstraction, DL implies a hierarchy of more straightforward concepts. A recurrent neural network (RNN) is a popular SA algorithm. In baseline RNNs, tokens are fed into recurrent units in a vector of fixed length (e.g., a sentence or document). This method can capture the inherent sequentially of language, i.e., words extend their meaning through their preceding words. RNN models differ from CNN models in that the RNN output is based on prior calculations instead of previously computed data. As a result, it captures text-to-speech dependencies and can model different text lengths.

In ABSA, sentiment classification aims to identify individuals' sentiments according to the words in the text using the interaction among the words and the entire sentence for the model. Typical encoder-decoder architectures, such as RNNs, may have the problem of encoding irrelevant data, mainly when the input contains much information. The attention mechanism can be used to focus on the part of the original text that is of interest to the model. In this mechanism, attention weights are computed from the lower level, and the weighted vectors are aggregated to represent the higher level.

Some neural models, such as long short-term memory (LSTM), capture contextual data implicitly and cannot explicitly represent primary aspect tokens. Only a subset of context words is required to infer sentiment towards an aspect. For instance, in the sentence "great food but the service was dreadful!", "dreadful" is an essential clue to "service," but "great" is not necessary. The standard LSTM algorithm operates sequentially and manipulates each context word using the same function, i.e., it cannot explicitly determine the significance of each context word. As part of an optimal solution, context words should be explicitly identified and used in the generation of features after an aspect-based word.

The approach proposed by Wang et al. [14] addresses the challenge of RNN-based SA. This approach comprises multiple computational layers with standard metrics. Layers are content- and location-based models that learn the importance/weight of context words and then compute continuous text representations based on these data. The text representation is viewed as a sentiment classification feature in the last layer. In gradient descent, which uses the cross-entropy error of sentiment classification as a loss function, the entire model can be efficiently trained using differentiable components. The presented approach uses the Amazon dataset to generate an SA model and the Facebook dataset to test the model and identify the social network users' sentiments for a marketing issue. After developing the model using the Amazon dataset, the Facebook data preprocessed and prepared to enter the network using the Word2Vec technique is fed into the generated model.

3.7. Product marketing

Social media is a collective term for communication applications and websites facilitating information interaction, generation, exchange, and dissemination among companies and their customer communities [23, 24]. On the other hand, Sajid [25] suggests that social media is an Internet-based source for sharing and discussing details among people. Finally, Delerue et al. [26] define social media as a group of Internet-based applications that allow the creation and exchange of user-generated content based on the technological foundations of Web 2.0.

Although social marketing is developing compared to commercial marketing, its role in social development has received intense public attention [24]. Most senior corporate managers consider developing, implementing, and revising competitive marketing strategies and policies their primary challenge. Therefore, companies' demand for high-rank marketing managers is increasing. Marketing should not be confused with its classical meaning, i.e., sales. Marketing, in a new sense, means meeting customer needs. Online reviews of business products and services allow users to share their viewpoints about products purchased or services received, which is of great value to other buyers. This trend has provided new contexts for companies that want to strengthen their presence in the market by investing in digital marketing and gaining a suitable position.

Due to the increasing attraction in SMM, many commercial tools have recently been generated to help organizations analyze the blogosphere on a large scale to extract data related to opinions about their products. However, many existing tools and research efforts are confined to assessing polarity or classifying sentiments based on a minimal set of sentiments [27-29].

Marketing aims to provide products based on individuals' attitudes and interests. These attitudes are expressed on social media platforms and selling sites and are an outstanding source for comment marketing. SA and detecting positive and negative opinions about a product help companies and marketers market their target items and sell them through appropriate offerings in the minimum possible time based on people's interests. After sentiment classification and SA processes, the proposed approach determines the number of context sentences in which the desired sentiment is expressed. Then, it calculates the average sentiment expressing the desired aspect using Eq. (2):

$$\mathrm{score}(a, R_i) = \left(\sum_{i=1}^{S_r} \frac{\mathrm{pos}(a, R_i)}{\mathrm{pos}(a, R_i) + \mathrm{neg}(a, R_i)}\right) + \mathrm{IDF}(a, R_i) \tag{2}$$

where R_i represents a set of reviews for product i, and S_{R_i} denotes the number of sentences in context r. The functions $\mathrm{pos}(a, R_i)$ and $\mathrm{neg}(a, R_i)$ indicate the positivity and negativity of the aspect in the review dataset. These values are the outputs of the sentiment classification. $\mathrm{IDF}(a, R_i)$ shows the number of occurrences of the aspect in the total review dataset. It is one of the functions of the term frequency-inverse document frequency (TF-IDF) technique, which is used to retrieve contextual information. Its value is determined using Eq. (3):

$$\mathrm{IDF}_i = \log \frac{N}{n_i} \tag{3}$$

where N indicates the total number of texts, and n_i denotes the number of texts containing the intended word.

