BiGRU-ANN based hybrid architecture for intensified classification tasks with explainable AI

Article *in* International Journal of Information Technology · October 2023

DOI: 10.1007/s41870-023-01515-0

5 authors, including:

Sovon Chakraborty
European University of Bangladesh

Jannatun Noor
BRAC University

Jia Uddin
Woosong University

Int. j. inf. tecnol. (December 2023) 15(8):4211–4221
https://doi.org/10.1007/s41870-023-01515-0

ORIGINAL RESEARCH

BiGRU-ANN based hybrid architecture for intensified classification tasks with explainable AI

Sovon Chakraborty[1] · Muhammad Borhan Uddin Talukder[2] · Mohammad Mehadi Hasan[3] · Jannatun Noor[4] · Jia Uddin[5]

Received: 4 June 2023 / Accepted: 3 September 2023 / Published online: 4 October 2023

Abstract Artificial Intelligence (AI) is increasingly being employed in critical decision-making processes such as medical diagnosis, credit approval, criminal justice, and many more. However, many AI models exploit complex algorithms that are difficult for humans to see through, which can lead to concerns about accountability, bias, and the ability to trust the outcomes. With the increasing demand for AI systems to be transparent, interpretable, and reliable, the field of Explainable AI (XAI) has gained attention of the researchers. This paper presents a robust hybrid architecture that combines Bidirectional Gated Recurrent Units (BiGRU) and Artificial Neural Networks (ANN) for the classification of texts and sentiment analysis. Interpretable Model Agnostic Explanation (LIME) has been employed with our proposed model to enhance confidence in the outcomes. The proposed architecture is found to be effective for sentiment analysis from texts, and classifying images containing handwrit- ten characters. It leverages the BiGRU to model the sequential dependencies in the data, while the ANN is used for the final classification. Evaluations on both Bengali and English datasets show that the proposed architecture outperforms state-of-the-art models in various performance metrics, providing meaningful and interpretable explanations for its predictions. The model can be used in systems that require the architectures to be computationally less demanding, yet a decent accuracy is secured.

Keywords Sentiment analysis · Explainable AI · BiGRU · ANN · LIME

✉ Jia Uddin
jia.uddin@wsu.ac.kr

Sovon Chakraborty
sovonchakraborty2014@gmail.com

Muhammad Borhan Uddin Talukder
borhan.talukdar4466@gmail.com

Mohammad Mehadi Hasan
hasanbdmehadi@eub.edu.bd

Jannatun Noor
jannatun.noor@bracu.ac.bd

1 Department of Computer Science and Engineering, University of Liberal Arts Bangladesh, Dhaka, Bangladesh

2 Department of Computer Science and Engineering, Daffodil International University, Birulia, Bangladesh

3 Department of Computer Science and Engineering, European University of Bangladesh, Dhaka, Bangladesh

4 Department of Computer Science and Engineering, BRAC University, Dhaka, Bangladesh

5 AI and Big Data Department, Woosong University, Daejeon, Korea

1 Introduction

Technology and scientific practices have experienced a paradigm shift over the past few decades. Artificial Intelligence (AI) has facilitated many aspects of our lives. As AI gradually gets more specialized, so does the field of Natural language processing (NLP). The field of NLP has been receiving unceasing research attempts in recent years, which helped it advance all aspects it is concerned about. Sentiment analysis is such a concern that focuses on the identification and classification of human expressions using certain polarity values [1, 2].

Although the literature is quite abundant when it comes to sentiment analysis tasks in English, Bengali has not received adequate research exposure yet. Hence, Bengali sentiment analysis has still a lot to explore and contribute to. Various machine learning and deep learning models have been utilized to perform sentiment analysis tasks both in Bangla and

English [3–9]. However, while some yield expected results to some extent, others fail to render the desired outcomes. Long Short-Term Memory (LSTM) is such a deep learning model, which sometimes suffers from certain shortcomings despite having extensive applicability. The main disadvantages of this architecture include consuming prolonged time for training, which makes it a slow candidate for some tasks. Although the Bidirectional LSTM (BiLSTM) can perform classification tasks faster compared to LSTM, it is computationally expensive when dealing with large amounts of sequential data.

Gated Recurrent Unit (GRU), however, has an advantage over LSTM in the case of being computationally less expensive. When we combine two such forward and backward-gated recurrent units, a BiGRU is formed. This approach helps to balance out some of the difficulties that traditional LSTM and BiLSTM suffer from [10–12].

Artificial Neural Networks (ANN) are a type of feedforward neural networks that are known for their ability to learn complex non-linear relationships between inputs and outputs [13]. When combined to work as a hybrid model, BiGRUs and ANNs can together make the most out of both models to achieve an improved classification performance. The BiGRU component captures the temporal dependencies of sequential data, while the ANN component enables the network to learn complex non-linear relationships. Thus the BiGRU can be used to extract features from sequential data, which can then be fed into an ANN for classification.

1.1 Literature review

Sentiment analysis is a concern of NLP where the underlying sentiment of a text is determined and labeled in terms of certain polarity values. The literature has an abundant supply of sentiment analysis tasks. A wide range of methods has been adopted to perform different feats. Mikolov et al. introduced Word2Vec [14] to represent the underlying meaning of words. Maximum Bengali NLP tasks have utilized the Word2vec approach to adapt the proper representation of words [15–17]. Later, Word embedding preprocessing technique was proposed to overcome one of the major concerns of Word2vec which is not handling the out-of-vocabulary words properly [18]. The language structure in Bengali makes it somewhat intricate to categorize them based on their inherent emotions. The negative and neutral polarities are often challenging to draw a line between.

Besides, in Bengali sentiment categorization, Machine learning and deep learning models offer a wide range of approaches [19]. Some of the endevors in the English and Bengali sentiment analysis related to this field are summarized below.

A hybrid BiLSTM-ANN model is proposed in [29], where authors were not focused on providing any original work. Kowsher et al. demonstrated information extraction from human names and showed the distinctive performance of LSTM- based models [30]. But the number of trainable parameters is too high. The recent studies are shown in Table 1.

