A Memory-Based Neural Network Model for English to Telugu Language Translation on Different Types of Sentences

Bilal Bataineh[1], Bandi Vamsi[2], Ali Al Bataineh[3], Bhanu Prakash Doppala[4]
Department of Computer Science, Jadara University, Irbid, Jordan[1]
Department of Artificial Intelligence, Madanapalle Institute of Technology & Science, Madanapalle, Andhra Pradesh 517325, India[2]
Artificial Intelligence Center, Norwich University, Northfield, VT 05663, United States[3]
Data Analytics, Generation Australia, 88 Phillip St, Sydney, NSW 2000, Australia[4]

***Abstract*—In India, regional languages play an important role in government-to-public, public-to-citizen rights, weather forecasting and farming. Depending on the state the language also changes accordingly. But in the case of remote areas, the understanding level becomes complex since everything nowadays is presented in the English Language. In such conditions, the regional language manual translation consumes more time to provide services to the common people. The automatic translation of one language to another by maintaining the meaning of the given input sentence there by producing the exact meaning in the output language is carried out through Machine Translation. In this work, we proposed a Memory Based Neural Network for Translation (MBNNT) model on simple, compound and complex sentences for English to Telugu language translation. We used BLEU and WER metrics for identifying the translation quality. On applying these metrics over different type of sentences LSTM showed promising results over Statistical Machine Translation and Recurrent Neural Networks in terms of the quality and performance.**

Keywords—Machine translation; English-Telugu translation; RNN; LSTM

I. Introduction

India is considered as the world's second highest number of languages. In India there are various regions based on their traditions and culture. These regions use their regional languages [1]. The main aim is that the Language translation system which is used in translation of different type of languages. This process makes it easy for people to communicate in their regional language [2]. The form of communicating among people is done through various sources like interchanging their thoughts, exchanging their ideas etc. Telugu language is considered as one of the ancient regional languages in India that is being used from centuries. Most of the southern states of India use Telugu as a commonly spoken language [3].

In English language, general conversations play an important role in understanding the language. Usually there are three type of sentences that are commonly used in these conversations, namely: simple, complex and compound sentences [4]. The simple sentences are those in which every other sentence is constructed upon them. These have only single independent clause that completes the sentence and hence called as simple sentences [5]. During translation to desired language the simple sentence can be easily translated since it has a single clause. In case of complex and compound sentences, they have two different independent clauses [6].

A. Compound Sentences

The compound sentences are constructed from the simple sentences, but have two independent clauses. This explains that each clause in compound sentence can be explained itself completely irrespective of the second clause [7]. The two clauses can be linked together by coordinating conjunctions. Those are FOR, AND, NOR, BUT, OR, YET, and SO which are explained below with an example each in Table I.

TABLE I. Conjunction words in compound sentences

English Sentence	Conjunction	Explanation	Telugu Connection
Arun doesn't speak Telugu, for he is from United States	For	Here both sentences are complete	కోసం
He lives in Ontario, and his friend's lives in Dubai.	And	Both the sentences are given equal importance	మరియు
I don't watch television, nor do I like movies.	Nor	Both are speaking negatively	లేదా
He is a professor, but his brother is a doctor.	But	Here the usage of 'but' is opposite	కాని
You can play cricket, or you can go for shopping.	Or	This clause shows an option	లేదా
They were not hungry, yet they went out for lunch	Yet	It refers that they have done something even they do not have to	ఇంకా
Sheetal was not well, so he visited the doctor.	So	First clause is the result of the second	కాబట్టి
English Sentence	Conjunction	Explanation	Telugu Connection

B. Complex Sentences

The complex sentences that are frequently used in the conversations also have two different clauses, where one is independent referring that it is a complete clause and the other is dependent clause [8]. This depends completely on one clause without the other the sentence cannot be completely understood. In complex sentences these two clauses are linked together through subordinating conjunctions [9]. There are six types of these subordinating conjunctions namely: contrast, cause, condition, time, place and relative pronouns [10]. During translation into Telugu Language, the subordinate clause always comes at the beginning of complex sentence [11]. The independent clause is always followed by a dependent clause as shown in Fig. 1.

Source Language: I like to eat chocolate, but I don't like to eat sweets.

Target Language: నేను చాక్లెట్ తినడానికి ఇష్టపడతాను, కానీ నేను స్వీట్లు తినడానికి ఇష్టపడను.

Dependent clause	Independent clause
నేను చాక్లెట్ తినడానికి ఇష్టపడతాను	కానీ నేను స్వీట్లు తినడానికి ఇష్టపడను

Fig. 1. Category 1 - Complex sentence in Telugu language.

In Telugu, the subordinate conjunction that appears in the dependent clause always comes at the end of the subordinate clause. Fig. 2 shows how the clauses are rearranged while translating.

Source Language: I will watch a movie, after I have gone to my native place.

Target Language: మా ఊరికి వెళ్లిన తర్వాత, సినిమా చూస్తాను.

Dependent clause	Independent clause
మా ఊరికి వెళ్లిన తర్వాత	సినిమా చూస్తాను

Fig. 2. Category 2 - Complex sentence in Telugu language.

C. Types of Conjunctions

Conjunctions are represented as parts of speech that connect the clauses, words and phrases in a sentence [12]. These conjunctions are divided into four types, namely coordinating, correlative, subordinating and conjunctive adverb. A few examples explaining the type of conjunctions are mentioned in Table II.

