*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 7***
***DOI: https://doi.org/10.17762/ijritcc.v11i7.8002***
***Article Received**: 10 May 2023 **Revised**: 24 June 2023 **Accepted**: 25 July 2023*

# A Hybrid Optimization Approach for Neural Machine Translation Using LSTM+RNN with MFO for Under Resource Language (Telugu)

**Srisudha Garugu[1*], D. Lalitha Bhaskari[2]**
[1]Research Scholar, Department of Computer Science & System Engineering, srisudha.garugu@gmail.com, Andhra University College of Engineering (A), Andhra University, Visakhapatnam-530003, India[1]
[2]Professor Department of Computer Science & System Engineering, Coordinator, IQAC, lalithabhaskari@yahoo.co.in, Andhra University College of Engineering (A), Andhra University, Visakhapatnam-530003, India[2]

**Abstract:** NMT (Neural Machine Translation) is an innovative approach in the field of machine translation, in contrast to SMT (statistical machine translation) and Rule-based techniques which has resulted annotable improvements. This is because NMT is able to overcome many of the shortcomings that are inherent in the traditional approaches. The Development of NMT has grown tremendously in the recent years but NMT performance remain under optimal when applied to low resource language pairs like Telugu, Tamil and Hindi. In this work a proposedmethod fortranslating pairs (Telugu to English) is attempted, an optimal approach which enhancesthe accuracy and execution time period.A hybrid method approach utilizing Long short-term memory (LSTM) and traditional Recurrent Neural Network (RNN) are used for testing and training of the dataset. In the event of long-range dependencies, LSTM will generate more accurate results than a standard RNN would endure and the hybrid technique enhances the performance of LSTM. LSTM is used during the encoding and RNN is used in decoding phases of NMT. Moth Flame Optimization (MFO) is utilized in the proposed system for the purpose of providing the encoder and decoder model with the best ideal points for training the data.

**Keywords:** Neural Machine Translation, statistical machine translation, Recurrent Neural Network, Moth Flame Optimization, Long short-term memory.

## I. Introduction

The term machine translation (MT) implies that it is the practice of using computer programs to make translations from one natural language to another, either with or without the assistance of a human translator [1]. As computational endeavors gain in popularity and the internet becomes accessible to a larger heterogeneous and international audience, studies in MT is continuing to rise at a rapid pace [2]. Computer and linguist scientists were researching as a team on MT for the last couple of decades, and as a result of their combined efforts, significant headway has been accomplished on this difficult assignment. There have been a variety of MT methodologies created, some of which include rule-based, example-based, statistical, andNMT( neural machine translation) techniques [1].

Technological breakthroughs developments in machine learning have led to an increase in the use of approaches that rely on neural networks a range of artificial intelligence (AI)-related subfields, one of which is NMT [3]. NMT is the method that is currently considered to be the most advanced in the field of MT. This method acquires its knowledge from a vast dataset that contains phrases written in the source language as well as their translations written in the language that is targeted [4]. Deep neural approaches, such as NMT, require a large amount of data and are not able to be effectively trained in contexts with limited resources. Although though MT has made significant progress in the past decades, It has only been applied to languages with a high availability of resources. This is due to the fact that neural networks require huge corpus in order to train effectively.

Hence, For only a small subset of language families is it feasible to acquire the extended parallel data necessary to maximize the effectiveness of NMT techniques.This restricts the amount of language pairs that may be studied [5]. This is because extended parallel data cannot be obtained for all language combinations. This is because extended parallel data is difficult to collect. Excellent translation quality was obtained for high-resource language pairings, like English and Portuguese [8], English and Russian [6,9], English and French [6,8,9], and English and Spanish [6–8]. NMT systems struggle to function well in situations with limited resources as a consequence of a deficit of sufficient training data for the languages in question [10]. People who speak low-resource languages face challenges in utilizing the most recent

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 7*
*DOI: https://doi.org/10.17762/ijritcc.v11i7.8002*
***Article Received**: 10 May 2023 **Revised**: 24 June 2023 **Accepted**: 25 July 2023*

technologies in their day-to-day lives, including dependable natural language processing (NMT) systems. This is due to the fact that low-resource languages aren't as effectively depicted in digital spaces as high-resource languages. This is because digital spaces are dominated by high-resource languages. Researchers have come up with a number of potential solutions to this issue, some of which include meta-learning [17], filtered pseudo-parallel datasets [16], data augmentation [15], multimodal NMT [14], exploiting related languages [13], transfer learning [12], and multilingual NMT [11].

Nevertheless, the majority of the suggested methods based their tests on high-resource languages. Based on prior literature study, no research workcollected utilizing monolingual source data with the goal of bettering NMT for languages with few resources. We suggest an effective strategy for exploiting relevant data to boost the effectiveness of NMT in low-resource languages; for the sake of this proposal, we employ the A case study in Wolaytta-English interpreting and translation. We gathered forty thousand Wolaytta monolingual data from a variety of sources, including Wolaytta textbooks, WolayttaWogeta Radio, WolayttaFana Radio and utilized them in conjunction in two experiments with a parallel Wolaytta–English dataset: (i) training from the baseline the required model on the groupings of the authentic as well as synthetic datasets, and (ii) model training on the present parallel Wolaytta–English dataset.

## II. Literature Review

Recent years have seen a surge in the number of suggestions made by researchers for the process to enhance NMT for languages that have limited resources. Using monolingual data and parallel datasets is one technique to enhance the NMT system. This is especially helpful for languages that have a limited amount of available resources. In this section, we investigate similar works undertaken for various languages utilizing monolingual datasets as an extra source for low-resource languages to enhance NMT systems. These studies were performed with the aim of improving NMT systems for those languages.

In order to solve the issue of parallel datasets that arises in bidirectional Tamil–Telugu neural machine translation (NMT), Laskar et al. [20] suggested making use of monolingual data in transformer model-based neural machine translation by means of pretrained word embeddings. The authors pre-trained on the monolingual corpora word embedding with GloVe [21] and then utilized those results within the context of the transformer model when the system is undergoing training.Both the translation from Telugu into Tamil and the translation from Tamil into Telugu both achieved a BLEU score of 4.05. The given model was used for both translations. Marie et al. [22] presented a novel way for generating huge parallel synthetic data by exploiting extremely in a particular domain little monolingual data. This would allow for the generation of parallel data in both languages simultaneously. They started with a modest amount of in-domain monolingual data and utilized a pre-trained GPT-2 model to fine-tune the model using the data. After that, they made use of the algorithm to create a substantial quantity of fabricated in-domain data in a single language. Then, in order to produce synthetic in-domain parallel data, a back translation was carried out by the team. When they trained NMT with synthetic data, their results demonstrated enhancements for all configurations in BLEU across all five domains and three language pairs (English to German, English to French, and English to Japan). Tars et al. [23] devised a multilingual training strategy, that could be enhanced upon utilizing the back translation technique for creating synthetic bilingual corpora by utilizing monolingual data. This can be done in order to improve the accuracy of the multilingual training strategy. Something can be done to increase the quality of machine translations using limited resources. Their multilingual learning technique and synthetic corpora both contributed to an improvement in the quality of translation for language pairs originating in the geographical regions of Estonia and Finland.