Understanding complex patterns is often challenging for GRU and LSTM [31]. The integration of DL architectures in such cases allows one to understand the patterns quickly. Besides, the number of trainable parameters is an indispensable criterion in classification tasks [32]. Despite having multiple gates, BiGRU has the capability to process data precisely with less trainable parameters after fine-tuning of hyperparameters [33]. With the help of ANN, complicated patterns and prediction issues can be modeled.

Although the NLP research works in Bengali suffer from less exposure to hybrid models, the literature, however, shows works in other languages where the researchers proposed hybrid models that showed outstanding performances [34, 35]. Being motivated by various hybrid models from the literature, we propose a hybrid deep learning architecture combining BiGRU and ANN. The model is capable of discerning complicated patterns where the number of trainable parameters is pretty modest. The experimental findings demonstrate that the model outperforms any cutting-edge architecture in a number of measures.

1.2 Contributions

The contributions of this paper can be summarized as follows:

- Two specialized datasets named JobSn and BdHikes are constructed.
- A Hybrid model comprising BiGRU and ANN is proposed. The model is validated using multiple datasets. Different statistical methods have been used to analyze the results. The model is found to be fruitful for both text and handwritten character image data.
- Explainable AI (XAI) is incorporated with the architecture to create a sense of how the model makes decisions,

Table 1 Recently utilized architectures in the correspond ing field

Article	Proposed architecture
Seyyar et al. [20]	BERT
Madsen et al. [21]	Support vector machine
Lauriola et al. [22]	Artificial Neural Network
Sharif et al. [23]	Multinomial Naive Bayes
Chakraborty et al. [24]	LSTM ANN
Salch et al. [25]	CNN LSTM
Tripathi et al. [26]	Artificial Neural Network
Youbi et al. [27]	Deep Neural Network
Kowsher et al. [28]	BiLSTM ANN

thereby enhancing the transparency, intelligibility, and interpretability of the proposed model.

- The proposed hybrid model outperforms previous state-of-the-art models.

The organization of this paper is further segmented into dif- ferent sections. Section 1 establishes the ground for this work, identifies the most recent research in the corresponding field, addresses some background studies that help to understand the research work, and mentions the contributions the authors of this paper made. Section 2 discusses the detailed methodology of this study. Experimental results are analyzed and discussed in Sect. 3. Section 4 draws the conclusion and addresses the scope of future research related to this study.

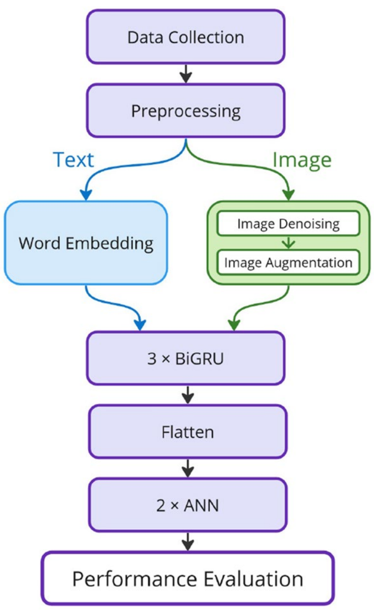


Fig. 1 A concise workflow of the study

2 Research methodology

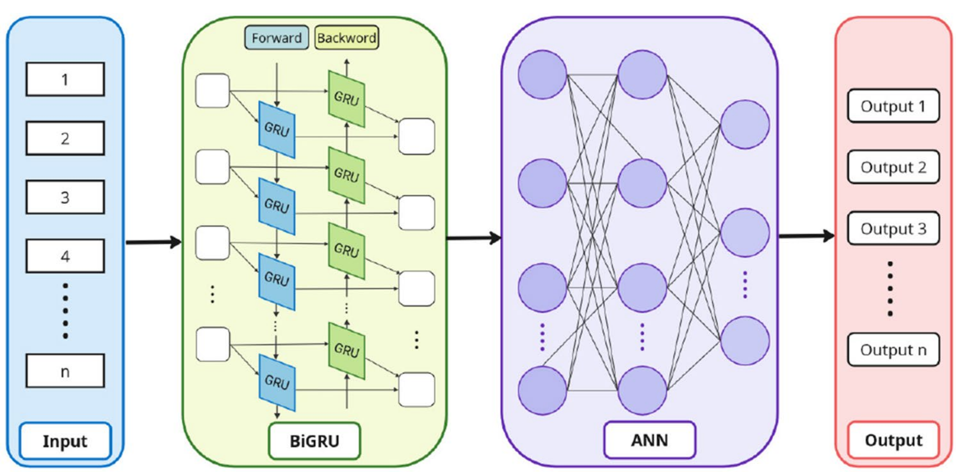
In this paper, we have proposed a hybrid architecture that includes BiGRU and ANN architectures. Figure 1 accounts for the concise workflow of the study, where subsequent steps are depicted. This research utilized two specialized datasets built by the authors of this paper, and another external dataset named NumtaDb, containing images of Bengali Handwritten Digits. The datasets were preprocessed first for preparing them to yield the optimum outcome. After preprocessing, the texts are embedded into numerical data, whereas the image data are denoised and augmented, before feeding them both into the model. The performance rendered by the architecture is then finally analyzed and discussed. Algorithm 1 shows the methodology of this research algorithmically.

2.1 The proposed hybrid model

We propose a hybrid deep learning architecture composed of BiGRU and ANN. While employed together as a hybrid model, they can be utilized to make the most out of both models and produce optimal outcomes. The BiGRU layers capture both the forward and backward contexts of the input data and ANN layers are then used to perform classification tasks. Figure 2 shows a brief outline of the proposed architecture. This study utilizes three datasets, two of which were built by the authors of this paper. These two datasets contain texts collected from various sources. Before feeding the datasets into the proposed architecture for training, datasets are preprocessed first, and the train-test ratio is set to 70:30 (Fig. 3).

