

Sentiment Analysis of Nepali COVID-19 Tweets using BERT-LSTM

Mamata Tharu¹, Sitaram Pokhrel¹, Badri Raj Lamichhane¹¹Department of Electronics and Computer Engineering, IOE, Pashchimanchal Campus, Tribhuvan University, Nepal

(Manuscript Received: 12/09/23; Revised: 26/10/2023; Accepted: 02/11/2023)

Abstract

The global impact of COVID-19 has significantly reshaped the day-by-day lives of individuals worldwide. COVID-19 is one of the top deadly diseases and has tragically claimed the lives of millions across the globe. The people are affected not only by the physical infection but also mentally. Among the various social media platforms, Twitter is a widely utilized medium, reflecting a substantial surge in discussions about the coronavirus. These discussions encompass a spectrum of positive, negative, and neutral sentiments. The sentiments acknowledged by individuals, encapsulated in their posts and tweets across this platform, offer valuable insights into their emotional states and perspectives. In this investigation, the people's sentiments using Nepali COVID-19-related Twitter datasets are inspected. For this, the approach involves a two-step process. Initially, the multilingual BERT(m-bert) model will utilize whose output is used for subsequent downstream tasks. Secondly, m-BERT's output is connected to the LSTM layer to categorize the people's sentiments. The model was trained and tested using publically available NepCov19Tweets datasets. These tweets were split into three groups (positive, negative, and neutral). The appraisal outcomes for NepCov19Tweets demonstrate that the proposed model comes up with outstanding performance when compared to the existing model, achieving an average accuracy of 76.04%, 80.03% recall, a precision of 77.12%, and an F1-score of 76%.

Keywords: Sentiment analysis; COVID-19; BERT; BERT-LSTM; LSTM; Machine learning; Nepali

1. Introduction

Sentiment analysis is a text classification technique that can be used to classify text at the aspect, sentence, or document level in natural language processing research. This field of study analyzes people's views, attitudes, feelings, and appraisals towards realities and their attributes expressed in written text. Businesses, governments, and experiments have found sentiment analysis to be a powerful tool for extracting and analyzing public moods and views, gaining business insights, and making better decisions [1]. Sentiment classification is the process of automatically classifying or identifying a text document's polarity such as neutral, negative, or positive.

In the past, due to COVID-19, social media posts, especially tweets, increased rapidly. Much research [2-4] has been conducted in the area of sentiment analysis related to COVID-19 tweets for languages such as English and other widely spoken languages. Still, there has been a relative lack of research on the sentiment analysis of tweets connected to the COVID-19 in Nepali language. Existing work [5-8] used BERT for high-level languages such as English. The model

trained for the English language did not apply to low-resource languages like Nepali, a Devnagari Script. There is no study of sentiment analysis on a combination of BERT-LSTM for Nepali COVID-19-related tweets to analyze people's sentiments in literature; therefore, the ideas generated for sentiment analysis work for the Nepali language. The approach discussed in [9] fails to capture the sequential patterns within sentences and fastText's pre-trained method falls short of effectively understanding the context between words in complex tweets. Additionally, it tends to overlook important subword details, which are crucial for capturing the finer nuances of language.

To address the issue mentioned earlier, this research utilizes LSTMs, which can help to solve the challenges of sequences in data by understanding long patterns and contextual word embedding BERT which handles subwords information better by using WordPiece tokenization. The BERT's bidirectional nature allows it to comprehend how words relate in context. BERT is strong in NLP tasks due to its complex structure and strong ability to learn nonlinear representations from large amounts of text data [10]. LSTM plays a fundamental role in enhancing effectiveness by memorizing as well as recognizing essential information patterns [8]. This is why, this research integrates contextualized word representations from

*Mamata Tharu. Tel.: +977- 9867232773,
E-mail address: tharumamata0123@email.com

BERT into LSTM, effectively enhancing the sentiment of people. Their remarkable ability to capture meanings and distant connections in tweets greatly contributes to this improvement.

The key contributions of the work are outlined below:

- This research is the first to combine BERT and LSTM for analyzing sentiments in Nep-COV19Tweets datasets.
- The stated method is evaluated against the best current approaches using a same dataset, and a proposed approach performs better.

