International Journal of Computing Sciences Research (ISSN print: 2546-0552; ISSN online: 2546-115X)
Vol. 8, pp. 2635-2676
doi: 10.25147/ijcsr.2017.001.1.179
https://stepacademic.net

Short Paper

A Systematic Review of the Sentiment Analysis Models Used in Handling Polarity Shift

Millicent K. Murithi
Computer Science Department, Murang'a University of Technology, Kenya
mkathambi@mut.ac.ke
(corresponding author)

Aaron M. Oirere
Computer Science Department, Murang'a University of Technology, Kenya

Rachael N. Ndung'u
Information Technology Department, Murang'a University of Technology, Kenya

Date received: January 2, 2024
Date received in revised form: February 5, 2024; February 8, 2024
Date accepted: February 19, 2024

Recommended citation:

Murithi, M. K., Oirere, A. M., & Ndung'u, R. N. (2024). A systematic review of the sentiment analysis models used in handling polarity shift. I*nternational Journal of Computing Sciences Research*, *8*, 2635-2676. https://doi.org/10.25147/ijcsr.2017.001.1.179

Abstract

Purpose – This comprehensive review aims to analyze sentiment analysis models, focusing on their effectiveness in managing polarity shifts (positive, neutral, and negative sentiments). The investigation delves into specific aspects, including feature selection/extraction techniques, data augmentation, input datasets, and evaluation metrics. The goal is to provide valuable insights for both theoretical understanding and practical applications, guiding advanced research.

Method – Following Kitchenham's framework, the review conducts an exhaustive literature search, emphasizing entities embedded within sentiment analysis models. This methodological approach explores intricate challenges associated with polarity shifts, revealing opportunities for future research.

Results – Key findings highlight the efficacy of sentiment analysis models in handling polarity shifts. Recursive Feature Elimination emerges as the most effective wrapper-based feature selection technique, with Word2Vec standing out for word embeddings. Notably, negation is identified as a significant polarity shifter. The study also identifies common datasets, machine learning, and deep-learning models, emphasizing the effectiveness of hybrid models and ensembles.

Conclusion – This systematic review offers a comprehensive analysis of sentiment analysis models, providing insights into their status and key trends. The findings contribute to advancing the field, offering valuable guidance for researchers, practitioners, and developers working on polarity shift challenges, particularly in addressing implicit negation.

Recommendations – Future research should prioritize refining models addressing implicit negation, requiring empirical investigations to assess their effectiveness. Additionally, efforts should focus on developing standardized evaluation metrics to accurately capture the intricacies of polarity shifts, ensuring a comprehensive assessment of model performance. Implementing these recommendations will advance the state-of-the-art.

Research Implications – The review suggests practical and theoretical implications by highlighting effective models and their embedded entities. Recognition of explicit negation underscores the need for models capable of discerning and addressing negated sentiments. Practitioners are encouraged to adopt diverse strategies, and research opportunities lie in exploring implicit negation for ongoing innovation in sentiment analysis research.

Keywords – sentiment analysis, positive polarity, negative polarity, polarity shift, sentiment classification

INTRODUCTION

Online activities, using social networking platforms, have proliferated in recent years. This has enabled users to point out their comments regarding products, companies, and services (Singh & Paul, 2021). These activities provide an enormous amount of both structured and unstructured data because of their high volume, velocity, diversity, value, and variability (Farooq, 2017). This data can be mined and made useful through natural language processing (Kafi et al., 2019). Sentiment analysis is a natural language processing tool used to determine the emotional undertone of text (Kafi et al., 2019). Sentiment analysis can accurately extract the opinions of individuals and determine their sentiments from huge unstructured review texts. Sentiment analysis can classify opinions into either positive, negative, or neutral.

Machine learning, deep learning models, and hybrid / ensemble models are the main approaches to sentiment analysis (A. Kumar & Garg, 2023);(S. Kumar & Pathak, 2021). However, with these models, problems arise when a positive polarity is reversed to a negative polarity, and vice versa, resulting in a polarity shift problem (Singh & Paul, 2021). Factors that affect polarity shift are referred to as polarity shifters because they can change the sentiment's polarity (Zirpe & Joglekar, 2017b). Several polarity shifters exist (Zirpe & Joglekar, 2017a). These include negation, contrast, sarcasm and irony, context, ambiguity, evolving language, emoticons and emojis, subjectivity, cultural and regional differences, news and events, authors' tones, review fraud, pragmatics and domain-specific knowledge (Eke et al., 2021); (R. Xia et al., 2016). Based on these polarity shifters, machine learning models need effective training strategies to recognize and adapt to the various polarity shifters. Such specialized models, as well as fine-tuning on specific datasets, can help improve a model's ability to handle polarity shifts, providing more accurate sentiment analysis results. In written text, polarity shift is critical because it frequently changes the sentiment polarity of text from positive to negative and vice versa(R. Xia et al., 2016). This has the effect of giving the sentence a meaning that is different from what the linguistic model based on machine learning would have anticipated. Polarity shift can mislead businesses and consumers into making wrong decisions whereby a customer ends up purchasing the wrong product or the business ends up picking the wrong product perception from the customers. In the current digital age, where automated opinion mining is possible, it is necessary to handle the polarity shift problem more effectively (Zirpe & Joglekar, 2017b). In these models, there is a need to detect and handle polarity shifters to achieve a high classification performance. To this end, this review aims to yield key results findings regarding the effectiveness of the sentiment analysis models used in handling the polarity shift problem.

The rest of the paper is organized as follows: Related work on the state-of-the-art models used in detecting and handling the polarity shift problems is presented in the second section. The research methodology employed in this review study is presented in the third section. The fourth section presents the results of the review process. Lastly, in the fifth section, the conclusion and future research direction are presented. This is drawn from the systematic review process.

RELATED WORK

This section focuses on the common polarity shifters that contribute to the challenge of polarity shift; the feature selection processes used with the sentiment analysis models for handling polarity shift as well as the metrics used in the evaluation process on these models. The discussion on the trends of the entities intertwined in these models is also done with a focus on developing models that have improved resilience and efficacy.

Common Polarity Shifters in Sentiment Analysis

The polarity shift is one of the major problems with the existing sentiment analysis models. The main causes of polarity shifts are referred to as polarity shifters. Polarity shifters are the factors that can change a word's prior polarity in one of three ways i.e. rise, decrease, or neutral (R. Xia et al., 2016). According to Madhuli and Rahuli, there is a need to consider the most significant factors that are key to polarity shifts (Yadav & Katarya, 2020). The effectiveness of the sentiment classification models can be enhanced by the automated detection and handling of such factors. If not detected and handled effectively, these polarity shifters can have a detrimental effect on the classification performance of these models (R. Xia et al., 2016).

Several polarity shifters exist (Yadav & Katarya, 2020); (Zirpe & Joglekar, 2017a). These include negation, contrast, context, ambiguity, evolving language, subjectivity, cultural and regional differences, news and events, authors tones, review fraud, pragmatics and domain-specific knowledge (Abdi et al., 2019) ;(Eke et al., 2021); (Japhne & Murugeswari, 2020). The presence of negation words like "not" and "never" can invert the sentiment. For example, "not bad" means "good" (Singh & Paul, 2021). Negative words like "not", "neither", and other similar phrases can change the polarity of a sentence, thus making them crucial for sentiment analysis (R. Xia et al., 2016). Positive expressions include "This movie is good" and negative ones include "The movie is not good". At most times, certain approaches eliminate negation terms since they are termed as stop-word lists or disregarded in entirety (Singh & Paul, 2021).

The polarity shift of a sentence is based on two essential components of negation. These include the cue and the scope. The covering of negation's effects on the sentence is its scope (Singh & Paul, 2021). According to Nikhil, the classification of negation cues is divided into three different categories i.e. linguistic negations, structural and valence shifters (Suchetha et al., 2019). Negation in linguistics can either be explicit or implicit. While implicit negation uses phrases like rarely, avoid, etc., the explicit negation on the other hand expresses itself clearly (Zirpe & Joglekar, 2017a). Sarcasm and irony are forms of figurative speech where the speaker means the opposite of what they say. Detecting sarcasm and irony in sentiment analysis can be challenging for sentiment analysis models, resulting in polarity shift problems (Poria et al., 2016); (H. M. K. Kumar & Harish, 2018). The sentiment expressed in a statement can change depending on the context. For example, a phrase that might seem positive in one context can turn negative in a different context (H. M. K. Kumar & Harish, 2018);(Japhne & Murugeswari, 2020). Ambiguous language or statements, interpreted in multiple ways, may lead to polarity shifts, as different interpretations can carry different sentiments (Rajabi et al., 2020).

Language evolves, and the meaning of words or phrases can change. Those that were considered positive in the past might presently be considered negative, leading to polarity shifts (Rajabi et al., 2020). Regarding subjectivity, different people may have different perspectives on what is positive or negative, resulting in discrepancies in sentiment analysis (R. Xia et al., 2016); (Zirpe & Joglekar, 2017a). Sentiments can be cultural and regional-specific and what is considered positive in one culture or region might be perceived differently in another, leading to polarity shifts (Koncar et al., 2020). Sentiments can change based on current news, events, or trends. For instance, a well-received product might suddenly receive negative sentiments due to a recall or controversy (Japhne & Murugeswari, 2020). The tone and style of the author can influence how a text is interpreted. For example, a sarcastic or critical tone may make a generally positive statement appear negative (Poria et al., 2016).

Regarding review fraud, there can be fake or fraudulent online reviews and ratings that are intentionally designed to mislead sentiment analysis systems, resulting in polarity shifts (Cunha et al., 2018). Pragmatic aspects of language, such as implicates and presuppositions, can lead to polarity shifts. The meaning can depend not just on the words used but on the speaker's intended meaning (Davis, 2016). Lastly, in domain-specific sentiment analysis, the absence of domain-specific knowledge may lead to polarity shifts, as certain terms or expressions may have unique connotations in a particular domain (Wang et al., 2020). Sentence contrast occurs when the neighboring sentences' polarities diverge from their intended direction. Explicit contrast is the term used for this change in polarity. If the contrast is taken into consideration, the foremost half of the review's positive sentiment will be reversed by the presence of a contrast, but the entire review will take on a negative emotion (Japhne & Murugeswari, 2020). Connecting one phrase, sentence, or paragraph to another one expresses contradictions that fall under the intra-sentences, extra-sentences, and extra-paragraph categories (Zirpe & Joglekar, 2017b). When there are inconsistencies or conflicting sentiment patterns in the reviews, a form of polarity shift occurs. Such a change can take place in lengthy evaluations if the reviewer presents opposing perspectives about many aspects of the good or service (Shahnawaz & Astya, 2017). In this instance, the reviewer's opinion on one side of the product differs from their opinion of the product.

The factors discussed above contribute to sentiment classification accuracy because they reverse the polarity of a sentence from what was initially intended. The performance of the sentiment classification model is significantly impacted since not all polarity shifters have been incorporated into the existing sentiment analysis models used in detecting and handling the polarity shift problem.

Feature Selection Process in Sentiment Analysis

Through the use of numerous technologies, the internet has become more accessible in the modern era, making it possible for people to read product and service reviews and share ideas (Shahnawaz & Astya, 2017). A set of features from meaningful words obtained from data are used in sentiment analysis. The selection of subsets is a process for narrowing down a large corpus's relevant characteristics to enhance classification efficiency (Goularas & Kamis, 2019).