Every product we want to market has certain features, including price, quality, size, and others. These data are contained in a secondary dataset representing a company's products. We intend to identify people's sentiments towards different features and aspects of a product based on the intended product. The most critical issue is the negative sentiments towards the products in the review texts. We aim to market these products based on negative and positive sentiments towards different tools. In this step, we use cosine similarity to match negative review text and product-related text. We compute the cosine similarity between two texts as Eq. (4):

$$\mathrm{Score}\ (d1, d2) = \mathrm{Cos}\ (W_{d1}, W_{d1}) = \frac{W_{d1} . W_{d2}}{\|W_{d1}\|_2 . \|W_{d2}\|_2} = \frac{\sum_{i=1}^{k} W_{i,d1} . W_{i,d2}}{\sqrt{\sum_{i=1}^{k} {W_{i,d1}}^2} \sqrt{\sum_{i=1}^{k} {W_{i,d1}}^2}} \tag{4}$$

In summary, cosine similarity is a measure that quantifies the similarity between two vectors, commonly used in text analysis tasks. It is based on the cosine of the angle between the vectors and provides a measure of their directional similarity, regardless of their magnitudes.

4. EVALUATION RESULTS

Python programming language was used to implement the proposed approach that facilitates the design and implementation of ML and DL algorithms. Anaconda was used for the Python programming environment. It is a free and open-source distribution of the Python and R programming languages for scientific computing (e.g., data science, ML applications, big data processing, predictive analytics, and others) to simplify package management and deployment. More than 15 million users use the Anaconda distribution. It includes over 1,500 popular data science packages. This section compares the performance of the proposed approach, which uses the attention-based RNN algorithm, with the baseline RNN algorithm.

4.1. Evaluation metrics

The performance evaluation metrics of the ML algorithm are based on the confusion matrix (Table 3). Therefore, the supervised ML algorithm uses the confusion matrix to evaluate its performance. Regarding classification, the terms "true positive (TP)," "false positive (FP)," "true negative (TN)," and "false negative (FN)" were used to compare the class labels. TP represents the number of samples whose true class is positive, and the classification algorithm recognizes their class as truly positive. FP indicates the number of samples whose true class is negative, and the classification algorithm recognizes their class as falsely positive. FN represents the number of samples whose true class is positive, and the classification algorithm recognizes their class as falsely negative. TP denotes the number of samples whose true class is negative, and the classification algorithm recognizes their class as truly negative.

Table 3 Confusion matrix for performance evaluation

		Actual Values	
		Positive	Negative
Prediction	Positive	TP	FP
	Negative	FN	TN

Model accuracy is the primary performance evaluation metric of a sentiment classification algorithm in academic research. This metric measures the overall accuracy of the classifier and is widely used for assessing the performance of classification algorithms. It indicates the percentage of correctly classified instances in the test dataset. The classification accuracy is calculated using Eq. (5), where TP and TN represent the true positives and true negatives, respectively. P is the number of positives in the experimental samples, and N is the number of negatives in the experimental samples. Maximizing TP and TN is crucial in binary or two-class classification problems.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{P+N}} \tag{5}$$

Recall and precision are two additional performance evaluation metrics commonly used in sentiment analysis research. Recall, calculated as the true positive rate divided by the sum of the true positive rate and false negative rate, reflects the model's ability to identify relevant data points in a dataset correctly. Precision, on the other hand, is calculated by dividing true positives by the sum of true positives and false positives. Precision indicates the proportion of data points classified as relevant by the model that are indeed relevant in the dataset. Precision decreases as recall increases.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{6}$$

In certain scenarios, researchers may prioritize maximizing recall or precision based on their specific objectives. The F1 score or F-measure is often employed as a combined metric to strike a balance between precision and recall. Equation (7) represents the combination of recall and precision to calculate the F1 score.

$$\text{F1 Score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (7)$$

In the study, an attention-based RNN algorithm was utilized with five hops. Tables 4 and 5 present the weights assigned to different aspects (e.g., quality, price, weight, size, service, design, and appearance) in each hop of the algorithm for the Facebook and Amazon datasets, respectively. The Amazon dataset was divided, with approximately 90% of samples used for training purposes and 10% for testing.