The remaining part of this paper is assembled as follows: In section 2, we look at what others have done with COVID-19 tweets, BERT, and BERT-LSTM for sentiment analysis. In Section 3, the proposed method's concept is explained in detail. In section 4, describe the data used, share the results of the research's experiments, and compare the performance of the presented approach against the state-of-the-art methods. Finally, the study wraps up in section 5, and considers potential developments in the future.

2. Related Work

Many studies [1], [5-7] have explored sentiment analysis using BERT in the English language on various datasets. However, there is a lack of research concerning COVID-19 tweet datasets, especially in the context of the Nepali language. Examination of sentiment in the English and Nepali language using machine and deep learning has been the subject to numerous studies [1], [4], [5], [7], [9], [11], [12].

In this [7] study, the author investigates sentiment analysis using two fine-tuned BERT models on the "vreviews" and "ntc-sv" datasets. The first method employs only the [cls] token with a neural network, while the second method utilizes all BERT output vectors for classification. Fine-tuning BERT with LSTM or TextCNN did not significantly enhance performance against the base fine-tuning strategy. However, employing the proposed fine-tuning approach in the BERT-RCNN model led to improved accuracy in sentiment analysis on the "vreviews" datasets. For the "vreviews" dataset, the BERT-RCNN model achieved an F1 score of 88.22 percent, while for the "ntc-sv" dataset, the F1 score reached 91.15 percent.

The model proposed in [8] focuses on detecting fake news using the FakeNewsNet dataset, comprising two sub-datasets: "Politifact" and "gossipcop". This method employs BERT with an LSTM layer to create a content-based classifier for news titles. By integrating contextualized word representations from

BERT into the LSTM, the model enhances its ability to classify false news, benefitting from their strong capacity to understand the meaning and long-distance connections within news titles. This novel approach is compared against other classification methods and the vanilla BERT model. The BERT-LSTM model surpasses the vanilla BERT model, exhibiting an accuracy of 88.75 percent for "Politifact" and 84.10 percent for "Gossipcop" datasets.

Mingzheng et.al designed a model to analyze sentiments in Chinese stock reviews using BERT with limited labeled data. As with limited labeled data, traditional methods do not perform well. Therefore, to solve this challenge, they used the BERT pretraining model to understand sentence-level stock reviews. Then, feed this understanding into a classifier for categorization. Among different methods, the BERT+FC approach works best with a small dataset. It outperforms BERT+LSTM and BERT+CNN. The BERT+FC approach achieves a remarkable F1 score of 92.50 percent. A novel approach was formulated by leveraging BERT's intermediate layers to improve the fine-tuning process. While earlier BERT-based methods focus solely on the final output layer and overlook the semantic insights within intermediate layers, the proposed technique addresses this limitation. This approach is also extended to a natural language inference task. The three methods were developed: BERT base, BERT-attention, and BERT-LSTM. Among these, the BERT-LSTM method achieves the highest performance, boasting an accuracy of 90.79 percent [6].

A bidirectional LSTM network called the RU-BiLSTM that is attention-based model was introduced to classify the people's sentiments, specifically tailored for the Roman Urdu datasets (RUECD). The team curated the RUCED datasets and found that their model's performance outperformed. The complexities of sparsity and high dimensionality in Roman Urdu text presented considerable challenges in capturing sentence semantics. To overcome these challenges, they proposed the RU-BiLSTM model to analyze Roman Urdu sentiment. In order to record contextual information in both directions, their model used bidirectional LSTM. Additionally, the attention mechanism was incorporated to emphasize crucial features. The neural word embedding and attention layers bolstered the model's semantic understanding and enhanced its capability to extract patterns. Three different word embeddings were employed: word2vec, GloVe, and fastText. Notably, the model's performance was notably better when using word2vec embedding. When evaluated in comparison to other techniques such as LSTM, RNN, RCNN, and CNN, the proposed model

showcased significant superiority with an impressive 8 percent boost in accuracy. For the RU-BiLSTM model with word2vec embedding, the accuracy reached 67 percent, while recall and F1-score were at 67 percent, and precision stood at 68 percent [1].