With feature selection techniques, sentiment analysis algorithms execute faster, making predictions and data analysis processes more accurate. Feature selection is very important, and choosing a suitable set of feature selection techniques improves the models' effectiveness (Kurniawan et al., 2022). To extract the features from several unstructured texts, data preprocessing, and feature extraction techniques are carried out. In subsequent sub-sections data preprocessing, feature extraction, feature selection, cross-validation, and lastly dimensionality reduction stages of the feature selection processes are described.

Data Preprocessing

Data preprocessing in the feature selection process comprises tokenization, lowercasing, removing punctuation, stemming, and lemmatization. This is performed to reduce words to their base form (Pradha et al., 2019); (Zirpe & Joglekar, 2017a).

In this review, the data preprocessing on the reviewed models includes the elimination of unnecessary noise such as hashtags, spaces, and uniform resource locators (URL). This process aids in text extraction because some symbols, URLs, and words are not necessary for classification features. To accomplish this, several steps are taken in the conversion process of unstructured text documents into word vectors.

Feature Extraction

Feature extraction is the process of transforming random text and images into numerical that could be used by sentiment analysis models (Hegde & Seema S., 2017). Since the mined user reviews are mostly in the form of text, there is a need to represent them numerically to be processed by the sentiment analysis algorithms. Several vectorization techniques, discussed in the subsequent section, are used in this process.

TF-IDF (Term Frequency and Inverse Document Frequency) Approach

The term frequency and inverse document frequency scores are used to highlight key phrases in a written form. It weighs the importance of a word based on its occurrence in a document as well as how rare the word is within the corpus matrix (Mee et al., 2021). This approach tries to reduce the influence of irrelevant words in a document (Mohd Nafis & Awang, 2021). The term's phrase frequency (tf) is clearly defined as the frequency of a term concerning the overall word count of the text. The IDF (Inverse Document Frequency) formula is used to determine the significance of a phrase (Mohd Nafis & Awang, 2021). Vectorization involves an assessment of how frequently a word appears throughout the text. (Ahuja et al., 2019). The "Term Frequency and Inverse Document Frequency" is a feature extraction method that has been used in several sentiment analysis models. Japhne and Murugeswari, in their study on the opinion mining complex model for handling polarity shift, use the TF-IDF technique to determine the important word (Japhne & Murugeswari, 2020). To achieve this, the technique considers the number of documents that the word appears. Ahmad and Aftab, in their work, use TF-IDF to evaluate the frequency of the most useful words in the document (Ahmad & Aftab, 2017). The process makes their sentiment detection process easy.

Kumar & Garg, in their empirical study, use TF-IDF to extract the most important terms in a text (A. Kumar & Garg, 2023). They also assessed the importance of a word in the corpus using the TF-IDF vectorizers. These are later used as the input features for training the classifier.

However, Alzami et al. compared the effects of the various feature extraction methods on the performance of the sentiment classification models. They compared Word2Vector, Word Bags, TF-IDF, and a hybrid of TF-IDF and Word2Vector (Alzami et al., 2020). It is noted that the main drawback of TF-IDF is its inability to capture text position, text semantics, and co-occurrence in many documents.

Use of Bag-of-words

Using the bag-of-words(BOW) method in NLP, a text document can be represented as an array of fixed-length vectors (Hegde & Seema S., 2017). It represents each document as a vector of word frequencies, typically using a fixed vocabulary (Zirpe & Joglekar, 2017b). The vector's length is equal to the document's vocabulary size. This fixed-length vector in the input document holds a value dependent on the word's occurrence (Hegde & Seema S., 2017). The technique employs a fixed-length vector format, making it appropriate for classification and clustering problems involving classification. The BOW technique proves as an ineffective model for classification since it doesn't care about the semantics of the sentence. The use of BOW as a feature extraction method has been applied in several sentiment analysis models used in handling polarity shift problems. Thevar et. al. created a sentiment analysis model for handling polarity shifts using the BOW feature extraction method (Thevar et al., 2017).

Word Embedding

This feature extraction technique represents words as dense vectors (S. Kumar & Pathak, 2021). Some examples of word embedding include; Word2Vec, GloVe, and FastText. Word2Ve uses a shallow two-layered neural network of the word embedding model (S. S. Kumar & Rajini, 2019). A word2Vec embedding is a term used to describe a word's numerical representation as a vector. Word2ec takes into account the related semantics of words to produce embeddings including the context of the word in a text. In the same or similar circumstances, the related words are mathematically classified into a vector space (S. S. Kumar & Rajini, 2019).

The "word embeddings" feature extraction methods are applied to several models. Avinash & Sivasankar, in their study, noted that Doc2vec is an extension of the word2vec and doc2vec models (Avinash & Sivasankar, 2019). Doc2vec improves the word-to-word sequences learning process of the embeddings that represent documents as fixed-length vectors with low dimensions. Singh and Paul, in their study, use word and character embedding for feature extraction. They conducted pre-training on the word embeddings to get the semantic and syntax information for each word in a sentence (Singh & Paul, 2021). Additionally, character embedding is used to capture negations that are morphological. As this usually assesses the frequency of helpful phrases, this offers significant and valuable information during the pre-processing stage and ultimately simplifies the sentiment detection process. The identification of significant information is highly influenced by the term frequency, as earlier explained by (Kafi et al., 2019).

Deho et al. generate high-dimensional word vectors that can recognize word context using Word2Vec (Oscar Deho et al., 2018). The results of the word vectors are then used to train the models. The results show improved classification accuracies. Rajabi et al. use a word embedding model to represent text semantically (Rajabi et al., 2020). They use batch numbers as per the concepts from the "ConceptNet" semantic network. Arora and Kansal propose a deep convolutional character level embedding for text normalization. Sentiment analysis is performed using a neural network model (Arora & Kansal, 2019). This model performs well when managing the small memory space difficulties in word-level embedded learning as well as in the processing of noisy sentences. The unstructured data performs precisely. Xia et al. suggest character-level embedding as a solution to the issue of the current text sentiment analysis algorithms' inability to extract sentiment words. Additionally, it resolves the pre-training word vector issue caused by the vocabulary gaps (H. Xia et al., 2020).

Singh and Paul identify a lack of contextual information as the main problem associated with one hot encoding vector (Singh & Paul, 2021). For instance, the phrases "Have a nice day" and "Have a great day". Even though the two sentences are fairly similar, they are unable to find any similarities between them when utilizing one hot encoding approach. This is because each term's vectors are orthogonal and do not share any contextual information. Rezaeinia et al. indicate, in their study, that among the most efficient and practical word embedding feature extraction techniques are the Word2Vec and GloVe. Nevertheless, they point out that these approaches overlook text sentiment information, and that to produce precise vectors during training, a sizable corpus of texts is required (Rezaeinia et al., 2019).

Count Vectorizer

This involves a simple vectorizer that converts every data token into a feature, through simple vectorization. Each text in the document is taken into consideration, and each relevant ID and word repetition is taken into account (Hegde & Seema S., 2017). The classifier can accept this multiset of words as input. The dataset is referred to as sparse because each file has a large number of zeros for every text that is absent from the dataset (Hegde & Seema S., 2017). The count vectorizer methods have been applied in several sentiment analysis models used in handling polarity shift problems. Rajat et al., in their study, carried out a sentiment analysis on a sizable real dataset from Amazon using the term frequency-inverse document frequency and the counter vectorizer (Rajat et al., 2021). Irawaty et al., in their study, conducted a comparison of the effectiveness of the following vectorizers, namely TFIDFVectorizer, HashingVectorizer, and CountVectorizer (Irawaty et al., 2020). Their results show that the best vectorizer is TFIDF because of its relatively higher accuracy in the prediction of negative values.

Additionally, there are more positive predictive values as compared to other vectorizers. However, Kalaivani et al., in their review, note that count-based embeddings do not exactly represent syntactic and semantic words in a document (Kalaivani et al., 2020).

Continuous bag of words (CBOW)

CBOW uses keywords in the context of predicting a word. Appropriate words are used as the input, and they are utilized for predicting the desired term. The corpus's vocabulary must be created to create a CBOW model (Mohd Nafis & Awang, 2021). Target and context generators are created afterward, and inputs are passed into the embedding layer before the lambda and dense layers. It is noted that a continuous bag of words guarantees all the vectors for each term orthogonally and shares all contextual information. It is therefore it is the most appropriate for neural network-based approaches (Singh & Paul, 2021). On the other hand, it provides contextual information in addition to representing a word as a vector. The continuous bags of words methods are applied to several sentiment analysis models used in handling polarity shift problems. Liu, in his study, uses a continuous bag of words (CBOW) for vector representation of text (Liu, 2020). For text sentiment analysis, feedforward neural networks serve as the foundation for the CBOW language model.

N-grams

N-grams, also referred to as bi-grams or tri-grams are used to capture word combinations that carry sentiment (Ahuja et al., 2019). The N-grams feature extraction methods are applied to several sentiment analysis models used in handling polarity shift problems. Ahuja et al. in their study, compare the N-Gram feature extraction method and the TF-IDF word level on the SS-Tweet dataset (Ahuja et al., 2019). For classification, they employed the Naïve Bayes machine learning algorithm, decision trees, support vector machines, and logistic regression. They concluded that using TF-IDF word level improves sentiment analysis performance by 3-4% when compared to N-grams. Kaur et al., in their study, used N-gram to perform feature extraction and post-tagging on the sentences (Kaur et al., 2018). They used the kNN classifier for sentiment analysis. Mutinda et al., in their work on sentence-level sentiment analysis, use a fixed hybrid N-gram window for feature extraction and a minimum redundancy maximum relevance feature selection algorithm (Mutinda et al., 2021). Their methodology enhances the N-gram feature extraction method by creating a hybrid vector from words, the semantic orientation of the words, and Part of Speech (POS) tags.

The most widely used feature extraction methods in sentiment analysis are a bag of words, N-gram, and Term Frequency-Inverse Document Frequency (Mutinda et al., 2021). However, it is noted that these methods face difficulties when they fail to take word relationships into account. They have a high feature dimensionality issue in addition to disregarding the properties of words.

Trends in Feature Selection Techniques for the Models Used in Handling Polarity Shift

The purpose of the feature selection techniques in sentiment analysis models is to minimize the dimensionality of the feature space through the identification of the most important features (Mhatre et al., 2017); (Pongthanoo & Songpan, 2020); (Prastyo et al., 2020). The specific phases of the feature selection process vary greatly. Some use metaheuristic techniques while others employ statistical and filter techniques (Miao & Niu, 2016). Feature selection techniques are becoming more popular in improving sentiment classification models. There are three types of feature selection techniques in sentiment analysis i.e. filter, wrapper, and hybrid approaches. In the subsequent sections, the feature selection methods used by the sentiment analysis models for handling the polarity shift problem are presented. They have been thematically categorized as filter methods, wrapper methods, and hybrid feature selection approaches.

Review of the Filter-based Approaches

In this approach, statistical measures are used to rank the features where the highest ranking is selected. Firstly, each feature is assigned a score while the top-scored features are then selected and ranked (Akbarian & Boroujeni, 2020). According to (Bommert et al., 2020), the filter-based approach is suitable for datasets with a high number of features besides working independently from the classifier used. Some commonly used filter-based approaches are Information Gain (IG) (Tubishat et al., 2019), Chi-square (CHI) (Prastyo et al., 2020), Document Frequency (DF) (Prastyo et al., 2020), and Mutual information (Prastyo et al., 2020). Several researchers have applied the filter-based feature selection approaches in their works.