Table 4 Weights of different aspects of the Facebook dataset

Aspect	Hop 1	Hop 2	Hop 3	Hop 4	Hop 5
Quality	0.2	0.15	0.14	0.13	0.23
Price	0.2	0.1	0.1	0.12	0.13
Weight	0.11	0.07	0.08	0.12	0.16
Size	0.07	0.08	0.12	0.17	0.078
Service	0.03	0.07	0.08	0.12	0.06
Design	0.08	0.07	0.08	0.12	0.06
Appearance	0.2	0.45	0.45	0.28	0.4

Table 5 Weights of different aspects of the Amazon dataset

Aspect	Hop 1	Hop 2	Hop 3	Hop 4	Hop 5
Quality	0.22	0.15	0.14	0.13	0.23
Price	0.21	0.11	0.1	0.11	0.12
Weight	0.03	0.11	0.08	0.11	0.06
Size	0.1	0.08	0.12	0.13	0.11
Service	0.11	0.11	0.08	0.11	0.06
Design	0.04	0.11	0.08	0.11	0.06
Appearance	0.22	0.32	0.45	0.32	0.43

Table 4 shows that "quality," "price," and "appearance" play the same role in the first hop. "Appearance" weight increases after the second hop. Finally, this model predicts a truly negative polarity for Facebook data. It represents the effect of the multi-hop steps. However, as shown in Table 5, the model predicts polarity for "appearance" and "price" in the first hop and then "appearance" in the subsequent hops for the Amazon data.

Fig. 3 illustrates the detection accuracy of different aspects of the attention-based RNN model for the Facebook dataset. All of these aspects of classification accuracy can be consistently improved by using multiple computational layers. This model works relatively well for all detections. The detection accuracy for "appearance" is higher than other aspects, which varies between 74% for the first and 90% for the ninth hop. On the other hand, the "design" has minimum detection accuracy.

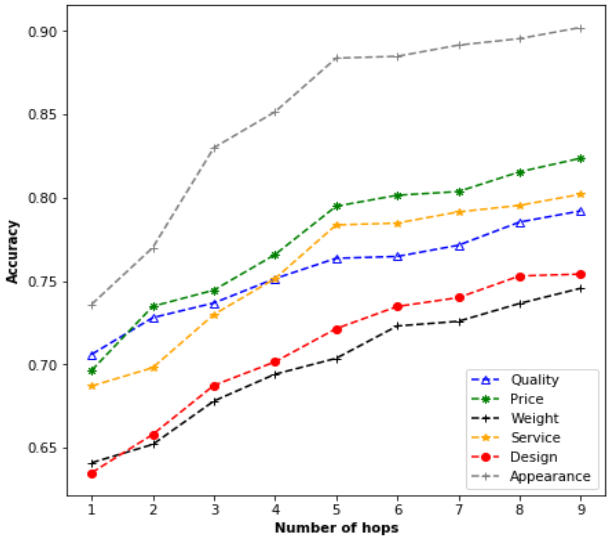


Fig. 3 The detection accuracy of aspects with different numbers of hops for attention-based RNN algorithm

This section evaluates the performance of the proposed approach compared to the baseline RNN-based approach for two-class classification between positives and negatives. Fig. 4 illustrates the confusion matrix for the approaches being assessed. Figs. 4 (a) and (b) show the confusion matrix for attention-based RNN and baseline RNN approaches, respectively.

Therefore, the attention-based RNN algorithm can detect samples with an average accuracy of 97.5% and 96% for the baseline RNN.

