Sentiment Analysis on Banking Feedback and News Data using Synonyms and Antonyms

Aniruddha Mohanty, Ravindranath C. Cherukuri
Department of Computer Science and Engineering
CHRIST (DEEMED TO BE UNIVERSITY)
Bangalore, India 560074

Abstract*—Sentiment analysis is crucial for deciphering customers' enthusiasm, frustration, and the market mood within the banking sector. This importance arises from financial data's specialized and sensitive nature, enabling a deeper understanding of customer sentiments. In today's digital and social marketing landscape within the banking and financial sector, sentiment analysis is significant in shaping customer insights, product development, brand reputation management, risk management, customer service improvement, fraud detection, market research, compliance regulations, etc. This paper introduces a novel approach to sentiment analysis in the banking sector, emphasizing integrating diverse text features to enable dynamic analysis. This proposed approach aims to assess the sentiment score of distinct words used within a document and classify them as positive, negative, or neutral. After rephrasing sentences using synonyms and antonyms of unique words, the system calculates sentence similarity using a distance control mechanism. Then, the system updates the dataset with the positive, negative, and neutral labels. Ultimately, the ELECTRA model utilizes the self-trained sentiment-scored data dictionary, and the newly created dataset is processed using the SoftMax activation function in combination with a customized ADAM optimizer. The approach's effectiveness is confirmed through the analysis of post-bank customer feedback and the phrase bank dataset, yielding accuracy scores of 92.15% and 93.47%, respectively. This study stands out due to its unique approach, which centers on evaluating customer satisfaction and market sentiment by utilizing sentiment scores of words and assessing sentence similarities.

***Keywords*—ELECTRA; Synonyms and antonyms; sentiment analysis; datasets; sentiment score; control distance**

I. Introduction

Enhanced comprehension of customer perceptions regarding various banking products and services entails evaluating customers' sentiments, opinions, and attitudes [1]. In the age of digitalization and advanced data analytics, the analysis of social media feedback data has become prevalent. The Banking, Financial Services, and Insurance (BFSI) sector utilizes Customer Relationship Management (CRM) to make informed business decisions based on customer feedback received daily. Nowadays, this feedback data is accessible on social media platforms, significantly aiding in analyzing customer sentiments [2]. Organizations are investing significant resources in research to create tools and strategies for analyzing customer feedback data and aligning products with current market trends [3]. Likewise, banking news reflects the general atmosphere or attitude conveyed in articles, reports, and discussions rooted in economic sentiment. Researchers increasingly favor text-based economic activities [4] due to their advantages over surveys in terms of cost, scope, and timeline.

Constraints imposed and integrity issues associated with text-based sources like news, microblogs, and organizational product disclosures have limited research in this field when analyzing their impacts on various market aspects [5]. Sentiment analysis represents a distinctive feature within the financial sector, enabling sentiment analysis that gauges customer confidence in banking products and services. Banks can pinpoint recurring issues or concerns raised by customers and promptly address them, thus enhancing customer satisfaction and loyalty. Utilizing feedback data allows for enhancing existing products and developing new services tailored to customer preferences while bolstering the organization's reputation and competitive offerings. Identifying unusual patterns and sentiments also serves the purpose of detecting fraud and can be harnessed for customer awareness [6]. Similarly, favorable banking news increases investor confidence, contributes to an upward market trend, and enhances an organization's reputation. Conversely, negative information can have the opposite effect [4]. Public sentiment regarding regulations and policies can sway lawmakers and regulators, shaping the formulation of crisis management strategies.

Contextual comprehension significantly influences sentiments; for instance, 'interest' typically evokes positive sentiment, while 'loans' often yield negative sentiments. Challenges arise in sentiment analysis due to sarcasm, subjectivity, and multilingual nuances in language. Within financial institutions, data often holds sensitive information, making balancing privacy regulations vital when handling customer feedback, reviews, and media releases. Leveraging the integration of machine learning and Natural Language Processing (NLP) [7] techniques proves instrumental in mitigating these challenges.

In this paper, various sources known for their integrity gather the data. After pre-processing and clustering the data, the system computes sentiment scores. The analysis system ensures the integrity of the document's content, focusing on preserving form rather than just content. This approach helps to prevent different classes resulting from the synonyms and antonyms in the text document. The conservation block compares the normalized sentences of the analysis system with newly generated sentences containing synonyms and antonyms, utilizing a control distance mechanism. This process results in an updated dataset for the experiment. The analysis system provides the updated dataset as input to the ELECTRA, a self-supervised language representation learning model for classifying responses within the context of the proposed customer-based banking analysis system.

The rest of the paper includes Section II, which elaborates on banking products and services. Section III presents related work on sentiment analysis, focusing on customer feedback and news from diverse sources. Section IV delves into the proposed system model and its architecture. The research methodology of the model is discussed in Section V, followed by result analysis and discussion of the proposed hypothesis with novelty, strength, and implication in Section VI. Finally, Section VII summarizes the paper, providing conclusions and future scope prospects.

II. BANKING PRODUCTS AND SERVICES

The market offers a variety of banking products, each differing from one organization to another. These products and services undergo regular updates and enhancements. The banking system has two primary categories: retail banking and corporate banking. Table I provides an overview of the products and associated services offered by retail and corporate banking sectors.

TABLE I. PRODUCTS OF RETAIL AND CORPORATE BANKING

Retail Banking	Corporate Banking
Checking and saving account, Certificates of deposit, Mortgage, Automobile financing, Credit cards, Lines of credit (Home equity lines of credit and personal credit products), Foreign currency and remittance services, Stock brokerage, Insurance, Wealth management, Private banking.	Loans and other credit products, Treasury and cash management services, Equipment lending, Commercial real estate, Trade financing, Employer service.

III. RELATED WORK

Sentiment analysis assesses customers' responses to products, services, and situations by analyzing texts, posts, reviews, news, and other digital content. This analysis aids business leaders in comprehending customer attitudes and market perceptions over time.

A. Sentiment Analysis on Banking Data

The bank ontology facilitates extracting text feedback data features from websites like Mouthshut.com and myBankTracker.com. The experiment then applies sentiment classification to this data [8]. Rule-based classifier helps to analyze the sentiments of Russian review texts to classify sentiments [9]. The Vader Aware Dictionary and Sentiment Reasoner (VADER) [10] is a lexicon and rule-based sentiment analysis tool specifically crafted to discern sentiments within social media data related to UniCredit bank's European region. Retail banks in South Africa employ both lexicon-based and machine learning-based methods [11] to assess customer feedback sentiments. The results from a fine-tuned DistilBERT model are fed into machine learning classifiers such as Random Forest, Decision Tree, Logistic Regression, and Linear Support Vector Classifier (SVC) [12] to categorize sentiments in news related to Indian banking, governmental, and global topics. Nine classifiers, including Naïve Bayes, Logistic Regression, K-Nearest Neighbours, Support Vector Machines, Random Forest, Decision Tree, Adaptive Boosting, eXtreme Gradient Boosting, and Light Gradient Boosting Mechanism [13], have been employed to detect customer satisfaction levels within Indonesian banks, such as Jenius, Jago, and Blu. The BERTopic architecture utilizes a combination of Kernel Principal Component Analysis (KernelPCA) and K-means Clustering to generate coherent topics, similar to the Latent Dirichlet Allocation (LDA) [14] approach. This method calculates coherence scores for Nigerian bank data, facilitating sentiment analysis. Word representation has evolved by integrating static and contextual words to handle language ambiguity, encompassing semantics, and syntax within a given context. These word representations are fed into a Convolutional Neural Network (CNN) [15] to capture sentiments within financial news contexts.

