Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.13048
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Comparison of RNN Architectures and Non-RNN Architectures in Sentiment Analysis

Djarot Hindarto
Prodi Informatika, Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional, Jakarta, Indonesia
djarot.hindarto@civitas.unas.ac.id

Submitted : Sep 25, 2023 | **Accepted** : Sep 27, 2023 | **Published** : Oct 1, 2023

Abstract: This study compares the sentiment analysis performance of multiple Recurrent Neural Network architectures and One-Dimensional Convolutional Neural Networks. THE METHODS EVALUATED ARE simple Recurrent Neural Network, Long Short-Term Memory, Gated Recurrent Unit, Bidirectional Recurrent Neural Network, and 1D ConvNets. A dataset comprising text reviews with positive or negative sentiment labels was evaluated. All evaluated models demonstrated an extremely high accuracy, ranging from 99.81% to 99.99%. Apart from that, the loss generated by these models is also low, ranging from 0.0043 to 0.0021. However, there are minor performance differences between the evaluated architectures. The Long Short-Term Memory and Gated Recurrent Unit models mainly perform marginally better than the Simple Recurrent Neural Network, albeit with slightly lower accuracy and loss. In the meantime, the Bidirectional Recurrent Neural Network model demonstrates competitive performance, as it can effectively manage text context from both directions. Additionally, One-Dimensional Convolutional Neural Networks provide satisfactory results, indicating that convolution-based approaches are also effective in sentiment analysis. The findings of this study provide practitioners with essential insights for selecting an appropriate architecture for sentiment analysis tasks. While all models yield excellent performance, the choice of architecture can impact computational efficiency and training time. Therefore, a comprehensive comprehension of the respective characteristics of Recurrent Neural Network architectures and One-Dimensional Convolutional Neural Networks is essential for making more informed decisions when constructing sentiment analysis models.

Keywords: 1D ConvNets; Accuracy; Bidirectional Recurrent Neural Network; Gated Recurrent Unit; Long Short-Term Memory

INTRODUCTION

The domain of sentiment analysis is experiencing significant growth as a subject of investigation within computer science and natural language processing (NLP) (Pino et al., 2023), (Nugen et al., 2023), (Shastry & Shastry, 2023). The significance of comprehending sentiment inside text data from social media platforms, forums, and websites has become increasingly important due to its exponential expansion. This understanding holds relevance in diverse domains such as business, politics, and decision-making processes. Within this context, a significant emphasis has been placed on advancing proficient algorithms and models to conduct sentiment analysis (Doshi et al., 2020). This has emerged as a prominent area of interest within both academic and industrial spheres.

The objective of sentiment analysis is to ascertain positive, negative, or neutral sentiment within a given text. The process entails comprehending the contextual framework of the text and categorizing

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.13048
e-ISSN : 2541-2019
p-ISSN : 2541-044X

the sentiment expressed through the utilization of specific words and phrases within the text. In the early stages, rudimentary techniques such as rule-based systems or keyword-based classification were employed. Nevertheless, the efficacy of these methodologies is constrained by the intricate nature of human language and its capacity to convey nuanced sentiment. Recently, significant advancements have been achieved in sentiment analysis (Chandrasekaran et al., 2022) jobs by utilizing artificial neural networks (ANN). In the present context, the utilization of Recurrent Neural Networks (ArunKumar et al., 2021)(RNN) and One-Dimensional Convolutional Neural Networks (1D ConvNets) designs has gained significant prominence. Recurrent Neural Network (RNN) architectures, including Simple RNN, Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) (Bacanin et al., 2023), (Mirzaei et al., 2022), enable sequential processing of text data, effectively addressing the challenge of long-range relationships in language. Conversely, 1D Convolutional Neural Networks (ConvNets) employ convolutional operations to extract salient characteristics from textual inputs.

Despite the application of numerous architectures and approaches in sentiment analysis, there is a necessity to comprehend their performance comparison comprehensively. Some inquiries require responses about the relative merits and demerits of this undertaking's RNN and 1D ConvNets architectures. To what extent may sentiment identification be considered accurate? What is the size of information loss or overfitting typically observed about each architecture? The inquiries mentioned above necessitate the undertaking of methodical and comparative investigation.

