Development of English-Hindi Interactive Machine Translation

Nidhi Sharma, Shambhavi Seth, Meenal Jain, Mehvish Syed, Nivedita Bharti

Abstract: ***Machine Translation systems are still far from being perfect and to improve their performance the concept of Interactive Machine Translation (IMT) was introduced. This paper proposes an IMT system, which uses Statistical Machine Translation and a bilingual corpus on which several algorithms (Word error rate, Position Independent Error Rate, Translation Error Rate, n-grams) are implemented to translate text from English to Indian languages. The proposed system improves both the speed and productivity of the human translators as found through experiments.***

Keywords: ***Machine Translation, Statistical Machine Translation, Computer Aided Translation, Interactive Machine Translation.***

I. INTRODUCTION

Since the very advent of languages, translations emerged as a very important aspect to minimize the communication gap. With time and computational capacity, a vast number of natural languages are being digitized. In India alone, we have such large multilingual diversity. It becomes a necessity to have translators for efficient communication and having human translators was tedious and both resource and time-consuming. This gave birth to Machine Translation Systems. In such a fast-paced world where time plays a significant role, MT requires certain enhancements which were introduced as CAT (Computer Aided Translation) tools. The current state-of-the-art MT systems do not produce ready- to-use translations; thus, human post-editing is required to achieve high-quality translations. To address to this issue, the MT systems are augmented with human translators to give birth to a new technology named Interactive Machine Translation. In this paper we have shown the development and the results of an English-Hindi Interactive Machine Translation (IMT) System. The rest of the paper is organized as: Section 2 describes the related work done in this area. Section 3 discusses our methodology where we have shown the development process of IMT. Section 4 evaluates the system and Section 5 concludes the same.

Revised Manuscript Received on October 30, 2019.

* Correspondence Author

Nidhi Sharma*, Department of Computer Science, Banasthali Vidyapith, Rajasthan, India. 304022.

Shambhavi Seth, Department of Computer Science, Banasthali Vidyapith, Rajasthan, India. 304022.

Mehvish Syed, Department of Computer Science, Banasthali Vidyapith, Rajasthan, India. 304022.

Meenal Jain, Department of Computer Science, Banasthali Vidyapith, Rajasthan. 304022.

Nivedita Bharti, Department of Computer Science, Banasthali Vidyapith, Rajasthan. 304022.

II. LITERATURE REVIEW

Nagao (1984) [1] proposed the idea of machine translation based on the examples, which were made available to the system prior to the translation process. A bilingual corpus was used which was comprised of source language text and their corresponding target language text. This approach was based on how humans process a sentence in source language and translate it to target language. Machine translation system use natural languages, which are highly complex (use of homonyms, different grammatical rules), to make decisions. In RBMT, defining rules is a difficult task and cost of building dictionaries and making changes is high. To overcome the problem of knowledge acquisition Corpus based MT comprising of SMT and EBMT was used. Languages with varying word orders proved to be a challenge for SMT. However, for production of dependency trees for sentence analysis and example database EBMT requires analysis and generation modules. Hybrid approach minimizes the challenges of other approaches (Okpor, 2014) [2]. Block transpositions (Leusch et al., 2003) [3] are used to extend edit distance to define inversion edit distance as a metric of the cost of parsing for a sentence pair within an inversion grammar. Experiments showed that at system level, correlation of automatic evaluation with human judgment is appropriate. DerivTool (DeNeefe et al., 2005) [4], an Interactive Translation Visualization Tool provides with a myriad number of options for the user to choose from and allows her to look into the decoding process where syntax-based framework for translation has been used. Moses (Koehn et al., 2007) [5], an open source toolkit for SMT featured the phrase-based translation with factors and confusion network decoding, which allowed translation of ambiguous input, along with the capabilities of Pharaoh decoder (Koehn et al., 2004) [6]. Caitra, developed by Koehn et al., (2009) [6] based on the TransType Project (Langlais et al., 2000) [7], a web-based IMT tool, with the help of Moses decoder, featured making suggestions for sentence completion, alternative word provision and phrase translation and giving post editing options with key stroke logging for detailed analysis. Cunei (Phillips et al., 2009) [8] proposed a hybrid system comprising of features of both EBMT and SMT, by using the example base to model each phrase pair at run time and perform recombination to get the translations. The system was tested using three language pairs Finnish to English, French to English and German to English against the Moses model, where Cunei displayed unexpectedly better results for German to English.

