NeuraLingo: A Multilingual Chat Application

A Project Work Synopsis

Submitted in the partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

IN

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Submitted by:

APRATIM SADHU
SAWAN SIHAG
ARSHIYA SARMAI
NEEL MADHAB ROY
ANMOL CHOPRA

19BCS6041
19BCS6059
19BCS6060
19BCS6064
19BCS6069

Under the Supervision of:

DR. MONIKA SINGH

CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413, PUNJAB

FEBRUARY 2020

Table of Contents

Abstract

Machine Translation bridges communication barriers and eases interaction among people having different linguistic backgrounds. Machine Translation mechanisms exploit a range of techniques and linguistic resources for translation prediction. Neural machine translation (NMT), in particular, seeks optimality in translation through training of neural network, using a parallel corpus having a considerable number of instances in the form of a parallel running source and target sentences. This project uses neural machine translation to translated English into some of the most spoken languages around the world such as Spanish, French, German, Japanese, Hindi and others. NMT has been implemented in this project using two types of networks: simple encoder-decoder networks and transformer networks. The predicted translation performance of the model for different languages have been evaluated using Bilingual Evaluation Understudy(BLEU) score and by human evaluators to assess the quality of translation in terms of its adequacy, fluency, and correspondence with human-predicted translation.

List Of Figures

List Of Tables

Timeline/Gantt Chart

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12	Week 13	Week 14
	12 Feb-27 Feb		28 Feb-13 Mar		14 Mar-03 Apr			04 Apr-24 Apr			25 Apr- 15 May			16-21 May
Project Proposal														
Project Synopsis														
Data Collection and pre-processing														
Modelling of LSTM-Encoder-Decoder														
Modelling of Transformer														
Model Evaluation and retraining														
Model Deployment and Android App Development											**Submission of Records**			
Writing of Research Paper														
Internal Presentation														
External Assessment														

Timeline of the Project

1 INTRODUCTION

1.1 Problem Definition

Communication and information exchange between people is necessary not only for business purposes but also for the people to share feelings, thoughts, opinions and facts. But language barriers between different countries pose a significant problem for effective exchange of information. This language barrier is a primary reason for ineffective communication. Sharing information between people is necessary not only for business purposes but also for sharing their feelings, opinions, and acts. To this end, translation plays an important role in minimizing the communication gap between different people.

India, for instance is a multilingual country with people from different states speaking in different regional languages. It has 23 constitutionally recognized official languages and several hundreds unofficial local languages. Despite Indian population is approximately 1.3 billion, only approximately 10% of them English speak English. Some studies say that out of these 10% English speakers only 2% can speak, write, and read English well, and rest 8% can merely understand simple English and speak broken English with an amazing variety of accents. Considering a significant amount of valuable resources is available on the web in English and most people in India can not understand it well, it is essential to translate such content in to local languages to facilitate people.

Considering the vast amount of information, it is not feasible to translate the content manually. Also it is not feasible to have human translators everywhere, we need effective approaches which do this job with as little human effort as possible. Hence, it is essential to translate text from one language to another language automatically. Machine Translation (MT) can be described as the task of translating text or speech from one natural language to another, with as little human effort as possible. MT aims to achieve quality translations which are semantically equivalent to the source sentence and syntactically correct in the target language. MT performs simple substitution of words on a ground level, but that alone is not enough, as recognition of whole phrases and their closest counterparts in the target language are necessary.

1.2 Project Overview

This project aims at accurate and effective translation of English to 6 most commonly spoken languages around the world using neural machine translation. This project implements NMT using two types of network: conventional encoder-decoder network and transformer network.

We have worked with 6 languages: English-German, English-Spanish, English-French, English-Hindi, English-Japanese, English-Bengali. This project is one of the most difficult application of NLP. The types of neural networks for this purpose comes under the class of Seq2Seq models. The use of transformer network for NMT is better than that of encoder-decoder models using LSTMs. The purpose of this project is to evalute the performance of these two types of networks. Easy availability of parallel corpora of the mentioned languages was an useful for creation of training dataset for implementation. The corpus has been cleaned and necessary preprocessing is carried out prior to modelling. The encoder-decoder and transformer networks are implemented using Keras API of Tensorflow framework.

