International Journal of Machine Learning and Cybernetics (2021) 12:2365–2376
https://doi.org/10.1007/s13042-021-01317-5

ORIGINAL ARTICLE

Machine translation using deep learning for universal networking language based on their structure

Md. Nawab Yousuf Ali[1] · Md. Lizur Rahman[1] · Jyotismita Chaki[2] · Nilanjan Dey[3] · K. C. Santosh[4]

Received: 11 May 2020 / Accepted: 30 March 2021 / Published online: 27 April 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract
This paper presents a deep learning-based machine translation (MT) system that translates a sentence of subject-object-verb (SOV) structured language into subject-verb-object (SVO) structured language. This system uses recurrent neural networks (RNNs) and Encodings. Encode embedded RNNs generate a set of numbers from the input sentence, where the second RNNs generate the output from these sets of numbers. Three popular datasets of SOV structured language i.e., EMILLE corpus, Prothom-Alo corpus and Punjabi Monolingual Text Corpus ILCI-II are used as two different case-study to validate. In our experimental case-study 1, for the EMILLE corpus and Prothom-Alo corpus dataset, we have achieved 0.742, 4.11 and 0.18, respectively as Bilingual Evaluation Understudy (BLEU), NIST (metric) and tertiary entrance rank scores. Another case-study for Punjabi Monolingual Text Corpus ILCI-II dataset achieved a BLEU score of 0.75. Our results can be compared with the state-of-the-art results.

Keywords Machine translation · Deep learning · Recurrent neural network · Encodings · Sequence-to-sequence learning

1 Introduction

Machine translation (MT) system has become an important area of research for many researchers as communication with the people of various nations, countries and cultures via social networks. MT translates the source language sentence into a targeted language sentence [1]. Thus, the application of MT of various languages has become a part of our daily life. Basically, MT systems, analyzes the inputted sentence and builds its grammatical structure to generate the translation into the target language structure [2–4]. Google translate is the most popular translator website which supports over 100 different human languages. According to the website of May 2017, more than 500 million people use Google translator daily [5]. The technology used to develop this translator is statistical MT. It changed the world by allowing people to communicate even after they do not know the language. Although it is a popular MT system, it does not provide efficient result for every sentence. The aim of this research is to build an efficient MT system that overcomes some of the limitations of previous MT systems. One of the major limitations for translating a SOV structured sentence into SVO structured language is semantic ambiguity. Semantic ambiguity refers to different meanings of one word. For this reason, translators provide logically meaningless output for each sentence. Moreover, language experts are required for adding new language or update the existing language rules to the translator. But our system overcomes these problems.

There are a lots of subject-object-verb (SOV) structured language in the world such as- Bengali, Hindi, Punjabi, Marathi, Japanese etc. Although our proposed MT system

✉ K. C. Santosh
santosh.kc@ieee.org

Md. Nawab Yousuf Ali
nawab@ewubd.edu

Md. Lizur Rahman
lizur.sky@gmail.com

Jyotismita Chaki
jyotismita.c@gmail.com

Nilanjan Dey
neelanjan.dey@gmail.com

[1] Department of Computer Science and Engineering, East West University, Dhaka 1212, Bangladesh

[2] School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India

[3] Department of Computer Science and Engineering, JIS University, Kolkata, West Bengal 700109, India

[4] KC's PAMI Research Lab, Computer Science, University of South Dakota, Vermillion, SD 57069, USA

is shown as SOV structural language for Bengali and English as SVO structural language, the process is similar for the same structural language. Bengali language follows subject-object-verb (SOV) structure whereas English follows subject-verb-object (SVO) structure. The general population in Bangladesh and two states (West Bengal and Tripura) in India use Bengali as their first dialect. According to a recent survey around 1/6th populace of the world is talking in Bengali [6]. Although 230 million people speak in Bengali, but only a few numbers of resources and tools are available for translating this language. On the other side, semantic ambiguity is one of the major problems in Bengali language [7]. More researches have been conducted to solve this problem over the past years. Some of the researches provide the best result for some particular parts, but none of them works efficiently for every part of a sentence. There are various approaches for machine translation such as- word-by-word MT, Rule Based MT, statistical MT, etc., [8]. Word-by-word MT is an approach which translates a sentence in word-by-word sequence. It does not provide efficient result for some languages like- English to Bengali, English to Hindi, English to Japanese, etc., [9]. Rule based MT system contains some set of rules which translate a sentence efficiently [10–12]. But it has limitations because the human language rule is not fixed, it's changing continuously. New things added to the language and updated the language continuously day by day. Statistical MT system solves those limitations of word-by-word and rule-based MT. Basically, Statistical MT analyzes the existing human translations to translate a sentence [13]. This translation system is popular as it generates more than one translated sentence for each translation and shows the most suitable one as a result [14–16]. Furthermore, the accuracy of this system is much better [17, 18]. But this technique also contains some disadvantages. This system requires a huge number of training data. To build and maintain this translation system is complicated compares to others. Moreover, new language experts are required for a new pair of languages translation [19–21].

In this study we have established an efficient system that will analyze, understand, and generate languages, which humans use naturally. Our proposed MT system contains four main parts—(1) analysis, (2) encoding, (3) decoding, and (4) generation. In the analysis part, source sentence is divided into meaningful words and every word is checked lexically. Two deep learning techniques—recurrent neural networks (RNNs) and encodings are used in encode part. One word at a time is sent into the RNNs and it generates a set of numbers using Encoding algorithms. In the decode part, another RNNs is trained to generate the word into the targeted language from the source language. In the generation part, those words are classified into syntactic categories based on verb, adverb, noun, pronoun, etc. Those words can be swapped their positions and combined with other words to create meaningful sentences in the targeted language. So, if we are able to generate the set of numbers from source language sentences using RNN and Encoding algorithm, it will be easier for us to translate the sentence into targeted language.