B. Sentiment Analysis Approach on Text Data

The RCNN [16] analyzes each word's context in the document and then applies the max-pooling layer to identify the text's crucial elements for classification, along with the SoftMax layer. Clustering techniques, such as the K-means-type algorithm [17], exhibit improved performance when applied to balanced data, whereas the designed weighting model delivers exceptional results for both balanced and unbalanced datasets. The word embedding layer [18] converts sentences into words, preserving the contextual information of each word, and then applies a CNN for sentiment analysis. Support Vector Machine (SVM), Deep Learning (DL), and Naïve Bayes (NB) [19] classifiers utilize sentiment scores and the associated weights of hashtags to classify sentiments within social media data. The classification of sentiments [20] involves employing the Bidirectional Long Short-Term Memory (BiLSTM) layer, a global pooling mechanism, and a sigmoid layer. The Collaborative and Bidirectional Gated Recurrent Unit (BiGRU) can also adopt this approach to assess performance. Hierarchical Attention Networks (HAN) [21] enable the identification of sentiment polarity in customer communications, thereby enhancing the efficiency of Customer Relationship Management (CRM) operators in terms of response time. The integration of the BERT model with Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) algorithms [22] facilitates the analysis of positive, negative, and neutral sentiments.

The literature analysis above demonstrates various modelling approaches employed in sentiment analysis of banking text data, incorporating different types of text data and data corpora. Examining synonyms and antonyms within a data corpus is instrumental in simplifying the identification of customer sentiments toward financial organizations in diverse regions. Further research is necessary to develop a more efficient approach capable of swiftly discerning customers' intent, serving as a driving force to create a more streamlined sentiment analysis context for optimizing banking business strategies.

IV. SYSTEM MODEL

The proposed sentiment analysis approach comprises two key components, as illustrated in Fig. 1. The initial phase involves collecting banking data from various sources, including customer service teams, visual feedback tools, review sites, net promoter sources, online surveys, social media, and news media facilitated by banking organizations. These data are then gathered and stored in a database. A "Stop-word" data cloud developed helps to preprocess the dataset documents.

The filtered and pre-processed text data identifies the sentiment scores for unique words. The subsequent step involves determining synonyms and antonyms for these unique words and their corresponding sentiment scores. Replacing each unique word with synonyms and antonyms forms new sentences, as depicted in Fig. 1. This process allows considering a maximum of three synonyms and antonyms for each identified word during replacement.

The experiment employs the Control Distance approach to assess the similarity between the original and newly generated sentences by rephrasing with synonyms and antonyms. Based on their degree of similarity, the experiment uses the resulting similarity score to categorize sentences as positive, negative, or neutral.

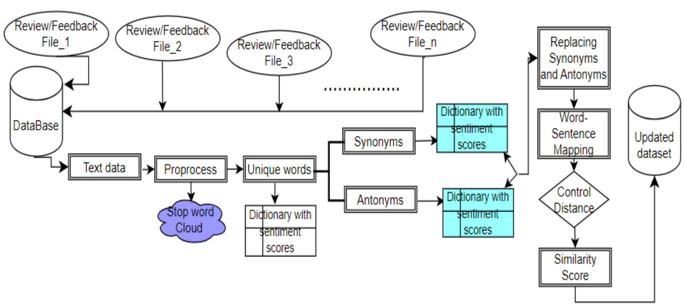

Fig. 1. System model to create an updated dataset.

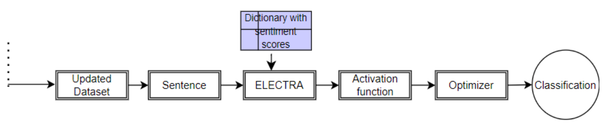

Fig. 2. System model to identify sentiment.

A self-supervised language representation learning model, ELECTRA, trains itself using all words from the data dictionary with sentiment scores. This model is then applied to the updated dataset, as Fig. 2 illustrates. Subsequently, individual sentences are classified and identified for sentiment, enhancing the efficiency of the Sentiment Analysis framework. ELECTRA operates efficiently on sample words, making it faster and requiring fewer resources. The following provides a summary of the detailed implementation:

- The initial step involves gathering datasets from a variety of banking data sources.
- Design and develop the 'Stop-word' word cloud that does not impact sentiments.
- Then, pre-processed the text data using the specially designed 'Stop-word' word cloud.
- Create a data dictionary by identifying unique words and their corresponding sentiment scores.
- The positive, negative, and neutral sentiment scores determine synonyms and antonyms of unique words and establish a distinct data dictionary for these synonyms and antonyms.
- A control distance approach, as proposed, assists in measuring the similarity score between the original and updated sentences. The resulting dataset encompasses positive, negative, and neutral sentiments.
- The ELECTRA model is self-trained and then applied to the updated dataset by merging the data dictionary containing synonyms, antonyms, unique words, and their sentiment scores. SoftMax activation and a modified Adam optimizer help to identify the sentiment for each feedback statement.

V. Research Methodology

This section delves into the analysis of sentiments from banking feedback and media data, encompassing a discussion of the experiments and the results obtained. Firstly, it covers the specifics of the dataset employed for training, cross-validation, and testing, followed by a detailed examination of the approach's implementation. Following this, the analysis delves into the experiment results, revealing the sentiments and information patterns present within the datasets. The study evaluates the effectiveness and performance of the proposed approach by comparing it with other established methods. Furthermore, the analysis includes applying the predefined approach to the post banking customer feedback and phrase banking datasets to determine the accuracy score, given their absence of prior implementation using the ELECTRA model.

A. Dataset Collection

Central banks worldwide gather customer feedback on banking services, encompassing satisfaction levels, complaints, and insights on product usability. The sources of post bank customer data [23] are from the Russian finance website "www.banki.ru" spanning 2013 to 2019. Reviews are rated from one to five, with one denoting negativity and five indicating positivity. The "responses_header" column comprises feedback messages from 16,659 customers.