Machine learning models (SVM, RandomForest, XGBoost) were constructed using distinct feature extraction methods like term-frequency-inverse document frequency (TF-IDF) and n-grams. These models tried to categorized data in binary (positive, negative) and ternary (positive, negative, and neutral) ways using datasets made up of English tweets concerning COVID-19. A remarkable accuracy of 90% was attained by the XGBoost (XGB) classification algorithm when combined with bigram and unigram features for binary classification. This method was designed to delve into public sentiment concerning COVID-19. By examining social perceptions shared on Twitter, a widely used social platform, the model was able to decipher the sentiments surrounding the pandemic [4]. Manguri et al. [3] gathered a week's worth of tweets (from 09-04-2020 to 15-04-2020) focused on COVID-19. These tweets were then subjected to sentiment analysis using TextBlob. Additionally, Nemes et al. [2] employed a recurrent neural network (RNN) to categorize the tweets into those that are either positive or negative sentiments. They curated a dataset of tweets centered around COVID-19, classifying them into four distinct levels of positivity and negativity: weak positive, strong positive, weak negative, and strong negative.

In the initial stages, the author [13] introduced the utilization of the TF-IDF approach to represent news documents, followed by three algorithms; Support Vector Machine (SVM), Multilayer Perceptron Neural Networks and Naïve Bayes. Among this algorithm, the SVM with RBF kernel demonstrated superior performance with an accuracy of 74.65%, precision of 75.4%, recall of 74.6%, and 74.4% f-score. Nonetheless, their methodology was developed on small-sized datasets, which may necessitate substantial adaption for larger datasets to ensure its applicability. Additionally, their approach primarily captures syntactical information, disregarding the significance of semantic or contextual information. This omission is pivotal for effectively distinguishing complex documents or tweets, where factors like higher interclass similarity and intraclass dissimilarity come into play. The author [12] established aspect-based sentiment analysis using SVM and Naïve Bayes classifiers. The algorithm gathers Nepali text data from several sources and uses part-of-speech tagging to locate relevant aspects and sentiment characteristics, which was done using the TF-IDF

method. Among the algorithms tested, the Bernoulli Naïve Bayes classifier outperforms with an accuracy of 77.5%. Similar to findings [13], this approach predominantly captures syntactical information while disregarding vital semantic and contextual meaning. To enhance the model's performance, expansion to more substantial datasets is imperative due to the current limitation of only 1576 sentences.

Deep learning algorithms for sentiment analysis were presented by Sitaula et al. [9] for Nepali COVID-19-related tweets. "NepCov19Tweets" dataset, which is a benchmark collection of Nepali COVID-19 tweets, was also introduced and is freely available on the Kaggle website. The authors developed and trained three distinct Convolutional Neural Network (CNN) models, each implementing a different approach for text representation: fastText(ft), domain-specific (ds), and domain-agnostic (da). These models were subsequently combined into an ensemble CNN for tweet sentiment classification. The resultant model achieved a classification accuracy of 68.7% when applied to the NepCov19Tweets dataset. However, their approach does encounter a couple of limitations. Firstly, their CNN models are intricate, potentially requiring a lot of processing power to implement. Secondly, as their techniques solely rely on semantic characteristics, they might not effectively capture crucial syntactic information. Recently, Shahi [11] introduced a hybrid-based feature extraction technique for analyzing the tone of Nepali tweets about COVID-19. They used the TF-IDF and fastText to create the hybrid feature which is capable of capturing distinctive attributes effectively. The experiment used three distinct feature presentation techniques: TF-IDF, FastText, and Hybrid, as well as nine well-known machine learning classifiers, including Logistic Regression, SVM, Naïve Bayes, Extreme Tree Classifier, K-Nearest Neighbor, AdaBoost, Decision Trees, and Random Forest.