A lot of work has been done on NMT. Previous works in NMT have been done in these particular languages. There are benchmarks available for English-French and English-German on WMT'14 datasets (which are publicly available) with about 38.95 and 24.67 BLEU score respectively [1]

The approach in this project is to develop a NMT model with better translation accuracy compared to previously developed models. The models developed will be compared using BLEU score as use of accuracy alone is not an ideal metric for seq2seq models.

The final outcome of the project is an android chat application. The purpose of this application is to help peers-to-peer communication in their the language of their choice. The application also has a translation feature which will help to find the translation of English to a language of user's choice.

1.3 Software Specification

The following software specifications are required for the successful completion of the proposed project:

- Python 3.7 or higher
- Jupyter/Colab Notebook
- Numpy, Scikit-Learn, NLTK ad other machine learning libraries
- TensorFlow/Keras deep learning framework
- Flask web development framework
- VSCode code editor for editing the python script.
- Browser

1.4 Hardware Specification

The hardware specifications required for the proposed project includes:

and GPU.

- A minimum of 8GB RAM. And GPU with at least 2GB RAM.
- A minimum of 20GB hard drive space.
- Internet connection for accessing online GPU such as Kaggle or Google Colab.

2 LITERATURE REVIEW

Rule based machine translation is the first approach towards machine translation. This strategy utilizes a lot of human efforts for part-of-speech tagging, syntactic parsers and bilingual dictionaries. Ghosh et al.[2] in their paper have used such a rule based translation strategy to translate Telugu to Marathi and vice-versa. Their research focuses on idioms and proverbs of both languages. Direct translation was used to translate these two languages as they have the same grammatical arrangement of sentences. Their approach is based on POS tagging and the authors have concluded that many complex sentences have their words interchangeable in order to get translation in the final language in direct translation.

Zhixing Tan et al.[3] in their paper have discussed methods related to NMT and have mentioned strategies related to architectures, decoding and augmentation. The authors have mentioned two previously developed techniques: statistical machine translation(SMT) and neural machine translation(NMT). However their research is focused on NMT and its associated architects along. They have mentioned beam search algorithm used for NMT. Their research have mentioned the use of attention mechanism[4] in transformer networks. Unsupervised NMT was also mentioned by the authors for translation of languages whose parallel translated corpora is not available. They have also summarized open-source NMT tools and tools for evaluation and analysis. The authors have concluded their paper by discussing the challenges and future scope for NMT tasks.

Kyunghyun Cho et al. [5] in their paper have compared and discussed two models for NMT, encoder-decoder model using RNN and their proposed model: gated CNN. The models are implemented using beam-search algorithms for translation of English to French. They have evaluated the models and using BLEU score. Their proposed model, grConv concsists of 2,000 neurons as compared to 1,000 neurons of RNNenc model. On training the models on English-French pairs, the authors have concluded that the performance of NMT suffers greatly from the length of associated sentence length. They have concluded that both the models are able to translate the sentences with similar accuracy, however they their performance decreases when the sentence length increases.

Felix Stahlber [6] in this research have presented a review on all the present NMT strategies and techniques. They have discussed in detail the available NMT architectures and associated algorithms. They have discussed the use of transformer networks with attention mechanism and greedy and beam search algorithms. They have reviews commonly used architectures such as recurrence, convolutions and attention. They have also discussed the advantages and disadvantages of different strategies along with the factors affecting the performance of mentioned models. They have concluded the paper discussing the future aspects of machine translation.

Amarnath Pathaket al.[7] in their paper have implemented NMT using encoder-decoder network with attention for translation of English to Hindi, Punjabi and Tamil from a parallel corpora of Indian languages. They have evaluated their model using BLEU score. The experimental setup used by the authors in their research used BLEU to evaluate the translation performance on different epochs, data size and on different sentence length. The authors have concluded that the translation performance depends largely on the size of training corpora and the performance of the models are heavily enhanced on using attention mechanism in the encoder-decoder network.