The phase bank dataset [24] includes positive, negative, and neutral sentiments of 5,000 customers about companies listed in OMX Helsinki. The data originates from the LexisNexis database, including 10,000 randomly selected articles from limited financial and economic resources. Analyzing these sentiments assists the marketing team in improvising the banking products and services. The details of collected datasets are

TABLE II. The Datasets in the Experiment

sl#	Dataset	Provided Ratings
1	Post Bank Customer Review	Positive: 2160
2		Negative: 2472
3		Neutral: 1391
4	Phrase Banking	Positive: 1853
5		Negative: 861
6		Neutral: 3131

shown in Table II and saved in CSV files to calculate the sentiment content of each text sentence. These datasets are divided into 65% for training, 15% for cross-validation, and 20% for testing.

B. Text Data

Each dataset includes a variety of text messages shared by different banking customers and media sources, extracting subjective information from the shared data. The goal is to discern the attitude, emotional tone, market sentiment, and expressions within a text, allowing for the analysis of overall sentiments from individual feedback and media statements. Various methods, such as rule-based, machine learning, and deep learning techniques, are employed to assess these sentiments.

The article employs a rule-based approach to measure the sentiment scores of unique words in the dataset file. It creates data dictionaries that assign sentiment scores to unique words, synonyms, and antonyms. These dictionaries help to determine the overall sentiment scores of sentences or paragraphs.

C. Pre-processing the Text Data

Text pre-processing involves cleaning text data by eliminating irrelevant information like URLs, numbers, and punctuation marks. This step ensures the availability of more pertinent texts for conducting sentiment analysis activities. Then, the process [25] involves tokenizing, normalizing, removing stop words, stemming, and lemmatization.

In this implementation, aside from general stop words, specific banking-related stop words such as currency names (Rupee, dollar, etc.), Roman numerals (I, II, III, etc.), and auditing firms (KPMG, Deloitte, etc.) are utilized. These stop words significantly impact sentiment determination and play a significant role in determining sentiment. They are used to generate the word cloud, which, in the process, assists in removing these words from the dataset files, as shown in Fig. 3. In pre-processing and stop word removal, the process

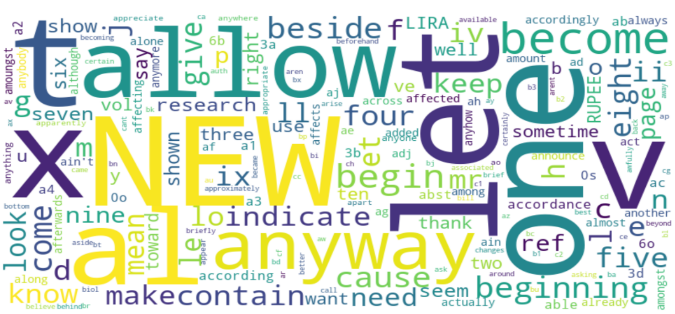

Fig. 3. Generated word cloud for stop words.

extracts unique words, facilitating the identification of context and reducing words to their root form. Identifying distinct words in a dataset facilitates the analysis of individual words for information retrieval purposes. The extraction of unique words serves several functions, including:

- Analyzing the richness of vocabulary in the data corpus.
- Constructing a data dictionary.
- Accessing the diversity and complexity in the written content.

WordNet, a computational lexical database with linguistic and physiological features [26], identifies English verbs, nouns, adjectives, and adverbs through part-of-speech tagging. Table III illustrates some unique words extracted from the data corpus.

TABLE III. Few Unique Words from the Data Corpus

solutions	powerful	multimedia	company	relevant
sales	search	technology	moved	doubled
turnover	platform	location	leverage	strengthened
content	communities	domestic	profit	loss
period	project	recorded	ownership	generated

D. Sentiment Score and Data Dictionary

Determining the significance of specific words within a group of relevant words is vital in determining the sentiment scores. The newly proposed approach calculates this importance by assigning sentiment scores, following the formulation, as depicted in Eq. (1).

$$SC = WP * \log(n/N) \tag{1}$$

where, SC is the sentiment score, WP is the word presence, n is the frequency of the unique word W_i, and N is the number of words present in the document. As $\log()$ is used in the formulation, SC value will be less and vice versa if the frequency of the word is more. WP takes the value as 0 or 1 for the presence or absence of the words in the document.

It can be challenging to obtain unique words and their corresponding sentiment scores for the next step. Therefore, a metadata repository, termed a data dictionary, is established to store words and their associated sentiment scores (SC). This data dictionary elucidates each word's context, offering crucial information for future reference without the need for in-depth analysis of the raw data. The SC value categorizes words as positive, negative, or neutral based on predefined threshold values, denoted as L_{value} and H_{value} in Eq. (2).

$$\begin{cases} positive & \text{if } SC \leq L_{value} \\ negative & \text{if } L_{value} < SC > H_{value} \\ neutral & \text{if } SC \geq H_{value} \end{cases} \tag{2}$$

The final step involves clustering the banking data document's positive, negative, and neutral words to serve as input for the subsequent phase.

In this implementation, words with a threshold value below and equal to 0.003 are classified as "positive," while words between 0.003 and 0.007 are considered "negative," and words exceeding the threshold of 0.007 are "neutral." These values are pivotal in creating a new dataset sorting statements into positive, negative, and neutral categories and establishing a data dictionary for positive, negative, and neutral categories containing unique words and their corresponding sentiment scores. Table IV presents a selection of unique words with their sentiment scores.

TABLE IV. UNIQUE WORDS WITH SENTIMENT SCORES

Sl#	Unique words	Sentiment Score
1	profit	0.0021
2	search	0.0014
3	relevant	0.0043
4	multimedia	0.0055
5	platform	0.0081

E. Synonyms and Antonyms with Sentiment Scores

Synonyms are words or phrases sharing identical meanings, allowing their replacement in a specific context without altering the meaning. Antonyms represent words with opposite meanings, conveying contrasting ideas in specific contexts. Comprehending synonyms and antonyms is vital for language and communication, fostering a diverse vocabulary and enhancing effective expression. This knowledge equips writers and speakers to convey messages precisely, highlighting contrasts and differences between ideas and concepts. WordNet [27] discerns connections between word meanings, identifying relationships such as synonyms and antonyms. For instance, “move, drive, impel” are synonymous words, while “stay, stop, discourage” are antonyms.

In this implementation, the process extracts synonyms and antonyms from unique words. It then allocates positive sentiment scores to synonyms, retains the same scores for unique words, and assigns antonyms a negative sentiment score with an identical value. Table V presents a selection of synonyms and antonyms with their sentiment scores corresponding to unique words.

TABLE V. SYNONYMS AND ANTONYMS WORDS WITH SENTIMENT SCORES

Sl#	Unique words	Synonyms	SC	Antonyms	SC
1	profit	financial gain	0.0021	loss	-0.0021
2	search	explore	0.0014	ignore	-0.0014
3	relevant	pertinent	0.0043	irrelevant	-0.0043
4	multimedia	multimodal	0.0055	monomodal	-0.0055
5	platform	podium	0.0081	ground	-0.0081

F. Sentence Similarity and Updated Dataset

Substituting synonyms and antonyms results in an updated dataset derived from the existing one. The newly proposed method, Control Distance, calculates sentence similarity after rephrasing these words, as depicted in Fig. 4.

