Journal of
Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7691383

A NEURAL NETWORK BASED MACHINE TRANSLATIONMODEL FOR ENGLISH TO AHIRANI LANGUAGE

Dr. Uday C. Patkar
Bharati Vidyapeeth's college of Engineering, lavale, Pune, (MS), India.
uday.patil@bharatividyapeeth.edu

Mrs. Pallavi Tekade
JSPM's rajarshi shahu College of Engineering, Pune, (MS), India.
tekade.pallavi@gmail.com

Mrs. N. S. Patil
Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, 411043, India.
nspatil@bvucoep.edu.in

Dr. A. M. Pawar
Bharati Vidayapeeth's College of Engineering for Women, Pune (MS), India.
avinash.m.pawar@bharatividyapeeth.edu

Dr. Uday Patil
Bharati Vidyapeeth's college of Engineering, lavale, Pune, (MS), India.
uday.patil@bharatividyapeeth.edu

Mrs. Gauri Rao
Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, 411043, India.
grrao@bvucoep.edu.in

Mrs. S. S. Patil
Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, 411043, India.
sspatil@bvucoep.edu.in

Dr. K. D. Bamane
D Y Patil College of Engineering, Akurdi Pune (MS), India.
kdbamane@dypcoeakurdi.ac.in

ABSTRACT: Machine Translation is the process of converting one natural language into another through the use of computer-assisted translation. The fundamental goal is to bridge the linguistic gap that exists between two distinct languages speaking people, communities, or countries. India is a multilingual country, with different territorial languages spoken in different parts of the country. However, not all Indians are polyglots. There are 18 constitutional languages and ten widely used scripts in the world. The majority of the Indians, particularly the peasants in distant areas, do not understand, read, or write English, making the use of an efficient language translator necessary and desirable. Machine translation systems that translate content from one language to another will contribute to the advancement of the enlightened civilization of Indians, regardless of their native tongue. Despite the fact that Marathi is the most widely spoken language in the state of Maharashtra, Ahirani is also extensively spoken. Taking advantage of the fact that English is a universal language and that Ahirani is the language spoken by the vast majority of Khandeshi people, we present a proposal for an English to Ahirani machine translation system using Recurrent neural networks (RNN).

KEY WORDS*: Indian Languages, Machine Translation, Natural Language Processing, Lexical Analysis, Computational Linguistics, Rule Based Translation*

INTRODUCTION

Since 1940, machine translation has always been in the early stages of research and development. Machine translation has been on the rise since the 1940s and will continue to do so. Machine translation systems can be used to translate text or speech from one natural language to another. To translate a document or text from another commonly known language into your native language, you must use machine translation. You can overcome language barriers. NLP is a branch of computer science that seeks to fill this gap. Neural machine translation is conceptually simple, requiring only a rudimentary understanding of the domain. Large neural networks have been trained to generate very long word strings. Unlike regular machine translation systems, this model explicitly stores large phrase dictionaries and language models, making it more efficient. Georgetown University and IBM worked together and in 1965 he made the first deployment of the MT system. An important factor contributing to the importance of machine translation in multilingual societies is the social and political importance of translation in multilingual societies. Additionally, the concept of attention mechanisms is used.

Machine translation (MT) has difficulty dealing with the issue of translation discrepancies. Differences in translation can be defined by differences in grammar from one language to the next in its grammar. According to numerous studies, divergence occurs when sentences in one language (L1) are translated into sentences in another language (L2) in significantly different forms. Many different perspectives were considered, and various techniques were presented to address the problem. A system for identifying and resolving translation discrepancies is essential for machine translation systems to ensure that translations are accurate and reliable. Translators encounter translation discrepancies at various levels, and the degree of complexity of each translation discrepancy has a corresponding impact on the overall quality of the translation. Some types of translational differences are universal in the sense that they exist in all languages, while other types of translational differences are specific to Particular pairs of translated languages, as found in the literature. Therefore, to fully understand translation discrepancies, it is necessary to examine them from both an inter-language and a language-specific perspective. This research examines the translation language pair of English and Ahirani from the perspective of discovering language-specific differences. Due to the significant differences between English and the Ahirani dialect, this translation language pair provides an excellent resource for exploring translation differences in machine translation.