Nurhayati et al. conducted a study to find out how the Naïve Bayes algorithm performed when Chi-square feature selection was used (Nurhayati et al., 2019). According to the results, Chi-square improved the recall, accuracy, and precision of the Naïve Bayes sentiment analysis model. In the studies conducted by (Paudel et al., 2019) and (Falasari & Muslim, 2022) the researchers concluded that Chi-square could increase the efficiency of the computation process while also increasing the accuracy of the classification method. However, in the study conducted by (Kurniawan et al., 2022), the researchers pointed out that although word frequency can have an impact on the classifier's performance, Chi-square did not take this into account. Different results occur because each has an exceptional case and dissimilar datasets. In (Manek et al., 2017), researchers proposed a feature selection technique based on the Gini index that is evaluated with a Support Vector Machine classifier.

Review of Wrapper-based Approaches

This technique combines machine learning algorithms with optimization of search algorithms to identify the most crucial features (Prastyo et al., 2020). The wrapper-based approach evaluates the features chosen based on the classification accuracy using a function called fitness. The iteration process is conducted to achieve the best features and feature subsets for a specific sentiment classification model. According to (Tubishat et al., 2020), the wrapper-based methods improve classification besides reducing the original feature space. Several researchers, in their works, have applied the wrapper-based methods for the feature selection. The researchers (S. Kumar & Pathak, 2021) use Particle Swarm Optimization (PSO) in their study to enhance the classification outcomes using the Naïve Bayes classifier. This algorithm operates on the feature weight-increasing principle. The application of the algorithm presents several factors, including the number of instances, the weight of inertia, and the constant speed of each feature to determine the optimal value.

The algorithm fails when used for datasets with high dimensions. Shang et al. propose a binary-based Particle Swarm Optimization (PSO) method for sentiment analysis feature selection (Shang et al., 2016). Their method aims at tackling the updated formula of velocity challenges experienced by the old PSO algorithm. Kristiyanti and Wahyudi compare three wrapper feature selections to enhance the classification performance of the Support Vector Machine i.e. the Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Principal Component Analysis (PCA). The datasets gathered from Amazon's product reviews for cosmetics are utilized. SVM outperforms the other feature selection techniques and increases the accuracy when using PSO, according to the experimental results (Prastyo et al., 2020). In their comparative study (Kristiyanti et al., 2019), the PSO performed better than GA. It was determined that the PSO is an effective wrapper technique that can get over a wide range of challenges. In (Ghosh et al., 2020), the Ant colony optimization (ACO), a meta-heuristic algorithm, is employed by the researchers to choose features. Gokalp et al. propose a wrapper-based feature selection technique for sentiment analysis. Their model is developed using a Greedy Algorithm that makes use of six distinct metrics based on filters (Gokalp et al., 2020). The Vortex Search Algorithm (VSA) model is implemented by (T.M. et al., 2022). It selects the best characteristics and discards the rest. They used the Naive Bayes algorithm to categorize tweets. Based on the experimental results, the VSA feature selection method increases the accuracy of the model by an average of 15%. With emoticon-based functions, though, this is not the case as they conclude that the model can be enhanced to accommodate the merging of the sentiment and textual diffusion data.

Review of Hybrid-based Approaches

The hybrid approach combines the use of filter and wrapper-based methods to obtain the best possible feature subset (Mohd Nafis & Awang, 2021). Hybrid approaches typically achieve high performance and accuracy by utilizing the advantages of the combined approaches. The top-scoring features are chosen and ranked in the filter method, which lowers the dimensionality of the original feature space. High-efficiency features from the filter-based approach and the best performance features from the wrapper-based approach are chosen via the hybrid approach. The hybrid feature selection approach has been applied by several researchers. Syafiqah and Suryanti suggest an enhanced machine learning-based hybrid feature selection method that improves sentiment classification accuracy.

They employed recursive feature elimination, word frequency-inverse document frequency, and support vector machines for sentiment classification (Mohd Nafis & Awang, 2021). The SVM score and recursive feature elimination were utilized to evaluate and rank the features through the identification of the most important features. The TF-IDF was used to ascertain the value of the feature in a text document. The features are iteratively evaluated and prioritized, and for sentiment classification, the SVM-RFE only uses the best features. However, in this model, high-dimensional datasets are unable to be addressed. The main challenge here is that the hybrid feature selection is inconsistent with the technique with different sets of datasets.

Elangovan and Subedha, suggest the Firefly (FF) and Levy flights (FFL) Models in their study as an approach that extracted features from reviews obtained from a website (Elangovan & Subedha, 2020). The model selects optimal features within a high-dimensional feature subset. The firefly approach aims at finding the best solution by modeling the behavior of a set of fireflies and computing each firefly's value based on its exact position. It employs an iterative procedure to update the positions of fireflies to ascertain the ideal value of the fitness function. Since the fireflies in this algorithm move toward their nearby neighbors' positions based on greater attraction values, the set stops over the best and most optimal response during sequential repetition (Elangovan & Subedha, 2020). The model is unable to support the fine-tuning of its features to generate the feature that would be most useful for classification purposes.

Patcharanikarn and Wararat, suggest using a hybrid of the Information Gain and Synthetic Minority Oversampling Techniques (SMOTE) for feature selection to obtain appropriate features for each class (Pongthanoo & Songpan, 2020). The information gain (IG) technique is employed in their study to lower the factor, and the SMOTE technique is used to correct the imbalance class. Before performing sentiment analysis and an accuracy evaluation, SMOTE allows the model to obtain an appropriate number of suitable samples for each class. J48, Naïve Bayes, k-Nearest Neighbor, and Support

Vector Machine (SVM) have their efficiency in every class evaluated. This technique needs improvement to facilitate feature selection during the data preparation stage. Deniz et al., suggest a hybrid feature selection method that combines evolutionary algorithmic techniques with an entropy-based metric (Deniz et al., 2021). They combine a wrapper-based method based on the Non-dominated Sorting Genetic Algorithm II with an information gain metric. By enhancing performance in real-world datasets and the benchmark dataset, Stanford Sentiment Treebank, the hybrid feature selection model increases the classification accuracy. Unfortunately, the feature vectorization process that makes sentiment analysis easier to perform is not controlled by the model. However, it is pointed out that the characteristics of the dataset and the exceptional difficulties in managing polarity shift in the sentiment analysis task should inform the choice of feature selection technique. It is therefore critical to experiment with multiple feature selection techniques and evaluate the effectiveness in improving model performance in handling polarity shift in the sentiment analysis.

Cross-Validation

Cross-validation is done after the vectorization process. To determine the set of data that performs better with the respective models, the process allows one set of data to train the model and another one to test the model. The different feature selection strategies are evaluated using a cross-validation strategy to ensure that there is no overfitting to a particular dataset or feature set (Japhne & Murugeswari, 2020). In (Japhne & Murugeswari, 2020), the researchers divide the dataset into 5 folds or subsets (k=5) of equal size, and each fold serves as a testing set once and a training set k-1 times. This 5-fold grid search cross-validation technique is applied in this work.

Dimensionality Reduction Methods

Principal Component Analysis (PCA) and t-distributed Stochastic Neighborhood and Embedding (t-SNE) are the two most popular dimensionality reduction methods in machine learning. They can be used to visualize or reduce the original feature space while retaining important information in cases where the high-dimensional feature space is extremely high (Mhatre et al., 2017). Researchers have applied various dimensionality reduction methods to improve the feature space. In (Mhatre et al., 2017), data preprocessing techniques are proposed to be used as dimensionality reduction techniques. To determine the preprocessing techniques with the highest accuracy, the researchers combine a number of them on the "Bag of Words Meets Bags of Popcorn" dataset from the Kaggle repository. They predicted sentiments using a random forest classifier. Based on the feature weighting, the researchers put forth an enhanced semi-supervised dimensionality reduction framework (Kim, 2018). They use a linear feature extraction that offers mapping and feature weighting to overcome the limitations of sentiment classification. They take into account the structural information and the labels in a dataset. In (Raunak et al., 2019), propose an effective dimensionality reduction method for word embeddings. They combine PCA based method and post-processing algorithm. The post-processing projects the embeddings away from the most dominant directions, therefore improving their performance and making them more discriminative.

Models Used in Handling Polarity Shift in Sentiment Analysis

Sentiment analysis aims to classify words as positive, negative, or neutral by identifying those conveying sentiment in unstructured text inputs. Among the main problems with sentiment analysis is the polarity shift. According to research, the accuracy of sentiment analysis is improved by polarity shift detection (Singh & Paul, 2021). Researchers have proposed models for addressing the problem of polarity shifts in sentiment analysis (Abirami & Gayathri, 2017). For handling polarity shifts, machine learning and deep learning models are the two main types of sentiment analysis models currently in use. The subsequent sections highlight the various models for handling polarity shifts based on their various categories.

Popular Machine Learning-based Models Used in Handling Polarity Shift

Cruz et al. (2016) present an automated approach for machine learning to detect negations in review texts (Cruz et al., 2016). In their two-step methodology, the cues for speculation and negation are first recognized. The classifier finds the words at the sentence level that are affected by the cues found in the first phase in the second step. The Support Vector Machine is used in the training of the classification algorithms. The model is unable to clarify all of the polarity shifters, sentence inconsistencies, and contrasts.

A stacking ensemble technique for handling polarity shift comprising Naive Bayes, Logistic Regression, and Support Vector Machine classifiers is proposed (Zirpe & Joglekar, 2017b). Using the maximum score of each classifier as the input and the output of the base classifiers as the output, the stacking ensemble approach determines the result. The approach combines statistical and rule-based techniques to identify and eliminate polarity shifts. The method is unable to handle the more complex polarity shift structures.

The researchers propose reinforcement learning as a technique for recognizing, understanding, and interpreting negations in natural language (Pröllochs et al., 2020). The authors can avoid the need for expensive word-level annotations on financial datasets by handling negation at the document level. The performance of their model is comparable to human interpretations of negation, and it outperforms the rule-based approaches. This model can be enhanced to take into consideration all the different polarity shifters.

Kamal et al. (2022), suggest an improved negation handling method that makes use of Part-of-Speech (POS) tagging and the Naïve Bayes algorithm for sentiment analysis of Twitter data. The suggested method seeks to identify words that directly negate something, like "not" and "no." (Kamal et al., 2022).

Popular Deep Learning Models Used in Handling Polarity Shift in Sentiment Analysis

The authors propose a model for labeling sequences in recursive neural networks for tasks that involve handling negations (Fei et al., 2020). This model is a neural network with a tree structure that can represent fully global dependency trees. Syntactic information is automatically learned by the model from a global dependency tree. The primary weakness of this model is its heavy reliance on dependency-tree structure parsing data. Therefore, if the parser provides the model with inaccurate parsing information for a sentence, the model yields an inaccurate output. It is crucial to look into a neural model based on trees that can incorporate characteristics from dependency and constituency trees automatically for negation detection.

A deep learning approach to handle negation in sentiment analysis is proposed by Singh and Paul (2021). The proposed deep neural network model for the negation handling problem is based on long short-term memory. It automatically learns negation attributes from labeled input training datasets. The primary weakness of this model is that it only remembers the events of each batch without passing on information to the ones that follow. Only negations polarity shifters are accounted for in this model. The researchers propose a hierarchical multi-task learning approach that incorporates the information about negation by using cascading and hierarchical neural architecture (Barnes et al., 2021).

A recursive neural network sequence labeling model for handling polarity shifts is proposed (Fei et al., 2020). They use the global dependency structure whereby the model first learns a high-level representation of each word in the context of each sentence before capturing the cue word and its target scope. A deep learning approach for negation detection is proposed by (Montenegro et al., 2021). This model looks for negation in Spanish-language product reviews using a transformer-based approach.