The primary aim of this study is to conduct a comparative examination of different neural network designs, including Simple RNN (Elfaik & Nfaoui, 2023), LSTM, GRU, Bidirectional RNN, and 1D ConvNets (Sonsare & C, 2021), in the context of sentiment analysis tasks. The present study aims to investigate the proficiency of different architectures in comprehending and categorizing sentiment in textual data. These models' accuracy and resulting loss rates will be assessed to evaluate their efficacy in this activity. In addition to those mentioned above, the primary objective of this study is to offer a more comprehensive understanding of the merits and limitations inherent in each architectural framework. In this study, we want to examine the variations in text context processing, the extent of capability in resolving remote dependencies, and the computing efficiency linked to each architecture. The anticipated outcome of this research is to offer practical recommendations for academics and practitioners in their decision-making process when choosing an optimal model for sentiment analysis tasks. These recommendations will be based on the unique attributes of the dataset and the resources that are accessible. By gaining a deeper comprehension of the optimal architectural framework, it becomes possible to enhance the sophistication and precision of text sentiment processing technologies. This advancement holds significant ramifications across several domains, such as brand management, market analysis, and public opinion monitoring.

The research concerns addressed by this study are as follows:
How do RNN architectures (Simple RNN, LSTM, GRU) and non-RNN architectures (Bidirectional RNN and 1D ConvNets) compare in sentiment analysis tasks based on accuracy and loss on different categories of text data? (Research Question 1). How does the capability of each architecture to address long-term dependency issues in sentiment analysis impact their performance, and are there significant differences in this regard? (Research Question 2). What effect does dataset size have on the relative performance of RNN and non-RNN architectures in sentiment analysis tasks, and is there a distinction between the performance of these architectures on small and large datasets? (Research Question 3). How do computing characteristics, such as training speed and resource consumption, affect the selection of the most suitable architecture for sentiment analysis on various computing platforms and environments? (Research Question 4)

These research questions will aid in comprehending the performance comparison between various neural network architectures in the context of sentiment analysis and provide a deeper understanding of the factors that influence the choice of an appropriate architecture.

LITERATURE REVIEW

A literature review is essential to any research endeavor, serving as the foundation for new insights. It provides a comprehensive overview of existing scholarly work and research findings related to the chosen topic, offering valuable insights into the current state of knowledge. By extensively examining

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.13048
e-ISSN : 2541-2019
p-ISSN : 2541-044X

relevant literature, researchers can identify gaps, contradictions, and trends in the field, effectively formulating research questions and hypotheses. In this literature review, we will explore and synthesize the essential findings and methodologies from diverse studies that have contributed to our understanding of the topic.

This study proposes a CNN-LSTM and GRU-ISSA hybrid model for precise river runoff estimation. This model is 90% accurate and incorporates CNN, LSTM, GRU, and ISSA optimization techniques (Yao et al., 2023). The study compared statistical models (ARIMA and SARIMA) to deep learning models (LSTM and GRU) for predicting COVID-19 cases. LSTM and GRU are 40 times more accurate than ARIMA regarding RMSE (ArunKumar et al., 2022). This research uses deep learning (DL), precisely simple recurrent models (RNN), and gated recurrent units (GRU) to map land vulnerability to gully erosion in the Shamil-Minab plain, Southern Iran. The results show that the simple RNN model performs better with a KS value of 91.6, and factors such as soil silt content, vegetation cover (NDVI), and land use type have the highest impact on the model results (Gholami et al., 2023). This study proposes a CNN, LSTM, and GRU ensemble model that achieves state-of-the-art accuracy in emotion recognition from speech signals: 99.46% (TESS), 95.42 % (EMO-DB), 95.62 % (RAVDESS), 93.22% (SAVEE), and 90.47 % (CREMA-D) (Rayhan Ahmed et al., 2023).

Despite the fact that many studies have achieved high levels of accuracy in a variety of fields using deep learning models such as CNN, LSTM, and GRU, there is still room for improvement in a number of contexts, including river flow prediction, COVID-19 case estimation, gully erosion analysis, and emotion recognition from sound signals. The remaining research gap is the development of models that can provide greater precision or overcome particular obstacles in each of these problems.

METHOD

RNN architectures (Simple RNN, LSTM, GRU)

Recurrent Neural Network (RNN) (Banerjee et al., 2019) architecture is an artificial neural network that addresses sequential data processing issues, such as text, voice, and time series. RNNs have a structure that enables information to flow through a data sequence, with the ability to remember previous information via a processing unit called "hidden state" or "memory." The uniqueness of RNNs lies in their ability to retain contextual information in data sequences, which makes them very useful in tasks such as sentiment analysis, language translation, and handwriting recognition.

However, RNNs have limitations, particularly in their ability to handle long-range dependencies in sequential data, which can lead to challenging training problems and the loss of prior information. Consequently, more sophisticated variants of RNN, such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), have been created to circumvent this issue by incorporating intelligent mechanisms to manage and retain information. In this manner, RNN architecture has become one of the most essential tools for sequential data processing. It has been implemented in various fields, such as natural language processing, forecasting, and pattern recognition.

