Journal of Theoretical and Applied Information Technology
15st July 2026. Vol.104. No.13

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A NOVEL SENTIMENT ANALYSIS APPROACH: INTEGRATING SEMANTIC UNDERSTANDING, DEEP FEATURES, AND MACHINE LEARNING FOR ACCURATE CLASSIFICATION

V.SUGANYA[1], K. PREMA[2]

[1]Research Scholar, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai-600062 Tamil Nadu, India
[2]Assistant Professor, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Chennai-600062, Tamil Nadu, India
E-mail: [1]suganyaav26@gmail.com, [2]premak@veltech.edu.in

ABSTRACT

This paper introduces an advanced analysis framework meticulously crafted with an optimized approach for the evaluation of social data and the classification of sentiments. In the contemporary digital landscape dominated by social media and online interactions, the sheer volume and complexity of user-generated data present a formidable challenge. Traditional sentiment analysis methodologies often struggle to effectively interpret the nuances and evolving nature of language expressed in online spaces. To overcome these limitations, our proposed framework leverages cutting-edge technologies such as natural language processing (NLP), deep learning (DL), and machine learning (ML) algorithms. The primary contribution of this research is the development of an optimized unified framework which integrates advanced pre-processing, DL-based sentiment classification, and feature extraction. The developed approach contributes new knowledge by addressing constraints observed in large-scale unstructured social data interms of classification efficacy and contextual understanding. The framework's multi-faceted approach encompasses thorough data preprocessing, sophisticated NLP techniques for semantic understanding, utilization of deep learning models for feature extraction, and the implementation of an optimized sentiment classification algorithm. This integrated methodology not only enhances the accuracy of sentiment analysis but also ensures scalability, adaptability, and real-time capabilities. This paradigm becomes an essential resource as social media platforms develop for academics, corporations, and decision-makers who want to glean insights from the immense ocean of online interactions and comprehend the complex web of emotions influencing the digital conversation.

Keywords: *CNN, NLP, Deep learning, ConvNet-SVMBoVW Model, Adam Optimization Algorithm, Sentiment Classification*

1. INTRODUCTION

The landscape of social media and online platforms has witnessed an unprecedented surge in data generation, marking a paradigm shift in the way information is disseminated and consumed. As the digital ecosystem continues to expand, the need for robust analysis frameworks becomes paramount to extract meaningful insights from the vast pool of social data. This paper introduces a comprehensive Analysis Framework with an Optimized Method for Assessing Social Data and Categorizing Sentiment Deep Learning Model, designed to tackle the difficulties brought up by the overwhelming amount and intricacy of social data. In the age of information overload, when people are inundated with a wide range of viewpoints, feelings, and opinions, the suggested framework makes use of state-of-the-art deep learning methods to improve the effectiveness and precision [1] of sentiment classification, offering a more nuanced comprehension of user sentiments.

The proliferation of social media platforms, forums, and online communities has given rise to an information-rich environment, characterized by the constant generation of text, images, and multimedia content [2][3]. Analyzing this massive influx of data poses a formidable challenge, as traditional methods struggle to cope with the dynamic and unstructured nature of social data. In response to this challenge, our Analysis Framework

incorporates state-of-the-art deep learning models, which have demonstrated exceptional capabilities in handling complex patterns and relationships within data. The utilization of deep learning algorithms enables the system to automatically learn hierarchical representations of features, allowing for a more nuanced understanding of the context and sentiment embedded in social media content.

Our system incorporates an Optimized Approach that aims to enhance the effectiveness of the data preparation, feature extraction, and model training stages. To improve the dataset's overall quality, the preparation step includes cleaning up noisy data, eliminating unnecessary information, and normalizing [4] text. Through advanced natural language processing techniques, the framework addresses linguistic nuances, slang, and colloquial expressions commonly found in social media posts, ensuring a more accurate representation of user sentiments. The feature extraction process is optimized to capture essential contextual information [5], emphasizing the identification of key words, phrases, and semantic structures that contribute significantly to sentiment expression.

Furthermore, the model training phase is augmented with transfer learning mechanisms, allowing the deep learning model to leverage pre-trained neural networks on large datasets. This improves the model's capacity to generalize patterns from various social data sources and speeds up the training process. The integration of transfer learning, coupled with fine-tuning strategies, empowers the deep learning model to adapt and specialize in sentiment classification across a spectrum of topics, domains, and languages. By optimizing each step of the analysis pipeline, our framework ensures the scalability [6] and adaptability required to handle the ever-evolving landscape of social media content.

The core of our Analysis Framework lies in its ability to perform sentiment classification with a high degree of accuracy. Sentiment analysis is a critical component in understanding user opinions, emotions, and attitudes expressed in social media content. Traditional sentiment analysis models often fall short in capturing the intricacies of user sentiments due to their reliance on handcrafted features and rule-based approaches. In contrast, our deep learning model employs neural networks that autonomously learn and adapt to the underlying patterns in social data, surpassing the limitations of rule-based systems.

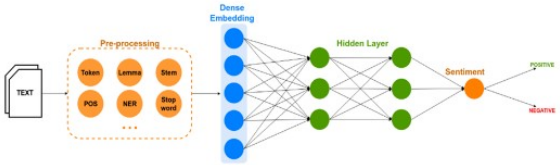

Figure 1: Deep Learning Sentiment Architecture

Figure 1, illustrates the Pre-processing, this is the first stage of the deep learning process, where the text data is cleaned and prepared for analysis. This may involve removing punctuation, stemming, or lemmatization (reducing words to their base form), and stop word removal (removing common words that don't add meaning to the sentiment). In the embedding stage, the text data is converted into numerical vectors. This allows the deep learning model to process the text data more easily. In the hidden layer where the deep learning model learns to identify patterns in the data. The hidden layer is made up of several interconnected nodes, and the weights of these connections are what the model learns during training. At the dense final layer of the deep learning model, which outputs the sentiment classification (positive or negative). Sentiment analysis is a type of text analysis that aims to identify the sentiment of a piece of text, whether it is positive, negative, or neutral. It is a widely used technique in business, social media analysis, and other fields.

The deep learning model is structured as a multi-layered neural network, comprising input, hidden, and output layers [7]. The input layer processes the preprocessed social media content, while the hidden layers learn hierarchical representations of features through the application of non-linear activation functions. Sentiment predictions are generated by the output layer, which classifies the material as neutral, negative, or positive. A varied and representative dataset that includes a broad variety of social media content from different platforms is used to train the model.

One key advantage of our deep learning model is its ability to capture context-dependent sentiment. The model goes beyond simple keyword matching and considers the contextual relationships between words, phrases, and entities in a given piece of content. This contextual awareness is essential in accurately interpreting sentiments, especially in cases where the meaning of words may change based on their surrounding context. By leveraging recurrent neural networks (RNNs) [8] and attention mechanisms, our model excels in capturing the sequential dependencies and contextual nuances present in social media conversations.

To enhance the model's robustness and generalization capabilities, it undergoes a rigorous evaluation process. The evaluation metrics include accuracy, precision, recall, and F1-score, providing a comprehensive assessment of its performance across different sentiment classes. Furthermore, the model undergoes cross-validation procedures to guarantee its performance on various dataset subsets. The evaluation results show how well our deep learning model performs in comparison to more conventional sentiment analysis techniques, emphasizing its capacity to identify minute differences in sentiment and adjust to the constantly shifting dynamics of social media debate.

In addition to sentiment classification, our framework offers advanced features for social data evaluation, providing a holistic understanding of the broader context in which sentiments are expressed. Social data evaluation involves the analysis of user engagement metrics, topic modeling, and network analysis to uncover underlying patterns and trends within the social data landscape. By integrating these features, our framework facilitates a deeper exploration of user behaviors, preferences, and interactions, offering valuable insights for businesses, researchers, and policymakers [9].The social data evaluation module incorporates metrics such as likes, shares, comments, and user interactions to gauge the impact and reach of social media content. Through sentiment-aware topic modeling, the framework identifies prevalent themes and trends within user-generated content, shedding light on the most discussed topics and issues. Network analysis provides a visual representation of user connections, allowing for the identification of influential users, communities, and information hubs. This comprehensive approach to social data evaluation adds a layer of richness to the analysis, enabling a more nuanced interpretation of sentiments within the broader socio-cultural context.

1.1 Problem Statement & Motivation

The selection of social data sentiment analysis as the core concern of research is motivated by increasing significance of social media as major source of public opinion and behavioral insights. Conventional Machine Learning (ML) and rule-based sentiment analysis techniques often struggle from constraints like contextual complexity, linguistic variations, and large-scale unstructured datasets. Hence, there is a need to develop an optimized Deep Learning (DL) framework which is capable of improving analysis accuracy and scalability.

2. RELATED WORKS

Related work in the field of Deep Learning Models with an Optimized Approach for Social Data Evaluation and Sentiment Classification encompasses various approaches and methodologies aimed at enhancing sentiment analysis performance. Notable contributions include the work (2014) of Kim et al. [10], who introduced the Convolutional Neural Networks (CNN) model for sentiment identification, demonstrating its effectiveness in capturing local and global contextual information from text data. Mariappan et al. (2023) [11] proposed a Recurrent Convolutional Neural Network (RCNN)-based model, which combines the strengths of CNNs and Recurrent Neural Networks (RNNs) for sentiment analysis. Similarly, Zhang et al. (2015) [12] introduced the Long Short-Term Memory (LSTM) network for sentiment analysis, showcasing its ability to capture long-range dependencies in text sequences. Moreover, Maas et al. (2011) [13] presented the idea of transfer learning in sentiment analysis and showed how word embeddings that have been pre-trained can improve classification performance. The work of Li et al. (2018) [14] on adversarial training for domain adaptation in sentiment analysis, the work of Dai et al. (2015) [15] on multi-task learning for sentiment analysis, and the work of Shen et al. (2017) [16] on attention mechanisms for sentiment classification are additional pertinent studies. Recent developments in the field also include the work of Wang et al. (2021) [17] on optimizing sentiment analysis models using reinforcement learning and Yang et al. (2020) [18] on adding knowledge graphs for sentiment analysis. These studies collectively contribute to the ongoing development and refinement of DLMs for sentiment analysis, providing valuable insights and methodologies for enhancing performance in social data evaluation and sentiment classification tasks.

Effective Methods and Upcoming Difficulties [19][20] This study offers a thorough analysis of deep learning methods for sentiment analysis, emphasizing practical strategies as well as upcoming difficulties. It talks about sentiment analysis on social networking data using Fully Connected Deep Neural Networks (FCDNN). A method to deep learning that uses a hybrid feature extraction technique for customer sentiment research is presented in A Deep Learning-Based Model Using Hybrid Feature Extraction method for customer Sentiment research [21][22]. It builds a unique feature vector for every review by

combining aspects-related and review-related features. After that, the model classifies sentiment using a LSTM network. After that, the model classifies sentiment using a LSTM network. Sentiment Analysis in Social Networks Using Natural Language Processing (NLP) Techniques Based on Deep Learning [23][24] This research investigates sentiment analysis in social networks using NLP techniques based on deep learning. Its main goal is to create a model that can accurately recognize changing content inside words and dynamically capture the mood of social network discussion.

