Deep Learning Approach in Complex Sentiment Analysis: A Case Study on Social Problems

Bambang Nurdewanto[1], Kukuh Yudhistiro[*2,] , Dani Yuniawan[3], Himawan Pramaditya[4], Mochammad Daffa Putra Karyudi[5], Yulia Natasya Farah Diba Arifin[6] , Puput Dani Prasetyo Adi[7]
Faculty of Information Technology, Universitas Merdeka Malang, Malang City, East Java, Indonesia[1, 2, 3, 4, 5, 6]
National Research and Innovation Agency (BRIN), Bandung Indonesia[7]

***Abstract*—This scholarly investigation examines the utilization of artificial intelligence (AI) technology in the analysis and resolution of intricate societal challenges in many countries. The originality of this study resides in the employment of deep learning algorithms, particularly Convolutional Neural Network (CNN), to execute sentiment analysis with an elevated degree of complexity. The examination encompasses three principal dimensions of sentiment: Sentiment, Tone, and Object, with the intention of offering profound insights into public perceptions regarding various social challenges. The fundamental sentiment is categorized into three classifications: Positive, Neutral, and Negative. Moreover, the Tone analysis introduces an additional layer of comprehension that encompasses Support, Suggestion, Criticism, Complaint, and Others, thereby delineating a more precise communicative context. The Object dimension is employed to ascertain the target of the sentiment, whether it pertains to an Individual, Organization, Policy, or other entity. This inquiry applied the analysis to several clusters of social issues, including Poverty and Economic Disparity, Health and Wellbeing, Education and Literacy, Violence and Security, as well as Environment and Social Life. The findings are anticipated to aid the government in devising policies that are more effective and responsive to the exigencies of society, through an enhanced understanding of public sentiment.**

Keywords*—*Component; sentiment analysis; deep learning; artificial intelligence; social case

I. Introduction

Sentiment analysis plays a critical role in understanding and quantifying the emotions or sentiments expressed in textual data, such as reviews, comments, and social media posts. It is an essential tool for businesses and organizations aiming to comprehend public opinion, improve customer service, and make data-driven decisions in product development and marketing strategies [10]. However, traditional sentiment analysis methods struggle with capturing the complexities of language, including sarcasm, ambiguous expressions, and nuanced opinions [1]. This limitation has led to a growing need for more sophisticated techniques to accurately analyze sentiment, especially when dealing with large volumes of unstructured data from sources like social media [7].

To address these challenges, recent research has increasingly employed deep learning models, which offer significant improvements over traditional methods in sentiment classification. Neural networks, particularly deep learning architectures such as capsule-based RNNs and Graph Convolutional Networks (GCNs), have shown remarkable success in handling complex textual data [6] [14]. These models not only enhance sentiment classification but also allow for the integration of multimodal data—such as text, audio, and visual cues—to capture deeper insights, especially in the case of complex emotions like sarcasm [8] [12]. Moreover, transfer learning techniques have been implemented to overcome data scarcity challenges, enabling models to be fine-tuned for smaller, domain-specific datasets [11] [16].

This study contributes to the field by employing Convolutional Neural Networks (CNNs), a deep learning methodology traditionally associated with image processing, for sentiment analysis [9]. CNNs have demonstrated effectiveness in analyzing text by treating it as spatial data, extracting local features, and offering a more straightforward yet efficient approach compared to other models [18]. By leveraging CNNs, this research aims to improve sentiment analysis performance, particularly in understanding nuanced social issues [20], and to provide more precise and actionable insights [21] [22].

In line with the abstract, this research focuses on three principal dimensions of sentiment analysis: Sentiment, Tone, and Object. Sentiment is categorized into three primary classifications—Positive, Neutral, and Negative—while Tone introduces additional layers of interpretation, including Support, Suggestion, Criticism, Complaint, and Other. The Object dimension further refines the analysis by identifying the target of the sentiment, such as an Individual, Organization, or Policy. This comprehensive approach enables a more detailed understanding of public perceptions regarding various social challenges.

Specifically, the research addresses key societal clusters such as Poverty and Economic Disparity, Health and Wellbeing, Education and Literacy, Violence and Security, and Environment and Social Life. The insights generated through this analysis are expected to support government agencies in developing policies that are more responsive to public sentiment and aligned with societal needs [23]. Through the integration of CNNs in sentiment analysis, this study offers a novel approach to addressing the limitations of traditional methods and aims to enhance decision-making processes by providing deeper insights into public opinion on critical social issues.

II. Literature Review

A. Sentiment Analysis

Sentiment analysis plays a crucial role in shaping social justice movements by providing insights into public opinion and emotional responses [24]. By analyzing sentiments expressed on social media platforms, activists can better understand the dynamics of support and opposition [30], which can inform strategies for advocacy and mobilization [27].

Impact on Public Sentiment: Black Lives Matter Movement: A study utilizing a BERT model analyzed over one million tweets related to the BLM movement, revealing themes of social justice and police brutality [29]. Positive sentiments were linked to significant events, guiding activists in their messaging and outreach efforts [13]. Sentiment assessment, akin to employing BERT on Twitter datasets pertaining to Black Lives Matter, elucidates public sentiment regarding social justice initiatives, accentuating motifs and occurrences linked with affirmative or adverse sentiment [25]. The BERT model, meticulously adjusted for Twitter sentiment, surpassed alternative models in performance. Retweet and lexical count frequencies underscored significant BLM motifs/occurrences.

EndSARS Movement: The sentiment analysis of the #EndSARS protests in Nigeria highlighted eight basic emotions, providing insights into public sentiment that could inform governmental responses and future activism strategies [14]. Sentiment analysis, as demonstrated in the research concerning the EndSARS demonstration in Nigeria, facilitates the comprehension of collective sentiments, directing governmental reactions, and potentially mitigating social turbulence in forthcoming movements [26]. Recognized eight fundamental emotions articulated during the EndSARS demonstration. Suggested methodologies for the government to tackle subsequent activist movements.

Emotional Dynamics: Chinese Online Movements: Research indicates that emotions like anger and anxiety are pivotal in mobilizing online social movements [28]. Understanding these sentiments can help predict movement trajectories and prevent potential social unrest [15]. Sentiment evaluation in digital social movements, particularly concentrating on outrage and apprehension, facilitates comprehension of sentiment progression to avert societal calamities, augmenting the efficacy of social justice campaigns. The suggested BERT-based framework surpasses foundational models. A sentiment evaluation dataset for Chinese digital social movements has been established.