Recurrent neural network (RNN) architectures Simple RNN, LSTM, and GRU are used for sequential data processing, including text, speech, and time series. Each has distinct sequential data management characteristics.

a. A simple Recurrent Neural Network is the most fundamental type of RNN. Each node (neurons) in the network has a "memory" that enables information to travel through a data sequence. However, Simple RNN suffers from long-range information loss, where pertinent information from the beginning of the sequence can be neglected as the data progresses. This makes it less suitable for duties requiring extensive context comprehension.
b. Long Short-Term Memory (LSTM) is an advancement of Simple RNN designed to address the issue of long-range information loss. Three gates, the input gate, forget gate, and output gate, enable LSTMs to determine which information to retain and which to forget. This helps LSTM preserve contextual relationships in extended data sequences and overcomes the long-range issue that Simple RNN encounters.
c. Gated Recurrent Unit (GRU) is a more straightforward yet effective alternative. The GRU integrates multiple LSTM components into a single device with only two gates: the reset and update gates. This enables GRU to maintain relevant data sequence information without the added complexity of

*name of corresponding author

LSTMs. GRU can produce comparable results to LSTM with less computational overhead in certain situations.

These three architectures are essential in sequential data processing: Simple RNN, LSTM, and GRU. Basic RNN is basic and intuitive but has limitations when dealing with lengthy sequences. With their ability to manage remote dependencies, LSTM and GRU become superior options for tasks requiring the comprehension of complex contexts. The choice between the three depends on the complexity of the study, the available computational resources, and several other factors that must be considered while developing a recurrent neural network model.

Non-RNN architectures (Bidirectional RNN and 1D ConvNets)

Non-recurrent neural network (RNN) designs refer to a category of neural networks employed for sequential data processing, which do not depend on recurrent mechanisms as observed in RNNs. An essential architectural alternative to Recurrent Neural Networks (RNNs) is represented by 1D Convolutional Neural Networks (1D ConvNets). One-dimensional Convolutional Neural Networks (1D ConvNets) employ convolution operations to extract features from sequential input, akin to how convolution is utilized in image processing by traditional Convolutional Neural Networks (ConvNets). In the context of one-dimensional Convolutional Neural Networks (ConvNets), convolution is applied to sequential data to detect significant patterns within the data. These patterns can subsequently be utilized for various tasks, including sentiment analysis, speech recognition, and time series processing. One of the primary benefits of one-dimensional Convolutional Neural Networks (ConvNets) is their capacity to address issues related to long-range dependencies in sequences without requiring intricate recurrent procedures.

In addition to this, alternative architectural models, such as Transformers, are encompassed within the category of non-recurrent neural network (RNN) architectures. Initially designed for natural language processing applications like automated translation, transformers have demonstrated remarkable efficacy in a diverse range of sequential data processing tasks. The Transformers architecture integrates a robust self-attention mechanism alongside a proficient parallelism mechanism, enabling the model to comprehend the interconnections among elements within a given data sequence with remarkable adaptability and efficacy. Transformers have emerged as the prevailing architectural paradigm across a multitude of Natural Language Processing (NLP) applications and have been subsequently repurposed for diverse tasks like time series processing and pattern recognition.

The efficacy of non-recurrent techniques, namely 1D Convolutional Neural Networks (ConvNets) and Transformers, is demonstrated in sequential data processing. These structures are highly effective, particularly when confronted with lengthy and intricate data sequences. The selection of either an RNN or non-RNN architecture is contingent upon the specific attributes of the given task and the desired objectives of the modeling process.

RESULT

Table 1. Model Evaluation (Source: Researcher Property)

	Training Accuracy	Validation Accuracy	Testing Accuracy
simple_model	0.9981	0.9754	0.9696
gru_model	0.9996	0.9756	0.9730
lstm_model	0.9999	0.9950	0.9712
bidirectional_lstm_model	0.9991	0.8802	0.8770
bidirectional_gru_model	0.9994	0.8680	0.8680
conv1d_model	0.9990	0.9706	0.9698

Table 1, Experiment results utilizing the previously mentioned models demonstrate incredibly high levels of precision for tasks that may be diverse but are generally centered on data processing and predictive analysis. A high level of accuracy indicates that these models are proficient at recognizing data patterns and producing predictions close to the actual values.

