A Novel Framework for Sentiment Analysis: Dimensionality Reduction for Machine Learning (DRML)

Dhamayanthi N., Lavanya B*
Department of Computer Science, University of Madras, Chennai, India

Abstract*—Sentiment analysis is vital for understanding public opinion, but improving its performance is challenging due to the complexities of high-dimensional text data and diverse user-generated content. We propose a novel framework based on Dimensionality Reduction for Machine Learning (DRML) that enhances the classification performance by 21.55% while reducing the dimension of the feature matrix by 99.63%. Our research addresses the fundamental question of whether it is possible to reduce the feature space significantly while improving sentiment analysis performance. Our approach employs Principal Component Analysis (PCA) to effectively capture essential textual features and includes the development of an algorithm for identifying principal components from positive and negative reviews. We then create a supervised dataset by combining these components. Furthermore, we integrate a range of state-of-the-art machine learning algorithms (Decision Tree, K-Nearest Neighbours, Bernoulli Naïve Bayes, and Majority Voting Ensemble) into our framework, along with a custom tokenizer, to harness the full potential of reduced-dimensional data for sentiment classification. We have conducted extensive experiments using gold standard multi-domain benchmark datasets from Amazon to show that DRML outperforms other state-of-the-art approaches. Our proposed methodology gives superior performance with an average performance of 98.38% which is a significant increase in performance by 21.55% compared to the baseline methodology using Bag of Words (BoW). In terms of individual evaluation parameters, DRML shows an increase of 21.84% in Accuracy, 20.4% in Precision, 21.84% in Recall, and 22.11% in F1-score. In comparison with the state-of-the-art (SOTA) methodologies applied to the same benchmark dataset in recent years, our framework demonstrates a significant average increase in Accuracy for Sentiment Analysis by 10.96%. This substantial improvement underscores the effectiveness of our approach. To conclude, our research contributes to the field of sentiment analysis by introducing an innovative framework that not only improves the efficiency of sentiment analysis but also paves the way for the analysis of extensive textual data in diverse real-world applications.

Keywords—Machine learning; text mining; natural language processing; sentiment analysis; opinion classification

I. Introduction

In the era of information abundance, understanding and extracting valuable insights from the vast amount of text data available on the internet has become paramount. Sentiment analysis, a critical component of natural language processing, plays a pivotal role in this endeavour. Sentiment analysis also referred to as opinion mining, employs computational techniques to identify and categorize subjective information present in textual data. The field of sentiment analysis has witnessed a rapid evolution, driven by the relentless efforts of researchers to enhance the accuracy, adaptability, and efficiency of sentiment classification methods. Sentiment analysis finds practical utility across various domains, spanning market research, social media monitoring, customer feedback analysis, and even the development of political campaign strategies. By discerning and quantifying sentiment from text data, organizations and individuals can make informed decisions, refine products and services, track public perception, and craft effective communication strategies. In light of this dynamic and multifaceted research landscape, our work embarks on a novel framework for sentiment analysis, leveraging dimensionality reduction techniques and machine learning algorithms to further improve accuracy, adaptability, and efficiency. This research builds on the foundational knowledge amassed by previous works and addresses the identified challenges and opportunities in the field, positioning itself as a valuable contribution to the ever-evolving landscape of sentiment analysis.

Ensemble learning, on the other hand, offers a promising solution. By combining the outputs of individual classifiers in a manner that compensates for each other's weaknesses, ensemble techniques can enhance the overall classification scheme's robustness and predictive accuracy. Thus, the integration of feature extraction and ensemble techniques in this research paper is motivated by the need to overcome the intrinsic challenges of sentiment analysis, including high-dimensional data, variable review characteristics, and computational complexity. By doing so, this study seeks to pave the way for a more effective and efficient sentiment analysis framework, ultimately contributing to improved sentiment polarity determination.

A. Research Objective

The primary objective of our research is to investigate whether it is feasible to achieve a significant reduction in dimensionality while simultaneously enhancing the performance of sentiment analysis. This objective is guided by the fundamental question: "Can we achieve a substantial reduction in dimensionality while simultaneously improving the performance of sentiment analysis?" In pursuit of this objective, we aim to address the critical question of whether it is possible to streamline the feature space used in sentiment analysis, thereby making the analysis more efficient, without compromising the quality of sentiment analysis results. The goal is not merely to reduce dimensionality but to do so without sacrificing the accuracy and reliability of sentiment analysis outcomes.

Through this investigation, our research aspires to provide valuable insights into the field of sentiment analysis, especially in the context of practical applications that involve large and complex datasets. Our aim is to offer solutions and methodologies that empower the analysis of extensive textual data, ensuring that sentiment analysis remains accurate and efficient even in real-world scenarios with substantial data volumes. In summary, our research objectives revolve around the dual goal of dimensionality reduction and performance enhancement, with the ultimate aim of delivering practical and valuable contributions to sentiment analysis, particularly in applications that rely on the analysis of large datasets.

Our main contributions are listed below:

- We introduce an innovative framework, Dimensionality Reduction for Machine Learning (DRML), which yields a remarkable 21.55% enhancement in classification performance while reducing the feature matrix's dimension by an impressive 99.63%.
- We employed Principal Component Analysis (PCA) to effectively capture the essential features of the review text. We devised an algorithm to identify principal components from positive reviews and negative reviews separately and then prepared a supervised dataset with a mix of these components.
- We conducted several experiments with varying principal components and found that PCA with 50 components is ideal for high performance. This reduces the dimension of the feature matrix by 99.63%.
- We established a baseline using the BOW (Bag of Words) methodology and performed a comprehensive experimental analysis to compare it with our proposed methodology. This evaluation utilized three gold standard multi-domain benchmark datasets from Amazon. For each benchmark dataset, 4 classifiers are trained, and their performance is compared with the performance of DRML which gives a superior performance of 99.38% than the baseline with an increase of 21.84% in Accuracy, 20.4% in Precision, 21.84% in Recall and 22.11% in F1-score.
- In comparison with state-of-the-art (SOTA) methodologies applied to the same benchmark dataset in recent years, our framework demonstrates a significant average increase in accuracy for sentiment analysis by 10.96%. This substantial improvement underscores the effectiveness of our approach.

The structure of this paper is as follows: Section II provides a literature survey, summarizing existing work in sentiment analysis, dimensionality reduction, and machine learning. Section III details our methodology, covering both the baseline approach and our proposed Dimensionality Reduction for Machine Learning (DRML) method, including a visual framework representation and pseudocode, as well as dataset descriptions. Section IV presents the results and provides an in-depth discussion of our findings. Finally, Section V offers our conclusions, summarizes key contributions, and outlines potential directions for future research in this domain.

II. Literature Survey

Sentiment analysis, a pivotal component of natural language processing, has evolved significantly, with researchers continuously striving to enhance the accuracy and adaptability of sentiment classification methods. In the age of digital reviews, sentiment analysis is crucial for categorizing customer reviews [47]. User reviews, especially in e-commerce and social media, have become increasingly significant. Semantic features like sentence level features (SLF) and domain-sensitive features (DSF) are used to improve supervised sentiment analysis, leading to favourable performance gains [33]. Research on Aspect-Based Sentiment Analysis (ABSA) focuses on inferring sentiment with respect to a certain aspect [5, 15, 20, 21, 39, 50, 51]. Researchers have worked on several sentiment classification models to enhance sentiment analysis performance [7, 17, 22, 38]. Sailunaz et al. [35] describe a novel approach for detecting sentiment and emotion in Twitter posts. Ligthart et al. [23] offer an effective means to identify and filter out spam content within online reviews and comments. The use of fuzzy logic in sentiment analysis has been explored by Serrano-Guerrero et al. [37] with an extensive review of its applications in opinion mining. Sivakumar et al. [39] focused on aspect-based fuzzy logic.

Cross-domain and Multimodal sentiment analysis, a challenging aspect of the field, has received considerable attention. Innovations such as hierarchical attentional networks, Topic Driven Adaptive Network, Hierarchical Attention-BiLSTM model and pre-trained language models have showcased the depth of current research into understanding sentiments in complex varied data sources [14, 45, 46, 49]. By extracting sentiment lexicons from domain-specific corpora using active learning strategies, researchers have achieved higher accuracy in sentiment classification [27]. Multilingual sentiment analysis is a significant focus, with tailored models, leading to improved accuracy [4, 9, 24, 36, 40]. Systematic reviews have shed light on the Arabic aspect-based sentiment analysis techniques and resources [5]. Additionally, feature-based sentiment analysis for Arabic addresses challenges posed by colloquial language and dialects [2].

Many researchers have used machine learning techniques for sentiment analysis [1, 31, 34, 41]. Additionally, sentiment analysis models have evolved beyond conventional Bag-of-Words (BOW) techniques. The dual sentiment analysis (DSA) model introduced a novel approach by incorporating sentiment-reversed reviews and dual training algorithms, enabling classification into three classes: "positive, negative, and neutral" [44]. Blending neural networks with sentiment lexicons reduces the need for extensive labelled data while adapting word polarities to the target domains [8]. Moreover, divide-and-conquer approaches have been introduced to sentence-level sentiment classification, improving sentiment classification by categorizing sentences based on the number of sentiment targets and employing distinct neural network models for sentiment analysis within each group [12].

Meanwhile, deep learning models have revolutionized sentiment analysis by significantly improving accuracy while reducing training time, particularly relevant in the era of digital reviews [18, 25, 30, 32, 42, 43]. Ensemble techniques and a classification model taxonomy have been introduced, enhancing classification accuracy and offering a more nuanced analysis of sentiment [6]. To tackle challenges related to dimensionality and feature importance, Onan [26] has introduced an architecture that incorporates a bidirectional convolutional recurrent neural network with group-wise enhancement. The use of deep learning models has led to high average recall values for in-domain and out-of-domain data, demonstrating scalability and effectiveness in large-scale topic modelling and sentiment analysis [28]. Heterogeneous ensemble techniques offer median performance gains across various domains, highlighting their efficiency in sentiment analysis tasks [19]. Ensemble methods, including the Voting ensemble method, Bagging, Boosting, and classifiers like Random Forest, and Bayesian Ensemble Learning have been recognized as powerful tools in sentiment analysis, achieving exceptional results in various scenarios [3, 13].

Feature selection techniques like Information Gain, Chi Square, and Gini Index have been instrumental in refining sentiment analysis, leading to substantial improvements in accuracy when thoughtfully combined and applied with classifiers like SMO [16]. Feature definition based on entropy and semantic context [29], Feature selection methods using novel term weighting [49] are being used for enhancing sentiment classification results. Moreover, Many researchers reviewed the challenges and opportunities in sentiment analysis research [10].

While these techniques have addressed challenges and limitations, the research on sentiment analysis continues to evolve paving the way for a more accurate, adaptable, and efficient field of study. Our novel framework, DRML (Dimensionality Reduction for Machine Learning) adds to the literature as a valuable contribution due to its impeccable performance compared to the baseline and state-of-the-art methodologies.

III. Methodology

A. Baseline Methodology for Evaluating DRML Framework

In order to assess the effectiveness of our novel framework, DRML, for sentiment analysis, we have implemented a baseline methodology, as depicted in Fig. 1.