During the training phase, data dropout is introduced to reduce the chances of overfitting. For each BiGRU layer, we have used a dropout of 0.01. For each epoch, two BiGRU layers are considered. 256 memory units are considered in the first layer where 128 are for the second

Fig. 2 Outline of the proposed hybrid architecture



Algorithm 1 : Methodology of the Study

1: Preparing datasets
   i) Gathering data to build the BdHikes and JobSn datasets
   ii) Preparing the NumtaDB dataset consisting Bengali handwritten digits
2: Preprocessing of the datasets
   i) JobSn and BdHikes: Annotating; Extracting salient data; Word embedding
   ii) NumtaDb: Image Resizing, Denoising, and Augmentation
3: Building, Training, and validation of the proposed architecture
   i) Building a BiGRU-ANN based hybrid architecture
   ii) Hyperparameter tuning
   iii) Running for a variable number of epochs and observing the performance
4: Performance evaluation of the model on the test set
5: Interpretation of the outcomes using XAI
   i) Analyzing the most significant words for a particular outcome
   ii) Highlighting the regions of the images that influence different outcomes

Fig. 3 Algorithm for the proposed workflow

layer. In the BiGRU layers, we have considered ReLU activation for introducing non-linearity. The kernel regularizer and bias regularizer are set to 0.025 at the L2 level. We utilized the flattened layer to provide to the ANN after passing three BiGRU layers. Any neuron in the dense layer of a neural network receives feedback from every neuron in the network due to the layer's strong connections. Two fully connected dense layers were employed to create the ANN in this instance. The first layer's unit is 256 while the second layer's unit is 128. We utilized the activation function "Relu" in the first dense layer and "Softmax" in the last dense layer. As the image dataset contains multiclass data, therefore the Softmax activation function is also deployed. All the hyperparameters are fine-tuned. The regularizer function kernel regularizer is applied to the kernel weights matrix for the first layer. Moreover, a bias regularizer function was applied to the bias vectors. The parametric details are shown in Table 2.

Table 2 Parametric details of the proposed model

Name of the parameters	Estimated value
Batch Size	256
Learning rate	0.0025
Optimizer	Adam optimizer
Vector dimensionality	256
Epsilon	1e-05
Activation function	ReLU, Softmax
Loss function	Categorical cross-entropy
Recurrent dropout	0.005
Number of epochs	10
Train and test ratio	70: 30
Number of trainable parameters	458,152
Number of Bi-GRU units	256,128

2.2 Dataset description

A total of three datasets are utilized in this research. The following two out of those three are built by the authors of this paper-

- JobSn—Data collected through scrapping from glassdoor.com
- BdHikes—Social media comments regarding the recent price hikes in Bangladesh.

The following dataset is collected from external sources:

- NumtaDb—Bangla handwritten character dataset.
- *The JobSn Dataset:* We gathered current or former employee reviews from glassdoor.com, which were written to express satisfaction or dissatisfaction with some renowned Bangladeshi IT organizations that they worked for. We col- lected the data by means of web scraping. The scraped data required further preprocessing to make the data usable. Figure 4 shows the data scraping process:

The dataset contained a total of 11,206 reviews. The "Rating", "Pros", and "Cons" columns are the three main data attributes in the dataset. Ratings are given on a scale of 1–5. Sometimes the ratings are found as fractional numbers. However, the numbers are rounded to the closest decimal points in the preprocessing stage. The gathered information are then categorized into five polarity values, where the annotation is carried out based on the ratings. A review marked with a five-star rating is noted as "Super Positive", whereas a one-star rating is deemed as "Super Negative". There are also four more levels of polarity in between. The annotated polarity for each rating along with their value counts are demonstrated in Table 3.

(2) *The BdHikes Dataset:* Bangladesh is currently undergoing an enormous wave of price hikes. A greater portion of the population is baffled by the increased prices of daily commodities. This crisis left people with no clue, as even the most basic expenses now sum up to cost more than they get to earn. People have been constantly discussing the crisis and sharing their frustrations regarding this issue. We have built a dataset that contains those comments that people posted in the

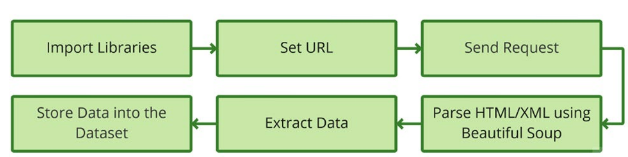


Fig. 4 Data scraping steps using Python

Table 3 Polarity value counts for the JobSn dataset

Corresponding polarity	Rating	Total value counts
Super positive	5.00	534
Positive	4.00	1363
Neutral	3.00	3121
Negative	2.00	3534
Super negative	1.00	2654

comment section of the most popular electronic and print media of Bangladesh. We initially pick up some of the most popular media in Bangladesh. Then the data is gathered from the Facebook and YouTube comment sections of those media. Figure 5, represents a partial view of the collected dataset. The dataset contains only three sentiment polarities—Positive, Negative, and Neutral.

Table 4 shows the polarity counts for each sentiment in the gathered dataset. In total 12,234 comments are present.

(3) *The NumtaDb Dataset:* This dataset contains handwritten bangla digits ranging from 0 to 9. Some example images of this dataset are provided in Fig. 6. There are in total 746 images in the NumtaDb dataset with varying styles. This dataset is used to evaluate how robust and precise our model can be in terms of image classification. The source of these images are kaggle.com

Comments	Polarity	News Source
"আর কত এভাবে চলা যায়"	Negative	Collected from BBC Bangla facebook page.
"বেঁচে থাকাই দায় হয়ে পড়েছে"	Negative	Collected from SomoyTV online portal.
"চিরদিন তো দাম একই থাকবেনা। দাম বাড়বেই"	Positive	Collected from The Daily Star facebook page.
"জীবন আর চলেনা"	Negative	Collected from New Age online portal.
"মানিয়ে নেয়া ছাড়া উপায় নেই"	Neutral	Collected from Independent TV online portal.

Fig. 5 Partial view of the BdHikes dataset

Table 4 Polarity value counts for the BdHikes dataset

Sentiments	Polarity counts
Positive	1353
Negative	7637
Neutral	3244

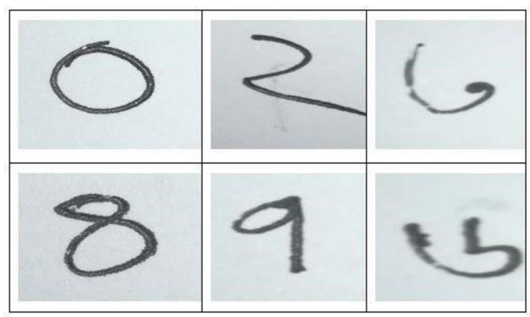


Fig. 6 Some instances from the NumtaDb dataset

2.3 Data preprocessing

Both text and image datasets are used in order to test what our proposed model is capable of. Data preprocessing tasks are carried out differently as we need to deal with both text and image data.

