A Project Report

On

MACHINE TRANSLATING HINDI TO ENGLISH

Submitted in partial fulfillment of the

Requirement for the award of the degree of

Bachelor of Technology

Computer Science and Engineering

Under The Supervision of
Dr. S. Srinivasan
Professor

Submitted By

Kaushki kumari (18021011510)

Lakshay Ahlawat (18021011600)

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA
INDIA
DECEMBER, 2021

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA

CANDIDATE'S DECLARATION

We hereby certify that the work which is being presented in the project, entitled **"MACHINE TRANSLATION HINDI TO ENGLISH"** in partial fulfillment of the requirements for the award of the submitted in the School of Computing Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of month, **JULY-2021 to DECEMBER-2021**, under the supervision of **Dr S.SRINIVASAN**, **Professor Department of Computer Science and Engineering**, of School of Computing Science and Engineering , Galgotias University, Greater Noida

The matter presented in the project has not been submitted by us for the award of any other degree of this or any other places.

Lakshay Ahlawat
Kaushki Kumari

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Dr .S. Srinivasan
Professor

CERTIFICATE

The Final Project examination of **Lakshay Ahlawat (18SCSE1010368)** & **Kaushki kumari(18SCSE1010273)** has been held on December 2021 and the work is recommended for the award **of BACHELOR OF TECHNOLGY IN COMPUTER SCIENCE AND ENGINEERING.**

Signature of Examiner(s) **Signature of Supervisor(s)**

Signature of Project Coordinator **Signature of Dean**

Date: December, 2021
Place: Greater Noida

Acknowledgement

We would like to express my special thanks to our college “GALGOTIAS UNIVERSITY” as well as our guide **Dr. S. Srinivasan Sir** .who gave us the golden opportunity to do this wonderful project on the topic “MACHINE TRANSLATION HINDI TO ENGLISH”, which also helped me in doing a lot of research and I come to know about so many thing

Secondly we would also like a thank my teammate .who do so many things to finishing this project

ABSTRACT

Multilingual Machine Translation is a computerized system that is designed to translate source text from various natural languages into target text of another natural languages. Due to increased need of global communication, multilingual machine translation is the propel for researchers. With the advancement in technology, now we have computerized systems that can replace the human experts in particular domains. One such popularly adopted domain by researchers is artificial intelligence (AI) and NLP .Natural language processing (NLP) is a field of computer science concerned with the interaction between computer and human (natural) languages. It is becoming one of the most active area in the interaction between human and computer. These includes text system that integrate text and natural language. Machine translation (MT) system is available mostly for a specific language pair. Recently many MT systems are being designed for bilingual pairs. And the research is heading towards developing MT systems that are capable of handling more than two language pairs i.e. multilingual. MT systems still has so many issues to deal with. Recently developed multilingual systems are laden with the enormous size corpus .In this proposed system we are analyzing machine translation system using NLTK

Contents

List of Table

List of Figures

Acronyms

B.Tech.	Bachelor of Technology
NLP	Natural language processing
SOV	Subject objective verb
MTS	Machine translation system
NLTK	Natural language tool kit
SMT	Statistical machine translation
RMBT	Rule-based machine translation

CHAPTER-1

1.1 INTRODUCTION

The main purpose of Natural Language Processing is for a Hindi sentence to be interpreted by the computer and appropriate action taken. Despite the challenges, natural language processing is widely regarded as a promising and critically important set out in the field of computer research. The applications that will be possible when NLP capabilities are fully realized are impressive computers would be able to process natural language, translating languages accurately and in real time, or extracting and summarizing information from a variety of data sources, depending on the users' requests. This system focuses on the solution of the problems arising in the analysis or generation of Natural language text, such as syntactic and semantic analysis or compilation of dictionaries and grammars necessary for such analysis. It proposes the architecture for translating Natural language Hindi to English

1.2 FORMULATION OF PROBLEM

This section reports some inherent challenges in Hindi to English translation systems. Linguistically, morphological manifestation and structural divergences are important characteristics on which the two language can be differentiated. English is based on Subject-Verb-Object (SVO) structure, but Hindi is a Subject-Object-Verb (SOV) type of language. Hindi is morphologically more rich than English. In general, these divergences are the factors which make the translation process difficult and error-prone. Furthermore, Hindi also have some inherent challenges in translating to English Lack of articles in Hindi makes the translation imprecise. Multiple contextual meaning of English prepositions makes it difficult to predict them accurately The major challenge in the machine translation (MT) between two languages is to identify an inherent translation divergence exist between source and target language. To elaborate more, the divergence can be observed when a sentence in a source language L1 translated to target language in a quite different form. For a robust machine translation (MT) system, it is crucial not only to identify the type of translation divergences but also to resolve them in order to obtain more accurate translation In this paper, several Hindi-English translation divergences have been studied in order to identify language specific divergences and further to be incorporated in EBMT and RBMT phases during the translation. In terms of configurational characteristics, English is more rigid and restrictive that follows fixed word order patterns as opposed to Hindi. For instance, one of the translation divergences related to specific word-order pattern is the interpretation of a Hindi question