1) Data processing and feature extraction: The initial step involves the extraction of text from customer reviews on Amazon. To accomplish this, we employed the BeautifulSoup XML parser. Subsequently, a comprehensive pre-processing procedure was applied before extracting tokens from the text. These tokens were segregated into two distinct arrays for positive and negative reviews. To facilitate further analysis, we constructed a dictionary for word-index mappings. Using this dictionary, we transformed the textual data into numerical format, generating feature vectors.

2) Sentiment Prediction: The resultant feature vectors were then used as input for various machine learning models to predict sentiment. The evaluation of our model's performance is based on key metrics, including Accuracy, Precision, Recall, and F1-score.

3) Baseline Model: Our baseline model employed a straightforward Bag-of-Words (BoW) approach to convert the textual data into numerical format. In this approach, the dimensionality of the feature vectors is equivalent to the size of the vocabulary.

B. Proposed Methodology

We introduce a novel sentiment analysis framework called "DRML" (Dimensionality Reduction for Machine Learning), designed to effectively analyze sentiment by harnessing dimensionality reduction techniques and machine learning algorithms. This section details the development and structure of our framework.

1) Framework overview: The block diagram in Fig. 2 along with the pseudocode in this section explains our framework. Our approach is rigorously evaluated using three widely recognized gold standard multi-domain benchmark datasets sourced from E-Commerce reviews on Amazon. We compare the performance of DRML against published results, emphasizing its effectiveness.

2) Framework structure: The DRML framework is composed of several key components, each contributing to its efficacy in sentiment analysis. The following breakdown illustrates the structural aspects of DRML:

a) Data Collection: In the data collection layer, we employ the BeautifulSoup XML parser to extract text from positive and negative reviews across various domains from Amazon. These reviews are stored in separate files, ensuring data integrity.

b) Text Pre-Processing: Our custom pre-processing module includes stemming, lemmatization, noise removal, stop word elimination, and domain-specific word removal. We then extract tokens from both positive and negative reviews and store them in distinct lists.

c) Feature Extraction: We compile a word index map to preserve all unique tokens. Using this map, we create N-dimensional feature vectors of size A x B, where A represents the number of reviews, and B signifies the length of the word index map.

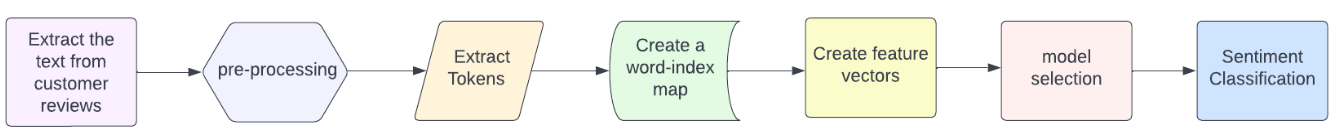

Fig. 1. Baseline methodology

d) Dimensionality Reduction: To reduce dimensionality effectively, we employ Principal Component Analysis (PCA) for feature extraction. Extensive experimentation is conducted with varying numbers of components (50, 100, 150, 200, 250) for both positive and negative feature vectors.

e) Feature Vector Transformation: The application of PCA results in a significant dimension reduction. For example, in the DVD dataset, the feature vector dimension is reduced from “2000 x 21344” to "2000 x 51" for 50 components. This compact representation facilitates efficient sentiment analysis.

f) Supervised Dataset Creation: We combine both the positive and negative feature vectors to create a supervised dataset. Afterwards, the dataset is randomly shuffled to mix positive and negative reviews.

g) Training and Classification: The dataset is divided into two parts, with 70% for training and 30% for testing. The model is trained using three distinct classifiers: Bernoulli Naïve Bayes, Decision Tree Classifier, and K-Nearest Neighbour. A Max Voting Ensemble classifier is then applied to these three classifiers.

h) Model Evaluation: The performance of the model is assessed using the 30% test dataset, with performance measures including Precision, Recall, Accuracy, and F1-score serving as evaluation metrics.

3) Framework Development: Our framework is developed in Python, with BeautifulSoup XML parser employed for data extraction. This comprehensive methodology ensures the accuracy and efficiency of sentiment analysis.

By employing DRML, our goal is to advance sentiment analysis techniques and provide more accurate and insightful results.

C. *Dataset Description*

For our research, we have chosen three gold standard benchmark datasets, each sourced from Amazon and initially published by Blitzer et al. [11]. These datasets are widely recognized and have been extensively used in the realm of sentiment analysis research. The datasets encompass product reviews gathered from three distinct domains: DVD, Electronics, and Kitchen. These domains were chosen to ensure diversity in the types of products and reviews included, contributing to the robustness and applicability of our analysis.

The fields in the datasets are depicted in Table I. In these datasets, customers have assigned ratings to the product reviews using a scale of 1 to 5 stars. Reviews receiving ratings of 4 or 5 stars are categorized as positive, reflecting favourable sentiment towards the products. Conversely, reviews receiving ratings of 1 or 2 stars are categorized as negative, indicating less favourable sentiment or dissatisfaction.

To maintain the integrity of our datasets and ensure the balance between positive and negative instances, we have thoughtfully selected 1000 positive reviews and 1000 negative reviews from each of the three domains. This rigorous selection process guarantees that our dataset is both comprehensive and representative, allowing for robust sentiment analysis.

In summary, our dataset consists of three Amazon benchmark datasets, encompassing reviews from three diverse domains. The reviews are categorized into positive and negative sentiments based on customer ratings, and the dataset is carefully balanced to support our sentiment analysis research effectively.

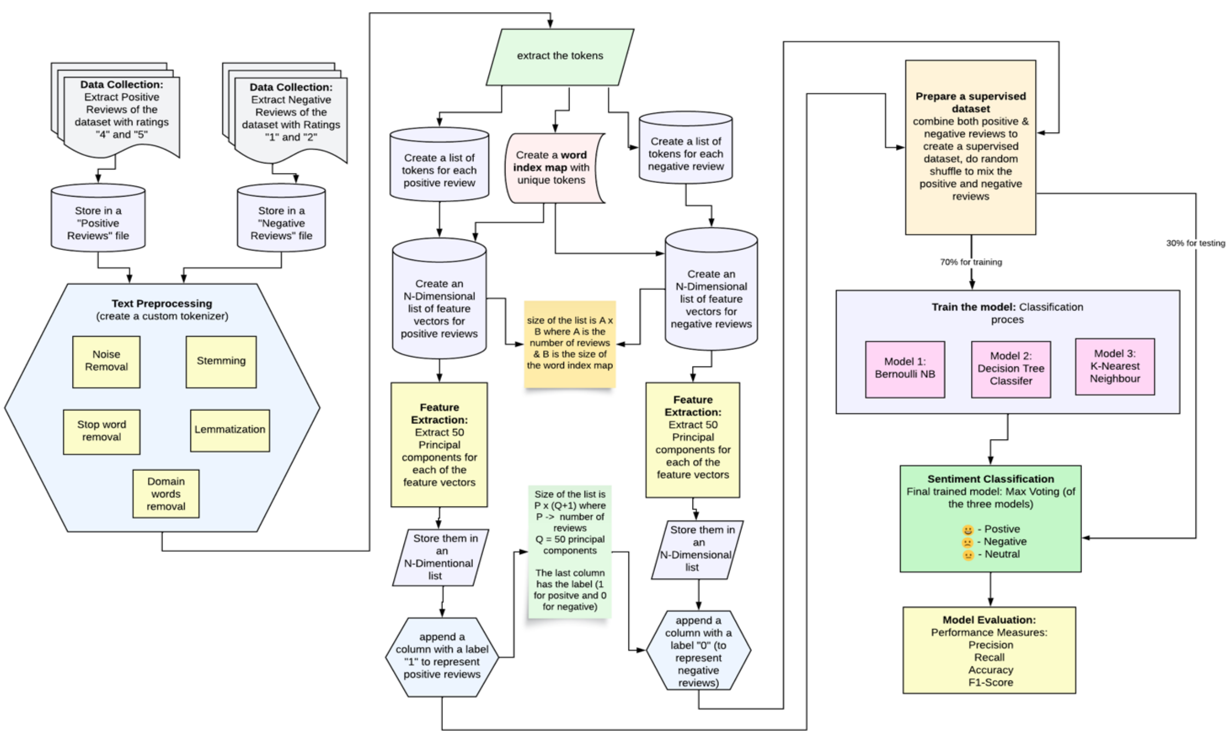

Fig. 2. Proposed methodology.

The detailed process is explained in the pseudocode below:

```
Input:                        Amazon Multi domain datasets: DS = {DVD, Electronics, Kitchenware}
Input:                        Machine Learning Algorithms:  ML_Alg =  {Decision_Tree, KNN, BernoulliNB, Max_voting}
Dimensionality Reduction:             Principal Component Analysis (PCA)
PCA components:               PCA_comp = {50, 100, 150, 200, 250}
Proposed framework:           DRML (Dimensionality Reduction for Machine Learning)
Performance Measures:         Accuracy, Precision, Recall, F1-score

#Pseudocode for the proposed framework DRML

For each dataset in DS do
        Create a custom tokenizer CT for Stemming, Lemmatization, Noise removal, stop word removal, domain word removal
        For each algorithm in ML_Alg do
        1.   Load the reviews from DS using the BeautifulSoup XML Parser
             1.1   Find the Review Text from reviews with ratings "4" and "5" and store it in positive_reviews
             1.2   Find the Review Text from reviews with ratings "1" and "2" and store it in negative_reviews
        2.   Create two empty lists positive_tokenzied and negative_tokenized
        3.   Initialize word_index_map //word_index_map is a dictionary of key-value pairs
        4.   For each Review Text in positive_reviews and negative_reviews do
             4.1   Pass the Review Text to CT to extract all the tokens and save them in positive_tokenized and negative_tokenzied
                lists // Each element is a list of tokens for that review
             4.2   Save all the unique tokens in word-index_map
             Endfor
        5.   Write a function, "tokens_to_vector (tokens)" that will take "tokens" as parameters and returns an N-dimensional list of feature vectors of size A x B where
             A is the number of reviews & B is the length of word_index_map
        6.   For each record in positive_tokenized do
             6.1   pass the tokens from positive_tokenized to tokens_to_vectors & save the feature vectors in data_pos, which is N-
                dimensional list // list size is A x B, where A is the number of reviews and B is the length of word_index_map
             EndFor
        7.   For each record in negative_tokenized do
             7.1   pass the tokens from negative_tokenized to tokens_to_vectors & save the feature vectors in data_neg, which is N-
                dimensional list // list size is A x B, where A is the number of reviews and B is the length of word_index_map
             EndFor
        8.   While there are entries in PCA_comp do
             8.1   For each comp in PCA_comp do
                   # Create a dataset for positive reviews with "comp" Principal Components
                   8.1.1    Create a StandardScalar object and fit it to the data_pos data
                   8.1.2    Transform the data_pos to obtain a standardized version of the data
                   8.1.3    Apply PCA to the standardized data by creating a PCA object with comp components and fitting it to the standardized version of data_pos
                   8.1.4    Transform data_pos using PCA to obtain a dataset pc_results_pos with comp principal components //pc_results_pos is P x comp list with P
                            being the number of positive reviews
                   # Create a dataset for negative reviews with "comp" Principal Components
                   8.1.5    Create a StandardScalar object and fit it to the data_neg data
                   8.1.6    Transform the data_neg to obtain a standardized version of the data
                   8.1.7    Apply PCA to the standardized data by creating a PCA object with comp components and fitting it to the standardized version of data_neg
                   8.1.8    Transform data_neg using PCA to obtain a dataset pc_results_neg with comp principal components //pc_results_neg is P x comp list with P
                            being the number of negative reviews
                   #Prepare a supervised dataset with a mix of positive & negative reviews
                   8.1.9    Append a column to pc_results_pos with a label "1" for all rows and create final_pos list //the last column has a label "1" for positive
                            reviews
                   8.1.10   Append a column to pc_results_neg with a label "0" for all rows and create final_neg list //the last column has a label "0" for negative
                            reviews
                   8.1.11   Create a final_data dataset by concatenating pc_results_pos and pc_results_neg lists // final_data has a dimension of P x (comp+1), where P
                            is the total reviews
                   8.1.12   Random shuffle the final_data to mix positive and negative reviews
                   8.1.13   Split the final_data into Xtrain, Ytrain, Xtest & Ytest with 70% for training and 30% for testing
                   #Train the model
                   8.1.14   If ML_Alg is Max_voting, then
                            8.1.14.1       Assign BernoulliNB Classifier to model1
                            8.1.14.2       Assign Decision_Tree to model2
                            8.1.14.3       Assign KNN to model3
                            8.1.14.4       Create a final_model with Max_voting combining model1, model2 and model3 using Xtrain &
                                           Ytrain
                            8.1.14.5       Use final_model for the testing Xtest & Ytest
                            8.1.14.6       Compute Performance Measures with confusion matrix, Accuracy, Precision, Recall & F1-score
                            Else
                            8.1.14.1       Create a final_model with ML_alg using Xtrain & Ytrain
                            8.1.14.2       Use final_model for the testing Xtest & Ytest
                            8.1.14.3       Compute Performance Measures with confusion matrix, Accuracy, Precision, Recall & F1-score
                            EndIf
                Endfor
             EndWhile
        EndFor
```