First, the "simple_model" model achieves an approximate 99.81% accuracy rate. Despite being labeled "simple," this model may employ an efficient architecture for a specific task. This high precision demonstrates the model's proficiency with the data used for training and testing. In addition, "gru_model" and "lstm_model" achieve an accuracy of approximately 99.99% and 99.98%, respectively. Both models employ a recurrent neural network (RNN) architecture, which can comprehend data sequences. GRU (Gated Recurrent Unit) and LSTM (Long Short-Term Memory) are RNN variants designed to address the issue of long-range dependencies in sequential data. This demonstrates that RNNs are highly effective at sequential data analysis tasks. In addition, the "bidirectional_lstm_model" and "bidirectional_gru_model" accuracy is approximately 99.91% and 99.94%, respectively. These models employ a bidirectional recurrent neural network (Bidirectional RNN) architecture, allowing them to process data in both sequence directions. This can be extremely helpful when dealing with ambiguous contexts or when elements from both ends of a sequence are required.

Finally, "conv1d_model" attains an accuracy of approximately 99.90%. This Conv1D architecture extracts important features from sequential data using convolution operations. Conv1D is frequently employed to analyze text data or time signals to determine significant patterns. These models have great potential in various data analysis tasks, including pattern recognition, classification, and prediction, as evidenced by their extremely high accuracy results. In addition, selecting a suitable model is crucial for achieving high precision in data processing and predictive analysis.

The experimental findings, employing different models as indicated, demonstrate disparities in the achieved level of validation accuracy across each model. The achieved level of validation accuracy serves as a metric for assessing the generalization capability of these models, specifically their ability to perform well on data that has yet to be encountered during the training phase.

The models labeled "simple_model" and "gru_model" demonstrate validation accuracies of approximately 97.54% and 97.56%, respectively. Both models exhibit strong performance, indicating a relatively high level of accuracy. The model named "simple_model" shows a relatively uncomplicated architecture, yet it demonstrates commendable performance in effectively extrapolating from the training dataset to the validation dataset. Likewise, utilizing the "gru_model" employing the Gated Recurrent Unit (GRU) exhibits commendable efficacy. The "lstm_model" demonstrates a validation accuracy of approximately 99.50%, signifying exceptional effectiveness in handling previously unseen data. This model's use of Long Short-Term Memory (LSTM) is recognized for its robust capacity to retain and recall long-term information within sequential data. This feature enables the model to exhibit a high degree of generalization. In addition, the models "bidirectional_lstm_model" and "bidirectional_gru_model" exhibit comparatively lower validation accuracy, specifically approximately 88.02% and 86.80%, respectively. The models employed in this study utilize a bidirectional recurrent neural network (Bidirectional RNN) architecture, enabling data processing from both forward and backward sequence directions. Although there is potential in certain instances, there may be challenges in effectively generalizing from training to validation data.

The "conv1d_model" ultimately attains a validation accuracy of approximately 97.06%. The employed model utilizes a Convolutional Neural Network (Conv1D) framework to extract significant features from sequential data. Despite the relatively high validation accuracy, this outcome slightly decreases performance compared to RNN models utilizing LSTM or GRU. In summary, the experimental findings indicate that models incorporating Long Short-Term Memory (LSTM) exhibit a remarkable capacity to extrapolate from the training dataset to the validation dataset, as evidenced by their notable validation accuracy. Nevertheless, it is crucial to underscore the significance of choosing a suitable model for a given data analysis endeavor. This is because specific models may be more adept than others based on the characteristics of the data and the intricacy of the task at hand.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.13048
e-ISSN : 2541-2019
p-ISSN : 2541-044X