The rest of the paper organized as follows. We start with a discussion of various previous works along with various machine translation techniques such as word-by-word translation, rule-based translation, and statistical machine translation of MT in Sect. 2. The overall discussion of proposed model and deep learning algorithms (RNN and Encodings) are shown in Sect. 3. Section 4 focuses on dataset collection and experimental result analysis of this system whereas Sect. 5 shows the overall discussion of the system. Finally, we summarized the paper with some concluding remarks and future direction in Sect. 6.

2 Previous works

Many researchers contributed various methods for translating sentences from one language to another. More often, deep learning, machine learning, and artificial intelligence have been used. In [22], the contributors proposed a recurrent neural network (RNNs) based statistical machine translation model. They improved the quality of result by comparing the phrase table of statistical machine translation and the phrase table of their proposed method. The authors used Moses toolkit to develop the machine translation (MT) model. In [23], the authors have shown a tense and phrase-based English to Bengali transfer architecture using fuzzy Rule-Based approach. Based on the attributes of sentences they have categorized each sentence and organized them into a pattern. Their proposed system separated meaningful words from each sentence and arranges them according to the rules. After that, the system uses morphological analysis to reconstruct the sentence into the target language. They have also shown the efficiency of their work by comparing their experimental result with Google translator. In [24], authors proposed a Bengali DeConverter for translating Universal Networking Language to Bengali Language. They tested their DeConverter on UNL expressions of 300 Bengali sentences using a Russian and English Language Server. And found that their system generates 90% syntactically and semantically correct Bengali sentences with a UNL Bilingual Evaluation Understudy (BLEU) score of 0.76.

In [25], another phrase based statistical machine translation model has been discussed in English to Bengali sentence. Since there is not much training data available for Bengali language, authors have used a transliteration module to handle this problem. Finally, the effectiveness of this system has shown by BLEU, NIST and TER scores. They have claimed that the overall BLEU score of their proposed

model is 11.7 and for the short sentence, it is 23.3. In [26], the authors build a syntactic structure of Bengali and English sentence using context-free grammars in English assertive sentences. In this system, they have used a bilingual dictionary for the contextual information and the morphological properties of English to Bengali words. The transferring English language structure of the corresponding Bengali language structure with lexical meaning is required to build this system. They have also generated a prototype of their system and used a huge number of English assertive sentences. They have found that their system shows the efficient results compare to other systems.

As per the above discussion, some researchers have been conducted in the development of automatic translation of Bengali noun-based compound sentence in universal networking language (UNL) documents and develops some automatic software for it. In [27], the authors have shown the method of morphological analysis in Bengali words into UNL. They have discussed the techniques to develop a mediator language from Bengali language which can easily be converted into various languages and vice versa. In this paper, the authors have explained the morphological rules. Based on tense, subject, preposition etc., their morphological rules can modify the parts of speech of a sentence. In [28], the authors have endeavored to create machine translation, Bengali word references that address the association, substance and subtleties of the data. In [29], contributor grew minimal effort English to Bengali (E2B)—ANUBAD making an interpretation of English content into Bengali content, utilizing both guidelines based and change based MT conspires alongside three dimensions of parsing. Another endeavor by [30] was to build up a factual Bengali to English interpretation motor utilizing just basic Bengali sentences that contains a subject, an item and an action word.

The various MT systems have been proposed by various researchers. Those are word-by-word MT, Rule Based MT, statistical MT, etc. All of the systems perform well in some particular areas, but none of them does not provide efficient results in every translation. Here, we have discussed about various machine translation techniques along with figures and examples.

2.1 Word-by-word translation

This is the easiest and the simplest translation process. The main concept of this system is to replace each word of a sentence with the translated word in the target language [31]. Figure 1 shows an example of translating from Bengali to English language using word-by-word technique. The implementation process of this technique is easy because it requires a dictionary for word translation. But the resultant accuracy of this technique is poor, because it does not care any grammatical rule and word order of the targeted language.

2.2 Rule-based translation

To make the previous system more accurate, language specific grammars and rule-based translation system should be added. For example, the order of verbs and nouns might be swapped because in Bengali language verbs usually come after the noun on the basis of structure of subject-object-verb (SOV) unlike English where English follows SVO structure. And sometimes two words might be translated into one single word. Figure 2 shows an example of translating from Bengali to English language using grammars and context. If more grammars and rules are added, the system will be more efficient [32, 33].

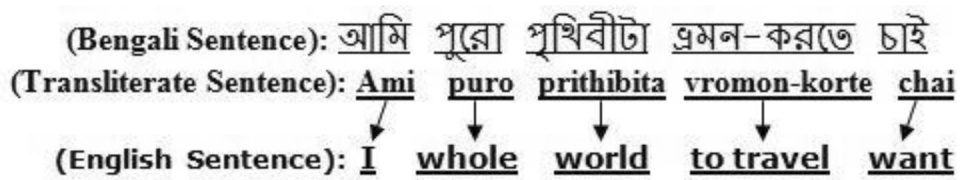

Fig. 1 Word-by-Word translation system

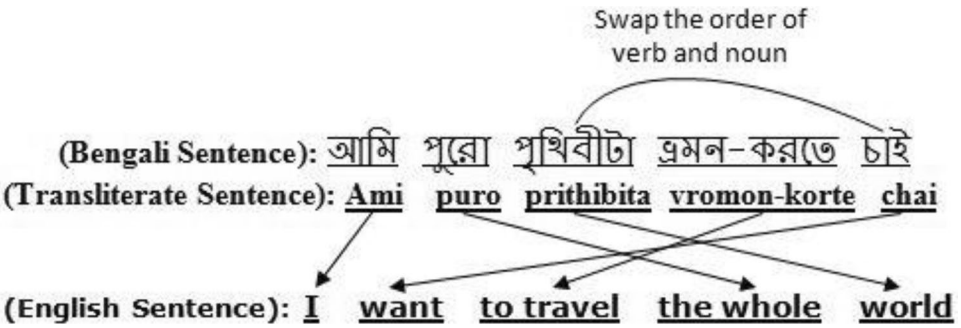

Fig. 2 Rule based translation system

Some rules for universal networking language (UNL) Bengali word dictionary are shown below:

The mapping of Bengali words to UW and their corresponding lexical-semantic attributes are stored in UNL Bengali word dictionary. Basically, the collection of entries in the UNL Bengali word dictionary is known as lexicon. Every entry is made out of three sorts of components: Universal Word (UW), Headword (HW) and Grammatical Attribute (GA). A HW is a documentation of a word in a local dialect making the information sentence. It is utilized as a trigger in acquiring equal UWs from a Word Dictionary during the time spent in re-conversion. A UW communicates the importance of a word is utilized in making UNL systems (i.e., UNL articulations) of yield. GAs is the data on how words carry on in a sentence and are utilized in re-conversion rules. Every lexicon passage has the accompanying configuration partner with any local dialect word [10].