EndFor

Output:

a. Reduced feature set for multi-domain datasets using DRML
b. Performance Measures for 4 Classifiers on multi-domain datasets
c. Trained sentiment classification model using DRML

TABLE. I. DATASET DESCRIPTION

Field Name	Description
Unique Id	Unique Identifier for reviews
Asin	List of Amazon identifiers used for the product
Product Name	Name of the product purchased
Product Type	Type of product
Helpful	Number of customers that found this review useful
Rating	Contains 1 to 5 stars to rate the product
Title	Title for the review
Date	Date of the review
Reviewer	Name of the reviewer
Review Location	Location of the reviewer
Review Text	Reviews shared by the reviewer about the product purchased

IV. RESULTS AND DISCUSSION

We validate our proposed methodology, DRML against the gold standard multi-domain datasets, DVD, Electronics and Kitchenware explained in Section 3.3, and with four machine learning algorithms, Decision Tree, K-Nearest Neighbour (KNN), Bernoulli Naïve Bayes (BNB) & Max Voting Ensemble explained in section 4.1. For Dimensionality reduction, we have used Principal Component Analysis (PCA) described in section 4.2. We have developed an algorithm that separates principal components from positive and negative reviews and subsequently created a supervised dataset by combining these components. We have conducted experiments by selecting different components for PCA to identify the right feature set.

A. Baseline Comparisons

In our study, we have conducted a comprehensive comparison of our proposed methodology with both a traditional Bag of Words (BoW) baseline approach and recently published state-of-the-art (SOTA) research. This comparison is pivotal to assess the effectiveness and performance of our approach in the context of sentiment analysis.

1) Bag of words (BoW) baseline: To establish a foundational baseline, we implemented the Bag of Words technique, a traditional and widely recognized approach in sentiment analysis. Our aim is to benchmark our methodology against this well-established method, providing a clear point of reference.

2) Comparison with recently published state-of-the-art (sota) research: Additionally, we have selected and evaluated our methodology against eight recent research papers that have employed the same benchmark dataset from Amazon for sentiment analysis. These research papers have published their accuracy results, allowing us to conduct a comparative analysis. By choosing recent studies, we ensure that our baseline comparisons are current and relevant to the state of the field.

3) Dataset consistency: To ensure the validity of our baseline comparisons, it is crucial to note that we have conducted these comparisons on the same dataset(s) as those used in the selected research papers. This practice enables an equitable and accurate assessment of our approach against the selected baselines.

To facilitate comparison, we employed four evaluation metrics: Accuracy, Precision, Recall, and F1-score.

B. Performance Metrics of Baseline Methodology

Table II presents performance metrics for the baseline methodology using Machine Learning algorithms with the DVD dataset. BNB achieved the highest accuracy at 76.4%, closely followed by Max Voting Ensemble at 73.2%. Precision scores were led by BNB at 78.12%. Recall scores matched accuracy, and the highest F1-score was 75.61 for BNB. Table III outlines performance metrics for the Electronics dataset. BNB led with the highest accuracy at 81.2%, followed by Max Voting at 78%. BNB also achieved the highest precision at 81.25, and the highest F1-score at 81.16. Table IV presents performance metrics for the Kitchenware dataset. BNB maintained its lead with 83.6% accuracy, 85.46% precision, and 83.54% F1-score.

TABLE. II. PERFORMANCE OF MACHINE LEARNING ALGORITHMS WITH DVD DATASET

Machine Learning Algorithms	Accuracy	Precision	Recall	F1-score
Decision Tree	66.4	66.29	66.4	66.24
K-Nearest Neighbors (KNN)	61.2	67.47	61.2	55.46
Bernoulli Naïve Bayes (BNB)	76.4	78.12	76.4	75.61
Max voting Ensemble	73.2	74.2	73.2	73.07

TABLE. III. PERFORMANCE OF MACHINE LEARNING ALGORITHMS WITH ELECTRONICS DATASET

Machine Learning Algorithms	Accuracy	Precision	Recall	F1-score
Decision Tree	74.4	74.47	74.4	74.39
K-Nearest Neighbors (KNN)	60.4	69.12	60.4	55.07
Bernoulli Naïve Bayes (BNB)	81.2	81.25	81.2	81.16
Max voting Ensemble	78	78.83	78	77.73

TABLE. IV. PERFORMANCE OF MACHINE LEARNING ALGORITHMS WITH KITCHENWARE DATASET

Machine Learning Algorithms	Accuracy	Precision	Recall	F1-score
Decision Tree	72.8	72.78	72.8	72.77
K-Nearest Neighbors (KNN)	61.6	69.97	61.6	56.69
Bernoulli Naïve Bayes (BNB)	83.6	85.46	83.6	83.54
Max voting Ensemble	78.4	81.01	78.4	77.97

Fig. 3 illustrates bar graphs for Accuracy, Precision, Recall, and F1-score. BNB achieved the highest scores in all parameters. Max Voting was chosen as the baseline for benchmarking due to its usage in our proposed framework, DRML.

In Fig. 4, we present performance metrics for the baseline methodology using multi-domain datasets (MDS). These metrics, accuracy of 76.53%, precision of 78.01%, recall of 76.53%, and an F1-score of 76.26 will serve as a benchmark for evaluating our proposed methodology, DRML.

C. Performance Metrics of Proposed Methodology

Table V presents the performance metrics for the DVD dataset with various PCA components. Notably, the Max Voting Ensemble outperforms other algorithms across all PCA components. The highest accuracy of 99.0% is achieved with 50 PCA components using Max Voting, followed by Decision Tree with 97.0%. The lowest accuracy is observed with KNN at 84.0%.

In Table VI, we compare the performance of the Electronics dataset. The highest accuracy score is 98.5% with 50 PCA components using Max Voting. Decision Tree achieves the second-highest accuracy at 97.3%. In contrast, KNN attains the lowest accuracy of 84.4%.

Table VII showcases the results for the Kitchenware dataset, which exhibits the lowest scores among the three domains. Here, Max Voting achieves the highest accuracy of 97.6% with 50 PCA components, while KNN records the lowest accuracy at 86.4%.

The series of graphs titled "Accuracy %, Precision %, Recall %, and F1-Score % with Different PCA Components for Various Datasets and Algorithms", Fig. 5 to Fig. 8 provides a comprehensive analysis of the performance of four machine learning algorithms—Decision Tree, K-Nearest Neighbors (KNN), Bernoulli Naive Bayes, and Max Voting—across three datasets (DVD, Electronics, and Kitchenware) with varying numbers of Principal Component Analysis (PCA) components (50, 100, 150, 200, and 250). Each graph illustrates the impact of PCA on a specific performance metric, namely accuracy, precision, recall, and F1-score.

Fig. 3. Accuracy, precision, recall & F1-score for baseline methodology.

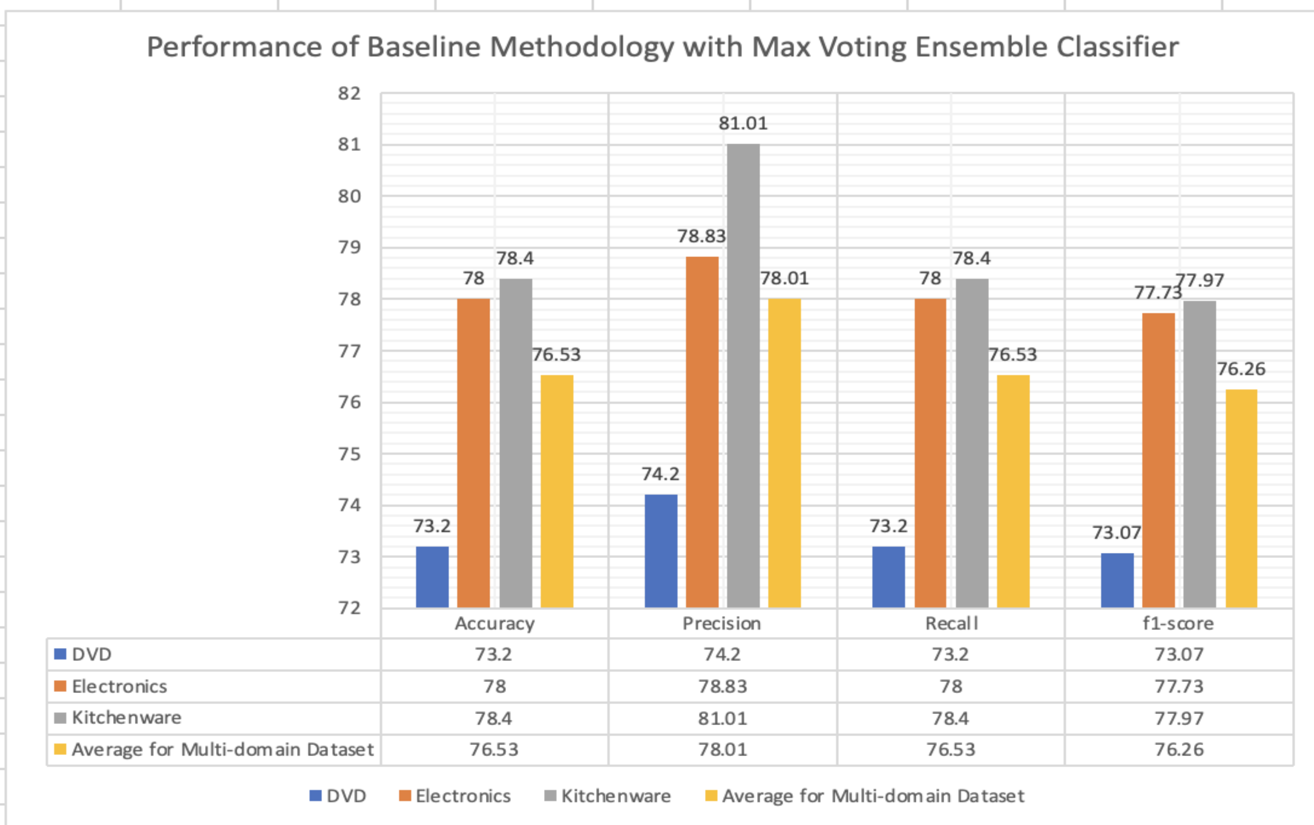

Fig. 4. Performance of baseline methodology with max voting ensemble classifier.