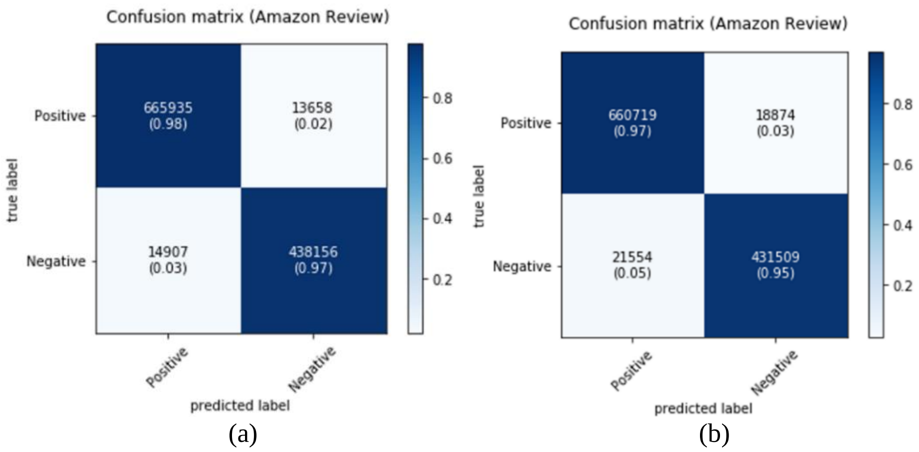


Fig. 4 Confusion matrix for SA on Amazon dataset: (a) Attention-based RNN algorithm; (b) Baseline RNN algorithm

Fig. 5 illustrates the results obtained using the confusion matrix in Fig. 4 according to four metrics: accuracy, recall, precision, F-measure, or F1 score. As shown in Fig. 5, the proposed algorithm outperforms the baseline RNN algorithm. This outperformance is noticeable for accuracy because the higher error rate for minority-class samples increases it compared to the RNN algorithm. The classification accuracy of the proposed approach is 97.5% versus 94.4% for the baseline RNN approach.

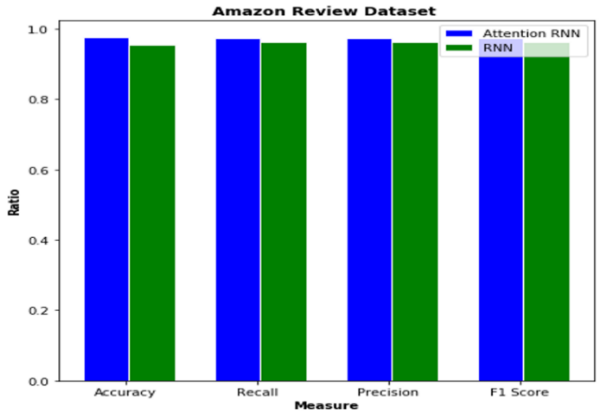
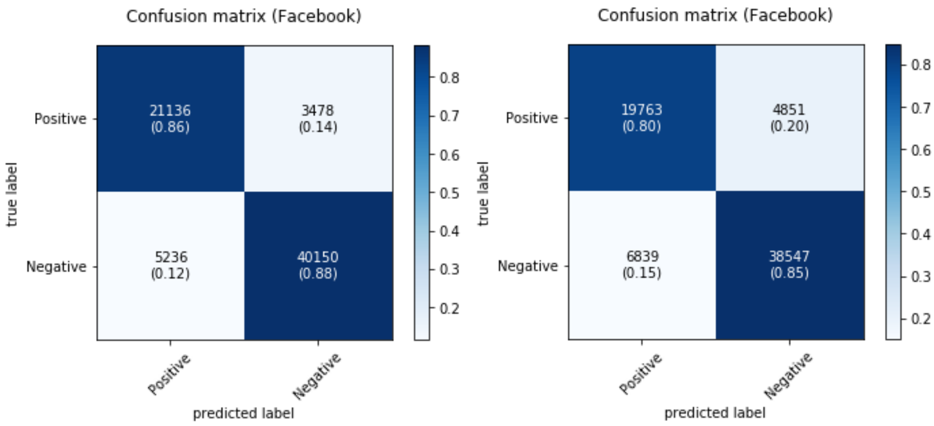


Fig. 5 The performance of attention-based RNN and baseline RNN on the Amazon dataset

The confusion matrix presented in Fig. 6 compares the performance of attention-based RNN and baseline RNN algorithms on Facebook data. It is important to note that these models were initially trained on Amazon data. Due to the differences in the data distribution between Amazon and Facebook, the sentiment classification performance of the models on Facebook data is comparatively lower. The Facebook dataset used in this evaluation consists of Facebook posts pertaining to various products, but it lacks the ability to differentiate between different product categories.