Broader Implications: Social Justice and Well-being: The correlation between feelings of social justice and subjective well-being suggests that sentiment analysis can also reflect broader societal perceptions, influencing public policy and community engagement [19]. The investigation examines the robust association between perceptions of social equity and subjective welfare, proposing that sentiments of justice within a community affect personal life contentment and well-being [31]. Welfare is impacted by financial resources, social connections, and perceptions of equity. Significant association between life contentment and societal justice interpretations.

While sentiment analysis offers valuable insights, it is essential to consider the potential for misinterpretation or manipulation of data, which can skew public perception and impact the effectiveness of social justice movements.

Previous CNN research related to sentiment analysis for social problems, Sentiment analysis using deep learning, particularly Convolutional Neural Networks (CNNs), has gained traction in addressing social issues through the analysis of public sentiment on platforms like Twitter and Reddit. This approach leverages advanced methodologies to extract meaningful insights from vast datasets, enhancing understanding of societal concerns [32].

Deep Learning Techniques in Sentiment Analysis: CNNs and RNNs: Recent studies highlight the effectiveness of CNNs and hybrid models like CRDC (Capsule with Deep CNN and Bi-structured RNN), achieving accuracy rates up to 98.02% in sentiment classification tasks [6] The manuscript presents an innovative hybrid methodology, CRDC, amalgamating Capsule, Deep CNN, and Bi-RNN for sentiment analysis, surpassing conventional CNN or RNN frameworks with elevated precision across diverse datasets. Although the Capsule-based RNN methodology exceeds the performance of CNN or RNN models, the CRDC model demonstrates exceptional efficacy in sentiment analysis.

LSTM Models: While CNNs are prominent, LSTM models have also shown significant results, with an 88.7% accuracy in analyzing sentiments regarding animal testing (Ismail, 2024). This document explores the realm of affective assessment utilizing Long Short-Term Memory (LSTM) networks rather than relying on Convolutional Neural Networks (CNN). It does not tackle the application of CNN in social contexts. A preponderance of tweets articulated adverse sentiment regarding animal experimentation. The LSTM architecture attained an accuracy of 88.7% in sentiment categorization.

Applications in Social Contexts: Crisis Response and Governance: AI/ML tools, including sentiment analysis, are utilized to gauge public opinion on governance issues, as demonstrated in a study analyzing 114,390 Reddit comments, achieving an accuracy of 89% with the SiEBERT model [8]. The manuscript centers on affective evaluation employing SiEBERT, a pre-trained transformer architecture, attaining 89% precision, signifying the efficacy of deep learning artificial intelligence in comprehending the apprehensions of citizens via social media sentiment assessment. BERTopic with K-means attained the utmost coherence rating.

Consumer Insights: Automated sentiment analysis models, such as ASASM-HHODL, have been developed to process social media data effectively, enhancing brand monitoring and customer feedback analysis [5]. The manuscript advocates for an Automated Sentiment Evaluation in Social Media (ASASM) employing Harris Hawks Optimization and Deep Learning (HHODL) methodologies, concentrating on Attention-oriented BiLSTM for sentiment categorization, without explicitly referencing CNN. The ASASM-HHODL framework surpasses contemporary sentiment analysis methodologies. Deep Learning models autonomously derive characteristics from datasets for sentiment evaluation.

While deep learning models, particularly CNNs, show promise in sentiment analysis, challenges remain in data preprocessing and model optimization, necessitating ongoing research to refine these techniques for broader applications.

How do Convolutional Neural Networks impact sentiment analysis?, Convolutional Neural Networks (CNNs) significantly enhance sentiment analysis by effectively capturing local patterns in text data, improving classification accuracy and interpretability. Their ability to process n-grams and semantic features allows for nuanced sentiment detection, as evidenced by various studies [33].

CNNs facilitate understanding of sentiment classification through the analysis of convolutional filters, revealing that certain parts of speech (POS) are more influential in determining sentiment polarity. The preference for shorter n-grams in negative sentiment classification highlights the model's focus on critical linguistic features [4]. CNNs impact sentiment analysis by revealing the relevance of certain parts of speech tags and the preference for shorter n-grams in classifying negative sentiment sequences, indicating the potential for smaller architectures.

Hybrid Models and Performance: Integrating CNNs with other architectures, such as BERT and BiLSTM, enhances the extraction of both local and global semantic features, achieving high accuracy rates (e.g., 92.35%) in sentiment classification tasks [2]. The analysis showcases a cutting-edge hybrid system that integrates two pathways in neural networks, employing an attention mechanism to expertly combine Convolutional Neural Networks (CNNs) and Bidirectional Long Short-Term Memory networks (BiLSTMs), hence significantly boosting sentiment analysis via the proficient acquisition of both local and global semantic elements. The suggested framework attains a sentiment classification accuracy of 92.35%. Moreover, the outlined system reaches an F1 score of 91.59% regarding sentiment classification.

Graph-based approaches, like hierarchical dual graph convolutional networks, further improve aspect-based sentiment analysis by effectively modeling syntactic and semantic dependencies [17]. Hierarchical dual graph convolutional network (HD-GCN) enhances sentiment analysis by extracting syntactic and semantic information, outperforming state-of-the-art models on ABSA tasks, as shown in the study.

While CNNs excel in sentiment analysis, their black-box nature raises concerns about interpretability and understanding of internal representations, which can hinder trust in automated systems [15]. Convolutional Neural Networks (CNNs) play a crucial role in sentiment analysis by extracting features from text data, enhancing sentiment classification accuracy, and addressing interpretability challenges in neural network models.

III. Research Method

A. Sentiment Aspects and the Labels

Sentiment aspect prediction refers to the process of identifying and categorizing various components of sentiment expressed within a text. It involves breaking down the text to determine not just the overall sentiment (positive, neutral, negative) but also other relevant aspects that give more context to the sentiment. Here’s a breakdown of how predictions for each sentiment aspect typically work: Basic Sentiment: [Positive, Neutral, Negative], According to Pang and Lee (2008), basic sentiment analysis focuses on classifying text into these three main categories. This classification is a starting point in many sentiment analysis applications.

- *Positive*: Indicates a supportive or optimistic attitude toward the topic.
- *Neutral*: Indicates a neutral or impartial attitude.
- *Negative*: Indicates an opposing or pessimistic attitude toward the topic.

Furthermore, Tone: [Support, Suggestion, Criticism, Complaint, Others], According to Liu (2012), understanding tone in sentiment analysis provides deeper insights into how an entity or issue is perceived, not only from the polarity perspective but also from the communication context.