RESEARCH METHODOLOGY

The various steps of undertaking the systematic review are discussed in this section. The review adheres to Kitchenham's five-step process to conduct a systematic literature review (Kitchenham et al., 2009). This process involved framing the questions for the review guided by the five research questions. Secondly, the identification of the relevant literature is done. Thirdly, inclusion and exclusion criteria for the publications which included identification of the terms and resources to be used are done. A study of the quality assessment procedure is identified before finally discussing the research findings.

Research Questions

This systematic review study is guided by the following research questions:

RQ1: Which sentiment analysis models are used in detecting and handling the polarity shift problem?

RQ2: What polarity shifters are commonly detected and handled by the existing sentiment analysis models?

RQ3: What is the nature of the input datasets used with the sentiment analysis models for detecting and handling polarity shift problems?

RQ4: What feature selection techniques, feature extraction techniques, and data augmentation strategies are commonly applied in the sentiment analysis models used for detecting and handling the polarity shift problem?

RQ5: What metrics are used to evaluate the effectiveness of the sentiment analysis models used for detecting and handling the polarity shift problem?

Search string and Query space

A combination of specific keywords is used to search for the articles from the online libraries. The five research questions, which include sentiment analysis, models, and polarity shift directed the selection of the precise keywords. Following that, a search query was finalized with the following keywords: "Models" OR "Techniques" AND "polarity shift handling in sentiment analysis". These articles were mainly from IEEE, ACM, Research Gate, Google Scholar (scholar.google.com), and Springer databases which are the biggest collection of citations and abstracts. These sources were chosen because they had published papers on the topic.

Criteria for Selection

The inclusion and exclusion criteria are guided by the research questions. The identification of the literature meta-search on articles dedicated to the "polarity handling in sentiment analysis" and "models" as the main keywords. The articles published between 2017 and 2023 are identified. This choice indicates that a full-date review of the literature has not been undertaken. A total of 80 papers were obtained through an exhaustive search before the quality assessment was conducted. Quality assessment and selection criteria further ensured that only the articles that have a greater relevance to the research question were chosen.

Assessment on Quality

Quality assessment parameters take the following into account to achieve good quality results that answer the research questions:

- To extract the relevant research materials, only reliable electronic scientific libraries mentioned above are utilized.

- Only the most recent peer-reviewed journal paper or published conference proceedings are considered to ascertain the good quality results
- The process of selecting the relevant articles is balanced and uses the following criteria for assessment; if the paper was a primary, if the study objective which is a polarity shift in sentiment analysis, and if the paper discussion had clarity.
- Kitchenham's five-step process to conduct a systematic literature review is strictly followed.

RESULTS

The main purpose of this systematic review is to identify the most important data from the publications that have been released within the last five years. The main focus of the articles is the deep learning and machine learning models used to handle the polarity shift problem in sentiment analysis. Table 1 presents the sentiment analysis models used for detecting and handling the polarity shift problem, their key strategies as well as their limitations. Table 2 presents polarity shifters that are commonly detected and handled by the existing sentiment analysis models as well as their impact on the performance. Table 3 presents the name of the input datasets used, their nature, dimensionality levels as well as their number of instances. Table 4 presents feature selection techniques, feature extraction techniques, and data augmentation strategies for enhancing training data in the existing sentiment analysis models. Their strategies as well as their performance gaps are also presented. Table 5 presents metrics used to evaluate the effectiveness of the sentiment analysis models used for detecting and handling the polarity shift problem, including their suitability. The results' findings, in each of these tables are aligned to the research questions in this study.

Table 1. Sentiment analysis models used in the detection and handling of the polarity shift problem

Author, Date	Machine learning techniques used	Description	Limitations
Singh and Paul, 2021 (Singh & Paul, 2021)	LSTM-based Deep Neural Network Approach	This model is trained to first recognize negation cues in sentences. Bidirectional LSTM is then employed to determine the scope of the cue in sentences by determining the relationship between the cue and other words. To ascertain the correct polarity of the sentence, the model is trained using word-level features.	The model needs more data preprocessing to allow for improved results. The model only handles negation polarity shifter.

Table 1. Sentiment analysis models used in the detection and handling of the polarity shift problem (cont.)

Author, Date	Machine learning techniques used	Description	Limitations
Barnes et al. 2020 (Barnes et al., 2021)	Hierarchical Multi-task Learning Approach	The approach incorporates the information about negation by using cascading and hierarchical neural architecture. The model is trained with negations as an auxiliary task.	The model experimented with negation, shows the need to include others like linguistic and paralinguistic phenomena which also affect sentiment classification
Fei et al., 2020 (Fei et al., 2020)	Recursive Neural Network Based on Sequence Labeling Model	The cue word and its target scope are identified using the global Dependency structure after this first a high-level representation for each word in the context of each sentence. The representation from the recursive neural network is then fed into the Conditional Random Fields (CRF) layer, which collaborates to jointly decode the labels.	The constituency tree and dependency tree features are not integrated by the model to enable scope detection.
Prollochs et al. 2020 (Pröllochs et al., 2020)	Reinforcement Learning Framework	This model uses a dynamic learning framework that detects negations. It uses document-level labels as input and learns negation policy. The model aims to replicate human perceptions.	This model can be enhanced to take into consideration all the most significant polarity shifters.
Japhne and Murugeswari, 2020 (Japhne & Murugeswari, 2020)	Opinion Mining based Complex Polarity Shift Pattern	To determine which machine learning model best fits the dataset, this model tests the datasets using nine different models. Polarity shifters, contrast transitions, intensifiers, diminishers, and negatives are all handled. Because random forest employs an ensemble approach, it performs better.	Word sense disambiguation polarity shifters that can enhance the sentiment classification process are not fully addressed by the model. The use of Deep Learning techniques can further increase the accuracy.

Table 1. Sentiment analysis models used in the detection and handling of the polarity shift problem (cont.)

Author, Date	Machine learning techniques used	Description	Limitations
Zirpe and Joglekar, 2017 (Zirpe & Joglekar, 2017a)	Stacking Ensemble Method	Preprocessing, polarity shift detection and elimination, sentiment classification, and stacking ensemble method are the four modules that make up the suggested model.	This model can be improved to handle more complex polarity shift structures like sentence inconsistencies and contrasts.
Mathapati et al. (Mathapati et al., 2019)	Semi-supervised Domain Adaptive Dual Sentiment Analysis	This model converts the reviews to their opposite review. The Bayesian classifier is used to train both the original and reversed reviews. Probabilities of the tuples or the specified class are used by the classifier. The reviews are put to the test using later dual prediction. LSTM handles the classification while CNN handles feature extraction.	This model is trained to handle only domain-specific polarity shifters.
Montenegro et al. (Montenegro et al., 2021)	Negation Detection based on Deep Learning Approach	This model looks for negation in Spanish-language product reviews using a transformer-based approach. It performs negation detection using a BERT-based model and transfer learning techniques.	The model does not perform the sentiment analysis tasks. It only detects the negation in the sentences.
Khandelwal and Sawant (Khandelwal & Sawant, 2019)	Transfer Learning Approach to handle negations	The model uses Neg BERT architecture which employs a series of tokenized and encoded sentence tokens as the input. For all three datasets, cue detection and scope resolution are carried out; one dataset is used for training, and the other two for testing.	The model may require a bigger dataset to extract the maximum generalizability from the architecture.

Table 1. Sentiment analysis models used in the detection and handling of the polarity shift problem (cont.)

Author, Date	Machine learning techniques used	Description	Limitations
Lazib et al. (Lazib et al., 2019)	Negation Scope Detection based on Recurrent Neural Networks Technique	Sequence labeling is implemented in the model, whereby each word in a sentence is designated as either inside (I) or outside (O) of the scope. The RNN model classifies every word in a sentence as belonging to the scope or not. The RNN layer then receives as inputs the word embedding model that is produced as well as the index representation on the sentence. It is done to predict each label's probability. Lastly, the labels that each word in the sentence corresponds to are the model's output.	The model only performs negation detection. There is a need to perform sentiment analysis.
Chen (Chen, 2019)	Negation and Assertion Detection based on Attention-based Deep Learning System	To identify assertions and negations, the model employs word embeddings and attention-based BiLSTM networks. This approach focuses on scope resolution to solve NLP issues.	The model is domain-specific and handles the negation only.
Thevar et al. (Thevar et al., 2017)	Opinion Mining-based Model for Handling Polarity Shift	The negation in the sentence and the negator's proximity to the associated word are ascertained by this model using the Stanford parser. Governor word extraction follows. The Reverse Term Document Matrix is subtracted from the Term Document Matrix to create the Term Document Matrix, which is then fed to the Random Forest Classifier.	The model only handles negation polarity shifters.

Figure 1 below shows the frequency of the different types of models used for handling the polarity shift problem.

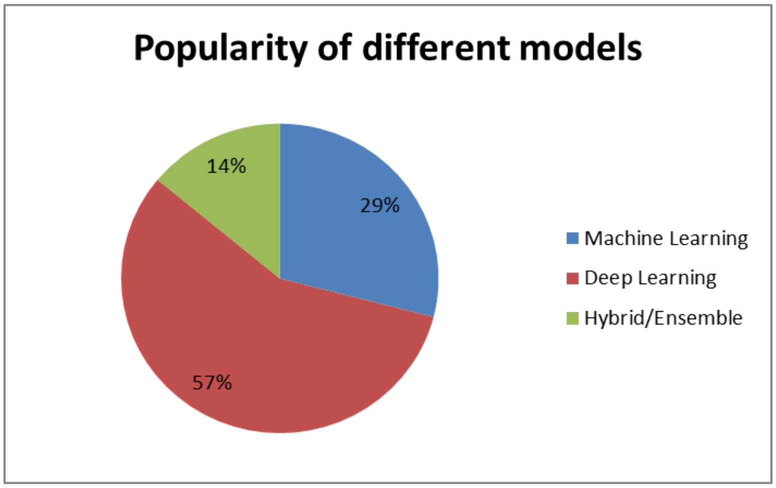

Figure 1. Frequency of the different types of models used in polarity shift management

Table 2. Common polarity shifters in the existing models and their significance in the sentiment analysis process

Author, Date	Model Name	Key polarity shifter detected and handled	Significance of the polarity shifter in the sentiment analysis process
Singh and Paul, 2021 (Singh & Paul, 2021)	LSTM-based Deep Neural Network Model	Negation	Negation plays a significant role in written text. Negations are inherently dynamic because of their association with linguistic complexity. They can be explicit or implicit, and their use is highly domain-specific.
Barnes et al. 2020 (Barnes et al., 2021)	Multi-task Approach	Explicit negation with a larger scope	Sentiment analysis is directly impacted by compositional phenomena, the most common of which is negation. To enhance the precision of sentiment classification, the classifier must possess the ability to recognize negation and separate the impact of its extent on the ultimate polarity of a given text.

Table 2. Common polarity shifters in the existing models and their significance in the sentiment analysis process (cont.)

Author, Date	Model Name	Key polarity shifter detected and handled	Significance of the polarity shifter in the sentiment analysis process
Fei et al., 2020 (Fei et al., 2020)	Recursive Neural Network Sequence Labeling Model	Negation and speculation scope	Negation and speculation scope detection is a sequence labeling task. Therefore, there is a need to employ a labeling scheme for each token. Both speculative scope and negativity will benefit from one another.
Prollochs et al. 2020 (Pröllochs et al., 2020)	Reinforcement Learning Framework	Negation	The presence of negatives makes it challenging to interpret the narrative's meaning. Negation use is domain-specific. Therefore, there is a need to use dynamic learning to detect negations.
Zirpe and Joglekar, 2017 (Zirpe & Joglekar, 2017a)	Stacking Ensemble Method	Negation, contrast and sentiment inconsistencies	Typical polarity shifters include negativity, contrast, and sentiment inconsistency. The use of negation words and contrasts causes models to classify the text as positive or negative when it isn't.
Japhne and Murugeswari, 2020 (Japhne & Murugeswari, 2020)	Opinion Mining based Complex Polarity Shift Pattern	Negations, intensifiers, diminishers, and contrast transitions	Complex linguistic structures known as "polarity shifters" can change or reverse the text's sentiment polarity.
Mathapati et al. (Mathapati et al., 2019)	Semi-supervised Domain Adaptive Dual Sentiment Analysis	Domain-specific shifters	Classifiers trained on one domain fail to perform well in another domain. This requires the model to be trained once more for the other domain.