TABLE. V. PERFORMANCE OF MACHINE LEARNING ALGORITHMS AFTER DIMENSIONALITY REDUCTION WITH DVD DATASET

Machine Learning Algorithms	PCA Dimension	Accuracy	Precision	Recall	F1-score
Decision Tree	50	95.6	95.63	95.6	95.6
	100	93.2	93.34	93.2	93.2
	150	95.4	95.44	95.4	95.4
	200	96.5	96.53	96.5	96.5
	250	97.2	97.2	97.2	97.2
K-Nearest Neighbors (KNN)	50	92.8	92.84	92.8	92.8
	100	91.2	91.19	91.2	91.19
	150	87.2	87.53	87.2	87.22
	200	84.8	84.83	84.8	84.81
	250	84	84.06	84	84.01
Bernoulli Naïve Bayes (BNB)	50	93.2	93.21	93.2	93.2
	100	89.2	89.76	89.2	89.16
	150	87.6	88.73	87.6	87.37
	200	88	89.69	88	87.85
	250	90.8	91.99	90.8	90.74
Max voting Ensemble	50	99	99.02	99	99
	100	95.33	95.43	95.33	95.34
	150	94	94.31	94	93.98
	200	91	91.88	91	90.99
	250	96	96.26	96	95.97

TABLE. VI. PERFORMANCE OF MACHINE LEARNING ALGORITHMS AFTER DIMENSIONALITY REDUCTION WITH ELECTRONICS DATASET

Machine Learning Algorithms	PCA Dimension	Accuracy	Precision	Recall	F1-score
Decision Tree	50	97.2	97.2	97.2	97.2
	100	94.9	94.9	94.9	94.9
	150	96.4	96.43	96.4	96.4
	200	97.33	97.34	97.33	97.33
	250	96.8	96.83	96.8	96.8
K-Nearest Neighbors (KNN)	50	92	92	92	92
	100	90.8	90.8	90.8	90.8
	150	85.2	85.19	85.2	85.2
	200	84.4	84.41	84.4	84.39
	250	87.2	87.23	87.2	87.19
Bernoulli Naïve Bayes (BNB)	50	92	92.51	92	91.96
	100	85.6	88.14	85.6	85.39
	150	90.4	91.66	90.4	90.3
	200	91.6	92.35	91.6	91.5
	250	96	96.17	96	95.98
Max voting Ensemble	50	98.5	98.54	98.5	98.5
	100	96.15	96.22	96.15	96.13
	150	95	95.45	95	94.98
	200	95.2	95.23	95.2	95.19
	250	95.2	95.63	95.2	95.2

TABLE. VII. PERFORMANCE OF MACHINE LEARNING ALGORITHMS AFTER DIMENSIONALITY REDUCTION WITH KITCHENWARE DATASET

Machine Learning Algorithms	PCA Dimension	Accuracy	Precision	Recall	F1-score
Decision Tree	50	97	97.01	97	97
	100	96.2	96.2	96.2	96.2
	150	96.8	96.8	96.8	96.8
	200	96.67	96.69	96.67	96.67
	250	96.4	96.4	96.4	96.4
K-Nearest Neighbors (KNN)	50	88.8	88.82	88.8	88.81
	100	87.6	87.8	87.6	87.58
	150	86.4	86.4	86.4	86.38
	200	87.6	87.61	87.6	87.59
	250	87.2	87.22	87.2	87.2
Bernoulli Naïve Bayes (BNB)	50	91.2	91.62	91.2	91.2
	100	90	90.14	90	90
	150	88	88.73	88	87.89
	200	87.6	88.2	87.6	87.59
	250	91.2	91.73	91.2	91.18
Max voting Ensemble	50	97.6	97.66	97.6	97.6
	100	94.58	94.73	94.58	94.57
	150	94.2	94.29	94.2	94.2
	200	95.6	95.84	95.6	95.59
	250	96.4	96.41	96.4	96.4

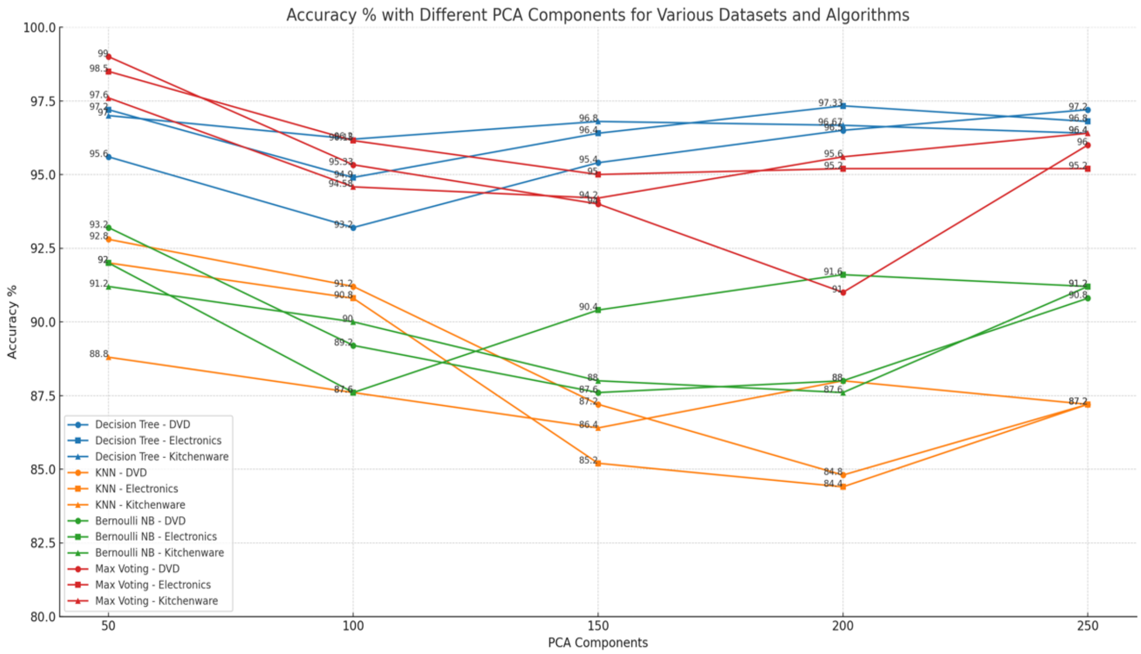

Fig. 5. Comparison of accuracy % with different PCA components for 3 datasets and 4 machine learning algorithms.

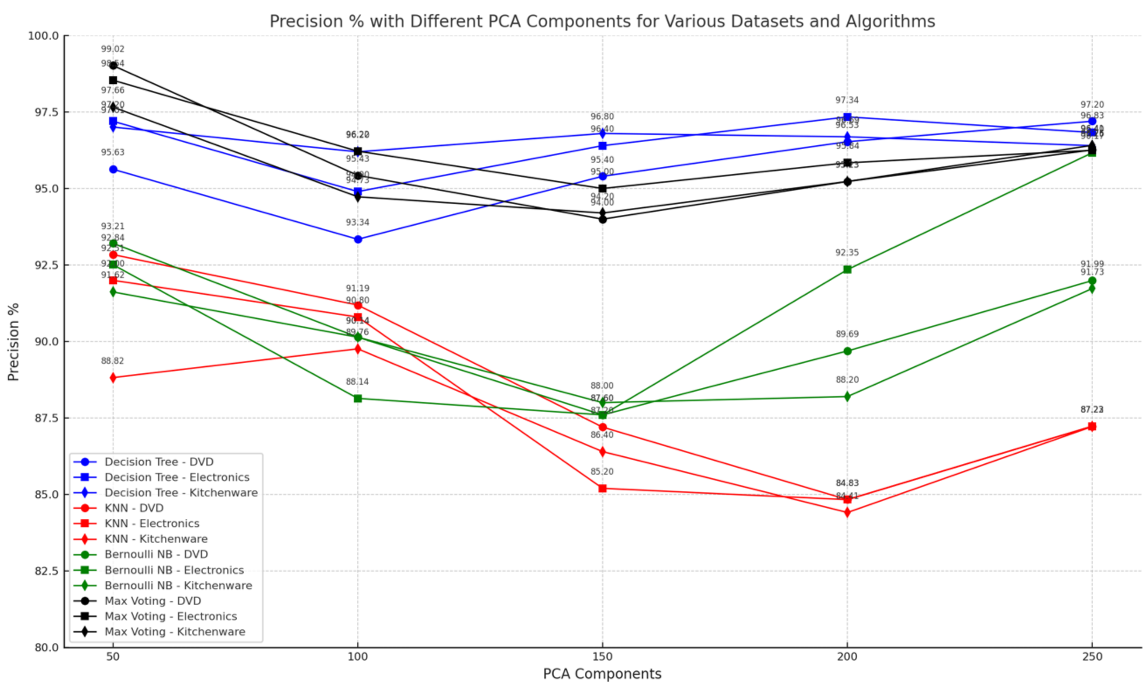

Fig. 6. Comparison of precision % with different PCA components for 3 datasets and 4 machine learning algorithms.

