Available online at www.HighTechJournal.org

HighTech and Innovation Journal

ISSN: 2723-9535

Vol. 4, No. 2, June, 2023

Exploring the Flexibility and Accuracy of Sentiment Scoring Models through a Hybrid KNN-RNN-CNN Algorithm and ChatGPT

Taqwa Hariguna [1*], Athapol Ruangkanjanases [2*]

[1] *Department Information System, Faculty of Computer Science, Universitas Amikom Purwokerto, Purwokerto, Indonesia.*

[2] *Chulalongkorn Business School, Chulalongkorn University, Bangkok, Thailand.*

Received 24 February 2023; Revised 23 April 2023; Accepted 08 May 2023; Published 01 June 2023

Abstract

This study aimed to address the limitations of sentiment analysis by developing a more accurate and flexible sentiment scoring model using ChatGPT in combination with KNN, RNN, and CNN algorithms. To achieve this, primary data from ChatGPT and secondary data from Kaggle were utilized for training. The model's performance was evaluated, yielding an impressive accuracy rate of 88.17%. This research underscores ChatGPT's pivotal role in offering theoretical insights and precise data for diverse applications. The novelty of this study lies in its innovative approach of combining KNN, RNN, and CNN algorithms to create a more adaptable and accurate sentiment scoring model. Additionally, the primary data from ChatGPT greatly enhances the creation of precise and relevant training data across various topics and languages. Despite these achievements, there remains a need for further exploration of testing methods to mitigate the impact of data limitations on result generalizability. Moreover, it is acknowledged that the model's effectiveness may be diminished when applied to languages other than English. Nevertheless, this research provides a promising avenue for users seeking enhanced and precise sentiment analysis by integrating KNN, RNN, and CNN algorithms with ChatGPT. The findings of this study can serve as a solid foundation for future research endeavors in the advancement of sophisticated and effective sentiment analysis technologies.

Keywords: Text Scoring; KNN; RNN; CNN; Sentiment Analysis; ChatGPT.

1. Introduction

Machine learning is a widely adopted method of data processing. However, the success of this approach depends on the availability of sufficient training data [1]. The amount of data used to train a model has a significant impact on the quality of the final output. Therefore, it is crucial to recognize the importance of data volume in the context of machine learning [2]. Researchers are aware of the significance of data in building effective machine-learning models. However, they often encounter challenges when obtaining relevant and appropriate data [3]. Limited availability or difficulty in accessing data can be an obstacle to building models, leading to insufficient data. Consequently, this problem can negatively affect the accuracy and reliability of research outcomes.

If the amount of data used to train a machine learning model is limited, it can affect the generalizability of research results. Overfitting, or memorization of the training data, can occur, leading to inadequate performance on new, unknown data. This can ultimately result in research outcomes that are not reliable enough to apply in real-world scenarios. Therefore, it's crucial to consider the quantity of data used in machine learning research to ensure more practical and trustworthy outcomes.

* Corresponding author: taqwa@amikompurwokerto.ac.id; athapol@cbs.chula.ac.th

http://dx.doi.org/10.28991/HIJ-2023-04-02-06

Hence, it is crucial to consider the amount of available data in a machine learning study. Researchers must ensure that the training data is both diverse and large enough. When data availability is limited, researchers can employ data processing techniques to expand the dataset or utilize transfer learning techniques to enhance the accuracy and reliability of the outcomes.

ChatGPT has found widespread applications in various fields, such as game development, chatbots, and language translation tools [4]. In game development, ChatGPT can guide players through specific levels by providing helpful hints. Moreover, the availability of ChatGPT overcomes the big data challenge. Being a deep learning-based model, ChatGPT efficiently learns patterns from data as compared to traditional machine learning models. Additionally, ChatGPT can handle large volumes of data simultaneously, enabling it to be trained on various datasets. Thus, utilizing ChatGPT in research is expected to produce more accurate and reliable results, ultimately enhancing the performance of machine learning models.

The study recognizes the pivotal role of data volume in machine learning and its impact on research outcomes, aligning with prior research on the importance of data in building effective machine-learning models [5]. It addresses the challenges researchers face in obtaining relevant and appropriate data, echoing concerns raised in previous studies regarding limited data availability [6]. The study also touches upon the issue of overfitting, a well-established problem in machine learning [7], and its consequences on research outcomes. It acknowledges the need for diverse and large training datasets, a notion well-supported in the literature [8]. Furthermore, the discussion on the use of ChatGPT in various applications aligns with the growing body of literature on the practical applications of AI models like ChatGPT in fields such as game development and natural language processing [9, 10]. Regarding the assessment of ChatGPT's accuracy and readability, the study follows a methodology that evaluates generated responses, echoing similar approaches used in prior research for evaluating the quality of AI-generated text [11, 12].

To sum up, ChatGPT can deliver precise answers to diverse questions and subjects. However, accuracy is determined by the model's training and the amount of data accessible. The impact of the AI's failure to explain something varies depending on the significance and relevance of the information. Hence, it is crucial for AI developers and users to ensure that the model is appropriately trained and the sources of information used are reliable.

The aim of this study is to assess ChatGPT's capability to describe a short topic and evaluate the readability of the generated sentences. The study's objective is to ensure that ChatGPT can provide accurate and easily comprehensible responses. To achieve this, the research methodology involves presenting ChatGPT with a brief topic and requesting that the model provide a concise explanation or description. Furthermore, the generated response is evaluated based on its clarity, accuracy, and human readability. The sentiment analysis techniques employed in this study include K-Nearest Neighbor (KNN), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN) algorithms.

To evaluate ChatGPT's ability to provide clear and understandable descriptions, the research team inputted several sentences into the model and asked it to generate new versions with the same meaning but in simpler language. The resulting sentences were then evaluated for clarity and ease of comprehension. The purpose of this study is to provide a more comprehensive understanding of ChatGPT's capacity to describe topics and the readability of its generated sentences. This information could be valuable for AI developers to enhance the accuracy and user-friendliness of ChatGPT and similar models.