Due to German compounding, it becomes difficult for the one to many alignment and phrase extraction, but Cunei had advantage due to its runtime modeling as it adapts itself over time by updating weights. The complex lexicalized distortion model gave nearly the same result for the reordering model which was informed only about how far and frequently the phrases were moved during decoding. Joshi et al [9] developed a mechanism to write in Hindi using English. They used statistical machine learning to predict a word when some of the initial characters are typed. Using this Joshi et al. [10] also developed an Example Based Machine Translation System. Joshi et al. [11] also evaluated the system developed. They also compared the performance of this system with other popularly available MT engines. Gupta et al. [12] developed a rule-based stemmer for Urdu. They developed several rules to implement this stemmer. They further used this stemmer in evaluation of some English-Urdu MT systems [13]. Singh et al. [14] developed a POS tagger for Marathi using Statistical Machine Learning. Bhalla et al. [15] developed a procedure of transliteration of name entities from English to Punjabi. Joshi et al [16] evaluated several open domain MT engines. Gupta et al. [17] did the same for English-Urdu MT engines. Singh et al. [18] developed a POS tagger for Marathi using supervised learning. Joshi et al. [19] further developed a technique to using machine learning in evaluating MT engines. Tyagi et al. [20] [21] developed an approach of translating complex English sentences by first simplifying them and then translating into Hindi. Yogi et al. [22] developed an approach to identify candidate translation which are good for post editing. Gupta et al. [23] further extended their stemmer by adding derivational rules to the inflectional stemmer. Asopa et al. [24] developed mechanism for chunking Hindi sentences using a rule-based approach. Gupta et al. [25] developed a rule based lemmatizer for Urdu which was an extension to their stemmer. Kumar et al. [26] developed several machines learning-based classifiers for identifying different senses to a word in Hindi. Joshi et al. [27] developed a mechanism to estimate the quality of English-Hindi MT engines. Chopra et al. [28] [29] developed a name entity recognition and tagging tool for Hindi using several machine learning approaches. Gupta et al. [30] developed a POS tagger for Urdu using machine learning approach. Mathur et al. [31] developed an ontology matching evaluation using tool which used the MT engine developed by Joshi et al. Chopra et al. [32] developed a mechanism for rewriting English sentence and then translating them into Hindi. This significantly improved the performance of their MT engine. Joshi et al. [33] investigated some approaches to classifying documents and further suggested an approach for effective classification of text documents.