TABLE II. CONJUNCTION WORDS IN COMPOUND SENTENCES

English Sentence	Conjunction	Explanation	Telugu Connection
Arun doesn't speak Telugu, for he is from United States	For	Here both sentences are complete	కోసం
He lives in Ontario, and his friend's lives in Dubai.	And	Both the sentences are given equal importance	మరియు
I don't watch television, nor do I like movies.	Nor	Both are speaking negatively	లేదా
He is a professor, but his brother is a doctor.	But	Here the usage of 'but' is opposite	కాని

In this work, we consider Machine Translation (MT) for handling simple, compound and complex sentences from English language to Telugu Language [13-14]. The main aim of this translation is to generate the meaningful sentence without any grammatical errors obtained from source to desired target language. MT is a major study in the field of natural language processing (NLP) which is used in the translation of regional languages technically [15]. Form the past few years, neural machine translation (NMT) gained success and became the major part in applied MT. This mechanism depends on the availability of large corpus data and memory-based methods which gather verbal details through sequence-to-sequence information [16].

NLP is a language translation approach that supports in understanding the grammar and verbals of a language with its accuracy in analyzing the language. It is related to the advancement of methods that automatically strive to translate a language. Also, support in utilizing this language for conveyance among common people. This approach needs some information while translating and also needs some data related to the language including its grammar [17].

There are various methods which are developed to attain greater accuracy in translations namely: Statistical Machine Translation (SMT), Knowledge, Rule, and Corpus. These methods individually have their advantages and disadvantages. Of them, SMT is a sub-field of corpus translation and is commonly utilized. Since it shows good results when compared with other methodologies [18]. In recent days one of the popular methods called neural networks in MT is used all over the World. This has become a unique method of MT with the help of these neural nodes commonly termed NMT. The process of NMT to convert the source language to the target language is shown in Fig. 3.

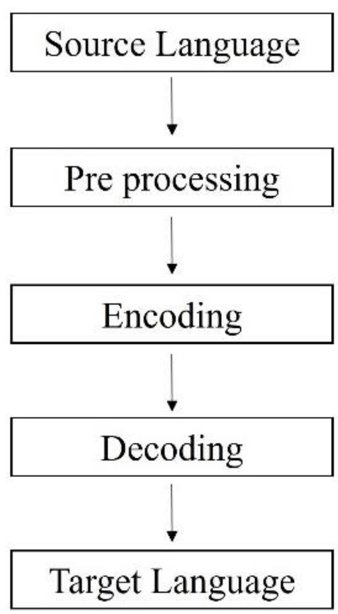

Fig. 3. Work flow of neural machine translation.

In this research work, we used a memory based neural machine translation models such as Recurrent Neural Network (RNN) and Long Short-term Memory (LSTM) for translation of English language to Telugu Language. The objectives of this work are arranged as follows:

- The given input English language sentence is divided into phrases and clauses.
- The generated phrases are identified depending upon the type of sentence.
- The type of sentence is identified at the pre-processing phase.
- Based on the type of sentence, the phrases are rearranged.
- These rearranged phrases are given to memory-based Neural Machine Translation to obtain the target Telugu language.

The remaining sections of this work are organized as follows: Section II discusses Related Works considered for this study, Section III discusses the methodology of this work, Section IV deals with the results and performance evaluation metrics, and Section V describes the conclusion and link to future work.

II. Related Work

An automated grading model for English translation based on NLP is created in order to decrease the burden of conventional grading and increase the effectiveness of scoring. It is suggested to use an attention LSTM of the English MT model. Initially, an English MT model based on LSTM focus encoding is defined; the framework level of the translation evaluation system is developed. This is done in accordance with the properties of the basic LSTM network, which uses resolved layered vectors to demonstrate phrases in the modelling phase [19].

The posterior distribution of a specific phrase or word combination in the translation process is statistically intended by using the defined linguistic structure of the language translation for calculating scores. The outcomes demonstrate that, when comparison to English MT models built using existing NN such as standard LSTM and RNN models, the LSTM focus embedding-based model developed in this study can improve the recognition of the input language relevant information and increase the accuracy of the English MT model and the reliability of the translation [19].

Researchers now understand the value of data with the need to examine this large amount of data according to the expanding digital environment. In required to conduct content analysis, it is necessary to classify a sizable volume of multilingual textual information. In this case, NMT is used to suggest a labelling with the use of annotated English sample, which is available in plenty. Labeling data for multilingual character recognition is difficult to retrieve [20]. In order to create context and categorize text into positive and negative sentiments groups, the proposed method would use vectors as embedding that are given to RNN and LSTM. The system for labelling texts that is utilized to collect tagged bilingual textual information is the main selling point. This model can categorize content analysis part of the text [20].

The effectiveness and desirability of English MT in application areas are both constrained by linguistic diversity, the restricted capacity to represent semantic features, and the scarcity of parallel corpus data. Due to highly parallel computing computational capability, which shortens the model training and enables it to record the lexical significance of all phrases in the sentence, the self-attention method has drawn a lot of interest in English MT. Moreover, because the self-attention process disregards the location and information about the structure among word vectors, its effectiveness is different from that of RNN. In order to exploit location information among phrases, the English MT model focused on self-process encode the actual location data of words [21].