1) Control Distance: Calculating the sentence similarity between two strings involves determining the minimum number of edits (insertions, deletions, or substitutions) necessary to transform one string into another. This method also handles topographical inconsistencies in string data. In this implementation, the proposed expression assesses the similarity or dissimilarity between the original sentence and the sentences with substituted synonyms and antonyms.

$$SS_v = n * \frac{\sum_{i=1}^{n} SC_i}{N} \tag{3}$$

Here, SS_v is the Sentence similarity, and n is the words substituted in a sentence or a paragraph. N is the sum of

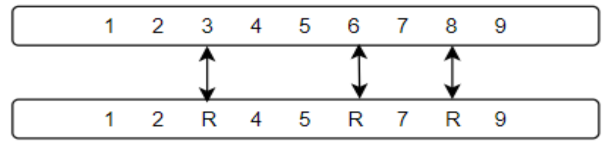

Fig. 4. Replaced words on original statement.

the words present in a sentence or a paragraph. SC_i is the sentiment score of the substituted words.

After determining the Sentence Similarity (SS_v) between the original sentence and the sentences substituted with synonyms and antonyms, the SS_v is utilized to classify the sentences as similar or dissimilar by applying a standard threshold value, T_v.

$$Similarity = \begin{cases} Similar & SS_v \geq T_v \\ Dissimilar & SS_v < T_v \end{cases} \tag{4}$$

The similarity of the sentences determines the corresponding positive, negative, and neutral sentiments assigned to the original sentences.

Standard threshold values help measure sentences’ positive, negative, and neutral sentiments. Sentences with similarity values ranging from 0 to 3.0 are classified as positive, while those with similarity scores between 0 and -3.0 are negative. The remaining sentences with different similarity scores are considered neutral. Table VI presents a few of the sentence similarity scores.

TABLE VI. SIMILAR AND DISSIMILAR SENTENCES

Sl#	Statement	Similarity Score	Sentiment
1	L&G still paying price for dividend cut during crisis, chief says	Nill	Positive
2	L&G still rewarding price for dividend cut during emergency, chief says	2.26	Positive
3	L&G still remunerating price for dividend cut during disaster, chief says	1.93	Positive
4	L&G still compensating price for dividend cut during hardship, chief says	2.67	Positive
5	L&G still forgiving price for dividend cut during normalcy, chief says	-1.63	Negative
6	L&G still withholding price for dividend cut during certainty, chief says	-0.98	Negative
7	L&G still deferring price for dividend cut during harmony, chief says	-2.16	Negative

G. Model Implementation with Optimizer

ELECTRA (Efficiently Learning an Encoder that Classifies Token Replacements Accurately) [28] stands as a self-supervised language representation learning model. Its pretraining objective resembles the traditional Masked Language Model (MLM) but incorporates binary classification objectives. ELECTRA is a binary classifier during pretraining but can help adapt for multiclass classification tasks through finetuning.

ELECTRA operates through a Discriminator (D) designed to differentiate between “real” and “forged” tokens within a sentence. Additionally, it includes a Generator (G) network that substitutes certain input tokens with incorrect ones. Both the generators and discriminators are composed of transformer encoder layers, following the formulation below.

- Masked token replacement with sentiment score: The sequence token $T = [t_1, t_2, t_3, t_4, ..., t_i, ..., t_n]$ with associated sentiment score $SC = [SC_1, SC_2, SC_3, SC_4, ..., SC_i, ..., SC_n]$ where t_i represents the i^{th} token with corresponding sentiment score SC_i. The process of generating masked token operates as follows: for token t_i, it is replaced by a mask token $[MASK]$ with sentiment score SC_i.

- Generator (G): The generator (G) attempts to predict original token t_i from the $[MASKED]$ token, and its output is a probabilistic distribution over the vocabulary of each position i. Generator Loss (L_G) is calculated as

$$L_G = \sum_{i=1}^{n} SC_i * \sum_{j=1}^{V} t_i \log G([MASK], t_{connect}) \quad (5)$$

Here, V represents the vocabulary size, $G([MASK], t_{connect})$ signifies the output probabilities of the generator for masked taken $[MASK]$ given the context $t_{connect}$.

$$\begin{aligned} L_D = \sum_{i=1}^{n} SC_i * t_i \log D(t_i, t_{connect}) + \\ (1 - t_i) \log(1 - D([MASK], t_{connect}))) \end{aligned} \quad (6)$$

Here, $D(t_i, t_{connect})$ represents discriminator prediction for the token t_i with given context $t_{connect}$. Both the generator and discriminator are trained to minimize their respective loss based on the Sentiment scores.

To produce a probabilistic distribution over masked token, the SoftMax Function [29] is employed in the generator component. The function converts raw score denoted as $[S_1, S_2, S_3, S_4, ..., S_i, ..., S_v]$ into probabilities $[P_1, P_2, P_3, P_4, ..., P_i, ..., P_v]$ using below formula

$$SoftMax(P_i) = \frac{e^{S_v}}{\sum_{j=1}^{V} e^{S_j}} \quad (7)$$

The raw score generated by the ELECTRA generator is passed through the SoftMax function to obtain the probability of each token being the correct replacement for the masked token.

An optimizer is a crucial algorithm employed to fine-tune neural network attributes, such as learning rates and weights, to minimize loss during training. Its main purpose is to identify the optimal parameters that minimize the disparity between predicted and actual values. In this particular implementation, using the Adaptive Moment Estimation (ADAM) optimizer [30] proves invaluable for obtaining optimal results from the ELECTRA model. Additionally, integrating the L2 regularizer [31] into the modified Adam optimizer enhances convergence, addresses sparse gradient issues, prevents computation in local minima, and handles imbalanced data effectively.

H. Sentiments

Banking customers convey feelings about specific media data, brands, products, or services, encompassing attitudes, opinions, and emotions. Analysts examine sentiments based on written or voice feedback provided by customers and the mass media data, typically classifying them into three categories: positive, negative, and neutral [17], reflecting different emotional tone.

- **Positive:** Expresses satisfaction, joy, love, excitement, etc.
- **Negative:** Conveys feelings of sadness, anger, disappointment, displeasure, etc.
- **Neutral:** Lacks strong positive or negative emotions, presenting factual text without any specific emotional tone.

An automated sentiment analysis approach that identifies these emotions in feedback and media text documents offers valuable insights into the emotional content, proving particularly useful in banking data analysis.