The NMT algorithm needs to be made more efficient for Turkish-English and English-German translations, a number of optimizations have been made.Sennrich et al. [24] made use of data that was only available in a single language. This was accomplished by improving the model's ability to translate from Turkish to English. They were successful in reaching their objective of producing synthetic parallel data by performing a translation of monolingual data into the source language from the target language.

After that, their first target–source MT system using the parallel data was trained that was readily available. Upon training, the system obtained the back-translated data by translating the monolingual dataset into the target language from the source language. This was done so that the data could be back-translated. Following this, By combining information from the original parallel with information from the back-translated parallel, Sennrich et al. [24] are training the final source-target NMT system. This was done so as to enhance the system's accuracy.

In conclusion, the researchers were successful in achieving considerable advances on increases ranging from +2.1 to 3.4 BLEU on the Turkish-English IWSLT-14 low-resource task, and from +2.8 to 3.7 BLEU on the English-German WMT-15 task. Jiao et al. [25] provided in a single language for NMT purposes uncertainties self-training

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 7*
*DOI: https://doi.org/10.17762/ijritcc.v11i7.8002*
***Article Received**: 10 May 2023 **Revised**: 24 June 2023 **Accepted**: 25 July 2023*

sampling as a means of enhancing the sampling technique by picking the highly monolingual informative sentences with which to supplement the parallel data. This was done in order to improve NMT.

They determined the degree of ambiguity associated with monolingual sentences by employing the bilingual lexicon that was produced from the parallel data. They came up with the notion that putting more of an emphasis on the acquisition of doubtful monolingual phrases would improve of high-uncertainty sentences, its translation quality and would be advantageous to the prediction on the target side of the equation of low-frequency terms. According to the outcomes of their research of enormous WMT Datasets, which compared English and Chinese as well as English and German, the proposed technique improved the performance of NMT.

By utilizing the copied corpus techniques, back translation, and sub word segmentation, Dione et al. [26] were able for bidirectional Wolof–French translation to increase efficiency of the NMT in practice. The accuracy of the translation to French from Wolof demonstrated signs of enhancement in every direction when the copied corpus as well as back translation were utilized together. This improvement was noticeable in every direction. Pham [27] suggested using the Google Translate program as a method of reverse the quality of back-translated monolingual texts from English to Vietnamese could be enhanced with the help of MT. This would be accomplished by translating the texts from Vietnamese into English. This was done using the Google Translate language pair. Their recommended approach increased the BLEU score by 16.37 points in comparison to the strategy that was used as the baseline for English–Vietnamese NMT. The authors Ngo et al. [28] presented the extraction of vocabulary from the target text of the production of artificial translation units as well as the original bilingual corpus by identifying each standard translation. They argued for these two practices. They did this by marking each standard translation with a unique identifier. After that, they trained NMT systems using a combination of the native corpus and the synthetic corpus that they created. The results of their technique showed an increase in BLEU scores of +1.8 for Chinese–Vietnamese translation assignments and +1.9 for Japanese–Vietnamese translation tasks, respectively.

## III. Methodology

Encoding and decoding are both considered to be phases of the planned NMT because they are both involved in the process.In most cases, the standard RNN will be used for both the encoding and decoding parts of the process. The traditional RNN is likely to have some trouble when it comes to managing the longer-range dependencies that are present in the source phrase.

Encoder and decoder in the proposed model are both implementations of LSTM(long short-term memory), rather than the more traditional RNN. In the event of long-range dependencies, LSTM will generate more accurate results than a standard RNN would. LSTMs are required to be utilized during the encoding and decoding phases of NMT.

It tackles both OOV and data replication difficulties at the pre-processing phase of the operation. Whilst the MFO will be utilized for the purpose of providing the encoder and decoder model with the best ideal points for training the data.The NMT process can be broken down into three distinct stages: pre-processing, encoding, and decoding. The NMT is a method that is driven by data. Hence, it is contingent upon the parallel corpus.

In order to provide more accurate translations, NMT relies on the use of massive parallel corpora. Collecting a parallel corpus might be difficult for languages with few resources, such as Telugu and English, which are both examples. The generation of parallel corpora can be accomplished manually or with the use of various tools.

The use of tools will result in the production of noises. It is challenging to contend with these sounds. So, carrying out the preparation of the corpus by hand would be preferable; nonetheless, the data may still contain noise and inconsistencies. When there is inconsistency or noise in the Parallel corpus, the accuracy of the translation will suffer as a result. During the pre-processing phase, both inconsistencies and noises found in a parallel corpus will be dealt with. In addition, NMT has issues with data replication and out-of-vocabulary (OOV).

The identical source sentences, but with different translations, will make up the data replication in the parallel corpus. And vice versa. During the pre-processing phase, NMT will handle any issues pertaining to data replication as well as OOV.As of right now most of the NLP related RNN based model or model that is follow the base structure of RNN at its core like BERT, LSTM, DymemNN etc.

most of these models uses ADAM optimizer for the optimum point while training the data which is not the best optimizer there is, So Now for Improving this model we will change its core optimizer from ADAM to MFO which will Increase model training Process by help model train only the optimum data.

Moth Flame Optimization, or MFO for short, is one of the most effective Swarm-like optimization techniques currently available, alongside PSO and ABC. It will be

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 7*
*DOI: https://doi.org/10.17762/ijritcc.v11i7.8002*
***Article Received**: 10 May 2023 **Revised**: 24 June 2023 **Accepted**: 25 July 2023*

utilizing libraries such as Tensorflow, pyTorch, MealPy, NiaPy, and Sklearn, amongst others, in order to create this Model.For the purpose of this implementation's preprocessing, we will be utilizing strategies such as TFIDF, stop Words, Corpus, count Vector tokenizer, and the OpenNMT-py toolkit. In addition, it will be utilizing a number of essential NLTK library components, such as word2vec, tokenizer, porter streamer, and word clouds, amongst other preprocessing strategies.