An English translation system prototype has been created for mapping English and Chinese phrases through knowledge vectors, and it leverages RNN for both encodings. It is examined how well the activation function-based model performs. According to the study results, the decoding layers of the activation function and the encoder layers both exhibit the best result. According to the effectiveness of the LSTM and GRU stages, the GRU layer performs better. The nonlinear activation functions are used to set the attention layers [22].

Unlike the classic SMT, this neural machine translation focuses at developing a single neural model which can be collaboratively modified to maximize the translation quality. With the help of the LSTM approach, 47 multilingual food recipes from Spanish to English and English to Spanish language translation. This work produced new insights and useful guidance for developing and improving NMT. The BLEU metric is used to evaluate this model. According to the comparison results, the conversion of food recipes into English to Spanish has achieved a value of 0.998426 for BLEU with a ratio of 70% and 30% [23].

The complete LSTM analysis of the previous state is necessary for the subsequent LSTM phase. For a series of n nodes, this needs to be calculated ‘n’ times. The main cost components in this are the linear transforms needed for the LSTM gates and condition calculations. This approach comprehensive LSTM contextual analysis by calculating hidden layers and gates with an input signal and a straightforward bag-of-words encoding of the previous tokens contexts in order to allow sequential parallel computation of LSTMs. As a result, we can effectively measure each information step in comparison rather than the previously expensive sequenced linear functions. Then, using computationally affordable element wise procedures, we link the results of each concurrent phase [24].

In order to represent the semantic relationship between distant phrases in a text, a tree-based conversion models are used. However, it has issues with costly manual annotating costs and inaccurate automated annotations. This research focused on how to encode an input text language into a matrix in an un-supervised way to decode the target language. A Gumbel Tree-LSTM can learn to create tree hierarchies to attain sequence-to-sequence model by using both spoken media corpora [25].

By studying and refining real datasets, the SCN-LSTM (Skip Convolutional Network and Long Short Term Memory) language translation model is developed. In order to provide conceptual foundation for the study and use of the SCN-LSTM require similar in English instruction, the performing viability, translating accuracy, and scalability of the model are examined. To be more precise, the rate of translated ambiguity is reduced by 39.21% relative to LSTM model, and the scalability is 0.4 times of the N-tuple prototype in the SCN-LSTM language conversion [26].

From the research works [19-26], it can be concluded that while translating a sentence from one language to another the importance of grammar like finding the clauses, and maintaining the subject is mainly essential. Whereas, the above-mentioned works only highlighted the normal translation process using various neural network models. Based on this limitation, the aim of this work is to design a memory-based neural network by maintaining the quality of the sentence during translation.

III. METHODOLOGY

This section deals with the memory-based neural models for translation of English to Telugu language of simple, compound and complex sentences.

A. Preprocessing

During this phase, the type of the sentence is identified. The given input sentence may contain subject, verbs, main clause, conjunctions and punctuations. In order to identify the sentence, the main clause and conjunctions are stored in a memory for easy translation of a sentence. Usually, a simple sentence can be identified easily since it has only a single independent clause. In similar to it, the compound sentence can be identified by two main clauses. In the case of complex sentences, the main clause is combined together by one or more dependent clauses. In such cases the conjunctions come into the picture in joining a complex sentence. The memory associated with the networks acts as a connecting bridge in retrieving the information in translating such complex sentences.

B. Dataset

The data set used in this research work is collected from “Indian-Parallel-Corpa” [27], which contains 1263 English and Telugu sentences. For model validation purpose along with this dataset we have also created our own synthesized dataset which includes simple, compound and complex type of sentences. The synthesized dataset contains 150 simple, 480 compound and 970 complex sentences. The summary of the datasets used in this work is given in Table III. Based on the benchmark and synthesized data we have a total of 2,863 sentences making our model to achieve the high translation quality.

C. Neural Machine Translation

Neural Machine Translation (NMT) is commonly used in language translations to attain meaningful sentences by maintaining the quality. This uses only a small amount of memory like Statistical Machine Translation (SMT) approach. For extending the performance of translation every node is connected directly to the previous nodes in the entire neural network.

This neural network structure contains three basic layers namely: input, hidden and output as shown in Fig. 4. The English sentence is divided into words or phrases and is given to input layer. Each word in input layer is connected to different nodes within the neural network. These nodes are interconnected with each other in the hidden layer. This traditional network provides good translation quality while translating word-to-word or phrase-to-phrase in case of simple sentences.

The language translation not only depends on word to word translation but also re-ordering of the translated words in a particular language is also to be maintained correctly for a meaningful sentence. For example, the English sentence “I like watching movies” is translated into Telugu sentence as “నాకు సినిమాలు చూడటం ఇష్టం” . The word ‘like’ appeared as the second word in the English sentence and after translation, this word appeared at the last position in the Telugu language. For translating such simple sentences small amount of memory is required that do not make the network complex. In case of compound and complex sentences the amount of memory is to be increased depending on the complexity of the sentence by using Recurrent Neural Network (RNN) and Long-short Term Memory (LSTM).