Dzmitry Bahdanau et al. [8] in their research have proposed a novel approach towards neural machine translation. They have translated English to French with a parallel corpora of 61M words. The authors in their proposed model RNNsearch have extended the basic encoder–decoder by letting a model (soft-)search for a set of input words. The authors have shown that this approach frees the model to encode a fixed length vector and lets the model focus on relevant information for generation of the next word. Their model outperforms the traditional RNNencdec model. Their proposed model performed well while predicting longer sentences. The authors concluded that their model achieved results comparable to existing phrase-based statistical machine translation.

Kartik Revanuru **[**9**]** in their research paper have created a system which has various models and they have applied the Neural Machine Translation techniques for the creation of this system. It has been applied to six Indian language pairs. In their research paper, they have demonstrated that they were able to achieve good accuracy with less data and shallow networks of two layers. After comparing their test set with the Google translate, their models for Urdu-Hindi, Gujarati-Hindi and Punjabi Hindi outperformed Google translate with the BLUE score of 17, 30 and 29 respectively. The authors concluded the research paper discussing the future aspects how it can also be extended for real-time speech to speech translation.

Shuangzhi Wu **[**10**]** in their research paper have proposed a novel Sequence-to-Dependency Neural Machine Translation Method. In this method, the construction and the modeling of the target word sequence and its dependency structure will be done together. This structure will be used as context to simplify word generations. According to the authors, their proposed method can improve the quality of translation of Neural Machine Translation systems. The authors have concluded that their method is able to outperform baselines of state-of-art in Japanese-English and Chinese-English translations.

Himanshu Choudhary **[**11**]** in their research paper have proposed a novel Neural Machine Translation technique which uses Byte-Pair-Coding along with using word embedding. The proposed NMT technique has been applied to English-Tamil pair language.They have shown that this technique performed better than complex techniques specifically for the Indian languages. The main aim to propose this technique was to overcome the Out-of-Vocabulary (OOV) problem for the languages whose translations are not much available online. The authors have concluded that their proposed MIDAS translator was able to outperform Google Translate with a BLUE score margin of 4.58.

Translation is an open vocabulary problem, but neural machine translation usually works with fixed vocabulary. There are some previous works which addresses the out-of-vocabulary translation problem by using backing off to a dictionary. Rico Sennrich et at.[12] proposed a simpler and more effective approach to resolve out-of-vocabulary problems, by encoding rare and unknown words as a sequence of subword units and the author did this without using a back-off model for rare words. The authors used encoding (rare) words via subword units and byte pair encoding (BPE) for word segmentation, to perform open-vocabulary neural machine translation. Their research concluded that subword models achieve better accuracy than large-vocabulary models and back-off dictionaries for the translation of rare words. The author also mentioned that their model is able to generate new words which were not present in training time.

Nadeem Jadoon Khan et at.[13] in their research paper they used Phrase-based Statistical Machine Translation (SMT) to analyze the performance of multiple Indian Languages. The author mentioned the performance of Indian languages (Bengali, Guajarati, Hindi, Malayalam, Punjabi, Tamil, Telugu and Urdu) to English language with an average accuracy of 10% on baseline system translation. The language used by the author in their research paper has sparse-resources; due to that they carry out a low BLEU score with a mean of 0.12. The author also mentioned the different BLEU scores for different language pairs. They concluded the paper by discussing the future aspects of Statistical Machine Translation (SMT) by using discrete approaches to develop quality language models.

K Hans et at.[14] in their paper have compared and discussed five different models for English to Tamil language translation pair. The models are Statistical Machine Translation (SMT), Phrase-Based SMT, RNNSearch, RNNSearch with Word2Vec and RNNMorph. It was observed that the performance of Phrase-Based SMT model was inferior to the RNNSearch Model in terms of BLEU score, when RNNSearch model is been used with word2vec vectors there has been slightly increment in the BLEU score as compared to the BLEU score of RNNSearch, use of morphological segmentation enhance the performance of RNNMorph neural machine translation by 7.05 BLEU points on top of RNNSearch Model. They concluded their paper by discussing future aspects to carry out an end-to-end translation methodology for morphologically rich languages.