(1) *Text Data Preprocessing:* Both JobSn and BdHikes datasets are preprocessed first, while most of the preprocessing steps are common for them. These preprocessing steps are undertaken to increase the accuracy of the model along with a reduction in the execution time. Figure 7 shows the preprocessing steps performed for text datasets.

Punctuation Removal: When it comes to giving expressions an emotive vibe, punctuation marks might be crucial. However, punctuation marks don't provide much value when it comes to understanding the text. Moreover, data are more stan- dardized and focused as the punctuation marks are removed, making it easier to analyze and process.

Stopwords Removal: Stopwords are some common and frequently seen words that are often used in a sentence without having any significant role in creating the central theme of the text. These words can cause a nuisance in the model if not taken care of timely. As the stopwords get removed from the data, the model now just needs to focus on the words that are significant. Thus, the removal of stopwords can serve to reduce the amount of noise and enhance focus, thereby enabling the model to produce better efficiency and accuracy.

Hyperlink and Emoji Removal: When conversing with someone on social media, emoticons are essential for conveying situational emotion. However, the emoticons are often misleading to convey the underlying emotion when

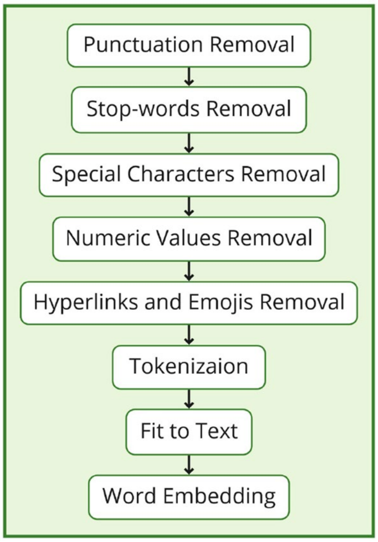


Fig. 7 Preprocessing steps for JobSn and BdHikes datasets

used arcastically or ironically and hence do not bear the essence of the text. Hyperlinks in the comments also do not offer anything useful for training a dataset. Hence both hyperlinks and emojis should be disposed of before the remnant text goes ahead for training the model.

Digits and Special characters Removal: More often than not, numerical values do not offer contexts that can be utilized to categorize contents. In this study, the digits in the collected data do not seem to bear much significance, hence discarded. Special characters are also eradicated for the same reasons.

Tokenization: Tokenization is the process of breaking down a large piece of text into smaller, significant units called tokens. This is a crucial step in natural language processing (NLP) as it allows the computer to understand and analyze the meaning of words within a text. By dividing a large block of text into smaller, more manageable chunks, NLP algorithms can more easily identify patterns and relationships between words, which can be used for sentiment analysis, text classification, and many more.

Word Embedding: Word embedding technique is used in NLP to represent words as numerical vectors in a high-dimensional space. The idea is to capture the meaning and context of words in a way that can be easily processed by machine learning models. In a word embedding, each word is assigned a vector of real-valued numbers. It is described as a type of numeric vector input that enables words with related meanings to share a common representation. There are sev- eral methods for generating word embeddings. The methods include neural network-based techniques like Word2Vec and GloVe, as well as probabilistic methods like Latent Dirichlet Allocation (LDA).

To demonstrate all the preprocessing steps clearly, a review is considered-

"Good Environment with a decent salary!"

At first, the punctuation marks are removed from the sentence. Then the sentence turns into the following-

"Good Environment with a decent salary"

Now all the characters are converted into lowercase charac- ters. After that, the sentence yields the following pattern-

"good environment with a decent salary"

Stopwords, hyperlinks, and emoticons need to be discarded as they don't bear much significance for our hybrid architec- ture. After eradicating stopwords, hyperlinks, and emoticons, the sentence turns into-

"good environment decent salary"

Unnecessary whitespaces are also removed from the sentence. Finally, the sentence is converted into the below format after applying tokenization:

[["good"], ["environment"], ["decent"], ["salary"]]

After completing all these steps, we have applied word embedding into the sentence and FitToText. This enables assigning unique numerical values for each unique word.

[[2342], [1242], [4353], [435]]

In the case of BdHikes dataset, same techniques are applied to convert raw text into numerical values.

(2) *Image Data Preprocessing:* Before providing the dataset to our proposed model, necessary preprocessing is required, so that the proposed model can extract features properly.

Image Resizing: All the images contained in the Num-taDb dataset are not of the same size. To get rid of this problem, all the images are resized into 256 × 256 pixels.

Image Denoising: Several image denoising techniques are available for removing unnecessary noises from images. Gaus- sian filter blurs images along with removing some details of an image. Taking this thing into account, Gaussian filter is applied to the image dataset as it contains images with unnecessary details.

Image Augmentation: Image Augmentation is a technique that can increase the amount of data available for training.

It performs various transformation techniques on existing images such as scaling, flipping, shifting, modifying brightness or contrast, etc. The new set of images has some variations with the original data, but still contains the essence of them. A bigger number of training samples enables the model to learn better about the features and thus avoid overfitting. Table 5 represents the image augmentation parameters for this research.

After performing all the necessary preprocessing, the image dataset is converted into vectors. Supervised labelling is ap- plied in the dataset. All the images are represented in 28 × 28 vectors.

2.4 Training set up

Tensorflow and Keras libraries are used in this research to implement the BiGRU and ANN layers. For visualization purposes, both Seaborn and Matplotlib.pyplot are used. In the case of scraped data, HTML/XML is parsed with the help of Beautiful Soup. For images, OpenCV library is utilized so that all the necessary preprocessing tasks can be accomplished precisely. Pandas library is utilized for manipulating and maintaining corpus. For training purposes, the authors have used an Intel Core i7 processor with a speed of 3.8 GHz. The hardware system also consists of 16 GB RAM along with 1 TB SSD. For GPU, NVIDIA GTX 1660TI is utilized.