Singh et al. [34] developed an approach to automatically generate transfer grammar rules. This approach significantly improved the development process of their transfer-based MT engine. Singh et al. [35] developed an approach for text processing of Hindi documents using deep neural networks. They further developed this approach to mine textual data from web documents [36]. Singh et al. [37] developed a translation memory tool which worked as a sub-system in their transfer-based MT system. This further improved the accuracy of their system. Gupta et al. [38] further showed how fuzzy logic can be used in developing NLP applications. Gupta et al. [39] used several NLP tools in preprocessing the tweets that they extracted from web. They found that this approach improves the accuracy of their machine learning model which classifies the tweets. Gupta et al. [40] developed an approach which helped in identification and classification of multiword expressions from Urdu documents. Nathani et al. [41] developed a rule based inflectional stemmer for Sindhi which was written in Devanagari script. Asopa et al. [42] developed a shallow parser for Hindi using conditional random fields. Gupta et al. [43] showed the use of machine learning approached in developing NLP applications. Gupta et al. [44] used fuzzy operations in analyzing sentiments of tweets on several topics. This approach showed very promising results over traditional approaches. Sharma and Joshi [45] developed a rule-based word sense disambiguation approach for Hindi. It gave an accuracy of 73%. Katyayan and Joshi [46] studied various approaches of correct identification of sarcastic phrases in English documents. Gupta and Joshi [47] showed show tweets can be classified using NLP techniques. They showed how negative sentences can be handled using NLP approaches. Shree et al. [48] showed how there is difference between Hindi and English languages what problems the current state of the art MT system face while translating text. Ahmed et al. [49] showed how MT system can be developed by using an intermediate language which is related to both the languages. They developed a Arabic-Hindi MT system using Urdu as the intermediate language. They further performed the same study using English and found that if we have a large sized corpus then English which in unrelated to Arabic and Hindi, can be used for developed a MT system [50]. Seal and Joshi [51] developed a rule based inflectional stemmer for Assamese. This system showed very good results. Singh and Joshi [52] showed the developed of POS taggers for Hindi using different markov models. They concluded that hidden markov model-based tagger produced the best results among several markov based POS taggers. Pandey et al. [53] showed how NLP approached can help in develop a better ranking model for web documents. They used particle swam optimization and NLP approaches in improving the performance of their ranking model. Singh and Joshi [54] developed a rule-based approach for identifying anaphora in Hindi Discourses. Sinha et al. [55] developed a sentiment analyzer for Facebook post using the methods developed by Gupta et al.

Sharma et al. [56] [57] used some of the markov model-based approaches used by Singh et al. to develop their association classification model. Similar approaches we used by Goyal et al. [58] [59] for their models.

CASMACAT (Alabau et al., 2014) [60], a modular, web-based translation workbench with advanced functionality and editing features provided us with TM for raw match and automatic translations from MT server for post editing, consisting of a GUI, backend, an MT server and a CAT server.

The main features included Interactive Translation Prediction (ITP), confidence measures, prediction length control, search and replace, word alignment information, one clicks rejection, replay mode and logging function, and e-pen for handwritten interaction. The system was evaluated at Celer Soluciones, where 9 professors participated to carry out post editing for English-French translations. Three workbench features were tested, Post Editing, Intelligent Autocompletion (IA) and ITP where ITP had a slower rate due to the unfamiliarity with the IA system. Also due to the lack of visual aid to control IA, double checking had to be done even for small post editing tasks.

III. PROPOSED METHODOLOGY

To implement IMT, we integrated out translator with an existing machine translation system, which has been trained statistically and is based on Moses machine translation toolkit. In order to train this statistical system, we used a mix-domain corpus which comprises of translations from tourism, health, agriculture and administrative domains. We preferred mix-domain training to achieve open domain translations. Another important factor we incorporated in our system is a bilingual corpus on which our translator is trained. The chunks in the corpus are taken from the tourism domain. When source text is provided to the system, it first looks for the sentences in the corpus, evaluates the sentences one at time, to check for their similarity with the input text provided. In case, the similarity score is greater than the set threshold value, the target text corresponding to that source text as well as the output from SMT system is joined and the first order output is proposed to the user. As we claim our system to be interactive, we allow the user to make changes as required in the target text of most suitable candidate target text and further, add this new source-target pair as a new tuple in the corpus so that the forthcoming translations can make a use of it. IMT calculates the similarity score of hypotheses (translation produced by system) and reference sentence (example translation available in knowledge base) using various algorithms. These algorithms are named as Word Error Rate, Position Independent Error Rate, Translation Error Rate, n-grams similarity. Our IMT system made use of these algorithms to get the example translations and ranked them according to similarity. It also took the translation of an MT system and add to the list of possible translations. These all are provided to the human translation who can select the one which the human translator feels most appropriate and perform post editing to generated a final high-quality translation. The working of this entire system is shown in figure 1.