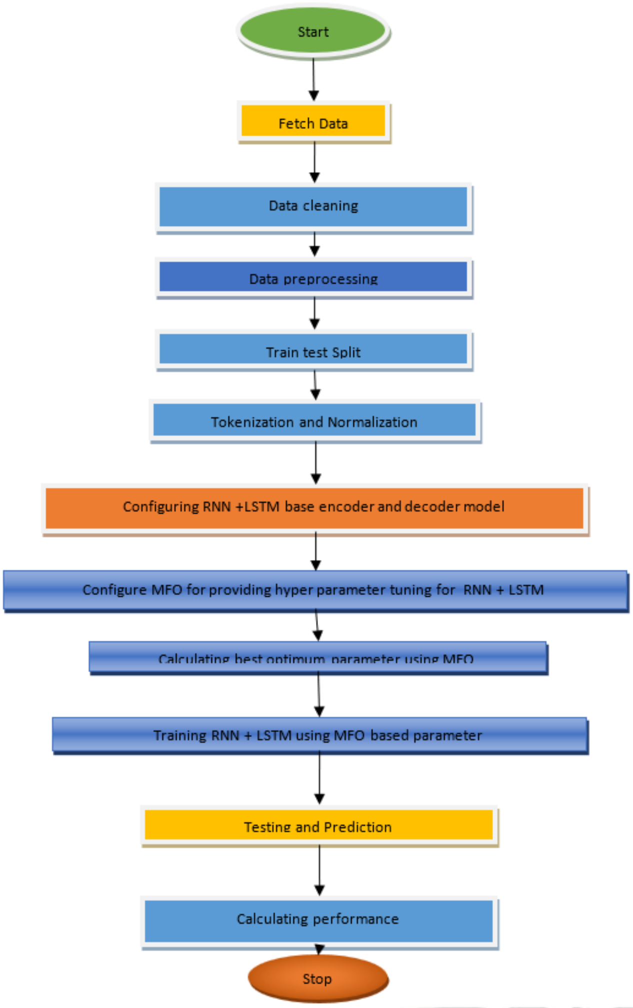


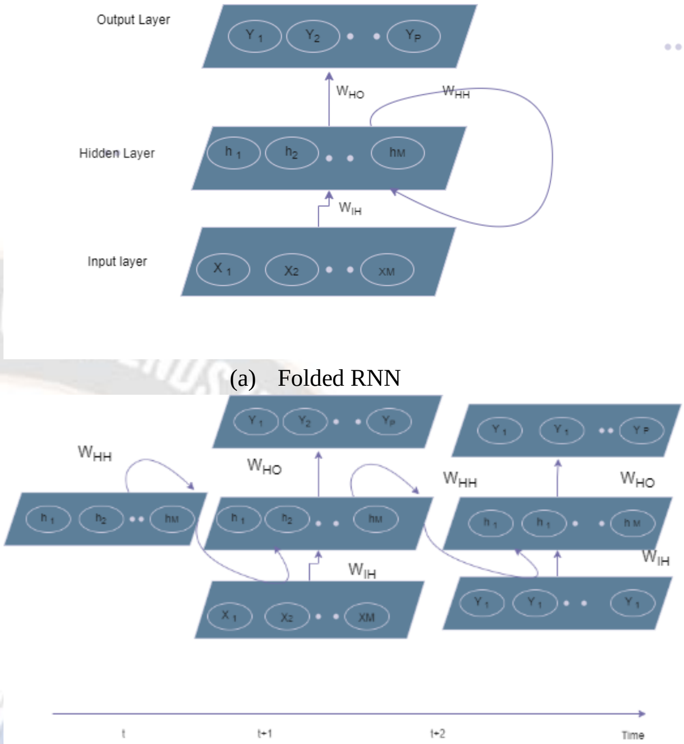

Figure 1: Flowchart of the Proposed Research

### 3.1 Recurrent Neural Network

RNNs, or recurrent neural networks, are an assortment of supervised deep learning methods and multiple feedback loops [7]. RNNs are also known as recurrent neural networks.

Figure 1 illustrates the feedback loops, which are repeated cycles across sequence or time. In order to train an RNN in a supervised manner, a train dataset consisting of input target pairings is required.

The purpose is to optimize the weights of the network in order for minimizing the variations that exists among the target pairs (also known as the loss value) and the output pairs.




(a) Folded RNN

(b)Unfolded

Figure 2: A Simple RNN and its Unfolded Structure through time t. Each arrow represents a complete network of interlayer connections.

As can be seen in Figure 1a, a straightforward RNN is composed of three layers, labeled respectively as three layers: the input layer, the recurrent hidden layer, and the output layer. There are N input units located in the layer below the input layer. This layer receives its input in the form of a vectors sequence across time t, for instance $\{\ldots, \mathbf{x}_{t-1}, \mathbf{x}_t, \mathbf{x}_{t+1}, \ldots\}$, where $x_t$ equals the current time (x1, x2, ..., xN ). The connections between the input units and the hidden units in a fully connected RNN are described by a weight matrix called WIH. The input units are linked to the hidden units in the hidden layer via connections made by the input units.

Figure 1b shows that the hidden layer has M hidden units that are related to one another through time by recurrent connections. These units are denoted by the notation ht = (h1, h2,..., hM). It is possible to increase the overall performance of the network as well as its stability by employing modest non-zero components to execute the initialization of hidden units [9]. The state space of the system, sometimes known as its "memory," is defined by the hidden layer as

$$h_t = f_H(0_t), \ldots.(1)$$

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 7*
*DOI: https://doi.org/10.17762/ijritcc.v11i7.8002*
***Article Received**: 10 May 2023 **Revised**: 24 June 2023 **Accepted**: 25 July 2023*

Where

$$o_t = W_{IH}x_t + W_{HH}h_{t-1} + b_h, \ldots\ldots (2)$$

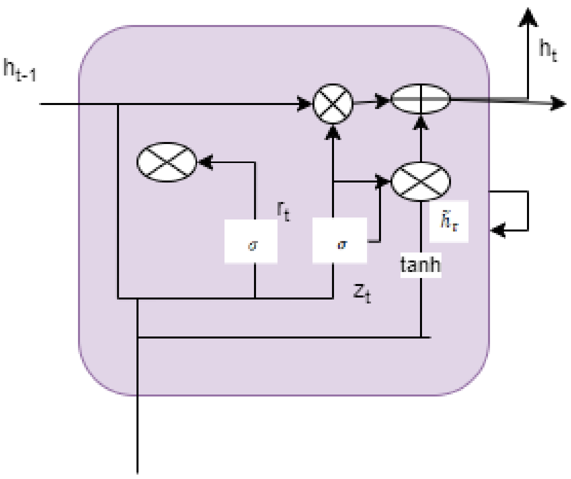
Of the hidden units, $b_h$is the bias vector and the hidden layer activation function is $f_H(\cdot)$. The hidden units are given connection whose WHO connections are weighted and lead to the output layer. The P units of the output layer, denoted yt= (y1, y2,..., yP), are being calculated as

$$y_t = f_O(W_{HO}h_t + b_o), \ldots\ldots (2)$$

Where$b_o$in the output layer is the bias vector and $f_O(\cdot)$ is the activation functions and. Because the input-target pairs through time are continuous, the initial steps are given repetition subsequently t = (1,..., T) during the course of time. The equations (1) and (2) together model an RNNconsists of a small number of non-linear state equations that, as time passes, can be iterated. A prediction is provided at the output layer by the hidden states at the end of each and every time step, and this prediction is on the basis of the vector input. A RNN's concealed state is information that, independent of the impact of any externally associated elements, summarizes all of the one-of-a-kind information that is crucial regarding the states that the network has been in across a number of time steps. In other words, the hidden state of an RNN is a state that is not visible to the outside world. Because of the integration of these pieces of information, it is possible to describe the network behavior in the future and to enable accurate prediction at the output layer [5]. In each unit of an RNN, a straightforward nonlinear activation function is utilized. However, despite its apparent simplicity, such a structure has the potential to accurately describe complex dynamics provided it is adequately trained using successive time steps.