Prakash and Sasikumar [15] presented a English to Indian language machine translation that poses the challenge of structural and morphological divergence. We used pre-ordering and suffix separation for translation. Pre-ordering or reordering transforms the source sentence into a target-like order using syntactic parse tree of the source text. One of the main issue in this translation was English uses the Subject-Verb-Object (SVO) order and most of the Indian languages, including the ones under study, primarily use Subject Object-Verb (SOV). Out of all the models Factored SMT with suffix separation and reordering performs better. Transliteration as postprocessing further helps to improve the translation quality. However, there were problems while translation between English to Malayalam and Punjabi.

Roee, Melvin and Orhan [16] presented a research on Multilingual neural machine translation (NMT) that enables training a single model that supports translation from multiple source languages into multiple target languages. However, it was shown in a somewhat extreme case with more than 100 languages trained jointly, where we saw that in some cases the joint training may harm the performance for some language pairs (i.e. German-English above).

Rico and Bios [17] demonstrated the performance of neural machine translation (NMT) that drops starkly in low-resource conditions, underperforming phrase-based statistical machine translation (PBSMT) and it requires large amounts of auxiliary data to achieve competitive results. Results show that low-resource NMT is very sensitive to hyperparameters such as BPE vocabulary size, word dropout, and others, and by following a set of best practices, we can train competitive NMT systems without relying on auxiliary resources.

Minh-Thang Luong et al.[18] in their research have implemented neural machine translation with two attention types; local and global. The models developed the authors are trained on the WMT'14 training data consisting of 4.5M sentences pairs (116M English words, 110M German words). Their local attention approach yielded a gain of 5.0 BLEU over non-attentional models. Their English-German approach using attention model has achieved state-of-the-art results for both WMT'14 and WMT'15 and outperformed existing models by more than 1 BLEU.

Lucia Benková et al.[19] in their paper have discussed two most common approach towards neural machine translation; statistical machine translation(SMT) and neural machine translation(NMT). The author discussed advantages and disadvantages of both approaches and have concluded that NMT provides better translation as compared to SMT. The have also mentioned that SMT fulfills certain shortcomings of NMT.

Sree Harsha Ramesh et al.[20] in their research have implemented SMT AND NMT on the augmented parallel corpora of two languages: English-Hindi and English-Tamil. Initially, they have extracted parallel corpora from Wikipedia pages using Siamese BiRNN encoder using GRU as the activation function. The models implemented yielded a percentage increase in BLEU scores of 11.03% and 14.7% for en–ta and en–hi pairs respectively, due to the use of parallel sentence pairs extracted from comparable corpora using the neural architecture.

2.1 Literature Review Summary

Table 2.1 Literature review summary

Year of Publication	Title	Author	Dataset	Methodology	Reference
2014	Translation of telugu-marathi and vice-versa using rule based machine translation	Dr. Siddhartha Ghosh , Sujata Thamke Kalyani U.R.S	-	POS tagging, syntactic parsing	2
2020	Neural machine translation: A review of methods, resources, and tools	Zhixing Tan, Shuo Wang, Zonghan Yang, Gang Chen, Xuancheng Huang, Maosong Sun, Yang Liu	WMT20, IWSLT20, WAT20	Beam Search, Transformers	3
2014	On the Properties of Neural Machine Translation: Encoder–Decoder Approaches	Kyunghyun Cho, Bart van Merrienboer, Dzmitry Bahdanau, Yoshua Bengio	news-test2012 and news-test2013 and news-test2014	Beam Search, RNN-encoder-decoder, Gated CNN	5
2020	Neural Machine Translation: A Review and Survey	Felix Stahlberg	-	Beam Seach, Greedy Search, Attention mechanism	6
2019	Neural Machine Translation for Indian Languages	Amarnath Pathak and Partha Pakray	MTIL corpora	Encoder Decoder with attention	7
2016	Neural machine translation by jointly learning to align and translate	Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio	WMT ’14	RNN-encoder-decoder model	8