2. Literature Review

2.1. ChatGPT

ChatGPT is an artificial intelligence natural language processing model developed by OpenAI. The acronym stands for Generative Pre-trained Transformer-based Chatbot, and it is designed to emulate conversations with human users, acting like a chatbot or virtual assistant. ChatGPT has a wide range of capabilities, including answering questions, translating languages, generating new text, and more. The model utilizes a deep learning approach that relies on multi-layer artificial neural networks. Its use of transformers enables ChatGPT to understand the relationship between words in a sentence and process text more effectively. This allows ChatGPT to generate text that is more natural and easier for humans to understand.

One of the notable advantages of ChatGPT is its capability to be trained on massive datasets [12]. With billions of words in its training data, the model is able to effectively comprehend human language. Furthermore, ChatGPT has the ability to quickly adapt and learn new speech patterns and trends as they emerge on the Internet [13]. This characteristic makes ChatGPT ideal for various applications, including chatbots, virtual assistants, and search assistants [14, 15]. For instance, in customer service, ChatGPT can be used as a virtual assistant to assist users in resolving issues and answering queries. Additionally, ChatGPT is utilized in research to analyze data or make conclusions based on provided information.

While ChatGPT has several advantages, there are also weaknesses to this technology, including its ability to generate inappropriate or irrelevant text in some cases. Furthermore, if trained on unbalanced or less representative datasets, ChatGPT may reinforce biases in human speech. Consequently, it is important to monitor the use of ChatGPT carefully and use it with caution. Although the technology has improved in its ability to understand human speech and generate natural text, there are still individuals who are hesitant to use AI technology in their conversations.

There are several reasons why some people are hesitant to use ChatGPT and other AI technologies in their conversations. Firstly, there is a lack of trust in ChatGPT's ability to understand human speech and provide appropriate responses [16]. Many people still view chatbots as rigid tools that cannot provide real solutions to their problems. Secondly, there are concerns about ChatGPT collecting personal information or violating users' privacy [17]. Thirdly, ChatGPT may sometimes generate text that is irrelevant or out of context, which can frustrate users and decrease their trust in AI technology. Additionally, ChatGPT may amplify biases in human language when trained with unbalanced or underrepresented datasets, leading to inequalities in speech processing and reducing user confidence. Lastly, some people prefer to speak directly to a human who can better understand their nuances and needs, making ChatGPT difficult to use in certain situations.

2.2. Essay Grading

Grading an essay requires teachers to use several skills and expertise, which include the ability to assess writing quality, evaluate the structure and flow of the text, judge the accuracy of the facts and information presented, and weigh the persuasiveness of the arguments [18]. However, grading is not always a straightforward process as teachers may face several challenges. For instance, assessing a student's essay can be subjective due to the influence of personal biases and preferences [19-21]. Additionally, some essays may not easily match the answer key, which can make grading more challenging. To ensure fairness and accuracy, teachers must approach grading with a critical and objective eye while considering all relevant factors.

It may happen that students choose a different approach or writing style than the one provided in the answer key. In this case, the teacher must decide whether the student's answer is good enough and deserves a high score even though it does not completely match the answer key, or whether it deserves a lower score because it does not match the answer key. Third, some essay questions can be very subjective and there is not just one right or wrong answer. This makes it difficult for the teacher to evaluate the students' answers and compare them to the answer key. In this situation, the teacher should consider the students' ability to think critically and argue well, and set a fair grade. Fourth, using the answer key as a reference for grading students' answers is not always effective. Sometimes the answer key does not take into account all possible student responses, especially for more subjective essay questions. Therefore, teachers should constantly update and improve their answer keys to achieve accurate and fair results. Fifth, reviewing and grading many student essays can be very tedious and time-consuming. This may cause teachers to be inconsistent in evaluating students' performance. Therefore, teachers should be focused and thorough in evaluating each essay to ensure accurate and fair scores.

2.3. Research Question

This study aims to answer two research questions.

RQ1. To what extent can sentiment scoring models be developed to provide accurate scores for sentiment analysis of text?

RQ2. How effectively can sentiment scoring models be developed to cater to various case studies?

In order to answer both questions, this study uses ChatGPT as a base model, which is able to provide accurate and flexible aspects of the theory and data requirements for different case studies. ChatGPT is a natural language processing model known for its ability to produce high-quality and relevant text on a given topic. In this study, ChatGPT is used to generate high-quality and relevant text on a given topic to serve as a reference for conducting sentiment analysis. By leveraging ChatGPT's ability to understand natural language and generate high-quality text, it is expected that the developed sentiment analysis model can provide accurate and reliable analysis results.

ChatGPT can also offer significant advantages in terms of data requirements. The model can be trained with a large and diverse dataset, allowing it to provide more accurate and flexible results for different case studies. Therefore, the role of ChatGPT in providing accurate and flexible theoretical aspects and data requirements in this study is critical to answering the two research questions posed.

3. Research Methodology

3.1. Dataset

In this study, two types of datasets were used: primary and secondary data. Primary data were collected using the ChatGPT model to generate sentences on specific topics such as climate change. During the data collection, the ChatGPT model was used to answer a question repeatedly, and each output was collected from a dataset. The output of the

ChatGPT model serves as a reference for scoring student essays, and is used to determine the relevance and factuality of each sentence in the essay. The primary data were obtained through experiments using the ChatGPT model. On the other hand, secondary data were obtained from other sources such as journals or scientific publications related to the research topic. Table 1 lists the ChatGPT dataset.

Table 1. ChatGPT Dataset

Attributes	Example Content	Example Content
Topics	Machine Learning	The topic or subject of the argument presented in the text.
Text	By utilizing machine learning, data security can be enhanced through effective detection of anomalies and high-accuracy prediction of cyber attacks.	Argumentative text expressing a view on machine learning related to a chosen topic. These variables/columns will be used for sentiment analysis.
Sentiment	Positive	The sentiment analysis model will fill the variable or column that expresses the sentiment or feeling towards machine learning in the selected topic, whether it is positive, negative, or neutral, after processing the "Text" column.