**3.2 LSTM**

LSTM is a novel recurrent network architecture that is combined appropriately with gradient-based learning algorithm. LSTM was developed specifically to address issues related to error back-ow. Even in the presence of noisy, incompressible input sequences, it is capable of learning to bridge temporal spans of more than 1000 steps without suffering a reduction in its small time lag capabilities. So as to obtain this objective, it is needed to have an effective algorithm which is based on gradients and an architecture across the internal states of special units that guarantees continuous (and, as a result, either vanishing nor exploding) error flow. The long-term error flow will not be affected by this as long as the computation of the gradient is terminated at specified locations that are relevant to the design.




$$z_t = \sigma(W_z[h_{t-1}, x_t])$$

$$r_t = \sigma(W_r[h_{t-1}, x_t])$$
$$\tilde{h}_t = tanh(W[r_t * h_{t-1}, x_t])$$
$$h_t = (1 - z_t)\big(W * h_{t-1} + z_t * \tilde{h}_t\big)$$

Figure 2 : LSTM Block Diagram

In the proposed research, based on CNN training, the LSTM is trained with MFO.

**3.3 Moth Flame Optimization**

A fresh optimization paradigm that takes its cues from nature has been conceptualized and given the name Moth-Flame Optimization (MFO). The strategy of moths in properties is called as transverse motion, and it served as the main reason for the enablement of this Optimization. Moths are capable of flying at night by maintaining a constant angle in relation to the moon, resulting in an excellent method for advancing in a single direction throughout long distances. Nevertheless, these elegant insects are destined to die since they must adhere to a lethal and pointless course around artificial lights. This behavior is mathematically modeled in this work so that optimization can be performed.

The moth-flame optimization algorithm, also known as the MFO algorithm, has a few advantages, including the fact that it is simple to grasp and put into practice, that it converges quickly, and that it has a small number of setup parameters. Moth Flame Optimization (MFO) (Mirjalili, 2015)

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 7*
*DOI: https://doi.org/10.17762/ijritcc.v11i7.8002*
*Article Received: 10 May 2023 Revised: 24 June 2023 Accepted: 25 July 2023*

is a new solution that was recently developed for the resolution of global optimization issues and applications in the real world. In addressing a wide range of issues, it has previously been demonstrated and demonstrated effectively that MFO is effective in terms of population diversity and convergence rate.

**The following is a list of the most important parts of the MFO algorithm:**

**Moth component:**The MFO algorithm can locate global minima for any given problem. It is composed of three pieces that work together to perform this task. The following is one possible definition of it:

$$MFO = (I, P, T)$$

I is the component that determines the initialization in search space from a population of moths; P is the function that moves moths closer to the fire, and T is the function that is true if and only if the requirement for termination is satisfied.

The array below displays an example of how the fitness value is stored of every solution:

$$OM = \begin{bmatrix} OM_1 \\ OM_2 \\ . \\ . \\ . \\ OM_n \end{bmatrix}$$

The value n in this context refers to the total number of moths. The location vector of the moth is an input that is taken into consideration by the fitness function, and it returns fitness value as output.

**Flame component:** Similar to the way the moth matrix is handled, the F matrix when attempting to explain the flames, Additionally, the fitness value of the F matrix is calculated in a manner that is analogous to that of the previous method.In this case, the moths and the flames are both viable options variation between them is in the technique by which they are updated.

In addition, we are able to say that flames are the best solution that has been found so far, and the moths are the search agents investigating the potential outcomes. As moths are searching for something, they leave behind flames, which might be thought of as flags.

**Transverse orientation of moths:**Through the duration of the flame, the moth's location is updated utilizing the equation in order to imitate the behavior of the moth in the mathematical model (2)

$$M_i = S(M_i, F_j)$$

In this case, S for the spiral function, Fj for the jth flame, and Mi stands for the ith moth, which describes how The moth moves "around" the flame, rather than always between the flames.

Within the MFO algorithm that has been proposed, we make the assumption that the candidate solutions are moths, and the locations of the moths within the room serve as the problem's variables.

Resultant, the moths are capable of flying in one dimension. 2-D, 3-D, or with their position vectors constantly changing, even hyper-dimensional space. Since the MFO technique is a population-based technique, and the set of moths is indicated in the following way in a matrix:

$$\begin{bmatrix} m_{1,1} & m_{1,1} & \cdots & \cdots & m_{1,d} \\ m_{2,1} & m_{2,2} & \ldots & \cdots & m_{2,d} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{n,1} & m_{n,2} & \cdots & \ldots & m_{n,d} \end{bmatrix}$$

Where d represents the total quantity of different factors (dimension) and n represents the total quantity of moths.

## IV. Results and Discussion

### 4.1 Dataset used

The proposed research is utilizing two datasets. In the first dataset, the Hindi_English_Truncated_Corpus(with Telugu Translation using Google API) is considered and second dataset is created by taking Tenth Class Telugu Biology Text Book

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 7***
***DOI: https://doi.org/10.17762/ijritcc.v11i7.8002***
***Article Received**: 10 May 2023 **Revised**: 24 June 2023 **Accepted**: 25 July 2023*

Table 1: Representation of the dataset used in the proposed research

| Source | English_sentence | hindi_sentence | Telugu_sentence |
|---|---|---|---|
| ted | politicians do not have permission to do what needs to be done. | राजनीतिज्ञोंकेपासजोकार्यकरनाचाहिए , वहकरनेकिअनुमतिनहींहै . | చేయాల్సినపనిచేయడానికిరాజకీయనాయకులకుఅనుమతిలేదు. |
| ted | I'd like to tell you about one such child, | मईआपकोऐसेहीएकबच्चेकेबारेमेंबतानाचाहूंगी , | అలాంటిపిల్లలగురించినేనుమీకుచెప్పాలనుకుంటున్నాను |
| ted | what we really mean is that they're bad at not paying attention. | हमयेनहींकहनाचाहतेकिवोध्याननहींदेपाते | మనంనిజంగాఅర్థంచేసుకున్నదిఏమిటంటేవారుశ్రద్ధచూపకపోవడంలోచెడ్డవారు. |
| ted | And who are we to say, even, that they are wrong | औरहमहोतेकौनहैंयहकहनेभीवालेकिवेगलतहैं | మరియువారుతప్పుఅనిచెప్పడానికిమనంఎవరు |
| ted | So there is some sort of justice | तोवहाँन्यायहै | కాబట్టిఒకవిధమైనన్యాయంఉంది |
| ted | This changed slowly | धीरेधीरेयेसबबदला | ఇదినెమ్మదిగామారింది |
| ted | were being produced. | उत्पन्ननहींकिजातीथी. | ఉత్పత్తిచేయబడుతున్నాయి. |
| ted | And you can see, this LED is going to glow. | औरजैसाआपदेखरहेहै, येएल.ई.डी. जलउठेगी। | మరియుమీరుచూడగలరు, ఈ LED మెరుస్తుంది |
| ted | to turn on the lights or to bring him a glass of water, | लाईटजलानेकेलिएयाउनकेलिएपानीलानेकेलिए , | లైట్లువేయడానికిలేదాఅతనికిఒకగ్లాసునీరుతీసుకురావడానికి |
| ted | Can you imagine saying that? | क्याआपयेकल्पनाकरसकतेहै | అలాచెప్పడంమీరుఊహించగలరా? |

The Hindi_English_Truncated_Corpusdataset has a total of **127606** sentences (with Telugu Translation using Google API). The translated English sentences were produced with the help of the Google Translate API and the dataset presented above.

https://www.kaggle.com/datasets/umasrikakollu72/hindi-english-truncated-corpus.