This machine translation system provides good accuracy in written and plainly structured documents, such as simple article, and weather report [34]. It cannot work efficiently for the real-world documents. The main reason is human language does not follow a fixed set of rules [35]. Human languages are full of regional variations, special cases, and new rules. New rules are continuously coming into the language and old rules are continuously changing [35]. So, a supervised automated system is required for efficient translation of language, which can easily resolve those problems.

2.3 Statistical machine translation

Since rule-based translation systems contain lots of limitation, new translation systems were developed. Instead of grammars and rules of a sentence, the new translation system uses the statistics and probability to translate a sentence.

Information Format: *[HW]{ID}"UW"(ATTRIBUTE1, ATTRIBUTE2,...)<FLG, FRE, PRI>*
Here,
HW: Head Word (Bangla Word)
ID: Head Word Identification (omissible)
UW: Universal Word
ATTRIBUTE: Head Word Attribute
FLG: Language Flag
FRE: Head Word Frequency
PRI: Head Word Priority
Some examples of Bangla Word dictionary are presented below:
[**kukil**(কোকিল)] {} "Cuckoo (icl>bird)" (N, COMN, CEND, BIRD, ANI, WILD)
[**shiuli** (শিউলি)] {}" Jasmine (icl>plant)" (N, COMN, VEND, FLWR, PLNT)
[**Se** (সে)] {} "he(icl>person)"(PRON, HPRON, SUB, MALE, ANI,3SG)

In the dictionary entries, attributes N denotes noun, COMN for common nouns, ANI for animal object, CEND means consonant ended word, VEND means vowel ended word, FLWR for flower, PLNT for plant, PRON alludes to Pronoun, 3SG for third person singular number, HPRON to human pronoun, respectively.

This statistical machine translation system requires a huge number of training data [36]. The interesting thing of statistical machine translation systems is that unlikely previous two translation systems they do not generate only one translation for one sentence. Instead of generating only one translation they generate all possible translations and placed them in terms of rank [36]. Finally, outputs the lowest

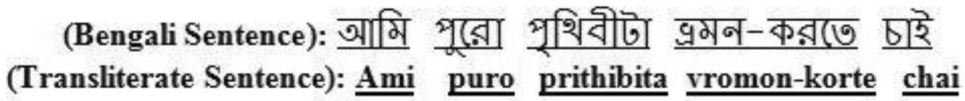

Fig. 3 Sentence is divided into meaningful words

Fig. 4 Meaningful word translation into targeted language

ranked translated sentence as the result. The whole process has been completed by three steps [13].

Step 1: Divide the sentence into meaningful words In the first step, divide the original sentence into simple, meaningful words that can be translated easily. Figure 3 shows how original sentence is divided into simple meaningful words. Here each underlined word is a meaningful word.

Step 2: Translate each meaningful word into the targeted language in all possible ways In the second step, find out how humans translated those words with all possible ways. Not only has the simple translation dictionary involved in looking up these words, but also how actual human translate these same words are involved. This process is much helpful to find out all the possible ways of translating each word. Figure 4 shows all possible ways of translating each word of a sentence.

Since each word contains lots of translated words in targeted language, some translated words are more meaningful and frequent than others. A score has been provided for each translated word based on how frequently these translations appearing the training data. For example, it is more common to everyone that "puro" means "the whole" rather than "the every". Based on how frequently "puro" translates to "the whole" in the training data, the system provides a score to each word.

Step 3: Make all possible sentences and take the suitable one Finally, the third step uses all possible combinations of translated words to find out the possible sentence. More than 2000 different sentences can be generated only from the translated words shown in Fig. 4. Some examples are shown below:

I | would like | to visit | the entire | universe
I | want | to travel | the whole | world
I | quest | to see | the full | earth
I | desire | to look over | the whole | universe

But in the real world, people do not use the same order to expose a sentence. They may use different order of words for the same sentence. Therefore, more possible sentences might be come. Some examples are shown below:

I | wish | to travel | the entire | earth
I | would like | to sally out | the total | world
I | want | to visit | the entire | terra
I | hope | to travel | the whole | glove

Now, the main challenge is to find out the sentence, which is more humanistic. To make sure of this the system compares every translated sentence with millions of real-world sentences. These real-world sentences are written on different types of books, magazines, newspapers, etc. Based on this technique, the system will generate a probability score for each translated sentence. For example, consider this possible translation:

I | quest | to see | the full | earth

It seems that nobody has ever written a sentence like this way and any similarity of this sentence will not be found in the dataset. Therefore, a low probability score has been given in this sentence. But look at this possible translation:

I | want | to travel | the whole | world

Since this is much meaningful sentence and many similar sentences will be found in the datasets, so a high probability score has been given in this sentence. Finally, the system will output a translated sentence, which contains the most probable score. If enough training data have given to this system, it will perform more accurately compared to other two translation systems. In the early of 2000s, Franz Josef Och used those ideas to make Google Translator [5]. And still now Google Translator is performing well.

Although statistical machine translation systems perform well, they have some limitations too. To build and maintain a statistical machine translation system is complicated compares to others. New language experts are required for multi-step translation pipeline when you want to translate new pair of languages. Building a new pipeline is not as easy as it contains lots of work. For example, if you want to translate Sinhala to Bosnian with Google translate, firstly Sinhala will translate into English as an intermediate language and then English to Bosnian translation will take place. Because there is not enough translation happens between Sinhala to Bosnian language pair. To make a translation directly from Sinhala to Bosnian, new language experts and multi-step translation pipeline are required.