Table 2. Common polarity shifters in the existing models and their significance in the sentiment analysis process (cont.)

Author, Date	Model Name	Key polarity shifter detected and handled	Significance of the polarity shifter in the sentiment analysis process
Montenegro et al. (Montenegro et al., 2021)	Negation Detection based on Deep Learning Approach	Negation	A negation trigger affects a set of tokens, known as scope, and negation detection is divided into two subsets: scope resolution and trigger detection.
Thevar et al. (Thevar et al., 2017)	Opinion Mining-based Model for Handling Polarity Shift	Negation	The BOG model produces a different overall impact because it pays little attention to polarity shifts. Even if the associated terms do not immediately follow the negation modifier, they still need to be negated.

Table 3. Commonly used Datasets

Author, Date	Model Name	Name of the input dataset	Nature and description of the datasets
Singh and Paul, 2021 (Singh & Paul, 2021)	LSTM-based Deep Neural Network Model	Conan Doyle (CD) Story Corpus	Negation Cues, Negated Events, and Scope are annotated on the CD dataset. It is taken from three novels.
Barnes et al. 2020 (Barnes et al., 2021)	Multi-task Approach	i. SFU Review Corpus ii. ConanDoyle-neg (CD) iii. Stanford Sentiment Treebank (SST)	Eight domains (books, cars, computers, cookware, hotels, movies, music, and phones) comprise 400 reviews found in SFU reviews. ConanDoyle-neg dataset contains 3640 sentences of the training set, Out of the 848 sentences, 144 are negated, and the remaining 787 sentences comprised a development set.

Table 3. Commonly used Datasets (cont.)

Author, Date	Model Name	Name of the input dataset	Nature and description of the datasets
Fei et al., 2020 (Fei et al., 2020)	Recursive Neural Network Sequence Labeling Model	BioScope and CNeSp datasets	Sentences with speculative and negative cues annotated along with their scopes in the biomedical domain constitute the BioScope datasets. The domains of scientific literature, financial articles, and product reviews are among the sentences that make up CNeSp, which is annotated with negative and speculative cues. Out of the 16,841 sentences in all, over 20% are negation or speculation.
Prollochs et al. 2020 (Pröllochs et al., 2020)	Reinforcement Learning Framework	Loughran-McDonald dictionary	It is a finance-specific dictionary that contains positive entries of 354 and negative entries of 2350.
Zirpe and Joglekar, 2017 (Zirpe & Joglekar, 2017a)	Stacking Ensemble Method (combines rule and statistic-based method)	Airline Reviews.	The reviews are taken from the Skytrax website and categorized into airline, airport, lounge, and seat.
Japhne and Murugeswari, 2020 (Japhne & Murugeswari, 2020)	Opinion Mining based on complex Polarity Shift Pattern	Standard labeled dataset	The dataset includes 2000 movie reviews, 1000 of which are positive and 1000 of which are negative.
Xia et al. ,2015 (R. Xia et al., 2016)	Three-stage Cascade Model (Polarity Shift Detection, Elimination and Ensemble)	Simon Fraser University Review Corpus (SFU)	There are 2000 movie reviews in the datasets, 1000 of which are positive and 1000 of which are negative.

Table 3. Commonly used Datasets

Author, Date	Model Name	Name of the input dataset	Nature and description of the datasets
Mathapati et al. (Mathapati et al., 2019)	A Dual Sentiment Analysis based on Semi-supervised Domain Adaptive	Multi-Domain dataset	The dataset comprises of reviews written in English across four categories: books, DVDs, electronics, and kitchens. There are 1000 positive and 1000 negative entries are present in each domain. The reviews are obtained from Amazon.com.
Montenegro et al. (Montenegro et al., 2021)	Negation Detection based on Deep Learning Approach	SFU ReviewSP-NEG	9455 Spanish sentences from user-generated product reviews are included in the datasets. Eight label categories are included in it, including those for cell phones, books, music, washing machines, computers, cars, and hotels. Negation cues and their ranges are annotated into sentences. 4327 cues total in the corpus; 386 do not negate and 3941 negate.
Khandelwal and Sawant (Khandelwal & Sawant, 2019)	Transfer Learning Approach to Handle Negations	The Sherlock dataset, the SFU Review Corpus, and the BioScope Corpus	The BioScope Corpus and SFU Review Datasets lack annotations for affixes, whereas the Sherlock dataset does have.
Chen (Chen, 2019)	Attention-based Deep Learning System for Negation and Assertion Detection	Patient discharge summaries and progress notes	170 discharge summaries and 256 progress notes are used. In addition, sentences that correspond to the clinical conditions are taken from 116 clinical notes.
Thevar et al. (Thevar et al., 2017)	Opinion Mining-based Model for Handling Polarity Shift	It contains 25,000 reviews for Bag of Words meets Bag of Pop-corn.	1800 IMDB movie reviews—900 positive and 900 negative. They are divided into partial datasets for this model's use. They include sentiment and review attributes.

Table 4. Feature selection techniques, Feature extraction techniques, and data augmentation strategies used with the models used in sentiment analysis

Author, Date	Sentiment Model Used	Intertwined entities	Limitations
Nafis and Awang, 2021 (Mohd Nafis & Awang, 2021)	Support Vector Machine (SVM) Classifier	Term Frequency-Inverse Document Frequency and Supports Vector Machine –Hybrid Approach	To determine the technique's effectiveness and efficiency, a larger dataset must be tested. Unable to choose the appropriate threshold for features from the TF-IDF feature list.
Elangovan and Subedha, 2020 (Elangovan & Subedha, 2020)	Multilayer Perceptron (MLP)	Firefly (FF) and Levy Flights (FFL)	The main challenge with this technique is the disparity between light intensity and brightness formulation. This is because the brightness of FF is considered with attractiveness that is connected by encoding the objective function.
Deniz et al., 2021 (Deniz et al., 2021)	Logistic Regression and Support Vector Machine	Entropy-based metric and an Evolutionary Algorithm-Hybrid Model	The technique does not control the feature vectorization step. It needs to be expanded to more datasets from different domains.
Patcharanikarn and Wararat, 2020 (Pongthanoo & Songpan, 2020)	Machine Learning Algorithms -Hybrid	SMOTE and Information Gain	Before testing using the classification method, the feature selection can be used during the data preparation process, which is crucial.

Table 4. Feature selection techniques, Feature extraction techniques, and data augmentation strategies used with the models used in sentiment analysis

Author, Date	Sentiment Model Used	Intertwined entities	Limitations
Abdur et al., 2020 (Rasool et al., 2020)	C4.5, SMO Classifier, Naïve bayes, and KNN	GAWA- utilized the Wrapper Approaches (WA) to select the premier features and the Genetic Algorithm (GA) to reduce the size of the premier features	The technique can be examined with multiple datasets from a variety of sources to select the best features with various categories of syntactic and stylistic features.
Joy and Dilip, 2022(Gorai & Shaw, 2022)	Used an Ensemble of SVM, KNN, DT	Chi-square-based feature selection	The technique is not tested using a deep-learning model
Wang et al,2022(Wang et al., 2022)	OL-DAWE mode	Synonym Replacement	the synonym replacement is the data augmentation technique that has been used in this paper
Wang et al., 2023(Wang et al., 2023)	Back Translation	Neuromorphic Data Augmentation - NDA	Sentiment Classification Based on RoBERTa and Data Augmentation.

Table 5. Metrics used to evaluate the performance of the models used in handling polarity shift in sentiment analysis

Author, Date	Machine learning techniques used	Evaluation Metrics used	Performance scores
Cruz et al., (Cruz et al., 2016)	A Machine Learning Approach to Negation and Speculation Detection	F1 measure, Geometric(G)-mean, percentage of correct relaxed scopes (PCRS), PCR	F1 of 84.07%.G-mean 90.42%, PCS 57.86%, and PCRS 79.13% values
Barnes et al. 2021 (Barnes et al., 2021)	Multi-task approach	Accuracy	Model Accuracy of 86.04%
Singh and Paul, 2021 (Singh & Paul, 2021)	LSTM-based Deep Neural Network model	F1-Score	F1-Score of 93.34%
Fei et al., 2020 (Fei et al., 2020)	Recursive Neural Network Sequence Labeling Model	F1 score and Percentage of Correct Scopes (PCS)	F1 score 90.4% and 88.5% (PCS) for detection of negation scope. 93.5% F1 and 91.7% PCS in speculation scope detection.
Japhne and Murugeswari, 2020 (Japhne & Murugeswari, 2020)	Opinion Mining based Complex Polarity Shift Pattern	Confusion Matrix, Accuracy, Precision, F1 Score, and Recall	3.5% improvement in accuracy.
Zirpe and Joglekar, 2017 (Zirpe & Joglekar, 2017a)	Stacking Ensemble Method	Recall, Precision, and Accuracy	97% Accuracy, 71% Precision, and 98% recall.
Montenegro et al. (Montenegro et al., 2021)	Deep Learning Approach for Negation Detection	F1 score	F1 score achieved 95.4% in the cue detection and scope resolution achieved 91.5%.
Chen (Chen, 2019)	Attention-based Deep Learning System for Negation and Assertion Detection	Precision, Recall, and F1 scores on negation were done	The micro-F1 score of 0.922 for assertion detection.
Thevar et al. (Thevar et al., 2017)	Opinion Mining Model for Handling Polarity Shift	10-fold cross-validation	Accuracy before was 80.16% and 90.88% after handling the polarity shift.

Figure 2 below shows the frequency of the different feature selection techniques used with the sentiment analysis models used in polarity shift management.

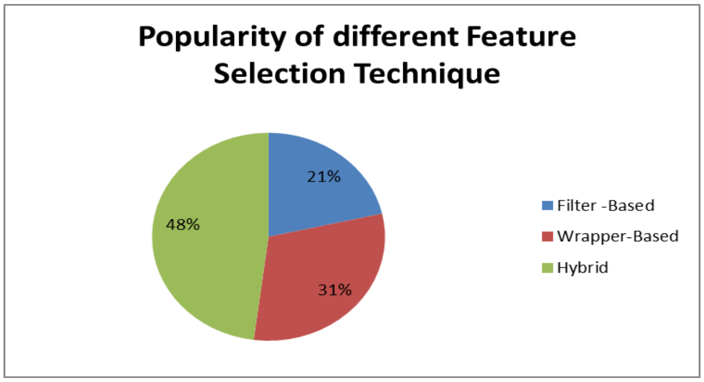

Figure 2. Frequency of the different evaluation metrics

DISCUSSIONS

Model's Diversity in Polarity Shift Management, Efficacy, and Future Opportunities

The review reveals a diverse landscape of sentiment analysis models employed for polarity shift management. Different models showcase varying approaches to handling the dynamic nature of sentiment polarity. The commonly applied traditional machine learning models include Support Vector Machines, Naive Bayes, and Logistic Regression. These models demonstrate robustness in certain contexts. Furthermore, deep learning models, notably LSTM, GRU, and BERT, are also deployed. The latter offers advanced capabilities in capturing nuanced sentiment variations and subsequently improved polarity shift management. Hybrid models and ensemble techniques, combining the strengths of different approaches, have also been applied. These models exhibit a rich diversity that contributes to enhanced efficacy in managing polarity shifts. The efficacy of sentiment analysis models in polarity shift management is evident through their ability to adapt to changing sentiments across different contexts. Recursive Feature Elimination stands out as the most effective wrapper-based feature selection technique, while the Term Frequency-Inverse Document Frequency proves to be a powerful filter-based method. Random Forest emerges as a proficient hybrid feature selection technique. In terms of feature extraction, Word2Vec and BERT demonstrate superior efficacy as word and contextual embeddings-based techniques, respectively. The findings underscore the importance of model diversity and highlight the nuanced effectiveness of specific techniques in managing polarity shifts within sentiment analysis.