These languages are also very syntactically, structurally, and sociocultural different from each other, and these differences need to be thoroughly investigated. The purpose of this work is to analyze some features of English and Ahirani grammar that may be related to translation discrepancies between Ahirani and English machine translations. We discuss the issue of syntactic and structural divergence in English-Ahirani machine translation, analyzing the same translation pair with back-translation to determine the nature of the divergence in each situation separately.

This paper examines language pairs in terms of Dorr's taxonomy of the problem of translation differences between English and Ahirani, which is the starting point for our investigation. Section 2 reviews the literature on the classification of translation discrepancies. This section describes various lexical-semantic and syntactic differences. Section 3 details how to implement a recurrent neural network for machine translation. Section 4 of the study is devoted to discussion of the results obtained, after which the work is completed.

REVIEW OF LITERATURE

Natural language processing is often performed when systems are deployed to manipulate or manipulate text or speech. As part of current research [1-5], a text-to-speech system was developed that synthesizes text by incorporating natural language processing (NLP) before proceeding to digital signal processing (DSP) to generate synthesized words. I was able to develop a simple but useful text-to-speech application that converts written text to speech and saves the result as an mp3 file.

Bridging the gap between linguistic scripts and languages for the deaf will have a positive impact on the promotion and growth of a closely domesticated population of civilians. Most of the literature on English text-to-speech mentions how difficult this is, but with limited success. If you're ever curious about text-to-speech conversion, researching the local language should give you results. In their study [6], they claim that this technique can be used to convert optional text to speech (2). In this case, instead of constructing a concatenated set of values from the actual language data, we convert directly from the registered language parameters. This algorithm also participates in a global comparison compared to another synthesizer commonly used for language learning purposes and responsible for most communication problems. A person who is not fluent in English can simply look at a grid of symbols enlarged on a computer screen to see how the content would be rendered in the original language if they were not fluent in English. This study presents data from 53 different languages and provides an opportunity to apply the results to a variety of other languages, thus providing a new research area. users can retrieve information in images based on their preferred language settings [7]. When we talk about Graphing, we are talking about the process of extracting text from images and translating the recovered text into multiple languages for viewing Google Speech API [Application Program Interface] and Tesseract OCR [Optical Character Recognition] were used in conjunction with each other. Therefore, travelers can spend their time relaxing by listening to audio in their preferred language, English. It is also accessible to people with visual impairments. This device is especially useful for those who want to read text in its original language.

Building on rule-based text-to-speech synthesis that has proven effective, this work combines natural phonetics and lexical formant approaches [8] to transform the final product into word and concatenated to the concatenation of phrases. Simulations show that in Marathi, the current system manages words, phrases, sentences and paragraphs better than the proposed approaches for better speech recognition. With 91% accuracy, the entire system can predict how long the process will take with 91% accuracy. Further efforts are needed to improve the accuracy of detecting stressed syllables and vocal strength.

To generate the output, the article breaks the text into its constituent parts and assembles them into speech. Voices are stored in a database and retrieved as needed. As a result, the considered work [9] used concatenated-input-based speech synthesis on the MATLAB 2010 platform, which is considered to have good and clear sound. It is widely used and gaining popularity as an aid for people with dyslexia and dyslexia, especially those with visual impairments. The software can be used for many other purposes, but its main purpose is to improve foreign language skills and pronunciation. These include reducing eye strain when reading (whether in paper or digital format), reducing document production costs through digital text printing, English translation, digital writing/editing, and enhancing listening comprehension.

Abstract: Translates simple English sentences into assertive Marathi sentences using rule-based techniques described in research [10]. Again, the algorithm accepts a simple sentence as input and once again outputs a phrase describing the output element or situation (lexeme). The tokenizer identifies each newly formed English letter and each letter found in the English dictionary according to a tokenizer (lexicon). If a lexical item has tokens associated with it, you can optionally access the morphological features associated with those tokens. Instead of extracting anything before finding a valid root, morphemes are classified based on features present in the word itself, saving time. Groups established in specific places provide information about how written sentences are grammatically correct. When using bottom-up parsing, syntax is verified by looking at the depth of parsing. In addition, this program's search dictionary is enriched with syntactic tokens (Marathi). This allows you to accurately identify the target language words in a sentence. If the required Marathi word cannot be found word by word and cannot be connected to the corresponding English word, then create a Marathi sentence with someone who can make an English translation of the text using the remaining words in the text. Topics covered in this article include the principles that form English sentences, and various ways of rearranging vocabulary and proofreading the intricacies of the language.