3 Proposed method

In this paper, we have proposed a deep learning-based translation system, which can translate between two languages without human intervention. Our proposed translation system uses RNN and encoding algorithms. This section shows the proposed model along with an overview on RNN, and encodings algorithms.

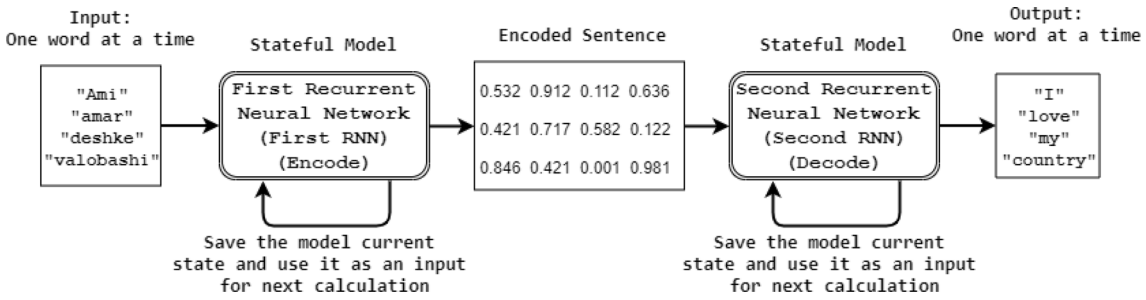

Fig. 5 Description of our proposed model

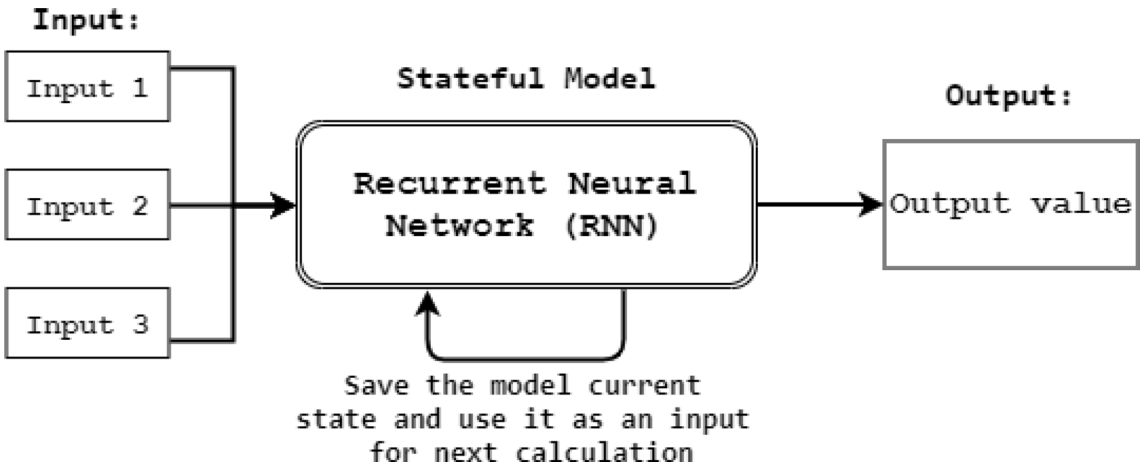

Fig. 6 Block diagram of RNN

3.1 Proposed model

Our proposed system is categorized into three stages.

1. We used the encoding technique to generate a series of unique numbers from a given sentence. The input sentence to the encodings is in subject-verb-object (SOV) structured language form.
2. We added RNN with encodings. In addition to encoding algorithm, we used a RNN so it can generate unique numbers only for one word at a time. The final result (series of unique numbers) for a given sentence generates after the last word was processed. In the process, we are not required to know the meaning of each encoded number because unique numbers are generated for each unique sentence.
3. Lastly, we added another RNN with the previous stage. Two RNNs were used to translate a sentence. The first RNN embedded with Encodings generates a series of measurement numbers only if the input sentences are SOV structured. While the second RNN decodes those generated numbers and generates the translated sentence. Input a sentence and encodes it, then decodes into the same language is not a useful idea. Therefore, we train our second RNN to decode the sentence into SVO structured language (English) instead of SOV (Bengali). Figure 5 shows the overall description of our proposed model.

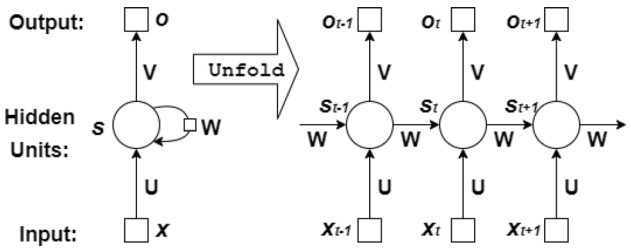

Fig. 7 Structure of a RNN being unfolded into the network [38]

3.2 Recurrent neural network (RNN)

To be able to use the sequential information in the calculation is the main idea of RNNs [37]. In a neural network, all

inputs and outputs are independent of each other. But this is not efficient for most of the systems. Suppose we want to build a system which can predict the next word we are going to write. In that case, we need to know the current word. Otherwise, it will be difficult of a system to predict the next word. In these types of applications, neural networks cannot perform well. To solve these problems a system is required, which can use the previous states result for the next state calculations. RNN is a little bit updated version of a neural network, where the result of previous state is one of the inputs for the next state. Unlike neural networks, next output depends on the previous output in RNNs. Figure 6 shows the block diagram of RNN and Fig. 7 shows the structure of a RNN being unfolded into a network.

Unfold means to write the complete sequence of the network. Suppose we have a sentence of four words, the network will be unfolded into a neural network of 4-layers, one layer per word. The working procedure of Fig. 7 is described as follows.

At the time step t, x_t denotes input, s_t denotes hidden state, and o_t denotes output of a sentence. The current state input and previous hidden state are required to calculate the next hidden state. The equation to calculate the hidden state is:

$$S_t = f\left(Ux_t + Ws_{t-1}\right). \tag{1}$$

Here, f is a nonlinear function and s_{t-1} is the previous hidden state. s_t is also known as the memory of this network. s_t store the previous state information at the time step t. The point to be noted is that s_t cannot store all previous state information, it can store only a few previous state information. Although a normal deep neural network uses various parameters on each layer, RNN uses the same parameters (here, *U, V, W*) in above all layers [39–42].