1.1.2 TOOL AND TECHNOLOGY USED

SYSTEM REQUIREMENT:

1.	**OPERATING SYSTEM (WINDOW)**
2.	**ANACONDA**
3.	**PYTHON IDE**

Table 1

TECHNOLOGY USED:

1.	**MACHINE LEARNING**
2.	**ARTIFICIAL INTELLIGENCE**
3.	**NATURAL LANG. PROCESSING(NLP) , NATURAL LANG. TOOL KIT (NLTK)**

Table 2

CHAPTER – 2

2.1 LITERATURE REVIEWS

According to the traditional Systems and prior research, the methodologies and technologies used for translation A Survey of Translation Quality of English to Hindi Online Translation Systems (Google and Bing) , “” India has declared a large assembly in Hindi is the language they speak and in several areas it works in all type of official and study. Many online translator technologies today use different machine translation approach like each translation approaches diverse characteristics; the outcome of the translation would be different. Bing Translator and Google Translate free online machine translators is using statistical machine translation. Both translators are of the most accepted. It keeps growing the language option and increasing its usability. The procedure Due to the major reverse characteristics of both the translator services and their important role for the growth of machine translator’s particular in internet platform, it is determined to have a study about their similarity. The idea of this study is to create understanding about the special performance of the two online translation services due to the same procedures they include. The testing designed is intended to show how the two online translation services have its own benefit and disadvantage which can influence their performance. From this investigation it can learned that the characters of two online machine translator application, Bing machine translators and Google machine translator. They have difference in basic features, where it leads to differences in its structure architecture. Bing and Google machine translator exercise statistical information of earlier translation to learn about the language which is to be translated. Statistical machine translators are separated into three kinds, word-based, syntax based and phrase-based. They chose Bing Tend Google Translator represent English to Hindi translation as it is the most known online machine translators. Mining Hindi-English Transliteration pair from Online Hindi Lyrics, “Approximate string-matching algorithm” explains a procedure to mined Hindi-English transliteration pairs from online Hindi song lyrics. The technique is based on the clarification that lyrics are transliterated word-by-word, maintaining the precise word order. The mining task is

Nevertheless difficult because the Hindi lyrics and its transliterations are usually available from different, often unrelated, websites. Thus, it is a non-trivial assignment to match the Hindi lyrics to their transliterated counterparts. Moreover, there are different types of noise in lyrics data that needs to be appropriately handled before songs can be aligned at word level. The supply data of 30823 unique Hindi-English transliteration pair with a precision of more than 92% is accessible publicly. Even though the present work information mining of Hindi-English word pairs, the same procedure can be easily adapted for supplementary languages for which song lyrics are accessible online in local and Roman scripts. In this paper explain a technique for extracting transliteration correspondent using online song lyrics. The mined data is of high quality and comparatively noise-free. This data can be used for training English-Hindi backward transliteration engine for building IME for Hindi and for other general applications of transliteration engines in IR and MT. The mined pairs will as well be useful for linguistic learning on spelling difference and typing patterns in text.

2.2 PROJECT DESIGN

FLOWCHART:

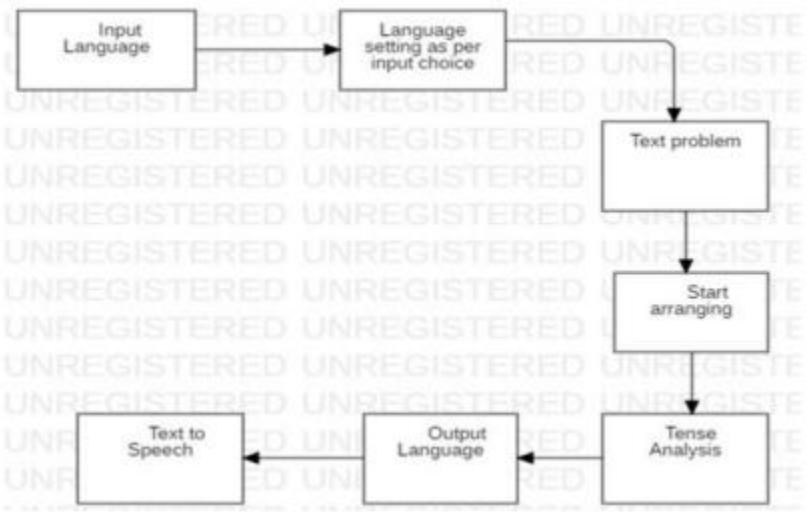

Figure 1

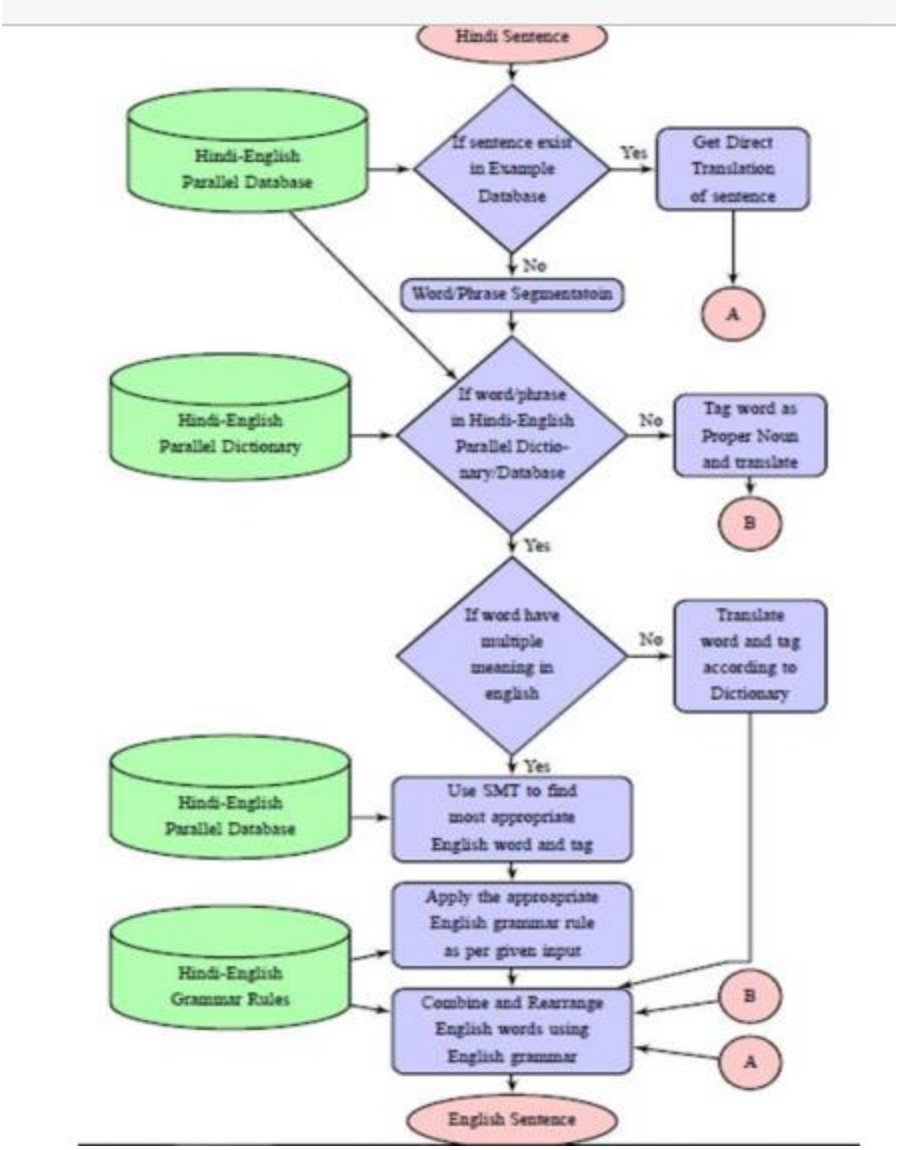
Hindi Sentence
Hindi-English Parallel Database
If sentence exist in Example Database
Yes
Get Direct Translation of sentence
No
Word/Phrase Segmentatoin
A
Hindi-English Parallel Dictionary
If word/phrase in Hindi-English Parallel Dictio-nary/Database
No
Tag word as Proper Noun and translate
B
Yes
If word have multiple meaning in english
No
Translate word and tag according to Dictionary
Yes
Hindi-English Parallel Database
Use SMT to find most appropriate English word and tag
Apply the appropriate English grammar rule as per given input
Hindi-English Grammar Rules
Combine and Rearrange English words using English grammar
B
A
English Sentence

Figure 2

CHAPTER - 3

3.1 FUNCTIONALITY OF PROJECT

The major challenge in the machine translation (MT) between two languages is to identify an inherent translation divergence exist between source and target language. To elaborate more, the divergence can be observed when a sentence in a source language L1 translated to target language L2, in a quite different form. For a robust machine translation (MT) system, it is crucial not only to identify the type of translation divergences but also to resolve them in order to obtain more accurate translation. In this paper, several Hindi-English translation divergences have been studied in order to identify language specific divergences and further to be incorporated in EBMT and RBMT phases during the translation. In terms of configurational characteristics, English is more rigid and restrictive that follows fixed word order patterns as opposed to Hindi. For instance, one of the translation divergences related to specific word-order pattern is the interpretation of a Hindi question particle "Èया". "Èया' can be used both as a question particle in yes/no type of question sentences and as a type of interrogative pronoun in content question sentences as shown
In example 1.