2017	Neural Machine Translation of Indian Languages	Karthik Revanuru, Kaushik Turlapaty, Shrisha Rao	Technology Proliferation & Deployment Centre (TDIL-DC)	LSTM, Bidirectional LSTM, Deep BiLSTM with residual connection and Attention	9
2017	Sequence-to-Dependency Neural Machine Translation	Shuangzhi Wu, Dongdong Zhang, Nan Yang, Mu Li, Ming Zhou	LDC datasets, NIST2005, NIST2006, NIST2008 and NIST2012.	Sequence-to-Dependency Neural Machine Translation (SD-NMT)	10
2018	Neural Machine Translation for English-Tamil	Himanshu Choudhary, Aditya Kumar Pathak, Rajiv Ratn Shah, Ponnurangam Kumaraguru	EnTam V2.0 and Opus	BiLSTM encoder-decoder with attention	11
2016	Neural Machine Translation of Rare Words with Subword Units	Rico Sennrich, Barry Haddow, Alexandra Birch	WMT 2015	Byte Pair Encoding with BiLSTM encoder-decoder model	12
2017	Machine Translation Approaches and Survey for Indian Languages	Nadeem Jadoon Khan, Waqas Anwar, Nadir Durrani	EMILLE (Enabling Minority Language Engineering).	Moses System	13
2017	Improving the Performance of Neural Machine Translation Involving Morphologically Rich Languages	K Hans, Milton R S	EnTam v2 corpus and TDIL (Technology Development for Indian Languages)	Phrase-Based SMT, RNNSearch, RNNSearch + Word2Vec, RNN with Morphology	14
2018	Machine Translation in Indian Languages: Challenges and Resolution	Raj Nath Patel, Prakash B. Pimpale, Sasikumar M	MTIL-2017	Transliteration in the postprocessing phase	15
2019	Massively Multilingual Neural Machine Translation	Roee Aharoni, Melvin Johnson, Orhan Firat	TED Talks parallel corpus, WMT news translation	NMT models by different authors	16

2019	Revisiting Low-Resource Neural Machine Translation: A Case Study	Rico Sennrich, Biao Zhang	TED data from the IWSLT 2014	Phase based SMT, NMT baseline, NMT optimized	17
2015	Effective Approaches to Attention-based Neural Machine Translation	Minh-Thang Luong, Hieu Pham, Christopher D. Manning	WMT’14 t	Attention based encoder-decoder model	18
2020	Neural Machine Translation as a Novel Approach to Machine Translation	Lucia Benková, Ľubomír Benko	-	NMT and SMT	19
2018	Neural Machine Translation for Low Resource Languages using Bilingual Lexicon Induced from Comparable Corpora	Sree Harsha Ramesh. Krishna Prasad Sankaranarayanan	Wikipedia	Siamese BiRNN encoder using GRU as the activation function	20

3 PROBLEM FORMULATION

Human communication is a social interaction process. It is an essential part of our daily life. It is a process of creating, exchanging, sharing ideas, information, opinions, facts, feelings, and experiences between a sender and a receiver. Communication is fundamental to the existence and survival of individuals, groups, societies, and nations. Language is the most common tool of communication. And multiplicity of languages plays a major role in dividing factors in world society, reinforcing geographical, socio-economic (especially caste or class), political, ideological, professional and religious separatism. Language barriers are the root causes of many problems or obstacles in health care, aviation, maritime,business, and education.

A German study titled "Language Barriers in Different Forms of International Assignments" [21] has connected language barriers to a series of organizational behavior phenomena. The results showed that language barriers have effects on the multinational corporation as follows: effects on employees' emotions, social identity formation, trust formation, power relations.

Here are some language barriers

- Language difference, where a person interacts with someone speaking a different native language
- Dialects and accents, where two people may share a common language but they speak it differently (based on a particular region)
- Lack of clear speech, where people speak too soft or too fast; either way, it’s unclear what they’re saying
- Use of technical words or jargon, where someone communicates using specific terms that are highly technical and subject-specific
- Word choice, where someone uses words with two meanings or says it sarcastically that may be misinterpreted by the listener.