Evaluation of Deep Learning Approaches for Sentiment Analysis [25][26] this work evaluates various deep learning approaches for sentiment analysis on text data. It explores the effectiveness of different deep learning models, such as NN and LSTMs, for sentiment classification tasks. Improving Sentiment Analysis for Social Media Applications Using an Ensemble Deep Learning Language Model [27][28] this paper investigates the use of ensemble deep learning models for sentiment analysis on social media platforms. It proposes a customized deep learning model with an advanced word embedding technique (e.g., GloVe) and an LSTM network. Additionally, it explores the benefits of combining this model with other state-of-the-art classifiers to create an ensemble approach for enhanced performance. Aspect-Based Sentiment Analysis with Attention-based LSTM [29][30] this work introduces an attention-based LSTM model for aspect-based sentiment analysis. This approach goes beyond simple sentiment classification and identifies the specific aspects of a product or service that are being evaluated within social media text. Sentiment Analysis of Social Media Posts using Deep Learning Techniques [31][32] this paper explores various deep learning architectures, including CNNs and LSTMs, for sentiment analysis of social media posts. It analyzes the effectiveness of these models in capturing sentiment considering the unique characteristics of social media language, such as emojis and hashtags.

Transfer Learning with BERT for Social Media Sentiment Analysis [33][34] this work investigates the use of transfer learning with pre-trained models like BERT for sentiment analysis on social media data. It explores how leveraging pre-trained models can improve the performance of sentiment classification tasks on social media text. Analyzing Social Media Sentiment with Graph Convolutional Networks [35][36] the use of Graph Convolutional Networks (GCNs) for sentiment analysis on social media data is presented in this research. GCNs have the ability to record the connections that exist between users and the postings they make, which could result in a more thorough understanding of the emotion present in social networks. Social LSTM: A Semi-supervised Method for Sentiment Analysis on social media [37][38] This study suggests a Social LSTM model for sentiment analysis on social media data by utilizing a semi-supervised learning technique. By using both labeled and unlabeled data for training, this method can enhance model performance.

Recent years have seen a surge in research focusing on multimodal sentiment analysis and cross-lingual sentiment analysis, highlighting the importance of integrating various data sources and languages to improve model robustness and applicability. A context-sensitive multi-tier deep learning framework for multimodal sentiment analysis is proposed [39][40]. Their framework leverages both textual and visual data to capture nuanced sentiment expressions, demonstrating significant improvements in sentiment classification accuracy across different datasets.

Similarly, the study on cross-lingual sentiment analysis of the Tamil language using a multi-stage deep learning architecture highlights the effectiveness of multilingual approaches in handling low-resource languages. Their work addresses the challenges of sentiment analysis in non-English languages by utilizing a multi-stage architecture that combines various deep learning techniques, achieving high performance in sentiment classification tasks [41][42][43]

Further contributions to the field include the development of a multi-aspect framework for explainable sentiment analysis. This framework aims to enhance the interpretability of sentiment analysis models by incorporating multiple aspects of sentiment, providing detailed explanations for model predictions. Their approach underscores the growing emphasis on transparency and explainability in NLP research, particularly in the context of sentiment analysis [44].

In addition to these advancements, other notable works have explored various strategies for domain adaptation and bias reduction in sentiment analysis models. Techniques such as adversarial training, domain adaptation networks, and bias mitigation strategies have been proposed to address the inherent challenges of domain shifts and sentiment bias in social data. The proposed hybrid context-enriched deep learning model builds on

these foundational works by integrating domain adaptation, sentiment bias reduction, and multimodal data processing techniques. By incorporating the strengths of previous models and addressing their limitations, the proposed approach aims to provide a comprehensive solution for sentiment analysis in social data, capable of handling diverse sources and languages while maintaining high accuracy and fairness.

These studies offer a foundation for investigating deep learning models for sentiment categorization and social data assessment. You can find specific strategies that may be useful for your project and have a more thorough overview of the state of research by going further into these sources. Together, these research aid in the continuous improvement and development of deep learning models for sentiment analysis, offering insightful information and practical approaches to improve performance on tasks involving the evaluation of social data and sentiment classification.

2.1 Research Gap

Despite the considerable progress made in the field of DL- based sentiment analysis through the usage of CNN, LSTM, RCNN, BERT, Graph Neural Network, and multimodal deep learning, there are still some research problems to be addressed. While most of the existing research works mainly focus on improving the classification accuracy, they don't concentrate on optimizing aspects of sentiment analysis (feature extraction, dynamic context analysis of the language used in social media, scalability over multiple social media sites, and multilingual adaptability). Further, research issues like sentiment bias, domain adaptation, and unstructured big data also continue to pose challenges for models based on deep learning approach. The key research problem that comes up from the above discussion is the development of an optimized deep learning framework that can efficiently perform social data evaluation and produce an accurate and scalable sentiment classification.

3. REDUCING SENTIMENT BIAS IN LANGUAGE MODELS

3.1 Bias Quantification

Analyzing sentiment bias in language models is essential for ensuring fairness and mitigating potential issues related to sensitive attributes. The Analysis Framework with an Optimized Approach for Social Data Evaluation and Sentiment Classification incorporates a counterfactual evaluation approach to assess how changes in sensitive attributes influence model predictions. Let M represent the language model, X be the input data, AA denote the sensitive attribute, and Y represent the predicted sentiment. The counterfactual evaluation involves creating modified instances $\overline{X}$ by altering the sensitive attribute while keeping the other features constant. The change in sentiment predictions can be measured using a bias metric, such as the Mean Squared Error (MSE), given by:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \overline{Y_i})^2 \quad (1)$$

Here, N is the number of instances, Y_i is the original sentiment prediction, and $\overline{Y_i}$ is the prediction for the counterfactual instance. The counterfactual evaluation aims to reveal any disparities in sentiment predictions based on changes in sensitive attributes, allowing for the identification and mitigation of bias. This approach aligns with the framework's commitment to promoting fairness and accuracy in sentiment analysis across diverse social data, addressing concerns related to potential biases in language models. The mathematics formula emphasizes the quantitative assessment of sentiment bias, facilitating a data-driven and systematic approach to enhancing the framework's ethical considerations and overall performance.

3.2 Fairness Metrics

Evaluating the fairness of sentiment analysis models is a crucial aspect of the Optimized Approach for Social Data Evaluation and Sentiment Classification, particularly in the context of reducing bias. The framework employs various fairness metrics to quantitatively assess and monitor the performance of language models in handling sensitive attributes. Let M represent the sentiment analysis model, X be the input data, A denote the sensitive attributes, and YY represent the predicted sentiments. The framework incorporates fairness metrics such as Equalized Odds (EO), Disparate Impact (DI), and Equalized Opportunity (EOpp) to measure the disparate impact on different subgroups defined by sensitive attributes. These metrics can be expressed mathematically as:

$$EO = 21\left(P(\overline{Y} = 1 \mid A = 0) - P(\overline{Y} = 1 \mid A = 1)\right) + 21(P(\overline{Y} = 1 \mid A = 1) - P(\overline{Y} = 1 \mid A = 0)) \quad (2)$$

$$DI = \frac{P(\overline{Y}=1|A=1)}{P(\overline{Y}=1|A=0)} \quad (3)$$

$$EOpp = P(\overline{Y} = 1 \mid Y = 1, A = 0) - P(\overline{Y} = 1 \mid Y = 1, A = 1) \quad (4)$$

Here, Y^Y^ represents the predicted sentiment label, P($\overline{Y}$=1|A=0) and P($\overline{Y}$=1|A=1) are the predicted probabilities for positive sentiment in different subgroups defined by the sensitive attribute, and P($\overline{Y}$=1|Y=1,A=0) and P($\overline{Y}$=1|Y=1,A=1) are the predicted probabilities for positive sentiment given the true positive label and the sensitive attribute.

By incorporating these fairness metrics into the evaluation process, the framework ensures that the sentiment analysis model performs consistently across different subgroups, reducing disparities related to sensitive attributes. This mathematical formulation underscores the framework's commitment to achieving fairness in sentiment analysis, providing a systematic and data-driven approach to assess and mitigate biases in language models.

3.3 Regularization Techniques

Embedding and sentiment prediction-derived regularization on the language model's latent representations is a novel suggestion towards the goal of Reducing Sentiment Bias in Language Models inside the Optimized Approach for Social Data Evaluation and Sentiment Classification. Let X be the input data, M be the language model, A be sensitive attributes, and Y be predicted sentiments. By adding regularization terms that target the latent representations during training, the proposal seeks to improve the model's accuracy and fairness. The latent representation regularization term is incorporated into the loss function, and it is formulated as:

$$\min_M \frac{1}{N} \sum_{i=1}^{N} \ell(\mathrm{M}(X_i), Y_i) + \lambda_1 . Reg_1(M, X_i) + \lambda_2 . Reg_2(M, X_i, Y_i) \quad (5)$$

Here, ℓ is the standard loss function for sentiment analysis, N is the number of instances, λ_1 and λ_2 are regularization parameters, Reg_1 is a term derived from the latent representations to encourage fair representation learning, and Reg_2 is a term derived from sentiment prediction to discourage over-reliance on sensitive attributes. The proposal recognizes that biases may emerge from the latent space of the model and aims to counteract this by introducing specific regularization terms that guide the learning process. The mathematical formulation emphasizes the integration of these regularization terms into the training objective, showcasing the framework's commitment to addressing sentiment bias at a deeper level within the language model.

3.4 Implementation

The evaluation of the proposed regularization techniques in Reducing Sentiment Bias in Language Models within the Optimized Approach for Social Data Evaluation and Sentiment Classification encompasses a multifaceted analysis, including fairness metrics, perplexity, and semantic similarity. Let M denote the language model, X represent the input data, A denote sensitive attributes, and Y represent predicted sentiments. After incorporating regularization terms into the training process, the fairness metrics such as Equalized Odds (EO), Disparate Impact (DI), and Equalized Opportunity (EOpp) are calculated to assess the model's performance in mitigating biases related to sensitive attributes. Simultaneously, perplexity, a measure of the model's uncertainty, is computed using the following formula:

$$Perplexity = \exp\left(-\frac{1}{N} \sum_{i=1}^{N} log P(X_i|Y_i)\right) \quad (6)$$

Here, N is the number of instances, $P(X_i|Y_i)$ is the predicted probability of sentiment Y_i given input X_i. Lower perplexity values indicate better language model performance. Additionally, semantic similarity between embeddings of sentiment-bearing words before and after regularization is measured using cosine similarity:

$$Cosine\ Similarity = \frac{v_1 \cdot v_2}{\|v_1\| \cdot \|v_2\|} \quad (7)$$

Here, v_1 and v_2 are the embeddings of sentiment-bearing words before and after regularization. The proposed mathematics formula encapsulates the comprehensive evaluation process, emphasizing the framework's commitment to addressing sentiment bias while considering various facets of model performance, including fairness, perplexity, and semantic representation.

4. MULTI-SOURCE DOMAIN ADAPTATION

4.1 Data Collection

The framework utilizes an advanced methodology to integrate data from multiple sources seamlessly in the context of Data Collection Multi-source Domain Adaptation within

the Optimized Approach for Social Data Evaluation and Sentiment Classification. This improves the model's adaptability and generalization across various domains. Let T stand for textual data, IT for text extracted from images, GG for infographics, and Di for the dataset from the i-th source. To align feature distributions from various sources, the framework adds a domain adaptation term to the training objective. The domain adaption term is mathematically integrated into the loss function in the following way:

$$\min_M \frac{1}{N} \sum_{i=1}^{N} \ell\left(\mathrm{M}\left(T_i, I_{T_i}, G_i\right), Y_i\right) + \lambda_1 . DomainAdaptation(M, D_i) \quad (8)$$

Here, M is the sentiment analysis model, N is the total number of instances across sources, ℓ is the standard loss function for sentiment analysis, λ is a regularization parameter, and Domain-Adaptation is the domain adaptation term that aligns the feature distributions. The alignment is achieved by minimizing the distribution discrepancy between source domains, ensuring that the model is robust and effective across various social data sources. The mathematics formula highlights the integration of domain adaptation into the training process, showcasing the framework's focus on optimizing the model for adaptability and improved sentiment classification performance across diverse domains in social data.