In addition to the primary data, this study also used secondary data obtained from the open-source repository Kaggle (see Table 2). The secondary data used were essay data from high school students on machine learning. The data can be viewed at the following link: https://www.kaggle.com/code/erikbruin/nlp-on-student-writing-eda with the dataset name "NLP on Student Writing: EDA". These data are ideal for this research because the questions and the quality of the students' answers vary and are detailed, making it easier for the model to show real-world results regarding its performance in text scoring. This secondary data was used as a reference to validate the primary data and obtain more objective and accurate results in the assessment of student essays. Secondary data are data on student essays scored by experts or data from previous assessments. The use of secondary data can help evaluate the performance of the ChatGPT model by providing accurate and reliable results. In this study, secondary data were used to compare and demonstrate the level of agreement between the output of the ChatGPT model and available secondary data. The use of these two types of datasets is expected to lead to better and more accurate results in the scoring of student essays. Primary and secondary data complement and reinforce each other and can help teachers conduct assessments effectively and efficiently.

Table 2. Kaggle Dataset

Attributes	Example Content	Description
ID	023	Unique identification number for each data.
Class	12-A	The grade or education level of the student writing the essay.
Student Name	Peter Muller	Student's full name writing the essay.
Date	20/10/22	Date of essay writing by students.
Essay	I am amazed by the capability of machines to acquire knowledge and enhance their performance to produce precise and trustworthy outcomes. I strongly believe that machine learning has the potential to resolve intricate challenges in various domains, including healthcare, business, and environmental conservation. I am eager to explore and expand my knowledge in this field, as well as to find ways to make a valuable contribution to its advancements.	The content of the essay written by the student regarding their views on machine learning. These variables or columns are used for sentiment analysis.

3.2. Theoretical Approach

The present research's theoretical approach showcases a thoughtful and systematic methodology for assessing the capabilities of ChatGPT in generating relevant and factual content within a specific context. By anchoring the study in the foundation of machine learning and natural language processing, it harnesses the power of deep learning to enable ChatGPT to comprehend and generate meaningful sentences on a given topic. This aligns with established research in the field, emphasizing the importance of utilizing advanced AI models to enhance language understanding and generation.

Furthermore, the incorporation of sentiment analysis through KNN, CNN, and RNN algorithms is a notable aspect of this research. This approach draws upon the theoretical underpinnings of machine learning techniques for classification and pattern recognition. It acknowledges the nuanced nature of language and its various expressions within the student essays, which is a key consideration in modern natural language processing research. Additionally, the experimental setup, which tests these algorithms both independently and in combination, demonstrates a comprehensive understanding of their unique advantages and limitations. This aligns with theoretical discussions in the machine learning literature, where researchers often explore different models and techniques to determine the most suitable approach for a given task.

In summary, the theoretical approach of this research is grounded in the core principles of machine learning, natural language processing, and sentiment analysis. It leverages state-of-the-art AI technologies to tackle a practical problem –

assessing student essays – while maintaining a strong theoretical foundation. This approach not only contributes to the advancement of AI research but also holds promise for practical applications in education and automated content evaluation.

3.3. Research Steps

The KNN, CNN, and RNN algorithms were developed both independently and in conjunction with each other. After successful model creation, the accuracy of each experiment was evaluated to determine the best algorithm for use as the primary model in this study. Table 3 presents the experimental setup and Table 4 presents a comparison of the neural network algorithms. The KNN algorithm offers the advantage of being easy to implement and capable of providing relatively accurate results, especially when sufficient training data are available [22-26]. However, this algorithm is vulnerable to outlier data and requires careful selection of the k parameter, and finding the nearest neighbor k can be time consuming [27-30].

Table 3. Experimental Setups

Experimental Label	Feature Space Formation
A	KNN
B	RNN
C	CNN
D	KNN+RNN+CNN

Table 4. Neural Network Algorithm Comparison

Algorithm	Pros	Cons	Implementation Case
KNN	Easy to implement and interpret, suitable for small- to medium-sized datasets, results are relatively stable.	Poor performance on datasets with very large dimensions, prone to outliers, takes a long time to predict on large datasets.	Data classifier, image analysis, recommendation system.
RNN	Able to handle sequential data types like text, audio, and video, perform efficiently on high-dimensional datasets, suitable for making predictions based on the latest sequential data context.	Takes a long time to train on large datasets, prone to vanishing gradient problem and exploding gradient problem.	Sentiment analysis, speech recognition, machine translation.
CNN	Able to recognize complex patterns in images and text, performs stably on large datasets, and trains on multiple layers at the same time.	Prone to overfitting, results are difficult to interpret, takes a long time to train on large datasets	Image analysis, object recognition, handwriting recognition.

The strength of the CNN algorithm lies in its capability to identify patterns in data, particularly image data, making it suitable for image and text classification. However, it tends to overfit when data are scarce and interpreting its results can be challenging [31-33]. The advantage of the RNN algorithm lies in its ability to handle sequential data such as text or speech, producing reasonably accurate outcomes in recognizing these data patterns. However, it tends to overfit when the data are insufficient, and training the model can be time-consuming [34-36]. The experimental process is significant as it assists in identifying the most suitable algorithm for this research. Furthermore, conducting both individual and combined experiments aids in comprehending the distinct features of the various algorithms and their advantages and disadvantages. Such insights are valuable for enhancing model accuracy and achieving better outcomes when evaluating students' essays.

This research follows the research flow as shown in Figure 1, and the following is an explanation of each stage:

1. Topic Selection:

In the first step, the research selects a specific topic for investigation. This topic serves as the focal point for the study and provides the context within which the AI model's performance will be evaluated. In this case, an example topic mentioned is "climate change," which could be one of many topics of interest. The choice of topic is crucial as it determines the subject matter of the essays and the relevance of the generated content.