(a) (b)

Fig. 6 Confusion matrix for SA on Facebook dataset: (a) Attention-based RNN algorithm; (b) Baseline RNN algorithm

Despite these challenges, the attention-based RNN achieves an average classification accuracy of approximately 87%, while the baseline RNN achieves around 82.5%. This means that the attention-based RNN performs better than the baseline RNN on the Facebook data. The attention-based RNN leverages the mechanism of attention, which allows it to focus on relevant parts of the input during the classification process. This ability to attend to important information contributes to its higher accuracy compared to the baseline RNN.

Fig. 7 presents additional metrics derived from the confusion matrix shown in Fig. 6. The results reveal that the attention-based RNN achieves an average performance of 84%, while the baseline RNN achieves approximately 79%. Consequently, the proposed attention-based algorithm outperforms the baseline RNN by a margin of about 5%. These metrics provide further evidence of the superior performance of the attention-based RNN in sentiment classification tasks on Facebook data.

However, it is worth noting that both algorithms exhibit lower rates in comparison to the other metrics presented. Specifically, the attention-based RNN achieves a rate of 80%, and the baseline RNN achieves a rate of 74%. These rates likely correspond to specific subsets or categories within the Facebook data that pose challenges for accurate sentiment classification. Identifying and addressing the factors contributing to these lower rates could be a potential area for future research and improvement of the models' performance.

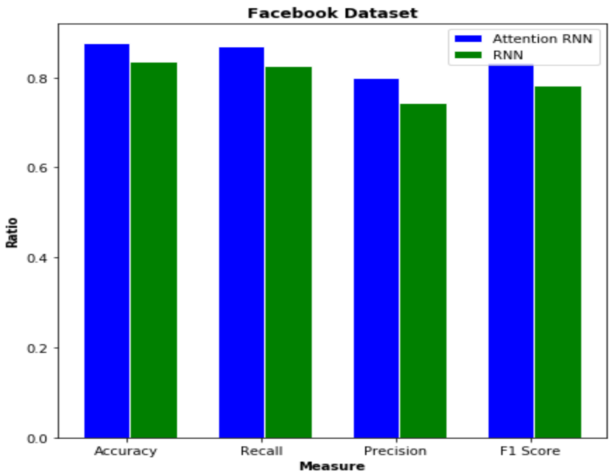


Fig. 7 The performance of attention-based RNN and baseline RNN on the Facebook dataset

Table 6 reports the weight of various aspects with different product categories. The results indicate that the product's "appearance" has the maximum weight for both algorithms. "Quality" and "price" get the subsequent ratings. These weights for cars and luxury goods are more than other product categories, indicating people's sentiment and attention to these aspects. In addition, quality ranks lower than appearance and price for luxury goods. Another significant aspect is the "service" for household appliances, representing one of the most important aspects for consumers.

Table 7 presents the rating of different product categories based on various marketing aspects. These values are the average weights calculated for the goods listed for marketing. Each product is evaluated using Eq. (2). Higher ratings indicate the importance of the aspect and people's sentiment towards the intended outcome. Since the positive sentiments towards the aspect are in the numerator, the higher the number of positive sentiments, the larger the fraction's output. The second part of Eq. (2), related to IDF, indicates that the more the desired aspect is repeated in more posts, the higher the importance of that aspect and (both positive and negative) sentiments towards it.

5. LIMITATIONS AND ETHICAL CONSIDERATIONS

While the proposed approach for sentiment analysis in social media marketing demonstrates promising results, there are several limitations that should be acknowledged.

Firstly, the evaluation of the proposed approach relies primarily on Amazon data, encompassing specific product categories. Although the results indicate its effectiveness within this context, the generalizability of the approach to other domains and datasets remains to be explored. Future research should consider evaluating the approach on diverse datasets to assess its performance in different social media marketing scenarios. Secondly, the study addresses the issue of imbalanced data distribution by employing a resampling technique. However, it is important to note that the effectiveness of this approach may vary depending on the characteristics of the dataset. The resampling technique used in this study may not yield optimal results in scenarios with significantly imbalanced classes or when dealing with extremely large datasets. Alternative or more sophisticated techniques for handling imbalanced data should be explored in future work. Thirdly, the proposed approach utilizes deep neural networks, which often require significant computational resources, particularly for large-scale datasets. In this study, the computational time required for sentiment analysis using the neural network algorithm was manageable. However, it is important to consider the scalability and computational efficiency of the approach when applying it to big data scenarios involving datasets of terabytes or more. Future research should explore strategies to optimize the computational requirements of the approach, such as parallel computing or distributed processing.