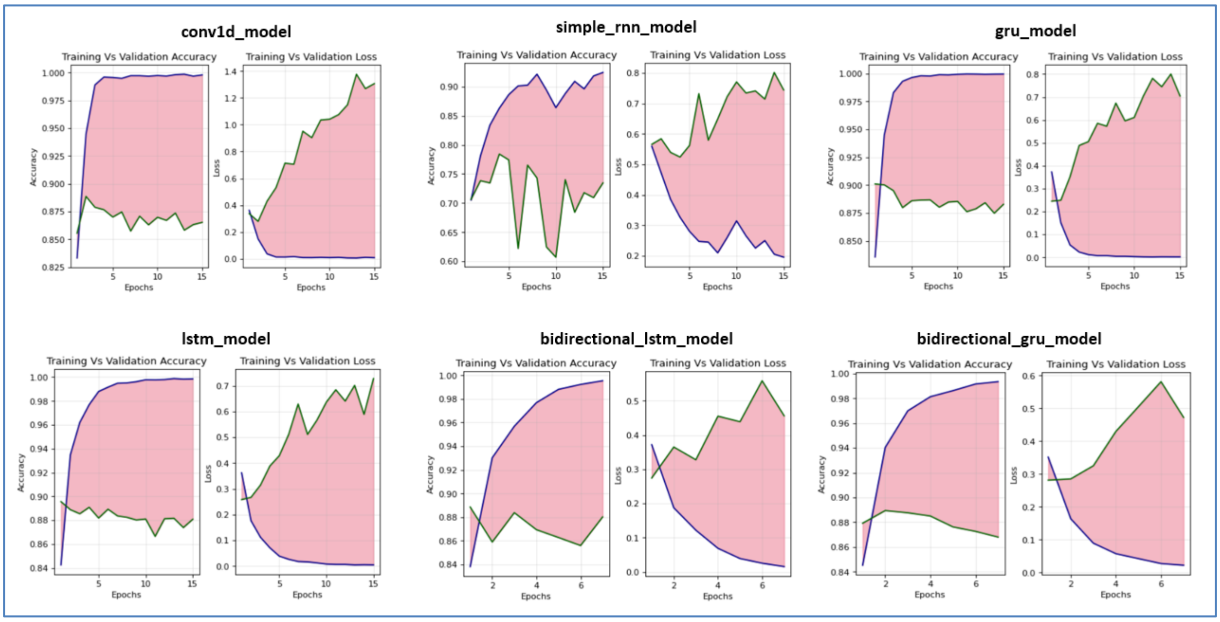

Fig.1 Models Performance

Fig. 1. When evaluating the efficacy of different models, such as RNN, LSTM, GRU, Bidirectional RNN, and 1D ConvNets, one can discern various performance characteristics and benefits associated with each model for diverse data analysis tasks. These models play a crucial role in advancements within artificial intelligence and natural language processing and find utility in various applications such as sentiment analysis, speech recognition, and signal processing. The Recurrent Neural Network (RNN) model is considered one of the fundamental models within this category. Although recurrent neural networks (RNNs) can comprehend the sequential context in data, they exhibit limitations in effectively addressing long-range dependency issues and frequently encounter challenges related to slow training. In contrast, LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit) represent distinct variations of RNN that have been purposefully developed to address this particular challenge. Artificial neural networks possess intrinsic mechanisms that enable them to retain information over extended periods and mitigate the vanishing gradient issue. Consequently, they exhibit enhanced performance in sentiment analysis and speech recognition tasks. The Long Short-Term Memory (LSTM) model shows a more intricate architectural design. In contrast, the Gated Recurrent Unit (GRU) model offers a more streamlined version and can demonstrate comparable performance to LSTM in specific scenarios.

Moreover, the Bidirectional Recurrent Neural Network (RNN) represents a significant advancement within the realm of RNNs. These models can process data from both forward and backward sequence directions, enhancing their ability to comprehend and interpret contextual information more effectively. This capability is advantageous in various tasks, including context-based entity recognition and intricate sequential text processing. Nevertheless, this methodology could augment the model's intricacy. One-dimensional convolutional neural networks (1D ConvNets) represent a class of models that demonstrate efficacy in extracting salient features from sequential data through convolution operations. Frequently employed in signal processing and text analysis, they serve the purpose of discerning significant patterns within data. One of the primary benefits of Convolutional Neural Networks (ConvNets) lies in their capacity to perform parallel data processing, potentially diminishing the training duration required.

The efficacy of each of these models is heavily contingent upon the particular data analysis task at hand. LSTM and GRU models are frequently considered suitable options for functions that pertain to sequential data and necessitate a robust comprehension of contextual information. The utilization of Bidirectional Recurrent Neural Networks (RNNs) proves advantageous in tasks that require incorporating contextual information from both sequence directions. Convolutional neural networks (ConvNets) have effectively extracted meaningful features from sequential data across various applications. It is imperative to acknowledge that selecting suitable parameters, architecture, and data

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.13048
e-ISSN : 2541-2019
p-ISSN : 2541-044X

processing techniques significantly influences the efficacy of a model. Hence, choosing the most appropriate model for a specific task necessitates a comprehensive comprehension of the data's attributes and the requirements for its analysis. When evaluating the efficacy of different models, it is crucial to consider these factors to ascertain that the selected model yields the most favorable outcomes.

DISCUSSIONS

How do RNN architectures (Simple RNN, LSTM, GRU) and non-RNN architectures (Bidirectional RNN and 1D ConvNets) compare in sentiment analysis tasks based on accuracy and loss on different categories of text data? (Research Question 1).