Similarly Second dataset is created using tenth Class Biology Text Book has a total of **6740** sentences.Initially for the first dataset used in the proposed research, the two datasets are merged together to create a single dataset and deleted Hindi Sentences Column. The Original The information can be found in table 1.

The shape of the First dataset is **(3000,4)**. The second dataset is represented as given in table 1. The shape of the second dataset is (2051, 3).For the first dataset, initially the column Selection is done and then the grammar Selection is completed. The steps used in the proposed methodology is represented for the first dataset.

The second dataset also has similar steps of implementation. The final results of both the datasets are compared. The Dataset-1 consists of 5064 tides, 3982 ted and 3714 indic2012. Final Dataset Sample with Newly Translated Data Column.

In the final dataset, the unnameddata, source, english_sentence, telugu_sentenceare all zero with dtype in the int64. The shape of the dataset is (3982, 4).

Table 2: Final Dataset Sample with Newly Translated Data Column(Data set 2)

| | Telgu_sentence | English_sentence | source |
|---|---|---|---|
| 0 | జీవులన్నింటికీపెరుగుదలవంటిజీవక్రియలనునిర్వర్తించుకోడంకోసంఆహారంఅవసరం. | All living things need food to carry out metabolic processes such as growth. | Ted |
| 1 | జీవులుశారీరకఉష్ణోగ్రతనుక్రమబద్ధీకరించుకోడానికికూడాఆహారంఅవసరం. | Organisms also need food to regulate body temperature. | Ted |
| 3 | మానవశరీరంలోవివిధకణాలకుతమవిధులనునిర్వర్తించడానికివేరువేరురకాలపోషకాలుఅవసరంఅవుతాయి. | Different cells in the human body require different types of nutrients to perform their functions. | Ted |
| 7 | స్వయంపోషకాలుఅనగానేమి? అవితమఆహారాన్నిఎలాపొందుతాయి? | What are auto nutrients? How do they get their aura? | Ted |
| 13 | అవినేలలోనినీటినిమరియుఖనిజలవణాలతోపాటుగాగాలిలోనికొన్నియువులనుకూడావినియోగించుకుంటాయి. | They consume water and mineral salts from the soil as well as some nutrients from the air. | Ted |
| 23 | కిరణజన్యసంయోగ(క్రియలోచిట్టచివరిగాఏర్పడేఉత్పన్నాలుఏమిటుంటాయి? | Photosynthetic (What are the final derivatives of a verb? | Ted |
| 26 | కిరణజన్యసంయోగక్రియ | Photosynthesis | ted |
| 30 | ఇదిచాలాసంక్లిష్టమైనప్రక్రియ . ఇందులోఅనేకరకాలచర్యలుక్రమపద్ధతిలోజరగటంతోపాటుగాఅనేకమధ్యస్థసమ్మేళనాలు (intermediary compounds) కూడాఏర్పడుతుంటాయి. | It is a very short process. In this, many types of actions take place systematically and many intermediate compounds are also formed. | Ted |
| 32 | పటం-1: కిరణజన్యసంయోగక్రియసూచించడానికిప్రయత్నిస్తున్నారు | Figure-1: Attempting to describe photosynthesis. | Ted |
| 35 | కిరణజన్యసంయోగక్రియావిధానంచాలాసంక్లిష్టమైనదైనప్పటికీప్రస్తుతంమనమందరంసులభమైన, సరళమైనఈసమీకరణాన్నే | Although the mechanism of photosynthesis is very complex, we all know this simple, simple equation. | Ted |
| 37 | పోషణ–ఆహారసరఫరావ్యవస్థ | Nutrition – Food supply system | ted |

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 7*
*DOI: https://doi.org/10.17762/ijritcc.v11i7.8002*
***Article Received**: 10 May 2023 **Revised**: 24 June 2023 **Accepted**: 25 July 2023*

### 4.2 Data preprocessing

Checking and Dropping Null Values takes place initially in the preprocessing step. Next, Applied Text Based Data Cleaning Pre-Processing steps are taking place. The applied test based data cleaning Pre-Processing steps are:

- Lowercase all characters
- Remove quotes
- Remove all numbers from text
- Remove extra spaces

The data After Pre-processing is as given in figure 4. (data set 2)

Table 3:Data After Pre-processing for Dateset-1

| Index | Telugu_sentence | English_sentence | Source |
|---|---|---|---|
| 2588 | తరువాత ఈ ద్రవం x అనే ద్వారం ద్వారా పోతుంది ast | START_ then this fluid drains through an orifice called x ast _END | ted |
| 2075 | ఎఫిడ్ తన తొండాన్ని మొక్కలో లోనికి చొప్పించి రసాన్ని పీలుస్తుంది | START_ the aphid inserts its stem into the of the plant and sucks the sap _END | ted |
| 1391 | ఆరోగ్యవంతమైన జీవనానికి దోహదపడేలా శ్వాసక్రియను  నియం తించుకోగలగడం ఒక్క మానవునికే సాధ్యమవుతుంది | START_ it is possible for only a human being to regulate breathing to contribute to a healthy life _END | ted |
| 823 | గ్రసని గాలిని వెచ్చ చేయడం గాలికి తేమను చేర్చడం వంటి కార్యక్రమాలు శ్వాస జీర్ణ వ్యవస్థలు రెండింటికీ సంబంధించిన ఈ భాగంలో కొనసాగుతాయి | START_ grasani the functions of warming the air and adding moisture to the air continue in this part of both the respiratory and digestive systems _END | ted |
| 4201 | ఆకలి వేసినట్లనిపించే భావన మనల్ని ఆహారం తీసుకునేందుకు ప్రేరేపిస్తుంది | START_ feeling hungry motivates us to eat _END | ted |

Table 4:Data After Pre-processing for Dateset-1

| Index | Source | English_sentence | Hindi_sentence | Telugu_sentence |
|---|---|---|---|---|
| 1395 | ted | START_ only one out of senators was willing to vote _END | तो 100 सीनेटर में से सिर्फ़ एक व्यक्ति वोट करने को तैयार था | మంది సెనేటర్లలో ఒకరు మాత్రమే ఓటు వేయడానికి సిద్ధంగా ఉన్నారు |
| 10677 | ted | START_ by number of books sold mentions in media _END | बेची गई पुस्तकों की संख्या से मीडिया में उल्लेख है, | అమ్మిన పుస్తకాల సంఖ్య మీడియాలో ప్రస్తావనలు |
| 2239 | ted | START_ fn and i also had an interest in dangerous inventions _END | FN: और मुझे खतरनाक आविष्कारों में भी रुचि थी. | fn మరియు నాకు ప్రమాదకరమైన ఆవిష్కరణలపై కూడా ఆసక్తి ఉంది |
| 10177 | ted | START_ applause _END | (तालिया) | చప్పట్లు |
| 7616 | ted | START_ we did a design competition _END | हमने एक डिजाइन प्रतियोगिता रखवाई, | మేము డిజైన్ పోటీ చేసాము |

Converting Data into Vectors, English Text Vectors Size is 4249 and Telugu Text Vectors Size is 6151. Adding Token Length Columns is given in Table 4. No. Of Columns is (0, 7).