Example 1. a. "Èया आप ्लख रहे ह¢?" ⇒ Are you writing?

b. "आप Èया ्लख रहे ह¢?" ⇒ What are you writing?

Like most South Asian languages, Hindi also shows the replication phenomena of the lexical items to express different grammatical metaphors. In counterpart, English translation doesn't exhibit any such replicative structures. In example 2, the replicative element "चलते चलते" is an adverbial clause which is captured lexically in Hindi. The English counterpart for this example is obtained by a gerundive propositional phrase.

Example 2. "वह चलते चलते थक गया।" ⇒ He got tired of walking.

The existence of expressive words also responsible to make the translation process error-prone, due to lack of exact parallel counterpart in the target language. Generally, expressive words come from the sound associated with the semantics of the action verb such as टपटपाना (drip), खटखटाना (knock) etc. In example 3, the word धड़ाम is only distantly mapped by 'bump'.

Example 3. "वह धड़ाम से ग़ार४।" ⇒ She fell with a bump.

Another source of divergence, in Hindi-English translation is associated with the use of different conjunction and particles in Hindi. In Hindi some of these particles are such as {क,ना and या etc.} as shown in example 4.

Example 4. "राम èकूल गया है क मंÖदर।" ⇒ Ram went to school or temple.

The above discussion clearly represent the importance of study to exhaust all type of translation divergences in order to build a robust Hindi-English MT. As, it is difficult to process and deal with the all type of translation divergences simultaneously, we have tried to incorporate hybridized Example -based and rule-based techniques to sort out some of them. Since last several decades, people have developed a number of translation approaches to transform one language content to Another ranging from simple word-to-word translation systems to corpus based statistical models as in figure 1. Marcus et al. hypothesizes the translation as "If a sentence to be translated or a very similar one can be found in the TMEM1, an EBMT system has a good chance of producing a good translation. However, if the sentence to be translated has no close matches in the TMEM, then an EBMT system is less likely to succeed. In contrast, an SMT system may be able to produce perfect translations even when the sentence given as input does not resemble any sentence from the training corpus."

A. Rule Based Machine Translation

Jordi et al. have used rule based machine translation for Chinese to Spanish Machine translation [9]. In this work, they have used Pentium platform which is a toolbox for shallow transfer MT. For the generation of translation rules, a bilingual Chinese-Spanish dictionary is constructed consisting of almost 9000 distinctive words. Grammatical transfer-rules were developed manually. They test this system on different domains with average accuracy of 82%. Pratik et al. have worked on rule based English-to-Hindi machine translation, but their system is domain restricted they have used Dependency Parsing as an intermediate representation in the translation. During the translation, the phases of classical analysis, transfer, and generation strategies are replaced with a syntax planning algorithm that directly linearizes the dependency parse of the source sentence as per the syntax of the target language.

B. Statistical Machine Translation

A statistical approach based English-to-Hindi machine translation system is developed consisting of three processing units Language Model, Translation Model and Decoder [6]. Language model calculates the probability of a sentence in target language. Translation model designed to compute the target sentence probability for the given source sentence. Decoder's job is to select the target sentence which maximizes the probability. The SMT model is trained on the parallel dataset of 5000 sentences pairs. Google translator which is a worldwide renowned and mostly used bilingual translator, is also based on the SMT approach. Google translator learns the SMT parameters from their huge corpus collected from all over the web. The SMT accuracy depends on the corpus Quality and the parameter estimation needed be to learn. A group from IBM T.J Watson Research Center work on the Mathematics of SMT and the Parameter Estimation.