[11] The texts reviewed here show how the authors approved the Marathi language used in this study. We found two unique corpora in the original text. Each represents the author of the most frequently occurring term and an interpretation of the text based on patterns identified by us in the original text (his SMORDT sequential minimum optimality using statistical similarity models and rules transformation-based decision tree approach). To demonstrate that feature extraction is important for any model, we tested and validated the feature extraction method used in this experiment to demonstrate that it is. The proposed method has been tested over a long period of time, considering three main criteria: recall, accuracy and precision.

With the urbanization of civilization, the demand for clean air is increasing. Since it was first recorded, the technology has expanded the data volume, doubling his data volume every two years since it was first released. As a result, computers are expected to acquire, analyze, interpret, and apply data with minimal human input, unlike previous methods. Many of these text segments in source code are difficult to interpret. However, some were included in "code mixing". Or "machine code-like" languages that are particularly hard to parse. This has already been done in this area and this work [12] is an active contributor and applying comparability to classify and evaluate these documents to classify and evaluate these documents into classes of learning systems (Bag of Words and Human Priors (NB, SVM)), Marathi and Hindi text translation documents (data We train and test a classification that examines the subjectivity of (using support vector machines and naive Bayes) to decide which classification method to use. On the other hand, the results of basic machine learning algorithms are often better than or often comparable to those of analytical approaches.

When writing sentences in Marathi, he can start each phrase with two words to clarify word breaks [13]. This is an option for those who speak the language. Marathi does not contain the beginning-of-sentence markers found in English, so it is difficult to distinguish between the beginning and the end of a sentence. To correctly determine the end of an order, a reliable system considers a large amount of information.

In particular translation apps [14-15] (tools to use language to complement English) translate Malayalam texts into their native languages and are used by language minorities who wish to engage with the culture and language of Kerala is intended for Machine translation and text-to-speech accuracy are improved by using various combined approaches that include both machine translation and text-to-speech. If Malayalam is the source language, the translation is completed using a grammar-based translation method that expands the English part, morphological analysis, and a Malayalam-Tamil-English dictionary. Individuals come from each of the translated Malayalam syllables and words and are distinct from each other (one in every). The syllable database contains large number of syllable recordings that have been created and recorded. The concatenation and synthesis of Malayalam words leads to the creation of synthesized words. In the first two iterations of Malayalam machine translation, the accuracy is 70%, increasing to 73% in three stages of unfolding. To ensure the accuracy of the speech output, the device should take advantage of naturalness and intelligibility characteristics .87% of the sentences are well formed during the grading process.

RNN FOR MACHINE TRANSLATION

RNN is a type of neural network with loops that can persist information across network steps. A loop tells the neural network to look at what happened in all previous words before deciding what the current word really means. RNNs can be thought of as copying and pasting the same network over and over again. Each new copy and paste adds slightly more information than the previous one. RNN applications differ greatly from traditional NN applications in that they do not have an output and input set as a concrete value; instead, they take sequences as input or output.

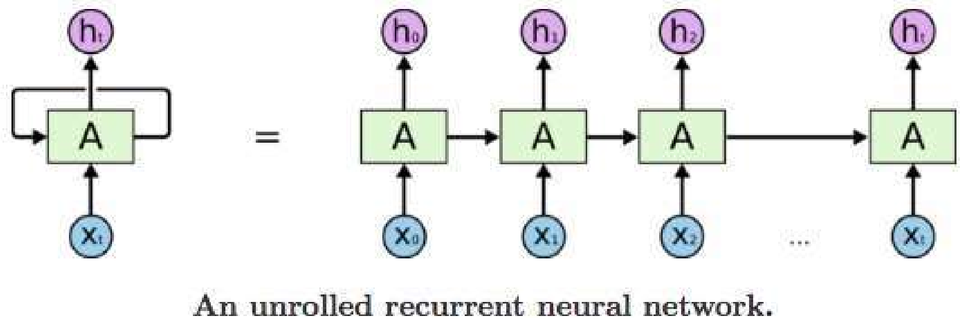

Figure1. An unrolled recurrent neural network

This issue can arise in any network that employs gradient-based optimization techniques. When back-propagation (calculating the gradients of loss with respect to the weights) is calculated, the gradients become extremely small as the back propagation algorithm moves through the network. As a result, the earlier layers learn much slower than the later ones. This reduces the effectiveness of RNNs because they are frequently unable to fully account for long sequences. An RNN becomes less and less effective as the gap between needed information grows larger.