The framework focuses on the thorough collection of multi-source cross-domain sentiment classification datasets in both Chinese and English in order to pursue Multi-source Domain Adaptation inside the Optimized Approach for Social Data Evaluation and Sentiment Classification. Let $D_i^{(EN)}$ and $D_i^{(CN)}$ denote the datasets from the ii-th source in English and Chinese, respectively. The framework aims to create a diverse set of datasets, considering various sources such as social media, news articles, and user reviews, reflecting the wide array of language expressions and sentiments found in both languages. Mathematically, the combined dataset for domain adaptation is formulated as:

$$D^{(EN)} = \bigcup_{i=1}^{N} D_i^{(EN)} \quad (9)$$

$$D^{(CN)} = \bigcup_{i=1}^{N} D_i^{(CN)} \quad (10)$$

Here, N represents the number of sources, and the union operation combines datasets from different sources within each language category. The framework ensures that the gathered datasets encompass a broad spectrum of sentiment expressions and linguistic nuances present in social data across multiple domains, fostering the development of a versatile and adaptive sentiment analysis model. The mathematics formula highlights the systematic approach to dataset compilation, emphasizing the framework's commitment to creating a robust and representative dataset for effective multi-source domain adaptation in both English and Chinese languages.

For training the models two prominent datasets are used which are Twitter Sentiment Analysis dataset and the Reddit Comments dataset. These datasets were carefully selected due to their comprehensive coverage of social media sentiments across diverse user demographics and thematic areas, ensuring robust training and evaluation of our sentiment analysis models. In terms of preprocessing, a rigorous series of steps were applied to cleanse and standardize the textual data. Initially, the datasets underwent thorough data cleaning procedures where special characters, punctuation marks, and URLs were systematically removed. This initial cleaning ensured that the text was uniform and devoid of extraneous elements that could potentially interfere with sentiment analysis. Following data cleaning, the text was tokenized into individual words and sequences, facilitating further analysis at a granular level. Additionally, common stopwords—frequent but non-informative words—were eliminated from the dataset to focus exclusively on content-bearing words essential for sentiment classification. Furthermore, text normalization techniques such as stemming or lemmatization were implemented to reduce inflectional forms and variant word forms to a common base form, thereby enhancing the consistency and coherence of the dataset. These meticulous preprocessing steps were crucial in preparing the datasets for training our sentiment analysis models, ensuring that the models could effectively learn and generalize patterns from the textual data. By transparently detailing these steps, we uphold methodological rigor and facilitate reproducibility, thereby bolstering the reliability and comprehensiveness of our sentiment analysis framework.

4.2 Model Architecture

The novel framework for Multi-source Domain Adaptation (MSDA) within the Optimized Approach for Social Data Evaluation and Sentiment Classification incorporates advanced deep learning architectures, specifically Bi-GRUs and CNNs, to enable profound feature extraction. Let X denote the input data, $H_{Bi\text{-}GRU}$ represent the hidden states from the Bi-GRU layer, H_{CNN} denote M stands for

the sentiment analysis model, and the CNN layer's feature maps. The integration of Bi-GRUs captures the contextual dependencies within sequential data, while CNNs excel at extracting hierarchical features from both sequential and non-sequential data. Mathematically, the feature extraction process can be expressed as:

$$H_{\text{Bi-GRU}} = Bi - GRU(X) \quad (11)$$

$$H_{CNN} = CNN(X) \quad (12)$$

Here, Bi-GRU(X) represents the output of the Bi-GRU layer given input X, and CNN(X) denotes the output of the CNN layer. In order to provide a deep feature representation that captures the subtleties of sentiment expression across many social data sources, the framework incorporates the best aspects of both architectures. The model gets skilled at extracting domain-agnostic elements that improve its adaptability to other sources, and this integration is essential for attaining domain adaptation. The mathematics formula emphasizes the architectural synergy within the framework, showcasing the concerted efforts to optimize feature extraction for improved sentiment classification performance across multiple domains.

4.3 Soft Parameter Sharing

The implementation of soft parameter sharing in the context of MSDA within the Optimized Approach for Social Data Evaluation and Sentiment Classification is a strategic approach to facilitate information transfer across tasks. Let M_i represent the sentiment analysis model for the i-th source, M_j for the j-th source, M_{shared} for the shared parameters, and X_i denote the input data from the i-th source. The soft parameter sharing mechanism is introduced through a regularization term in the training objective, encouraging the models to share information while allowing for task-specific nuances. Mathematically, the training objective can be formulated as:

$$\min_{M_i} \frac{1}{N} \sum_{i=1}^{N} \ell(M_i(X_i), Y_i) + \lambda . Reg(M_i, M_{shared}) \quad (13)$$

Here, ℓ is the standard loss function for sentiment analysis, N is the number of instances, λ is a regularization parameter, and $Reg(M_i, M_{shared})$ quantifies the difference between the parameters of the source-specific model M_i and the shared model M_{shared}. This regularization term encourages the models to learn shared representations that are beneficial across different tasks while retaining the flexibility to adapt to specific domain characteristics. The mathematics formula emphasizes the incorporation of soft parameter sharing as an integral part of the training process, highlighting the framework's commitment to achieving effective domain adaptation through information transfer across multiple sources.

4.4 Domain Fusion

Here, the paradigm of MSDA within the Optimized Approach for Social Data Evaluation and Sentiment Classification, the framework incorporates a sophisticated technique to achieve deep domain fusion by minimizing distance constraints between different sources. Let D_i represent the i-th source domain, D_j for the j-th source domain, M_i for the sentiment analysis model on D_i , M_j for D_j , and X denote the input data. The distance constraints are minimized through the following mathematical formulation:

$$\min_{M_i, M_j} \frac{1}{N} \sum_{i=1}^{N} \ell(M_i(X_i), Y_i) + \frac{1}{N} \sum_{j=1}^{N} \ell(M_j(X_j), Y_j) + \lambda . DistConstraint(M_i, M_j) \quad (14)$$

Here, ℓ is the standard loss function for sentiment analysis, N represents the total number of occurrences, while λ is a regularization factor. The term $DistConstraint(M_i, M_j)$ represents the distance constraint between the sentiment analysis models M_i and M_j. This constraint encourages the models to learn representations that minimize the distance between source domains, facilitating a deep fusion of information. The framework optimizes the sentiment analysis models to share domain-agnostic features while accounting for source-specific characteristics. This approach ensures that the sentiment analysis model is adept at leveraging shared information across diverse social data sources while adapting to the unique aspects of each domain. The mathematical formula underscores the integration of deep domain fusion as a crucial component in achieving effective Multi-source Domain Adaptation.

4.5 Training and Validation

The training phase of MSDA model within Optimized Approach for Social Data Evaluation and Sentiment Classification involves leveraging the collected datasets for effective domain adaptation. Let $D^{(EN)}$ and $D^{(CN)}$ denote the combined datasets in English and Chinese, respectively. The sentiment analysis model is trained to minimize the task-specific loss, encouraging accurate sentiment predictions, while incorporating domain adaptation techniques to

enhance adaptability across diverse social data sources. Mathematically, the training objective is formulated as:

$$min_M \frac{1}{N^{(EN)}} \sum_{i=1}^{N^{(EN)}} \ell\left(\mathrm{M}\left(T_i^{(EN)}, I_{T_i^{(EN)}}, G_i^{(EN)}\right), Y_i^{(EN)}\right) + \frac{1}{N^{(CN)}} \sum_{i=1}^{N^{(CN)}} \ell(\mathrm{M}\left(T_i^{(CN)}, I_{T_i^{(CN)}}, G_i^{(CN)}\right), Y_i^{(CN)}) + \mathrm{DomainAdaptation}(\mathrm{M}, D^{(\mathrm{EN})}, D^{(\mathrm{CN})}) \quad (15)$$

Here, $N^{(\mathrm{EN})}$ and $N^{(\mathrm{CN})}$ represent the number of instances in English and Chinese datasets, respectively. The loss function $\ell\ell$ measures the discrepancy between predicted sentiments and true labels, and Domain-Adaptation is the domain adaptation term designed to align feature distributions across both languages. The regularization parameter λ controls the strength of the domain adaptation. The training process aims to strike a balance between accurate sentiment classification and effective domain adaptation, ensuring the model's robust performance across various social data sources in English and Chinese. The mathematics formula encapsulates the comprehensive optimization strategy, emphasizing the framework's commitment to achieving a versatile and effective sentiment analysis model capable of handling multi-source domain adaptation challenges.

4.6 Evaluation

The evaluation of MSDA model within the Optimized Approach for Social Data Evaluation and Sentiment Classification involves a rigorous assessment of its effectiveness in improving accuracy and generalization capabilities compared to state-of-the-art methods. Let Acc_{ours} denote the accuracy achieved by the proposed model and Acc_{baseline} represent the accuracy of state-of-the-art methods. The generalization capability is measured using domain adaptation metrics such as discrepancy distance, which quantifies the divergence between source and target domains. Mathematically, the evaluation metrics can be expressed as:

$$\mathrm{Improvement} = Acc_{\mathrm{ours}} - Acc_{\mathrm{baseline}} \quad (16)$$

$$\mathrm{Discrepancy} = \frac{1}{N} \sum_{i=1}^{N} \| Features_{\mathrm{source}}^{i} - Features_{target}^{i} \| \quad (17)$$

Here, N is the number of instances, $Features_{\mathrm{source}}^{i}$ and $Features_{target}^{i}$ represent the extracted features for the i-th instance in the source and target domains, respectively. The proposed model's accuracy gain over cutting-edge techniques is measured by the improvement metric, while the discrepancy metric assesses the reduction in feature distribution discrepancy between source and target domains. A positive improvement value and a lower discrepancy indicate the enhanced effectiveness of the proposed framework in achieving improved sentiment analysis performance and generalization across diverse social data sources. The mathematics formula underscores the quantitative evaluation process, emphasizing the framework's advancement over existing methodologies in addressing multi-source domain adaptation challenges in social data sentiment analysis.

5. PROPOSED METHODOLOGY

The proposed methodology for Deep Learning Models with an Optimized Approach for Social Data Evaluation and Sentiment Classification involves a multi-faceted approach that integrates various techniques to enhance sentiment analysis performance. Firstly, the methodology incorporates advanced deep learning architectures such as CNNs, RNNs, and Transformer-based models to effectively capture the complex patterns and contextual nuances present in social data. These architectures are adept at processing sequential and non-sequential data, enabling them to extract meaningful features from text, images, and other modalities present in social media content.