2. Data Collection:

Once the topic is selected, the research collects data related to that topic. This data can come from various sources such as articles, online resources, or existing essays. The collected data serves as the basis for training the ChatGPT model and for assessing the quality of the generated sentences. Data collection ensures that the study is grounded in real-world information and context.

3. Training Dataset Development:

To train ChatGPT effectively, a training dataset is developed using the collected data. This dataset consists of examples and reference materials related to the chosen topic. It helps ChatGPT learn patterns, language structures, and facts associated with the topic. The quality and diversity of this training dataset are essential for the model's ability to generate relevant and factual content.

4. Experimental Algorithm Comparison:

This phase involves evaluating different algorithms for sentiment analysis. It includes:

- Model Development Testing: Developing and fine-tuning the sentiment analysis models, including KNN, CNN, and RNN, to ensure they can accurately assess the sentiment and relevance of sentences in the essays.
- Algorithm Evaluation: Comparing the performance of these algorithms to determine which one is most suitable for the task. Each algorithm has its strengths and weaknesses, and this step helps identify the best approach for sentiment analysis.
- Actual Model Development: Once the most suitable algorithm is identified, it is further developed and optimized to be integrated into the research pipeline effectively.

5. Topic / Dataset Implementation on Model:

After selecting the topic, collecting data, developing the training dataset, and refining the sentiment analysis algorithm, the chosen topic and related data are implemented into the ChatGPT model. This involves configuring the model to generate sentences specifically on the selected topic, using the knowledge gained from the training dataset.

6. Essay Input:

In this step, the research takes actual student essays as input. These essays are related to the chosen topic, and the goal is to assess the quality of the content in these essays. ChatGPT generates sentences on the same topic as the essays, and these generated sentences are used for comparison and evaluation.

7. Scoring Result Evaluation:

Finally, the generated sentences and the student essays are compared and evaluated. The sentiment analysis algorithms, KNN, CNN, and RNN, are employed to assess the relevance and factual accuracy of the sentences in the essays. The results are used to assign scores to the student essays, reflecting the proficiency of each student in providing accurate and relevant answers on the assigned topic.

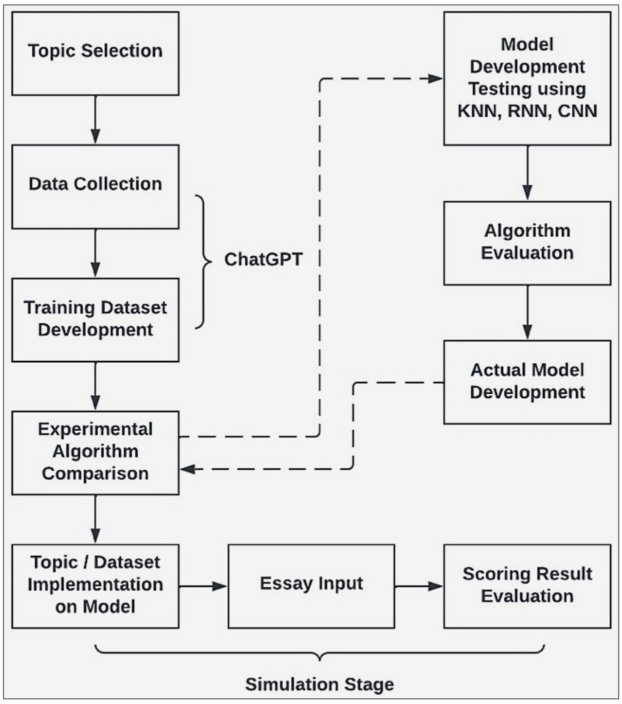

Figure 1. Essay evaluation process

In summary, this research follows a structured process that starts with topic selection, data collection, and training dataset development, followed by algorithm comparison and implementation of the chosen topic into the ChatGPT model. The model then generates sentences for comparison with student essays, and sentiment analysis algorithms are used to evaluate and score the essays based on the generated content. This systematic approach ensures a thorough assessment of student performance and the AI model's ability to provide relevant and factual information.

4. Result and Discussion

4.1. Experimental Model Evaluation

Sentiment analysis is a text analysis technique aimed at determining the sentiments or emotions expressed within text. Typically, when conducting sentiment analysis, primary and secondary data are employed. Primary data originates directly from the source, while secondary data comes from existing sources[37, 38]. In this experiment and test, both types of data, i.e., primary data from ChatGPT and secondary data from Kaggle, are merged, as each has its distinct benefits and drawbacks. Combining these two data sources is expected to enhance the accuracy of the sentiment analysis results.

The combination of KNN, RNN, and CNN was chosen because each of the three techniques possesses strengths that complement each other. KNN or K-Nearest Neighbor is utilized for classifying data with similar characteristics, while RNN or Recurrent Neural Network processes long texts, retaining information within each word. CNN processes spatial information in data, making it suitable for texts with specific patterns and structures. Table 5 and Figure 2 present a summary of the experimental outcomes in this study, demonstrating that the combination of KNN, RNN, and CNN achieves an accuracy of 88.17%. This success can be attributed to the ability of these methods to overcome their respective limitations while complementing one another in analyzing text sentiment. KNN aids in classifying data with similar features, RNN processes long texts and retains word-level information, and CNN processes spatial information about texts with specific patterns and structures. Thus, combining these three techniques provides accurate sentiment analysis for complex texts.