Exceptionally, the hybrid feature extraction technique displayed greater performance than the two separate feature extraction approaches, across the spectrum of nine diverse machine learning algorithms. Specifically, SVM with an RBF kernel, utilizing hybrid features, attained the highest precision (71.4%), recall (72.1%), F1-score (70.1%), and accuracy (72.1%). This methodology has certain

Table 1: Nepali alphabets and numbers

Consonants	क, ख, ग, घ, ङ, च, छ, ज, झ, ञ, ट, ठ, ड, ढ, ण, त, थ, द, ध, न, प, फ, ब, भ, म, य, र, ल, व, श, ष, स, ह, क्ष, त्र, ज्ञ
Vowels	अ, आ, इ, ई, उ, ऊ, ऋ, ए, ऐ, ओ, औ, अं, अः
Numbers	०, १, २, ३, ४, ५, ६, ७, ८, ९

limitations: employs traditional machine learning algorithms, thus lacking the capability to capture sequential information from sentences.

3. Materials and Method

3.1 Data Collection

The dataset used in this study is the Nep-COV19Tweets dataset, which includes Nepali language tweets on COVID-19. The data was collected from February 11, 2020, to January 10, 2021, and is considered valid as it has been utilized by [9], [11] for the analysis. The dataset consists of a total of 35,789 Nepali sentiment texts, including 14,408 with a negative sentiment, 15,880 with a positive sentiment, and 5,501 with a neutral sentiment. To access the Nep-COV19 Tweets dataset, you can visit [14].

3.2 Methodology

The proposed method involves two steps; preprocessing, and hybridizing the BERT and LSTM algorithm to categorize the opinions expressed at COVID-19 in tweets. The proposed approach is presented in Figure 1.

3.2.1 Data Preprocessing

Data preprocessing refers to the process of converting the raw data into a suitable or understandable format for feature extraction. The data preprocessing steps include:

- Eliminating the special symbols, numbers, and white spaces
- Stopwords are remove like "अक्सर, अगाडि, अझै "

3.2.2 BERT

The BERT is a multilayered model consisting of Bi-directional Transformer encoder layers trained on raw text data. BERT builds contextualized representations by pre-training on a substantial corpus of text [6].