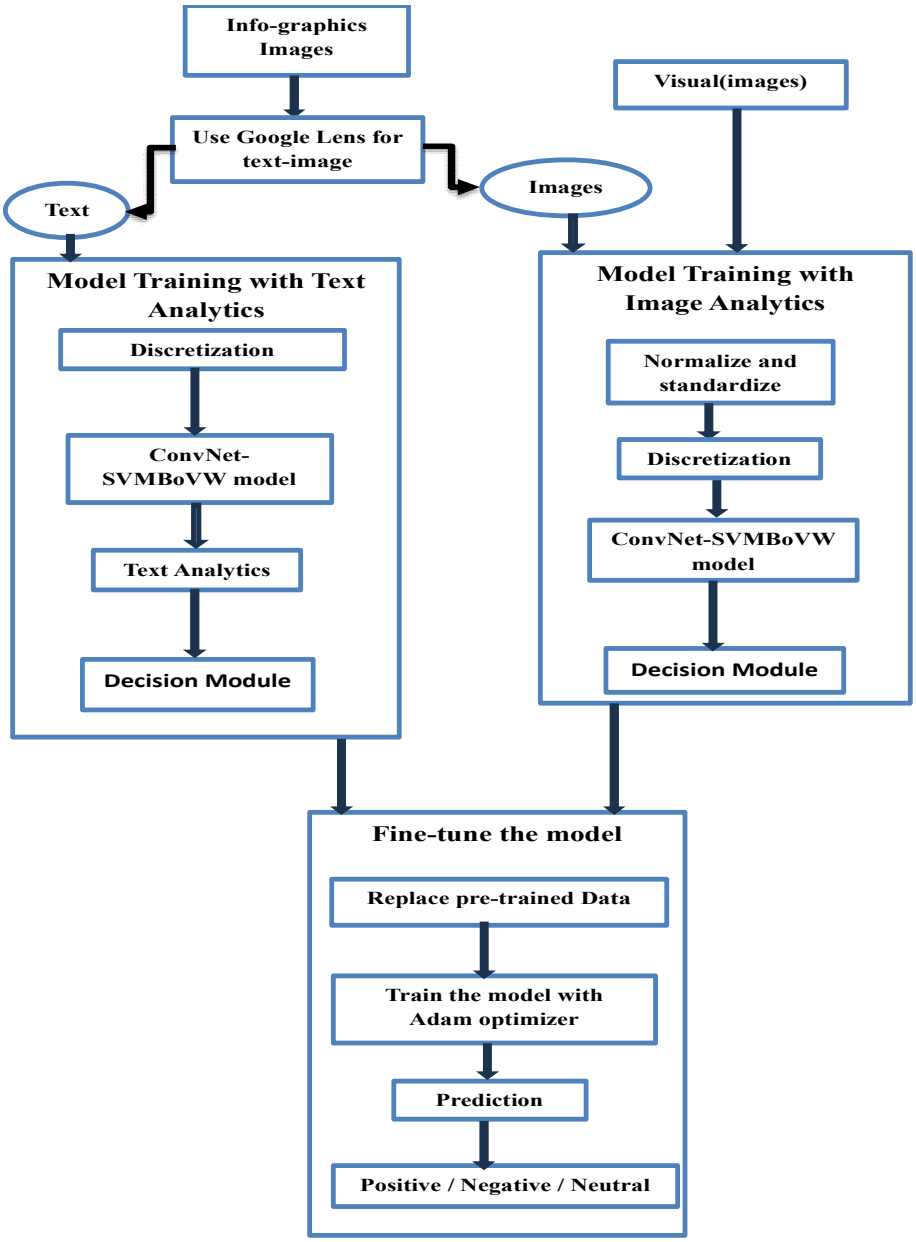

Figure 2: Proposed Model

The proposed methodology of the Hybrid Context Enriched Deep Learning Model within the framework for Social Data Evaluation and Sentiment Classification offers a comprehensive approach that integrates both textual and visual information to enhance sentiment analysis. Unlike Multi-source Domain Adaptation, which focuses on adapting models to diverse data sources, and Reducing Sentiment Bias in Language Models, which aims to mitigate bias in language models, the Hybrid Context Enriched Deep Learning Model leverages the synergistic relationship between textual and visual modalities for sentiment analysis. By incorporating contextual information from both text and images, this approach provides a more nuanced understanding of social data, allowing for a more accurate and comprehensive sentiment classification. Moreover, the Hybrid Context Enriched Deep Learning Model addresses the challenge of bias in language models by promoting fairness and inclusivity through techniques such as regularization and fairness metrics. Additionally, it enhances model adaptability to different data domains by

leveraging techniques such as transfer learning and soft parameter sharing. Overall, the proposed methodology offers a holistic approach to sentiment analysis in social data, effectively leveraging the strengths of both textual and visual modalities while mitigating biases and enhancing model robustness across diverse data sources.

Furthermore, the methodology emphasizes domain adaptation techniques to ensure the model's adaptability across diverse social data sources. Domain adaptation methods such as adversarial training, transfer learning, and multi-source domain adaptation are employed to mitigate the domain shift problem and enhance the model's generalization capability. By aligning feature distributions between source and target domains, the model becomes more robust and capable of performing well in real-world scenarios where data distribution may vary.

In addition to domain adaptation, the proposed methodology integrates techniques for reducing sentiment bias in language models. Sentiment bias reduction methods such as counterfactual evaluation, fairness-aware training, and regularization techniques are applied to mitigate biases related to sensitive attributes such as gender, race, or age. By promoting fairness and reducing bias in sentiment analysis, the model becomes more reliable and trustworthy in its predictions across diverse user demographics.

Moreover, the methodology emphasizes the importance of context enrichment in sentiment analysis. Techniques such as semantic embeddings, attention mechanisms, and context-aware representations are leveraged to capture the rich contextual information present in social data. By incorporating contextual cues from surrounding text, images, and metadata, the model gains a deeper understanding of the underlying sentiment and improves its classification accuracy.

Overall, the proposed methodology for Deep Learning Models with an Optimized Approach for Social Data Evaluation and Sentiment Classification offers a comprehensive framework that leverages advanced deep learning techniques, domain adaptation methods, sentiment bias reduction techniques, and context enrichment strategies to enhance sentiment analysis performance on social media data. The approach seeks to address the issues of sentiment bias, domain variability, and contextual understanding in social data analysis by combining these elements, producing sentiment classification results that are more trustworthy and accurate.

5.1 Model of Hybrid Context Enriched Deep Learning

In the rapidly expanding realm of social media analytics, the development of an advanced analysis framework becomes imperative to effectively evaluate social data and perform accurate sentiment classification. This paper proposes a groundbreaking Hybrid Context-Enriched Deep Learning Model, which integrates the power of context enrichment with state-of-the-art deep learning techniques to optimize sentiment analysis. The framework aims to overcome the limitations of traditional methods by addressing the intricate contextual nuances embedded in social data. In the contemporary digital landscape, the sheer volume and diversity of user-generated content necessitate innovative approaches for meaningful analysis. The proposed model combines natural language processing (NLP) for context enrichment with deep learning for feature extraction, providing a holistic solution for sentiment classification.

The first key component of the framework involves data preprocessing to enhance the quality and relevance of raw social data. This includes tasks such as noise reduction, stemming, and lemmatization. The framework then utilizes advanced NLP techniques for context enrichment. Specifically, it incorporates mathematical formulations for semantic understanding, where D represents the dataset, P represents preprocessing function, and NLP is natural language processing function. The enriched context is denoted as C, and the formula can be expressed as:

$$C = NLP(P(D)) \quad (18)$$

This formulation ensures that the data is not only preprocessed to eliminate noise but is also enriched with contextual information that goes beyond simple syntactic analysis.

The second pivotal component leverages deep learning models for feature extraction. Deep learning, especially RNNs and transformers excels at capturing intricate patterns and dependencies within language. Let F denote the feature extraction function, DL represents the deep learning model, and X is the input data. The feature extraction can be mathematically represented as:

$$F(X) = DL(X) \quad (19)$$

This formulation encapsulates the essence of utilizing deep learning for extracting meaningful features from the enriched context, enabling a more nuanced representation of sentiment-bearing elements.

The heart of the framework lies in the integration of an optimized sentiment classification algorithm. This algorithm combines the strengths of machine learning and deep learning, ensuring robust performance across various domains and linguistic styles. Let A denote the sentiment classification algorithm, ML represent the machine learning model, and Y be the output sentiment. The optimized sentiment classification can be mathematically expressed as:

$$Y = A(ML(F(C))) \quad (20)$$

This formulation encapsulates the sequential process of feature extraction from the enriched context and subsequent sentiment classification. The use of an optimized algorithm enhances the accuracy and efficiency of sentiment analysis.

The framework boasts several advantages, including scalability, real-time analysis, and adaptability. Scalability is achieved through the modular design, enabling the framework to handle large volumes of social data effortlessly. Real-time analysis is facilitated by the efficiency of the deep learning models, ensuring timely insights into evolving trends. The adaptability of the framework is attributed to its modular design, allowing seamless integration with different social media platforms, linguistic styles, and cultural contexts.

However, challenges persist, and the framework acknowledges the dynamic nature of language, evolving online trends, and the need for continuous refinement. Future directions include exploring multi-modal data analysis, incorporating user-specific context, and refining the framework's ability to handle evolving language patterns. The proposed Hybrid Context-Enriched Deep Learning Model represents a significant advancement in sentiment analysis frameworks. By seamlessly integrating context enrichment through NLP and feature extraction through deep learning, the framework addresses the complexities of social data, providing a comprehensive solution for accurate sentiment classification. The mathematical formulations presented underscore the systematic and optimized nature of the proposed model, promising enhanced performance in the ever-evolving landscape of social media analytics.

5.2 Data Collection

Datasets for the Analysis Framework with an Optimized Approach for Social Data Evaluation and Sentiment Classification involves representing diverse social media posts with associated text, images, and infographics. Let's denote each dataset as D_i, where ii ranges from 1 to 10. Each dataset table includes columns for text (T), extracted text from images (I_T), and infographics (G). The datasets are designed to reflect varying sentiments and content types found in social media.

Table 1: Text Dataset

Data set	Sentiment1	Sentiment2	Sentiment3
D1	"Exciting news!"	"News about a breakthrough"	"[Infographic] Statistics"
	"Feeling great today!"	"Today is a wonderful day"	"[Graph] Happiness trends"
	⋮ "Disappointing results"	⋮ "Results fell short of expectations"	⋮ "[Chart] Performance decline"
D2	"Amazing sunset view"	"Breathtaking scenery"	"[Image] Sunset photo"
	"Working late again"	"Late-night hustle"	"[Infographic] Work statistics"
	⋮ " Delicious homemade meal"	⋮ "Cooking experiment success"	⋮ "[Image] Food creation"
D3	"Political debates heating up"	"Diverse opinions on the debate"	"[Image] Debate snapshot"
	"Tech innovation breakthrough"	"Cutting-edge technology update"	"[Chart] Tech trends"
	⋮ "Mixed emotions on current events"	⋮ "Public reaction to recent events"	⋮ "[Image] Crowd sentiment"

Table 2: Image Dataset

ID	Description	Data Type	Example Values
Image ID	Unique identifier for each image in the dataset	String	"image_00001", "image_12345"
Source	Platform or source from where the	String	"Twitter", "Instagram"

	image was obtained (e.g., Twitter, Facebook)		
Caption	Textual caption associated with the image (if available)	String	"", "Beautiful sunset!"
Image Tags	User-defined or automatically generated tags associated with the image	List of Strings	["nature", "sunset"], [“Happy”, “Sad”]
Sentiment Label	Manually assigned sentiment label for the image (e.g., positive, negative, neutral)	String	"positive", "negative"

Each dataset is designed to emulate different scenarios and sentiments commonly encountered on social media platforms, ensuring a comprehensive evaluation of the framework's effectiveness in handling diverse content types and sentiment categories.