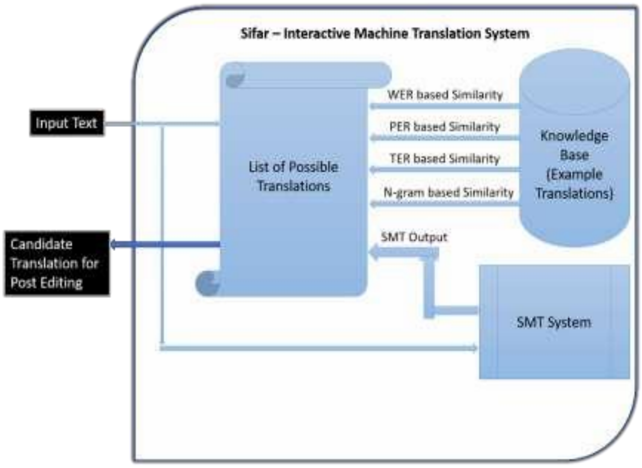

Figure 1: Working of IMT System

IV. EVALUATION

We performed our evaluation on 2 human translators who were asked to translate 1000 sentences using Sifar. For the first 500 sentences, we did not provide them with any suggestive translations and where are to do translations on their own. For the rest 500 sentences, we provided the translators with suggestions. The analyzed the results based on speed and productivity. The results are as follows:

A. Speed

All the five translators performed very well, when they were provided with the suggestive list of possible translations. This improved their speed to doing translations. Table 1 shows the results of this experiment. In this, we calculated the speed as the total time taken to complete the task divided by total no. of words in the documents. This is shown in equation (1). Here the calculated speed of doing translations when no suggestions were provided and when the suggestions were provided.

$$Speed = \frac{Total\ Time\ Taken}{Total\ No.of\ Words} \tag{1}$$

Table 1: Time Taken to Complete the Translations

Human Translations	Without Suggestive Translators	With Suggestive Translations	Difference
H1	6.3 sec/word	3.2 sec/word	3.1 sec/word
H2	5.3 sec/word	3.4 sec/word	1.9 sec/word

B. Productivity

In order to ascertain productivity, we calculated the score of HTER (Human Translation Edit Rate). In this, we calculated the total number of edits (inserts, deletes, substitutes, sifts) performed by human translator to correct the translation which was selected from the list of suggestive translations. This was done on only last 500 sentences, as these were the sentences on which suggestive translations were provided. In all the cases, the edit rate of the translation was very less. This confirms that the use of IMT improves the productivity of the human translators. Table 2 shows the results of this study.

Table 2: HTER Score

Human Translators	HTER Score
H1	0.3452
H2	0.1542

V. CONCLUSION

In this paper, we have shown the design of an Interactive Machine Translation system which combines the strengths of both human and machine and provides a high-quality machine assisted translation. Through experiments we have verified our claim, as with the help of our system the speed of the human translators improved. It also improved the productivity of the translators. As an extension to this study, we wish to improve the user friendliness of the system by providing several features like providing a human translator an option to add their example translations in the knowledge base, before doing actual translations. We also wish to perform a more thorough evaluation to understand the expectations of the human translators from this system. One of the possible expectations is to reduce the key in effort. We shall perform the usability evaluation of the system to analyze such expectations and incorporate the same in subsequent versions of our IMT system.