An RNN also has different models that it can follow, this study follows a sequence-to-sequence model. Here the RNN receives an input sequence and outputs a sequence.

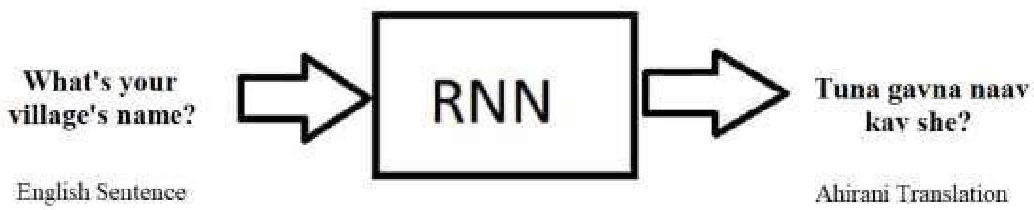

Figure 2. Machine Translation: an RNN reads a sentence in one language and then outputs it in another

The use of activation functions that do not cause vanishing gradients, such as the RNN reads a sentence in one language and outputs it in another problem. Long-Short Term Memory Networks are a better solution.

Setup of the RNN

Depending on your application, you may need to configure your RNN to handle inputs and outputs differently. As you can see below, this project uses many-to-many processes. The input is a sequence of English words and the output is a sequence of Japanese words (fourth from the left in the figure below).

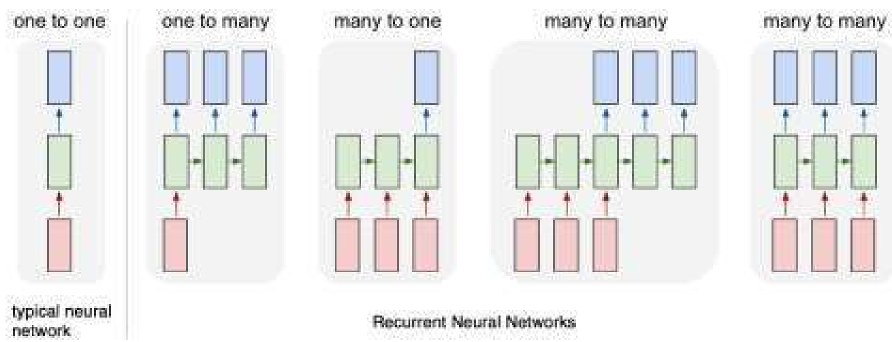

Figure 3. Recurrent Neural Network

The following section outlines various preprocessing and modeling steps. Here are the high-level steps:

- Preprocessing includes reading and inspecting data, cleaning, tokenizing, and padding.
- Modeling is the process of creating, training, and testing models.
- To be able to make predictions, we need to generate a specific translation from English to Ahirani and compare it with the ground truth translation.
- Continue to iterate on the model and experiment with different architectural configurations.

This study uses Keras for the frontend and Tensor Flow for the backend to achieve its goals. Using Keras on top of Tensor Flow is my preferred method as it has a simpler syntax and makes creating model layers more intuitive. The downside of Keras is that you lose the ability to make fine adjustments. This is a big advantage. However, this does not affect the model developed for this study.

PREPROCESSING

Load & Examine Data

Below is an example of data. The input is in English, and the output is a translation of the English text. The word count of the dataset indicates that the vocabulary of the dataset is relatively limited. This was designed specifically for this project.This allows the model to be trained in a reasonable amount of time.

Cleaning

No additional cleaning required at this time. Since the data has already been converted to lowercase and split, there are spaces between every word and punctuation as shown above.

Tokenization

Next, we need to tokenize the data. That is, it converts text to numbers. This allows the neural network to perform operations on the data it receives as input. Each word and punctuation mark is assigned a unique ID number for this project. Running to kenizer generates a word index, which is used to convert each sentence into vector form.

Padding

When entering a sequence of word IDs into a model, you must ensure that each sequence is the same size. Padding is applied to sequences shorter than the maximum length (that is, shorter than the longest sentence) to achieve this result.

Modeling

To understand the architecture of a RNN, we first need to look at it at a high level. As shown in the diagram above, there are several components of the model that you should be aware of:

Inputs-The input sequence enters the model word by word, with each time step represented by a different word. Each word is encoded as a single integer or as a one-hot encoded vector corresponding to the word in the lexical database of English records.