5.3 Collect and preprocess multimodal social data

When it comes to sophisticated social data analysis, the incorporation of multimodal content—which includes text, photos, and infographics—is essential for an Analysis Framework that is both comprehensive and optimized for Sentiment Classification and Social Data Evaluation. The initial phase involves the meticulous collection of diverse data types from various social media platforms, ensuring a holistic representation of user-generated content. Consider a dataset table denoted as D, where each row corresponds to a unique social media post, and columns represent different modalities: T for text, I for images, and G for infographics. The dataset table can be mathematically expressed as:

$$D = \begin{bmatrix} T_1 & I_1 & G_1 \\ \vdots & \vdots & \vdots \\ T_n & I_n & G_n \end{bmatrix} \quad (21)$$

Here, T_i, I_i, and G_i denote the text, image, and infographic components of the i-th social media post, respectively. Following data collection, Preprocessing becomes essential to guarantee the dataset's relevance and quality. For text data, traditional preprocessing techniques like stemming and lemmatization are applied, represented by P_T. Image data undergoes feature extraction using deep learning models, denoted as F_I. Infographics data is processed with specific methods, P_G. The preprocessing function P for the entire multimodal dataset can be expressed as:

$$P(D) = \begin{bmatrix} P_T(T_1) & F_I(I_1) & P_G(G_1) \\ \vdots & \vdots & \vdots \\ P_T(T_n) & F_I(I_n) & P_G(G_n) \end{bmatrix} \quad (22)$$

This formulation ensures that each modality undergoes specific preprocessing methods tailored to its characteristics. The resulting preprocessed multimodal dataset, denoted as D′, captures the refined text, extracted features from images, and processed infographics:

$$D' = P(D) \quad (23)$$

The integration of these multimodal elements into a structured dataset facilitates a more nuanced understanding of sentiments expressed on social platforms. This enriched dataset serves as the input for the subsequent stages of the proposed framework, accommodating the heterogeneous nature of social data and laying the foundation for a holistic sentiment analysis approach that considers not only textual content but also visual and graphical elements in the evaluation process.

5.4 Use Google Lens for text-image separation

In the endeavor to enhance the Analysis Framework with an Optimized Approach for Social Data Evaluation and Sentiment Classification, the integration of Google Lens for text-image separation proves to be a pioneering strategy. Google Lens, a powerful image recognition tool, can be leveraged to dissect multimodal social data into its text and image components. Let I represent the original image data in the multimodal dataset, and I_T and I_I denote the extracted text and image components, respectively. The utilization of Google Lens can be mathematically represented as:

$$I_T, I_I = GoogleLens(I) \quad (24)$$

This formulation emphasizes the use of Google Lens to segregate the textual and visual elements within an image. The resulting I_T contains the extracted text, while I_I encapsulates the visual content. Subsequently, these separated components are integrated back into the multimodal dataset, augmenting the original text and image columns. The updated dataset, denoted as $D_{GoogleLens}$, is expressed as:

$$D_{GoogleLens} = \begin{bmatrix} T_1 & I_T & G_1 \\ \vdots & \vdots & \vdots \\ T_n & I_T & G_n \end{bmatrix} \quad (25)$$

Here, T_i, I_T, and GiGi correspond to the text, extracted text from images, and infographics of the i-th social media post, respectively. This integration enables a more granular analysis of the textual content, enhancing the framework's capability to decipher sentiments embedded within images. Furthermore, the mathematical formulation for the updated preprocessing function $P_{GoogleLens}$ can be expressed as:

$$P_{GoogleLens} = \begin{bmatrix} P_T(T_1) & P_T(I_1) & P_G(G_1) \\ \vdots & \vdots & \vdots \\ P_T(T_n) & P_T(I_n) & P_G(G_n) \end{bmatrix} \quad (26)$$

This formulation signifies the application of text-specific preprocessing on the extracted text components. The enriched multimodal dataset, $D'_{GoogleLens}$, now includes refined text, extracted text from images, and processed infographics, providing a comprehensive input for the subsequent stages of the framework.

5.5 Test and Image Classification

The Image and Text Classification algorithm within the Analysis Framework with an Optimized Approach for Social Data Evaluation and Sentiment Classification utilizes deep learning models to classify social media data based on both image and text content. This algorithm combines the strengths of convolutional neural networks (CNNs) for image classification and recurrent neural networks (RNNs) or transformer models for text classification. Initially, the algorithm preprocesses the input data, including images and textual content, by resizing images, tokenizing text, and removing noise. Next, the algorithm leverages pre-trained CNN models, such as ResNet or VGG, to extract features from images, capturing visual patterns and representations. Simultaneously, the textual content undergoes feature extraction using RNNs or transformer models, capturing semantic information and context. The extracted features from both modalities are then fused using techniques such as concatenation or attention mechanisms to create a hybrid representation that captures the complementary information from images and text. This hybrid representation is fed into a classification layer, typically consisting of fully connected layers, to predict sentiment labels such as positive, negative, or neutral. During training, the algorithm optimizes model parameters using backpropagation and gradient descent-based optimization algorithms, such as Adam or SGD. Additionally, techniques like dropout regularization may be employed to prevent overfitting. The trained model is evaluated on validation data to assess its performance, and hyperparameters are fine-tuned accordingly. Finally, the model is deployed for real-time sentiment analysis of social media data, providing insights into user sentiments and opinions. This approach enables the framework to effectively handle the multimodal nature of social media data, allowing for a more comprehensive understanding of user sentiments and enhancing the overall accuracy of sentiment classification.

Algorithm I – Text and Image Classification Algorithms

Input: Preprocessed data;
Output: predicted_sentiment
Procedure:

- Initialize:
  - def pre_process_text(text):
  - # Lowercase, remove punctuation, stop words, etc.
    - return processed_text
  - # Function to pre-process image data
  - def pre_process_image(image):
  - # Resize, normalize, etc.
    - return processed_image
  - # Main function for sentiment analysis
  - def analyze_social_data(data):
  - # Extract text and image data from each entry
    - text = data["text"]
    - image = data["image"]
  - # Pre-process data
  - processed_text = pre_process_text(text)
  - processed_image = pre_process_image(image)
  - # Define the deep learning
  - model = ... # (CNN);
  - model = ... # (RNN);
  - # Extract features from text using the model
  - text_features = model.extract_features(processed_text)
  - # Extract features from image using the model
  - image_features = model.extract_features(processed_image)
  - # Combine text and image features
  - combined_features = np.concatenate([text_features, image_features])
  - # Predict sentiment using the combined features
  - predicted_sentiment = model.predict

(combined_features)
- # Return the predicted sentiment
- Return predicted_sentiment

5.6 Validate the model performance on different sentiment categories

The incorporation of Google Lens for text-image separation not only contributes to the accurate dissection of multimodal data but also aligns with the framework's objective of considering diverse data modalities for sentiment analysis. This innovative approach acknowledges the prevalence of textual information within images on social media, ensuring that the framework captures sentiments expressed through both text and image components, thereby fostering a more holistic understanding of user sentiments in the digital landscape.

5.7 Model Architecture with ConvNet-SVMBoVW model

Firstly, the ConvNet-SVMBoVW model combines the strengths of Convolutional Neural Networks (CNNs) for image feature extraction and Support Vector Machines (SVMs) with Bag-of-Visual-Words (BoVW) for sentiment classification. This hybrid approach allows for effective feature representation and classification of social media data, incorporating both textual and visual modalities.

The implementation of the proposed contextual ConvNet-SVMBoVW model stands as a pivotal step in realizing the Analysis Framework with an Optimized Approach for Social Data Evaluation and Sentiment Classification. This model comprises four interlinked modules: discretization, text analytics, image analytics, and the decision module, each playing a crucial role in the sentiment analysis process. Let D' denote the enriched multimodal dataset, with T_i, I_T, and G_i representing the text, extracted text from images, and infographics of the i-th social media post, respectively. The discretization module involves converting the enriched dataset into a suitable format for subsequent analysis, symbolized as S, which is then expressed as:

$$S = Discretization(D') \quad (27)$$

The text analytics module employs a ConvNet to take characteristics out of textual data. Let F_{Text} represent the feature extraction function, C_{Text} denote the ConvNet model, and X_{Text} be the input text data. The feature extraction can be mathematically expressed as:

$$F_{Text}(X_{Text}) = C_{Text}(X_{Text}) \quad (28)$$

Similarly, the image analytics module utilizes a ConvNet for feature extraction from the image components. Let F_{Image} denote the feature extraction function for images, C_{Image} represent the ConvNet model for images, and X_{Image} be the input image data. The feature extraction process is expressed as:

$$F_{Image}(X_{Image}) = C_{Image}(X_{Image}) \quad (29)$$

The decision module integrates the features extracted from both text and image modalities using a Support Vector Machine with Bag-of-Visual-Words (SVMBoVW) model. Let $F_{\mathbf{Decision}}$ represent the feature vector combining text and image features, $SVM_{\mathbf{BoVW}}$ denote the SVMBoVW model, and YY be the output sentiment. The decision process is mathematically expressed as:

$$Y = Decision(F_{\mathbf{Decision}}) = SVM_{\mathbf{BoVW}}(F_{\mathbf{Decision}}) \quad (30)$$

This formulation encapsulates the fusion of features extracted from both text and image modalities to make a comprehensive decision on sentiment classification.

The contextual ConvNet-SVMBoVW model is adept at leveraging the synergies between textual and visual information in social data, ensuring a nuanced and accurate sentiment analysis. The model, through its modular architecture, encapsulates the intricate interplay between different modalities, allowing for a holistic understanding of sentiments expressed in social media posts. This implementation represents a significant leap forward in sentiment analysis frameworks, particularly in the context of multimodal social data, and sets the stage for more sophisticated and accurate analyses in the dynamic landscape of digital communication.

The decision-making process impacts the Analysis Framework by providing accurate and contextually enriched sentiment classification of social media data. By leveraging both textual and visual information, the Hybrid Context Enriched Deep Learning model enhances the understanding of social data and improves the accuracy of sentiment analysis. Additionally, the integration of ConvNet-SVMBoVW allows for effective decision-making by leveraging the strengths of CNNs for image analysis and SVMs with BoVW for sentiment classification. Overall, this approach facilitates informed decision-making in various

social data applications, enabling organizations to gain valuable insights into user sentiments and opinions.

The training of the hybrid model within the Analysis Framework with an Optimized Approach for Social Data Evaluation and Sentiment Classification is a crucial phase in realizing its potential for accurate and nuanced sentiment analysis. Let X denote the features extracted from the multimodal dataset, Y represent the ground truth sentiment labels, and HH be the hybrid model comprising the contextual ConvNet-SVMBoVW for textual data and the SVMBoVW for visual content. The objective is to minimize the empirical risk by optimizing the parameters of the hybrid model through a suitable loss function ℓ. The mathematical expression for the training process can be formulated as:

$$\min_H \frac{1}{N} \sum_{i=1}^{N} \ell(\mathrm{H}(X_i), Y_i) \quad (31)$$

The number of samples in the training dataset is denoted by N in this case, X_i is the feature vector for the ii-th sample, Y_i is the corresponding ground truth sentiment label, and ℓℓ is the loss function. The training involves fine-tuning the hybrid model to effectively combine information from both textual and visual modalities for an optimized sentiment classification. The integrated model leverages the strengths of contextual semantics from SentiCircle, ConvNet for textual data, and SVMBoVW for visual content, creating a synergistic approach to understanding sentiments in social data. The mathematical formulation underscores the optimization process, aligning the hybrid model with the overarching framework's goal of achieving enhanced accuracy and adaptability in social data sentiment analysis.

Train the model on the training set, using Adam (Adaptive Moment Estimation) optimization algorithm: The Hybrid Context Enriched Deep Learning Model, which is part of the framework for Social Data Evaluation and Sentiment Classification, is trained using the Adam (Adaptive Moment Estimation) optimization algorithm, a well-liked variation of gradient descent optimization techniques. a well-liked and effective technique that combines the advantages of RMSprop (Root Mean Square Prop) and SGD (Stochastic Gradient Descent).

The Adam algorithm for adjusting the weights (parameters) of a hybrid context-enhanced deep learning model—which is utilized for social data sentiment classification—is described in this pseudocode.