The architecture of BERT is shown in Figure 2. The most used model of BERT is

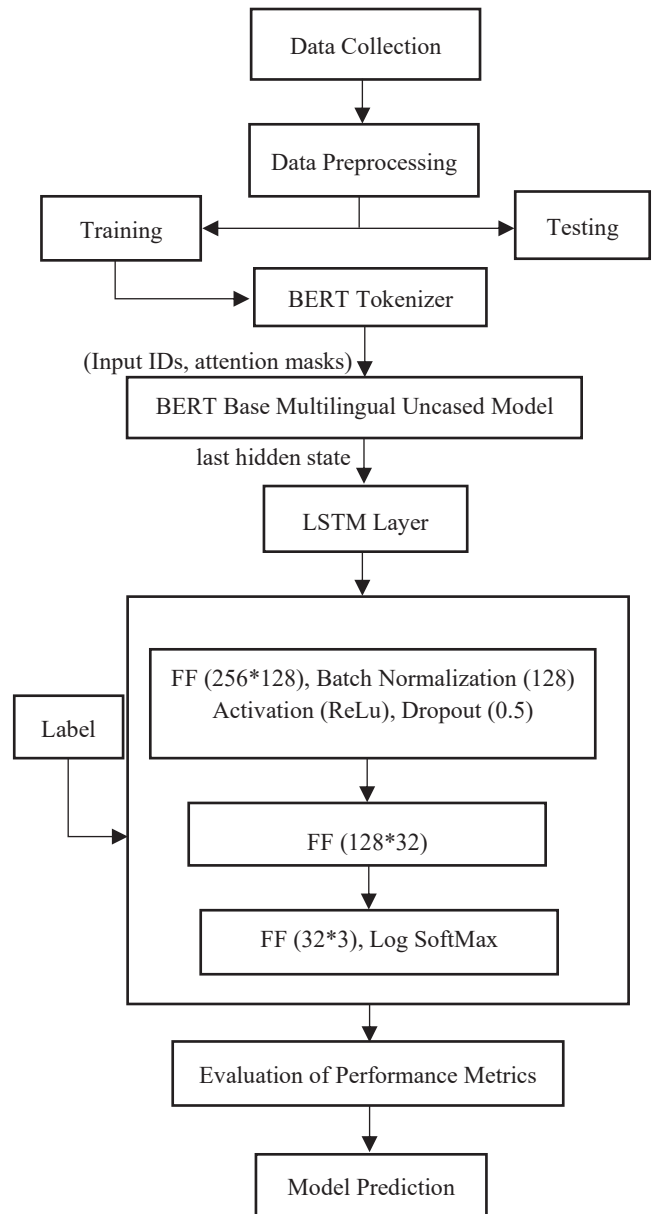


Figure 1: Block diagram of the proposed model

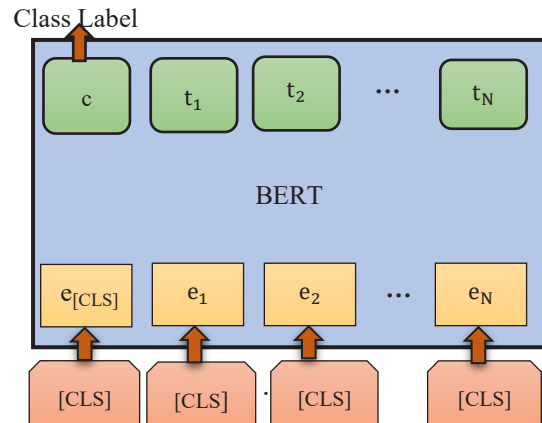


Figure 2: Architecture of BERT[8]

- BERT-base: 12 transformer layers, 768 dimensions, 12 multi-head attention, and 110 million parameters.
- BERT-large: 24 transformer layers, 1024 dimensions, 16 multi-head attention, and 340 million parameters.

Before applying the pre-trained model, the input data must be transformed into proper format i.e. add the special symbol [CLS] at the beginning and [SEP] at the end of each sentence. Each sentence has relevant embeddings that have been found. In this architecture, an inventory of embedded tokens and associated attention masks are input into each encoder layer. The output has been chosen to be equivalent amount of embeddings with the same hidden size. The entire input is represented by a single vector that is fed into the classifier for classification purposes. The entire phrase can be represented for classification purposes by the hidden state of the first token [CLS] of the model's output.

For example: The special symbol [CLS] and [SEP] are added to the sentences.

[CLS] शिक्षक शिक्षा क्षेत्रमा रमाएको मान्छे सामु सरकारी स्कूल कलेजमा पढेर करीब बर्ष बढी [SEP]

The embedded tokens of the above sentences is

[[CLS], [श', '##िक', '##षक', 'शिक्षा', 'कषतर', '##मा', 'र', '##मा', '##एको', 'मान', '##छ', 'स', '##ाम', 'सरकारी', 'सकल', 'कल', '##ज', '##मा', 'प', '##ढ', '##र', 'कर', '##ी', '##ब', 'ब', '##रष', 'ब', '##ढी'], 'SEP']]

Input IDs of the embedded tokens is

[[101, 577, 13559, 43749, 48496, 22003, 12572, 573, 12572, 49068, 91039, 22621, 579, 35877, 95405, 31034, 14372, 14256, 12572, 567, 30119, 11551, 15807, 10824, 17046, 569, 95181, 569, 71595, 102]]

3.2.3 LSTM

The LSTM is a widely used recurrent neural network (RNN) model that is capable of addressing the distance dependence issues of the classical RNNs. LSTM architecture is composed of repeating modules known as cells which consist of four interconnected neural networks. The LSTM cells pass two cell states (C_t) and a hidden state (h_t), to the subsequent cells. The ambition of the LSTM is to retain essential information and manipulate it through three mechanisms called gates i.e., forget gate, input gate, and output gate. The information that is no longer necessary for the LSTM to function is discarded by the forget gate. The input gate is responsible for incorporating new information into the current cell state. Which data from the