The review identifies several promising avenues for future research in the realm of sentiment analysis models and polarity shift management. For example, addressing the challenges associated with implicit negation represents a critical area for exploration, requiring models that can discern and handle subtle shifts in sentiment expression.

Exploring the impact of temporal aspects on sentiment dynamics and developing models that can adapt to evolving language trends presents an exciting avenue for future studies. Empirical analysis among models could provide deeper insights into their respective strengths and limitations. Additionally, there is a need for standardized evaluation metrics that comprehensively capture the complexities of polarity shift. Future research endeavors should focus on refining existing models, exploring innovative hybrid approaches, and developing models that can handle diverse and domain-specific datasets. These opportunities collectively pave the way for advancing the field of sentiment analysis and enhancing models' efficacy in managing polarity shifts.

Trends in Feature Selection, Feature Extraction, and Data Augmentation Techniques

This review unveils notable trends in feature selection, feature extraction, and data augmentation techniques employed in sentiment analysis models dedicated to managing polarity shifts. The trends indicate a growing preference for sophisticated feature selection techniques that enhance model adaptability. Recursive Feature Elimination stands out prominently, demonstrating effectiveness in preserving relevant features crucial for capturing nuanced sentiment changes. Additionally, the emergence of Term Frequency-Inverse Document Frequency as a popular filter-based method suggests an inclination towards techniques that consider both term frequency and document specificity, optimizing the selection of informative features. The review highlights a shift towards more intricate hybrid feature selection methods, with Random Forest gaining prominence for its ability to capture complex relationships within the data.

In the realm of feature extraction, the review identifies a shift towards leveraging advanced word embeddings and contextual representations. Word2Vec remains a consistent trend as an effective word embedding technique, showcasing its continued relevance in capturing semantic relationships. BERT, representing contextual embeddings, emerges as a dominant trend, highlighting the increasing importance of contextual information in understanding sentiment nuances. The use of such sophisticated embeddings signifies a trend towards more nuanced and context-aware sentiment analysis models capable of addressing polarity shifts. Data augmentation trends in sentiment analysis models reveal a focus on techniques that enhance model generalization and robustness in the face of polarity shifts. Synonym replacement remains a commonly used technique, ensuring diversity in expressions without compromising sentiment integrity. Back translation emerges as a rising trend, showcasing an interest in multilingual augmentation for a broader understanding of sentiment variations. Furthermore, there is a notable trend toward developing domain-specific and sentiment-specific augmentation strategies, emphasizing the need for specialized approaches catering to the intricacies of different domains and sentiment expressions. These identified trends collectively suggest a paradigm shift towards more intricate, context-aware, and adaptable sentiment analysis models for managing polarity shifts.

The integration of advanced feature selection and extraction techniques, coupled with diverse and domain-specific data augmentation strategies, reflects the ongoing pursuit of models that can effectively capture the dynamic nature of sentiment in evolving language contexts.

Dynamism in Handling Polarity Shift and the Common Polarity Shifter

The review underscores a pronounced trend toward enhancing the dynamism of sentiment analysis models in effectively managing polarity shifts. The dynamism is reflected in models' adaptability to changing sentiments across diverse contexts. Traditional machine learning models such as Support Vector Machines, Naive Bayes, and Logistic Regression exhibit a certain level of dynamism, particularly in capturing explicit shifts in sentiment. However, it is the deep learning models, including LSTM and BERT that excel in providing a more dynamic understanding of sentiment changes, especially in nuanced and context-dependent scenarios. The integration of advanced embeddings and contextual representations contributes to the dynamism of these models, enabling them to capture subtle variations in sentiment expressions associated with polarity shifts. The systematic review reveals that among various factors influencing polarity shift, negation emerges as the most significant polarity shifter.

Sentiment analysis models commonly encounter instances where the presence of negation alters the polarity of sentiment expressions. Explicit negation, such as the use of "not" or "but," is effectively handled by most models. However, there is limited evidence of the effectiveness of these models in managing implicit negation, indicating a potential area for improvement. The commonality of negation as a polarity shifter highlights its importance in influencing sentiment dynamics, necessitating models that can discern and appropriately adjust sentiments in the presence of negating language. Addressing the complexities of negation in sentiment analysis models is crucial for enhancing their overall dynamism and effectiveness in managing polarity shifts.

Domain-Specific considerations for the datasets and the Labeling Process

The review emphasizes the significance of domain-specific considerations in shaping effective sentiment analysis models for managing polarity shifts. The nature and characteristics of the dataset play a pivotal role in the adaptability and performance of sentiment analysis models. The review identifies a growing trend toward the utilization of domain-specific datasets that capture the intricacies of sentiment expressions within industries or topics. These datasets are designed to accommodate the nuances and unique linguistic features associated with specific domains, enabling sentiment analysis models to exhibit domain-aware behavior. This recognition of domain specificity contributes to the dynamism of the models, enhancing their ability to effectively manage polarity shifts within contextually distinct domains. In the realm of sentiment analysis models, the labeling process is a critical factor influencing their effectiveness in managing polarity shifts. The systematic review reveals a shift towards employing more meticulous and context-aware labeling processes. There is a growing emphasis on ensuring that labels accurately represent the sentiment expressed within the given context, particularly considering the dynamic nature of polarity shifts.

Domain experts and context-aware labeling strategies are increasingly employed to address the challenges posed by evolving language trends and changing sentiment dynamics. The consideration of multiple aspects within the domain during the labeling process further enhances the richness and relevance of labeled datasets, allowing sentiment analysis models to better navigate and manage polarity shifts within diverse contexts. In conclusion, domain-specific considerations for datasets and the meticulous labeling process are integral components in the development of effective sentiment analysis models for managing polarity shifts. Acknowledging the unique linguistic characteristics of specific domains and ensuring accurate contextual labeling contribute to the models' adaptability and robustness, enabling them to effectively capture and navigate polarity shifts within varied and nuanced contexts.

Performance Evaluation Metrics

The systematic review delves into the critical aspect of performance evaluation metrics, offering insights into the methodologies employed to assess the efficacy of sentiment analysis models in managing polarity shifts. The selection of appropriate metrics is paramount in gauging the models' effectiveness in capturing nuanced sentiment variations across diverse contexts. The review identifies several commonly used evaluation metrics, showcasing their relevance in assessing different aspects of sentiment analysis models. Accuracy, precision, recall, and F1-score are identified as the traditional metrics that measure the overall correctness and balance of sentiment classifications. However, the review emphasizes the importance of metrics tailored to polarity shift management. Cohens's Kappa and Matthew's Correlation Coefficient are highlighted for their ability to account for imbalanced datasets and measure the models' agreement with ground truth labels, providing a more comprehensive evaluation. Acknowledging the influence of domain specificity, the review notes the adoption of domain-centric evaluation metrics.

Domain-weighted sentiment Analysis Metrics and Domain-Adapted F1-score are examples of metrics tailored to assess model performance within specific domains, reflecting the need for nuanced evaluations aligned with diverse application areas. The systematic review underscores the limited evidence regarding the effectiveness of sentiment analysis models in handling implicit negation. As a response, there is a call for the development and adoption of metrics that specifically evaluate models' performance in recognizing and managing sentiments affected by implicit negations, thereby addressing a notable gap in current evaluation practices. In summary, the systematic review identifies a diverse set of performance evaluation metrics tailored to assess sentiment analysis models in the context of polarity shift management. The incorporation of temporal, domain-specific, and implicit negation metrics reflects the evolving landscape of sentiment analysis evaluation, providing a nuanced understanding of models' adaptability and effectiveness in capturing the complexities of sentiment changes.

Model's Generalizability, Embedding, and Representation Learning

Lastly, the review underscores the pivotal role of model generalizability in the effective management of polarity shifts within sentiment analysis. Generalizability refers to a model's ability to perform well across diverse datasets and real-world scenarios. The review reveals a trend towards developing sentiment analysis models that exhibit a higher degree of generalizability, enabling them to adapt to varying linguistic patterns, contexts, and domains. The integration of domain-specific considerations, robust feature selection techniques, and diverse datasets contribute to the enhanced generalizability of sentiment analysis models, allowing them to navigate and manage polarity shifts in a broader spectrum of applications. The systematic review recognizes the evolving landscape of sentiment analysis models with a distinct focus on embedding and representation learning techniques.

Word embeddings, such as Word2Vec, GloVe, and FastText, remain foundational in capturing semantic relationships among words, fostering a deeper understanding of sentiment nuances. The review highlights the increasing prevalence of contextual embeddings, with BERT emerging as a powerful representation learning model. These embeddings go beyond traditional word-level analysis, considering the contextual dependencies within sentences and documents. The adoption of such advanced embedding techniques enhances the models' ability to capture subtle changes in sentiment expressions associated with polarity shifts, contributing to their adaptability and effectiveness. The systematic review identifies a noteworthy trend in leveraging transfer learning as a mechanism to enhance generalizability. Pre-trained models, especially those based on deep learning architectures like BERT, are fine-tuned for sentiment analysis tasks. This transfer learning approach allows models to leverage knowledge gained from large, diverse datasets, improving their adaptability to specific sentiment analysis tasks and boosting generalizability across various domains. Despite advancements, challenges exist in achieving optimal generalizability. Implicit negation, domain-specific linguistic nuances, and dynamic shifts in sentiment expressions pose challenges that necessitate continued research and innovation. Future directions include exploring innovative representation learning techniques, refining transfer learning strategies, and addressing the adaptability of models to emerging language trends.

Additionally, investigating the effectiveness of sentiment analysis models in multilingual settings contributes to the broader goal of enhancing generalizability. In summary, the systematic review unveils a significant emphasis on enhancing model generalizability in the context of sentiment analysis for polarity shift management. Advanced embedding and representation learning techniques, coupled with transfer learning strategies, contribute to models' adaptability and effectiveness in capturing nuanced sentiment variations across diverse linguistic contexts and domains.

CONCLUSION AND FUTURE WORK

This review was tailored to sentiment analysis models used for handling polarity shift problems. Kitchenhams was used as the methodological framework because of its rigorous nature. The diverse landscape of these models included the traditional machine learning models, as well as the state-of-the-art deep learning models. Trends in feature selection, extraction, and data augmentation techniques underscored the importance of adaptability, context awareness, and more resilient models with effective capability for handling implicit negations. Negation was the most significant negation shifter.

Data sets were domain-specific and annotated with negations with several aspects. A review of evaluation metrics underscored the importance of using multiple metrics for comprehensive assessment as well as a call for more refined metrics tailored to implicit negation. These findings collectively contributed to shaping both the practical and theoretical implications. Among these implications, include an empirical investigation on the effectiveness of these models in handling the implicit negation besides explicit negation. This also involves development of the state-of-the-art hybrid models, ensemble models, and data augmentation strategies.