Problem with English as a base language: - English is now globalized as English stands at the very center of the global language system, and as international students from non-English speaking countries enter the western world, where speaking English is the norm, they are faced with the challenge of learning a new language as a prerequisite to successfully acculturation and further thriving. A study found that most international students from a non-English speaking background did not meet the expectations, they struggled with class participation, writing, comprehension, reading and may suffer from a lack of self-confidence.

Limitations of Human Translations: - The demand for translation jobs is literally skyrocketing. The US Bureau of Labor Statistics predicts, during the period 2012-2022, jobs in the translation industry will grow by 46 percent.

- Slow Speed - On an average, it is estimated that a human translator can translate 200-250 words in an hour. This corresponds to 1000-1500 words per day, considering breaks, checks etc. If your document is made up of 10,000 words or more, it would easily take 10 working days to be translated. Compare this with a machine translator, which will literally do the job in a few seconds.
- Higher Cost - A human translator roughly makes around $45,000 a year. However, depending on the work and how long it will take, it is very expensive. While using a machine translator you can work on as many documents as you want for free of cost.
- Confidentiality Risks: - When working with documents that are highly confidential, it is usually hard to keep them confidential if you work with a human translator, but that’s not the case while working with machine translators.
- Time and Resources: - Depending on your work and the target language you are trying to translate, finding a good translator has its own challenges. A human translator, translates max to max 10 languages whereas a machine translator can translate 108 languages.

4 RESEARCH OBJECTIVES

The proposed work is aimed to carry out work leading to the development of an android chat application which will have the feature for language translation in the application itself. The proposed aim will be achieved by dividing the work into following objectives:

1. Creation of the language model pairs using Deep Learning for Neural Machine Translation:
    - English-German Language model
    - English-Spanish Language model
    - English-French Language model
    - English-Japanese Language model
    - English-Hindi Language model
    - English-Bengali model
2. Implementation of simple encoder-decoder model using LSTM networks.
3. Implementation of transformer network using multi-head attention
4. Comparison of implemented models using Bilingual Evaluation Understudy score.
5. Deploying the best NMT models of each language using APIs for android chat application.

5 METHODOLOGY

The following methodology will be followed to achieve the objectives defined for the proposed research work:

1. Translate data of multiple languages will be used. They are as follows:
    - German to English
    - French to English
    - Spanish to English
    - Japanese to English
2. A detailed study and analysis of the datasets will be carried out in order to make the neural translation Model.
3. NLP preprocessing will be done on datasets, such as tokenization and vectorization will be done prior to modelling.
4. An encoder-decoder with LSTM network will be implemented using Keras API along with transformer network with multi-head attention.
5. The model is to be trained to a considerable number of epochs for convergence.
6. The models for different language translation will be evaluated using BLEU score.
7. On achieving the desired results, the best model will be evaluated again and the evaluations will be visualized.
8. The visualizations and a comparison between the models trained on the dataset will be done for research
9. The best model will be used for deployment using Django API Framework.
10. An android app will be created to access the model and translate multiple languages to English.
11. App will also have the feature to create a Chat room, people from different linguistic backgrounds will be able to Chat with each other in the room, as they will be provided with live translation in their own language of choice.

The proposed methodology of the project is summarised in Figure 1.