The probability distribution of a sequence can be learned by training to guess the next symbol by an RNN. In that case, the output of each stamp is conditionally distributed as:

$$p(xt|xt-1,\ldots,x1). \tag{2}$$

The probability of the sequence x can be measured by combining these probabilities using:

$$p(x) = \prod t = 1T \times p(xt|xt-1,\ldots,x1). \tag{3}$$

New sequence can be predicted easily by sampling a symbol in each time step.

3.3 Encodings

Encodings is a technique which represents something complicated into a simple way [43]. The most popular application of encodings is face recognition. For example, a system is required which can compare two different people's faces with a computer. Various measurement data from each face are needed to build this system. Those various measurements might be the spacing between eyes, the size of nose, lips and size of each ear, etc. Here, we can use a neural network to generate those measurements data. The system will compare those data of two faces to see if they are the same person. This idea of collecting a list of measurement data from a face is an example of encoding.

We can use the same concept of encodings to generate measurement values from a sentence. Figure 8 shows the list of measurement data F(X) from an inputted sentence X. This system is much more efficient, because the system only generates 128 numbers for each sentence.

As RNN reads each symbol, thus the hidden state of the RNN changes accordingly Eq. (1). After the end of each sequence reading, the hidden state of the RNN contains an overview c of the input sequence. In this proposed model, the decoder is another RNN which generates the output by predicting the next symbol o_t where the hidden state is s_t.

$$S_t = f\left(s_{(t-1)}, o_{(t-1)}, c\right). \tag{4}$$

Similarly, the conditional distribution of the next symbol for the activation functions given s and g is:

$$P\left(o_t|o_{t-1}, o_{t-2}, \ldots, o_1, c\right) = g\left(s_t, o_{t-1}, c\right). \tag{5}$$

The proposed RNNs two components encoder decoder are jointly trained to optimize the conditional log-likelihood,

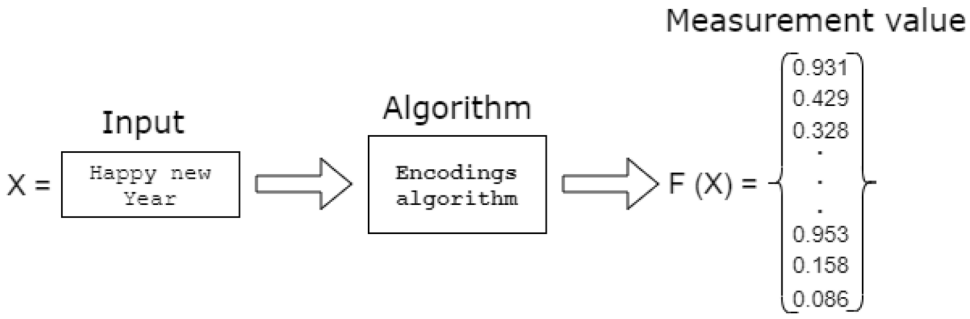

Fig. 8 List of measurements from Input Sentence

$$\max \frac{1}{N} \sum_{n=1}^{N} \log\, p(On\, V\, Xn), \tag{6}$$

where, Θ denotes the set of the model parameters, x_n denotes the input sequence, o_n denotes output sequence pair from the training set. When the RNN encoder-decoder trained, it can be used to predict scores of the given input and output sequences. The score shown in terms of probability $p\Theta(olx)$ from Eqs. (2) and (3). If the RNN encoder-decoder is equipped, we apply a new score to the current pair of sentences in each equation. This allows new scores to be inserted into existing tuning algorithms with minimal additional overhead for the calculation.

4 Experiments and results analysis

In this research, we have used two datasets for learning and testing purpose. One is an Enabling Minority Language Engineering (EMILLE) corpus [44] and another is Prothom-Alo corpus [45]. A corpus is the collection of written texts. EMILLE was developed by a joint research by Central Institute of Indian Languages (CIIL), India and Lancaster University, UK. Although EMILLE corpus contains three components (monolingual, parallel and annotated), but only two of them (monolingual and parallel) contains Bengali language data. The EMILLE monolingual corpus contains 1,867,452 words and EMILLE parallel contains 189,495 words of Bengali text data. Prothom-Alo corpus was developed by BRAC University, Bangladesh which contains 19,496,884 words of Bengali text data. We have also used "Facebook Graph API" tools for collecting the social network data by ourselves. This tool can extract data from any page or group such as restaurant, movie, etc. We have collected above 1800 sentences using this tool.

Before training the system, we encoded the files of the corpus to UTF-8. Then all the sentences extracted from XML to text. We have used an automatic sentence aligner to align those sentences. Since, a single evaluation tool is not enough to evaluate an MT system, we have used three evaluation tools in this system. BLEU (bilingual evaluation understudy) [46], NIST (metric) [47], and TER (tertiary entrance rank) [48] evaluation tools are used to evaluate our system.

The major problem for translating a sentence from one language to another is semantic ambiguity. Semantic ambiguity means different meanings of one word. For example, this Bengali word "khay" (খায়) has various meanings. The word "khay" (খায়) is a verb. In Bengali sentence, a verb comes after the noun. The word "khay" (খায়) changes its meaning based on the noun of a sentence. Similar case for some words like- "dekhar" (দেখার), "purbe" (পূর্বে), "suni" (শুনি) etc. Table 1 shows some examples of semantic ambiguity. For this reason, although a huge number of researches had been done for English to Bengali language translation, but a few research had been done for Bengali to English language. From those few researches, none of them address this problem. Even the world popular translator Google translates also cannot solve this problem. Our proposed system can solve this problem efficiently. Table 2 shows some Bengali sentences and their corresponding translated English sentences using various translation techniques. The evaluation result of our system describes the efficiency of it. Our system gets 0.742 BLEU score, 4.11 NIST score, and 0.18 TER score for the overall combined dataset. Table 3 shows the evaluation result. We have compared our evaluation result with Islam et al. [25] to find out the efficiency of our proposed system (Fig. 9).