Layers for Embedding- Each word is converted into a vector by embedding. Vocabulary complexity affects the size of the vector representing the vocabulary. Iterative shift is a type of shift that occurs repeatedly (encoder). This means that the context from the previous time step word vector is applied to the current word vector. Layers that are dense (Decoder) Typical fully connected layers that are used to convert the encoded input into the correct translation sequence are depicted in the diagram.

Outputs- The output is returned as a series of integers or as a one-hot encoded vector that can be mapped to the French record vocabulary after processing.

Embedding

Embeddings should be used to achieve very simple semantic and syntactic word relationships. In particular each word is projected onto an n-dimensional space, giving us this result. Similar-sounding words occupy this space in a similarity search. The closer two words are, the more similar they are in meaning. And often the equation between both proverbs represents an interconnected network involving gender, prepositions, and even strategic ties.

Embedding large datasets from scratch requires a large amount of data and computation, making it infeasible for most people. To avoid having to train the embeddings from scratch, we typically use pre-trained embedding packages such as Glove or word2vec. Used in this way, embeddings can be viewed as a form of transfer learning. However, due to the limited vocabulary and few syntactical variations in this project's dataset, we use the Keras machine learning framework to train the embeddings themselves.

Encoder & Decoder

Our sequence-to-sequence model connects two recurrent networks: an encoder and a decoder, which are connected by a recurrent network. The encoder summarizes the input and stores it in a context variable, also known as a state variable. Afterwards, the decoding context is obtained, and the output sequence is generated.

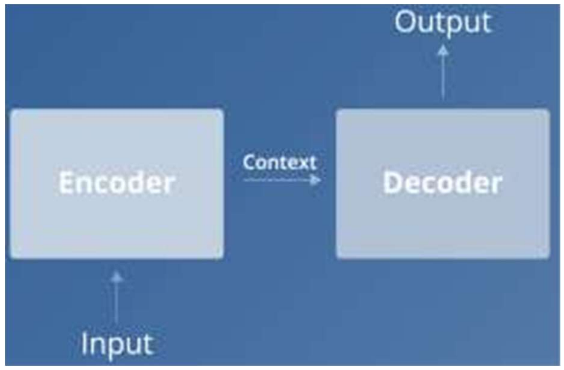

Figure 4. Encoder & Decoder

Due to the fact that both the encoder and decoder are recurrent, they both have loops that process each part of the sequence at a different time step. Unrolling the network to see what's going on at each time step is the most effective way to conceptualize this phenomenon.

To encode the entire input sequence, as shown in the following example, it takes four time steps. During each time step, the encoder "reads" the input word and applies a transformation to the hidden state of the input word. Then it moves on to the next time step with the hidden state in tow. Please remember that the hidden state is a representation of the relevant context that is passing through the network. The larger the hidden state, the greater the learning capacity of the model; however, the larger the hidden state also increases the computation requirements of the model.

The design of our suggested system is the same as that of the Moses SMT system, which allows us to compare the scores of the phrases in the phrase table obtained by the Moses SMT system. It is an RNN that progressively reads each symbol of an input sequence x, which is called the encoder. When the RNN reads a symbol, the hidden state of the RNN changes in accordance with the equation (1). Following the reading of the end of the sequence, the hidden state of the RNN is a summary c of the entire input sequence, as shown in the following figure. When the hidden state h is known, the proposed model's decoder is yet another RNN that has been trained to construct the output sequence by anticipating the next symbol yt when that state is known (t). The hidden state of the decoder at time t is calculated using the formula.

$$h_{(t)}=f(h_{(t-1)}, y_{t-1}, c)$$

And similarly, the conditional distribution of the next symbol is

$$P(y_t|y_{t-1}, y_{t-2}, \dots, y_1, c) = g(h_{(t)}, y_{t-1}, c)$$

For given activation functions f and g (the latter must produce valid probabilities, e.g. with a SoftMax).

$$\max \frac{1}{N} \sum_{n-1}^{N} \log p\Theta(y_n|x_n)$$

The two components of the proposed RNN encoder-decoder are jointly trained to maximize the conditional log-likelihood.