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 7***
***DOI: https://doi.org/10.17762/ijritcc.v11i7.8002***
***Article Received**: 10 May 2023 **Revised**: 24 June 2023 **Accepted**: 25 July 2023*

Table 5: Adding Token Length Columns for Dateset-1 & 2

| | **Telugu_sentence** | **English_sentence** | **Source** | **Length_eng_sentence** | **Length_tel_sentence** |
|---|---|---|---|---|---|
| **2588** | తరువాతఈద్రవం x అనేద్వారంద్వారాపోతుందిast | START_ then this fluid drains through an orifice called x ast _END | ted | 12 | 9 |
| **2075** | ఎఫిడ్తనతొండాన్నిమొక్కలోలోనికిచొప్పించిరసాన్నిపీలుస్తుంది | START_ the aphid inserts its stem into the of the plant and sucks the sap _END | ted | 16 | 8 |
| **1391** | ఆరోగ్యవంతమైనజీవనానికిదోహదపడేలాశ్వాసక్రియనునియంతించుకోగలగడంఒక్కమానవునికేసాధ్యమవుతుంది | START_ it is possible for only a human being to regulate breathing to contribute to a healthy life _END | ted | 19 | 9 |
| **823** | గ్రసనిగాలినివెచ్చచేయడంగాలికితేమనుచేర్చడంవంటికార్యక్రమాలుశ్వాసజీర్ణవ్యవస్థలురెండింటికీసంబంధించినఈభాగంలోకొనసాగుతా | START_ grasani the functions of warming the air and adding moisture to the air continue in this part of both the respiratory and digestive systems _END | ted | 26 | 17 |

Maximum Text Possibility on One Line is maximum length of Telugu Sentence is 15 and the maximum length of English Sentence is 20.

Table 6: Sample dataset after tokenization for Dataset-1

| | **Telugu_sentence** | **English_sentence** | **Source** | **Length_eng_sentence** | **Length_tel_sentence** |
|---|---|---|---|---|---|
| **5297** | ఆపక్షులకుఎలాఉపయోగపడుతుంది | START_ how is it useful for birds _END | ted | 8 | 4 |
| **2487** | ఉదాఆల్కలాయిడ్స్టానిన్లురెసిన్లుజిగురులుమరియులేటెక్సులు | START_ eg alkaloids tannins resins glues and latexes _END | ted | 9 | 7 |
| **198** | అమరేలాచూడండి | START_ see amarela _END | ted | 4 | 2 |
| **1897** | నీటిమట్టంరబ్బరుగొట్టంకంటేకాస్తపైకికనబడేవిధంగాఉండాలి | START_ the water level should be slightly above the rubber tube _END | ted | 12 | 9 |
| **4078** | జీవజాతులయొక్కజనాభాలోనిలకడశాశ్వతంగానిలుచుటకోసమైప్రత్యుత్పత్తిఎలాసహకరిస్తుంది | START_ persistencepersistence in populations of species how does reproduction contribute _END | ted | 11 | 9 |
| **3952** | ఎయిడ్స్దినరిపఎయిడ్స్వ్యాధికకారణమైనవైరస్ఏమిటి | START_ what is the virus that causes aids _END | ted | 9 | 7 |
| **391** | లాలాజలంలోఅమైలేజ్టయలిన్అనేఎంజైమ్ఉంటుంది | START_ saliva contains an enzyme called amylase tyalin _END | ted | 9 | 6 |

Tokenizing Vectors with Start/END Tokens is (6151, 4209). The dataset is split into training and testing dataset. The Train Test Split is given as ((2400,), (600,)). Stop Word and Word Index Tokenization of Training and Testing Data is given as in figure 3. For the purpose of conducting an analysis of the content of the text, tokenization is the process of breaking down huge amounts of textual material into their component pieces.

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 7***
***DOI: https://doi.org/10.17762/ijritcc.v11i7.8002***
***Article Received**: 10 May 2023 **Revised**: 24 June 2023 **Accepted**: 25 July 2023*

| | |
|---|---|
| '<OOV>' | **1,** |
| 'మరియు' | 2, |
| 'నేను' | 3, |
| 'మీరు' | 4, |
| 'ఇది' | 5, |
| 'ఈ' | 6, |
| 'చాలా' | 7, |
| 'మేము' | 8, |
| 'ఒక' | 9, |
| 'అది' | 10, |
| 'కాబట్టి' | 11, |
| 'వారు' | 12, |

Figure 3: Stop Word and Word Index Tokenization of Training and Testing Data for Dataset-1

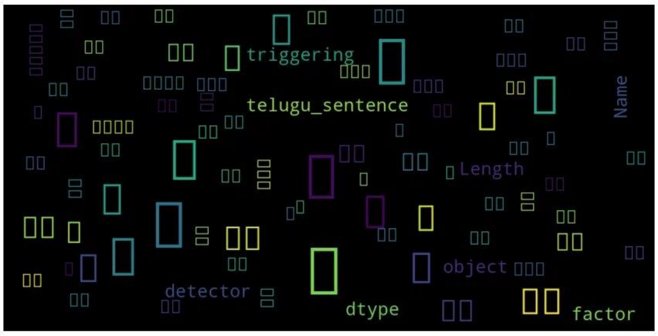



Figure 4: Word Cloud For DatSet-1

## 4.3 Creating RNN and LSTM Auto Encoder Model Architecture

Model Summary of the RNN and LSTMauto encoder model architecture is given in figure 9. In the process of running a classification or regression model, a model summary is given generation on its own in automation. The name of the model, the type of the model, and the formula for the model in the model summary are all displayed. When dealing with parametric models, extra summary statistics that are to the specific model type pertinent being discussed are also shown. These statistics can provide an indicator of how well the model fits the data, and they could also be utilized for comparing one model to another model that is of similar type.

Training of the RNN & LSTM model is given in figure 5. Model Loss Graph is given in figure 6. A loss is the inevitable outcome that will follow from an incorrect prediction. In other words, loss is a numerical figure that represents how erroneous the model's estimate was based on a single data point. This forecast was derived from a single point of data. In the case that the model's prediction is absolutely accurate, there won't be any loss; nevertheless, in any other scenario, the loss will be significantly greater.