THEORETICAL AND PRACTICAL IMPLICATIONS

This review gives rise to several theoretical and practical implications. Firstly, practitioners gain variable guidance in choosing models based on the context. Secondly, the emphasis on generalizability enhances the applicability of models in a variety of contexts. Thirdly, the identification of the key challenge of handling the implicit negation provides practitioners with a road map for advanced research. Such advanced research incorporates empirical investigation. Lastly, the insights into the handling of implicit negation through leveraging on transfer learning strategies broaden the conceptual framework for this specific sentiment analysis research.

ACKNOWLEDGEMENT

The researcher is indebted to my Ph.D. supervisors and the entire faculty members for their tireless guidance in writing this research work, my spouse, children, and my family at large for their financial and moral support.

FUNDING

The study did not receive funding from any institution.

DECLARATION

Conflict of Interest

The author declared that there is no conflict of interest.

Inform Consent

This may not be applicable because this is a review article, and respondents are not involved.

Ethics Approval

It is not applicable because this is a review article, and no respondents are required.

REFERENCES

Abdi, A., Shamsuddin, S. M., Hasan, S., & Piran, J. (2019). Deep learning-based sentiment classification of evaluative text based on Multi-feature fusion. *Information Processing & Management*, 56(4), 1245–1259. https://doi.org/10.1016/j.ipm.2019.02.018

Abirami, A. M., & Gayathri, V. (2017). A survey on sentiment analysis methods and approach. *2016 Eighth International Conference on Advanced Computing (ICoAC)*, 72–76. https://doi.org/10.1109/ICoAC.2017.7951748

Ahmad, M., & Aftab, S. (2017). Analyzing the Performance of SVM for Polarity Detection with Different Datasets. *International Journal of Modern Education and Computer Science*, 9(10), 29–36. https://doi.org/10.5815/ijmecs.2017.10.04

Ahuja, R., Chug, A., Kohli, S., Gupta, S., & Ahuja, P. (2019). The Impact of Features Extraction on the Sentiment Analysis. *Procedia Computer Science*, *152*, 341–348. https://doi.org/10.1016/j.procs.2019.05.008

Akbarian, F., & Boroujeni, F. Z. (2020). An Improved Feature Selection Method for Sentiments Analysis in Social Networks. *2020 10th International Conference on Computer and Knowledge Engineering (ICCKE)*, 181–186. https://doi.org/10.1109/ICCKE50421.2020.9303710

Alzami, F., Udayanti, E. D., Prabowo, D. P., & Megantara, R. A. (2020). Document Preprocessing with TF-IDF to Improve the Polarity Classification Performance of Unstructured Sentiment Analysis. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 235–242. https://doi.org/10.22219/kinetik.v5i3.1066

Arora, M., & Kansal, V. (2019). Character level embedding with deep convolutional neural network for text normalization of unstructured data for Twitter sentiment analysis.

Social Network Analysis and Mining, *9*(1), 12. https://doi.org/10.1007/s13278-019-0557-y
Avinash, M., & Sivasankar, E. (2019). A Study of Feature Extraction Techniques for Sentiment Analysis. In A. Abraham, P. Dutta, J. K. Mandal, A. Bhattacharya, & S. Dutta (Eds.), *Emerging Technologies in Data Mining and Information Security* (Vol. 814, pp. 475–486). Springer Singapore. https://doi.org/10.1007/978-981-13-1501-5_41
Barnes, J., Velldal, E., & Øvrelid, L. (2021). Improving sentiment analysis with multi-task learning of negation. *Natural Language Engineering*, *27*(2), 249–269. https://doi.org/10.1017/S1351324920000510
Bommert, A., Sun, X., Bischl, B., Rahnenführer, J., & Lang, M. (2020). Benchmark for filter methods for feature selection in high-dimensional classification data. *Computational Statistics & Data Analysis*, *143*, 106839. https://doi.org/10.1016/j.csda.2019.106839
Chen, L. (2019). Attention-Based Deep Learning System for Negation and Assertion Detection in Clinical Notes. *International Journal of Artificial Intelligence & Applications*, *10*(01), 1–9. https://doi.org/10.5121/ijaia.2019.10101
Cruz, N. P., Taboada, M., & Mitkov, R. (2016). A machine-learning approach to negation and speculation detection for sentiment analysis. *Journal of the Association for Information Science and Technology*, *67*(9), 2118–2136. https://doi.org/10.1002/asi.23533
Cunha, E., Magno, G., Caetano, J., Teixeira, D., & Almeida, V. (2018). Fake News as We Feel It: Perception and Conceptualization of the Term "Fake News" in the Media. In S. Staab, O. Koltsova, & D. I. Ignatov (Eds.), *Social Informatics* (Vol. 11185, pp. 151–166). Springer International Publishing. https://doi.org/10.1007/978-3-030-01129-1_10
Davis, W. A. (2016). *Irregular Negatives, Implicatures, and Idioms* (Vol. 6). Springer Netherlands. https://doi.org/10.1007/978-94-017-7546-5
Deniz, A., Angin, M., & Angin, P. (2021). Evolutionary Multiobjective Feature Selection for Sentiment Analysis. *IEEE Access*, *9*, 142982–142996. https://doi.org/10.1109/ACCESS.2021.3118961
Eke, C. I., Norman, A. A., & Shuib, L. (2021). Context-Based Feature Technique for Sarcasm Identification in Benchmark Datasets Using Deep Learning and BERT Model. *IEEE Access*, *9*, 48501–48518. https://doi.org/10.1109/ACCESS.2021.3068323
Elangovan, D., & Subedha, V. (2020). An Effective Feature Selection Based Classification model using Firefly with Levy and Multilayer Perceptron based Sentiment Analysis. *2020 International Conference on Inventive Computation Technologies (ICICT)*, 376–380. https://doi.org/10.1109/ICICT48043.2020.9112425
Falasari, A., & Muslim, M. A. (2022). Optimize Naïve Bayes Classifier Using Chi Square and Term Frequency Inverse Document Frequency For Amazon Review Sentiment Analysis. *Journal of Soft Computing Exploration*, *3*(1), 31–36. https://doi.org/10.52465/joscex.v3i1.68
Farooq, U. (2017). Negation Handling in Sentiment Analysis at Sentence Level. *Journal of Computers*, 470–478. https://doi.org/10.17706/jcp.12.5.470-478

Fei, H., Ren, Y., & Ji, D. (2020). Negation and speculation scope detection using recursive neural conditional random fields. *Neurocomputing*, *374*, 22–29. https://doi.org/10.1016/j.neucom.2019.09.058

Ghosh, M., Guha, R., Sarkar, R., & Abraham, A. (2020). A wrapper-filter feature selection technique based on ant colony optimization. *Neural Computing and Applications*, *32*(12), 7839–7857. https://doi.org/10.1007/s00521-019-04171-3

Gokalp, O., Tasci, E., & Ugur, A. (2020). A novel wrapper feature selection algorithm based on iterated greedy metaheuristic for sentiment classification. *Expert Systems with Applications*, *146*, 113176. https://doi.org/10.1016/j.eswa.2020.113176

Gorai, J., & Shaw, D. K. (2022). Two-phase feature extraction with chi square-based feature selection approach for the classification of online product reviews. *2022 7th International Conference on Communication and Electronics Systems (ICCES)*, 1433–1438. https://doi.org/10.1109/ICCES54183.2022.9835934

Goularas, D., & Kamis, S. (2019). Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data. *2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML)*, 12–17. https://doi.org/10.1109/Deep-ML.2019.00011

Hegde, R. & Seema S. (2017). Aspect based feature extraction and sentiment classification of review data sets using Incremental machine learning algorithm. *2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*, 122–125. https://doi.org/10.1109/AEEICB.2017.7972395

Irawaty, I., Andreswari, R., & Pramesti, D. (2020). Vectorizer Comparison for Sentiment Analysis on Social Media Youtube: A Case Study. *2020 3rd International Conference on Computer and Informatics Engineering (IC2IE)*, 69–74. https://doi.org/10.1109/IC2IE50715.2020.9274650

Japhne, A., & Murugeswari, R. (2020). Opinion Mining based complex polarity shift pattern handling for improved sentiment classification. *2020 International Conference on Inventive Computation Technologies (ICICT)*, 323–329. https://doi.org/10.1109/ICICT48043.2020.9112565

Kafi, A., Ashikul Alam, M. S., Bin Hossain, S., Awal, S. B., & Arif, H. (2019). Feature-Based Mobile Phone Rating Using Sentiment Analysis and Machine Learning Approaches. *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, 1–6. https://doi.org/10.1109/ICASERT.2019.8934555

Kalaivani, K. S., Uma, S., & Kanimozhiselvi, C. S. (2020). A Review on Feature Extraction Techniques for Sentiment Classification. *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, 679–683. https://doi.org/10.1109/ICCMC48092.2020.ICCMC-000126

Kamal, L. H., McKee, G. T., & Othman, N. A. (2022). Naïve Bayes with Negation Handling for Sentiment Analysis of Twitter Data. *2022 9th International Conference on Soft Computing & Machine Intelligence (ISCMI)*, 207–212. https://doi.org/10.1109/ISCMI56532.2022.10068474

Kaur, S., Sikka, G., & Awasthi, L. K. (2018). Sentiment Analysis Approach Based on N-gram and KNN Classifier. *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, 1–4. https://doi.org/10.1109/ICSCCC.2018.8703350
Khandelwal, A., & Sawant, S. (2019). *NegBERT: A Transfer Learning Approach for Negation Detection and Scope Resolution*. https://doi.org/10.48550/ARXIV.1911.04211
Kim, K. (2018). An improved semi-supervised dimensionality reduction using feature weighting: Application to sentiment analysis. *Expert Systems with Applications*, *109*, 49–65. https://doi.org/10.1016/j.eswa.2018.05.023
Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, *51*(1), 7–15. https://doi.org/10.1016/j.infsof.2008.09.009
Koncar, P., Fuchs, A., Hobisch, E., Geiger, B. C., Scholger, M., & Helic, D. (2020). Text sentiment in the Age of Enlightenment: An analysis of spectator periodicals. *Applied Network Science*, 5(1), 33. https://doi.org/10.1007/s41109-020-00269-z
Kristiyanti, D. A., Normah, & Umam, A. H. (2019). Prediction of Indonesia Presidential Election Results for the 2019-2024 Period Using Twitter Sentiment Analysis. *2019 5th International Conference on New Media Studies (CONMEDIA)*, 36–42. https://doi.org/10.1109/CONMEDIA46929.2019.8981823
Kumar, A., & Garg, G. (2023). Empirical study of shallow and deep learning models for sarcasm detection using context in benchmark datasets. *Journal of Ambient Intelligence and Humanized Computing*, *14*(5), 5327–5342. https://doi.org/10.1007/s12652-019-01419-7
Kumar, H. M. K., & Harish, B. S. (2018). Sarcasm classification: A novel approach by using Content Based Feature Selection Method. *Procedia Computer Science*, *143*, 378–386. https://doi.org/10.1016/j.procs.2018.10.409
Kumar, S., & Pathak, S. K. (2021). A Novel Approach for Sentiment Analysis and Optimization Using PSO. *2021 International Conference on Advances in Technology, Management & Education (ICATME)*, 43–47. https://doi.org/10.1109/ICATME50232.2021.9732752
Kumar, S. S., & Rajini, A. (2019). Extensive Survey on Feature Extraction and Feature Selection Techniques for Sentiment Classification in Social Media. *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1–6. https://doi.org/10.1109/ICCCNT45670.2019.8944391
Kurniawan, D., Yasir, M., & Venna, F. C. (2022). Optimization of Sentiment Analysis using Naive Bayes with Features Selection Chi-Square and Information Gain for Accuracy Improvement. *2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 153–160. https://doi.org/10.23919/EECSI56542.2022.9946510
Lazib, L., Zhao, Y., Qin, B., & Liu, T. (2019). Negation scope detection with recurrent neural networks models in review texts. *International Journal of High Performance Computing and Networking*, *13*(2), 211. https://doi.org/10.1504/IJHPCN.2019.097501

Liu, B. (2020). Text sentiment analysis based on CBOW model and deep learning in big data environment. *Journal of Ambient Intelligence and Humanized Computing*, *11*(2), 451–458. https://doi.org/10.1007/s12652-018-1095-6

Li, Y., Kim, Y., Park, H., Geller, T., & Panda, P. (2022, October). Neuromorphic data augmentation for training spiking neural networks. In European Conference on Computer Vision (pp. 631-649). Cham: Springer Nature Switzerland.