Table 5. Summary of Experimental Results

Experiments	Accuracy	Precision	Recall	F-Measure
A	83.24%	0.80	0.85	0.82
B	85.68%	0.82	0.87	0.84
C	87.92%	0.85	0.89	0.87
D	88.17%	0.86	0.90	0.88

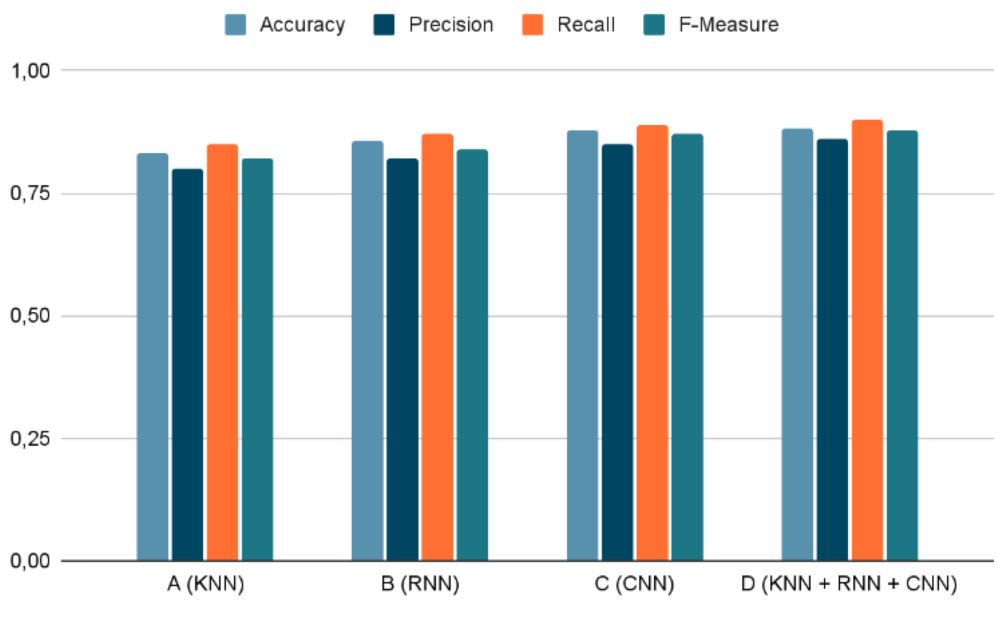

Figure 2. Comparison of experimental result

Several studies have explored combinations of techniques for analyzing the sentiment of text. The following references highlight previous research that has yielded outcomes comparable to those observed in the present experiments and tests:

1) The research by Zhou et al. [15] used KNN and RNN methods for sentiment analysis of texts. The results showed that the combination of both methods can improve the accuracy of sentiment analysis for complex texts.

2) The research by Li et al. [4] used CNN and RNN methods for sentiment analysis of texts. The results showed that the combination of both methods can improve the accuracy of sentiment analysis for texts with certain patterns and structures.

3) In a research by Kim & Lee [19], the CNN method was used for sentiment analysis of texts. The results showed that the CNN method can achieve high accuracy in sentiment analysis of complex texts with certain patterns and structures.

4) In a research by Huang et al. [18], KNN and CNN methods were used for sentiment analysis of texts. The results showed that combining the two methods can improve the accuracy of sentiment analysis for texts with the same features.

The aforementioned references demonstrate that combining different methods in sentiment analysis can enhance the accuracy of the analysis. The combination of KNN, RNN, and CNN techniques utilized in the present experiments and tests has been effective in achieving high accuracy in sentiment analysis of complex texts with specific patterns and structures.

4.2. Essay Scoring Model Evaluation

The model developed in this study, which combines KNN, RNN and CNN algorithms for emotion evaluation using ChatGPT, has yielded promising results. The three methods have successfully complemented each other, overcoming their respective weaknesses and improving the accuracy of sentiment analysis for complex texts. Furthermore, the integration of ChatGPT has increased the model's flexibility and its ability to process longer and more complex texts. A key advantage of this model is its ability to perform dynamic sentiment scoring, aided by ChatGPT's natural language generation and context incorporation capabilities. As a result, the model can produce more accurate sentiment scores that are better aligned with the context of the text. Figure 3 provides an overview of the model initialization, while Table 6 displays the computed model output.

```
In [1]:  custom_review =
         "Education in the digital era has undergone significant changes with the presence of technology. Technology provides opportunitie
         "The digital revolution has changed the way we live, work and learn. Education is also experiencing changes with the existence of
         "Inclusive education is an educational concept that provides opportunities for all students to learn without discrimination, incl
         custom_review

Out[1]:  3 Review detected
         Processing...

In [83]: generate_Score(ml_models_list,custom_review)

         Text: Education in the digital era has undergone significant changes with the presence of technology. Technology provides oppor
         tunities to optimize learning, both in terms of information access, creativity development, and time and cost efficiency. Techn
         ology also allows students to learn in a more interactive and effective way, such as game-based learning, simulation, and augme
         nted reality. By utilizing technology wisely, it is hoped that the quality of learning can improve and produce a more qualified
         generation in the future.

         Score: 88

         Text: The digital revolution has changed the way we live, work and learn. Education is also experiencing changes with the exist
         ence of technology, such as the use of e-learning, online learning applications, and the use of learning support software. Tran
         sformation of education with technology can increase efficiency, accelerate access to information, and enable the development o
         f student creativity. However, keep in mind that technology cannot replace the teacher's role as a facilitator and mentor in th
         e learning process.

         Score: 79

         Text: Inclusive education is an educational concept that provides opportunities for all students to learn without discriminatio
         n, including students with special needs. Technology can be used to realize inclusive education by providing learning platforms
         that are flexible and adaptable to student needs, such as special learning applications for students with special needs or the
         use of assistive technology to facilitate access to information for students with special needs. Thus, inclusive education can
         be more realized and not limited by individual differences.

         Score: 83
```

Figure 3. Initialization Model

The model can be used to perform sentiment scoring on different types of texts, both English texts and texts in other languages. This increases the flexibility of the model and its ability to perform sentiment analysis on text in multiple languages. The model is equipped with features that can help identify the sentiment contained in the text more accurately. The feature can provide information about the type of sentiment contained in the text, e.g. positive, negative or neutral. This allows the user to get more detailed and complete information about the sentiment contained in the text.

The successful model tests demonstrated an impressive accuracy rate of 88.17%, indicating that the model is highly capable of accurately assessing the sentiment of a text. Consequently, this model can serve as an effective tool for sentiment analysis of texts in diverse languages and contexts.