C. Example Based Machine Translation

Example based machine translation systems (EBMT) perform the translation of a given input sentence s in three consecutive phases (i) (matching) check for existence of the given input s in the bilingual corpus (ii) (retrieval) extraction of useful segments from the sentence that match in the bilingual corpus and (iii) (transfer) recombining the translated segments . EBMT practically based on the retrieval of source sentences similar to s in the bilingual corpus, hence EBMT is also known as source-similarity based translation. Harold et al. had worked on the Example based machine translation focused on various intuitive problems of the EBMT like the size of Parallel Corpora, Granularity of Examples, Quantity of Examples and Suitability of Examples [8]. Mainsheet al. proposed to apply EBMT with Fuzzy logic for English to Hindi machine translation. Fuzzy logic implemented in the matching and the alignment of segments during the translation

D. Hybrid Machine Translation

Hybrid machine translation is a method of machine translation that combines characteristics of multiple machine translation approaches within a single machine translation system. Paul et al. worked on a multi-engine hybrid approach to MT, utilizing the statistical models to generate the best possible output from multiple machine translation systems. He has found promising results for Japanese-English machine translation on applying a decision-tree method to select the best possible hypothesis obtained from multiple RBMT, EBMT and SMT decoders. Marcus et al. have also identified the benefits of hybrid MT approaches as coupled multiple MT systems have the precedence over utilizing each MT separately. Evidently, it is been clearly visible that multi-engine MT approaches are capable of surpassing the existing individual MT systems. A comparative study of various existing machine translation with the details of approaches they build on, a large number of works have been done on multiple Indian languages i.e. Marathi, Hindi, Sanskrit to English and vice versa at Center for Indian Language Technology (CFILT), IIT Bombay. A Hindi word-net is produced too by CFILT as a semantic relation among the words, which is useful for RBMT systems and helps structural disambiguation to resolve word and attachment ambiguities. Ondřej

Bo jar from Charles University in Prague has prepared the Hindi-English parallel Corpus which could be used in the Example Based Machine Translation systems

Overall approach has been implemented in a sequence of four primary steps: 1) Segmentation, 2) Translation, 3) POS Tagging and 4) Rearrangement. In figure 2, the flowchart depicts the working relation among multiple steps for translating a Hindi sentence "विकास विकास ने किया।" to English "Vikas did development.". A inputted sentence has to go through all these steps being transformed from one form to another and at last translated to corresponding English sentence. All these steps are explained briefly in the following subsections.

B. Algorithm

Algorithm is elaborated in detail in this section through all the steps as in figure 3 with suitable example. The first step in translation is to split the sentence into words or simple sentences (if the input sentence is complex/compound sentence or consists of phrases). If a part of the input is in the example database then we keep that part as it is and the remaining part is segmented into words. When the segmentation is done, the chunks or segments are actually translated and tagged independently. If the segment is example based, it is directly converted to English. Further, the remaining words are translated using the parallel Hindi-English Dictionary. place, or organization. For many languages, a word is known to be as proper noun if it doesn't belong to the dictionary of that language. But in Hindi most of the proper nouns also have the dictionary meaning. This makes the process of pronoun identification more complicated. The problem is solved through incorporating some rules based on the possible contextual morphological information required to denote a word as pronoun as in Example 6.

Example 6. Input1:- "मैं ओमकार विकास धारया हूँ"

Tagging1: ['<PRON>', '<Name>', '<Name>', '<Name>']

Input2:- "विकास ने विकास किया"

Tagging2:- ['<Name>', '<NOUN>', '<VERB>']
Tagging3: Tagging is the process to identify the linguistic properties of each individual textual unit. The parallel Hindi English dictionary contains the tag of each English word. Depending upon the assigned tag, the system finds the proper.

Grammatical structure for the sentence and rearranges the words to construct a grammatically correct sentence. Various type of tags used during the translation are shown in table II. 4) Translation: All the segments whether words or partial simple sentences are translated individually. Words are translated based on its assigned POS tag referring to parallel Hindi-English dictionary [22]. If the Hindi word exist in the dictionary, the corresponding English word will be retrieved And tagged accordingly. On the other hand, if the word denoted as proper noun, it would be transformed to English by the Hindi-English transliteration. The words which represent many English words in translation, will be selected through learned SMT.

Example 7. Input:- ‘विकास ने’ , ‘विकास’ , ‘किया’
Translation:- ‘vikas’, ‘development’, ‘did’

5) Rearrangement: Rule Based Translation: In this phase we combine translated words and segments. Some rule based approaches are used to apply morphological modifications to the translated words i.e. add ‘e’, ‘ed’ or ‘ing’ after the verb, include “’s” with noun as apostrophe. Depending upon the assigned tags to the words, a matching grammar rule is selected for the given input. According to the selected sentence structure, required tense type of the sentence will be incorporated. And using that a proper linking verb i.e. am, is, are, was, were etc. are also required to be added to the sentence. Finally as per the matched grammar rule, words and other segments are rearranged and put in proper order. i.e some time “s” require after noun. We get final output after

3.2 WORKING OF PROJECT

Working Process:

- The first step in translation is to split the sentence into words or simple sentences.
- If a part of the input is in the example database then we keep that part as it is and the Remaining part segmented into words.
- When the segmentation is done, segments are actually translated and tagged independently.
- If the segment is example based, it is directly converted to English.
- Further, the remaining words are translated using the parallel Hindi-English Dictionary.