Manek, A. S., Shenoy, P. D., Mohan, M. C., & R, V. K. (2017). Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier. *World Wide Web*, *20*(2), 135–154. https://doi.org/10.1007/s11280-015-0381-x

Mathapati, S., Nafeesa, A., Tanuja, R., Manjula, S. H., & Venugopal, K. R. (2019). Semi-supervised domain adaptation and collaborative deep learning for dual sentiment analysis. *SN Applied Sciences*, *1*(8), 907. https://doi.org/10.1007/s42452-019-0943-0

Mee, A., Homapour, E., Chiclana, F., & Engel, O. (2021). Sentiment analysis using TF–IDF weighting of UK MPs' tweets on Brexit. *Knowledge-Based Systems*, *228*, 107238. https://doi.org/10.1016/j.knosys.2021.107238

Mhatre, M., Phondekar, D., Kadam, P., Chawathe, A., & Ghag, K. (2017). Dimensionality reduction for sentiment analysis using pre-processing techniques. *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, 16–21. https://doi.org/10.1109/ICCMC.2017.8282676

Miao, J., & Niu, L. (2016). A Survey on Feature Selection. *Procedia Computer Science*, *91*, 919–926. https://doi.org/10.1016/j.procs.2016.07.111

Mohd Nafis, N. S., & Awang, S. (2021). An Enhanced Hybrid Feature Selection Technique Using Term Frequency-Inverse Document Frequency and Support Vector Machine-Recursive Feature Elimination for Sentiment Classification. *IEEE Access*, *9*, 52177–52192. https://doi.org/10.1109/ACCESS.2021.3069001

Montenegro, O., Pabon, O. S., & De Pinerez R., R. E. G. (2021). A Deep Learning Approach for Negation Detection from Product Reviews written in Spanish. *2021 XLVII Latin American Computing Conference (CLEI)*, 1–6. https://doi.org/10.1109/CLEI53233.2021.9640190

Mutinda, J., Mwangi, W., & Okeyo, G. (2021). Lexicon-pointed hybrid N-gram Features Extraction Model (LeNFEM) for sentence level sentiment analysis. *Engineering Reports*, *3*(8), e12374. https://doi.org/10.1002/eng2.12374

Nurhayati, Putra, A. E., Wardhani, L. K., & Busman. (2019). Chi-Square Feature Selection Effect On Naive Bayes Classifier Algorithm Performance For Sentiment Analysis Document. *2019 7th International Conference on Cyber and IT Service Management (CITSM)*, 1–7. https://doi.org/10.1109/CITSM47753.2019.8965332

Oscar Deho, B., William Agangiba, A., Felix Aryeh, L., & Jeffery Ansah, A. (2018). Sentiment Analysis with Word Embedding. *2018 IEEE 7th International Conference on Adaptive Science & Technology (ICAST)*, 1–4. https://doi.org/10.1109/ICASTECH.2018.8506717

Paudel, S., Prasad, P. W. C., Alsadoon, A., Islam, Md. R., & Elchouemi, A. (2019). Feature Selection Approach for Twitter Sentiment Analysis and Text Classification Based on Chi-Square and Naïve Bayes. In J. Abawajy, K.-K. R. Choo, R. Islam, Z. Xu, & M. Atiquzzaman (Eds.), *International Conference on Applications and Techniques in*

Cyber Security and Intelligence ATCI 2018 (Vol. 842, pp. 281–298). Springer International Publishing. https://doi.org/10.1007/978-3-319-98776-7_30

Pongthanoo, P., & Songpan, W. (2020). Feature Selection and Reduction Based on SMOTE and Information Gain for Sentiment Mining. *2020 5th International Conference on Computer and Communication Systems (ICCCS)*, 109–114. https://doi.org/10.1109/ICCCS49078.2020.9118467

Poria, S., Cambria, E., Hazarika, D., & Vij, P. (2016). *A Deeper Look into Sarcastic Tweets Using Deep Convolutional Neural Networks*. https://doi.org/10.48550/ARXIV.1610.08815

Pradha, S., Halgamuge, M. N., & Tran Quoc Vinh, N. (2019). Effective Text Data Preprocessing Technique for Sentiment Analysis in Social Media Data. *2019 11th International Conference on Knowledge and Systems Engineering (KSE)*, 1–8. https://doi.org/10.1109/KSE.2019.8919368

Prastyo, P. H., Ardiyanto, I., & Hidayat, R. (2020). A Review of Feature Selection Techniques in Sentiment Analysis Using Filter, Wrapper, or Hybrid Methods. *2020 6th International Conference on Science and Technology (ICST)*, 1–6. https://doi.org/10.1109/ICST50505.2020.9732885

Pröllochs, N., Feuerriegel, S., Lutz, B., & Neumann, D. (2020). Negation scope detection for sentiment analysis: A reinforcement learning framework for replicating human interpretations. *Information Sciences*, *536*, 205–221. https://doi.org/10.1016/j.ins.2020.05.022

Rajabi, Z., Valavi, M. R., & Hourali, M. (2020). A Context-Based Disambiguation Model for Sentiment Concepts Using a Bag-of-Concepts Approach. *Cognitive Computation*, *12*(6), 1299–1312. https://doi.org/10.1007/s12559-020-09729-1

Rajat, R., Jaroli, P., Kumar, N., & Kaushal, R. K. (2021). A Sentiment Analysis of Amazon Review Data Using Machine Learning Model. *2021 6th International Conference on Innovative Technology in Intelligent System and Industrial Applications (CITISIA)*, 1–6. https://doi.org/10.1109/CITISIA53721.2021.9719909

Rasool, A., Tao, R., Kamyab, M., & Hayat, S. (2020). GAWA–A Feature Selection Method for Hybrid Sentiment Classification. *IEEE Access*, *8*, 191850–191861. https://doi.org/10.1109/ACCESS.2020.3030642

Raunak, V., Gupta, V., & Metze, F. (2019). Effective Dimensionality Reduction for Word Embeddings. *Proceedings of the 4th Workshop on Representation Learning for NLP (RepL4NLP-2019)*, 235–243. https://doi.org/10.18653/v1/W19-4328

Rezaeinia, S. M., Rahmani, R., Ghodsi, A., & Veisi, H. (2019). Sentiment analysis based on improved pre-trained word embeddings. *Expert Systems with Applications*, *117*, 139–147. https://doi.org/10.1016/j.eswa.2018.08.044

Shahnawaz, & Astya, P. (2017). Sentiment analysis: Approaches and open issues. *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 154–158. https://doi.org/10.1109/CCAA.2017.8229791

Shang, L., Zhou, Z., & Liu, X. (2016). Particle swarm optimization-based feature selection in sentiment classification. *Soft Computing*, *20*(10), 3821–3834. https://doi.org/10.1007/s00500-016-2093-2

Singh, P. K., & Paul, S. (2021). Deep Learning Approach for Negation Handling in Sentiment Analysis. *IEEE Access*, *9*, 102579–102592. https://doi.org/10.1109/ACCESS.2021.3095412

Suchetha, N. K., Nikhil, A., & Hrudya, P. (2019). Comparing the Wrapper Feature Selection Evaluators on Twitter Sentiment Classification. *2019 International Conference on Computational Intelligence in Data Science (ICCIDS)*, 1–6. https://doi.org/10.1109/ICCIDS.2019.8862033

Thevar, K., Thakur, V., Shukla, T., Vishwakarma, D., & Ghag, K. (2017). Handling polarity shift for opinion mining. *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, 1006–1011. https://doi.org/10.1109/ICCMC.2017.8282621

Tubishat, M., Abushariah, M. A. M., Idris, N., & Aljarah, I. (2019). Improved whale optimization algorithm for feature selection in Arabic sentiment analysis. *Applied Intelligence*, *49*(5), 1688–1707. https://doi.org/10.1007/s10489-018-1334-8

Tubishat, M., Idris, N., Shuib, L., Abushariah, M. A. M., & Mirjalili, S. (2020). Improved Salp Swarm Algorithm based on opposition based learning and novel local search algorithm for feature selection. *Expert Systems with Applications*, *145*, 113122. https://doi.org/10.1016/j.eswa.2019.113122

Wang, S., Lv, G., Mazumder, S., & Liu, B. (2020). *Detecting Domain Polarity-Changes of Words in a Sentiment Lexicon*. https://doi.org/10.48550/ARXIV.2004.14357

Wang, X., Xue, S., Liu, J., Zhang, J., Wang, J., & Zhou, J. (2023, August). Sentiment Classification Based on RoBERTa and Data Augmentation. In 2023 IEEE 9th International Conference on Cloud Computing and Intelligent Systems (CCIS) (pp. 260-264). IEEE.

Xia, H., Ding, C., & Liu, Y. (2020). Sentiment Analysis Model Based on Self-Attention and Character-Level Embedding. *IEEE Access*, *8*, 184614–184620. https://doi.org/10.1109/ACCESS.2020.3029694

Xia, R., Xu, F., Yu, J., Qi, Y., & Cambria, E. (2016). Polarity shift detection, elimination and ensemble: A three-stage model for document-level sentiment analysis. *Information Processing & Management*, *52*(1), 36–45. https://doi.org/10.1016/j.ipm.2015.04.003

Yadav, M., & Katarya, R. (2020). Polarity Shift Detection and Handling in Polarity Classification: A Survey. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3562967

Zirpe, S., & Joglekar, B. (2017a). Negation Handling using Stacking Ensemble Method. *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, 1–5. https://doi.org/10.1109/ICCUBEA.2017.8463946

Zirpe, S., & Joglekar, B. (2017b). Polarity shift detection approaches in sentiment analysis: A survey. *2017 International Conference on Inventive Systems and Control (ICISC)*, 1–5. https://doi.org/10.1109/ICISC.2017.8068737

Author's Biography

Millicent Kathambi Murithi is a Tutorial Fellow at the Computer Science Department, Murang'a University of Technology, Kenya. She holds an MSc. Degree in Computer Systems from Jomo Kenyatta University of Science and Technology, Kenya. Her research interests include Machine Learning, software engineering, and Natural Language Processing.

Aaron Mogeni Oirere is a Lecturer at the Department of Computer Science, Murang'a University of Technology, Kenya. He holds a Ph.D. degree in Computer Science from Dr. Babasaheb Ambedkar Marathwada University, Maharashtra, India. His research interests include Automatic Speech Recognition, Human-computer Interaction, Information Retrieval, Database Management Systems (DBMS), Data Analytics and Hardware & Networking.

Rachael Njeri Ndung'u is a Lecturer at the Department of Information Technology, Murang'a University of Technology, Kenya. She holds a Ph.D. degree in Information Technology. Her research interests include Artificial Intelligence, Data Analytics ,and Blockchain.