TABLE III. SUMMARY OF DATASET

Type of dataset	Number of sentences	Type of sentence		
		Simple	Compound	Complex
Benchmark [27]	1263	253	379	631
Synthesized	1600	150	480	970

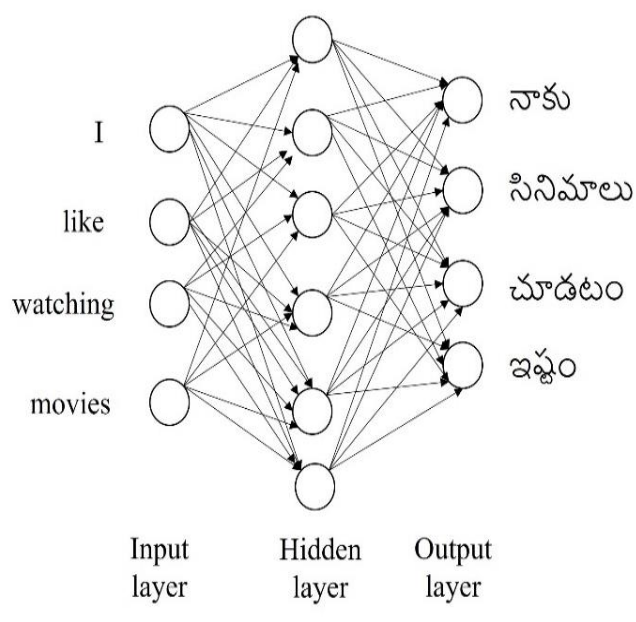

Fig. 4. Neural network for translation with hidden layers.

D. Proposed Model

From Fig. 5, our proposed model takes English sentence as input language. Based on the length of the sentence the number of phrases or words are identified. Depending on the clauses and type of conjunctions the input sentence is categorized into simple, compound and complex. If it is a simple sentence, the memory-based neural networks directly inputs the sentence and produce the desired translated sentence. If the given sentence is compound, the type of conjunction is identified and rearranged the sentence based on the conjunction and then given to memory-based neural network to produce the required output. If the given sentence is a complex sentence, the dependent and independent clauses are verified. Here the sentence is reordered to maintain the meaning of the language and the resultant sentence is given to memory-based neural network to attain the required translated sentence.

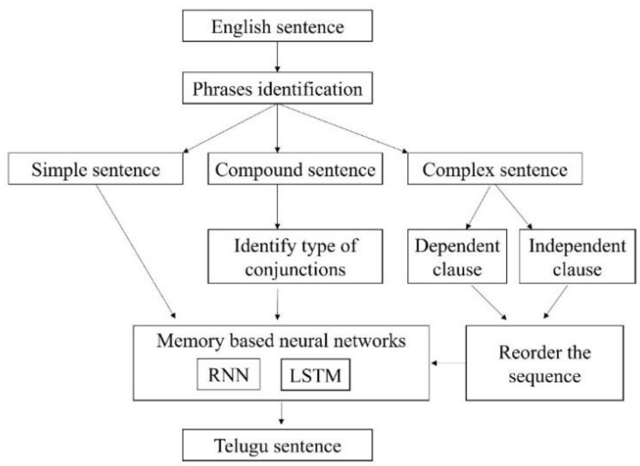

Fig. 5. Proposed model for a memory-based neural network for translation.

E. Recurrent Neural Network

A neural network is a differential function that maps one type of variable to another type of variable. Addressing the problem of predicting the type of event which is going to happen at every certain point is explained through Recurrent Neural Networks (RNN). These RNNs have loop like structures within them, which allow the information to pass. In RNNs the same weight is carried out in the entire recurrent unit. The RNNs have short-term memory that can store only limited amount of data. During translation process these RNNs use only nearby words that come in sequence. In this we use NLP task for completing the sentence formation.

In this work both the input and output sentences are sequences. The input is a certain word in a sentence and output is to predict the next word in the sequence with the help of proper training models. These are used in generating the sentence on its own. The sequence to sequence model had equal size of inputs and output, where most of them don't have such equal sequences. For instance, consider a 10 word sentence in English language as input. The final output may not have the 10 word sentence in Telugu Language. While at the time of texturization the input is a set of sentences, during translation process the training is carried out by the definition of the sentence given and is summarized to a group of sentences. In such cases, the encoder-decoder architecture is used in translating and producing the meaningful Telugu sentence.

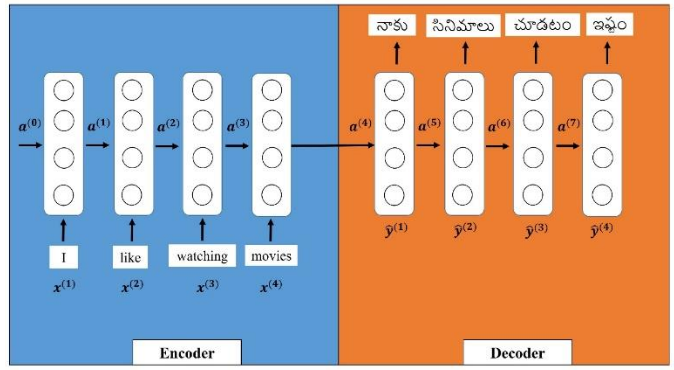

Fig. 6. Encoder and Decode in RNN.

Fig. 6, represents the encoder-decoder process for conversion of the sequence into a vector done by encoder whereas; the conversion of the vector into a sequence is done by the decoder. It takes the English sentence as input and converts it into its internal representation. This form of representation is a vector that holds the meaning of the English sentence. The decoder now takes this meaning of the vector and converts it into a sequence which is a Telugu sentence. The input is represented as $x^{(n)}$, the output is given by $\widehat{y^{(n)}}$ and the activation function is represented by $a^{(n)}$.

Internally when a sentence is passed as input every word is individually taken as node separately. The translation process is initiated only after the entire sequence of words is completed as shown in Fig. 7.