- *Support*: Text that shows approval or support for an issue or entity.
- *Suggestion*: Text that provides recommendations or solutions to a specific problem.
- *Criticism*: Text that expresses negative opinions or disagreement.
- *Complaint*: Text that conveys dissatisfaction or problems encountered.

Furthermore, Object: [Individual, Organization, Policy, Others]. According to Balahur et al. (2010), identifying the object in sentiment analysis helps understand the main target of the expressed sentiment, which is crucial in the context of social issues.

- *Individual*: Text referring to a person or a group of individuals.
- *Organization*: Text referring to organizational entities such as companies, institutions, or governments.
- *Policy*: Text referring to specific regulations, laws, or policies.

Furthermore, Cluster: [Poverty and Economic Inequality, Health and Well-being, Education and Literacy, Violence and Security, Environment and Social Life, Others], According to Cambria et al. (2013), clustering text based on topics or issues provides clearer context about the area of concern being analyzed, enabling deeper understanding and more targeted actions.

- *Poverty and Economic Inequality*: Text related to economic issues and social disparities.
- *Health and Well-being*: Text related to health issues, healthcare services, and social well-being.
- *Education and Literacy*: Text related to education systems, access to education, and literacy.
- *Violence and Security*: Text related to issues of violence, crime, and public security.
- *Environment and Social Life*: Text related to environmental issues and social interactions.

B. Data Collection

One essential parameter is Data Collection (see Fig. 1 and Fig. 2), Dataset that has been labeled for basic sentiment aspects

(Positive, Neutral, Negative). The dataset has been labeled for the Tone sentiment aspects (Support, Suggestions, Criticism, Complaints, Others).

```
import pandas as pd

data = {
 'text': [
  'I strongly support this program.',
  'This program should be improved.',
  'I am not satisfied with this service',
  'This service is very disappointing',
  'There is nothing special about this program'
 ],
 'label': ['Support', 'Suggestions', 'Criticism', 'Complaints', 'Other']
}

df = pd.DataFrame(data)
```

Fig. 1. Data Collection code 1.

```
import pandas as pd

data = {
 'text': [
  'I strongly support this program.',
  'This program should be improved.',
  'I am not satisfied with this service',
  'This service is very disappointing',
  'There is nothing special about this program'
 ],
 'label': ['Support', 'Suggestions', 'Criticism', 'Complaints', 'Other']
}

df = pd.DataFrame(data)
```

Fig. 2. Data Collection code 2.

Furthermore, The same method for the Tone aspect and basic sentiment is also applied to the Object sentiment aspects (Individual, Organization, Policy) and Cluster aspect (Poverty and Economic Inequality, Health and Well-being, Education and Literacy, Violence and Security, Environment and Social Life, Others).

C. *Text Preprocessing*

In this Text Preprocessing process there are several things that are done, namely Cleaning and preparing text data to be ready for analysis. And Steps. Steps consist of several parts, namely:

- Tokenization : Breaking down the text into smaller units, such as words or phrases. Example: "Poverty in this area is worsening due to economic inequality" -> ["Poverty", "in", "this", "area", "is", "worsening", "due", "to", "economic", "inequality"]
- Punctuation Removal: Removing unnecessary punctuation marks. Example: ["Poverty", "in", "this", "area", "is", "worsening", "due", "to", "economic", "inequality"]
- Stop Words Removal: Removing common words that do not carry significant meaning, such as "and," "or," "is." Example: ["Poverty", "area", "worsening", "due", "economic", "inequality"]
- Stemming or Lemmatization : Converting words to their base form.Example: ["Poverty", "area", "worsen", "due", "economic","inequality"]

D. *Feature Extraction*

Furthermore, Feature Extraction: Identifying and extracting important features from the text that will be used for classification. Steps: Using Term Frequency-Inverse Document Frequency (TF-IDF) to Calculate the importance of a word in a document compared to the entire corpus. Example: ["Poverty", "area", "worsening", "inequality", "economic"] -> [0.2, 0, 0.4, ..., 0.3, 0.5]

E. *Model Development Using CNN*

Training a machine learning or deep learning model to recognize cluster patterns in the text. Steps:

- Training the model using a Convolutional Neural Networks (CNN) architecture to recognize cluster patterns in the text.
- Training the model using data labeled.
- Model Validation: Evaluating the model using validation data to ensure its performance.

F. *Prediction for Each Sentiment Aspect*

Load the trained model: This step involves loading a pre-trained machine learning or deep learning model that has been specifically trained to analyze sentiment aspects. The model is typically saved in a file after training and can be loaded using various libraries such as TensorFlow or PyTorch. Loading the trained model allows us to utilize its learned parameters and weights for predicting sentiment aspects in new text data.

Make cluster predictions: Once the trained model is loaded, it can be used to make predictions on new, unseen data. This step involves feeding the input text data into the model to predict the specific clusters or categories of sentiment aspects. The output will indicate which sentiment cluster the text belongs to, such as "Poverty and Economic Inequality," "Health and Well-being," or other predefined clusters.

G. *Evaluation and Validation*

Evaluating the performance of the model using metrics such as accuracy, precision, recall, and F1-score. Steps:

- Confusion Matrix : Analyzing the model's prediction errors.
- Cross-validation: Performing cross-validation to ensure the model is not overfitting.

IV. Result and Discussion

A. *Text Preprocessing*

The implementation of the algorithm in the methodology section is implemented using the Python programming language with the necessary libraries such as TensorFlow.Text Preprocessing: For as example, in the following text as data input: 'This program aims to reduce poverty and improve community welfare.' Steps: Tokenization (see Fig. 3), Breaking down the text into smaller units, such as words or phrases. Fig. 4 shows text processing code.

```
from tensorflow.keras.preprocessing.text import Tokenizer
from tensorflow.keras.preprocessing.sequence import pad_sequences
from sklearn.preprocessing import LabelEncoder

# Tokenisasi teks
tokenizer = Tokenizer(num_words=5000)
tokenizer.fit_on_texts(df['text'])
sequences = tokenizer.texts_to_sequences(df['text'])

# Padding sequences
max_length = 50
padded_sequences = pad_sequences(sequences, maxlen=max_length)

# Encode label
label_encoder = LabelEncoder()
labels = label_encoder.fit_transform(df['label'])
```

Fig. 3. Tokenization code.

output: {'program': 1, 'aim': 2, 'reduce': 3, 'poverty': 4, 'improve': 5, 'community': 6, 'welfare': 7}

Punctuation Removal, Removing unnecessary punctuation marks.

Output: ["This", "program", "aims", "to", " reduce", " poverty", " and", " improve", " community", "welfare"]

Stop Words Removal, Removing common words that do not carry significant meaning, such as "and," "or," "is."