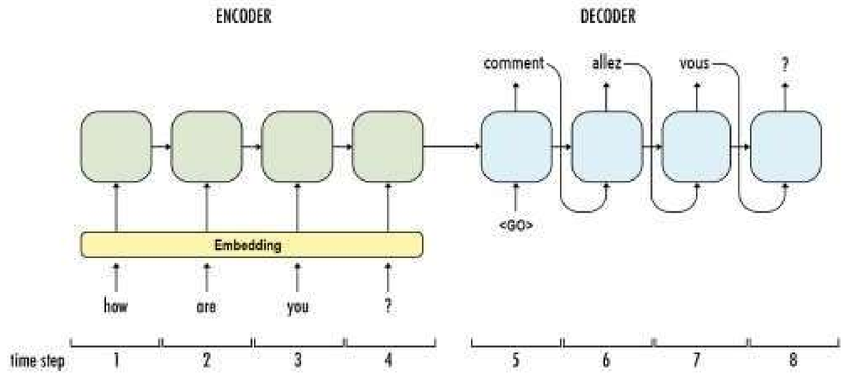

Figure 5. RNN encoder-decoder

Because the encoder and decoder are both recurrent, they both have loops that process every other part of the series at different time steps. To visualize this, we should unroll the network and examine what is happening at each time step.

Hidden Layer with Gated Recurrent Unit (GRU)

If you can choose which information from the hidden state is allowed to flow over the network, it's not desirable Maybe? Some information is more relevant than others, and some information should be removed entirely. The GRU, or Gated Repetition Unit, provides functionality similar to this. The Update Gate and Reset Gate are the only gates in the GRU and both are used together. A detailed explanation can be found in this essay by Simeon Kostadinov. In summary, the update gate (z) helps the model determine how much information is passed into the future from the previous time step. The reset gate (r), on the other hand, determines how much of the previous information is erased.

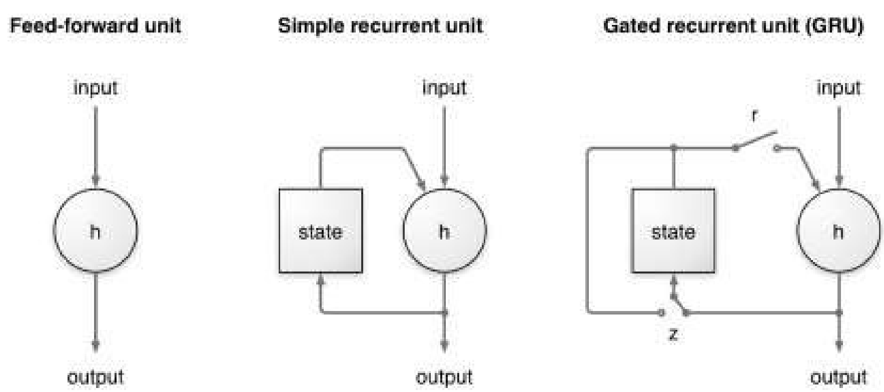

Figure 6. Gated Repetition Unit

RESULT ANALYSIS

Researchers Must Evaluate Their Research to Determine Success or Failure

This has its limitations, as most authors currently only assess their findings for precision and recall, or time lag and number of queries. Precision and recall can be just mathematical calculations, which can lead to unstructured research assessments. Therefore, there is a need to develop new parameters for evaluating research outcomes.

Using the Ahirani training corpus from the Ahirani dictionary (https://www.ahirani.in/en/books/ahirani-dictionary), consisting of parallel running source-target sentence pairs using English as the source language and the system was trained to work.

Commonly used in a variety of situations. Due to the limitations imposed on the corpus format by system training, data must be preprocessed prior to system training for a successful training session. Validation data, a subset of the training corpus containing 80% of the instances, is used to check the training convergence. The translation effectiveness of the trained and validated model was evaluated using a test corpus of approximately 1400 English sentences. Table 1 below shows the optimization parameters chosen for the implementation.

Table 1: RNN Algorithm Tunning Parameters

Number of Epochs	5
RNN Size	128
Batch Size	128
Epoch	2
Learning Rate	0.01
Encoding Embedding Size	128
Decoding Embedding Size	128
Keep Probability	0.8

Comparison of training and validation accuracy against number of epochs. Table 2 shows the Training and validation accuracy of RNN machine translation model of epochs. This analysis is also shown in figure 3.