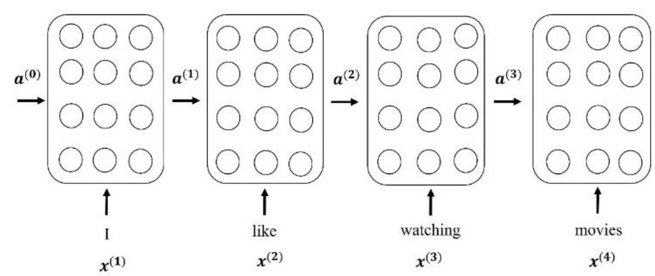

Fig. 7. Internal nodes of the hidden layer.

The sequence of the sentence can vary from 0 to infinity depending upon the sentence given. This is represented in Eq. (1) as follows:

$$x^{(n)} = W^n x^{(0)} \tag{1}$$

In the recurrent network having n units, initially $x^{(0)}$ is a scalar vector at a rate W. After n iteration units its value would be $x^{(n)}$. Since this is a dynamical system, we represent it in the form of Eq. (2).

$$W^n\, x^{(0)} \rightarrow \{\infty; W > 1 | 0; W < 1\} \tag{2}$$

For every n value the $x^{(n)}$ for $W > 1$ the value $W^n\, x^{(0)}$ explodes and for $W < 1$ value $W^n\, x^{(0)}$ becomes 0 or vanishes. If this happens, it leads to loss of the information which is given as input depending on their weights. This problem is called as vanishing gradient problem. The loss of information in these cases can be prevented in four ways.

- The usage of skip connections is required to eliminate the loss.
- By actively removing the connections having Length as 1 and replace them with longer connections. These force the network to move along in a modified path.
- Using the Leaky Recurrent Units, by adding a constant as shown in Fig. 8, over every edge that joins the network. This constant regulates the amount of information that the network which it has to remember over a certain period of time. If this constant is closer to value 1, more the memory is retained and if this is closer the value 0, memory of the previous state gets erased or vanishes.
- An enhancement to leaky recurrent network is the usage of Gated Recurrent Neural Networks (GRNN).

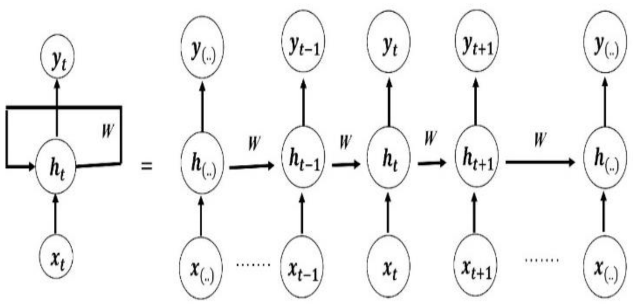

Fig. 8. Leaky RNN.

The expanded hidden layers of leaky RNN are represented in Eq. (3) which identifies any leakage or missing words within the network.

$$h^{(t+1)} = f(h^{(t)}, x^{(t+1)}) \tag{3}$$

Instead of assigning a constant value, we introduce a set of parameters one for every time lapse. Based on this, the network itself decides which sequence is to be remembered and what is to be erased. These set of parameters act as gates at every state of the network. In RNNs based on the beginning of any word, the ending of the sentence is predicted. These do not remember what appeared at the beginning of the sentence. They are fed with each word in the sentence depending on the weight's activation action is performed. It contains only a single layer, and relies on the time interval or time axis. In predicting the further words in a sentence, the initial words of the sentence are to be remembered. While in RNN, it is a short-term memory and hence is not recommended for auto completion of a sentence. One of the commonly used GRNN architecture is used in maintaining the memory of words for auto completion of a sentence that is Long Short Term Memory (LSTM). This is determined as the upgraded version of RNN which solve the problem of short term memory.

F. Long-short Term Memory (LSTM)

The name itself refers to extended memory during translation process called as LSTM. In this model we introduce a new state called long term memory along with short term memory in RNN. Every hidden layer is replaced with the LSTM or memory cell along with another connection for every cell called as cell state as shown in Fig. 9. Since there are many parameters in the layer, to avoid this we also use two gates namely 'update' and 'reset' gates.

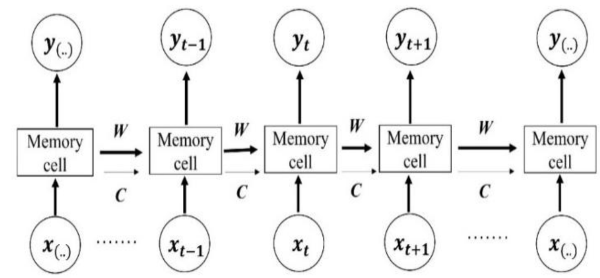

Fig. 9. Memory Cell Representation of LSTM.

The key words and main words in a given sentence are stored in long term memory for predicting the end words for auto completion of a sentence. These are stored until it finds the next keyword in the sequence. Depending on the current sentence certain keywords are added and previous sentence keywords are erased or deleted in auto completion process. This entire process is carried out in the training phase of the RNN. In training phase, the group of words is given to understand the structure in RNN based on which words are to be discarded and which words are to be stored.

LSTM contains three gates namely input gate, forget gate and output gate as shown in Fig. 10. The input gate control whether the memory cell is updated and is represented by Eq. (4).