Our proposed system is robust and generic for any SOV to SVO language translation. Our system provides good result in similar SOV structured language like Bengali e.g., Hindi, Sanskrit, Assamese, Oriya, Marathi, Punjabi, Nepali, Japanese, etc. Table 4 shows some sentences in various similar

Table 1 Examples of semantic ambiguity

Bengali sentence	English Sentence
সে ঘুষ **খায়**	He **takes** bribe
সে পানি **খায়**	He **drink** water
সে সিগারেট **খায়**	He **smoke** cigarettes
আমি চা **খাই**	I **take** tea
আমি টিভিতে খবর দেখার **সময় কফি** খাই	I **drink** coffee while **watching** the news on TV
আমি ঘুমানোর পূর্বে **গান** শুনি	I **listen** to music **before** I go to sleep

Table 2 Bengali to English translation in various translation techniques

Bengali sentence	English sentence using Word-by-Word translation	English sentence using Rule Based translation	English sentence using Google translate	English sentence using proposed method	Correct English sentence
সে ঘুষ **খায়**	He bribe eat	He eats bribe	He eats bribe	He takes bribe	He takes bribe
সে পানি **খায়**	He water eat	He eat water	He eats water	He drinks water	He drinks water
সে সিগারেট **খায়**	He cigarettes eat	He eat cigarettes	He ate cigarettes	He smokes cigarettes	He smokes cigarettes
আমি চা **খাই**	I tea eat	I eat tea	I eat tea	I take tea	I take tea
আমি টিভিতে খবর দেখার **সময় কফি** খাই	I TV news look time coffee eat	I eat coffee during looking news on TV	I eat coffee while watching news on TV	I drink coffee while watching the news on TV	I drink coffee while watching the news on TV
আমি ঘুমানোর পূর্বে **গান** শুনি	I sleep east song hear	I hear song before sleep	I listen music before sleep	I listen to music before I go to sleep	I listen to music before I go to sleep

Table 3 Evaluation of proposed method

Test Datasets	BLEU	NIST	TER
EMILLE	0.74	4.22	0.17
Prothom-Alo	0.765	4.97	0.13
Facebook	0.72	3.16	0.21
Combined	0.742	4.11	0.18

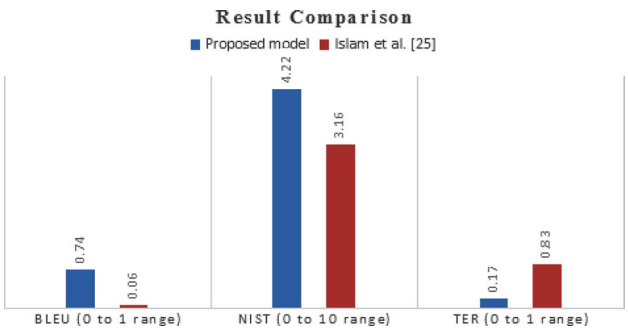

Fig. 9 Result comparison on EMILLE dataset of proposed method with Islam et al. [25]

structured languages and their corresponding translated English sentences using various translation techniques.

All we need to do is that to train the system with new language datasets. To find out the novelty, we have used Panjabi (SOV) and English (SVO) as another case study. We trained our system using Punjabi Monolingual Text Corpus ILCI-II dataset [49]. ILCI-II was initiated by Ministry of Electronics & Information Technology, Meity and collected by the Govt. of India, Jawaharal Nehru University, New Delhi. This corpus contains 30,000 sentences of general domain. We have found that our system got 0.75 BLEU score (Fig. 10).

In [24], authors proposed a DeConverter for Bengali language, where they used 300 Bengali sentences for check the efficiency of their system. Authors claimed that they have found 0.76 BLEU scores. To test the effectiveness of our system, we have trained our system with the exact same datasets and found that our system achieved a BLEU score of 0.79 (Fig. 11).

Table 4 Various Language to English translation in various translation techniques

Sentence (Language)	English sentence using Word-by-Word translation	English sentence using Rule Based translation	English sentence using Google translate	English sentence using proposed method	Correct English sentence
আমি চা খাই **(Bengali)**	I tea eat	I eat tea	I eat tea	I take tea	I take tea
मैं चाय खाता हूं (Hindi)	I tea eat	I eat tea	I eat tea	I take tea	I take tea
ਮੈਂ ਚਾਹ ਖਾਂਦਾ ਹਾਂ (Punjabi)	I tea eat	I eat tea	I eat tea	I take tea	I take tea
मी चहा खातो (Marithi)	I tea eat	I eat tea	I eat tea	I take tea	I take tea
म चिया खान्छु (Nepali)	I tea eat	I eat tea	I eat tea	I take tea	I take tea
わたしはお茶を食べます。 (Japanese)	I tea eat	I eat tea	I eat tea	I take tea	I take tea

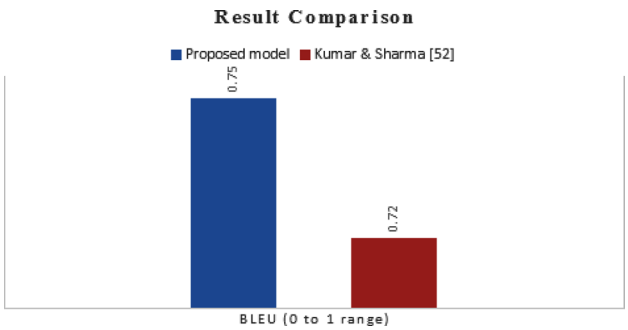

Fig. 10 Result comparison on Punjabi Monolingual Text Corpus ILCI-II dataset of proposed method with Kumar and Sharma [50]

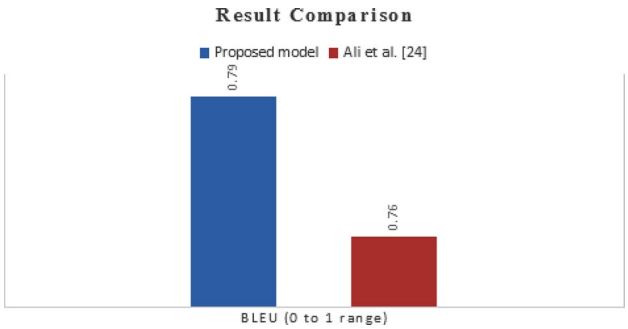

Fig. 11 Result comparison of proposed method with Ali et al. [24] based on same dataset

5 Discussion

We have evaluated the proposed machine translation system with only one other machine translation system because, for the Bengali language a few MT system has been proposed so far. Moreover, in [25] authors used EMILLE dataset to evaluate their proposed system. Basically, result comparison on the same dataset with two different systems is easier to find the best one.