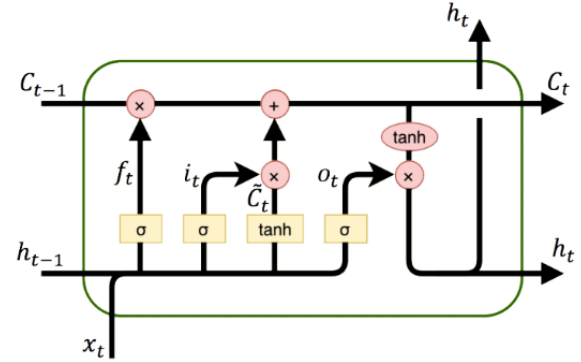


Figure 3: Architecture of LSTM [8]

current cell should be displayed as output is decided by the output gate [8]. The formula of input, forget, output gate, cell and hidden state is shown in Eq. (1), Eq. (2), Eq. (3) and Eq. (6) :

$$i_t = \sigma(A_i[h_{t-1}, x_t] + b_i) \quad (1)$$

$$f_t = \sigma(A_f[h_{t-1}, x_t] + b_f) \quad (2)$$

$$o_t = \sigma(A_o[h_{t-1}, x_t] + b_o) \quad (3)$$

$$\tilde{C}_t = \tanh(A_c[h_{t-1}, x_t] + b_c) \quad (4)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

$$h_t = o_t * \tanh(C_t) \quad (6)$$

3.2.4 Proposed Architecture

In this proposed method, the BERT-base-multilingual-uncased model (m-BERT) has been used with the LSTM layer and feed-forward network. The model was developed by collecting data which is outlined in Section 3.1. The collected information was then thoroughly preprocessed in subsection 3.2.1. The preprocessed text was split into training and testing sets after this preparation stage; the BERT tokenizer was used to efficiently tokenize the text data. The BERT Tokenizer extracts input IDs and attention masks, which are then fed into the m-BERT model and the result is an embedding vector for each token, with a hidden size of 768. The pre-trained m-BERT model is specified as the initial layer of the architecture. The last hidden state or the output of the m-BERT model containing batch size, sequence length, and hidden size information is input to the LSTM layer with an embedding dimension of 256. This LSTM layer processes the BERT output sequence, capturing sequential dependencies and contextual information within the data. BERT offers to contextualize sentence-level representations,

enhancing the LSTM's comprehension of sentence semantics. Thanks to BERT's bidirectional processing, the [CLS] token is encoded through a comprehensive multilayer procedure, encompassing representative details from all tokens. The [CLS] token functions as a consolidated representation, particularly beneficial for classification tasks. As a result, the embeddings associated with the [CLS] token can serve as a holistic sentence representation, which can be fed into the classifier.

The feed-forward (FF) linear layer follows the LSTM layer. The classifier incorporates a feed-forward linear layer with a dimension of 128. To standardize inputs, a batch normalization layer is applied. Additionally, a dropout layer with a rate of 0.5 is introduced to counter overfitting. A 32-layer feed-forward architecture is maintained after that. To categorize incoming tweets into categories that are positive, negative, or neutral, last feed forward layer with an output dimension of three is included. The model's performance metrics are then assessed. In order to determine if a tweet is positive, negative or neutral, unseen text was then supplied to the suggested model.

3.3 Evaluation metrics

In this research, performance evaluation relies on the utilization of accuracy (A), precision (P), recall (R), and f1-score (F) whose formulas are shown in Eq. (7), Eq. (8), Eq. (9) and Eq. (10):

$$A = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$P = \frac{TP}{TP+FP} \quad (8)$$

$$R = \frac{TP}{TP+FN} \quad (9)$$

$$F = 2 * \frac{TP}{TP+FP} \quad (10)$$

3. Results and Discussion

Table 2: Evaluation of NepCOV19Tweets based to classification performance (%) and the performance metrics of the suggested model

Methods	P	R	F	A
Hybrid feature +SVM [11]	71.4	72.1	70.1	72.1
Ensemble CNN [9]	67.2	56.4	56.4	68.7
Proposed method (BERT+LSTM)	77.12	80.03	76	76.04

Precision, Recall, F1-score, and Accuracy are all abbreviations for P, R, F and A

Table 3: Class-specific performance of the proposed model (%)

	P	R	F
Positive	74	68	71
Negative	73	75	74
Neutral	81	85	83

Notice that these models use hyperparameters with learning rates of 1e-5, a batch sizes of 32, 400 training epochs, and the AdamW optimizer.