REFERENCES

1. Alabau, V., Buck, C., Carl, M., Casacuberta, F., García-Martínez, M., Germann, U., ... & Mesa-Lao, B. (2014, April). Casmacat: A computer-assisted translation workbench. In Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics (pp. 25-28).
2. DeNeefe, S., Knight, K., & Chan, H. H. (2005, June). Interactively exploring a machine translation model. In Proceedings of the ACL 2005 on Interactive poster and demonstration sessions (pp. 97-100). Association for Computational Linguistics.
3. Dungarwal, P., Chatterjee, R., Mishra, A., Kunchukuttan, A., Shah, R., & Bhattacharyya, P. (2014, June). The iit bombay hindi-english translation system at wmt 2014. In Proceedings of the Ninth Workshop on Statistical Machine Translation (pp. 90-96).
4. Helcl, J., Libovický, J., & Variš, D. (2018). CUNI system for the WMT18 multimodal translation task. arXiv preprint arXiv:1811.04697.
5. Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., ... & Dyer, C. (2007, June). Moses: Open source toolkit for statistical machine translation. In Proceedings of the 45th annual meeting of the association for computational linguistics companion volume proceedings of the demo and poster sessions (pp. 177-180).
6. Koehn, P. (2004, September). Pharaoh: a beam search decoder for phrase-based statistical machine translation models. In Conference of the Association for Machine Translation in the Americas (pp. 115-124). Springer, Berlin, Heidelberg.
7. Langlais, P., Foster, G., & Lapalme, G. (2000). TransType: a computer-aided translation typing system. In ANLP-NAACL 2000 Workshop: Embedded Machine Translation Systems.
8. Joshi, N., Mathur, I. and Mathur, S., 2010. Frequency based predictive input system for Hindi. In Proceedings of the International Conference and Workshop on Emerging Trends in Technology pp. 690-693. ACM.
9. Joshi, N., Mathur, I. and Mathur, S., 2011. Translation memory for Indian languages: an aid for human translators. In Proceedings of the International Conference & Workshop on Emerging Trends in Technology pp. 711-714. ACM.
10. Joshi, N., Darbari, H. and Mathur, I., 2012. Human and automatic evaluation of english to hindi machine translation systems. In Advances in Computer Science, Engineering & Applications pp. 423-432. Springer, Berlin, Heidelberg
11. Gupta, V., Joshi, N. and Mathur, I., 2013. Rule based stemmer in Urdu. In 2013 4th International conference on computer and communication technology (ICCCT) pp. 129-132. IEEE.
12. Gupta, V., Joshi, N. and Mathur, I., 2013. Subjective and objective evaluation of English to Urdu Machine translation. In 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI) pp. 1520-1525. IEEE.
13. Singh, J., Joshi, N. and Mathur, I., 2013. Development of Marathi part of speech tagger using statistical approach. In 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI) pp. 1554-1559. IEEE.
14. Bhalla, D., Joshi, N. and Mathur, I., 2013. Improving the quality of MT output using novel name entity translation scheme. In 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI) pp. 1548-1553. IEEE.
15. Joshi, N., Mathur, I., Darbar, H., Kumar, A. and Jain, P., 2014. Evaluation of some English-Hindi MT systems. In 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI) pp. 1751-1758. IEEE.
16. Gupta, V., Joshi, N. and Mathur, I., 2014. Evaluation of English-to-Urdu Machine Translation. In Intelligent Computing, Networking, and Informatics pp. 351-358. Springer, New Delhi.
17. Singh, J., Joshi, N. and Mathur, I., 2014. Marathi Parts-of-Speech Tagger Using Supervised Learning. In Intelligent Computing, Networking, and Informatics pp. 251-257. Springer, New Delhi.
18. Joshi, N., Mathur, I., Darbari, H. and Kumar, A., 2015. Incorporating Machine Learning Techniques in MT Evaluation. In Advances in Intelligent Informatics pp. 205-214. Springer, Cham.
19. Tyagi, S., Chopra, D., Mathur, I. and Joshi, N., 2015. Comparison of classifier-based approach with baseline approach for English-Hindi text simplification. In International Conference on Computing, Communication & Automation pp. 290-293. IEEE.
20. Tyagi, S., Chopra, D., Mathur, I. and Joshi, N., 2015. Classifier based text simplification for improved machine translation. In 2015 International Conference on Advances in Computer Engineering and Applications pp. 46-50. IEEE.
21. Yogi, K.K., Joshi, N. and Jha, C.K., 2015. Quality Estimation of MT-Engine Output Using Language Models for Post-editing and Their Comparative Study. In Information Systems Design and Intelligent Applications (pp. 507-514). Springer, New Delhi.
22. Gupta, V., Joshi, N. and Mathur, I., 2015. Design & development of rule based inflectional and derivational Urdu stemmer 'Usal'. In 2015 International conference on futuristic trends on computational analysis and knowledge management (ABLAZE) pp. 7-12. IEEE.
23. Asopa, S., Asopa, P., Mathur, I. and Joshi, N., 2016. Rule based chunker for Hindi. In 2016 2nd international conference on contemporary computing and informatics (IC3I) pp. 442-445. IEEE.
24. Gupta, V., Joshi, N. and Mathur, I., 2016. Design and development of a rule-based Urdu lemmatizer. In Proceedings of International Conference on ICT for Sustainable Development pp. 161-169. Springer, Singapore.
25. Kumar, A., Mathur, I., Darbari, H., Purohit, G.N. and Joshi, N., 2016. Implications of Supervised Learning on Word Sense Disambiguation for Hindi. In Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies. p 54. ACM.
26. Joshi, N., Mathur, I., Darbari, H. and Kumar, A., 2016. Quality Estimation of English-Hindi Machine Translation Systems. In Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies. p 53. ACM.
27. Chopra, D., Joshi, N. and Mathur, I., 2016. Named Entity Recognition in Hindi Using Conditional Random Fields. In Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies p. 106. ACM.
28. Chopra, D., Joshi, N. and Mathur, I., 2016. Named Entity Recognition in Hindi Using Hidden Markov Model. In 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT) pp. 581-586. IEEE.
29. Gupta, V., Joshi, N. and Mathur, I., 2016. POS tagger for Urdu using Stochastic approaches. In Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies p. 56. ACM.
30. Mathur, I., Joshi, N., Darbari, H. and Kumar, A., 2016. Automatic Evaluation of Ontology Matchers. In Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies. p 56. ACM.