Table 7: Model summary

Model: "model"

| Layer(type) | Output Shape | Param # | Connected to |
|---|---|---|---|
| input_1(Input Layer) | [(Nome, None)] | 0 | [] |
| input_2(Input Layer) | [(Nome, None)] | 0 | [] |
| embedding(Embedding) | [(Nome,None,300)] | 1845300 | ['input_1[0][0]'] |
| embedding_1(Embedding) | [(Nome,None,300)] | 1274500 0 | ['input_2[0][0]'] |
| lstm(LSTM) | [(None,300),(None,300),(None,300)] | 7212100 | ['embedding[0][0]'] |
| lstm_1(LSTM) | [(None,none,300),(None,300),(None,300)] | 7212100 | ['embedding_1[0][0]','lstm[0][1]','lstm[0][1]'] |
| dense[Dense] | [(Nome,None,4250)] | 1279250 | [lstm_1[0][0]'] |

Total params : 5,841,950
Trainable params: 5,841,950
Non-Trainable params : 0

Training Epochs

```
Epoch 1/100
18/18 [==============================] - 2s 98ms/step - loss: 5.7669 - val_loss: 6.00
Epoch 2/100
18/18 [==============================] - 2s 95ms/step - loss: 5.7528 - val_loss: 6.01
Epoch 3/100
18/18 [==============================] - 2s 118ms/step - loss: 5.7428 - val_loss: 6.0
Epoch 4/100
18/18 [==============================] - 2s 137ms/step - loss: 5.7251 - val_loss: 6.0
Epoch 5/100
18/18 [==============================] - 2s 100ms/step - loss: 5.7015 - val_loss: 6.0
Epoch 6/100
18/18 [==============================] - 2s 99ms/step - loss: 5.6946 - val_loss: 6.00
.
.
.
Epoch 96/100
18/18 [==============================] - 2s 95ms/step - loss: 4.1229 - val_loss: 5.87
Epoch 97/100
18/18 [==============================] - 2s 94ms/step - loss: 4.0963 - val_loss: 5.87
Epoch 98/100
18/18 [==============================] - 2s 101ms/step - loss: 4.0815 - val_loss: 5.8
Epoch 99/100
18/18 [==============================] - 2s 114ms/step - loss: 4.0643 - val_loss: 5.8
Epoch 100/100
18/18 [==============================] - 3s 148ms/step - loss: 4.0413 - val_loss: 5.8
```

Figure 5: Training RNN & LSTM

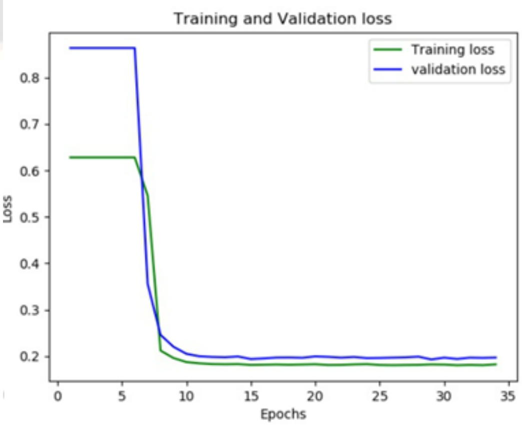



Figure 6: Model Loss Graph

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 7***
***DOI: https://doi.org/10.17762/ijritcc.v11i7.8002***
***Article Received**: 10 May 2023 **Revised**: 24 June 2023 **Accepted**: 25 July 2023*

Configuring MFO for RNN+LSTM Model is completed and the MFO Optimization process is as given in figure 7. Best Optimized Features and the **number of selected features is 10.**

```
INFO:mealpy.swarm_based.MFO.BaseMFO:Solving single objective optimization problem.

INFO:mealpy.swarm_based.MFO.BaseMFO:>Problem: P, Epoch: 1, Current best: 41.05323877069767, Global best: 41.05323877069767, Runtime: 0.00804 seconds

INFO:mealpy.swarm_based.MFO.BaseMFO:>Problem: P, Epoch: 2, Current best: 41.05323877069767, Global best: 41.05323877069767, Runtime: 0.00772 seconds

INFO:mealpy.swarm_based.MFO.BaseMFO:>Problem: P, Epoch: 3, Current best: 35.345877242601645, Global best: 35.345877242601645, Runtime: 0.00785 seconds
.
.
.
.
.
.
INFO:mealpy.swarm_based.MFO.BaseMFO:>Problem: P, Epoch: 999, Current best: 2.5083406062076947e-174, Global best: 2.5083406062076947e-174, Runtime: 0.01756 seconds
INFO:mealpy.swarm_based.MFO.BaseMFO:>Problem: P, Epoch: 1000, Current best: 6.06497857057715e-175, Global best: 6.06497857057715e-175, Runtime: 0.01137 seconds
```

Figure 7: MFO Optimization process

Table 8: Comparison of accuracy for both the datasets used in the Proposed Work and its Bleu Score

| Dataset Used | Accuracy (%) | Bleu Score |
|---|---|---|
| Dataset-1 (Kaggle) | 94.73 | 1.83621 |
| Dataset-2 (Biology Text Book) | 95.73 | 1.07732 |

### 4.4 Testing statements

The testing statements that were used on Dataset-1 and data set-2 for Telugu translation to English is given:

| | |
|---|---|
| Input Hindi sentence | కొన్ని సమయాల్లో ఇతరులకన్నా ఎక్కువ పని చేస్తుంది |
| Actual English Translation | higherfunctioning at times than others |
| Predicted English Translation | higherfunctioning at times than others |

Figure 8: Testing Statements 1(Data set 1)

| | |
|---|---|
| Input Hindi sentence | రాజకీయ నాయకుడికి ఇది చాలా కష్టం |
| Actual English Translation | It's very difficult for a politician |
| Predicted English Translation | it's very difficult for a politician |

Figure 9: Testing Statements 2(Data set 2) :

### 4.5 Comparative analysis

The proposed model LSTM with RNN and MFO is compared with independent LSTM and RNN for accuracy. The comparison graph is given in figure 13. It is seen that the proposed model has a higher accuracy of 95.73% when compared to the other models.

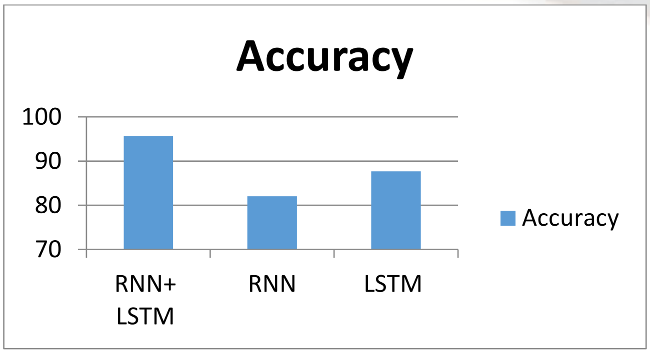



Figure 13: Comparison of accuracy of proposed model with independent LSTM and RNN mode

## V. Conclusion

The proposed research will use neural machine translation techniques to the Indian language pairs of English and Telugu, and evaluation metrics will be analyzed. In order to do this comparison, we will be using two separate datasets. When testing and training on the translated dataset, a hybrid approach involving Long Short Term Memory (LSTM), Recurrent Neural Networks (RNN), and Moth Flame Optimizer (MFO) is applied. Because of the precision with which our model was translating, it is clear that the method that has been proposed is of the utmost importance for the investigation of Neural Machine Translated datasets. Most notably, it is shown that a decent level of translation accuracy may be accomplished on Indian languages with less resources.

## References


[1] Madankar, M.; Chandak, M.B.; Chavhan, N. Information retrieval system and machine translation: A review. Procedia Comput. Sci. 2016, 78, 845–850.

[2] Kenny, D. Machine translation. In The Routledge Handbook of Translation and Philosophy; Routledge: Oxfordshire, UK, 2018; pp. 428–445.