Table 6 Weights of different aspects for attention-based RNN and baseline RNN on the Facebook dataset

Aspect	Algorithm	Cars	Jewelry and clothing	Household appliances	Luxury goods
Quality	Attention-based RNN	0.19	0.15	0.14	0.13
	RNN	0.27	0.12	0.17	0.06
Price	Attention-based RNN	0.15	0.2	0.1	0.12
	RNN	0.13	0.16	0.18	0.16
Weight	Attention-based RNN	0.11	0.07	0.08	0.12
	RNN	0.07	0.07	0.11	0.17
Size	Attention-based RNN	0.11	0.1	0.08	0.1
	RNN	0.14	0.09	0.06	0.18
Service	Attention-based RNN	0.08	0.07	0.2	0.07
	RNN	0.11	0.12	0.16	0.05
Design	Attention-based RNN	0.09	0.07	0.08	0.12
	RNN	0.12	0.14	0.06	0.18
Appearance	Attention-based RNN	0.32	0.18	0.18	0.38
	RNN	0.24	0.21	0.22	0.27

Table 7 Average product ratings by different aspects

Aspect	Cars	Jewelry and clothing	Household appliances	Luxury goods
Quality	1.45	0.475	1.54	1.33
Price	1.15	1.2	1.47	0.86
Weight	0.785	0.688	1.08	0.56
Size	0.48	0.63	0.47	0.68
Service	1.08	0.78	1.2	1.035
Design	1.09	0.68	1.08	1.12
Appearance	1.14	1.18	1.04	1.38

Furthermore, the study incorporates inversion and imitation techniques to generate synthetic samples for data balancing. While these techniques provide a means to address class imbalance, they may introduce certain limitations in terms of semantic coherence. The inversion of majority-class samples and the replacement of words in minority-class samples may result in synthetic data that lacks complete grammatical or semantic consistency. Future research should explore approaches that can generate synthetic samples while ensuring greater semantic coherency. Lastly, the evaluation of the proposed approach primarily focuses on sentiment analysis of Amazon data. Although sentiment analysis on Facebook data is briefly mentioned, further exploration and evaluation of the approach on other social media platforms, such as Twitter or Instagram, would provide a more comprehensive understanding of its effectiveness across different social media contexts.

In addition to the mentioned limitations, it is important to address the ethical implications and potential biases associated with sentiment analysis in social media marketing. These considerations should be discussed alongside the limitations to provide a comprehensive understanding of the challenges and ethical concerns in the proposed approach.

Sentiment analysis in social media marketing can have ethical implications and biases that need to be considered. The analysis of user-generated content on social media platforms raises concerns regarding privacy, consent, and data usage. It is crucial to handle user data responsibly, ensuring compliance with relevant regulations and obtaining appropriate consent for data collection and analysis. Transparency in data usage and providing clear explanations of the purpose and potential impact of sentiment analysis can help address ethical concerns and build trust with users. Furthermore, biases can emerge in sentiment analysis due to various factors. Cultural, demographic, and socioeconomic biases inherent in social media platforms can influence the sentiments expressed by users and introduce skewed results. The representativeness of the data used for sentiment analysis is crucial in minimizing biases. Efforts should be made to include diverse data sources that capture a wide range of perspectives. However, it is important to acknowledge that complete representativeness may not always be achievable.

Steps should be taken to ensure fairness and impartiality in sentiment analysis. The proposed approach can incorporate techniques such as data augmentation and resampling to address imbalances and increase inclusivity. However, it is essential to continuously monitor and evaluate the sentiment analysis model to identify and rectify any biases that may arise. Involving diverse and representative annotation teams can also help mitigate potential biases during the data labeling process. Additionally, it is important to recognize that biases may still persist in sentiment analysis results due to the inherent biases present in the data itself. Interpreting the analysis results with caution and considering the context and limitations of the data sources is essential in avoiding biased conclusions.