Output: ["program", " aims", " reduce", " poverty", " improve", " community", "welfare"]

Stemming or Lemmatization, Converting words to their base form.

Output: ["program", "aim", "reduce", "poverty", "improve", "community","welfare"]

```
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk.stem import WordNetLemmatizer
import string

# Inisialisasi NLTK
nltk.download('punkt')
nltk.download('stopwords')
nltk.download('wordnet')

def preprocess_text(text):
    # Tokenisasi
    tokens = word_tokenize(text)
    # Penghapusan tanda baca
    tokens = [word for word in tokens if word.isalpha()]
    # Penghapusan stop words
    stop_words = set(stopwords.words(english))
    tokens = [word for word in tokens if word.lower() not in stop_words]
    # Lemmatization
    lemmatizer = WordNetLemmatizer()
    tokens = [lemmatizer.lemmatize(word) for word in tokens]
    return tokens

processed_text = preprocess_text(input_text)
print(processed_text)
```

Fig. 4. Text Processing code.

B. *Feature Extraction*

Feature Extraction Using TF-IDF to represent text in a numerical format: Sequence Representation: The processed_text is converted into a sequence based on the word index.

```
output: [[1, 2, 3, 4, 5, 6, 7]]
```

Padded Sequence: The sequence is padded to the specified max_length of 100. Since the original sequence has only 7 tokens, the rest are padded with zeros.

```
output: [[1, 2, 3, 4, 5, 6, 7, 0, 0, 0, ..., 0]]
```

C. *Model Development Using CNN*

For each element of the sentiment aspect, including basic sentiment, tone, object and social problem clusters, the following are the settings for the function *model.add(Dense(…, activation='softmax'))*:

- Basic sentimen (Positive, Neutral, Negative) = *model.add(Dense(3, activation='softmax'))*
- Tone sentimen (Support, Suggestion, Criticism, Complaint, Others) = *model.add(Dense(5, activation='softmax'))*
- Object sentimen (Individual, Organization, Policy) = *model.add(Dense(3, activation='softmax'))*
- Cluster sentimen (Poverty and Economic Inequality, Health and Well-being, Education and Literacy, Violence and Security, Environment and Social Life) = *model.add(Dense(5, activation='softmax'))*

D. *Model Training*

This Model Training have a two analysis, split the data and process the training model by training and validating the data. Split the data into training data and validation data (e.g., 80% for training and 20% for validation). Fig. 5 shows the training test code.

```
from sklearn.model_selection import train_test_split X_train, X_val,
y_train, y_val = train_test_split(X, y, test_size=0.2, random_state=42)
```

Fig. 5. Training test code.

Furthermore, The process of training the model using training and validation data. Fig. 6 shows the validation data code.

```
history = model.fit(X_train, y_train, epochs=10, batch_size=32,
validation_data=(X_val, y_val))
```

Fig. 6. Validation data code.

Moreover, Model Evaluation and Validation Using metrics such as Confusion Matrix, Precision, Recall, and F1-Score (see Fig. 7).

```
from sklearn.metrics import classification_report, confusion_matrix
y_pred = model.predict(X_val)
y_pred_classes = np.argmax(y_pred, axis=1)
y_true = np.argmax(y_val, axis=1)

print(confusion_matrix(y_true, y_pred_classes))
print(classification_report(y_true, y_pred_classes))
```

Fig. 7. Confusion Matrix code.

Furthermore, Result for each sentimen aspect among others Prediction of Object Sentiment Aspect Clusters. The model generates predictions of Object sentiment aspect clusters (Individuals, Organizations, Policies, Others) based on learned patterns. Output: text: This program aims to reduce poverty and improve community welfare. Cluster/sentimen aspect Prediction: Policy.

Evaluation Model:

- Loss and Accuracy : Validation Loss: 0.4321
- Validation Accuracy : 0.8725
- Confusion Matrix :

$$\begin{bmatrix} 120 & 10 & 5 \\ 12 & 110 & 8 \\ 3 & 7 & 115 \end{bmatrix}$$

E. Classification Report Output

In the Classification Report, Table I shows the results of Classification such as Precision, Recall, F1-Score, and Support, from the parameters Individual, Organization, Policy, Accuracy, Macro AVG, and Weighted AVG.

TABLE I. CLASSIFICATION REPORT OUTPUT 1

Parameters	Classification Parameters			
	Precision	***Recall***	***F1-Score***	***Support***
Individuals	0.88	0.88	0.88	135
Organization	0.86	0.82	0.84	130
Policy	0.89	0.91	0.9	125
Accuracy			0.87	390
Macro AVG	0.87	0.87	0.87	390
Weighted AVG	0.87	0.87	0.87	390

Explanation of Table I:

1) *Precision:*

a) For individuals: 88% of predicted "Individual" cases were correct.

b) For organization: 86% of predicted "Organization" cases were correct.

c) For policy: 89% of predicted "Policy" cases were correct.

2) *Recall:*

a) For individuals: The model correctly identified 88% of all true "Individual" cases.

b) For organization: The model correctly identified 82% of all true "Organization" cases.

c) For policy: The model correctly identified 91% of all true "Policy" cases.

3) *F1-Score:*

a) For individuals: The F1-score is 0.88, indicating balanced precision and recall.

b) For organization: The F1-score is 0.84, slightly lower due to the recall being lower than precision.

c) For policy: The F1-score is 0.90, indicating strong performance in both precision and recall.

4) *Support:* The number of true instances for each category in the dataset. The model evaluated 135 instances for Individuals, 130 for Organization, and 125 for Policy.

- Overall Performance:

a) Accuracy: The overall accuracy of the model is 87%, meaning that 87% of all classifications were correct.

b) Macro AVG (Macro Average): The unweighted mean of precision, recall, and F1-score across all categories, indicating how the model performs equally across categories. In this case, the scores for precision, recall, and F1-score are all 0.87.

c) Weighted AVG (Weighted Average): The average of precision, recall, and F1-score, weighted by the number of instances (support) in each class. The weighted averages are also 0.87, meaning the model maintains similar performance when taking into account the distribution of instances.

- Summary:

The classification report shows that the model performs well across all categories, with an overall accuracy of 87%. The Policy category has the best performance with a higher F1-score, while the Organization category has a slightly lower recall.

F. Prediction of Tone Sentiment Aspect

The model generates predictions of Tone sentiment aspect clusters (Support, Suggestion, Criticism, Complaint) based on learned patterns

Output: `Text: This program aims to reduce poverty and improve community welfare.`

Predicted Tone: Support

Evaluation Model

Loss and Accuracy:

```
Validation Loss: 0.7241
Validation Accuracy: 0.85245
```

Confusion Matrix:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Moreover, the Classification Report Output can be seen in Table II.