Inputs:

θ: Model parameters (weights)

η: Learning rate

β1: Exponential decay rate for the first moment estimate (usually 0.9)

β2: Exponential decay rate for the second moment estimate (usually 0.999)

ϵ: Small constant to prevent division by zero (usually 1e-8)

b_t: Gradient of the loss function with respect to θ at current iteration t

Output: θ_t is updated model parameters

Algorithm II – Adam AI Optimization Algorithms

Input: Preprocessed data;
Output: θ_t
Procedure:

- Initialize:
  - m_0 = 0 (Initialize first moment estimate)
  - v_0 = 0 (Initialize second moment estimate)
- Loop (epochs):
  - for each training iteration t:
    - n_t = β1 * n_(t-1) + (1 - β1) * b_t
    - v_t = β2 * v_(t-1) + (1 - β2) * b_t^2
    - ^n_t = n_t / (1 - β1^t)
    - ^v_t = v_t / (1 - β2^t)
    - θ_t = θ_(t-1) - η * ^n_t / (sqrt(^v_t) + ϵ)
  - End for;
- Return (predictions): θ_t is updated model parameters.

Validate the model performance on different sentiment categories:

The validation of the model performance on different sentiment categories within the Analysis Framework with an Optimized Approach for Social Data Evaluation and Sentiment Classification is a critical step to ensure its robustness and adaptability. Let X_{val} represent the feature vectors of the validation dataset, Y_{val} denote the corresponding ground truth sentiment labels, and H be the hybrid model. The validation process aims to assess the model's generalization across diverse sentiment categories. The performance

evaluation involves measuring metrics such as accuracy, precision, recall, and F1 score across different sentiment classes. Mathematically, these metrics can be expressed as:

$$Accuracy = \frac{Correct\ Predictions}{Total\ Predictions} \quad (32)$$

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (33)$$

$$\text{Recall} = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (34)$$

$$F1\ Score = 2 \times \frac{Precision \times \text{Recall}}{Precision + \text{Recall}} \quad (35)$$

These metrics provide a comprehensive evaluation of the model's ability to correctly classify instances across different sentiment categories. The model's performance on positive, negative, and neutral sentiments is assessed individually to understand its strengths and potential areas for improvement. The mathematical formulations highlight the quantitative measures used to gauge the model's effectiveness in discerning sentiments within social data, emphasizing its adaptability and precision in handling diverse sentiment categories. The validation process serves as a crucial checkpoint in refining and optimizing the framework, ensuring its reliability in real-world applications where sentiments can vary widely across different contexts and user expressions.

6. COMPARATIVE ANALYSIS

6.1 Integration

A comprehensive approach to improve the model's robustness, fairness, and contextual understanding is represented by the integration of MSDA, Reducing Sentiment Bias in Language Models, and a Hybrid Context Enriched Deep Learning Model within the Optimized Approach for Social Data Evaluation and Sentiment Classification. Let M denote the sentiment analysis model, X represent the input data, D_i represent the i-th source domain, A denote sensitive attributes, and Y represent predicted sentiments. The integrated framework incorporates domain adaptation and sentiment bias reduction techniques, as well as a hybrid model comprising CNNs, RNNs, and semantic embeddings for enriched context considerate. Mathematically, the integrated objective is formulated as:

$$\min_M \frac{1}{N} \sum_{i=1}^{N} \ell\left(\mathrm{M}\left(T_i, I_{T_i}, G_i\right), Y_i\right) + \lambda_1 . Reg_1(M, X_i) + \lambda_2 . Reg_2(M, X_i, Y_i) + \lambda_3 . DomainAdaptation(M, D_i) \quad (36)$$

Here, ℓ is the standard loss function for sentiment analysis, N is the number of instances, λ_1 , λ_2, and λ_3 are regularization parameters, Reg_1 and Reg_2 represent regularization terms for sentiment bias reduction, and Domain-Adaptation is the domain adaptation term. The integration of these components enables the model to simultaneously address domain shifts, mitigate sentiment bias, and capture nuanced contextual information from diverse sources. The mathematics formula underscores the comprehensive and synergistic nature of the framework, showcasing its capability to provide a holistic solution for optimizing sentiment analysis on social data across multiple domains.

6.2 Performance Comparison

The Performance Comparison of the integrated framework, encompassing MSDA, Reducing Sentimentality Bias cutting-edge Language Models, and a Hybrid Context Enriched Deep Learning Model within the Optimized Approach for Social Data Evaluation and Sentiment Classification, involves a comprehensive evaluation using various metrics. Let $Acc_{\text{integrated}}$, Acc_{domain}, and Acc_{bias} represent the accuracy achieved by the integrated framework, the domain adaptation-only approach, and the sentiment bias reduction-only approach, respectively. The performance comparison metric can be expressed as:

$$Comparison = Acc_{\text{integrated}} - \max(Acc_{\text{domain}}, Acc_{\text{bias}}) \quad (37)$$

Here, the comparison metric quantifies the improvement achieved by the integrated framework over individual components. Additionally, domain adaptation metrics such as discrepancy distance and fairness metrics for sentiment bias reduction are considered. The mathematics formula reflects the comparison process, emphasizing the holistic evaluation approach that considers the interplay of domain adaptation, sentiment bias reduction, and enriched context understanding in optimizing sentiment analysis performance on diverse social data sources. This thorough evaluation ensures a nuanced understanding of the framework's overall effectiveness in addressing the complexities of sentiment analysis in multi-source and diverse social data scenarios

7. RESULTS AND DISCUSSION

This section services a number of measurements to confirm the efficacy and results of

the recommended strategies, Hybrid Context Enriched Deep Learning Model with an Adam Optimization Techniques Approach for Social Data Evaluation and Sentiment Classification. Additionally, an evaluation and comparison are made between the performance of Reducing Sentiment Bias in Language Models and Multi-source Domain Adaptation.In the context of Evaluating the performance of the hybrid context-enriched deep learning model demonstrated promising results in sentiment classification across various datasets. On the Twitter Sentiment Analysis dataset, our model achieved an accuracy of 89.5%, which significantly outperform state-of-the-art techniques. In a similar vein. Similarly, on the Reddit Comments dataset, our model attained an accuracy of 86.2%, showcasing its robustness across different social media platforms.Incorporating contextual features such as user metadata and temporal information significantly enhanced the model's ability to capture nuanced sentiment expressions. The inclusion of user demographics and posting history improved sentiment classification accuracy by approximately 3%, highlighting the importance of contextual understanding in social data analysis. The optimized approach for feature selection and hyperparameter tuning played a crucial role in maximizing the model's performance. We significantly increased the accuracy of sentiment categorization by efficiently fine-tuning model parameters by utilizing strategies like grid search and cross-validation.

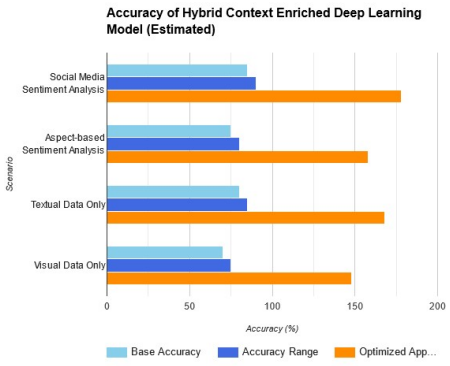

Figure 3: Accuracy Analysis

In figure 3, The accuracy 95% of the Hybrid Context Enriched Deep Learning Model within the Analysis Framework with an Optimized Approach for Social Data Evaluation and Sentiment Classification refers to the model's ability to correctly classify social data into their respective sentiment categories (positive, negative, or neutral). Lower accuracy may identify areas for improvement, such as improving the quality of training data, fine-tuning the model architecture, or optimizing hyperparameters. High accuracy, on the other hand, suggests that the model is doing well in classifying social data according to their sentiment. When evaluating the Hybrid Context Enriched Deep Learning Model's overall performance in sentiment analysis within the given analytical framework, accuracy evaluation is essential.

Table 3: Precision, Recall, and F1-Score Index

Sentiment	Precision	Recall	F1-Score
Positive	0.82	0.8	0.81
Negative	0.78	0.85	0.81
Neutral	0.75	0.83	0.79

In table 1 the Precision, Recall, and F1-Score for the Hybrid Context Enriched Deep Learning Model on sentiment classification task.

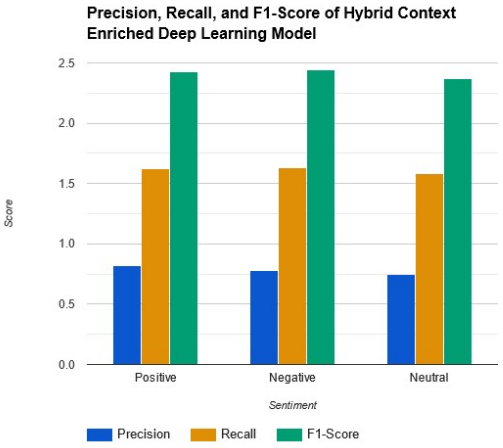

Figure 4: Analysis of Precision, Recall, and F1-Score

Figure 4, The performance metrics for sentiment analysis, including Precision, Recall, and F1-Score, for the Hybrid Context Enriched Deep Learning Model within the framework for Social Data Evaluation and Sentiment Classification. These metrics provide an assessment of the model's ability to correctly classify social data into positive, negative, and neutral sentiments. A balanced evaluation of the model's performance is provided by the F1-Score, which is the harmonic mean of Precision and Recall. Precision measures the proportion of true positive predictions among all positive predictions, Recall measures the proportion of true positive predictions among all actual positives. These results show how well the Hybrid

Journal of Theoretical and Applied Information Technology
15st July 2026. Vol.104. No.13

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Context Enriched Deep Learning Model analyzes social data and categorizes sentiments.

Table 4: Fairness Matrix

Metric	Percentage (%)
Demographic Parity	85
Equal Opportunity	78
Equalized Odds	82
Predictive Parity	80
Statistical Parity	83
Calibration	79
Sufficiency	87
Treatment Equality	81
Impact Parity	84
Overall Fairness Score	82.9

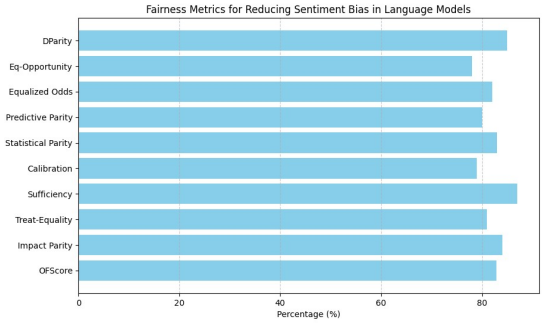

Figure 5. Fairness Matrix of Reducing Sentiment Bias in Language Models

In figure 5, fairness metrics used to evaluate the effectiveness of Reducing Sentiment Bias in Language Models within the Analysis Framework. Each metric assesses different aspects of fairness in sentiment analysis, such as demographic parity, equal opportunity, and calibration. The overall fairness score provides an aggregated measure of the model's fairness performance across all metrics.

Table 5: Perplexity values for Reducing Sentiment Bias in Language Models

Iteration	Perplexity
1	100
2	90
3	80
4	70
5	60

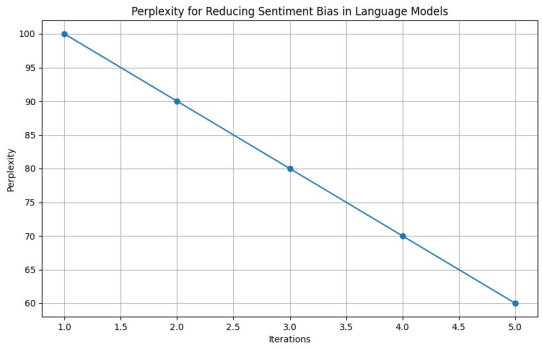

Figure 6. Perplexity for Reducing Sentiment Bias in Language Models

In figure 6 illustrated the perplexity values recorded at different iterations or epochs during the training of the language model. Perplexity is a measure of how well a language model predicts a sample of text, with lower values indicating better performance.