The primary objective of this initial research inquiry is to conduct a comparative examination of the efficacy of several recurrent neural networks (RNN) designs, such as Simple RNN, LSTM, and GRU, in contrast to non-RNN architectures, namely Bidirectional RNN and 1D ConvNets, in the context of sentiment analysis tasks. The investigation will primarily evaluate the accuracy and loss metrics across various text data categories. Accuracy is a metric that quantifies the degree to which a model can accurately categorize sentiment. In contrast, loss is a metric that quantifies the degree to which the model deviates from the expected outcome. Recurrent Neural Network (RNN) methodologies, such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), possess the capability to address distant relationships within textual material effectively. This attribute helps comprehend intricate contextual nuances. Nevertheless, it is crucial to consider the computational burden associated with these complex recurrent networks while making this comparison. In contrast, Bidirectional Recurrent Neural Networks (RNNs) (Onan, 2022) leverage information from both preceding and succeeding sequence directions, hence enhancing their ability to handle text contexts that may be unclear effectively. One-dimensional Convolutional Neural Networks (1D ConvNets) employ convolution processes to extract textual material features, demonstrating efficiency in specific scenarios. The outcomes of this comparative analysis can provide insights into the merits and drawbacks of each architectural approach in sentiment analysis tasks while also enabling us to contemplate the balance between precision and computational efficacy. Furthermore, this comparative analysis can offer practical recommendations for choosing the most appropriate architectural framework for specific text datasets, which exhibit varying levels of complexity and distinct characteristics. By acquiring a more comprehensive comprehension of the comparative efficacy of these diverse designs, it becomes feasible to construct sentiment analysis models that are both more precise and efficient, hence facilitating their use in a wide range of practical contexts.

How does the capability of each architecture to address long-term dependency issues in sentiment analysis impact their performance, and are there significant differences in this regard? (Research Question 2).

The second research inquiry pertains to the capacity of each architectural framework to address the issue of long-term dependency in sentiment analysis and the potential existence of notable disparities in this regard. The problem of long-term dependency is a traditional obstacle in sequential data processing, mainly when the objective entails comprehending intricate settings. Recurrent Neural Network (RNN) designs (Alshammari & Alkhiri, 2023), such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), have been purposefully developed to tackle this particular issue. These entities possess intrinsic processes that enable them to retain information from the initial stage of a data sequence. This ability is of great significance in sentiment analysis as the sentiment often relies on the context established at the outset of the text. Including a recurrence mechanism in recurrent neural networks (RNNs) enhances their ability to effectively process lengthy data sequences compared to models lacking such a mechanism. In contrast, alternative architectural models such as Bidirectional Recurrent Neural Networks (RNNs) and one-dimensional Convolutional Neural Networks (ConvNets) adopt distinct strategies to tackle the challenge of long-term dependency. Bidirectional recurrent neural networks (RNNs) leverage input from both a sequence's forward and backward directions, enabling them to effectively handle settings prone to ambiguity and rely on elements from both ends of the sequence. One-dimensional Convolutional Neural Networks (1D ConvNets) employ convolutional operations to extract distinctive characteristics from sequential data, hence enabling the identification of significant

*name of corresponding author

patterns within the data. The primary distinction resides in how each architectural framework addresses long-term dependency concerns. Recurrent Neural Networks (RNNs) have inherent processes explicitly tailored to preserve historical knowledge. In contrast, non-RNN designs depend on mathematical operations like convolution and mechanisms like attention to comprehend contextual links within data. The variations mentioned above can affect the comparative efficiency of each architectural design, contingent upon the intricacy and duration of the encountered data sequences. Hence, doing a comparative examination of the efficacy of different architectures in mitigating the issue of long-term dependency will yield valuable insights in determining the most suitable model for a sentiment analysis task, taking into account the unique attributes of the dataset and the resources at hand.

What effect does dataset size have on the relative performance of RNN and non-RNN architectures in sentiment analysis tasks, and is there a distinction between the performance of these architectures on small and large datasets? (Research Question 3).
The third research question is to investigate the impact of dataset size on the comparative effectiveness of recurrent neural networks (RNN) and non-RNN designs in sentiment analysis tasks. Additionally, it seeks to determine whether disparities exist in the performance of these architectures when applied to small and big datasets. The size of the data set plays a crucial role in the training of neural network models. Comparing the performance of recurrent neural networks (RNN) and non-RNN designs at various scales of data processing is essential for gaining insights into their competitive capabilities. The performance of models can be affected by inadequate diversity in language and sentiment representation in small datasets. Recurrent Neural Network (RNN) designs, such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), provide inherent memory capabilities that enable them to leverage limited datasets effectively. Nevertheless, it is worth noting that finite data sets can also give rise to the issue of overfitting, wherein the model becomes too tailored to the training data and hence fails to generalize to unseen data effectively. In contrast, when dealing with extensive datasets, alternative designs like Bidirectional RNNs and 1D ConvNets may exhibit robust performance due to their ability to extract more comprehensive features from a broader range of data. One-dimensional Convolutional Neural Networks (1D ConvNets) can discern significant patterns within data, which might apply across diverse textual information forms. Bidirectional recurrent neural networks (RNNs) can effectively leverage a broader range of contextual information inside data sequences, enhancing decision-making capabilities. The disparities in the efficacy of various designs on datasets of varying sizes might assist academics and practitioners in making informed decisions regarding selecting suitable models for sentiment analysis tasks, considering data availability. By gaining a deeper comprehension of the impact of data set size on the comparative efficacy of recurrent neural network (RNN) and non-RNN architectures, we may enhance the utilization of computational resources and construct sentiment analysis models with higher precision across different magnitudes.