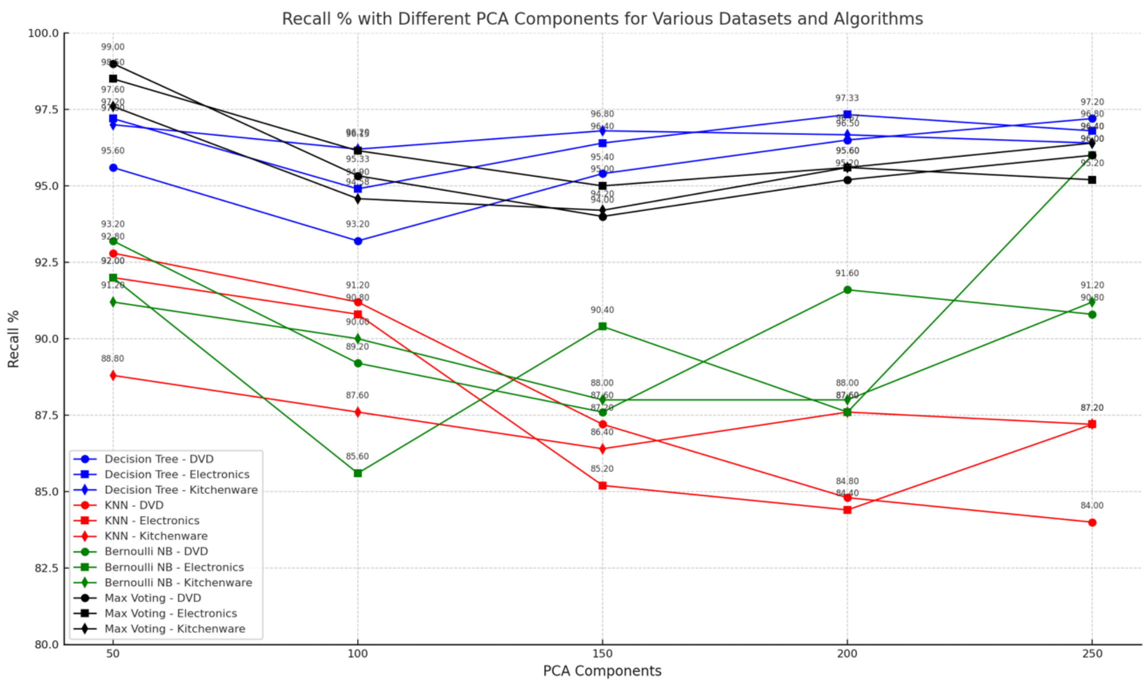

Fig. 7. Comparison of recall % with different PCA components for 3 datasets and 4 machine learning algorithms.

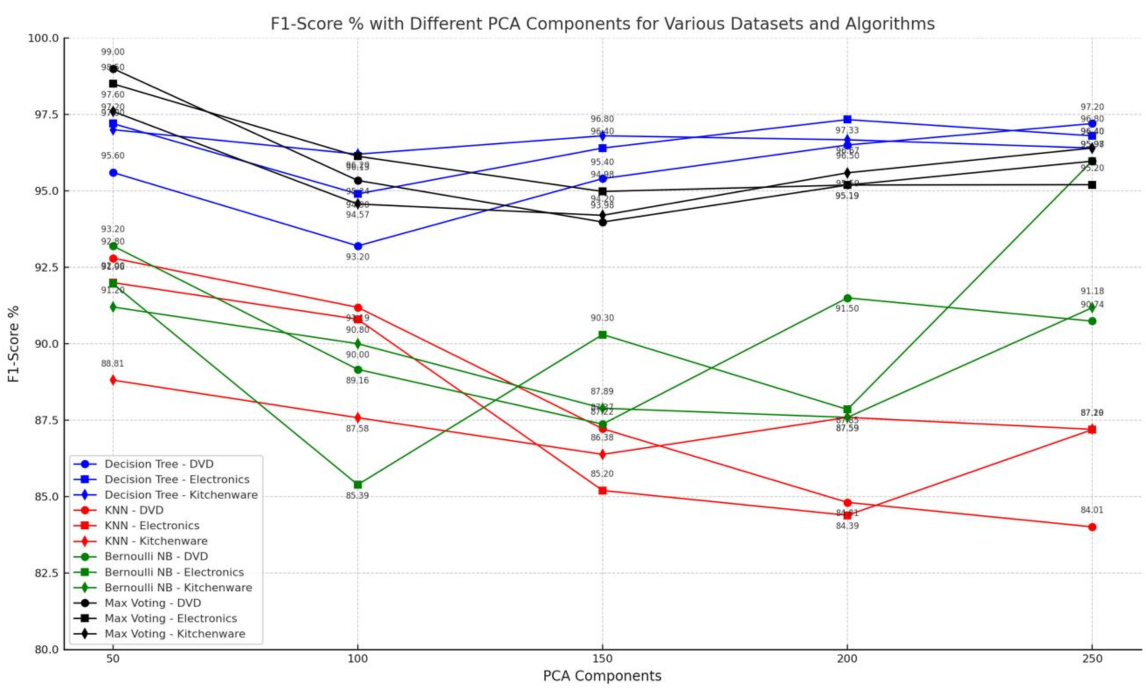

Fig. 8. Comparison of F1-Score % with different PCA components for 3 datasets and 4 machine learning algorithms.

The Decision Tree algorithm consistently demonstrates high performance across all metrics and datasets, particularly excelling with the Electronics dataset, achieving metrics above 96% across various PCA components. Max Voting also shows robust performance, especially with the DVD and Electronics datasets, reaching peak values of 99% in multiple metrics. KNN and Bernoulli Naive Bayes exhibit more variability, with KNN generally showing a decline in performance as the number of PCA components increases, particularly for the Kitchenware dataset. Conversely, Bernoulli Naive Bayes shows improvement in some cases, notably with the Electronics dataset, where it achieves high values in precision and F1-score with 250 PCA components.

Based on this comprehensive analysis, Max Voting with 50 PCA components emerges as a compelling choice. This conclusion is supported by its consistently high performance across all evaluated metrics—accuracy, precision, recall, and F1-score—particularly with the DVD and Electronics datasets. The Max Voting algorithm achieves peak values of 99% in both precision and recall for the DVD dataset and 98.5% in both metrics for the Electronics dataset with 50 PCA components. This demonstrates its robustness and reliability in maintaining high performance with reduced dimensionality, making it an efficient choice for real-world applications where computational resources and processing time are critical considerations. The stability and consistency of Max Voting with 50 PCA components across multiple datasets and metrics underscore its versatility and effectiveness as a classification model, providing a balanced trade-off between model complexity and performance.

In Fig. 9, comparing dimensionality reduction using various PCA components, it's evident that Max Voting consistently achieves the highest scores for multi-domain datasets with 50 PCA components, outperforming other configurations by 2.31% to 2.51%. This highlights the efficacy of Max Voting with 50 PCA components.

Fig. 10 displays the average performance of the Max Voting Ensemble with 50 PCA components across multi-domain datasets, including DVD, Electronics, and Kitchenware. The performance scores of 98.37 for Accuracy, 98.41 for Precision, 98.37 for Recall, and F1-score, achieved using the Max Voting Ensemble classifier with 50 PCA components, serve as a benchmark for comparing with the baseline methodology and existing research.

D. Comparison with the Baseline Methodology

After extensive experimentation, we have determined that employing the Max Voting classifier with 50 PCA components is the optimal approach. As shown in Fig. 11, the Max Voting classifier significantly enhances performance, achieving a 25.8% increase in Accuracy and Recall, 24.82% increase in Precision, and 25.93% increase in the F1-score for the DVD dataset. For the Electronics dataset, Accuracy reached 98.5, and for Kitchenware, it reached 97.6.

Table VIII highlights the impressive reduction in feature matrix size using DRML, resulting in a 99.76% reduction for the DVD dataset, 99.54% for Electronics, and 99.45% for Kitchenware datasets.

Fig. 12 illustrates the comparison between the baseline methodology and our proposed methodology using a multi-domain dataset. Table IX reveals significant improvements, with increases of 21.84% in Accuracy, 20.4% in Precision, 21.84% in Recall, and 22.11% in F1-score, demonstrating that DRML enhances sentiment analysis performance by 21.55%.

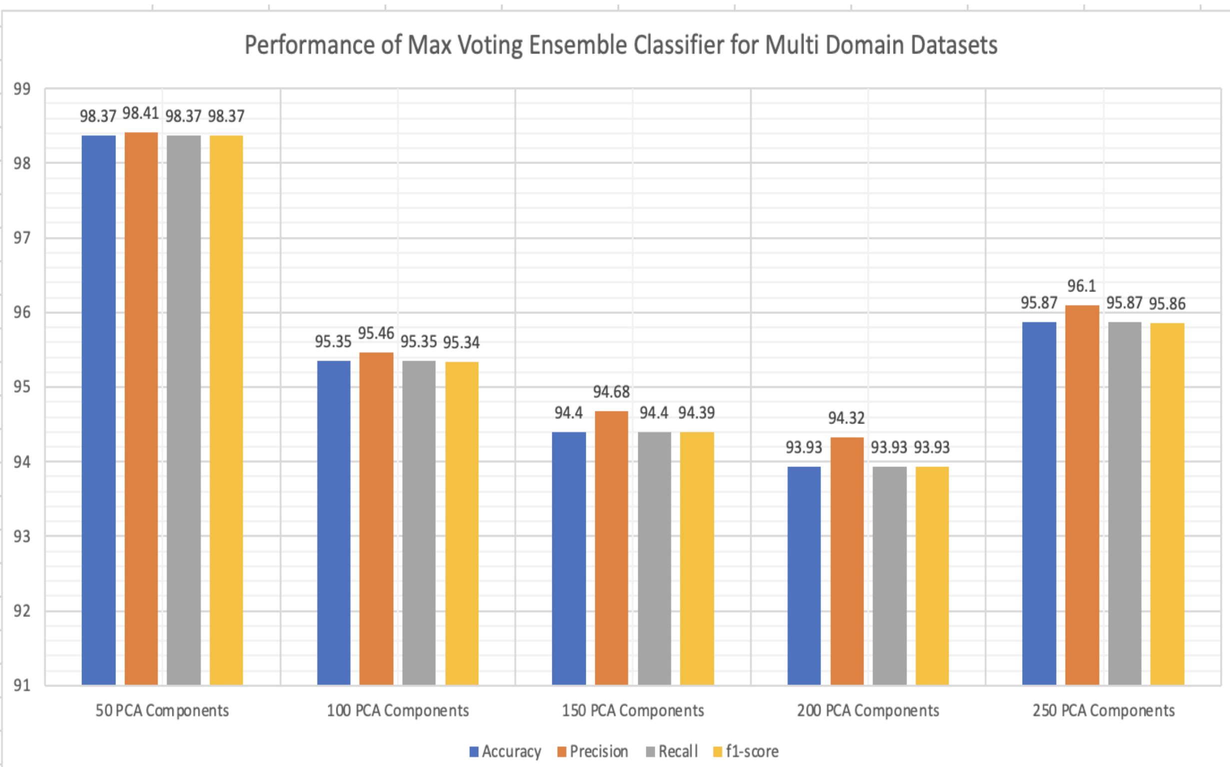

Fig. 9. Performance of max voting ensemble classifier for multi domain datasets.