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i) \tag{4}$$

The forget gate controls if the memory cell is set to 0 and is represented by Eq. (5). The output gate controls whether the information of the current cell state is made visible and is given by Eq. (6). Since RNN structure contains loops, to overcome this and to constitute the smooth curves in the range 0 to 1 the sigmoid function is used. Apart from these gates there is Vector $\bar{C}$ ', which modifies the cell state at the time of sigmoid activation as shown in Eq. (7). The 'tanh' has a 0 centered range performs sum operation is given by Eq. (8). This distributes the gradients equally among them. It also allows the cell state information to flow longer without erasing or exploding until the requirement of the certain keyword.

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f) \tag{5}$$

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o) \tag{6}$$

$$\bar{C}_t = tanh(W_c * [h_{t-1}, x_t] + b_c) \tag{7}$$

$$h_t = o_t * \tanh(C_t) \tag{8}$$

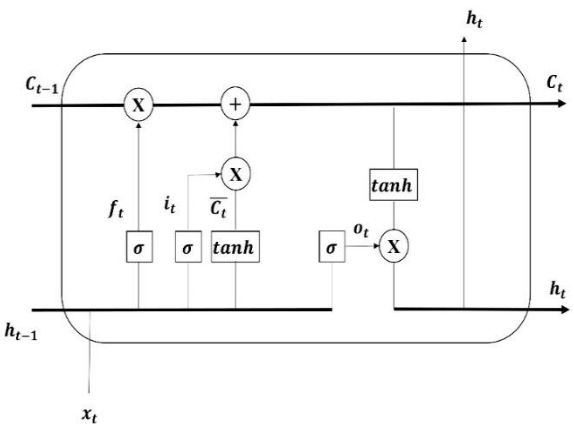

Fig. 10. Internal structure of LSTM memory cell.

It contains a previous hidden state current word by restricting the numbers between 0 and 1 during input. The restriction is maintained to discard the previous memory and gives a vector as output which has all the 0 and near to 0's. X defines the previous memory cell state that will be 0 and is determined by the forget gate. Every state has some weights attached while the sequence of words is given as input. Finally, the output is obtained by Eq. (9) which combines the sigmoid function along with the 'tanh' function and is given by the short-term memory.

$$C_t = f_t * C_{t-1} + i_t * \bar{C}_t \tag{9}$$

IV. Results and Discussion

This section represents the outcome of memory-based neural models for translation on simple, compound and complex sentences with BLEU and WER metrics.

A. Bilingual Evaluation Understudy Metric (BLEU)

The BLEU metric is used in analyzing the standard of the sentence which is translated through a machine from one language to another language. The term standard is determined as the reference between the output given by the machine with that of human translation. The BLEU scores are evaluated for each translated phrase by comparing them with a group of standard quality translation references available in the corpus is represented by Eq. (10). This score always represents a number between 0 and 1. The values which are closer to '1' represent the more common similar context between source and target language.

$$BLEU = \min\left(1, \frac{output\ context}{reference\ context}\right) * \prod_{i=2}^{4} precision_i \tag{10}$$

Where i ranges from 2 to 4, that represents the number of words such as 2-gram, 3-gram and 4-gram, $precision$represents the number of phrases in the given input sentence occurring in the reference sentence to total number of phrases in the input sentence.

B. Word Error Rate (WER)

The WER metric is also considered as one of the common metrics in evaluating the performance of Machine Translation models. It is used for distinguishing various systems and also for analyzing enhancements in one particular system. This is attained by arranging all the observed sequence of words with the reference sentence through dynamic sequence alignment. This can be calculated by Eq. (11).

$$WER = \frac{count(substitutions \cup deletions \cup insertions)}{Total\ number\ of\ words\ in\ the\ reference} \tag{11}$$

C. Performance of Proposed Model

The count of words in a sentence are represented as n-gram. In this work, the translation quality is measured on simple, compound and complex sentences of 2-gram, 3-gram and 4-grams. From Tables IV to VI, represents the translation quality on simple, compound and complex sentences among memory based neural networks and statistical model under BLEU and WER metrics. The visualization of these metric outcomes as shown in Fig. 11.

From Table IV, it can be observed that for a simple sentence the BLUE score for LSTM for all 2-gram, 3-gram and 4-gram is 0.89, 0.79 and 0.76 respectively. The BLEU score is higher for the LSTM model in all considered grams when compared with SMT and RNN. The rate of error obtained by LSTM is lesser with the values of 0.23, 0.21 and 0.20 when compared with other models. The model SMT doesn't have stored memory and obtained more error rate and lesser BLEU scores when compared to memory-based models like RNN and LSTM. Even though RNN and LSTM are memory-based models, with the help of our proposed model we can analyze similar outcomes between them in the case of simple sentences.

TABLE IV. Translation Quality of Simple Sentence

Model	2-gram		3-gram		4-gram	
	BLEU	WER	BLEU	WER	BLEU	WER
SMT	0.71	0.43	0.68	0.38	0.62	0.32
RNN	0.83	0.31	0.75	0.26	0.71	0.24
LSTM	0.89	0.23	0.79	0.21	0.76	0.20

Table V, represents the translation quality measures on 2-gram, 3-gram and 4-gram which are applied on compound sentences. On all these grams, the performance of the LTSM model is better performed with higher BLEU and lesser WER scores such as 0.91, 0.89, 0.82 and 0.14, 0.19, 0.20 respectively. This shows that LSTM considers every minute change of 0.01 in the given sentence and is calculated by showing better results when compared to RNN.