After comparing our work with [25], we have found that our work performs far better. The high in BLEU score means better results [46]. In Fig. 11 we can see that our proposed system got 0.74 BLEU score where [25] got 0.057. The lower in the TER score is, the better result [48]. Our system got 0.17 TER score, whereas [25] got 0.83. NIST score is 0 to 10 range where it contains three stages between this range-Low (0.0–3.9), Medium (4.0–6.9) and High (7.0–10.0) [47]. We can see that proposed system got Medium range score, whereas [25] got Low range score. Therefore, the efficiency of our proposed system much better than other system.

We have also trained our system with Punjabi Monolingual Text Corpus ILCI-II [49]. To find out the efficiency of our system, we have compared the result with a Punjabi Deconverter [50]. After comparing our work with [50], we have found that our work performs far better. In Fig. 10, we can see that our proposed system got 0.75 BLEU score where Punjabi Deconverter [49] got 0.72. Moreover, we have trained our system with the same datasets used in [24] and

compare the results with [24]. We have found that our proposed system achieves 0.79 BLEU score where [24] got 0.76 BLEU score. Figure 11 shows that our system outperforms [24].

Our proposed system does not require to be compelled to knowing any rules concerning human language. The system figures out those rules itself. Language experts are not required to tune each step of the translation pipeline. The system will do it automatically. One thing needs to remember is that this system required real word data for training purpose. Since rare words come often in a sentence, sometimes our system does not translate rear word correctly.

6 Conclusion

In this paper, we have presented a deep learning-based sequence-to-sequence statistical machine translation system which can translate any SOV language into SVO language. This paper also provides an overall description of other MT systems. Two popular deep learning algorithms named-recurrent neural network (RNN) and encodings are used in this system. Two case study is shown in this paper. Bengali (SOV) language and English (SVO) language as case study-1 where Punjabi (SOV) and English (SVO) language as case study-2. EMILLE corpus dataset and Prothom-Alo corpus dataset are used for case study-1 in this system. A very less research had been reported on Bengali to English language MT, but none of them solve the semantic ambiguity problem of Bengali sentence efficiently. Since a few numbers of MT system available for Bengali to English language, the efficiency of this system shown by comparing some translated sentence with other state-of-the-art translation system. On average, we have obtained satisfied scores for combined datasets (BLEU score: 0.742, NIST score: 4.11, TER score: 0.18). Punjabi Monolingual Text Corpus ILCI-II dataset is used for case study-2 which achieved a BLEU score of 0.75. Our immediate plan is to improve the accuracy of this system and we will also extend this work by adding some other languages.

References

1. Wu Y, Schuster M, Chen Z, Le QV, Norouzi M, Macherey W et al (2016) Google's neural machine translation system: bridging the gap between human and machine translation. arXiv:1609.08144
2. Russell SJ, Norvig P (2016) Artificial intelligence: a modern approach. Pearson Education Limited, Malaysia
3. Karaa WBA, Ashour AS, Sassi DB, Roy P, Kausar N, Dey N (2016) Medline text mining: an enhancement genetic algorithm based approach for document clustering. In: Applications of intelligent optimization in biology and medicine. Springer, Cham, pp 267–287
4. Santosh KC, Nattee C (2009) A comprehensive survey on on-line handwriting recognition technology and its real application to the Nepalese natural handwriting. Kathmandu University J Sci Eng Technol 5(I):31–55
5. https://en.wikipedia.org/wiki/Google_Translate. Accessed 27 Jul 2019
6. Eom YH, Aragón P, Laniado D, Kaltenbrunner A, Vigna S, Shepelyansky DL (2015) Interactions of cultures and top people of Wikipedia from ranking of 24 language editions. PLoS ONE 10(3):e0114825
7. Mridha MF, Saha AK, Das JK (2014) New approach of solving semantic ambiguity problem of Bangla root words using universal networking language (UNL). In: 2014 International Conference on Informatics, Electronics & Vision (ICIEV). IEEE, pp 1–6
8. Tripathi S, Sarkhel JK (2010) Approaches to machine translation
9. Vickrey D, Biewald L, Teyssier M, Koller D (2005) Word-sense disambiguation for machine translation. In: Proceedings of human language technology conference and conference on empirical methods in natural language processing, pp 771–778
10. Forcada ML, Ginestí-Rosell M, Nordfalk J, O'Regan J, Ortiz-Rojas S, Pérez-Ortiz JA et al (2011) Apertium: a free/open-source platform for rule-based machine translation. Mach Transl 25(2):127–144
11. Karaa WBA, Dey N (2017) Mining multimedia documents. Chapman and Hall/CRC, Boca Raton
12. Santosh KC, Nattee C (2006) Stroke number and order free handwriting recognition for Nepali. In: Pacific Rim International Conference on Artificial Intelligence. Springer, Berlin, Heidelberg, pp 990–994
13. Koehn P (2009) Statistical machine translation. Cambridge University Press, Cambridge
14. Koehn P (2005) Europarl: a parallel corpus for statistical machine translation. In: MT summit. vol 5, pp 79–86
15. Singh A, Dey N, Ashour AS (2017) Scope of automation in semantics-driven multimedia information retrieval from web. In: Web semantics for textual and visual information retrieval. IGI Global, pp 1–16
16. Santosh KC, Nattee C (2006) Structural approach on writer independent nepalese natural handwriting recognition. In: 2006 IEEE conference on cybernetics and intelligent systems. IEEE, pp. 1–6
17. Aiken M, Balan S (2011) An analysis of Google translate accuracy. Transl J 16(2):1–3
18. Maji P, Chatterjee S, Chakraborty S, Kausar N, Samanta S, Dey N (2015) Effect of Euler number as a feature in gender recognition system from offline handwritten signature using neural networks. In: 2015 2nd International conference on computing for sustainable global development (INDIACom). IEEE, pp 1869–1873
19. Farrús M, Costa-Jussa MR, Mariño JB, Poch M, Hernández A, Henríquez C, Fonollosa JA (2011) Overcoming statistical machine translation limitations: error analysis and proposed solutions for the Catalan-Spanish language pair. Lang Resour Eval 45(2):181–208
20. Pardeshi R, Chaudhuri BB, Hangarge M, Santosh KC (2014) Automatic handwritten Indian scripts identification. In: 2014 14th international conference on frontiers in handwriting recognition. IEEE, pp 375–380
21. Chaki J, Dey N, Shi F, Sherratt RS (2019) Pattern mining approaches used in sensor-based biometric recognition: a review. IEEE Sens J 19(10):3569–3580
22. Mahata SK, Das D, Bandyopadhyay S (2019) Mtil 2017: Machine translation using recurrent neural network on statistical machine translation. J Intell Syst 28(3):447–453
23. Mukta AP, Mamun AA, Basak C, Nahar S, Arif MFH (2019) A phrase-based machine translation from English to Bangla using rule-based approach. In: 2019 International conference on