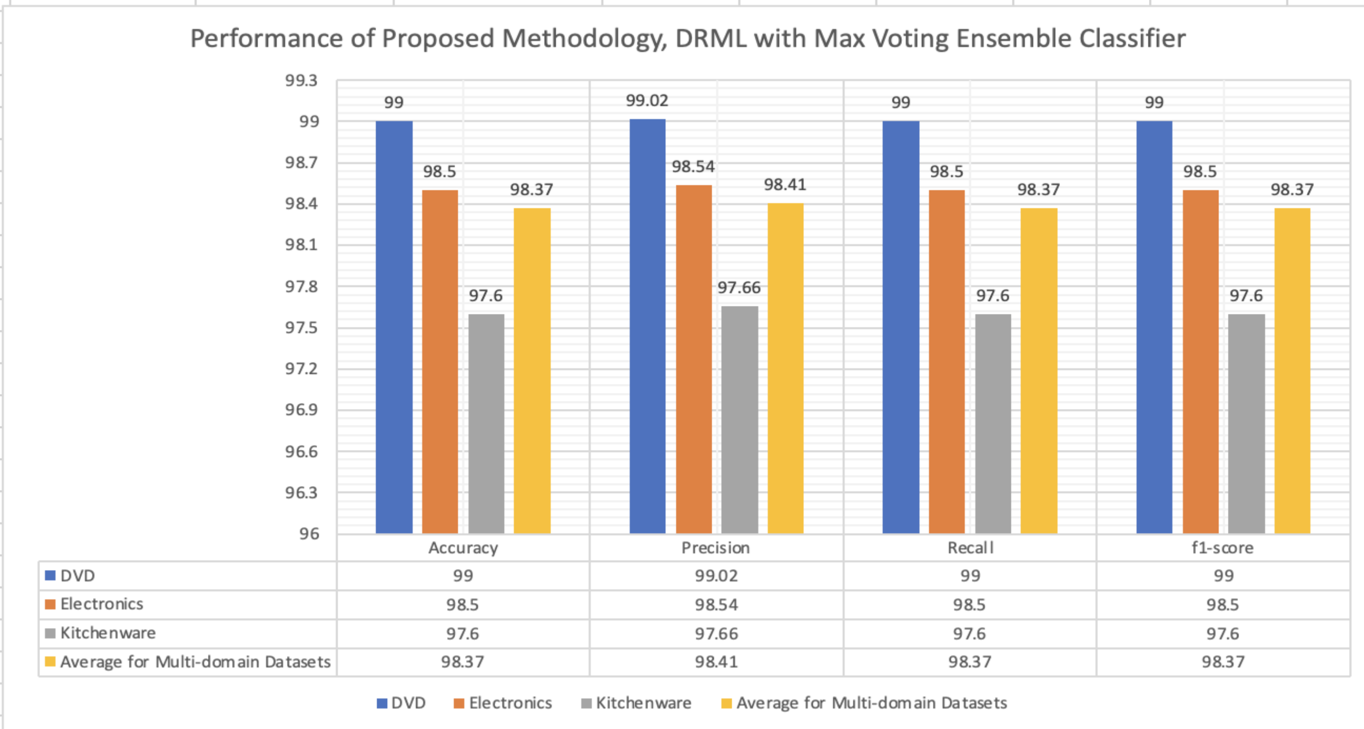

Fig. 10. Performance of Proposed Methodology, DRML with Max Voting Ensemble Classifier.

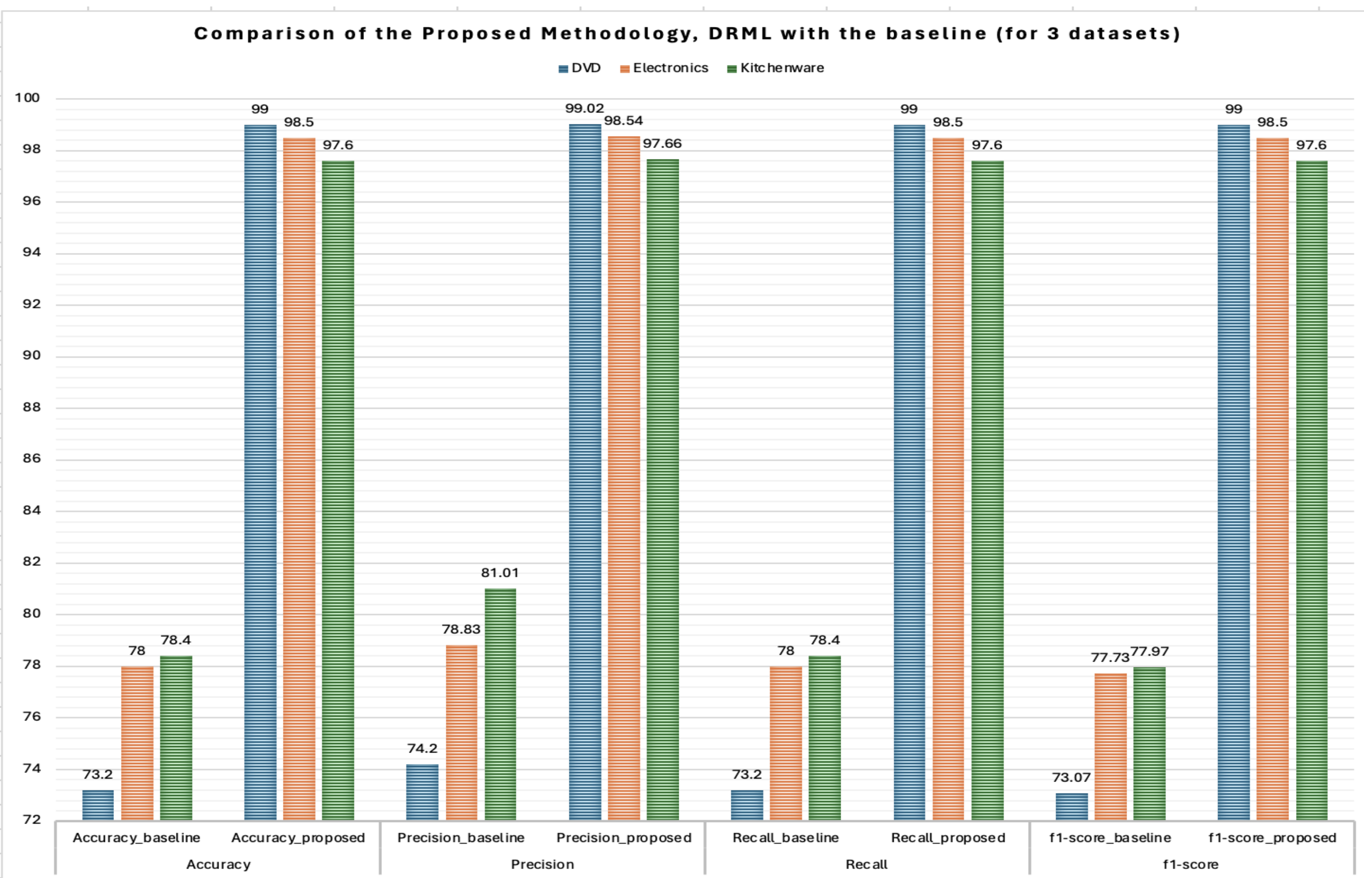

Fig. 11. Comparison of the proposed methodology, DRML with baseline across 3 datasets

TABLE. VIII. PERCENTAGE OF DIMENSIONALITY REDUCTION WITH DRML

Dataset	No. of Positive Reviews	No. of Negative Reviews	No. of features	Size of the feature matrix (baseline)	Size of the feature matrix using DRML	% of reduction in the size of the feature matrix (DRML)
DVD	1000	1000	21344	2000 x 21344	2000 x 51	99.76
Electronics	1000	1000	11150	2000 x 11150	2000 x 51	99.54
Kitchen	1000	1000	9268	2000 x 9268	2000 x 51	99.45

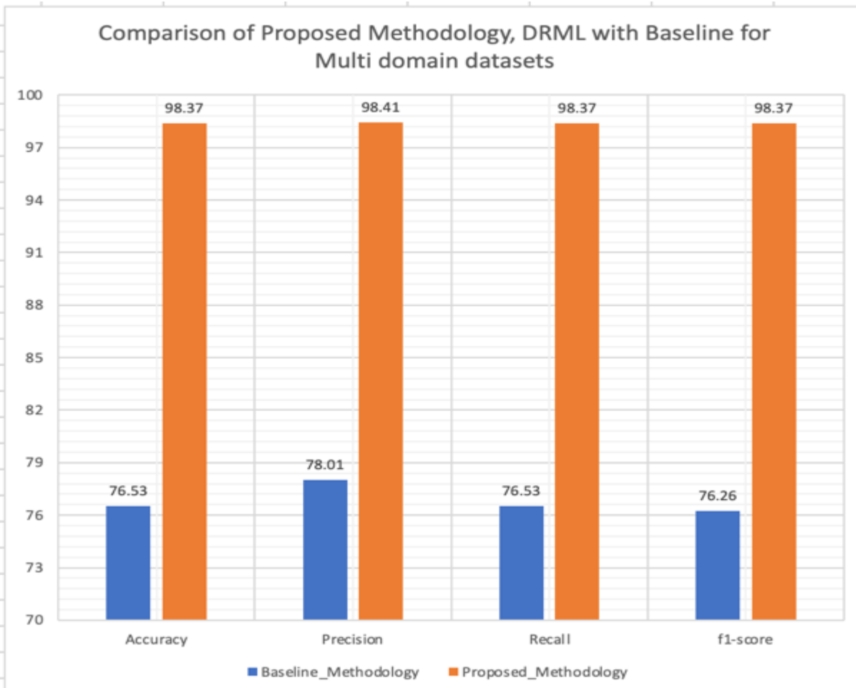

Fig. 12. Comparison of the proposed methodology with the baseline for multi-domain datasets

TABLE. IX. PERCENTAGE ENHANCEMENT BY THE PROPOSED METHODOLOGY VS. BASELINE METHODOLOGY

Methodology	No. of Reviews	Size of feature matrix	% of reduction in the size of feature matrix (DRML)	Accuracy	Precision	Recall	F1-Score	Overall % increase in performance
Baseline	2000	2000 x 13920		76.53	78.01	76.53	76.26	
Proposed	2000	2000 x 51		98.37	98.41	98.37	98.37	
% Enhancement			99.63	21.84	20.4	21.84	22.11	21.55

Table IX summarizes the overall outcome of the evaluation of DRML with the baseline methodology. It is heartening to note that our framework gives an average performance of 98.38%. There is an impressive improvement of 21.55% in performance while reducing the dimension by 99.63% in comparison with the baseline methodology. This demonstrates the remarkable impact of our proposed methodology on enhancing sentiment analysis.

E. Comparison with the State-of-the-Art (SOTA) Published Research

Upon comparing our proposed methodology with the baseline approach, it is now pertinent to evaluate its performance in relation to published research. We selected eight recent research papers that reported accuracy scores on benchmark multi-domain datasets using the same dataset as ours. Onan [26] incorporated GRU layers and bidirectional LSTM to reduce dimensionality while emphasizing significant features. Alrehili et al. [3] employed a Voting ensemble method with five classifiers, achieving high accuracy, with Random Forest leading at 89.87% in the unigram scenario. Geetha et al. [15] introduced the BERT Base Uncased model to improve sentiment analysis accuracy and reduce training time.