Retrieval Number: A1532109119/2019©BEIESP
DOI: 10.35940/ijeat.A1532.109119
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

31. Chopra, D., Joshi, N. and Mathur, I., 2016. Improving Quality of Machine Translation Using Text Rewriting. In 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT) pp. 22-27. IEEE.
32. Joshi, M.L., Mittal, N. and Joshi, N., 2017. An Insight into Role of Wordnet and Language Network for effective IR from Hindi Text Documents. In FIRE (Working Notes) pp. 158-163.
33. Singh, S.P., Kumar, A., Darbari, H., Singh, L., Joshi, N., Gupta, P. and Singh, S., 2017. Intelligent System for Automatic Transfer Grammar Creation Using Parallel Corpus. In International Conference on Information and Communication Technology for Intelligent Systems pp. 519-528. Springer, Cham.
34. Singh, S.P., Kumar, A., Darbari, H., Rastogi, A., Jain, S. and Joshi, N., 2017. Building Machine Learning System with Deep Neural Network for Text Processing. In International Conference on Information and Communication Technology for Intelligent Systems pp. 497-504. Springer, Cham.
35. Singh, S.P., Kumar, A., Darbari, H., Kaur, B., Tiwari, K. and Joshi, N., 2017. Intelligent Text Mining Model for English Language Using Deep Neural Network. In International Conference on Information and Communication Technology for Intelligent Systems pp. 473-486. Springer, Cham.
36. Singh, S.P., Kumar, A., Darbari, H., Tailor, N., Rathi, S. and Joshi, N., 2017. Intelligent English to Hindi Language Model Using Translation Memory. In International Conference on Information and Communication Technology for Intelligent Systems pp. 487-496. Springer, Cham.
37. Gupta, C., Jain, A. and Joshi, N., 2018. Fuzzy Logic in Natural Language Processing–A Closer View. Procedia computer science, 132, pp.1375-1384.
38. Gupta, I. and Joshi, N., 2017. Tweet normalization: A knowledge-based approach. In 2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS) pp. 157-162. IEEE.
39. Gupta, V., Joshi, N. and Mathur, I., 2017. Approach for multiword expression recognition & annotation in urdu corpora. In 2017 Fourth International Conference on Image Information Processing (ICIIP) pp. 1-6. IEEE.
40. Nathani, B., Joshi, N. and Purohit, G.N., 2018. A Rule Based Light Weight Inflectional Stemmer for Sindhi Devanagari Using Affix Stripping Approach. In 2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE) pp. 1-4. IEEE.
41. Asopa, S., Asopa, P., Mathur, I. and Joshi, N., 2019. A Shallow Parsing Model for Hindi Using Conditional Random Field. In International Conference on Innovative Computing and Communications pp. 295-302. Springer, Singapore.
42. Gupta, V., Joshi, N. and Mathur, I., 2019. Advanced Machine Learning Techniques in Natural Language Processing for Indian Languages. In Smart Techniques for a Smarter Planet pp. 117-144. Springer, Cham.
43. Gupta, C., Jain, A. and Joshi, N., 2019. A Novel Approach to feature hierarchy in Aspect Based Sentiment Analysis using OWA operator. In Proceedings of 2nd International Conference on Communication, Computing and Networking pp. 661-667. Springer, Singapore.