I. Algorithm for Sentiment Analysis

The suggested approach for analyzing sentiment from textual data involves four distinct phases. Algorithm 1 addresses text data pre-processing. Algorithm 2 outlines the identification of Synonyms and Antonyms associated with individual words and their respective Sentiment Scores. Algorithm 3 illustrates the generation of an enhanced Dataset. Lastly, Algorithm 4 demonstrates sentiment classification utilizing the ELECTRA model on the updated dataset.

Algorithm 1 Pre-Processing the text data

Notations: Banking_Feedback_Corpus (Tokens, Sentences, Sentiment, Threshold_1, Thrreshold_2)
INPUT: TextFile (Banking_Feedback_Corpus_File)
OUTPUT: Sentiment= (Positive, Negative, Neutral)

```
Begin
Read text data from Banking_Feedback_Corpus_File
Remove the URLs, Numbers, Punctuations
Words ← Split the Banking_Feedback_Corpus_File by Space
words ← lower (Words)
Add "words" to "processing_words" library
Create "Stop words" Word Cloud
Find distinct stemming words to the "stemmed_word" container
Add "stemmed_word" to "processing_text" library
```

VI. Results and Discussion

The Sentiment Analysis model uses Python, relying on various frameworks and machine learning libraries for implementation. Python (specifically Python $3.6.3rc1$) and the NLTK 3.0 library are used for text data processing, providing a range of built-in functions. The Seaborn and Matplotlib libraries for statistical data analysis and visualization help plot the graph. Additionally, the experiment utilizes machine

Algorithm 2 Extracting Synonyms and Antonyms from Unique words with Sentiment Score

unique_words ← set (processing_text)
unique_words_list ← list (unique_words)
Read the "processing_words" library
for word in unique_words_list **do**
  X ← n/N
  Y ← log (X)
  SC ← WP * Y
  if $SC \geq Threshold\_1$ **then**
    Positive Sentiment
  else if $SC \geq Threshold\_1$ AND $SC \leq Threshold\_2$ **then**:
    Negative Sentiment
  else
    Neutral Sentiment
  end if
end for
Create data Dictionary words with Sentiment Scores
synonyms, antonyms ← get_synonyms_antonyms (unique_words_list)
Create data dictionary Unique word, Synonyms, and Positive SC
Create data dictionary Unique word, Antonyms, and Negative SC

Algorithm 3 Creating Updated Dataset

Replace three synonyms and antonyms words on original sentence from text file
Original Sentence
sentence_to_compare= [sen_1, sen_2, sen_3, sen_4, sen_5, sen_6]
for sentence in sentence_to_compare **do**
  n ← number of replaced words in Original Sentence
  SS= []
  for i= 1 to n **do**
    Read SC_i
    Sum= Add (SC_i)
    SS = n * (Sum/N)
    if $SS > Similarity\_value$ **then**
      Similar
    else
      Dissimilar
    end if
  end for
end for
Update the sentiment of the sentences with similarity score

Algorithm 4 Sentiment Analysis using ELECTRA

Replace three synonyms and antonyms words on original sentence from text file
Segregate the dataset to train, cross-validation, and test
Self-Train the model with words, and the sentiment scores
ELECTRA Modelling applied Sentence/Paragraph
OUTPUT: Sentiment= Positive, Negative, Neutral
Measure Accuracy
End

learning libraries such as Keras (version 2.6.0), TensorFlow (version 2.6.0), and Scikit-learn.

A. Results

In this setup, the assessment of the model's validation performance uses various metrics from the confusion matrix, including accuracy, recall, precision, and F-score. A rigorous $10 - fold$ cross-validation approach helps to evaluate the sentiment scores. Multiple scenarios, including the original dataset, datasets with synonyms replaced, datasets with antonyms replaced, and combinations help to estimate the model's performance. As a constraint, the implementation includes only three synonyms and antonyms words. Table VII displays the model performance on the Postbank dataset, as well as datasets containing synonyms and antonyms rephrased sentences.

Fig. 5 illustrates a graphical representation of negative, neutral, and positive sentiments. The basis of this representation is on performance measures such as False Positive Rate (FPR) and False Negative Rate (FNR) values. The analysis includes the post-bank dataset, datasets with replaced synonyms and antonyms. The analysis of the post-bank dataset

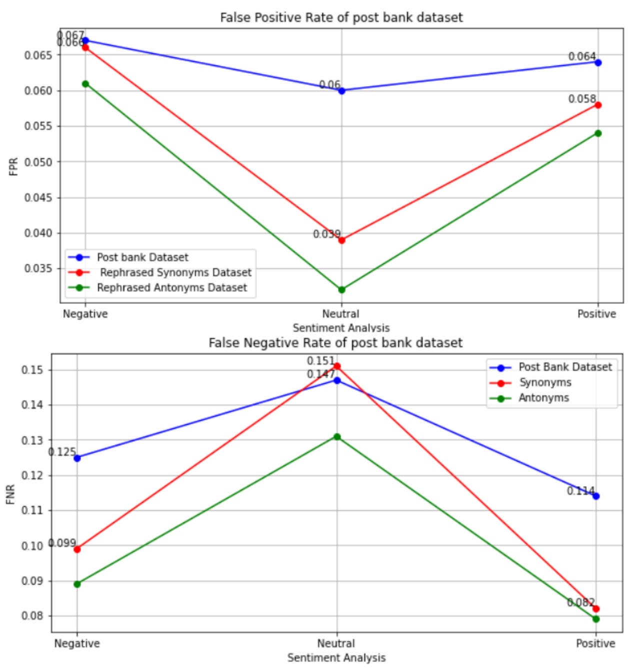

Fig. 5. FPR and FNR values for post bank, rephrase synonyms and antonyms dataset.

involves a combination of sentences with rephrased synonyms and antonyms, along with the original sentences. Table VIII presents the performance metrics for the post-bank dataset, including sentences with rephrased synonyms and antonyms.

Regarding the performance metrics, neutral sentiments are exhibiting lower performance compared to other sentiments. Similarly, the assessment of the performance of the Phrase bank dataset includes sentences with rephrased synonyms and antonyms. Table IX displays the model performance on the Phrase bank dataset, as well as datasets containing synonyms and antonyms rephrased sentences.