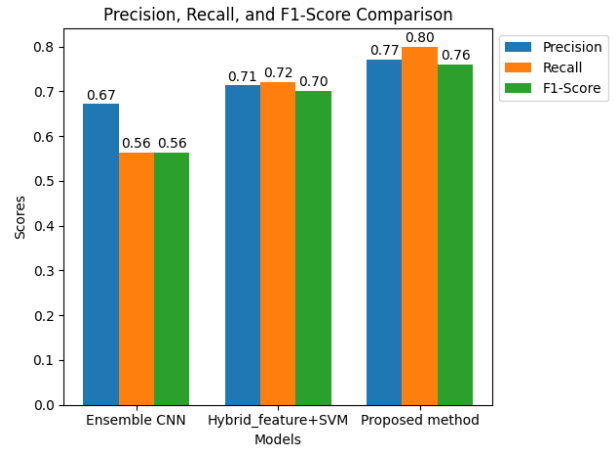


Figure 4: Bar Graph Comparison of the suggested model's performance measures with the present model

The proposed model outperformed the current model in terms of performance as demonstrated in Table 2 and Figure 4. The achieved results were compared with those of existing models such as hybrid feature + (SVM+RBF) [11] and ensemble CNN [9]. Table 2 makes it clear that the suggested model significantly enhances classification accuracy, reaching 76.04%, marking a 7.34% improvement over the least performing method (68.7%), and at least a 3.94% enhancement over the hybrid feature + SVM approach. While observing Table 2, a score improvement ranged from a maximum of 19.6% to a minimum of 5.1%, as compared to an existing model. Additionally, precision shows an enhancement of at least 5.52% and a maximum of 9.92%, while recall exhibits improvements of between 7.93% and 23.63% on a scale of lowest to highest when opposed to the existing model. The planned model outperformed the others by achieving the highest accuracy, precision, recall, and f1 score. This accomplishment was made possible because the proposed model is capable of comprehending the semantic and contextual complexities included in the text data and maintains the significance of word order, where the sequence of words was an essential factor for sentiment analysis.

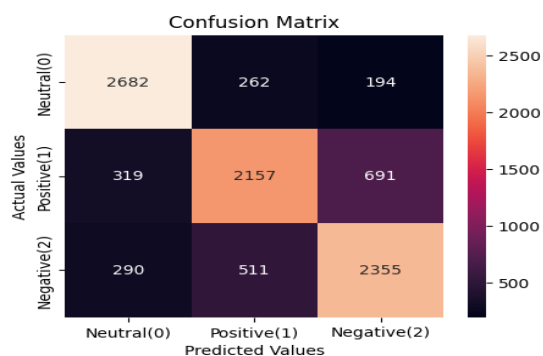


Figure 5: Confusion matrix of the presented model

In Table 3, the model attains the highest f1-score (83%), recall (85%), and precision (81%) for the neutral class as compared to other classes. The model generates a maximum f1-score of 83% for the neutral class and a minimum f1-score of 71% for the positive class. Similarly, the negative class of the model acquires a 74% f1-score.

The model successfully predicted 2,682 tweets were neutral, 2157 were positive and 2355 were negative as visualized in Figure 5. However, 194 neutral tweets were wrongly categorized as negative and 262 neutral tweets were categorized as positive. Additionally, 319 positive tweets were incorrectly classified as neutral and 691 positive tweets consider as negative, along with 290 negative tweets incorrectly classified as neutral and 511 negative tweets consider as positive.