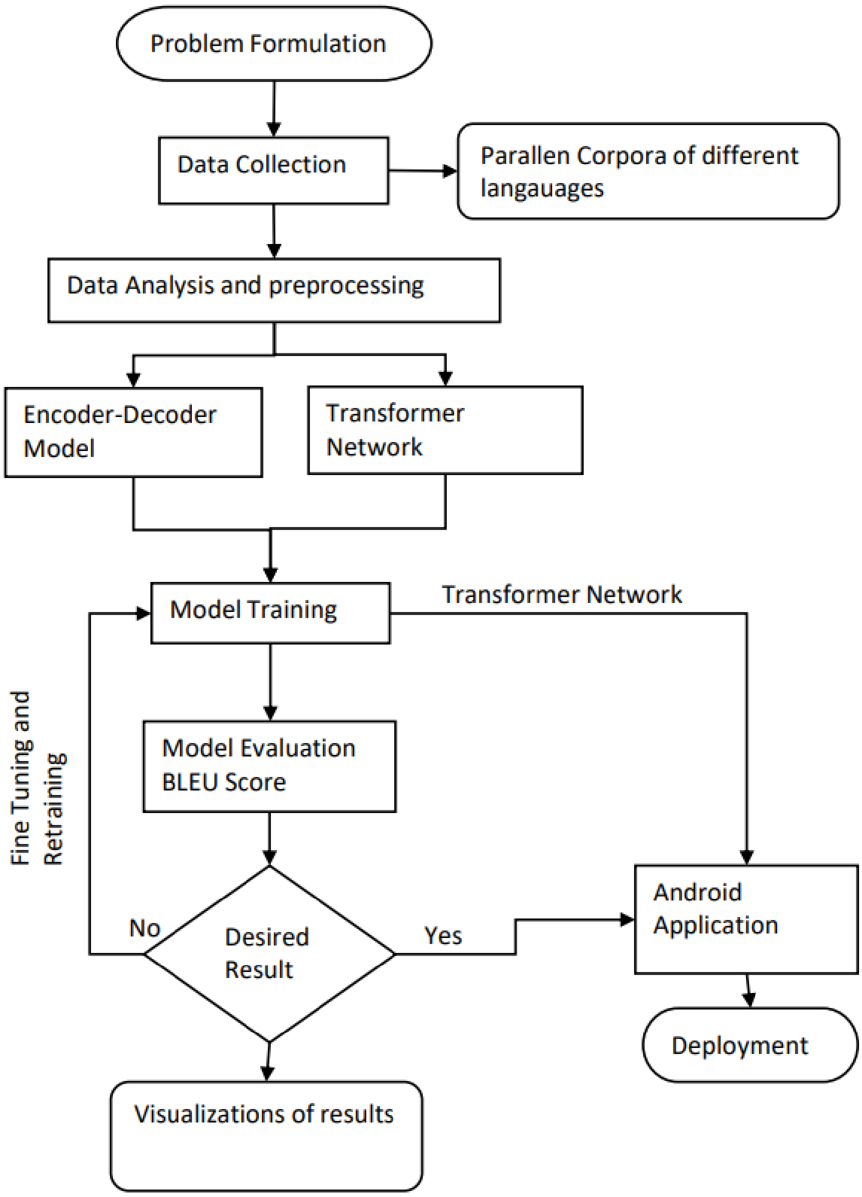

Fig 1. Flowchart of Proposed Methodology

6 TENTATIVE CHAPTER PLAN FOR THE PROPOSED WORK

CHAPTER 1: INTRODUCTION

This chapter will cover the overview and blueprint of the project.

CHAPTER 2: LITERATURE REVIEW

This chapter include the literature available for the proposed implementation of the project ideas. The findings of the researchers will be highlighted which will become basis of current implementation.

CHAPTER 2: BACKGROUND OF PROPOSED METHOD

This chapter will provide introduction to the concepts which are necessary to understand the proposed system.

CHAPTER 4: METHODOLOGY

This chapter will cover the technical details of the proposed approach.

CHAPTER 5: EXPERIMENTAL SETUP

This chapter will provide information about the subject system and tools used for evaluation of proposed method. Along with this the implementation detail of the networks will be discussed.

CHAPTER 6: RESULTS AND DISCUSSION

The result of proposed technique will be discussed in this chapter.

CHAPTER 7: CONCLUSION AND FUTURE SCOPE

The major finding of the work will be presented in this chapter. Also directions for extending the current study will be discussed.