TABLE VII. Performance Metrics of Post Bank, Synonyms and Antonyms Datasets

Sl#	Parameters	Original Dataset			Synonym Dataset			Antonym Dataset		
		negative	neutral	positive	negative	neutral	positive	negative	neutral	positive
1	Precision	87.26	84.80	88.98	87.16	90.41	89.81	88.23	92.01	90.50
2	Recall	87.40	85.22	88.50	90.05	84.83	91.72	91.07	86.81	92.07
3	F1 score	87.33	85.01	90.76	88.74	87.53	90.76	89.63	89.34	90.28
4	Support	1176	1068	1269	1176	1068	1269	1177	1062	1274

TABLE VIII. Performance Measure of Post Bank Customer Data on Testing Data

PÎndex	Negative	Neutral	Positive	Accuracy	MÂvg	WÂvg
Precision	89.8765	95.6587	91.5895	92.1577	92.3749	0.922754
Recall	93.1741	88.3003	94.5440	92.1577	92.0061	0.921577
f1 Score	91.4956	91.8323	93.0433	92.1577	92.1237	0.921528
Support	2344	2171	2511	92.1577	7026	7026

Fig. 6 depicts a graphical representation of negative, neutral, and positive sentiments based on performance metrics such as False Positive Rate (FPR) and False Negative Rate (FNR) values. The analysis of the Phrase bank dataset involves

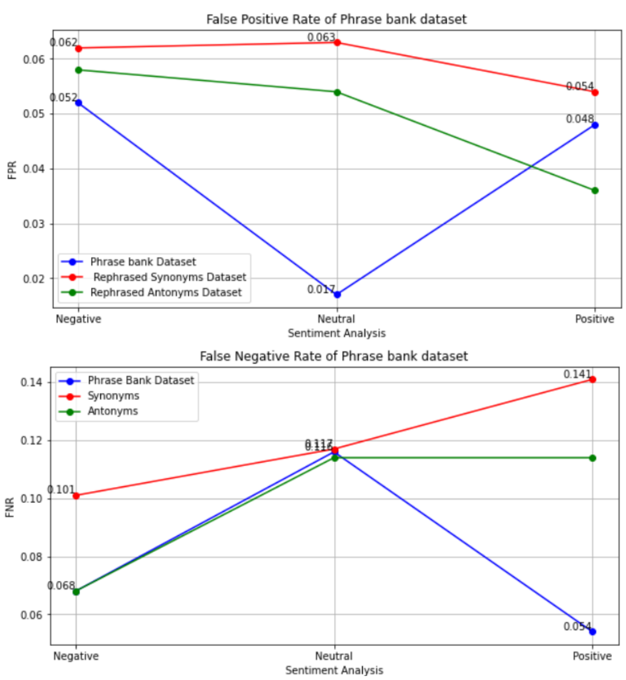

Fig. 6. FPR and FNR values for phrase bank, rephrase synonyms and antonyms dataset on testing data.

a combination of sentences with rephrased synonyms and antonyms, along with the original sentences. Table X presents the performance metrics for the Phrase bank dataset. The proposed model analyzes both the Post bank customer data and Phrase bank datasets. These datasets contain statements with rephrased synonyms and antonyms alongside original statements or paragraphs. The analysis utilizes the modified ADAM optimizer. Fig. 7 illustrates the overall performance of the proposed model for both datasets.

The implementation is validated using the ADAM optimizer and the modified Adam optimizer, incorporating L2 regularization. The modified Adam optimizer outperforms the

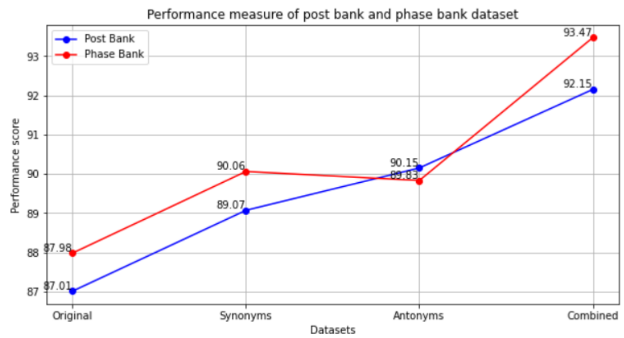

Fig. 7. Overall performance of the post bank and phrase bank dataset on testing data.

standard Adam optimizer significantly. Table XI shows the comparative performance results.

B. Performance Comparison

Utilizing the banking dataset in the proposed model yields superior results to conventional datasets such as general English sources (Wikipedia and BooksCorpus) and banking datasets like Phrase Bank.

The Name Entity Recognition (NER) task employs the financial language model based on ELECTRA [32]. In NER, a knowledge graph assists in comprehending the connections among various financial entities, such as individuals, organizations, and locations. The FLANG model achieves 82% accuracy in NER when applied to general English datasets like Wikipedia and BooksCorpus. However, the NER and FLANG concepts are also applied to the post-bank dataset, resulting in an accuracy of 82.9%.

The proposed approach is applied to analyze sentiments in the financial Phrase Bank dataset and financial tweets [33] utilizing the FinBERT model. Incorporating the CLS token, a special marker used for sequencing classification task representations, the model achieves an accuracy of 87.1%. Implementing the Phrase bank dataset [24] involves applying the ELECTRA model with Self-Attention and prediction layers, resulting in an accuracy of 83.87%, as shown in Table XII. This experiment represents the initial utilization of the Phrase bank dataset with ELECTRA and CLS tokens. Before this, ELECTRA had not incorporated the Phrase bank dataset.

C. Strength and Implication

The implementation results have numerous practical and theoretical implications. Theoretically, this research activity enhances our understanding of the sentiments conveyed by individual words in a sentence. It sheds light on market

TABLE IX. PERFORMANCE METRICS OF PHRASE BANK, SYNONYMS AND ANTONYMS DATASETS

Sl#	Parameters	Original Dataset			Synonym Dataset			Antonym Dataset		
		negative	neutral	positive	negative	neutral	positive	negative	neutral	positive
1	Precision	87.75	87.75	88.47	88.88	89.34	92.13	87.68	90.97	90.50
2	Recall	89.85	88.26	85.80	93.13	88.56	88.51	90.64	88.66	90.23
3	F1 score	88.79	88.01	87.11	90.96	88.95	90.29	89.13	89.80	90.59
4	Support	335	341	331	670	682	662	1005	1023	993

TABLE X. PERFORMANCE MEASURE OF PHRASE BANK DATA ON TESTING DATA

PÎndex	Negative	Neutral	Positive	Accuracy	MÂvg	WÂvg
Precision	93.7876	93.2171	93.4410	93.479	93.4819	93.4805
Recall	93.1343	94.0371	93.2528	93.479	93.4747	93.4790
f1 Score	93.4598	93.6253	93.3468	93.479	93.4773	93.4787
Support	1005	1023	993	93.3468	3021	3021

TABLE XI. COMPARISON OF MODEL PERFORMANCE ON ADAM AND MODIFIED ADAM OPTIMIZER

Sl#	Dataset	Parameters	ADAM	Modified ADAM
1	Post Bank Customer Review	Original Statements	81.9	87.01
2		Statements having Only Synonyms	83.19	89.07
3		Statements having Only Antonyms	82.45	90.15
4		Original + Rephrased statements	87.15	92.15
5	Phrase Banking	Original Statements	82.79	87.98
6		Statements having Only Synonyms	83.54	90.74
7		Statements having Only Antonyms	86.17	89.83
8		Original + Rephrased statements	87.36	93.47

sentiments and aids in comprehending the spread of misinformation within the Banking domain. This concept offers valuable insights into the psychological and social factors that influence the dissemination and reception of misinformation, informing the design of marketing strategies.