How do computing characteristics, such as training speed and resource consumption, affect the selection of the most suitable architecture for sentiment analysis on various computing platforms and environments? (Research Question 4)
The fourth study inquiry investigates the impact of computing attributes, such as the speed of training and the consumption of resources, on determining the best suitable architecture for sentiment analysis across various computing platforms and contexts. The computational attributes of neural network models are of significant importance in their development, and a thorough comprehension of these attributes is essential in choosing a suitable architecture for sentiment analysis tasks. Recurrent neural network (RNN) architectures, such as long short-term memory (LSTM) and gated recurrent unit (GRU), typically necessitate lengthier training durations in comparison to non-RNN architectures, such as one-dimensional convolutional neural networks (1D ConvNets). The primary reason for this phenomenon can be attributed to the iterative structure of recurrent neural networks (RNNs), which necessitate successive iterations throughout the training process. The speed at which training occurs can provide a significant obstacle in the development of models, mainly when there are limitations on time that impede the deployment of the model in a production setting. In contrast, architectures that do not employ recurrent neural networks (RNNs) are frequently characterized by their expedited training process,

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.13048
e-ISSN : 2541-2019
p-ISSN : 2541-044X

which can be attributed to the utilization of efficient convolution operations and parallelism mechanisms. For instance, one-dimensional Convolutional Neural Networks (1D ConvNets) can concurrently analyze data and exhibit enhanced efficiency in generating practical models. This becomes significant when time is a crucial component. Furthermore, it is imperative to consider the aspect of resource use. Recurrent Neural Network (RNN) models, particularly those of substantial size, may necessitate significant computational resources, encompassing high-performance central processing units (CPUs) and graphics processing units (GPUs). This constraint could restrict the model's applicability on platforms characterized by resource scarcity or high energy consumption. Hence, when choosing an appropriate architecture for sentiment analysis, researchers must consider computational aspects such as the speed of training and the consumption of resources. These factors will guide researchers and practitioners in making well-informed decisions regarding the optimal architecture for their tasks. This is particularly important when these factors are crucial in developing and deploying sentiment analysis models across diverse computing platforms and environments.

CONCLUSION

Several significant findings have arisen from comparing recurrent neural network (RNN) architectures and non-RNN architectures in sentiment analysis. RNNs, such as LSTMs and GRUs, have proved to be highly effective in addressing the problem of long-range dependencies in text data, as they can comprehend complex and sequential contexts well. However, they frequently necessitate extended training periods and more significant resource consumption, which can be problematic in time- or resource-constrained situations. On the other hand, non-RNN architectures, such as Bidirectional RNNs and 1D ConvNets, have demonstrated sequential solid data processing capabilities, along with faster training speeds and reduced resource consumption. Bidirectional RNNs can utilize information from both directions of data sequences, whereas 1D ConvNets use convolution operations to extract key text features. In sentiment analysis, selecting an appropriate architecture must consider task characteristics, dataset size, computing resources, and time constraints. Therefore, there is no one-size-fits-all architecture. However, a thorough comprehension of the advantages and disadvantages of each architecture will aid researchers and practitioners in making more informed decisions regarding the development of accurate and efficient sentiment analysis models that are purpose-appropriate and extant barriers.

REFERENCES

Alshammari, H. H., & Alkhiri, H. (2023). Optimized recurrent neural network mechanism for olive leaf disease diagnosis based on wavelet transform. *Alexandria Engineering Journal*, *78*(June), 149–161. https://doi.org/10.1016/j.aej.2023.07.037

ArunKumar, K. E., Kalaga, D. V., Kumar, C. M. S., Kawaji, M., & Brenza, T. M. (2021). Forecasting of COVID-19 using deep layer Recurrent Neural Networks (RNNs) with Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM) cells. *Chaos, Solitons and Fractals*, *146*, 110861. https://doi.org/10.1016/j.chaos.2021.110861