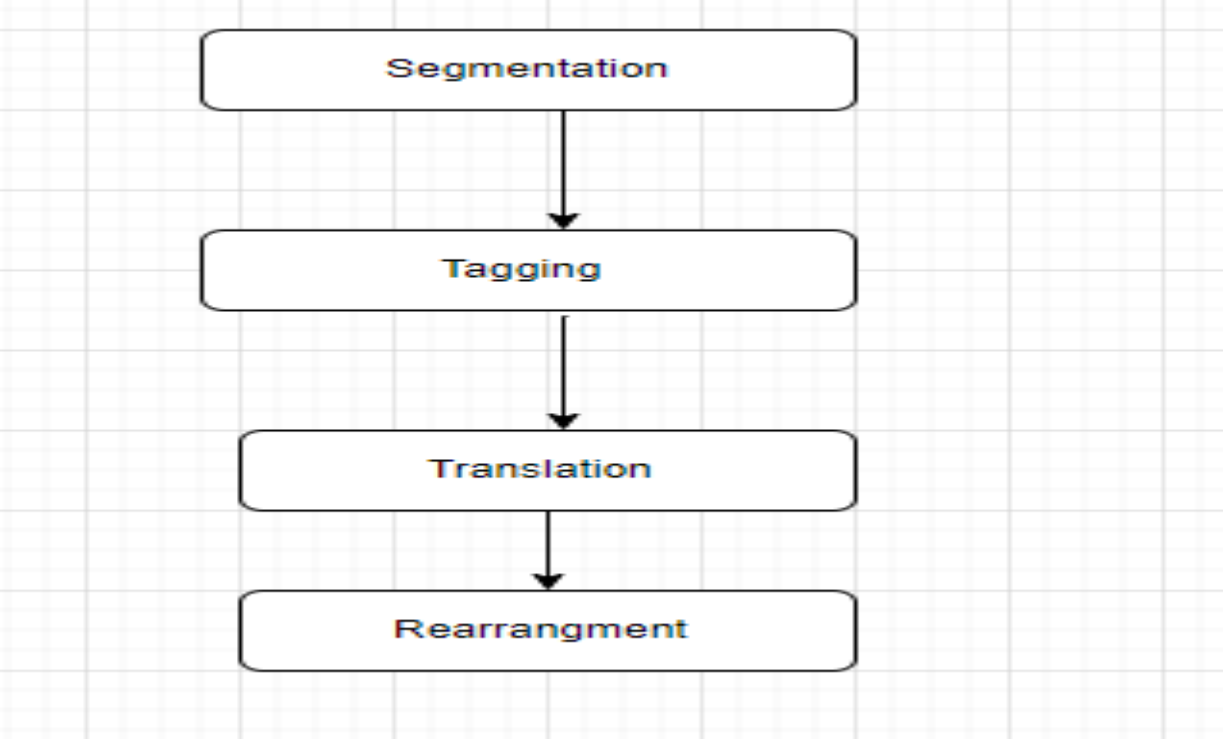

Figure 3

Cross language communication plays a pivotal role in building a favorable infrastructural environment for multifaceted benefits between two countries. In this internet era, machine translation fulfills the role of an agent to perform this cross language communication. Many countries have put forward enormous efforts for the development of several practical machine translation. Since last few decades, people have tried a significant number of approaches and resources to construct machine translation systems to be utilized in different applications ranging from simple textual translation to multilingual speech systems. Most of the machine translation systems follow the presumption that sentences will be grammatically correct and complete. Then such sentences are translated to target language with preserving the meaning in the source language. After Mandarin, Spanish and English, Hindi is the most natively spoken language in the world, almost spoken by 260 million people according to Ethnologies, 2014. Hence, there is vital requirement of many translators capable to translate sentences from Hindi language to other