electrical, computer and communication engineering (ECCE). IEEE, pp 1–5
24. Ali M, Yousuf N, Rahman M, Sorwar G (2019) Bangla DeConverter for extraction of BanglaText from Universal Networking Language. Information 10(10):324
25. Islam MZ, Tiedemann J, Eisele A (2010) English to Bangla phrase-based machine translation. In: proceedings of the 14th annual conference of the European association for machine translation
26. Ashrafi SS, Kabir MH, Anwar MM, Noman AKM (2013) English to Bangla machine translation system using Context-Free Grammars. Int J Comput Sci Issues 10(3):144
27. Ali MNY, Al-Mamun SA, Das JK, Nurannabi AM (2008) Morphological analysis of Bangla words for universal networking language. In: 2008 Third International Conference on Digital Information Management. IEEE, pp 532–537
28. Ali M, Ali MM (2002) Development of machine translation Dictionaries for Bangla language. In: 5th ICCIT, pp 272–276
29. Saha GK (2005) The E2B machine translation: a new approach to HLT. Ubiquity 2005(August):1–1
30. Uddin MG, Ashraf H, Kamal AHM, Ali MM (2004) New parameters for Bangla to English statistical machine translation. In: International Conference on Electrical and Computer Engineering. ICECE, pp 545–548
31. Luong MT, Manning CD (2016) Achieving open vocabulary neural machine translation with hybrid word-character models. arXiv:1604.00788
32. Simard M, Ueing N, Isabelle P, Kuhn R (2007) Rule-based translation with statistical phrase-based post-editing. In: Proceedings of the second workshop on statistical machine translation. Association for Computational Linguistics, pp 203–206
33. Subramanian CM, Cherukuri AK, Chelliah C (2018) Role based access control design using three-way formal concept analysis. Int J Mach Learn Cybern 9(11):1807–1837
34. Dey N, Ashour AS, Nguyen GN (2020) Recent advancement in multimedia content using deep learning
35. Pinker S (1991) Rules of language. Science 253(5019):530–535
36. Habash N (2007) Syntactic preprocessing for statistical machine translation. In: MT Summit XI. pp 215–222
37. Mikolov T, Kombrink S, Burget L, Černocký J, Khudanpur S (2011) Extensions of recurrent neural network language model. In: 2011 IEEE International conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 5528–5531
38. Gao M, Shi G, Li S (2018) Online prediction of ship behavior with automatic identification system sensor data using bidirectional long short-term memory recurrent neural network. Sensors 18(12):4211
39. Auli M, Galley M, Quirk C, Zweig G (2013) Joint language and translation modeling with recurrent neural networks
40. Ogata T, Murase M, Tani J, Komatani K, Okuno HG (2007) Two-way translation of compound sentences and arm motions by recurrent neural networks. In: 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, pp 1858–1863
41. Mikolov T, Karaiát M, Burget L, Černocký J, Khudanpur S (2010) Recurrent neural network based language model. In: Eleventh annual conference of the international speech communication association
42. Wang J, Liu F, Qin S (2019) Global exponential stability of uncertain memristor-based recurrent neural networks with mixed time delays. Int J Mach Learn Cybern 10(4):743–755
43. Zhang C, Ma Y (eds) (2012) Ensemble machine learning: methods and applications. Springer, New York
44. https://www.lancaster.ac.uk/fass/projects/corpus/emille/. Accessed 21 Jun 2019
45. Majumder KM, Arafat Y (2006) Analysis of and observations from a Bangla News Corpus
46. Papineni K, Roukos S, Ward T, Zhu WJ (2002) BLEU: a method for automatic evaluation of machine translation. In: Proceedings of the 40th annual meeting on association for computational linguistics. Association for Computational Linguistics, pp 311–318
47. Doddington G (2002) Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In: Proceedings of the second international conference on Human Language Technology Research. Morgan Kaufmann Publishers Inc, pp 138–145
48. Snover M, Dorr B, Schwartz R, Micciulla L, Makhoul J (2006) A study of translation edit rate with targeted human annotation. In: Proceedings of association for machine translation in the Americas, vol 200, no 6
49. https://tdil-dc.in/index.php?option=com_download&task=showresourceDetails&toolid=1890 (Punjabi Monolingual Text Corpus ILCI-II). Accessed 16 Apr 2020
50. Kumar P, Sharma RK (2013) Punjabi Deconverter for generating Punjabi from universal networking language. J Zhejiang Univ Sci C 14(3):179–196

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.