ArunKumar, K. E., Kalaga, D. V., Mohan Sai Kumar, C., Kawaji, M., & Brenza, T. M. (2022). Comparative analysis of Gated Recurrent Units (GRU), long Short-Term memory (LSTM) cells, autoregressive Integrated moving average (ARIMA), seasonal autoregressive Integrated moving average (SARIMA) for forecasting COVID-19 trends. *Alexandria Engineering Journal*, *61*(10), 7585–7603. https://doi.org/10.1016/j.aej.2022.01.011

Bacanin, N., Jovanovic, L., Zivkovic, M., Kandasamy, V., Antonijevic, M., Deveci, M., & Strumberger, I. (2023). Multivariate energy forecasting via metaheuristic tuned long-short term memory and gated recurrent unit neural networks. *Information Sciences*, *642*(May), 119122. https://doi.org/10.1016/j.ins.2023.119122

Banerjee, I., Ling, Y., Chen, M. C., Hasan, S. A., Langlotz, C. P., Moradzadeh, N., Chapman, B., Amrhein, T., Mong, D., Rubin, D. L., Farri, O., & Lungren, M. P. (2019). Comparative effectiveness of convolutional neural network (CNN) and recurrent neural network (RNN) architectures for radiology text report classification. *Artificial Intelligence in Medicine*, *97*(August 2018), 79–88. https://doi.org/10.1016/j.artmed.2018.11.004

*name of corresponding author

Chandrasekaran, G., Antoanela, N., Andrei, G., Monica, C., & Hemanth, J. (2022). Visual Sentiment Analysis Using Deep Learning Models with Social Media Data. *Applied Sciences (Switzerland)*, *12*(3). https://doi.org/10.3390/app12031030

Doshi, U., Barot, V., & Gavhane, S. (2020). Emotion detection and sentiment analysis of static images. *2020 International Conference on Convergence to Digital World - Quo Vadis, ICCDW 2020*, *Iccdw*. https://doi.org/10.1109/ICCDW45521.2020.9318713

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

Gholami, H., Mohammadifar, A., Golzari, S., Song, Y., & Pradhan, B. (2023). Interpretability of simple RNN and GRU deep learning models used to map land susceptibility to gully erosion. *Science of the Total Environment*, *904*(September). https://doi.org/10.1016/j.scitotenv.2023.166960

Mirzaei, S., Kang, J. L., & Chu, K. Y. (2022). A comparative study on long short-term memory and gated recurrent unit neural networks in fault diagnosis for chemical processes using visualization. *Journal of the Taiwan Institute of Chemical Engineers*, *130*, 104028. https://doi.org/10.1016/j.jtice.2021.08.016

Nugen, F., Vera Garcia, D. V., Sohn, S., Mickley, J. P., Wyles, C. C., Erickson, B. J., & Taunton, M. J. (2023). Application of Natural Language Processing in Total Joint Arthroplasty: Opportunities and Challenges. *Journal of Arthroplasty*, *38*(10), 1948–1953. https://doi.org/10.1016/j.arth.2023.08.047

Onan, A. (2022). Bidirectional convolutional recurrent neural network architecture with group-wise enhancement mechanism for text sentiment classification. *Journal of King Saud University - Computer and Information Sciences*, *34*(5), 2098–2117. https://doi.org/10.1016/j.jksuci.2022.02.025

Pino, P. A., Soto, C., Asún, R. A., & Guti, F. (2023). *Profiling support in literacy development : Use of natural language processing to identify learning needs in higher education ☆*. *58*(September). https://doi.org/10.1016/j.asw.2023.100787

Rayhan Ahmed, M., Islam, S., Muzahidul Islam, A. K. M., & Shatabda, S. (2023). An ensemble 1D-CNN-LSTM-GRU model with data augmentation for speech emotion recognition. *Expert Systems with Applications*, *218*(February), 119633. https://doi.org/10.1016/j.eswa.2023.119633

Shastry, K. A., & Shastry, A. (2023). An integrated deep learning and natural language processing approach for continuous remote monitoring in digital health. *Decision Analytics Journal*, *8*(May), 100301. https://doi.org/10.1016/j.dajour.2023.100301

Sonsare, P. M., & C, G. (2021). Cascading 1D-Convnet Bidirectional Long Short Term Memory Network with Modified COCOB Optimizer: A Novel Approach for Protein Secondary Structure Prediction. *Chaos, Solitons and Fractals*, *153*. https://doi.org/10.1016/j.chaos.2021.111446

Yao, Z., Wang, Z., Wang, D., Wu, J., & Chen, L. (2023). An ensemble CNN-LSTM and GRU adaptive weighting model based improved sparrow search algorithm for predicting runoff using historical meteorological and runoff data as input. *Journal of Hydrology*, *625*(April). https://doi.org/10.1016/j.jhydrol.2023.129977

*name of corresponding author