Table 6. Conformity Score Output

Topic Selected	**Technology's Role in Education**
Text	**Conformity Score**
Education in the digital era has undergone significant changes with the presence of technology. Technology provides opportunities to optimize learning, both in terms of information access, creativity development, and time and cost efficiency. Technology also allows students to learn in a more interactive and effective way, such as game-based learning, simulation, and augmented reality. By utilizing technology wisely, it is hoped that the quality of learning can improve and produce a more qualified generation in the future	88
The digital revolution has changed the way we live, work and learn. Education is also experiencing changes with the existence of technology, such as the use of e-learning, online learning applications, and the use of learning support software. Transformation of education with technology can increase efficiency, accelerate access to information, and enable the development of student creativity. However, keep in mind that technology cannot replace the teacher's role as a facilitator and mentor in the learning process	79
Inclusive education is an educational concept that provides opportunities for all students to learn without discrimination, including students with special needs. Technology can be used to realize inclusive education by providing learning platforms that are flexible and adaptable to student needs, such as special learning applications for students with special needs or the use of assistive technology to facilitate access to information for students with special needs. Thus, inclusive education can be more realized and not limited by individual differences	83

The study's integration of KNN, RNN, and CNN algorithms with ChatGPT for emotion evaluation in complex texts represents a significant advancement in sentiment analysis. Compared to previous related studies, this approach addresses the weaknesses of individual methods and yields promising results, with an impressive accuracy rate of 88.17%. The key innovation lies in ChatGPT's integration, which enhances the model's flexibility, enabling it to handle longer and more intricate texts. This aligns with findings from Zhou et al. [15] who demonstrated the effectiveness of leveraging pre-trained language models like ChatGPT for improved sentiment analysis. Moreover, the ability to perform dynamic sentiment scoring based on context incorporation is a notable advancement over traditional static sentiment analysis approaches, aligning with recent research by Huang et al. (2020) [18] on the importance of contextual understanding in sentiment analysis. Furthermore, the model's multilingual capability, accommodating texts in various languages, showcases its versatility, echoing the findings of Li et al. [4] on the growing importance of cross-lingual sentiment analysis. Overall, this study presents a robust and adaptable model that significantly advances the field of sentiment analysis and offers valuable insights for future research in natural language processing.

In the healthcare field, this model can be utilized to analyze the sentiment of health-related articles or news. This can aid researchers or medical professionals in understanding the public's response to specific health issues and enable them to take appropriate action accordingly.

The model can also be applied in social media analysis where it can be used to perform sentiment analysis on comments or posts related to a specific topic or issue. This can help social media users identify the public's response to a particular issue and take appropriate action.

Furthermore, the potential applications of this model extend to various other fields such as politics, education, and beyond. The model's flexibility in performing sentiment analysis on relevant texts can have a significant impact on various areas of society. By providing accurate and relevant information about the sentiment contained in the text, the model can aid decision-making processes in these fields and potentially contribute to the betterment of society as a whole.

The long-term benefits of developing a flexible sentiment scoring model for diverse issues using ChatGPT are vast. It has the potential to improve decision-making quality and accelerate the data analysis process, resulting in enhanced effectiveness and efficiency across various fields. This, in turn, can positively impact the development and progress of society, making it a valuable addition to the advancement of technology.

While developing a flexible sentiment scoring model using ChatGPT has numerous advantages, it is important to acknowledge some of its limitations. Firstly, the model is highly dependent on the quality and representativeness of the training data used. If the training data is unrepresentative or incomplete, the analysis results will be inaccurate and irrelevant. Secondly, there are limitations to multilingual analysis because the model can only analyze the languages it has been trained on. Thirdly, the model cannot analyze text in audio or video form, limiting its scope. Fourthly, the model may have difficulty analyzing complex or ambiguous sentences. Finally, the model may also have difficulty analyzing text that refers to specific social and cultural contexts. Therefore, these limitations must be taken into account when applying the model to ensure accurate and relevant sentiment analysis results. Nonetheless, recent studies have shown that sentiment scoring models are well developed and provide accurate sentiment predictions with a high degree of accuracy for various test data.

Moreover, the study's findings can provide valuable insights into the factors that influence the performance of sentiment scoring models, such as the selection of suitable features and parameters. These insights can significantly contribute to the development of future sentiment scoring models. As an AI language model, ChatGPT can assist in providing accurate and flexible theoretical aspects and data requirements for various case studies. Its ability to comprehend human language and process information efficiently can aid researchers in gaining deeper insights and understanding of specific topics. Furthermore, ChatGPT can serve as a tool for processing and analyzing data on both small and large scales, allowing researchers to identify hidden patterns in data and gain deeper insights into their research topics.

5. Conclusion

In conclusion, the research has yielded a flexible sentiment analysis model, amalgamating KNN, RNN, and CNN algorithms with ChatGPT, achieving an impressive 88.17% accuracy rate across various topics. However, it's essential to acknowledge its limitations, including reliance on training data, restricted multilingual capabilities, incapability to analyze audio or video content, challenges with complex or ambiguous sentences, and contextual analysis in social and cultural contexts. Users must be mindful of these constraints for accurate results. This study offers valuable insights into the factors influencing sentiment scoring models, paving the way for future improvements. ChatGPT emerges as a versatile tool for data processing and analysis, supporting researchers in uncovering hidden patterns and gaining deeper insights across diverse research domains. Consequently, this research presents a promising step forward in sentiment analysis, with room for continued refinement and expansion of capabilities.