TABLE V. Translation Quality of Compound Sentence

Model	2-gram		3-gram		4-gram	
	BLEU	WER	BLEU	WER	BLEU	WER
SMT	0.76	0.41	0.71	0.33	0.68	0.35
RNN	0.87	0.25	0.82	0.23	0.74	0.21
LSTM	0.91	0.14	0.89	0.19	0.82	0.20

Finally in case of complex sentence results shown in Table VI, it can be clearly examined that using the memory based LSTM model is very advantageous for obtaining the best results when compared to that of other models.

TABLE VI. TRANSLATION QUALITY OF COMPLEX SENTENCE

Model	2-gram		3-gram		4-gram	
	BLEU	WER	BLEU	WER	BLEU	WER
SMT	0.79	0.37	0.75	0.31	0.71	0.26
RNN	0.89	0.23	0.82	0.17	0.85	0.14
LSTM	0.94	0.19	0.89	0.14	0.92	0.11

On comparison of translation quality of different types of sentences, memory cell based neural network model LSTM attained higher BLEU metric score and least WER metric scores on 2-gram, 3-gram and 4-gram sentences from English to Telugu language translation. The visualization of neural network models performance based on BLEU and WER metrics as shown in Fig. 11.

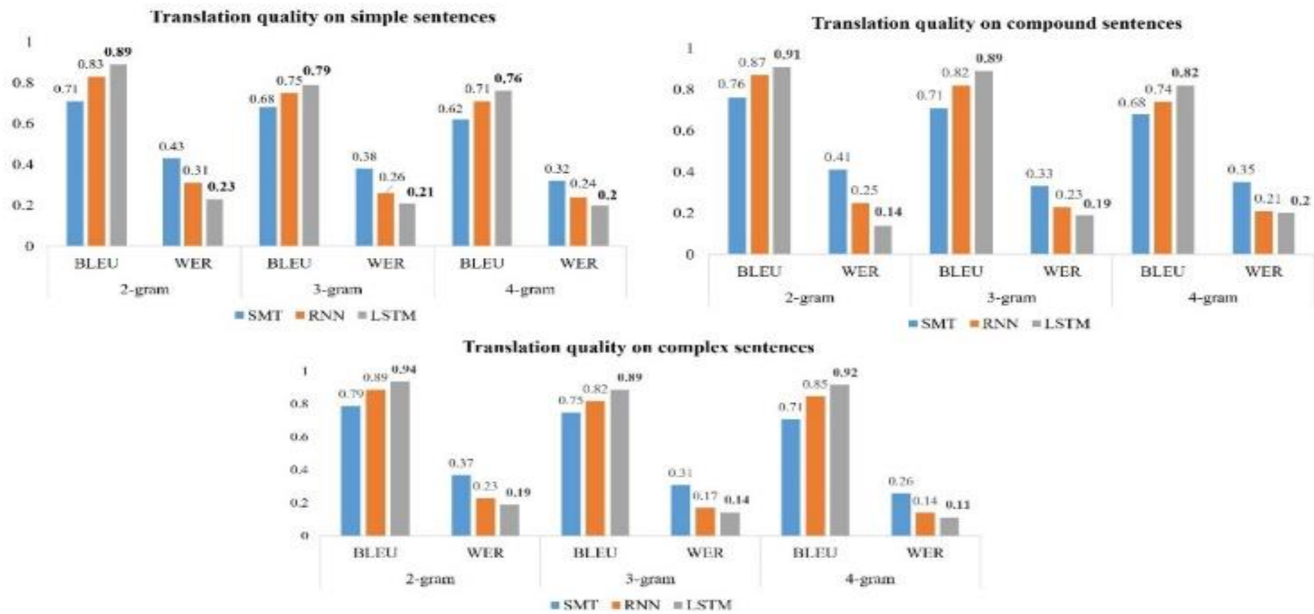

Fig. 11. Translation quality performance of neural network models.

V. CONCLUSION AND FUTURE SCOPE

This work is carried out on English to Telugu Language translation, since Telugu language is one of the fifth most common spoken language across India. During this process of translation, the word to word mapping becomes complex because this requires various reordering mechanisms and large corpus data. In SMT approach, the phrase to phrase mappings from input to output becomes complex due to the absence of memory cell and also the need for reordering becomes high. To overcome this limitation a memory based neural network is required to maintain proper syntax and semantics among source and target language. In this proposed work, based on the phrases the type of sentence is identified whether it is simple, compound or complex sentence. Depending on the type of sentence, the independent and dependent phrases are rearranged to obtain decent translation quality. In future work, there is a scope to reuse the existing memory cell for every word instead of creating a new memory every time. By using this approach, the complexity of designing the model becomes easier to handle any type of sentence.

REFERENCES

[1] Guo X, “Optimization of English Machine Translation by Deep Neural Network under Artificial Intelligence,” Computational intelligence and neuroscience, 2003411, 2022. https://doi.org/10.1155/2022/2003411.

[2] Ying, W., Li, L., & Yang, Z. X., “A Machine Translation Framework Based on Neural Network Deep Learning: from Semantics to Feature Analysis,” In 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), pp. 458-461, IEEE, 2022.

[3] Ren, B. “Neural Network Machine Translation Model Based on Deep Learning Technology,” In International Conference on Multi-modal Information Analytics," pp. 643-649, Springer, Cham, 2022.