TABLE II. CLASSIFICATION REPORT OUTPUT 2

Parameters	Classification Parameters			
	Precision	***Recall***	***F1-Score***	***Support***
Support	1.00	1.00	1.00	1
Advice	1.00	1.00	1.00	1
Critique	1.00	1.00	1.00	1
Complaint	1.00	1.00	1.00	1
More	1.00	1.00	1.00	1

Parameters	Classification Parameters			
	Precision	***Recall***	***F1-Score***	***Support***
Accuracy	1.00	1.00	1.00	5
Macro AVG	1.00	1.00	1.00	5
Weighted AVG	1.00	1.00	1.00	5

Explanation of Table II:

1) Precision: For Support, Advice, Critique, Complaint, and More, the precision is 1.00, indicating that all predictions made for each class were correct.

2) Recall: For all categories, the recall is 1.00, meaning the model identified all true instances for each category perfectly.

3) F1-Score: For all categories, the F1-score is 1.00, showing that the model performed perfectly across all classes.

4) Support: Each category has a support of 1, meaning the model was tested with only one instance per class.

- Overall Performance:

a) Accuracy: The model achieved perfect accuracy, as all predictions made were correct.

b) Macro AVG (Macro Average): The unweighted average of precision, recall, and F1-score across all classes. Since every class has perfect performance, the macro averages are all 1.00.

c) Weighted AVG (Weighted Average): Similar to the macro average but weighted by the number of instances in each class. Given that each class has only one instance, the weighted average is also 1.00.

- Summary:

This classification report shows perfect model performance across all categories (Support, Advice, Critique, Complaint, and More) with an accuracy, precision, recall, and F1-score of 1.00. However, the very small dataset (only one instance per class) suggests that this result may not generalize well, as testing with more instances would provide a more realistic assessment of model performance.

G. Prediction of Basic Sentiment Aspect

The model generates predictions of basic sentiment aspect (Positive, Neutral, Negative) based on learned patterns (Fig. 8).

```
from tensorflow.keras.models import load_model

# Memuat model yang sudah dilatih
model = load_model('model_object_cluster_cnn.h5')

prediction = model.predict(padded_sequences)
cluster = np.argmax(prediction, axis=1)
cluster_labels = ['Positive,Neutral, Negative']

predicted_cluster = cluster_labels[cluster[0]]
```

Fig. 8. Prediction of basic sentiment aspect code.

```
Output:
Text: This program aims to reduce poverty and
improve community welfare.
Predicted Tone: Positive
```

Evaluation Model:

- Loss and Accuracy:

```
Validation Loss: 0.7321
Validation Accuracy: 0.88245
```

- Confusion Matrix:

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Classification Report Output

TABLE III. CLASSIFICATION REPORT OUTPUT 3

Parameters	Classification Parameters			
	Precision	***Recall***	***F1-Score***	***Support***
Negative	1.00	1.00	1.00	2
Neutral	1.00	1.00	1.00	2
Positive	1.00	1.00	1.00	1
Accuracy			1.00	4
Macro AVG	1.00	1.00	1.00	4
Weighted AVG	1.00	1.00	1.00	4

Explanation of Table III:

1) Precision: For Negative, Neutral, and Positive, the precision is 1.00, meaning that all the predictions for each of these sentiment categories were accurate.

2) Recall: For all three categories, the recall is 1.00, showing that the model identified all the true instances of Negative, Neutral, and Positive sentiments without missing any.

3) F1-Score: Since both precision and recall are perfect (1.00) for each class, the F1-score is also 1.00 for all sentiment categories.

4) Support: There were two instances for both Negative and Neutral, and 1 instance for Positive.

- Overall Performance:

a) Accuracy: The overall accuracy is 1.00, meaning the model made correct predictions for all 4 instances in the test dataset.

b) Macro AVG (Macro Average): This is the unweighted average of precision, recall, and F1-score across all sentiment classes. Since every class has perfect scores, the macro average is 1.00.

c) Weighted AVG (Weighted Average): This is the weighted average of precision, recall, and F1-score, taking into account the number of instances in each class. As all classes have a perfect performance, the weighted average is also 1.00.

- Summary:

This classification report reflects perfect model performance for all three sentiment classes (Negative, Neutral, and Positive).

The model achieved an accuracy, precision, recall, and F1-score of 1.00 across the board.

V. Conclusion

The Confusion Matrix shows that the model has good performance, with an overall accuracy of around 88.46%. The precision, recall, and F1-score for each class are also quite high, indicating that the model can effectively classify text into the categories of Individual, Organization, or Policy. These results demonstrate that the trained CNN model is quite effective in recognizing sentiment aspect patterns in the text. Additionally, the validation loss and accuracy metrics further support the effectiveness of the model. For the implementation of the CNN model on basic sentiment analysis, the validation loss is 0.7321, with an accuracy of 0.8825. Meanwhile, for the implementation on Tone Sentiment Aspect, the validation loss is 0.7241, and the accuracy is 0.8524. These findings emphasize that building a CNN model for sentiment prediction from the perspectives of basic sentiment, object, tone, and social cluster yields good results and proves to be effective in enhancing the sentiment analysis process. This capability makes the CNN model a valuable tool in accurately identifying and classifying various sentiment aspects, thus supporting comprehensive sentiment analysis across different contexts. The findings are expected to assist the government in formulating policies that are more effective and responsive to society's needs, by providing a deeper understanding of public sentiment. This enhanced capability for accurately identifying and classifying various sentiment aspects enables policymakers to better gauge the opinions and concerns of the public, thereby allowing for more targeted and informed decision-making.

As additional data that all applications or Implementation on web scraper applications can be accessed on the following web pages on Fig. 9:

A. *Testing 1*

News URL:

https://edition.cnn.com/2024/08/28/climate/namibia-kill-elephants-meat-drought/index.html

Fig. 9. URL input view for web scraping.

Testing Evaluation

a) Data sets result: This section presents the results of applying the text processing algorithm in Fig. 3 and Fig. 4, specifically for Punctuation Removal, which involves removing unnecessary punctuation marks. The Fig. 10 represented the dataset output is in JSON format and is displayed on the web for testing 1.