[3] Bahdanau, D.; Cho, K.H.; Bengio, Y. Neural machine translation by jointly learning to align and translate. arXiv2014, arXiv:1409.0473.

[4] Forcada, M.L. Making sense of neural machine translation. Transl. Spaces 2017, 6, 291–309.

[5] Nekoto, W.; Marivate, V.; Matsila, T.; Fasubaa, T.; Kolawole, T.; Fagbohungbe, T.; Akinola, S.O.; Muhammad, S.H.; Kabongo, S.; Osei, S.; et al. Participatory research for low-resourced machine translation: A case study in african languages. arXiv2020, arXiv:2010.02353.

[6] Freitag, M.; Firat, O. Complete multilingual neural machine translation. arXiv2020, arXiv:2010.10239.

[7] Ahmadnia, B.; Dorr, B.J. Augmenting neural machine translation through round-trip training approach. Open Comput. Sci. 2019, 9, 268–278.

[8] Johnson, M.; Schuster, M.; Le, Q.V.; Krikun, M.; Wu, Y.; Chen, Z.; Thorat, N.; Viégas, F.; Wattenberg, M.; Corrado, G.; et al. Google's Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation. Trans. Assoc. Comput. Linguist. 2017, 5, 339–351.

[9] Kapil Sharma, Rajiv Khosla, Yogesh Kumar. (2023). Application of Morgan and Krejcie & Chi-Square Test with Operational Errands Approach for Measuring Customers' Attitude & Perceived Risks Towards Online Buying. International Journal of Intelligent Systems and Applications in Engineering, 11(3s), 280–285. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2685

[10] Aharoni, R.; Johnson, M.; Firat, O. Massively multilingual neural machine translation. arXiv2019, arXiv:1903.00089.

[11] Koehn, P.; Knowles, R. Six challenges for neural machine translation. arXiv2017, arXiv:1706.03872.

[12] Lakew, S.M.; Federico, M.; Negri, M.; Turchi, M. Multilingual neural machine translation for low-resource languages. Ital. J. Comput. 2018, 4, 11–25.

[13] Zoph, B.; Yuret, D.; May, J.; Knight, K. Transfer learning for low-resource neural machine translation. arXiv2016, arXiv:1604.02201.

[14] Goyal, V.; Kumar, S.; Sharma, D.M. Effificient neural machine translation for low-resource languages via exploiting related languages. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop, Online, 5–10 July 2020; pp. 162–168.

[15] Chowdhury, K.D.; Hasanuzzaman, M.; Liu, Q. Multimodal neural machine translation for low-resource language pairs using synthetic data. In Proceedings of the Workshop on Deep Learning Approaches for Low-Resource NLP, Melbourne, Australia, 19 July 2018; pp. 33–42.

[16] Fadaee, M.; Bisazza, A.; Monz, C. Data augmentation for low-resource neural machine translation. arXiv2017, arXiv:1705.00440.

[17] Imankulova, A.; Sato, T.; Komachi, M. Improving low-resource neural machine translation with fifiltered pseudo-parallel corpus. In Proceedings of the 4th Workshop on Asian Translation (WAT2017), Taipei, Taiwan, 27 November–1 December 2017; pp. 70–78.

[18] Gu, J.; Wang, Y.; Chen, Y.; Cho, K.; Li, V.O.K. Meta-learning for low-resource neural machine translation. arXiv2018, arXiv:1808.08437.

[19] Ueffing, N.; Haffari, G.; Sarkar, A. Transductive learning for statistical machine translation. In Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, Prague, Czech Republic, 25–27 June 2007; pp. 25–32.

[20] Tonja, A.L.; Woldeyohannis, M.M.; Yigezu, M.G. A Parallel Corpora for bi-directional Neural Machine Translation for Low Resourced Ethiopian Languages. In Proceedings of the 2021 International Conference on Information and Communication Technology for Development for Africa (ICT4DA), Bahir Dar, Ethiopia, 22–24 November 2021; pp. 71–76.

[21] Laskar, S.R.; Paul, B.; Adhikary, P.K.; Pakray, P.; Bandyopadhyay, S. Neural Machine Translation for Tamil–Telugu Pair. In Proceedings of the Sixth Conference on Machine Translation, Online, 10–11 November 2021; pp. 284–287.

[22] Pennington, J.; Socher, R.; Manning, C.D. Glove: Global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1532–1543.

[23] Marie, B.; Fujita, A. Synthesizing Monolingual Data for Neural Machine Translation. arXiv2021, arXiv:2101.12462.

[24] Tars, M.; Tättar, A.; Fišel, M. Extremely low-resource machine translation for closely related languages. arXiv2021, arXiv:2105.13065.

[25] Sennrich, R.; Haddow, B.; Birch, A. Improving neural machine translation models with monolingual data. arXiv2015, arXiv:1511.06709.

[26] Jiao, W.; Wang, X.; Tu, Z.; Shi, S.; Lyu, M.R.; King, I. Self-training sampling with monolingual data uncertainty for neural machine translation. arXiv2021, arXiv:2106.00941.

[27] Dione, C.M.B.; Lo, A.; Nguer, E.M.; Ba, S. Low-resource Neural Machine Translation: Benchmarking State-of-the-art Transformer for Wolof<-> French. In Proceedings of the Thirteenth Language Resources and Evaluation Conference, Marseille, France, 20–25 June 2022; pp. 6654–6661.

[28] Pham, N.L. Data Augmentation for English-Vietnamese Neural Machine Translation: An Empirical Study. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4216607.

[29] Ngo, T.-V.; Nguyen, P.-T.; Nguyen, V.V.; Ha, T.-L.; Nguyen, L.-M. An Efficient Method for Generating Synthetic Data for LowResource Machine Translation: An empirical study of Chinese, Japanese to Vietnamese Neural Machine Translation. Appl. Artif. Intell. 2022, 36, 2101755.

## About The Authors:

Mrs. G. Srisudha has obtained M.Tech in Computer Science and Engineering From Jawaharlal Nehru Technological University in the year 2011. At the moment, she is enrolled in the Ph.D. programme in Full-Time Study at the Department of Computer Science and Systems Engineering (CS&SE), Andhra University, India. She has 12 years of teaching experience and her area of interests are in the fields of Artificial Intelligence, Machine learning, deep learning, Natural language Processing, Artificial Intelligence, Python, Java and Neural Networks.

Email:srisudha.garugu@gmail.com

Prof. D. Lalitha Bhaskariis currently employed as a professor in the department of Computer Science and Systems Engineering (CS&SE) at Andhra University's AUCE(A) campus as well as the Coordinator of the IQAC Visakhapatnam. She has been a teacher for the past 26 years, and her research focuses on cryptography and network security, computer vision, cyber forensics, and deep learning. Her teaching career also spans 26 years. She has roughly 68 papers that have been published in reputable international journals in addition to 38 papers that have been published in the proceedings of international conferences. She has worked as a consultant on international projects pertaining to deep learning and picture segmentation. In addition to being a life member of the CSI, she is a member of several other reputable professional groups. She has authored a book in image processing and compiled a few books. She is a recipient of national awards like young scientist,bharatvikas etc.

Email: lalithabhaskari@yahoo.co.in