In practical terms, the suggested approach for evaluating sentiment scores assists in discerning the sentiments of individual meaningful words. Using synonyms and antonyms enhances the system's efficiency, enabling it to identify sentiments across a broader spectrum of arguments. Additionally, the control distance approach aids in recognizing similarities between sentences and their rephrased counterparts.

D. Novelty and Scope

The study centers on extracting sentiments from banking text data, including feedback and media reports. Previous literature has covered various studies on sentiment analysis. This research is purpose-built for improved efficiency and reduced latency. The paper's novelty can be summarized as follows:

- Proposed a sentiment-scoring approach for individual words, synonyms, and antonyms. Assign sentiment scores (positive, negative, and neutral) to synonyms and antonyms.
- Devised a Control Distance approach to validate the similarity score between the original sentence and the synonyms and antonyms used in sentences.
- Subsequently, design a self-trained model for the ELECTRA model, incorporating words with sentiment scores from the designed data dictionary.
- The implementation assessed performance using the ADAM optimizer and the modified ADAM optimizer, incorporating the L2 regularizer.

TABLE XII. PERFORMANCE COMPARISON OF THE PROPOSED MODEL

Sl#	Model Implementation	Accuracy
1	FLANG˙ELECTRA (NER) [32]	82.0%
2	FLANG_ELECTRA (Post Bank customer dataset, NER)	82.9%
3	FinBERT (Phrase bank, CLS token) [33]	87.1%
4	ELECTRA (Phrase bank, CLS token)	83.87%
5	Proposed Model (Post Bank Customer Dataset)	92.27%
6	Proposed model (Phrase bank dataset)	93.48%

VII. CONCLUSION AND FUTURE SCOPE

The study presents a comprehensive sentiment analysis, identifying positive, negative, and neutral sentiments within customer feedback, financial, and economic texts. This analysis aids in combating misinformation in the market and informs marketing strategies. The article employed several methodologies, including text data preprocessing, sentiment identification for each word, synthesizing synonyms and antonyms for unique words, labeling individual sentences or paragraphs, and modeling. The dataset also labeled using sentiment scores for individual words and the control distance technique, providing meaningful sentiment analysis for individual sentences, thereby enhancing the overall implementation.

The implementation aims for a profound understanding of sentiments across a diverse dataset. A normalized threshold range of sentiment scores facilitates categorizing unique words into positive, negative, and neutral labels. Synonyms and antonyms are assigned similar sentiment scores, distinguished by positive and negative signs, respectively. The control distance approach verifies the sentiment scores of individual words in a sentence, evaluating positive, negative, and neutral statements. Subsequently, the ELECTRA model is self-trained using words with sentiment scores to classify sentiments. The implementation evaluates the model's output using both Adam and modified Adam optimizers for comparison.

The implemented approach demonstrates exceptional performance compared to FLANG_ELECTRA with NER model and the FinBERT model, using both Post banking customer data and Phrase banking data. The performance improvement compared to traditional models is substantial.

This article summarizes exploring the sentiments expressed in banking and financial-related data from customer feedback and financial news. The findings provide valuable insights for policymakers in the banking and financial sectors, aiding in developing strategies for customer feedback analysis, brand monitoring, product and service evaluation, fraud detection, market research, and competitor analysis. Additionally, the research holds practical significance by informing the design of automated tools for identifying sentiments within banking and financial text data.

Implementation of the proposed approach focuses on capturing sentiment scores within the documents. However, future work can broaden the scope. Currently, this implementation uses only three synonyms and antonyms. Exploring more synonyms and antonyms for word replacement opens avenues for further research.

Moreover, incorporating diverse deep learning, machine learning, and other models can enhance sentiment accuracy. Despite these options, this approach is the foundation for future investigations involving banking datasets. Researchers can also explore different banking-related datasets for in-depth analysis.

REFERENCES

[1] Mittal, Divya and Agrawal, Shiv Ratan, "Determining banking service attributes from online reviews: text mining and sentiment analysis," "International Journal of Bank Marketing," vol. 40, no. 3, pp. 558–577, 2022.

[2] Gavval, Rohit and Ravi, Vadlamani and Harshal, Kalavala Revanth and Gangwar, Akhilesh and Ravi, Kumar, "CUDASelfOrganizing feature map based visual sentiment analysis of bank customer complaints for Analytical CRM," "arXiv preprint arXiv:1905.09598," 2019, https://doi.org/10.48550/arXiv.1905.09598.

[3] Ahmed, Alim Al Ayub and Agarwal, Sugandha and Kurniawan, IMade Gede Ariestova and Anantadjaya, Samuel PD and Krishnan, Chitra, Akhilesh and Ravi, Kumar, "Business boosting through sentiment analysis using Artificial Intelligence approach," "International Journal of System Assurance Engineering and Management," vol. 13, no. Suppl 1, pp. 699–709, 2022, https://doi.org/10.1007/s13198-021-01594-x.

[4] Shapiro, Adam Hale and Sudhof, Moritz and Wilson, Daniel J, "Measuring news sentiment," "Journal of econometrics," vol. 228, no. 2, pp. 221–243, 2022, https://doi.org/10.1016/j.jeconom.2020.07.053.

[5] Mattia Atzeni, Amna Dridi, and Diego Reforgiato Recupero, "Fine-Grained Sentiment Analysis on Financial Microblogs and News Headlines," Semantic Web Challenges: 4th SemWebEval Challenge at ESWC 2017, Portoroz, Slovenia, May 28-June 1, 2017, pp. 124–128, doi: https://doi.org/10.1007/978-3-319-69146-6_11.

[6] Gandhi, Ankita and Adhvaryu, Kinjal and Poria, Soujanya and Cambria, Erik and Hussain, Amir, "Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions," "Information Fusion," vol. 91, pp. 424–444, 2023, https://doi.org/10.1016/j.inffus.2022.09.025.

[7] Muhammad Taimoor Khan, Muhammad Taimoor Khan, Mehr Durrani, Armughan Ali, Irum Inayat, Shehzad Khalid and Kamran Habib Khan, "Sentiment analysis and the complex natural language," Complex Adaptive Systems Modeling, vol. 4, no. 1, pp. 1–19, 2016, doi: https://doi.org/10.1186/s40294-016-0016-9.

[8] Deepshikha Chaturvedi and Shalu Chopra, "Customers Sentiment on Banks," International Journal of Computer Applications, vol. 98, no. 13, 2014.

[9] Yuliya Bidulya and Elena Brunova, "Sentiment Analysis for Bank Service Quality: a Rule-based Classifier," 2016 IEEE 10th International Conference on Application of Information and Communication Technologies (AICT), 2016, pp. 1–4, doi: 10.1109/ICAICT.2016.7991688.fr.