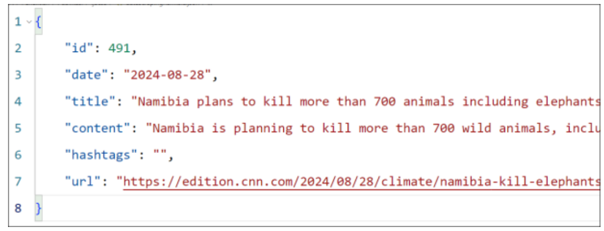
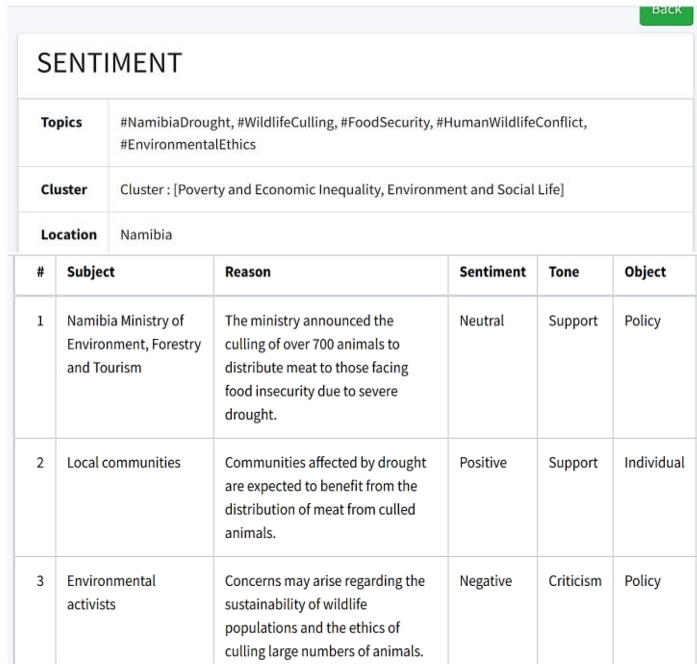


Fig. 10. The dataset output is in JSON format for testing 1.

b) Outcomes analysis result:



SENTIMENT

Topics	#NamibiaDrought, #WildlifeCulling, #FoodSecurity, #HumanWildlifeConflict, #EnvironmentalEthics
Cluster	Cluster : [Poverty and Economic Inequality, Environment and Social Life]
Location	Namibia

#	Subject	Reason	Sentiment	Tone	Object
1	Namibia Ministry of Environment, Forestry and Tourism	The ministry announced the culling of over 700 animals to distribute meat to those facing food insecurity due to severe drought.	Neutral	Support	Policy
2	Local communities	Communities affected by drought are expected to benefit from the distribution of meat from culled animals.	Positive	Support	Individual
3	Environmental activists	Concerns may arise regarding the sustainability of wildlife populations and the ethics of culling large numbers of animals.	Negative	Criticism	Policy

Fig. 11. Sentiment analysis results for testing 1.

The sentiment analysis and clustering of topics which is shown on Fig. 11 are accurately aligned with the context of the news. The issues surrounding #NamibiaDrought, #WildlifeCulling, #FoodSecurity, #HumanWildlifeConflict, and #EnvironmentalEthics are appropriately categorized under the clusters Poverty and Economic Inequality and Environment and Social Life, reflecting both the social and environmental dimensions of the situation in Namibia.

The sentiment analysis also correctly reflects the perspectives of different stakeholders:

1. Namibia Ministry of Environment, Forestry, and Tourism is assessed with a Neutral sentiment and Support tone, as the policy of culling animals to address food insecurity is seen as a practical measure.

2. Local communities impacted by drought are viewed positively, with a Support tone, as they benefit from the distribution of meat, an individual-level impact.

3. Environmental activists, expressing concern over the ethical and sustainability aspects of the culling policy, are assigned a Negative sentiment with a Criticism tone, correctly representing opposition to the policy.

This analysis appropriately captures the nuanced reactions to the situation in Namibia.

B. Testing 2

News URL

https://www.bbc.com/news/articles/c97w1x2deyvo

1) Data sets result: The Fig. 12 represented the dataset output is in JSON format and is displayed on the web for testing 2.


```
{
    "id": 493,
    "date": "2024-09-01",
    "title": "Illegal visa network making millions fleecing students",
    "content": "A global network has fleeced students out of tens of thousands
    "hashtags": "",
    "url": "https://www.bbc.com/news/articles/c97w1x2deyvo"
}
```


Fig. 12. The dataset output is in JSON format for testing 2.

2) Outcomes analysis result:

≡

SENTIMENT

Topics	#VisaFraud, #StudentExploitation, #UKImmigration, #CommunitySupport, #GovernmentPolicy
Cluster	Cluster : [Poverty and Economic Inequality, Health and Welfare]
Location	United Kingdom

#	Subject	Reason	Sentiment	Tone	Object
1	Taimoor Raza	Sold worthless visa documents to students, claiming they would enable employment in the UK.	Negative	Criticism	Individual
2	Students	Lost significant amounts of money and faced rejection of their visa applications due to invalid documents.	Negative	Complaint	Individual
3	Monty Singh and Sikh Advice Centre	Helped recover money for victims and raised awareness about the fraudulent activities.	Positive	Support	Organization

Fig. 13. Sentiment analysis results for testing 2.

The sentiment analysis and clustering of topics which is shown on Fig. 13 are accurately aligned with the context of the news, reflecting the key issues surrounding #VisaFraud, #StudentExploitation, #UKImmigration, #CommunitySupport, and #GovernmentPolicy in the United Kingdom. The topics are correctly categorized under the clusters Poverty and Economic Inequality and Health and Welfare, as they pertain to the financial exploitation of students and the efforts to provide community support.

The sentiment analysis is well-suited to the context:

1. Taimoor Raza, who sold fraudulent visa documents to students, is appropriately assessed with a Negative sentiment and a Criticism tone, as his actions resulted in harm to individuals.

2. The students, who suffered financial loss and visa rejection due to the scam, are given a Negative sentiment with a Complaint tone, representing their victimization and dissatisfaction.

3. Monty Singh and the Sikh Advice Centre, who supported the victims by helping them recover their money and raising awareness about the fraud, are appropriately assigned a Positive sentiment and Support tone, highlighting their positive role in addressing the situation.

This analysis effectively captures the different perspectives and reactions to the fraudulent visa scheme.

C. Testing 3

News URL:

https://www.bbc.com/news/articles/cze5x793569o

1) Data sets result: Fig. 14 represents the dataset output is in JSON format and is displayed on the web for testing 3.


```
{
    "id": 498,
    "date": "2024-09-01",
    "title": "Tens of thousands rally in Israel calling for hostage release deal",
    "content": "Tens of thousands of people have rallied across Israel after the bod
    "hashtags": "",
    "url": "https://www.bbc.com/news/articles/cze5x793569o"
}
```


Fig. 14. The dataset output is in JSON format for testing 3.