REFERENCES

7 REFERENCES

[1] Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, Jeff Klingner, Apurva Shah, Melvin Johnson, Xiaobing Liu, Lukasz Kaiser, Stephan Gouws, Yoshikiyo Kato, Taku Kudo, Hideto Kazawa, Keith Stevens, George Kurian, Nishant Patil, Wei Wang, Cliff Young, Jason Smith, Jason Riesa, Alex Rudnick, Oriol Vinyals, Greg Corrado, Macduff Hughes, and Jeffrey Dean. 2016. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. CoRR abs/1609.08144 (2016). http://arxiv.org/ abs/1609.08144

[2] Ghosh, Siddhartha & Thamke, Sujata & U R S, Kalyani. (2014). Translation Of Telugu-Marathi and Vice-Versa using Rule Based Machine Translation. Computer Science & Information Technology. 4. 10.5121/csit.2014.4501.

[3] Tan, Zhixing & Wang, Shuo & Zonghan, Yang & Chen, Gang & Huang, Xuancheng & Sun, Maosong & Liu, Yang. (2020). Neural machine translation: A review of methods, resources, and tools. AI Open. 1. 5-21. 10.1016/j.aiopen.2020.11.001.

[4] Vaswani, A., Shazeer, N.M., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., & Polosukhin, I. (2017). Attention is All you Need. *ArXiv, abs/1706.03762*.

[5] Cho, Kyunghyun & Merrienboer, Bart & Bahdanau, Dzmitry & Bengio, Y.. (2014). On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. 10.3115/v1/W14-4012.

[6] Stahlberg, F. (2019). Neural Machine Translation: A Review and Survey. *arXiv: Computation and Language*.

[7] Pathak, A., & Pakray, P. (2019). Neural Machine Translation for Indian Languages. *Journal of Intelligent Systems, 28*, 465 - 477.

[8] Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural Machine Translation by Jointly Learning to Align and Translate. *CoRR, abs/1409.0473*.

[9] Revanuru, K., Turlapaty, K., & Rao, S. (2017). Neural Machine Translation of Indian Languages. *Compute '17*.

[10] Wu, S., Zhang, D., Yang, N., Li, M., & Zhou, M. (2017). Sequence-to-Dependency Neural Machine Translation. *ACL*.

[11] Choudhary, H., Pathak, A.K., Saha, R.R., & Kumaraguru, P. (2018). Neural Machine Translation for English-Tamil. *WMT*.

[12] Sennrich, R., Haddow, B., & Birch, A. (2016). Neural Machine Translation of Rare Words with Subword Units. *ArXiv, abs/1508.07909*.

[13] Khan, N.J., Anwar, W., & Durrani, N. (2017). Machine Translation Approaches and Survey for Indian Languages. *ArXiv, abs/1701.04290*.

[14] Krupakar, H., & Milton, R.S. (2016). Improving the Performance of Neural Machine Translation Involving Morphologically Rich Languages. *ArXiv, abs/1612.02482*.

[15] Patel, R.N., Pimpale, P.B., & Sasikumar, M. (2019). Machine Translation in Indian Languages: Challenges and Resolution. *Journal of Intelligent Systems*, *28*, 437 - 445.

[16] Aharoni, R., Johnson, M., & Firat, O. (2019). Massively Multilingual Neural Machine Translation. *ArXiv, abs/1903.00089*.

[17] Sennrich, R., & Zhang, B. (2019). Revisiting Low-Resource Neural Machine Translation: A Case Study. *ArXiv, abs/1905.11901*.

[18] Luong, Minh-Thang & Pham, Hieu & Manning, Christopher. (2015). Effective Approaches to Attention-based Neural Machine Translation. 10.18653/v1/D15-1166.

[19] Benková, Lucia & Benko, Ľubomír. (2020). Neural Machine Translation as a Novel Approach to Machine Translation.

[20] Ramesh, Sree & Sankaranarayanan, Krishna. (2018). Neural Machine Translation for Low Resource Languages using Bilingual Lexicon Induced from Comparable Corpora. 112-119. 10.18653/v1/N18-4016.

[21] Tenzer, Helene & Schuster, Tassilo. (2017). Language Barriers in Different Forms of International Assignments. 10.1057/978-1-137-57406-0_3.