Desired target language. We chose English language as target language in this paper. Recently, most of the MT works were focused on English to Indian language translation systems. However, a few systems have been constructed for Hindi to English translation, but not matured enough to resolve all inherent ambiguity and uncertainties of the Hindi sentences. On the basic level, a machine translator simply converts sentences by substituting word to word from source language to target language. But only the word substitution would not be able to deliver desired results as it doesn't care about semantic and syntactic constraints of the target language. There are many approaches developed to takeover these limitations of automated machine translation such as SMT, EBMT and RBMT. These approaches have their own strengths and weaknesses. There are already existing many freely available Hindi English machine translation systems like Google Translator, MS-Bing and Babylon. These systems are developed based on different approaches i.e. Rule Based Machine Translation (RBMT), Example Based Machine Translation (EBMT), Statistical Machine Translation (SMT). But they all not well accurate in handling the challenges of word sense disambiguation, pronoun resolution and idioms translation. RBMT systems perform the translation based on the rules discovered by linguists which tell how the words, words sequences or any other structure from source language would be transformed to target language. Other two systems EBMT and SMT extract the rules themselves automatically instead from the parallel corpora developed manually between source and target language. This is why they are referred to as data driven approaches. Latest approaches of machine translation are the combination of multiple approaches, a "Hybrid Machine Translation". This approach delivers better quality and functionality from traditional approaches. But the problem with HMT is that computationally it's more complex than the traditional approaches. A new way of implementing the hybrid approach for machine translation (HMT) has been discussed in this paper that utilize the strength of EBMT, RBMT and SMT. We have also presented the results of experiments performed with our proposed experimental HMT system.

CHAPTER – 4

4.1 RESULT

Few translation results have given in table IV. The input is given in the Devanagari script with UTF-8 encoding. First example is just a simple translation comprising of word to-word conversion followed by grammatical rearrangement. Second example signifies the characteristic that how the Proposed hybrid model capable of handling proper noun. Third example shows the robustness of translator to handle ambiguity in the sentence. Fourth one is the example of translating a sentence consists of phrases/idioms taken care of by EBMT part of the hybrid model. Last one is an example of a complex sentence translation.

HINDI	ENGLISH
भारत मेरा देश है	India is my country
राकेश खेल रहा है	Rakesh is playing
मैं विकास हूँ	I am vikas
नमस्ते	Hello

Table 3 (result)

Word Error Rate (WER) based metric is used here to find the accuracy of the proposed approach. Fundamentally, WER is computed based on Leven stein distance also known as the edit distance calculated through summing up the minimum no of insertions (I), deletions (D) and substitutions (S) applied to make a sequence similar to other. Consequently, accuracy of the MT system will be calculated through averaging the Sent acc on all the testing sentences.

$$W\ ER = S + D + I/N$$

$$Sent\ acc = 1 - W\ ER$$

The proposed system's result is compared with Google, Microsoft BING and Babylonian translators on a set of 500 manually translated Hindi-English sentences which is made up of 150 complex sentences, 200 simple sentences, 75 idiom based sentences and 75 sentences with ambiguity. Google translator is basically an SMT type of translator which has to be learned on a big corpus for better efficiency and robustness. Likewise, MS-Bing is fundamentally based on both SMT and RBMT approaches for the translation. Similarly, Babylonian uses SMT and morphological operations to perform the translation. All in all, statistical learning plays a major role for a most accurate machine translator.

4.2 DISCUSSION

A statistical comparison of the proposed HMT based approach with Google, BING and Babylonian translators is presented in table V based on different sentence types. Analysis reveals that for simple sentences all MT systems show nearly same accuracy ranging between 83-90%. All system performance very well for simple and unambiguous sentence. For Sentences includes idioms, proposed system gives more accurate output compare to other systems. The reason is the Use of Example Based methods that is included in the proposed HMT system. The exact meaning of idioms is different than the actual meaning of word reside in it. Hence, the idioms could not be translated directly. We need the actual meaning or replacement of idioms in prior. Sentences with ambiguity are handled by the SMT part of the proposed approach. Probability based statistical parameters, learned from the Hindi-English corpus, and are used to resolve the ambiguity. To resolve the conflict between the proper noun and the dictionary word, various rules have been defined based on the morphological properties of sentence structure in Hindi.

CHAPTER – 5

5.1 CONCLUSION

The entire project has been developed from the requirements to a complete system alongside evaluation and testing. The system developed have achieved its aim and objectives. The client was happy with the overall performance of the system. However, though some challenges were encountered during implementation, they were addressed and implemented. Also, future work and strategies on how to improve the system are further in this section.

5.2 FUTURE SCOPE

Since last several decades, people have developed a number of translation approaches to transform one language content to another ranging from simple word-to-word translation systems to corpus based statistical models Latest approaches of machine translation are the combination of multiple approaches, a “Hybrid Machine Translation”. This approach delivers better quality and functionality from traditional approaches. But the problem with HMT is that computationally its more complex than the traditional approaches. A new way of implementing the hybrid approach for machine translation (HMT) has been discussed in this paper that utilize the strength of EBMT, RBMT and SMT. We have also presented the results of experiments performed with our proposed experimental HMT system.