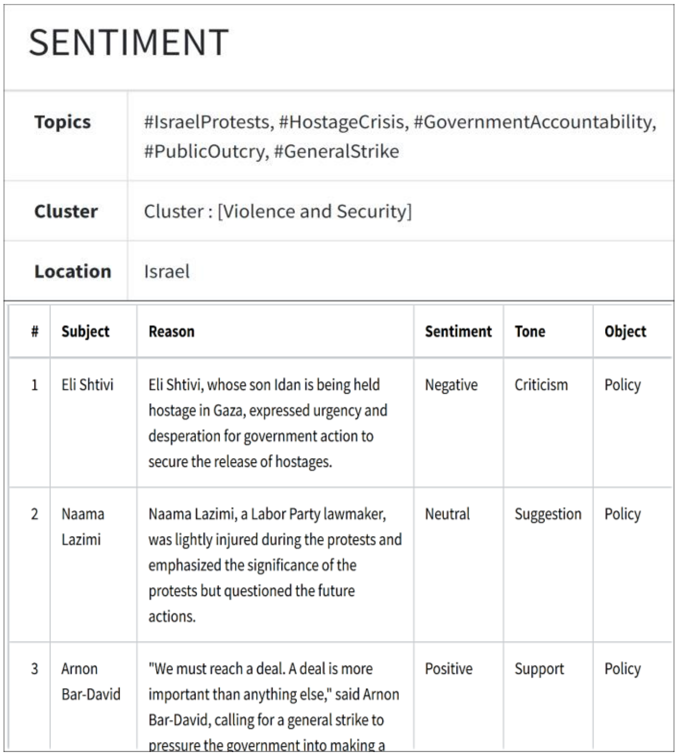
2) Outcomes analysis result: The sentiment analysis and clustering for the topics of #IsraelProtests, #HostageCrisis, #GovernmentAccountability, #PublicOutcry, and #GeneralStrike which is shown on Fig. 15 are highly appropriate, reflecting the Violence and Security cluster in the context of ongoing protests and government responses in Israel.

The sentiment analysis aligns well with the specific subjects:

1. Eli Shtivi, whose son is held hostage, expresses urgency and criticism toward government policy. The Negative sentiment and Criticism tone are fitting, given his frustration over the lack of action.

2. Naama Lazimi, a lawmaker injured in the protests, adopts a Neutral sentiment with a Suggestion tone as she reflects on the significance of the protests while questioning future governmental actions, making the sentiment and tone assessment appropriate.

3. Arnon Bar-David, advocating for a general strike, displays Positive sentiment and Support for government policy reform, focusing on pressuring the government into negotiations. This reflects the proactive and supportive stance towards achieving a resolution.



SENTIMENT

Topics	#IsraelProtests, #HostageCrisis, #GovernmentAccountability, #PublicOutcry, #GeneralStrike
Cluster	Cluster : [Violence and Security]
Location	Israel

#	Subject	Reason	Sentiment	Tone	Object
1	Eli Shtivi	Eli Shtivi, whose son Idan is being held hostage in Gaza, expressed urgency and desperation for government action to secure the release of hostages.	Negative	Criticism	Policy
2	Naama Lazimi	Naama Lazimi, a Labor Party lawmaker, was lightly injured during the protests and emphasized the significance of the protests but questioned the future actions.	Neutral	Suggestion	Policy
3	Arnon Bar-David	"We must reach a deal. A deal is more important than anything else," said Arnon Bar-David, calling for a general strike to pressure the government into making a	Positive	Support	Policy

Fig. 15. Sentiment analysis results for testing 3.

Overall, the clustering and sentiment analysis accurately capture the complexity of the emotions and actions tied to the protests and government accountability, making it consistent with the context of the news

ACKNOWLEDGMENT

This research titled "Deep Learning Approach in Complex Sentiment Analysis: A Case Study on Social Problems" was made possible through the support of The Ministry of Research, Technology, and Higher Education (Ristekdikti) of the Government of Indonesia under the 2024 Basic Research Grant Scheme for Higher Education. We would also like to extend our sincere gratitude to LPPM Universitas Merdeka Malang for their valuable assistance and support throughout this study. Without their funding and encouragement, this work would not have been possible.

REFERENCES

[1] A., A. G., & V., V. (2024). Sentiment analysis on a low-resource language dataset using multimodal representation learning and cross-lingual transfer learning. Applied Soft Computing, 157, 111553. https://doi.org/https://doi.org/10.1016/j.asoc.2024.111553

[2] Chen, N., Sun, Y., & Yan, Y. (2023). Sentiment analysis and research based on two-channel parallel hybrid neural network model with attention mechanism. IET Control Theory & Applications, 17(17), 2259–2267. https://doi.org/10.1049/cth2.12463

[3] Forsé, M., & Parodi, M. (2021). Sentiments de bien-être et de justice sociale. The Tocqueville Review, 42(1), 73–106. https://doi.org/10.3138/ttr.42.1.73

[4] Giménez, M., Fabregat-Hernández, A., Fabra-Boluda, R., Palanca, J., & Botti, V. (2024). A detailed analysis of the interpretability of Convolutional Neural Networks for text classification. Logic Journal of the IGPL. https://doi.org/10.1093/jigpal/jzae057

[5] Halawani, H. T., Mashraqi, A. M., Badr, S. K., & Alkhalaf, S. (2023). Automated sentiment analysis in social media using Harris Hawks optimisation and deep learning techniques. Alexandria Engineering Journal, 80, 433–443. https://doi.org/10.1016/j.aej.2023.08.062

[6] Islam, Md. S., Kabir, M. N., Ghani, N. A., Zamli, K. Z., Zulkifli, N. S. A., Rahman, Md. M., & Moni, M. A. (2024). "Challenges and future in deep learning for sentiment analysis: a comprehensive review and a proposed novel hybrid approach." Artificial Intelligence Review, 57(3), 62. https://doi.org/10.1007/s10462-023-10651-9

[7] Ismail, W. S. (2024). Emotion Detection in Text: Advances in Sentiment Analysis Using Deep Learning. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 15(1), 17–26. https://doi.org/10.58346/JOWUA.2024.I1.002

[8] Kumi, S., Snow, C., Lomotey, R. K., & Deters, R. (2024). Uncovering Concerns of Citizens Through Machine Learning and Social Network Sentiment Analysis. IEEE Access, 12, 94885–94913. https://doi.org/10.1109/ACCESS.2024.3426329