44. Sharma P., Joshi N., 2019. Design and development of a knowledge-based approach for word sense disambiguation by using wordnet for Hindi. International Journal of Innovative Technology and Exploring Engineering, Vol 8, pp 73-78.
45. Katyayan, P. and Joshi, N., 2019. Sarcasm Detection Approaches for English Language. In Smart Techniques for a Smarter Planet pp. 167-183. Springer, Cham.
46. Gupta I., Joshi N., 2019. Enhanced Twitter Sentiment Analysis Using Hybrid Approach and by Accounting Local Contextual Semantic. Journal of Intelligent Systems. Walter de Gruyter.
47. Shree, V., Mathur, I., Yadav, G., Joshi, N., 2019. Digital humanities: Can machine translation replace human translation. International Journal of Recent Technology and Engineering. Vol 7, pp 128-134.
48. Ahmed, S.A., Joshi, N., Mathur, I., Katyayan, P., 2019. Impact of related languages as pivot language on machine translation. International Journal of Recent Technology and Engineering. Vol 7, pp 1539-1546.
49. Ahmed, S.A., Joshi, N., Mathur, I., Katyayan, P., 2019. Implications of english as a pivot language in arabic-hindi machine translation. International Journal of Engineering and Advanced Technology. Vol 8, pp 271-277.
50. Seal, S., Joshi, N., 2019. Design of an inflectional rule-based assamese stemmer. International Journal of Innovative Technology and Exploring Engineering. Vol 8, pp 1651-1655.
51. Singh, A., Joshi, N., 2019. Part of speech tagging of Hindi using Markov model. International Journal of Innovative Technology and Exploring Engineering. Vol 8, pp 1723-1726.
52. Pandey, S., Mathur, I. and Joshi, N., 2019. Information Retrieval Ranking Using Machine Learning Techniques. In 2019 Amity International Conference on Artificial Intelligence (AICAI) pp. 86-92. IEEE.
53. Singh, V., Joshi, N., 2019. A rule based approach for anaphora resolution. International Journal of Innovative Technology and Exploring Engineering. Vol 8, pp 2652-2657.
54. Sinha, S., Saxena, K., Joshi, N., 2019. Sentiment analysis of facebook posts using hybrid method. International Journal of Recent Technology and Engineering. Vol 8, pp 2421-2428.
55. Kumar, S. and Joshi, N., 2016. Rule power factor: a new interest measure in associative classification. Procedia Computer Science, 93, pp.12-18.
56. Sharma, O., Kumar, S. and Joshi, N., 2019. SRPF Interest Measure Based Classification to Extract Important Patterns. In Proceedings of 2nd International Conference on Communication, Computing and Networking pp. 523-530. Springer, Singapore.
57. Goyal, H., Sharma, C. and Joshi, N., 2017, August. Estimation of Monthly Rainfall using Machine Learning Approaches. In 2017 International Conference on Innovations in Control, Communication and Information Systems (ICICCI) pp. 1-6. IEEE.
58. Goyal, H., Joshi, N. and Sharma, C., 2019. Feature Extraction in Geospatio-temporal Satellite Data for Vegetation Monitoring. In Emerging Trends in Expert Applications and Security pp. 177-187. Springer, Singapore.
59. Nagao, M.: A framework of a mechanical translation between Japanese and English by analogy principle. Artificial and human intelligence, 351-354 (1984).