Sharma et al. [38] introduced "SentiDraw", a novel approach that leverages probability distributions across reviews with different star ratings to calculate Sentiment Orientation (SO) scores. This hybrid approach, combining SentiDraw with supervised methods, achieved state-of-the-art performance in polarity determination for reviews.

Beigi et al. [8] presented an innovative approach blending neural networks and sentiment lexicons, adapting word polarities to target domains and outperforming unsupervised domain adaptation alternatives.

Zhao et al. [48] proposed the PTASM-BERT method, utilizing parameter transferring and attention sharing mechanisms to achieve state-of-the-art results on Amazon review cross-domain datasets. Fu et al. [14] introduced the Sentiment-Sensitive Network Model (SSNM), surpassing existing methods on the Amazon review dataset by transferring attention to emotions across domains. Xia et al. [44] introduced the dual sentiment analysis (DSA) model, effective in classifying sentiments into three categories (positive-negative-neutral) and constructing a corpus-based pseudo-antonym dictionary.

In Fig. 13, we provide a visual comparison of Accuracy % for sentiment analysis of our proposed methodology, DRML, with the aforementioned research papers. Table X details the percentage increase for each paper. Our methodology, DRML, achieved an impressive average increase of 10.96% in Accuracy for sentiment analysis, establishing its competitive edge in the field.

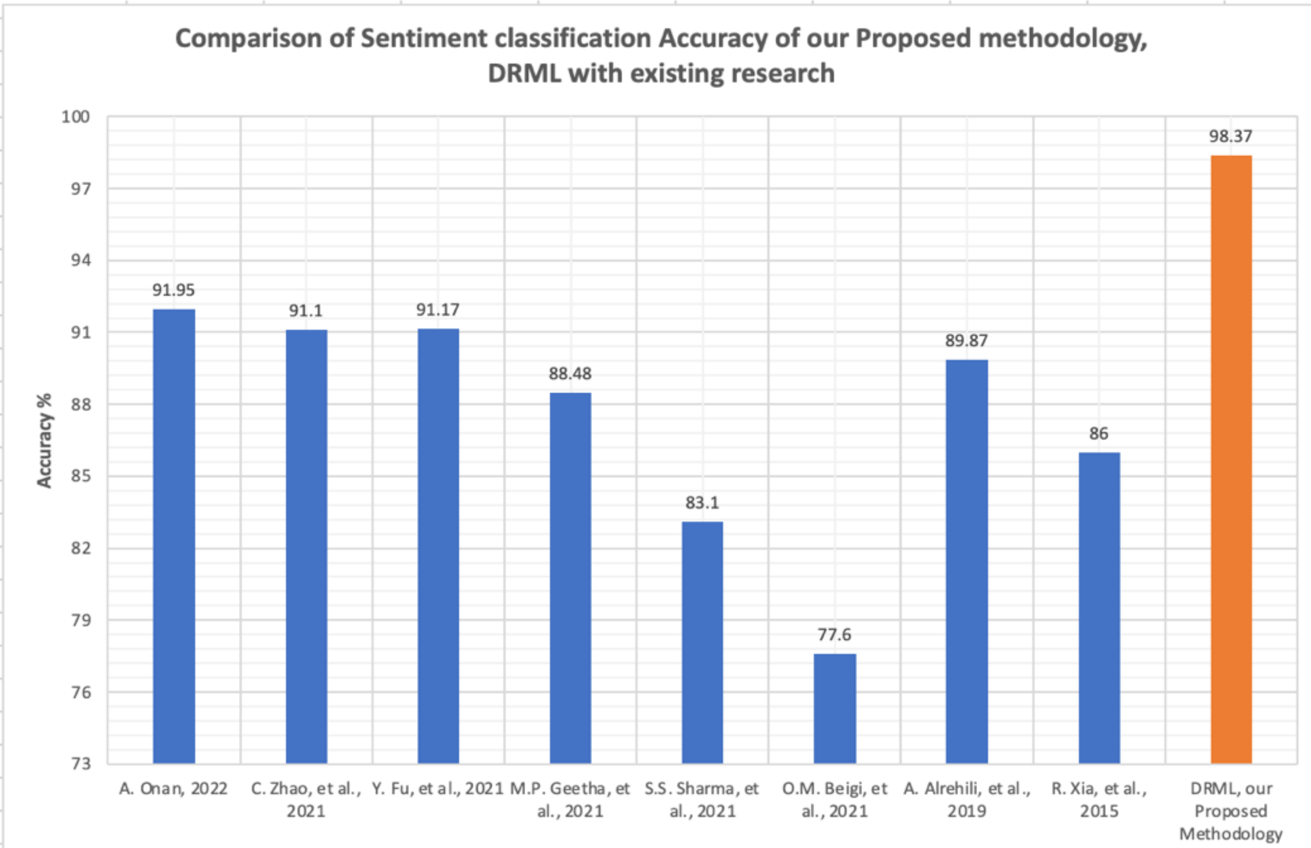

Fig. 13. Comparing sentiment classification accuracy: DRML vs. SOTA

TABLE. X. PERCENTAGE INCREASE OF ACCURACY: DRML VS. SOTA RESEARCH

Sentiment Analysis Research	Accuracy %	% Increase by DRML
Onan, 2022	91.95	6.42
Zhao et. al., 2021	91.1	7.27
Fu and Liu, 2021	91.17	7.2
Geetha & Renuka, 2021	88.48	9.89
Sharma & Dutta, 2021	83.1	15.27
Beigi & Moattar, 2021	77.6	20.77
Alrehili & Albalawi, 2019	89.87	8.5
Xia et. al., 2015	86	12.37
DRML, our Proposed Methodology	98.37	

F. *Discussion of Research Implications*

Sentiment analysis, a critical component of natural language processing, plays a pivotal role in understanding public opinion and user sentiment across various domains. Enhancing the performance of sentiment analysis presents a significant challenge due to the complexities of high-dimensional text data and the intricacies of user-generated content.

1) Theoretical implications: Our research introduces a novel framework named DRML for sentiment analysis, which leverages Principal Component Analysis for dimensionality reduction. This approach showcases theoretical advancements by significantly reducing the dimension of the feature matrix while improving classification performance. This reduction in dimensionality demonstrates the potential for streamlined sentiment analysis, an important contribution to the field's theoretical framework.

2) Practical implications: Our work offers valuable practical applications across industries. In the realm of e-commerce, DRML can enhance product recommendations and brand reputation management, leading to improved user experiences. Financial markets can harness data-driven trading decisions for better investment strategies. Businesses, irrespective of their domain, can optimize marketing campaigns, customer support, and decision-making based on accurate sentiment analysis, ultimately enhancing customer satisfaction and fostering growth. These practical implications underscore the potential for our research to drive real-world applications.

3) Distinguishing from existing work: While existing sentiment analysis methodologies are often limited by high-dimensional data, our approach, DRML, distinguishes itself by demonstrating the ability to significantly reduce dimensionality while simultaneously improving sentiment analysis performance. This sets it apart from traditional methods and contributes a unique perspective to the field.

The comparison of DRML against baseline methods and state-of-the-art research papers underlines its superiority. Our experiments have showcased an average increase of 21.55% in sentiment analysis accuracy, demonstrating the practical and theoretical significance of our work.

V. Conclusion

In this study, we have introduced a novel framework, Dimensionality Reduction for Machine Learning (DRML), aimed at enhancing the efficiency of sentiment analysis. Our research has successfully addressed the fundamental question of whether substantial feature space reduction can enhance sentiment analysis performance. Through a rigorous examination of well-established benchmark datasets from Amazon, including DVD, Electronics, and Kitchenware, we have demonstrated the efficacy of our approach.

Our findings provide crucial insights into the application of dimensionality reduction techniques in sentiment analysis. By utilizing Principal Component Analysis (PCA) to extract key features from product reviews, we have successfully reduced the dimension of the feature matrix by an impressive 99.63%. Simultaneously, our ensemble machine learning classifier, incorporating various algorithms, has boosted sentiment classification performance by an average of 21.55%. Furthermore, the comparisons with state-of-the-art (SOTA) methodologies and baseline approaches underscore the significance of our research. DRML consistently outperformed individual classifiers such as Decision Tree, K-Nearest Neighbors, and Bernoulli Naïve Bayes across various domains, achieving accuracy scores as high as 99%. These results exhibit the practical applicability of our approach in domains like e-commerce, financial markets, and beyond.

Looking to the future, the role of sentiment analysis in the workplace is poised for transformation. The explosion of user-generated content across online platforms, customer reviews, and social media necessitates advanced tools for understanding sentiment at scale. Our work, which combines dimensionality reduction and machine learning, sets the stage for more sophisticated techniques. Deep learning, with its capacity to capture intricate patterns in textual data, is an area ripe for exploration. The integration of deep neural networks into our sentiment analysis framework offers an exciting avenue for future research. By combining the power of dimensionality reduction, traditional machine learning algorithms, and cutting-edge deep learning networks, our research can continue to push the boundaries of sentiment analysis performance, adaptability, and scalability.

In conclusion, our research advances the field of sentiment analysis by presenting a novel framework that not only enhances the efficiency of sentiment analysis but also opens new avenues for the analysis of large-scale textual data in various real-world applications. As the landscape of textual data analysis continues to evolve, our approach offers a promising foundation for future research and applications in a world increasingly dominated by vast volumes of user-generated content.

References

[1] Ahmed C, ElKorany A, & Elsayed E (2023) Prediction of customer's perception in social networks by integrating sentiment analysis and machine learning. Journal of Intelligence Information Systems 60:829-851 https://doi.org/10.1007/s10844-022-00756-y

[2] Alhamad G, & Kurdy M (2020) Feature-Based Sentiment Analysis for Arabic Language. (IJACSA) International Journal of Advanced Computer Science and Applications, 11(11). http://dx.doi.org/10.14569/IJACSA.2020.0111158

[3] Alrehili A, & Albalawi K (2019) Sentiment Analysis of Customer Reviews Using Ensemble Method. International Conference on Computer and Information Sciences, (ICCIS). https://doi.org/10.1109/ICCISci.2019.8716454

[4] Alwehaibi A, Bikdash M, Albogmi M, & Roy K (2022) A study of the performance of embedding methods for Arabic short-text sentiment analysis using deep learning approaches. Journal of King Saud University – Computer and Information Sciences, 34:6140-6149. https://doi.org/10.1016/j.jksuci.2021.07.011

[5] Alyami S, Alhothali A, & Jamal A (2022) Systematic literature review of arabic aspect-based sentiment analysis. Journal of King Saud University – Computer and Information Sciences, 34(9):6524-6551. https://doi.org/10.1016/j.jksuci.2022.07.001

[6] Araque O, Corcuera-Platas I, Sanchez-Rada F, & Iglesias C.A (2017) Enhancing deep learning sentiment analysis with ensemble techniques in social applications. Expert Systems With Applications, 77:236-246. http://dx.doi.org/10.1016/j.eswa.2017.02.002

[7] Araque O, Zhu G, & Iglesias C.A (2019) A semantic similarity-based perspective of affect lexicons for sentiment analysis. Knowledge-Based Systems, 165:346-359. https://doi.org/10.1016/j.knosys.2018.12.005