By acknowledging these ethical implications and biases, researchers and practitioners can work towards developing more robust and fair sentiment analysis approaches in social media marketing.

6. CONCLUSIONS

This study developed a social media marketing approach based on neural networks. Therefore, an improved deep learning algorithm, called attention-based RNN, was used to classify sentiments into positive and negative categories. We used data from Amazon's website, including four product categories: cars, household appliances, luxury goods, and clothing and jewelry. An issue with this data that affects the sentiment classification performance is that the data has an imbalanced distribution between different classes and positive and negative samples. Therefore, it influences sentiment classification performance and increases the rate of misclassifications. We addressed this issue using a resampling strategy for data balancing, where minority-class samples increase and majority-class samples decrease. In addition, inversion and imitation techniques were used to generate new samples. The inversion technique inverts some majority-class samples and converts them into minority classes.

In contrast, the imitation technique increases minority-class samples by replacing some words. Alternative words for inversion or imitation were selected using the weighted lexicon graph technique. The method computes the semantic similarity between words and finds their antonyms or synonyms based on several seed words. The Facebook dataset was used for social media-based sentiment analysis after generating a neural network model using the Amazon dataset. Then, these data were applied to the model generated using Amazon data. Finally, the products were evaluated and rated based on various aspects such as "price," "quality," "appearance," and others. These ratings express people's sentiments on social media towards different aspects of a product. Hence, sellers and marketers can use these analytics for marketing their products based on their intended characteristics.

In addition, we performed different experiments to evaluate the proposed model. These experiments modeled the RNN-based algorithm using five hops. First, sentiment classification performance was assessed for both Facebook and Amazon datasets. The results also showed that sentiment analysis for Amazon data outperforms Facebook data. The neural network was modeled based on Amazon data with different product categories. However, the classifier's performance was acceptable for Facebook data (about 85%). The results of applying Facebook data to find sentimental aspects for different product categories demonstrated that "appearance" has the highest weight, and "price" and "quality" are in the subsequent ratings.

The amount of network data is increasing rapidly due to the high activity on social networking platforms. Therefore, it is suggested to use big data management techniques, including master-slave database architecture and CPU/GPU implementations. However, experiments showed that using neural network algorithms for contextual data requires high computing time and is not suitable for big data of terabytes and above. Therefore, it is suggested to use the proposed approach for other social media data, such as Twitter, for future researchers.