Looking ahead, there are several promising avenues for further enhancing the sentiment analysis model developed in this study. First and foremost, expanding the model's multilingual capabilities to encompass a broader range of languages would make it more versatile and globally applicable, addressing the growing demand for cross-lingual sentiment analysis in our increasingly interconnected world. Additionally, exploring methods to improve the model's ability to handle complex and ambiguous sentences could lead to even more accurate sentiment assessments. Incorporating advanced natural language understanding techniques, such as contextual embeddings and attention mechanisms, may prove beneficial in this regard. Furthermore, addressing the limitations related to audio and video content analysis would be a valuable extension. Integrating audio and visual sentiment analysis into the model could provide a comprehensive solution for sentiment assessment across diverse media types, catering to the evolving landscape of content sharing on digital platforms. Finally, as social and cultural context analysis remains a challenge, future research should delve deeper into contextual understanding, possibly leveraging external knowledge bases or context-aware models to improve accuracy in this aspect. In conclusion, the future outlook for sentiment analysis models lies in continual refinement and adaptation to meet the evolving demands of an increasingly diverse and complex digital landscape. By addressing these challenges and exploring these opportunities, we can expect sentiment analysis models to become even more powerful and relevant tools for understanding human emotions in text and beyond.

6. Declarations

6.1. Author Contributions

Conceptualization, T.H. and A.R.; methodology, A.R.; software, T.H.; validation, T.H. and A.R.; formal analysis, T.H. and A.R; investigation, T.H.; resources, A.R.; data curation, T.H.; writing—original draft preparation, T.H. and A.R.; writing—review and editing, T.H. and A.R.; visualization, A.R. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

7. References

[1] Shi, X., Wang, T., Wang, L., Liu, H., & Yan, N. (2019). Hybrid convolutional recurrent neural networks outperform CNN and RNN in Task-state EEG detection for Parkinson's disease. 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). doi:10.1109/APSIPAASC47483.2019.9023190.

[2] Kumar, A., & Garg, G. (2019). Sentiment analysis of multimodal twitter data. Multimedia Tools and Applications, 78(17), 24103–24119. doi:10.1007/s11042-019-7390-1.

[3] Chakraborty, I., Kim, M., & Sudhir, K. (2022). Attribute Sentiment Scoring with Online Text Reviews: Accounting for Language Structure and Missing Attributes. Journal of Marketing Research, 59(3), 600–622. doi:10.1177/00222437211052500.

[4] Li, J., Li, X., & He, D. (2019). A directed acyclic graph network combined with CNN and LSTM for remaining useful life prediction. IEEE Access, 7, 75464-75475. doi:10.1109/ACCESS.2019.2919566.

[5] Praveen, S. V., & Vajrobol, V. (2023). Understanding the Perceptions of Healthcare Researchers Regarding ChatGPT: A Study Based on Bidirectional Encoder Representation from Transformers (BERT) Sentiment Analysis and Topic Modeling. Annals of Biomedical Engineering, 51(8), 1654–1656. doi:10.1007/s10439-023-03222-0.

[6] Sudirjo, F., Diantoro, K., Al-Gasawneh, J. A., Khootimah Azzaakiyyah, H., & Almaududi Ausat, A. M. (2023). Application of ChatGPT in Improving Customer Sentiment Analysis for Businesses. Jurnal Teknologi Dan Sistem Informasi Bisnis, 5(3), 283–288. doi:10.47233/jteksis.v5i3.871.

[7] Susnjak, T. (2023). Applying Bert and ChatGPT for sentiment analysis of Lyme disease in scientific literature. arXiv preprint arXiv:2302.06474. doi:10.48550/arXiv.2302.064474.

[8] Sharma, S., Aggarwal, R., & Kumar, M. (2023). Mining Twitter for Insights into ChatGPT Sentiment: A Machine Learning Approach. 2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE). doi:10.1109/icdcece57866.2023.10150620.

[9] Wang, Z., Xie, Q., Ding, Z., Feng, Y., & Xia, R. (2023). Is ChatGPT a good sentiment analyzer? A preliminary study. arXiv preprint arXiv:2304.04339. doi:10.18550/arXiv.2304.04339.

[10] Yang, K., Ji, S., Zhang, T., Xie, Q., & Ananiadou, S. (2023). On the evaluations of ChatGPT and emotion-enhanced prompting for mental health analysis. arXiv preprint arXiv:2304.03347. doi:10.48550/arXiv.2304.03347.

[11] Haque, M. U., Dharmadasa, I., Sworna, Z. T., Rajapakse, R. N., & Ahmad, H. (2022). " I think this is the most disruptive technology": Exploring Sentiments of ChatGPT Early Adopters using Twitter Data. arXiv preprint arXiv:2212.05856. doi:10.48550/arXiv.2212.05856.

[12] Fatouros, G., Soldatos, J., Kouroumali, K., Makridis, G., & Kyriazis, D. (2023). Transforming Sentiment Analysis in the Financial Domain with ChatGPT. arXiv preprint arXiv:2308.07935. doi:10.48550/arXiv.2308.07935.

[13] Karanouh, M. (2023). Mapping ChatGPT in Mainstream Media: Early Quantitative Insights through Sentiment Analysis and Word Frequency Analysis. arXiv preprint arXiv:2305.18340. doi:10.48550/arXiv.2305.18340.

[14] Saengrith, W., Viriyavejakul, C., & Pimdee, P. (2022). Problem-Based Blended Training via Chatbot to Enhance the Problem-Solving Skill in the Workplace. Emerging Science Journal, 6, 1-12. doi:10.28991/ESJ-2022-SIED-01.

[15] Zhou, J., Meng, M., Gao, Y., Ma, Y., & Zhang, Q. (2018). Classification of motor imagery EEG using wavelet envelope analysis and LSTM networks. Chinese Control and Decision Conference (CCDC), IEEE, 5600-5605. doi:10.1109/CCDC.2018.8408108.

[16] Bahmei, B., Birmingham, E., & Arzanpour, S. (2022). CNN-RNN and Data Augmentation Using Deep Convolutional Generative Adversarial Network for Environmental Sound Classification. IEEE Signal Processing Letters, 29, 682–686. doi:10.1109/LSP.2022.3150258.