3 Experimental result analysis

After setting all the necessary tools, we have observed the results shown by the proposed BiGRU-ANN model. In both image and text cases, our proposed model has outperformed the state-of-the art architectures.

3.1 Classification performance analysis for the JobSn dataset

Table 6 demonstrates the outstanding performance of the proposed model for the JobSn dataset.

Table 5 Parametric details for image augmentation

Parameters	Value
Scaling	1/0.255
Zoom-range	0.2
Shear-range	0.2
Rotation-range	40
File-mode	Nearest
Horizontal-flip	True
Vertical-flip	True

The integration of BiGRU and ANN enables the model to understand complex data patterns. When the experimental results of the proposed model are compared against that of the other state-of-the-art ML and DL architectures, it gives a clearer view of how our model performs better than the pre- viously performed research. The state-of-art- architectures are focused on numerous deep learning architectures. A detailed comparison has been performed between the proposed and state-of-the-art architectures [36–41].

Figure 8 shows that, Our proposed hybrid model has shown an accuracy of 98.75% where the lowest accuracy has been shown by Maximum Entropy (ME). On the other hand, the closest accuracy has been shown by the Naive-Bayes.

Table 6 Performance analysis for the JobSn dataset

Set Name	Precision	Recall	F1-Score	Accuracy
Training	0.9646	0.9573	0.9464	0.9632
Validation	0.9732	0.9578	0.9686	0.9734
Testing	0.9693	0.9628	0.9784	0.9875

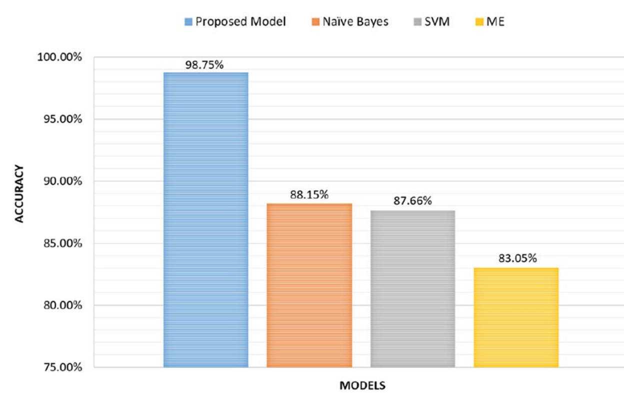


Fig. 8 Accuracy comparison of the proposed model with state- of-the-art ML architectures

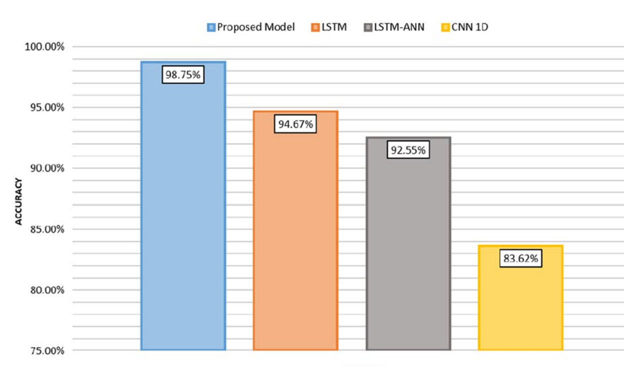


Fig. 9 Accuracy comparison of the proposed model with state- of-the-art DL architectures

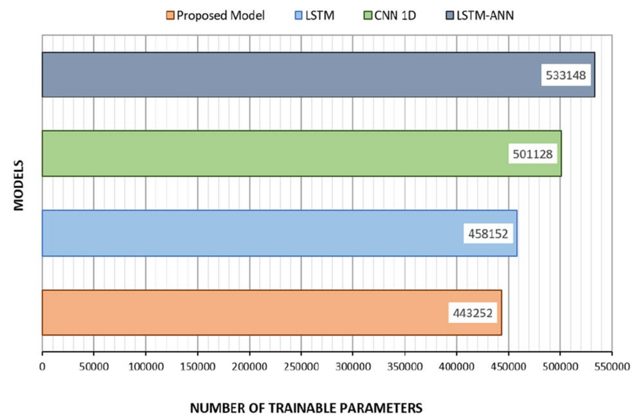


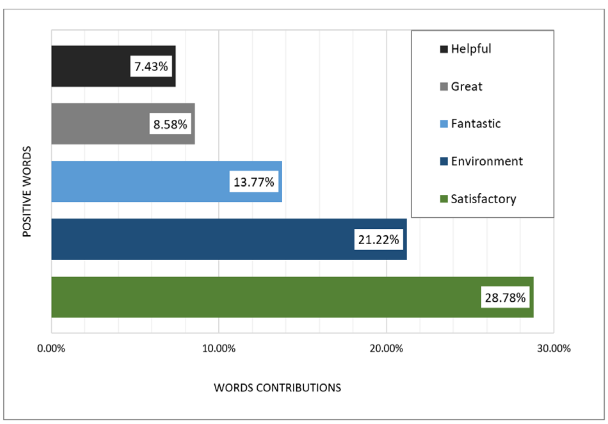
Fig. 10 Comparison for the number of trainable parameters with state-of-the-art DL models

Here we compare our model with the recent DL architectures proposed by other researchers. Figure 9, represents that the proposed BiGRU-ANN has also shown supremacy over other DL architectures in terms of the accuracy.

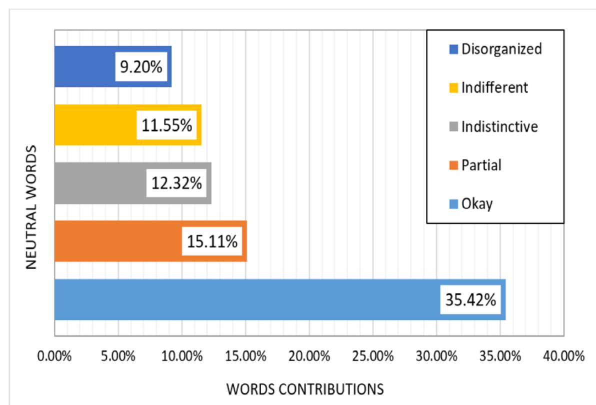
The proposed hybrid architecture not only produces higher accuracy in classification, it is computationally less expensive as well. A comparison of the number of trainable parameters is demonstrated in Fig. 10.