[8] Beigi O.M, & Moattar M.H (2021) Automatic construction of domain-specific sentiment lexicon for unsupervised domain adaptation and sentiment classification. Knowledge-Based Systems, 213:106423 https://doi.org/10.1016/j.knosys.2020.106423

[9] Bilal M, Israr H, Shahid M, & Khan A (2016) Sentiment classification of Roman-Urdu opinions using Naive Bayesian, Decision Tree and KNN classification techniques. Journal of King Saud University – Computer and Information Sciences, 28:330-344. http://dx.doi.org/10.1016/j.jksuci.2015.11.003

[10] Birjali M, Kasri M, & Beni-Hssane A (2021) A comprehensive survey on sentiment analysis: Approaches, challenges and trends. Knowledge-Based Systems, 226:107134 https://doi.org/10.1016/j.knosys.2021.107134

[11] Blitzer J, Dredze M, & Pereira F (2007) Biographies, Bollywood, Boomboxes and Blenders: Domain Adaptation for Sentiment Classification. Association of Computational Linguistics (ACL 2007), 440-447. https://aclanthology.org/P07-1056.pdf

[12] Chen T, Xu R, He Y, & Wang X (2017) Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. Expert Systems With Applications. 72:221-230. http://dx.doi.org/10.1016/j.eswa.2016.10.065

[13] Dhamayanthi N, & Lavanya B (2021) Sentiment Analysis Framework for E-Commerce Reviews Using Ensemble Machine Learning Algorithms. Data Engineering and Intelligent Computing. Advances in Intelligent Systems and Computing, 1407. https://doi.org/10.1007/978-981-16-0171-2_34

[14] Fu Y, & Liu Y (2021) Cross-domain sentiment classification based on key pivot and non-pivot extraction. Knowledge-Based Systems, 228:107280. https://doi.org/10.1016/j.knosys.2021.107280

[15] Geetha M.P, & Renuka D.K (2021) Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model. International Journal of Intelligent Networks. 2:64-69. https://doi.org/10.1016/j.ijin.2021.06.005

[16] Ghosh M, & Sanyal G (2018) An ensemble approach to stabilize the features for multi-domain sentiment analysis using supervised machine learning. Journal of Big Data, https://doi.org/10.1186/s40537-018-0152-5

[17] Jain D.K, Boyapati P, Venkatesh J, & Prakash M (2022) An Intelligent Cognitive-Inspired Computing with Big Data Analytics Framework for Sentiment Analysis and Classification. Information Processing and Management, 59:102758. https://doi.org/10.1016/j.ipm.2021.102758

[18] Kaur H, Ahsaan S.U, Alankar B, & Chang V (2021) A Proposed Sentiment Analysis Deep Learning Algorithm for Analyzing COVID-19 Tweets. Information Systems Frontier, 23:1417-1429. https://doi.org/10.1007/s10796-021-10135-7

[19] Kazmaier J, & Vuuren J.H (2022) The power of ensemble learning in sentiment analysis. Expert Systems With Applications. 187:115819. https://doi.org/10.1016/j.eswa.2021.115819

[20] Ke W, Gao J, Shen H, & Cheng X (2021) Incorporating explicit syntactic dependency for aspect level sentiment classification. Neurocomputing, 456, 394-406. https://doi.org/10.1016/j.neucom.2021.05.078

[21] Lengkeek M, Knaap F.V.D, & Frasincar F (2023) Leveraging hierarchical language models for aspect-based sentiment analysis on financial data. Information Processing and Management, 60:103435. https://doi.org/10.1016/j.ipm.2023.103435

[22] Li Z, Li X, Xie H, Wang F.L, Leng M, & Li Q (2023) A novel dropout mechanism with label extension schema toward text emotion classification. Information Processing and Management, 60:103173. https://doi.org/10.1016/j.ipm.2022.103173

[23] Ligthart A, Catal C, & Tekinerdogan B (2021) Analyzing the effectiveness of semi-supervised learning approaches for opinion spam classification. Applied Soft Computing Journal, 101:107023. https://doi.org/10.1016/j.asoc.2020.107023

[24] Liu X, Tang T, & Ding N (2022) Social network sentiment classification method combined Chinese text syntax with graph convolutional neural network. Egyptian Informatics Journal, 23:1-12. https://doi.org/10.1016/j.eij.2021.04.003

[25] Mukherjee P, Badr Y, Doppalapudi S, Srinivasan S.M, Sangwan R.S, & Sharma, R (2021) Effect of Negation in Sentences on Sentiment Analysis and Polarity Detection. Procedia Computer Science, 185:370-379. https://doi.org/10.1016/j.procs.2021.05.038

[26] Onan A (2022) Bidirectional convolutional recurrent neural network architecture with group-wise enhancement mechanism for text sentiment classification. Journal of King Saud University – Computer and Information Sciences, 34:2098-2117. https://doi.org/10.1016/j.jksuci.2022.02.025

[27] Park S, Lee W, & Moon I (2015) Efficient extraction of domain specific sentiment lexicon with active learning. Pattern Recognition Letters, 56:38-44. http://dx.doi.org/10.1016/j.patrec.2015.01.004

[28] Pathak A.R, Pandey M, & Rautaray S (2021) Topic-level sentiment analysis of social media data using deep learning. Applied soft Computing, 108:107440. https://doi.org/10.1016/j.asoc.2021.107440

[29] Pilar G, Isabel S, Diego P, & Luis G.J (2023) A novel flexible feature extraction algorithm for Spanish tweet sentiment analysis based on the context of words. Expert Systems With Applications, 212:118817. https://doi.org/10.1016/j.eswa.2022.118817

[30] Pimpalkar A, & Raj J.R (2022) MBiLSTMGloVe: Embedding GloVe knowledge into the corpus using multi-layer BiLSTM deep learning model for social media sentiment analysis. Expert Systems with Applications, 203:117581. https://doi.org/10.1016/j.eswa.2022.117581

[31] Prastyo P.H, Hidayat R, & Ardiyanto I (2022) Enhancing sentiment classification performance using hybrid Query Expansion Ranking and Binary Particle Swarm Optimization with Adaptive Inertia Weights. ICT Express, 8:189-197. https://doi.org/10.1016/j.icte.2021.04.009

[32] Ranjan Kumar Behera, Monalisa Jena, Santanu Kumar Rath, Sanjay Misra (2021) Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data. Information Processing & Management 58: 102435 https://doi.org/10.1016/j.ipm.2020.102435

[33] Rintyarna B. S, Sarno R, & Fatichah C (2019) Evaluating the performance of sentence level features and domain sensitive features of product reviews on supervised sentiment analysis tasks. Journal of Big Data, 6:84. https://doi.org/10.1186/s40537-019-0246-8

[34] Rybinski K (2023) Content still matters. A machine learning model for predicting news longevity from textual and context features. Information Processing and Management, 60:103398. https://doi.org/10.1016/j.ipm.2023.103398

[35] Sailunaz K, & Alhajj R (2019) Emotion and sentiment analysis from Twitter text. Journal of Computational Science, 36:101003. https://doi.org/10.1016/j.jocs.2019.05.009

[36] Savci P, & Das B (2023) Prediction of the customers' interests using sentiment analysis in e- commerce data for comparison of Arabic, English, and Turkish languages. Journal of King Saud University – Computer and Information Sciences 35:227-237. https://doi.org/10.1016/j.jksuci.2023.02.017

[37] Serrano-Guerrero J, Romero F.P, & Olivas J.A (2021) Fuzzy logic applied to opinion mining: A review. Knowledge-Based Systems, 222:107018. https://doi.org/10.1016/j.knosys.2021.107018

[38] Sharma S.S, & Dutta G (2021) SentiDraw: Using star ratings of reviews to develop domain specific sentiment lexicon for polarity determination. Information Processing and Management, 58:102412. https://doi.org/10.1016/j.ipm.2020.102412

[39] Sivakumar M, & Uyyala S.R (2021) Aspect-based sentiment analysis of mobile phone reviews using LSTM and fuzzy logic. International Journal of Data Science and Analytics, 12:355-367. https://doi.org/10.1007/s41060-021-00277-x

[40] Soumya S, & Pramod K.V (2020) Sentiment analysis of malayalam tweets using machine learning techniques. ICT Express, 6:300-305. https://doi.org/10.1016/j.icte.2020.04.003

[41] Tripathy A, Agrawal A, Rath S.K (2016) Classification of sentiment reviews using n-gram machine learning approach. Expert Systems With Applications, 57:117-126. http://dx.doi.org/10.1016/j.eswa.2016.03.028

[42] Ullah M.A, Marium S.M, Begum S.A, & Dipa N.S (2020) An algorithm and method for sentiment analysis using the text and emoticon. ICT Express, 6:357-360. https://doi.org/10.1016/j.icte.2020.07.003

[43] Vernikou S, Lyras A, & Kanavos A (2022) Multiclass sentiment analysis on COVID-19-related tweets using deep learning models. Neural Computing and Applications, 34:19615-19627. https://doi.org/10.1007/s00521-022-07650-2

[44] Xia R, Xu F, Zong C, Li Q, Qi Y, & Li T (2015) Dual Sentiment Analysis: Considering Two Sides of One Review. IEEE Transactions on Knowledge & Data Engineering, 27(8):2120-2133. https://doi.org/10.1109/TKDE.2015.2407371

[45] Yicheng Zhu,Yiqiao Qiu, Qingyuan Wu, Fu Lee Wang, Yanghui Rao (2023) Topic Driven Adaptive Network for cross-domain sentiment classification. Information Processing & Management 60: 103230 https://doi.org/10.1016/j.ipm.2022.103230

[46] Yuanqing Li, Ke Zhang, Jingyu Wang, Xinbo Gao (2021) A cognitive brain model for multimodal sentiment analysis based on attention neural networks. Neurocomputing 430: 159-173 https://doi.org/10.1016/j.neucom.2020.10.021

[47] Zhang Y, Wang J, & Zhang Z (2021) Personalized sentiment classification of customer reviews via an interactive attributes attention model. Knowledge-Based Systems, 226:107135. https://doi.org/10.1016/j.knosys.2021.107135

[48] Zhao C, Wang S, Li D, Liu X, Yang X, & Liu J (2021) Cross-domain sentiment classification via parameter transferring and attention sharing mechanism. Information Sciences, 578, 281-296. https://doi.org/10.1016/j.ins.2021.07.001

[49] Zhao H, Liu Z, Yao X, & Yang Q (2021) A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach. Information Processing and Management, 58:102656. https://doi.org/10.1016/j.ipm.2021.102656

[50] Zhou Z, & Liu F (2021) Filter gate network based on multi-head attention for aspect-level sentiment classification. Neurocomputing, 441:214-225. https://doi.org/10.1016/j.neucom.2021.02.041

[51] Zhu X, Zhu L, Guo J, Liang S, & Dietze S (2021) GL-GCN: Global and Local Dependency Guided Graph Convolutional Networks for aspect-based sentiment classification. Expert Systems With Applications, 186:115712. https://doi.org/10.1016/j.eswa.2021.115712