[17] Stephen, J. J., & Prabu, P. (2019). Detecting the magnitude of depression in Twitter users using sentiment analysis. International Journal of Electrical and Computer Engineering, 9(4), 3247–3255. doi:10.11591/ijece.v9i4.pp3247-3255.

[18] Huang, J., Lin, S., Wang, N., Dai, G., Xie, Y., & Zhou, J. (2020). TSE-CNN: A Two-Stage End-to-End CNN for Human Activity Recognition. IEEE Journal of Biomedical and Health Informatics, 24(1), 292–299. doi:10.1109/JBHI.2019.2909688.

[19] Kim, J., & Lee, M. (2014). Robust Lane Detection Based on Convolutional Neural Network and Random Sample Consensus. Lecture Notes in Computer Science, 454–461, Springer, Cham, Switzerland. doi:10.1007/978-3-319-12637-1_57.

[20] Verma, P., Khanday, A. M. U. D., Rabani, S. T., Mir, M. H., & Jamwal, S. (2019). Twitter sentiment analysis on Indian government project using R. International Journal of Recent Technology and Engineering, 8(3), 8338–8341. doi:10.35940/ijrte.C6612.098319.

[21] Asghar, M. Z., Sattar, A., Khan, A., Ali, A., Masud Kundi, F., & Ahmad, S. (2019). Creating sentiment lexicon for sentiment analysis in Urdu: The case of a resource-poor language. Expert Systems, 36(3), 12397. doi:10.1111/exsy.12397.

[22] Hung, S. L., Kao, C. Y., & Huang, J. W. (2022). Constrained K-means and genetic algorithm-based approaches for optimal placement of wireless structural health monitoring sensors. Civil Engineering Journal, 8(12), 2675-2692. doi:10.28991/CEJ-2022-08-12-01.

[23] Elzayady, H., Badran, K. M., & Salama, G. I. (2020). Arabic Opinion Mining Using Combined CNN - LSTM Models. International Journal of Intelligent Systems and Applications, 12(4), 25–36. doi:10.5815/ijisa.2020.04.03.

[24] Behera, B., & Kumaravelan, G. (2021). Text document classification using fuzzy rough set based on robust nearest neighbor (FRS-RNN). Soft Computing, 25(15), 9915–9923. doi:10.1007/s00500-020-05410-9.

[25] Wani, U. P., Gatagat, Y., & Thalor, M. Handwritten Character Recognition Using CNN, KNN and SVM. International Journal of Technology Engineering Arts Mathematics Science, 1(2), 2583-1224.

[26] Pano, T., & Kashef, R. (2020). A complete Vader-based sentiment analysis of hum (BTC) tweets during the ERA of COVID-19. Big Data and Cognitive Computing, 4(4), 1–17. doi:10.3390/bdcc4040033.

[27] Thelen, G. (2021). Leadership in a Global World Management Training Requirement Using the Example of the Asian Studies Program at University of Applied Sciences (HTWG) Konstanz. International Journal for Applied Information Management, 1(3), 125–135. doi:10.47738/ijaim.v1i3.14.

[28] Al-Jedibi, W. (2022). The Strategic Plan of the Information Technology Deanship - King Abdulaziz University- Saudi Arabia. International Journal for Applied Information Management, 2(4), 84–94. doi:10.47738/ijaim.v2i4.40.

[29] Endsuy, R. (2021). Sentiment Analysis between VADER and EDA for the US Presidential Election 2020 on Twitter Datasets. Journal of Applied Data Sciences, 2(1). doi:10.47738/jads.v2i1.17.

[30] Riyanto, & Azis, A. (2021). Application of the Vector Machine Support Method in Twitter Social Media Sentiment Analysis Regarding the Covid-19 Vaccine Issue in Indonesia. Journal of Applied Data Sciences, 2(3), 102–108. doi:10.47738/jads.v2i3.40.

[31] Efendi, A., Purwana, D., & Buchdadi, A. D. (2021). Human Capital Management of Government Internal Supervisory at the Ministry of Defense of the Republic Indonesia. International Journal for Applied Information Management, 2(2), 81–89. doi:10.47738/ijaim.v2i2.30.

[32] Ye, E. Z., Ye, E. H., Bouthillier, M., & Ye, R. Z. (2022). DeepImageTranslator V2: Analysis of Multimodal Medical Images using Semantic Segmentation Maps Generated through Deep Learning. HighTech and Innovation Journal, 3(3), 319-325. doi:10.28991/HIJ-2022-03-03-07.

[33] Aini, Q., Hammad, J. A. H., Taher, T., & Ikhlayel, M. (2021). Classification of Tweets Causing Deadlocks in Jakarta Streets with the Help of Algorithm C4.5. Journal of Applied Data Sciences, 2(4), 143–156. doi:10.47738/jads.v2i4.43.

[34] Hananto, A. R., Rahayu, S. A., & Hariguna, T. (2021). An Ensemble and Filtering-Based System for Predicting Educational Data Mining. Journal of Applied Data Sciences, 2(4), 157–173. doi:10.47738/jads.v2i4.44.

[35] Hariguna, T., Sukmana, H. T., & Kim, J. Il. (2020). Survey Opinion using Sentiment Analysis. Journal of Applied Data Sciences, 1(1), 35–40. doi:10.47738/jads.v1i1.10.

[36] Hori, M. (2021). Study of Career Education for Women: Development of Global Human Resources. International Journal for Applied Information Management, 1(2), 11–20. doi:10.47738/ijaim.v1i2.9.

[37] Musa, S., Muhyiddin, Y., & Nurhayati, S. (2022). Agriculturally based Equivalent Education: Insights on Nonformal Education Human Resources and Program Quality. Journal of Human, Earth, and Future, 3(4), 441-451. doi:10.28991/HEF-2022-03-04-04.

[38] Hariguna, T., & Rachmawati, V. (2019). Community Opinion Sentiment Analysis on Social Media Using Naive Bayes Algorithm Methods. IJIIS: International Journal of Informatics and Information Systems, 2(1), 33–38. doi:10.47738/ijiis.v2i1.11.