4. Conclusions

The model was constructed by combining a pre-trained multilingual BERT with an LSTM layer to effectively represent sentiment in Nepali COVID-19-related Tweets for classification purposes. Furthermore, the model's performance metrics were compared with those of an existing model utilizing the NepCOV19Tweets dataset. The classification model used to achieve this objective was a m-BERT model including an additional LSTM layer. Notably, the proposed model exhibited superior performance having an average accuracy 76.04% in evaluation to the existing models using NepCOV19Tweets dataset. However, during experimentation, the model's training process consumed a significant amount of time. While training the model, it occupied 6989 minutes and 8.1 seconds i.e., approximately 4.854 days. It may be possible to reduce training time by combining an LSTM sequence model with pre-trained word embeddings like Glove or Word2Vec. This research empowers companies to fine-tune marketing strategies and enhance their products and services. A business may create marketing messages that highlight safety and security in their goods or services if customers are showing anxiety or dread.

References

- [1] Chandio, B. A., Imran, A. S., Bakhtyar, M., Daudpota, S. M., and Baber, J. Attention-Based RU-BiLSTM Sentiment Analysis Model for Roman Urdu, *Appl. Sci.*, 12 (7) (2022).
- [2] Nemes, L. and Kiss, A. Social media sentiment analysis based on COVID-19, *J. Inf. Telecommun.*, 5 (1) (2021)1–15.
- [3] Manguri, K. H., Ramadhan, R. N., and Mohammed Amin P. R. Twitter Sentiment Analysis on Worldwide COVID-19 Outbreaks. *Kurdistan J. Appl. Res.*, (2020) 54–65.
- [4] Aljabri, M., Aljameel, S. S., Khan, I. U., Aslam, N., Chrouf, S. M. B., and Alzahrani, N. Machine Learning Model for Sentiment Analysis of COVID-19 Tweets, *Int. J. Adv. Sci. Eng. Inf. Technol.*, 12 (3) 1206–1214 (2022).
- [5] Li, M., Chen, L., Zhao, J., and Li, Q. Sentiment Analysis of Chinese stock reviews based on BERT model. (2021) 713–718.
- [6] Song, Y., Wang, J., Liang, Z., Liu, Z., and Jiang, T. Utilizing BERT Intermediate Layers for Aspect Based Sentiment Analysis and Natural Language Inference, (2020).
- [7] Nguyen, Q. T., Nguyen, T. L., Luong, N. H., and Ngo, Q. H. Fine-Tuning BERT for Sentiment Analysis of Vietnamese Reviews, *Proc. - 2020 7th NAFOSTED Conf. Inf. Comput. Sci. NICS 2020*, 302–307 (2020).
- [8] Rai, N., Kumar, D., Kaushik, N., Raj, C., and Ali, A. Fake News Classification using transformer based enhanced LSTM and BERT, *Int. J. Cogn. Comput. Eng.*, 2 (2022) 98–105.
- [9] Sitaula, C., Basnet, A., Mainali, and Shahi, T. B. Deep Learning-Based Methods for Sentiment Analysis on Nepali COVID-19-Related Tweets, *Comput. Intell. Neurosci.*, (2021).
- [10] Dunn, D. Understanding BERT and GPT: Using the Right Tool for the Job, (2023), see <https://www.linkedin.com/pulse/understanding-bert-gpt-using-right-tool-job-doug-dunn#:~:text=With its bidirectional context understanding, entity recognition%2C, and sentiment analysis.>
- [11] Shahi, T. B., Sitaula, C., and Paudel, N. A Hybrid Feature Extraction Method for Nepali COVID-19-Related Tweets Classification, *Comput. Intell. Neurosci.*, (2022).
- [12] Tamrakar, S., BK, B., and Thapa, R. B. Aspect Based Sentiment Analysis of Nepali Text Using Support Vector Machine and Naive Bayes, *Tech. J.*, 2 (1) 22–29 (2020).
- [13] Shahi, T. B. and Pant, A. K. Nepali news classification using Naive Bayes, Support Vector Machines and Neural Networks, *Proc. - 2018 Int. Conf. Commun. Inf. Comput. Technol. ICCICT 2018*,

1–5 (2018).

- [14] Sitaula, C., Basnet, A., Mainali, A., and Shahi, T. B., <https://www.kaggle.com/datasets/mathew11111/nep-cov19tweets>, *Comput. Intell. Neurosci.* (2021).