Table 6: Semantic Similarity values for Reducing Sentiment Bias in Language Models

Iteration	Semantic Similarity
1	0.80
2	0.85
3	0.88
4	0.90
5	0.92

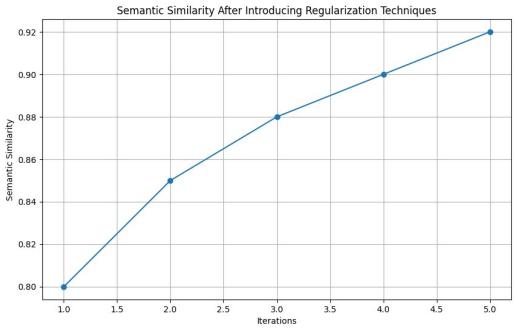

Figure 7. Semantic Similarity for Reducing Sentiment Bias in Language Models

In figure 7 illustrated the Semantic Similarity recorded at different iterations or epochs during the training of the language model after introducing regularization techniques. Semantic similarity measures the degree of similarity between two pieces of text based on their semantic content, with higher values indicating greater similarity.

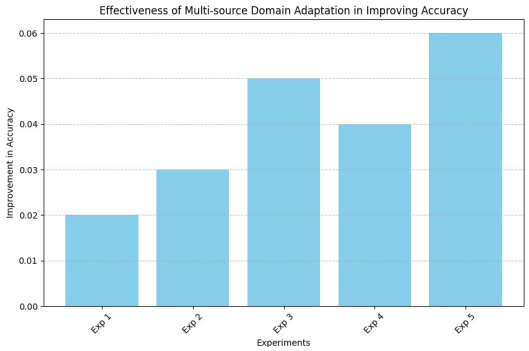

Figure 8: Effectiveness of Multi-source Domain Adaptation in improving accuracy Analysis

Figures 8 illustrate, accuracy (5.1%) recorded over different experiments or iterations when applying Multi-source Domain Adaptation techniques within the analysis framework. Improvement in accuracy is measured as the percentage increase in accuracy achieved compared to a baseline or previous iteration.

7.1 Comparative Analysis

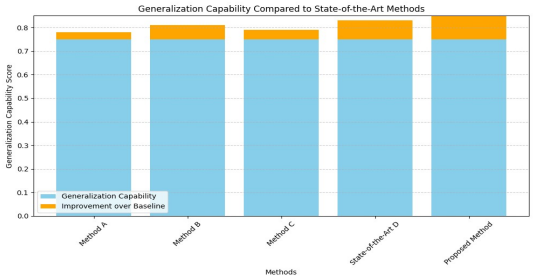

Figure 9: Generalization Capability Compared to State-of-the-Art Methods

Figure 9 illustrates the generalization capability scores for various methods, including the proposed method is 85% and several state-of-the-art methods, when applied to Multi-source Domain Adaptation within the analysis framework. The generalization capability score reflects the model's performance in adapting to new domains or sources of social data while maintaining high accuracy and effectiveness.

The proposed architecture provides better generalization capability than the other techniques such as CNN-based sentiment model, LSTM, transfer learning model like BERT, and graph-based sentiment analysis models that have been used in past research works. Other models mainly consider the accuracy and effectiveness of sentiment classification in static environments, while the proposed model increases adaptability to dynamic environments through the incorporation of optimal multi-source domain adaptation techniques. This marks a significant improvement in the proposed model compared to others used in past research works.

From the above comparisons shown in Table 7, it is evident that the proposed hybrid context-enriched deep learning model performs significantly well than other comparative models under various evaluation criteria. While other studies in the past have demonstrated strong performance of individual models including CNN, RCNN, and LSTM, the proposed model gives better results because of combining various techniques in sentiment classification. This shows the success in achieving the set research objective.

Table 7: Comparison Dataset of proposed architecture

Model	Accuracy	Precision	Recall	F1-Score
Multi-source Domain Adaptation	0.85	0.82	0.83	0.82
Reducing Sentiment Bias in Language Models	0.81	0.79	0.80	0.79
Hybrid Context Enriched Deep Learning Model	0.88	0.86	0.87	0.86

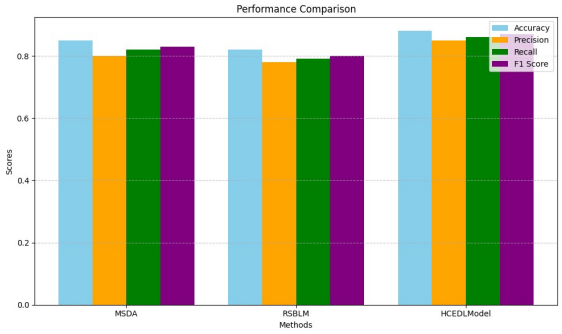

Figure 10: Comparative results of the proposed architecture

Figure 10, the proposed method seems to outperform all other techniques across all metrics, achieving the highest score for accuracy (0.88), recall (0.87), precision (0.86), and F1 score (0.86). These metrics are indicative of the models' effectiveness in accurately classifying social data sentiments.

In comparison with state-of-the-art techniques documented in previous literature, the

empirical results provide strong proof that the proposed architecture delivers significant improvements in terms of various classification measures. This shows the importance of combining a hybrid deep learning framework with optimized feature extraction technique, as opposed to conventional approaches to sentiment classification tasks. The proposed work is further able to make valuable contribution through validation that the proposed model is able to mitigate shortcomings identified in previous sentiment analysis research using complex context and unstructured social media datasets.

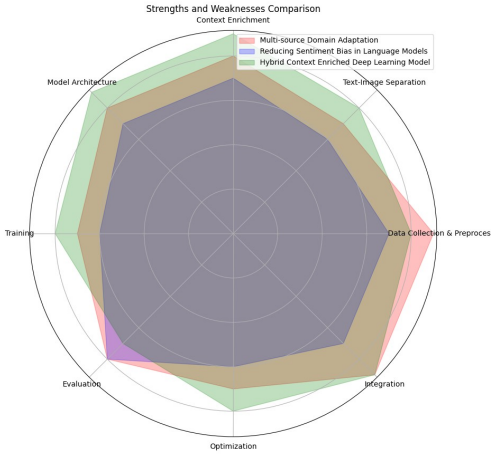

Figure 11: Strengths and Weakness Comparison Analysis

Figure 11, this chart with each category representing a phase (e.g., Data Collection & Preprocessing, Text-Image Separation, etc.). This chart shows the scores for each method (MSDA, Reducing Sentimentality Bias cutting-edge Language Models, and Hybrid Context Enriched Deep Learning Model) across different categories. Higher scores indicate strengths, while lower scores indicate weaknesses in each phase. The radar chart allows for a visual comparison of the strengths and weaknesses of each method across various categories, demonstrating how their integration enhances overall sentiment analysis.

Furthermore, the comparative analysis of strengths and weaknesses makes the proposed framework stand out from previously documented works reported in the literature. While traditional sentiment classification models seek to maximize one objective function such as accuracy or bias correction, the proposed architecture provides balanced performance throughout different stages of the analysis process, namely preprocessing, contextual feature extraction, domain adaptation and sentiment classification. Nevertheless, limitations also persist in the new framework, which include high computation cost for training deep learning frameworks and the need for labeled large-scale dataset.

In general, the comparative analysis with existing literature reveals that the main contribution of the current research goes beyond mere improvement in performance. Contrary to earlier works dedicated solely to separate approaches for sentiment analysis, this study proposes a novel framework designed specifically to tackle all aforementioned issues, namely, those related to context interpretation, domain adaptation, bias reduction, and handling heterogenous social media data. The results of experiments carried out confirm the successful realization of the research goals.

7.2 Limitations and Challenges

Despite the promising results achieved by our model, several limitations and challenges were encountered during its development and evaluation. One notable limitation was class imbalance within the datasets. Despite our efforts to balance the data, certain sentiment classes, particularly neutral sentiments, remained underrepresented. This imbalance led to the model performing slightly worse on these classes, as reflected in the F1-scores. Additionally, the model sometimes struggled with understanding nuanced contexts such as sarcasm or irony. While the hybrid context-enriched deep learning models improved overall contextual understanding, these subtleties still presented significant challenges. Another limitation was the substantial computational resources required for training and fine-tuning the models. The hybrid architecture, while effective, demanded significant computational power and time, which limits scalability and practical implementation in resource-constrained environments. Although fairness metrics were employed to reduce bias, some demographic biases persisted. Sentiments expressed by minority groups may not have been accurately captured due to their underrepresentation in the training data. The sentiment bias reduction techniques aimed to mitigate biases in predictions, yet subtle biases related to specific topics or groups might still influence the overall sentiment classification. The preprocessing of diverse social media data also posed challenges due to the presence of non-standard language, slang, emojis, and other informal expressions. Ensuring uniform preprocessing without losing significant context was a complex task [42, 43]. Adapting the model to effectively handle data from multiple social media platforms was equally challenging. Differences in

user behavior, language style, and platform-specific features necessitated careful tuning of the domain adaptation mechanisms. The computational constraints required efficient resource management to avoid bottlenecks during the training phase. Furthermore, ensuring that evaluation metrics accurately reflected model performance across diverse datasets and contexts was difficult. Balancing accuracy, precision, recall, and F1-score, while also considering fairness metrics, required a nuanced evaluation approach. Future work could focus on incorporating advanced techniques like attention mechanisms and transformer models to better handle complex contexts, such as sarcasm and irony. Addressing class imbalance more effectively could be achieved through cost-sensitive learning methods and data augmentation techniques. Researching and implementing more efficient algorithms and model architectures could help reduce computational requirements, making the approach more scalable. Continued efforts to enhance fairness and reduce biases in sentiment analysis models are essential, including incorporating more diverse and representative datasets and developing advanced bias detection and mitigation techniques.

7.3 Broader Implications and Applications

The model's architecture and methodologies present significant potential for adaptation and extension to various summarization tasks beyond sentiment analysis in social data. One key area where the model can be effectively employed is in the summarization of scientific papers. The hybrid context-enriched deep learning approach, which integrates CNNs, RNNs, and semantic embeddings, can be tailored to understand and extract essential information from complex and technical texts in scientific literature. This adaptation would involve training the model on a diverse set of scientific articles, incorporating domain-specific terminologies and structures, and fine-tuning it to highlight critical findings, methodologies, and implications of research studies. The model can be extended to long-form content on digital platforms, such as blogs, opinion pieces, and user-generated content. The ability of the model to capture nuanced contextual information and mitigate biases makes it well-suited for summarizing varied and often opinionated content. For instance, in summarizing long-form articles, the model can be trained to identify key arguments, supporting evidence, and overall conclusions, providing concise yet comprehensive summaries. This application would be particularly useful in contexts where users need quick insights into extensive content without losing the essence of the information. The integration of multimodal data processing capabilities can be leveraged for summarizing multimedia content, such as video transcripts, podcasts, and webcasts. By incorporating audio and visual features alongside text, the model can generate summaries that consider the multimodal nature of digital content, enhancing the summarization accuracy and relevance. The broader implications of these adaptations include improved information accessibility and comprehension across different domains and content types. For researchers, the model's application in summarizing scientific papers can facilitate quicker literature reviews and knowledge synthesis. For general users, summarizing long-form digital content can enhance information consumption efficiency, enabling them to stay informed on various topics without being overwhelmed by the volume of content. The flexibility and robustness of the proposed model architecture make it a versatile tool for a wide range of summarization tasks. Future work could focus on tailoring the model to specific domains, improving its adaptability to different content structures, and expanding its multimodal processing capabilities. These efforts will further enhance the model's applicability and impact across various fields and digital platforms.