Table 2: Number of Epoch Vs Training and Validation Accuracy

Epoch	Train Accuracy	Validation Accuracy
1	67.31	67.05
2	76.13	74.36
3	81.29	80.65
4	84.45	83.38
5	87.50	85.65

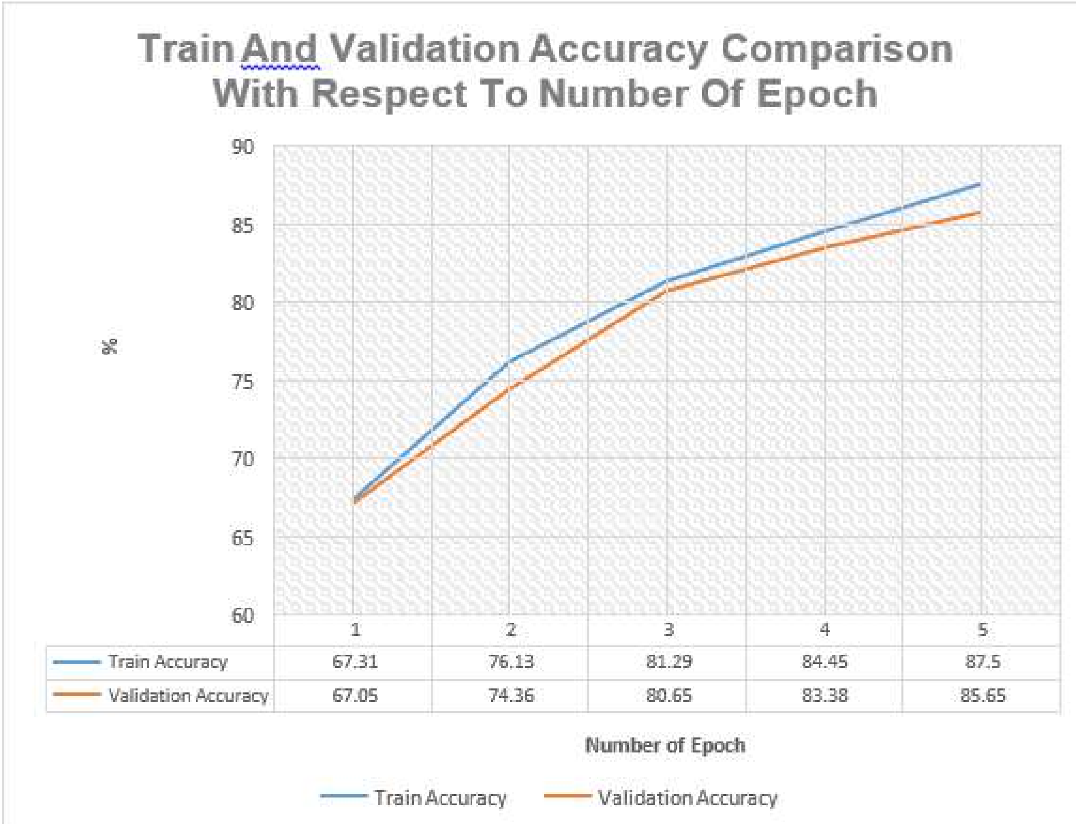

Figure 7. Training and Validation Accuracy for RNN Machine Translation Model

The losses occurred during the execution is shown in table 3 and represented in figure 4.

Table 3. Epochs Vs Training Losses

Epoch	**Training Loss**
1	0.4356
2	0.3316
3	0.2497
4	0.2070
5	0.1514

Figure 8. Training Losses for number of Epochs

CONCLUSION REMARKS

Network (RNN) system for predicting translations. Key benefits of using NMT in an Indic context include the ability to generate fluent translations, enhanced contextual analysis capabilities, and better performance than SMT systems. These are some of the main advantages that motivated us to study the use of NMT in the Indic context. In particular, the English-Ahirani pairing algorithm has been trained and tested by our team. Different experimental settings were also developed to investigate the relationship between translation efficiency of the English-Ahirani MT system and variations in the number of epochs, amount of training data, and test phrase length. After examining the predicted translations in detail, we came to the conclusion that the MT system produces fluent translations and performs better as the amount of training data and the length of the test phrases increase. Additionally, the translation performance against the epoch diagram helps determine whether the system training is converging.

The basic working concept of NMT is highly dependent on the size of the training corpus, so the number of training corpus instances should be increased for the training corpus to be effective. The attention mechanisms and scoring functions used to compute attention for each source state have the greatest impact on translation success. By changing the merit function, we can increase the interaction between the source state vector and the current hidden state of the decoder. Moreover, a deeper understanding of the nature of the target language components helps improve the comprehensibility, relevance and fluency of translations.