3.2 Performance analysis for the BdHikes dataset

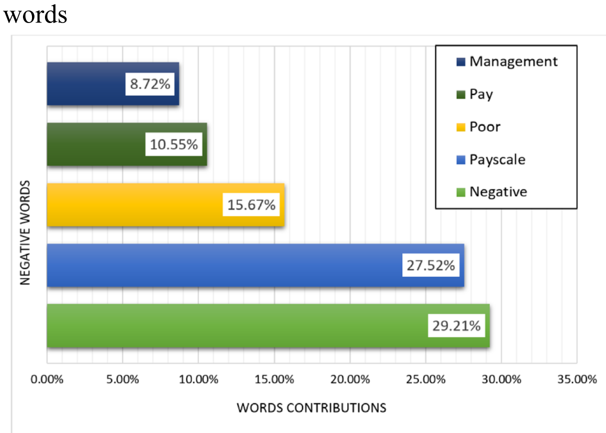
The proposed architecture also maintains its performance for the BdHikes dataset (Fig. 11). The data loss between the training set and the validation set is very small, which implies that the model learned the underlying patterns and relationships in the training data well, and has generalized well to unseen data in the validation set. This is a good sign, as it suggests that the model did not overfit the training data and learned to make nearly accurate predictions on new data. In Fig. 12, demonstrates the comparative analysis of the DL architectures with the proposed architecture according to F1-Score.



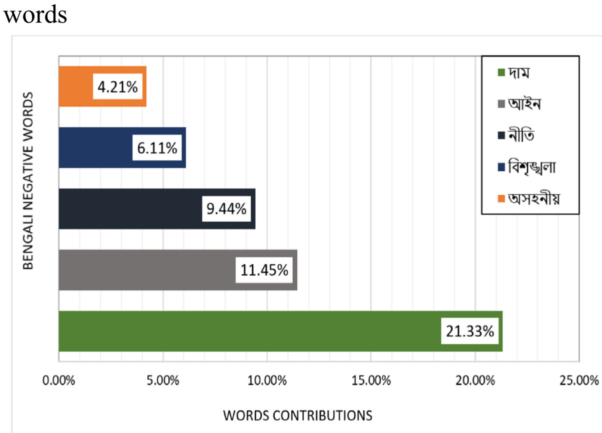
(a) Significant words analysis using LIME for positive words



(b) Significant words analysis using LIME for neutral words



(c) Significant words analysis using LIME for negative words



(d) Significant words analysis using LIME for Bengali words

Fig. 11 Interpretation of Bengali and English text data using LIME

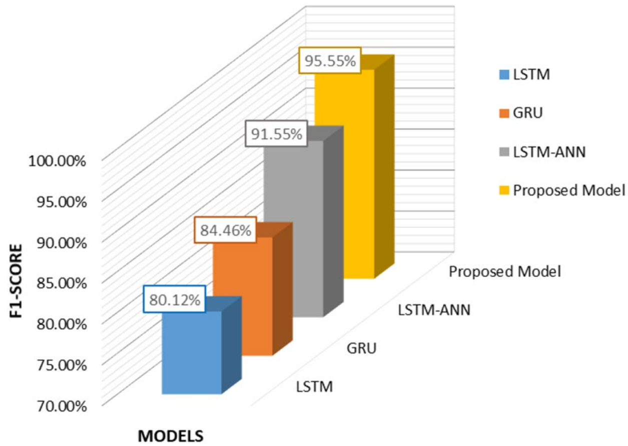


Fig. 12 Sentiment classification comparison with the DL architectures

Table 7 Classification report for the BdHikes dataset

Metrics	Positive	Negative	Neutral
Precision	0.9847	0.9375	0.9473
Recall	0.9943	0.9463	0.9373
F1-score	0.9978	0.9275	0.9453

One shortcoming of the proposed architecture is, sometimes it gets confused between Neutral and Negative polarity. As in Bengali, it is often difficult to discriminate between Neutral and Negative expressions. The classification report is shown in Table 7.

3.3 Image classification result analysis

The dataset contains images of the digits from 0 to 9. There are a total of ten classes. The proposed integrated model has shown impressive results in identifying images for each of these classes. Figure 13, plots the training and validation accuracy for each epoch. The precision and recall from the classification reports are presented in Table 8. Higher precision means that the model has a low false positive rate. High recall, on the other hand, means that the model has a low false negative rate.

3.4 Interpreting the outcomes using explainable AI (XAI)

Local Interpretable Model-agnostic Explanations (LIME) is a popular XAI technique [42] that generates explanations of the predictions made by AI models that are otherwise taxing to understand. LIME works by approximating the decision surface of a model in the vicinity of a prediction using a simple, explicable model such as a decision tree or a linear model. This local approximation provides a basis for the prediction in terms of the input features. The degree of detail in the explanation can be tweaked by varying the size of the neighborhood surrounding the prediction [43].

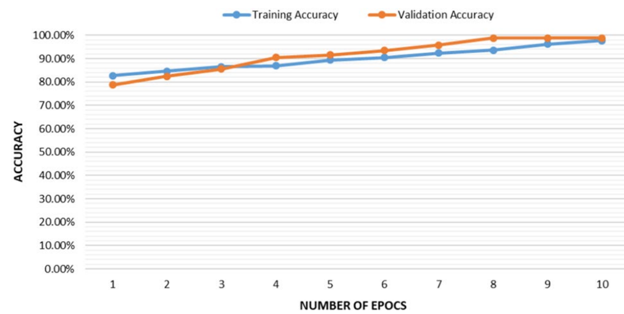


Fig. 13 Training and testing accuracy comparison for each of the epochs

Table 8 Precision and recall analysis for the NumtaDb dataset

Class name	Precision	Recall
0	0.9847	0.9375
1	0.9856	0.9764
2	0.9985	0.9763
3	0.9936	0.9746
4	0.9723	0.9385
5	0.9853	0.9854
6	0.8795	0.9275
7	0.9546	0.9854
8	0.9964	0.9964
9	0.9896	0.9785