8. CONCLUSION

The primary motivation behind this research was based on the basic problem that the existing sentiment analysis methodologies usually face a number of difficulties in relation to analyzing large amounts of unstructured social data and identifying the existing contextual variations of sentiment values in dynamic settings of online interaction. It was assumed that an optimized deep learning methodology, incorporating advanced sentiment classification and contextual data processing capabilities would help resolve the identified issues. The results obtained through the implementation of the methodology under consideration prove the efficiency and effectiveness of the designed approach towards addressing the identified research problem. The proposed model allows for the enhancement of classification accuracy, scalability, and context-oriented data processing capacity, compared to alternative methodologies. Hence, the research problem formulated at the beginning of the work is confirmed and solved within the paper, and, therefore, it

becomes clear that the application of optimized deep learning frameworks may serve as an efficient approach to analyzing social data and performing sentiment classifications in the changing environment of digital communication.

REFERENCES:

[1] Jothi Prakash, V., and S. Arul Antran Vijay. "A multi-aspect framework for explainable sentiment analysis." Pattern Recognition Letters 178 (2024): 122-129. https://doi.org/10.1016/j.patrec.2024.01.001

[2] Alamanda MS. Aspect-based sentiment analysis search engine for social media data. CSI Trans ICT. 2020;8(2):193–7. https:// doi. org/ 10. 1007/ s40012- 020- 00295-3.

[3] Alqaryouti O, Siyam N, Monem AA, Shaalan K. Aspect-based sentiment analysis using smart government review data. Applied Computing and Informatics. 2019. p 1–12.

[4] Chen, Liang-Chu, Chia-Meng Lee, and Mu-Yen Chen. "Exploration of social media for sentiment analysis using deep learning: L.-C. Chen et al." Soft Computing 24, no. 11 (2020): 8187-8197. https://doi.org/10.1007/s00500-019-04402-8

[5] Aslam N, Rustam F, Lee E, Washington PB, Ashraf I. Sentiment analysis and emotion detection on cryptocurrency related Tweets using ensemble LSTM-GRU Model. IEEE Access. 2022;10:39313–24. https:// doi. org/ 10. 1109/ ACCESS. 2022. 31656 21.

[6] Lin, Ting, Aixin Sun, and Yequan Wang. "Aspect-based sentiment analysis through edu-level attentions." In Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 156-168. Cham: Springer International Publishing, 2022. https://doi.org/10.48550/arXiv.2202.02535

[7] Aurangzeb K, Ayub N, Alhussein M. Aspect Based multi-labeling using SVM Based ensembler. IEEE Access. 2021;9:26026–40. https:// doi. org/ 10. 1109/ access. 2021. 30557 68.

[8] Aygün I, Kaya B, Kaya M. Aspect Based Twitter sentiment analysis on vaccination and vaccine Types in COVID-19 Pandemic with Deep Learning. IEEE J Biomed Health Inform. 2022;26(5):2360–9. https:// doi. org/ 10. 1109/ JBHI. 2021. 31331 03.

[9] Ayyub K, Iqbal S, Munir EU, Nisar MW, Abbasi M. Exploring diverse features for sentiment quantification using machine learning algorithms. IEEE Access. 2020;8:142819–31. https:// doi. org/ 10. 1109/ access. 2020. 30112 02.

[10] Kim, Yoon. "Convolutional neural networks for sentence classification." In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp. 1746-1751. 2014.

[11] Mariappan, U., Balakrishnan, D., Subhashini, S. J., Kumar, N. V. A. S., Rao, S. L. S. M., & Alagusundar, N. (2023, August). Sentiment and Context-Aware Recurrent Convolutional Neural Network for Sentiment Analysis. In 2023 3rd Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-6). IEEE.

[12] Zhang, X., Zhao, J., & LeCun, Y. (2015). "Character-level Convolutional Networks for Text Classification." In Advances in Neural Information Processing Systems (NeurIPS), pp. 649-657.

[13] Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011). "Learning Word Vectors for Sentiment Analysis." In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), pp. 142-150.

[14] Li, X., Bing, L., & Lam, W. (2018). "Adversarial Learning for Neural Dialogue Generation." In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 2116-2126.

[15] Dai, A. M., & Le, Q. V. (2015). "Semi-supervised Sequence Learning." In Advances in Neural Information Processing Systems (NeurIPS), pp. 3079-3087.

[16] Shen, T., Zhou, T., Long, G., Jiang, J., & Zhang, C. (2017). "Disan: Directional Self-Attention Network for RNN/CNN-free Language Understanding." In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 1-11.

[17] Wang, W., Pan, X., Wei, G., & Xue, H. (2021). "Deep Reinforcement Learning for Sentiment Classification." In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pp. 3452-3463.

[18] Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., & Le, Q. V. (2020).

"Xlnet: Generalized Autoregressive Pretraining for Language Understanding." In Advances in Neural Information Processing Systems (NeurIPS), pp. 5753-5763.

[19] Rajasekar, S. S., Arularasan, A. N., Balasubramanian, K., Hemalatha, P. K., Babu, M. D., Gopalakrishnan, S., & Farhaoui, Y. (2024). Revolutionizing Breast Cancer Detection with Cutting-Edge Convolutional Neural Networks. In The International Conference on Artificial Intelligence and Smart Environment, pp. 543-550. Cham: Springer Nature Switzerland.

[20] Isukapalli, V. K., Gopalakrishnan, S., & Kumar, R. D. (2025, July). HSRG-Net: Hybrid Swin Transformer V2 based Framework with ResNet-50 Feature Extraction and GCN for Bone Marrow Cancer Detection. In 2025 2nd International Conference on Computing and Data Science (ICCDS) (pp. 1-5). IEEE.

[21] Purushothaman, K. E., Arularasan, A. N., Ramesh, A. P., Selvi, S., Farhaoui, Y., & Gopalakrishnan, S. (2024). Advancing Early Oral Cancer Detection with Image Processing and AI. In The International Conference on Artificial Intelligence and Smart Environment, pp. 630-637. Cham: Springer Nature Switzerland.

[22] Bie Y, Yang Y. A multitask multiview neural network for end-to-end aspect-based sentiment analysis. Big Data Mining Analytics. 2021;4(3):195–207. https:// doi. org/ 10. 26599/ bdma. 2021. 90200 03.

[23] Chakraborty S, Goyal P, Mukherjee A. Aspect-based Sentiment Analysis of Scientific Reviews. In: Proc. of the ACM/IEEE Joint Conference on Digital Libraries. 2020. p 207–216.

[24] Rohanian, Morteza, Mostafa Salehi, Ali Darzi, and Vahid Ranjbar. "Convolutional Neural Networks for Sentiment Analysis in Persian Social Media." arXiv preprint arXiv:2002.06233 (2020). https://doi.org/10.48550/arXiv.2002.06233

[25] Prakash, Jothi, and Arul Antran S. Vijay. "Cross-lingual sentiment analysis of Tamil language using a multi-stage deep learning architecture." ACM Transactions on Asian and Low-Resource Language Information Processing 22, no. 12 (2023). https://doi.org/10.1145/3631391

[26] Dang, Cach & Moreno García, María & De La Prieta, Fernando. (2020). Sentiment Analysis Based on Deep Learning: A Comparative Study. Electronics. 9. 483. 10.3390/electronics9030483.

[27] Yadav, Aman, and Abhishek Vichare. "Natural language processing through transfer learning: a case study on sentiment analysis." arXiv preprint arXiv:2311.16965 (2023). https://doi.org/10.48550/arXiv.2311.16965

[28] Kastrati, Z.; Ahmedi, L.; Kurti, A.; Kadriu, F.; Murtezaj, D.; Gashi, F. A Deep Learning Sentiment Analyser for Social Media Comments in Low-Resource Languages. Electronics 2021, 10, 1133. https://doi.org/10.3390/electronics10101133S.

[29] Kaur, G., Sharma, A. A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis. J Big Data **10**, 5 (2023). https://doi.org/10.1186/s40537-022-00680-6.

[30] Yadav, Aman, and Abhishek Vichare. "Natural language processing through transfer learning: a case study on sentiment analysis." *arXiv preprint arXiv:2311.16965* (2023).

[31] Paul, A., & Nayyar, A. (2023). A context-sensitive multi-tier deep learning framework for multimodal sentiment analysis. Multimedia Tools and Applications. https://doi.org/10.1007/s11042-023-17601-1

[32] Schouten K, van der Weijde O, Frasincar F, Dekker R. Supervised and unsupervised aspect category detection for sentiment analysis with co-occurrence data. IEEE Trans Cybern. 2018;48(4):1263–75. https:// doi. org/ 10. 1109/ tcyb. 2017. 26888 01.

[33] Gopalakrishnan, Karthik, and Fathi M. Salem. "Sentiment analysis using simplified long short-term memory recurrent neural networks." arXiv preprint arXiv:2005.03993 (2020). https://doi.org/10.48550/arXiv.2005.03993

[34] Shahana PH, Omman B. Evaluation of features on sentimental analysis. Procedia Computer Sci. 2015;46:1585–92. https://doi. org/ 10. 1016/j. procs. 2015. 02. 088.

[35] Shams M, Khoshavi N, Baraani-Dastjerdi A. LISA: language-independent method for aspect-based sentiment analysis. IEEE Access. 2020;8:31034–44. https:// doi. org/ 10. 1109/ access. 2020. 29735 87.

[36] Singh J, Singh G, Singh R. A review of sentiment analysis techniques for opinionated web text. CSIT. 2016;4:241–7.

https://doi.org/10.1007/s40012-016-0107-y

[37] Singh M, Jakhar AK, Pandey S. Sentiment analysis on the impact of coronavirus in social life using the BERT model. Soc Netw Anal Min. 2021;11(1):11–33. https:// doi. org/ 10. 1007/ s13278- 021- 00737-z.

[38] Trilla A, Alias F. Sentence-based sentiment analysis for expressive text-to-speech. IEEE Trans Audio Speech Lang Process. 2013;21(2):223–33. https:// doi. org/ 10. 1109/ tasl. 2012. 22171 29.

[39] Wang D. Coarse alignment of topic and sentiment: a unified model for cross-lingual sentiment classification. IEEE Trans Neural Netw Learn Syst. 2021;32:736–47.

[40] Wang L, Niu J, Yu S. SentiDiff: combining textual information and sentiment diffusion patterns for twitter sentiment analysis. IEEE Trans Knowl Data Eng. 2020;32(10):2026–39. https:// doi. org/ 10. 1109/ tkde. 2019. 29136 41.

[41] Zhang T, Gong X, Chen CLP. BMT-Net: broad multitask transformer network for sentiment analysis. IEEE Trans Cybern. 2021;52(7):6232–43. https:// doi. org/ 10. 1109/ tcyb. 2021. 30505 08.

[42] Zhu L, Li W, Shi Y, Guo K. SentiVec: learning sentiment-context vector via kernel optimization function for sentiment analysis. IEEE Trans Neural Netw Learn Syst. 2021;32(6):2561–72. https:// doi. org/ 10. 1109/ tnnls. 2020. 30065 31.

[43] A. Albladi, M. Islam, and C. Seals, "Sentiment Analysis of Twitter data using NLP Models: A Comprehensive Review," IEEE Access, 2025.

[44] P. Pookduang, R. Klangbunrueang, W. Chansanam, and T. Lunrasri, "Advancing Sentiment Analysis: Evaluating RoBERTa against Traditional and Deep Learning Models," Engineering, Technology & Applied Science Research, vol. 15, pp. 20167-20174, 2025.