[4] Xiao, L, “Machine English Translation Scoring System Based on Deep Neural Network in the Internet of Things Environment,” In the International Conference on Cyber Security Intelligence and Analytics, pp. 593-600, Springer, Cham, 2022.

[5] Nagaraj, P. K., Ravikumar, K. S., Kasyap, M. S., Murthy, M. H. S., & Paul, J, “Kannada to English Machine Translation Using Deep Neural Network”, Ingénierie des Systèmes d Inf., 26(1), 123-127, 2021.

[6] Tian, Y., Khanna, S., & Pljonkin, A., “Research on machine translation of deep neural network learning model based on ontology”, Informatica, 45(5), 2021.

[7] Zhao, L., Gao, W., & Fang, J, “High-Performance English–Chinese Machine Translation Based on GPU-Enabled Deep Neural Networks with Domain Corpus,” Applied Sciences, 11(22), 10915, 2021.

[8] Zhao, J., “Optimization of machine online translation system based on deep convolution neural network algorithm,” Computational Intelligence and Neuroscience, 2021.

[9] Shi, Y., Lu, J., Gu, S., Wang, Q., & Zheng, X., “Semantic-Aware Deep Neural Attention Network for Machine Translation Detection,” In China Conference on Machine Translation, pp. 63-76, Springer, Singapore, 2021.

[10] Singh, M., Kumar, R., & Chana, I., “Corpus based machine translation system with deep neural network for Sanskrit to Hindi translation,” Procedia Computer Science, 167, 2534-2544, 2020.

[11] Zhou, L., Zhang, J., Kang, X., & Zong, C., “Deep Neural Network--based Machine Translation System Combination,” ACM Transactions

on Asian and Low-Resource Language Information Processing (TALLIP), 19(5), 1-19, 2020.

[12] Liu, Y. P., Ma, C. G., & Zhang, Y. N, "Hierarchical machine translation model based on deep recursive neural network," Chin. J. Comput, 40(4), 861-871, 2017.

[13] Qing-dao-er-ji, R., Su, Y. L., & Liu, W. W, "Research on the LSTM Mongolian and Chinese machine translation based on morpheme encoding," Neural Computing and Applications, 32(1), 41-49, 2020.

[14] Xiao, Q., Chang, X., Zhang, X., & Liu, X., "Multi-information spatial–temporal LSTM fusion continuous sign language neural machine translation," IEEE Access, 8, 216718-216728, 2020.

[15] Cui, Y., Wang, S., & Li, J., "LSTM neural reordering feature for statistical machine translation," arXiv preprint arXiv:1512.00177, 2015.

[16] Zhang, J., & Zong, C., "Deep Neural Networks in Machine Translation: An Overview," IEEE Intell. Syst., 30(5), 16-25, 2015.

[17] Bensalah, N., Ayad, H., Adib, A., & Ibn El Farouk, A., "LSTM vs. GRU for Arabic Machine Translation," In International Conference on Soft Computing and Pattern Recognition, pp. 156-165, Springer, Cham, 2020.

[18] Vamsi, B., Al Bataineh, A., Doppala, B.P., "Lexical based reordering models for English to Telugu machine translation," Revue d'Intelligence Artificielle, Vol. 37, No. 5, pp. 1109-1120, 2023. https://doi.org/10.18280/ria.370503.

[19] Mahanty, M., "A Corpus-Based Auto-encoder-and-Decoder Machine Translation Using Deep Neural Network for Translation from English to Telugu Language," SN COMPUT. SCI. 4, 354, 2023. https://doi.org/10.1007/s42979-023-01678-4.

[20] Baliyan, A., Batra, A., & Singh, S. P., "Multilingual sentiment analysis using RNN-LSTM and neural machine translation," In 2021 8th International Conference on Computing for Sustainable Global Development, pp. 710-713, IEEE, 2021.

[21] Pan, W., "English Machine Translation Model Based on an Improved Self-Attention Technology," Scientific Programming, 2021.

[22] Li, B., "Research on English Translation Based on Recursive Deep Neural Network," In 2021 3rd International Conference on Artificial Intelligence and Advanced Manufacture, pp. 483-487, 2021.

[23] Dedes, K., Utama, A. B. P., Wibawa, A. P., Afandi, A. N., Handayani, A. N., & Hernandez, L., "Neural Machine Translation of Spanish-English Food Recipes Using LSTM," JOIV: International Journal on Informatics Visualization, 6(2), 290-297, 2022.

[24] Xu, H., Liu, Q., van Genabith, J., Xiong, D., & Zhang, M., "Multi-head highly parallelized LSTM decoder for neural machine translation," In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Volume 1: Long Papers, pp. 273-282, 2021.

[25] Su, C., Huang, H., Shi, S., Jian, P., & Shi, X., "Neural machine translation with Gumbel tree-LSTM based encoder", Journal of Visual Communication and Image Representation, 71, 102811, 2020.

[26] Ren, B., "The use of machine translation algorithm based on residual and LSTM neural network in translation teaching", Plos one, 15(11), e0240663, 2020.

[27] Joshua-Decoder, "indian-parallel-corpora/te-en/dev.te-en.en.0 at master," joshua-decoder/indian-parallel-corpora, GitHub, https://github.com/joshua-decoder/indian-parallel-corpora/blob/master/te-en/dev.te-en.en.0