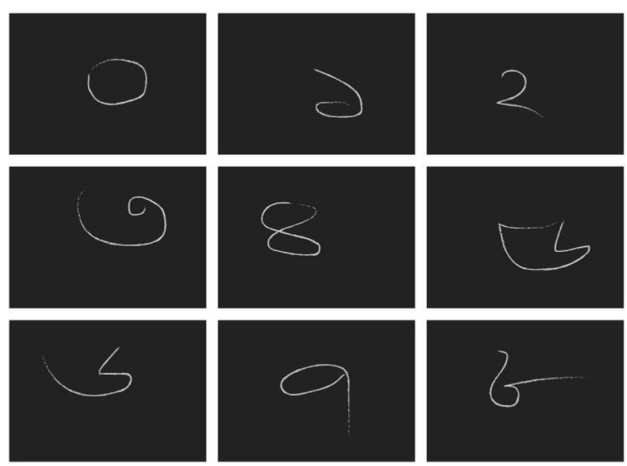


Fig. 14 Featured image interpretation with LIME

LIME has been found to be applicable to a wide range of domains with a comprehensive and useful level of explanations for the decisions made by machine learning models [44]. Findings by LIME on three polarity values namely Positive, Negative, and Neutral from the JobSn dataset are shown in Fig. 11. LIME extracts the most significant words for these polarity values. The same technique is applied on the BdHikes dataset which consists of Bengali words. This is useful for tasks such as sentiment analysis, and text or document classification.

LIME can also be used to explain predictions made by models that work with image data. In the case of image data, LIME can be used to generate heatmaps that highlight the regions of the image that are the most important for the prediction about to be made by the model. This can help to understand how the model is recognizing objects or features in the image.

Figure 14 represents the feature extraction of the images from the NumtaDb dataset using LIME. LIME creates a number of altered versions of the original image, each with a minor adjustment like noise addition or image cropping. In our case, LIME highlights the important regions of the handwritten digits, which influence the model's decision.

4 Conclusion and future work

This paper proposes a hybrid deep learning architecture combining two models named BiGRU and ANN. The model is capable of rendering remarkable outcomes in terms of text and image classification. we have used three datasets, where two of which are concerned with real-life expressions or comments from people. The datasets are gathered from Glassdoor.com and social media platforms respectively. A popular Bengali handwritten character dataset has also been used in this research. In all cases, the proposed hybrid architecture out- performed the state-of-the-art models. The proposed BiGRU-ANN architecture can be trained with more data in an attempt to achieve even better performance. For image classification, this model can further be integrated with Convolutional Neural Network (CNN).

Author contributions SC: conceptualization, writing—editing, methodology, coding. MBUT: writing—editing, methodology, data visualization. MMH: data collection, data preprocessing, data curation. JN: supervision, writing—reviewing. JU: supervision, writing—reviewing.

Funding This research is funded by Woosong University Academic Research 2023 (Grant JIA UDDIN).

Data and code availability The necessary data and code are available in this given github link. https://github.com/SovonChakraborty/Paper-Code-and-Data.

Declarations

Conflict of interest The authors affirm that none of their known financial conflicts of interest or close personal connections might have seemed to affect the research presented in this study.

Ethical approval Authors declare that, they did not violate any ethics of research while conducting the research.

Consent for publication If the paper is accepted for publication, the authors provide consent and maintain the policy of this journal.

References

1. Arias F, Núñez MZ, Guerra-Adames A, Tejedor-Flores N, Vargas Lombardo M (2022) Sentiment analysis of public social media as a tool for health-related topics. IEEE Access 10:74850–74872
2. Li W, Shao W, Ji S, Cambria E (2022) BiERU: bidirectional emotional recurrent unit for conversational sentiment analysis. Neurocomputing 467:73–82
3. Yu LC, Lee CW, Pan HI, Chou CY, Chao PY, Chen ZH et al (2018) Improving early prediction of academic failure using sentiment analysis on self-evaluated comments. J Comput Assist Learn 34(4):358–365
4. Xu G, Meng Y, Qiu X, Yu Z, Wu X (2019) Sentiment analysis of comment texts based on BiLSTM. Ieee Access 7:51522–51532
5. Hassan A, Amin MR, Al Azad AK and Mohammed N (2016) Sentiment analysis on bangla and romanized bangla text using deep recurrent models. In: 2016 International Workshop on Computational Intelligence (IWCI) (pp 51–56). IEEE
6. Yao Y, Sullivan T IV, Yan F, Gong J, Li L (2022) Balancing data for generalizable machine learning to predict glass-forming ability of ternary alloys. Scripta Mater 209:114366
7. Tuhin RA, Paul BK, Nawrine F, Akter M and Das AK (2019) An automated system of sentiment analysis from Bangla text using supervised learning techniques. In 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS) (pp 360– 364). IEEE
8. Abbasimehr H, Paki R, Bahrini A (2022) A novel approach based on combining deep learning models with statistical methods for COVID- 19 time series forecasting. Neural Comput Appl 34(4):3135–3149
9. Dollen DV, Neukart F, Weimer D, Bäck T (2023) Predicting vehicle prices via quantum-assisted feature selection. Int J Inf Technol. https://doi.org/10.1007/s41870-023-01370-z
10. Yan W, Zhou L, Qian Z, Xiao L, Zhu H (2021) Sentiment analysis of student texts using the CNN-BiGRU-AT model. Sci Program 2021:1–9
11. Kobayashi M, Nakaji K, Yamamoto N (2022) Overfitting in quantum machine learning and entangling dropout. Quant Mach Intell 4(2):1–9
12. Chiong R, Fan Z, Hu Z, Dhakal S (2022) A novel ensemble learning approach for stock market prediction based on sentiment analysis and the sliding window method. IEEE Trans Comput Soc Syst. https://doi.org/10.1109/TCSS.2022.3182375