REFERENCE

[1] E. Alpaydin. Introduction to Machine Learning. Adaptive Computation and Machine Learning Series, MIT Press. 2014.

[2] Lotter, Hempel. Lernen, Lernschwierigkeiten – Diagnostik der Lernvoraussetzungen. Regierung Oberbayern. 2008.

[3] S. Chetlur. cuDNN: Efficient Primitives for Deep Learning. arXiv preprint arXiv:1410.0759c2. 2014

[4] T. Brants et. al. Large Language Models in Machine Translation. EMNLP-CoNLL. 2007.

[5] W. Ding, et. al. "Theano-based Large-Scale Visual Recognition with Multiple GPUs" ICLR. 2015.

[6] P. Flach. Machine Learning: The Art and Science of Algorithms that Make Senso of Data. Cambridge University Press. 2012.

[7] H. Fröning. GPU Computing Slides University of Heidelberg, ZITI. 2014.

[8] Y. Jia et. al. Caffe: Convolutional Architecture for Fast Feature Embedding. arXiv preprint arXiv:1408.5093v1, 2014.

[9] F. Ren and H. Shi, "Parallel machine translation: principles and practice," in Engineering of Complex Computer Systems, 2001. Proceedings .Seventh IEEE International Conference IEEE, 2001, pp. 249–259.

[10] O. Abiola, A. Adetunmbi, and A. Oguntimilehin, “A review of thevarious approaches for text to text machine translations,” International Journal of Computer Applications, vol. 120, no. 18, 2015.

[11] U. Tiwary and T. Siddiqui, Natural language processing and information retrieval. Oxford University Press, Inc., 2008.

[12] Ethnologue, “Most Widely Spoken Languages in the World,”http://www.infoplease.com/ipa/A0775272.html, 2014, [Online; accessed 24-June-2016].

[13] P. K. Goswami and S. K. Dwivedi, “An empirical study on English to hindi e-contents machine translation through multi engines,” in Reliability, Infocom Technologies and

Optimization (ICRITO)(Trends and Future Directions), 2014 3rd International Conference on. IEEE,2014, pp. 1–6.

[14] N. Sharma, “English to hindi statistical machine translation system,”Ph.D. dissertation, THAPAR UNIVERSITY PATIALA, 2011.

[15] P. F. Brown, J. Cocke, S. A. D. Pietra, V. J. D. Pietra, F. Jelinek, J. D. Lafferty, R. L. Mercer, and P. S. Roossin, “A statistical approach to machine translation,” Computational linguistics, vol. 16, no. 2, pp. 79– 85, 1990.

[16] H. Somers, “Review article: Example-based machine translation,” Machine Translation, vol. 14, no. 2, pp. 113–157, 1999.

[17] J. Centelles and M. R. Costa-Jussa, “Chinese-to-spanish rule-based machine translation system,” 2014.

[18] D. Groves and A. Way, “Hybrid data-driven models of machine translation,” Machine Translation, vol. 19, no. 3-4, pp. 301–323, 2005.

[19] B. Nithya and S. Joseph, “A hybrid approach to english to malayalam machine translation,” International Journal of Computer Applications,vol. 81, no. 8, 2013.

[20] P. Dungarwal, R. Chatterjee, A. Mishra, A. Kunchu kuttan, R. Shah, and P. Bhattacharyya, “The iit bombay hindi-english translation system at wmt 2014,” ACL 2014, p. 90, 2014.

[21] B. J. Dorr, “Machine translation divergences: A formal description and proposed solution,” Computational Linguistics, vol. 20, no. 4, pp. 597–
633, 1994.

[22] R. Sinha and A. Thakur, “Disambiguation of

Krizhe sky. One weird trick for parallelizing convolutional neural networks. Ar Xiv:1404.5997 [cs.NE]. 2014.

Y. Le Cun, et. al. Backpropagation applied to handwritten zip code recognition AT&T Bell Laboratories. 1989.

Y. Le Cun, et. al. Efficient Back Prop Neural Networks: Tricks of the trade, Springer. 1998.

K. P. Murphy. Machine Learning: A Probabilistic Perspective. MIT Press. 2012.

M. A. Nielsen, Neural Networks and Deep Learning. Determination Press. 2015.

NVidia. User Guide – cuDNN Library NVidia. DU-06702- 001_v6.5. 2014.

T. Paine et al.. Gpu asynchronous stochastic gradient descent to speed up neural network training. arXiv preprint arXiv:1312.6186. 2013.

O. Russakovsky el. al.. ImageNet Large Scale Visual Recognition Challenge. arXiv preprint arXiv:1409.0575. 2014.

S. Shalev-Shwartz, S. Ben-David. Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press. 2014.

Publication/ Screen Shots