[10] Botchway, Raphael Kwaku and Jibril, Abdul Bashiru and Kwarteng, Michael Adu and Cho vancova, Miloslava and Oplatková, Zuzana Komínková, "A review of social media posts from UniCredit bank in Europe: a sentiment analysis approach," Proceedings of the 3rd international conference on business and information Management, pp. 74–79, 2019.

[11] Kazmaier, J and Van Vuuren, JH "Sentiment analysis of unstructured customer feedback for a retail bank," ORiON, vol. 36, no. 1, pp. 35–71, 2020.

[12] Varun Dogra, Aman Singh, Sahil Verma, Kavita N. Z. and M.N. Talib, "Analyzing DistilBERT for Sentiment Classification of Banking Financial News," Intelligent Computing and Innovation on Data Science: Proceedings of ICTIDS 2021, pp. 501–510, 2021, doi: https://doi.org/10.1007/978-981-16-3153-5_53.

[13] Bramanthyo Andrian, Tiarma Simanungkalit, Indra Budi and Alfan Farizki Wicaksono, "Sentiment Analysis on Customer Satisfaction of Digital Banking in Indonesia," International Journal of Advanced Computer Science and Applications, vol. 13, no. 3, 2022.

[14] Ogunleye, Bayode and Maswera, Tonderai and Hirsch, Laurence and Gaudoin, Jotham and Brunsdon, Teresa, "Comparison of topic modelling approaches in the banking context," Applied Sciences, vol. 13, no. 2, pp. 797, 2023.

[15] Surabhi, Adhikari, Surendrabikram Thapa, Usman Naseem, Hai Ya Lu, Nana Bharathy and Mukesh Prasad, "Explainable hybrid word representations for sentiment analysis of financial news," Neural Networks, vol. 164, pp. 115–123, 2023, doi: https://doi.org/10.1016/j.neunet.2023.04.011.

[16] Siwei Lai, Liheng Xu, Kang Liu and Jun Zhao, "Recurrent Convolutional Neural Networks for Text Classification," Proceedings of the AAAI conference on artificial intelligence, vol. 29, no. 1, 2015.

[17] Baojun Ma, Hua Yuan and Ye Wu, "Exploring performance of clustering methods on document sentiment analysis," Journal of Information Science, vol. 43, no. 1, pp 54–74, 2017, doi: https://doi.org/10.1177/0165551515617374.

[18] Mohammed Attia, Younes Samih, Ali Elkahky and Laura Kallmeyer†, "Multilingual Multi-class Sentiment Classification Using Convolutional Neural Networks," Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), 2018.

[19] Saurav Pradha, Malka N. Halgamuge and Nguyen Tran Quoc Vinh, "Effective Text Data Preprocessing Technique for Sentiment Analysis in Social Media Data," 2019 11th international conference on knowledge and systems engineering (KSE), pp. 1–8, 2019, doi: 10.1109/KSE.2019.8919368.

[20] Zabit Hameed and Begonya Garcia-Zapirain, "Sentiment Classification Using a Single-Layered BiLSTM Model," Ieee Access, vol. 8, pp. 73992–74001, 2020, doi: 10.1109/ACCESS.2020.2988550.

[21] Nicola Capuano, Lica Greco, Pierluigi Ritrovato and Mario Vento, "Sentiment analysis for customer relationship management: an incremental learning approach," Applied Intelligence, vol. 51, pp. 3339–3352, 2021, doi: https://doi.org/10.1007/s10489-020-01984-x.

[22] Amira Samy Talaat, "Sentiment analysis classification system using hybrid BERT models," Journal of Big Data, vol. 10, no. 1, pp. 1–18, 2023, doi: https://doi.org/10.1186/s40537-023-00781-w.

[23] Andrei Plotnikov, Alexey Shcheludyakov, Vadim Cherdantsev, Alexey Bochkarev and Igor Zagoruiko, "Data on post bank customer reviews from web," Data in Brief, vol. 32, pp. 106152, 2020, doi: https://doi.org/10.1016/j.dib.2020.106152.

[24] Malo, Pekka and Sinha, Ankur and Korhonen, Pekka and Wallenius, Jyrki and Takala, Pyry, "Good debt or bad debt: Detecting semantic orientations in economic texts," Journal of the Association for Information Science and Technology, vol. 65, no. 4, pp. 782–796, 2014, doi: https://doi.org/10.1002/asi.23062.

[25] Javed, Muhammad and Kamal, Shahid, "Normalization of unstructured and informal text in sentiment analysis," "International Journal of Advanced Computer Science and Applications," vol. 9, no. 10, 2018, http://dx.doi.org/10.14569/IJACSA.2018.091011.

[26] Atoum, Issa and Otoom, Ahmed, "Efficient hybrid semantic text similarity using WordNet and a corpus," "International Journal of Advanced Computer Science and Applications," vol. 7, no. 9, 2016, http://dx.doi.org/10.14569/IJACSA.2016.070917.

[27] Mukhiya, Suresh Kumar and Ahmed, Usman and Rabbi, Fazle and Pun, Ka I and Lamo, Yngve, "Adaptation of IDPT system based on patient-authored text data using NLP.," 2020 IEEE 33rd international symposium on computer-based medical systems (CBMS), pp. 226–232, 2020, doi: 10.1109/CBMS49503.2020.00050.

[28] Clark, Kevin and Luong, Minh-Thang and Le, Quoc V and Manning, Christopher D, "Electra: Pre-training text encoders as discriminators rather than generators," arXiv preprint arXiv:2003.10555, 2020, doi: https://doi.org/10.48550/arXiv.2003.10555.

[29] Hasan, Mahmud and Islam, Labiba and Jahan, Ismat and Meem, Sabrina Mannan and Rahman, Rashedur M, "Natural Language Processing and Sentiment Analysis on Bangla Social Media Comments on Russia–Ukraine War Using Transformers," Vietnam Journal of Computer Science, pp. 1–28, 2023, doi: https://doi.org/10.1142/S2196888823500021.

[30] Kingma, Diederik P and Ba, Jimmy, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014, doi: https://doi.org/10.48550/arXiv.1412.6980.

[31] Loshchilov, Ilya and Hutter, Frank, "Fixing weight decay regularization in adam," 2018.

[32] Shah, Raj Sanjay and Chawla, Kunal and Eidnani, Dheeraj and Shah, Agam and Du, Wendi and Chava, Sudheer and Raman, Natraj and Smiley, Charese and Chen, Jiaao and Yang, Diyi, "When flue meets flang: Benchmarks and large pre-trained language model for financial domain," "arXiv preprint arXiv:2211.00083" 2022.

[33] Leonardo Colacicchi, Comparison and fine- tuning of methods for Financial Sentiment Analysis. Department of Data Science and Artificial Intelligence, Maastricht University, 2022.