13. Kalarani P, Selva Brunda S (2019) Sentiment analysis by POS and joint sentiment topic features using SVM and ANN. Soft Comput 23:7067–7079
14. Mikolov T, Grave E, Bojanowski P, Puhrsch C & Joulin A (2017) Advances in pre-training distributed word representations. arXiv preprint. arXiv:1712.09405
15. Wadud MAH, Kabir MM, Mridha MF, Ali MA, Hamid MA, Monowar MM (2022) How can we manage offensive text in social media-a text classification approach using LSTM-BOOST. Int J Inf Manage Data Insights 2(2):100095
16. Faruque MA, Rahman S, Chakraborty P, Choudhury T, Um JS, Singh TP (2022) Ascertaining polarity of public opinions on Bangladesh cricket using machine learning techniques. Spatial Inf Res 30(1):1–8
17. Shamrat FMJM, Chakraborty S, Imran MM, Muna JN, Billah MM, Das P, Rahman OM (2021) Sentiment analysis on twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm. Indonesian J Electric Eng Comput Sci 23(1):463–470
18. Chinnasamy P, Suresh V, Ramprathap K, Jebamani BJA, Rao KS, Kranthi MS (2022) COVID-19 vaccine sentiment analysis using public opinions on Twitter. Mater Today 64:448–451
19. Lazaridis A, Vlahavas I (2023) GENEREIT: generating multi-talented reinforcement learning agents. Int J Inf Technol 15(2):643–650
20. Seyyar YE, Yavuz AG, Ünver HM (2022) An attack detection framework based on BERT and deep learning. IEEE Access 10:68633–68644
21. Madsen A, Reddy S, Chandar S (2022) Post-hoc interpretability for neural NLP: a survey. ACM Comput Surv 55(8):1–42
22. Lauriola I, Lavelli A, Aiolli F (2022) An introduction to deep learning in natural language processing: models, techniques, and tools. Neurocomputing 470:443–456
23. Sharif O, Hoque MM & Hossain E (2019) Sentiment analysis of Bengali texts on online restaurant reviews using multinomial Näıve Bayes. In: 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT) (pp 1–6). IEEE
24. Abdelgwad MM, Soliman THA, Taloba AI, Farghaly MF (2022) Arabic aspect based sentiment analysis using bidirectional LSTM based models. J King Saud Univ-Comput Inf Sci 34(9):6652–6662
25. Saleh H, Mostafa S, Gabralla LA, Aseeri AO, El-Sappagh S (2022) Enhanced Arabic sentiment analysis using a novel stacking ensemble of hybrid and deep learning models. Appl Sci 12(18):8967
26. Tripathi K, Khan FA, Khanday AMUD, Nisa KU (2023) The classification of medical and botanical data through majority voting using artificial neural network. Int J Inf Technol. https://doi.org/10.1007/s41870-023-01361-0
27. Youbi F, Settouti N (2022) Analysis of machine learning and deep learning frameworks for opinion mining on drug reviews. Comput J 65(9):2470–2483
28. Swathi T, Sudha S (2023) Crop classification and prediction based on soil nutrition using machine learning methods. Int J Inf Technol. https://doi.org/10.1007/s41870-023-01345-0
29. Kowsher M, Tahabilder A, Sanjid MZI, Prottasha NJ, Uddin MS, Hossain MA, Jilani MAK (2021) LSTM-ANN & BiLSTM-ANN: hybrid deep learning models for enhanced classification accuracy. Procedia Comput Sci 193:131–140
30. Kowsher M, Sami AA, Prottasha NJ, Arefin MS, Dhar PK, Koshiba T (2022) Bangla-BERT: transformer-based efficient model for transfer learning and language understanding. IEEE Access 10:91855–91870
31. Habbat N, Anoun H, Hassouni L (2022) Combination of GRU and CNN deep learning models for sentiment analysis on French customer reviews using XLNet model. IEEE Eng Manage Rev 51(1):41–51
32. Zhou X, Ma R, Zou Y, Chen X, Gui T, Zhang Q et al. (2022) Making parameter-efficient tuning more efficient: a unified framework for classification tasks. In: Proceedings of the 29th International Conference on Computational Linguistics. pp 7053–7064
33. Lindauer M, Eggensperger K, Feurer M, Biedenkapp A, Deng D, Benjamins C et al (2022) SMAC3: a versatile Bayesian optimization package for hyperparameter optimization. J Mach Learn Res 23:54–61
34. Alam F, Hasan A, Alam T, Khan A, Tajrin J, Khan N & Chowdhury SA (2021) A review of Bangla natural language processing tasks and the utility of transformer models. arXiv preprint arXiv:2107.03844
35. Sakiba SN, Shuvo MMU, Hossain N, Das SK, Mela JD & Islam MA (2021) A memory-efficient tool for Bengali parts of speech tagging. In: Artificial intelligence techniques for advanced computing applications: proceedings of ICACT 2020. Springer Singapore, pp 67–78
36. Wadud MAH, Mridha M, Rahman MM (2022) Word embedding methods for word representation in deep learning for natural language processing. Iraqi J Sci 63:1349–1361
37. Chakraborty S, Talukdar MBU, Adib MYM, Mitra S & Alam MGR (2022) LSTM-ANN based price hike sentiment analysis from Bangla social media comments. In: 2022 25th International Conference on Computer and Information Technology (ICCIT) (pp 733–738). IEEE
38. Mahdaoui AE, Ouahabi A, Moulay MS (2022) Image denoising using a compressive sensing approach based on regularization constraints. Sensors 22(6):2199
39. Timofte R, Tuytelaars T & Van Gool L (2012) Naive bayes image classification: beyond nearest neighbors. In: Asian Conference on Computer Vision. Berlin: Springer, pp 689–703
40. Subramanian AS, Weng C, Watanabe S, Yu M, Yu D (2022) Deep learning based multi-source localization with source splitting and its effectiveness in multi-talker speech recognition. Comput Speech Lang 75:101360
41. Abiodun OI, Jantan A, Omolara AE, Dada KV, Umar AM, Linus OU et al (2019) Comprehensive review of artificial neural network applications to pattern recognition. IEEE Access 7:158820–158846
42. Jiang P, Suzuki H, Obi T (2023) XAI-based cross-ensemble feature ranking methodology for machine learning models. Int J Inf Technol 15(4):1759–1768
43. Dieber J & Kirrane S (2020) Why model why? Assessing the strengths and limitations of LIME. arXiv preprint arXiv:2012.00093.
44. Das A & Rad P (2020) Opportunities and challenges in explainable artificial intelligence (xai): a survey. arXiv preprint arXiv:2006.11371.