REFERENCES

1. Rout, J.K., Choo, K.K.R., Dash, A.K., Bakshi, S., Jena, S.K., Williams, K.L., 2018, *A model for sentiment and emotion analysis of unstructured social media text*, Electronic Commerce Research, 18, pp. 181-199.
2. Liu, H., Chatterjee, I., Zhou, M., Lu, X.S., Abusorrah, A., 2020, *Aspect-based sentiment analysis: A survey of deep learning methods*, IEEE Transactions on Computational Social Systems, 7(6), pp. 1358-1375.
3. Trueman, T.E., Cambria, E., 2021, *A convolutional stacked bidirectional LSTM with a multiplicative attention mechanism for aspect category and sentiment detection*, Cognitive Computation, 13, pp. 1423-1432.
4. Sadr, H., Pedram, M.M., Teshnehlab, M., 2019, *A robust sentiment analysis method based on sequential combination of convolutional and recursive neural networks*, Neural Processing Letters, 50, pp. 2745-2761.
5. Silva, N.F.F.D., Coletta, L.F., Hruschka, E.R., 2016, *A survey and comparative study of tweet sentiment analysis via semi-supervised learning*, ACM Computing Surveys (CSUR), 49(1), 15.
6. Dang, N.C., Moreno-García, M.N., De la Prieta, F., 2020, *Sentiment analysis based on deep learning: A comparative study*, Electronics, 9(3), 483.
7. Rong, W., Peng, B., Ouyang, Y., Li, C., Xiong, Z., 2013, *Semi-supervised dual recurrent neural network for sentiment analysis*, In 2013 IEEE 11th International Conference on Dependable, Autonomic and Secure Computing (pp. 438-445), Chengdu, China, IEEE.
8. Evans, D., Bratton, S., McKee, J., 2021, *Social media marketing*, AG Printing & Publishing.
9. *Capatina*, A., Kachour, M., Lichy, J., Micu, A., Micu, A.E., Codignola, F., 2020, *Matching the future capabilities of an artificial intelligence-based software for social media marketing with potential users' expectations*, Technological Forecasting and Social Change, 151, 119794.
10. Mattila, M., Salman, H., 2018, *Analysing social media marketing on twitter using sentiment analysis*, KTH.
11. Kitsios, F., Kamariotou, M., Karanikolas, P., Grigoroudis, E., 2021, *Digital marketing platforms and customer satisfaction: Identifying eWOM using big data and text mining*, Applied Sciences, 11(17), 8032.
12. Chen, L.C., Lee, C.M., Chen, M.Y., 2020, *Exploration of social media for sentiment analysis using deep learning*, Soft Computing, 24, pp. 8187-8197.
13. Umer, M., Imtiaz, Z., Ahmad, M., Nappi, M., Medaglia, C., Choi, G. S., Mehmood, A., 2023, *Impact of convolutional neural network and FastText embedding on text classification*, Multimedia Tools and Applications, 82(4), 5569-5585.
14. Wang, G., Sun, J., Ma, J., Xu, K., Gu, J., 2014, *Sentiment classification: the contribution of ensemble learning*, Decision Support Systems, 57, pp. 77-93.
15. Ahmad, K. ed., 2011, *Affective computing and sentiment analysis: emotion, metaphor and terminology*, Springer Science & Business Media.
16. Liu, B., 2022, *Sentiment analysis and opinion mining*, Springer Nature.
17. Saumya, S., Singh, J.P., 2018, *Detection of spam reviews: a sentiment analysis approach*, CSI Transactions on ICT, 6(2), pp. 137-148.
18. Prusa, J., Khoshgoftaar, T.M., Dittman, D.J., Napolitano, A., 2015, *Using random undersampling to alleviate class imbalance on tweet sentiment data*, In 2015 IEEE International Conference on Information Reuse And Integration, pp. 197-202, San Francisco, CA, USA, IEEE.
19. Li, S., Wang, Z., Zhou, G., Lee, S.Y.M., 2011, *Semi-supervised learning for imbalanced sentiment classification*, In Twenty-Second International Joint Conference on Artificial Intelligence, Barcelona Catalonia Spain, AAAI Press.
20. Manning, C.D., Surdeanu, M., Bauer, J., Finkel, J.R., Bethard, S., McClosky, D., 2014, *The stanford CoreNLP natural language processing toolkit*, In Proceedings of 52nd Annual Meeting of the Association For Computational Linguistics: System Demonstrations, pp. 55-60, Baltimore, Maryland, Association for Computational Linguistics.
21. Hamilton, W.L., Clark, K., Leskovec, J., Jurafsky, D., 2016, *Inducing domain-specific sentiment lexicons from unlabeled corpora*, In Proceedings of the Conference on Empirical Methods In Natural Language Processing (Vol. 2016, p. 595), Austin, Texas, Association for Computational Linguistics.
22. Larson, K., Watson, R., 2011, *The value of social media: toward measuring social media strategies*, In ICIS 2011 Proceedings, Shanghai.
23. Foltean, F.S., Trif, S.M., Tuleu, D.L., 2019, *Customer relationship management capabilities and social media technology use: Consequences on firm performance*, Journal of Business Research, 104, pp. 563-575.
24. Sajid, S.I., 2016, *Social media and its role in marketing*, Business and Economics Journal, 7(1), 203.
25. Kaplan, A.M., Haenlein, M., 2012, *Social media: back to the roots and back to the future*, Journal of Systems and Information Technology, 14(2), pp. 101-104.
26. Bisio, F., Oneto, L., Cambria, E., 2017, *Sentic computing for social network analysis*, In Sentiment Analysis in Social Networks, Morgan Kaufmann, pp. 71-99.

27. Jamil, K., Dunnan, L., Gul, R.F., Shehzad, M.U., Gillani, S.H.M., Awan, F.H., 2022, *Role of social media marketing activities in influencing customer intentions: a perspective of a new emerging era*, Organizational Psychology, 12, 808525.
28. Ibrahim, B., 2022, *Social media marketing activities and brand loyalty: A meta-analysis examination*, Journal of Promotion Management, 28(1), pp. 60-90.