[9] Li, H., Ding, Y., Jiang, J., Deng, P., Yuan, D., & Yang, S.-H. (2022). BERT-based Sentiment Analysis of Chinese Online Social Movements. 2022 27th International Conference on Automation and Computing (ICAC), 1–6. https://doi.org/10.1109/ICAC55051.2022.9911123

[10] Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093–1113. https://doi.org/https://doi.org/10.1016/j.asej.2014.04.011

[11] Ogbuju, E., Mpama, I., Oluwafemi, T. M., Ochepa, F. O., Agbogun, J., Yemi-Peters, V., Owoeye, F. O., Idoko, M., & Taoheed, B. (2022). The Sentiment Analysis of EndSARS Protest in Nigeria. Journal of Applied Artificial Intelligence, 3(2), 13–23. https://doi.org/10.48185/jaai.v3i2.560

[12] Pandey, A., & Vishwakarma, D. K. (2024). Progress, achievements, and challenges in multimodal sentiment analysis using deep learning: A survey. Appl. Soft Comput., 152(C). https://doi.org/10.1016/j.asoc.2023.111206

[13] Peng, J., Fung, J. S., Murtaza, M., Rahman, A., Walia, P., Obande, D., & Verma, A. R. (2022). A sentiment analysis of the Black Lives Matter movement using Twitter. STEM Fellowship Journal, 8(1), 56–66. https://doi.org/10.17975/sfj-2022-015

[14] Shang, W., Chai, J., Cao, J., Lei, X., Zhu, H., Fan, Y., & Ding, W. (2024). Aspect-level sentiment analysis based on aspect-sentence graph convolution network. Information Fusion, 104, 102143. https://doi.org/https://doi.org/10.1016/j.inffus.2023.102143

[15] Xing, F., Schuller, B., Chaturvedi, I., Cambria, E., & Hussain, A. (2023). Guest Editorial Neurosymbolic AI for Sentiment Analysis. IEEE Transactions on Affective Computing, 14(3), 1711–1715. https://doi.org/10.1109/TAFFC.2023.3310856

[16] Zhao, Z., Liu, W., & Wang, K. (2023). Research on sentiment analysis method of opinion mining based on multi-model fusion transfer learning. Journal of Big Data, 10(1), 155. https://doi.org/10.1186/s40537-023-00837-x

[17] Zhou, T., Shen, Y., Chen, K., & Cao, Q. (2023). Hierarchical dual graph convolutional network for aspect-based sentiment analysis. Knowledge-Based Systems, 276, 110740. https://doi.org/10.1016/j.knosys.2023.110740

[18] Yadav, A., & Vishwakarma, D. K. (2020). Sentiment analysis using deep learning architectures: A review. Artificial Intelligence Review, 53(6), 4335-4385. https://doi.org/10.1007/s10462-019-09794-5

[19] Medhat, W., Hassan, A., & Korashy, H. (2021). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093-1113. https://doi.org/10.1016/j.asej.2014.04.011

[20] Lai, M., & Joty, S. (2019). Multi-task learning for cross-domain sentiment classification. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 1(2), 466-476. https://doi.org/10.18653/v1/P19-1044

[21] Chakraborty, T., Bhattacharyya, S., & Bag, R. (2020). A hybrid deep learning architecture for sentiment analysis. Journal of Network and Computer Applications, 144, 102738. https://doi.org/10.1016/j.jnca.2019.102738

[22] Sun, T., Qiu, X., Xu, Y., & Huang, X. (2019). Multi-channel CNN with hierarchical attention for sentiment classification. Information Processing

& Management, 56(3), 973-988. https://doi.org/10.1016/j.ipm.2018.02.008

[23] García-Pablos, A., Cuadros, M., & Rigau, G. (2020). W2VLDA: almost unsupervised system for aspect-based sentiment analysis. Expert Systems with Applications, 123, 207-221. https://doi.org/10.1016/j.eswa.2019.01.071

[24] Akhtar, M. M., Zamani, A. S., Khan, S., Shatat, A. S. A., Dilshad, S., & Samdani, F. (2022). Stock market prediction based on statistical data using machine learning algorithms. Journal of King Saud University - Science, 34(4), 101940. https://doi.org/10.1016/j.jksus.2022.101940

[25] Saeed, F., Ghani, S., & Nasir, M. H. (2021). Impact of deep learning on social media sentiment analysis: A systematic review. Sustainability, 13(8), 4385. https://doi.org/10.3390/su13084385

[26] Yu, B., Chen, Z., & Wang, H. (2021). Recurrent neural networks and sentiment analysis: Overview and applications. Neurocomputing, 439, 51-61. https://doi.org/10.1016/j.neucom.2020.10.020

[27] Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1253. https://doi.org/10.1002/widm.1253

[28] Zhou, Y., & Jiang, Y. (2022). Sentiment analysis using a deep learning model: A case study on the impact of COVID-19 on economic sectors. Applied Soft Computing, 114, 107893. https://doi.org/10.1016/j.asoc.2021.107893

[29] Alruily, M. (2023). Sentiment analysis for predicting stress among workers and classification utilizing CNN: Unveiling the mechanism. Alexandria Engineering Journal, 81, 360-370. https://doi.org/10.1016/j.aej.2023.09.040

[30] Abd, D. H., Khan, W., Khan, B., Alharbe, N., Al-Jumeily, D., & Hussain, A. (2023). Categorization of Arabic posts using artificial neural network and hash features. Journal of King Saud University - Science, 35(6), 102733. https://doi.org/10.1016/j.jksus.2023.102733

[31] Alayba, A. M., & Palade, V. (2022). Leveraging Arabic sentiment classification using an enhanced CNN-LSTM approach and effective Arabic text preparation. Journal of King Saud University - Computer and Information Sciences, 34(10, Part B), 9710-9722. https://doi.org/10.1016/j.jksuci.2021.12.004

[32] Atandoh, P., Zhang, F., Adu-Gyamfi, D., Atandoh, P. H., & Nuhoho, R. E. (2023). Integrated deep learning paradigm for document-based sentiment analysis. Journal of King Saud University - Computer and Information Sciences, 35(7), 101578. https://doi.org/10.1016/j.jksuci.2023.101578

[33] Elfaik, H., & Nfaoui, E. H. (2023). Leveraging feature-level fusion representations and attentional bidirectional RNN-CNN deep models for Arabic affect analysis on Twitter. Journal of King Saud University - Computer and Information Sciences, 35(1), 462-482. https://doi.org/10.1016/j.jksuci.2022.12.015