Apart from that, skillful and thorough selection of settings for system parameters such as number of epochs, hidden layers and GPU can improve the overall quality of the translated text.

REFERENCES

1. Pramod Salunkhe, Mrunal Bewoor, Suhas Pati, Shashank Joshi, Aniket kadam, "Summarization and Hybrid Machine Translation System for English to Marathi: A Research Effort in Information Retriveal System (H-Machine Translation" ,Discovery The International journal (ISI thomas retuers indexed).2015.
2. Pramod Salunkhe , Mrunal Bewoor , Dr.Suhas Patil "A Research Work on English to Marathi Hybrid Translation System", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (3) 2015, 2557-2560.s
3. S. B. Kulkarni, P. D. Deshmukh "Linguistic Divergence Patterns in English to Marathi Translation International Journal of Computer Applications (09 75 8887) Volume 87 No.4, February 2014.
4. J. Ramanand, Akshay Ukey, Brahm Kiran Singh, Pushapak Bhattacharyya "Mapping and Structural Analysis of MultilingualWordnets"
5. Bulletin of the IEEE Computer Society Technical Committee on Data Engineering.
6. Sreelekha.S, Raj Dabre, Pushpak Bhattacharyya "Comparison of SMT and RBMT, The Requirement of Hybridization for Marathi Hindi MT" [online pdf].
7. Amruta Godase, Sharvari Govilkar"Survey Of Machine Translation Development for Indian Regional Languages" International Journal of Modern Trends in Engineering and Research IJMTER- 2014.
8. M.L.Dhore & S.X.Dixit (2011) "English to Devnagari Translation for UI Labels of Commercial web based Interactive Applications, International Journal of Computer Applications (0975 – 8887) Volume 35 –No.10, December 2011.
9. Devika Pishartoy, Priya, Sayli Wandkar (2012) "Extending capabilities of English to Marathi machine Translator., IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 3, May 2012.
10. Priyanka-choudhary "A approach for interlingua example based EM Translation, IJCSIT, 2015.
11. Abhay A, Anita G, Paurnima T, Prajakta G (2013), "Rule based English to Marathi translation of Assertive sentence International Journal of Modern Trends in Engineering and Research 2014.
12. Aniket Kadam, A.D. Dept. Inf. Tech., BVDUCOEP, Pune, India ; Joshi, S.D.; Medhane, S.P, "Question Answering Search engine short review and road-map to future QA Search Engine", Electrical, Electronics, Signals, Communication and Optimization (EESCO), 2015 , 10.1109/EESCO.2015.7253949.
13. Kadam Aniket, Prof.S.D.Joshi, prof.S.P.Medhane, "QAS" International Journal of Application or Innovation in Engineering & Management IJAIEM , Volume 3, Issue 5, May 2014 May 2015.
14. Kadam Aniket, Prof.S.D.Joshi, Prof.S.P.Medhane, "QAS" International Journal of

Application or Innovation in Engineering & Management IJAIEM , Volume 3, Issue 5, May 2014 May 2015.

15. Kadam Aniket, Prof.S.D.Joshi, Prof.S.P.Medhane “Search Engines to QAS:Explorative Analysis”, International Journal of Application or Innovation in Engineering & Management IJAIEM , Volume 3, Issue 5, May 2014 May 2015 IJAIEM May 2015 .
16. Mishra, Sudha. "Heterogeneous Groups in Competition: Innovative Pedagogy for Teaching English as Second Language." International Journal of Educational Science and Research 4.2014 (2014): 1-6.
17. Jha, Sanjay Kumar. "Exploring the Causes of Sanskrit’s Decline (A Pilot Study)." International Journal of Educational Science and Research (IJESR) 8 (2018): 165-172.
18. Kumar, Manoj, and Sanjay Kumar. "SPEECH AS TEXT FOR TEACHING: AN ANALYSIS OF LINGUISTIC AND STYLISTIC MANEUVERINGS OF STEVE JOBS‟ FAMOUS SPEECH STAY HUNGRY STAY FOOLISH AND EXERCISES FOR ENHANCING STUDENTS‟ LANGUAGE SKILLS." International Journal of English, and Literature (IJEL) 6.1 (2016): 47-58.