TEST Engineering & Management

May – June 2020
ISSN: 0193-4120 Page No. 94 - 101

Impact of Discourse in Modern AI based NLP Applications

[1] **Mitali Halde, [2]Dr. Smitha Rao M Suresh,**
[1] Department of Computer Science Engineering, School of Engineering, Presidency University Bengaluru, Karnataka, India. Mail: mitalihalder@presidencyuniversity.in
[2]Department of Computer Science Engineering, School of Engineering, Presidency University Bengaluru, Karnataka, India. Mail: smitharao@presidencyuniversity.in

Article Info
Volume 83
Page Number: 94 - 101
Publication Issue:
May - June 2020

Article History
***Article Received**: 11August 2019*
***Revised**: 18November 2019*
***Accepted**: 23January 2020*
***Publication**:07May2020*

Abstract:
While Machine Translation (MT) has traditionally considered isolated sentence for word-to-wordtranslation, to improve the translation further, a multi-gram language model felicitated through phrase-table with inter-linguae mappings to improve the accuracy of translations. But this neglects the context essence of collocated and coherent sentences. Thus, the incorporation of Discourse in Machine Translation plays a major role to improve translation. These incorporations of Discourse in MT have significant State-of-the-art application such as Text Extraction, Emotion extraction, Sentiment analysis etc. In this paper, we have an overview of Discourse and Discourse Relations. Also conducted an exhaustive research on the impact of discourse key application of Natural Language Processing (NLP).

Keywords*—Discourse, Discourse Relation, Machine Translation, Application of Discourse, coherent sentences.*

I. INTRODUCTION

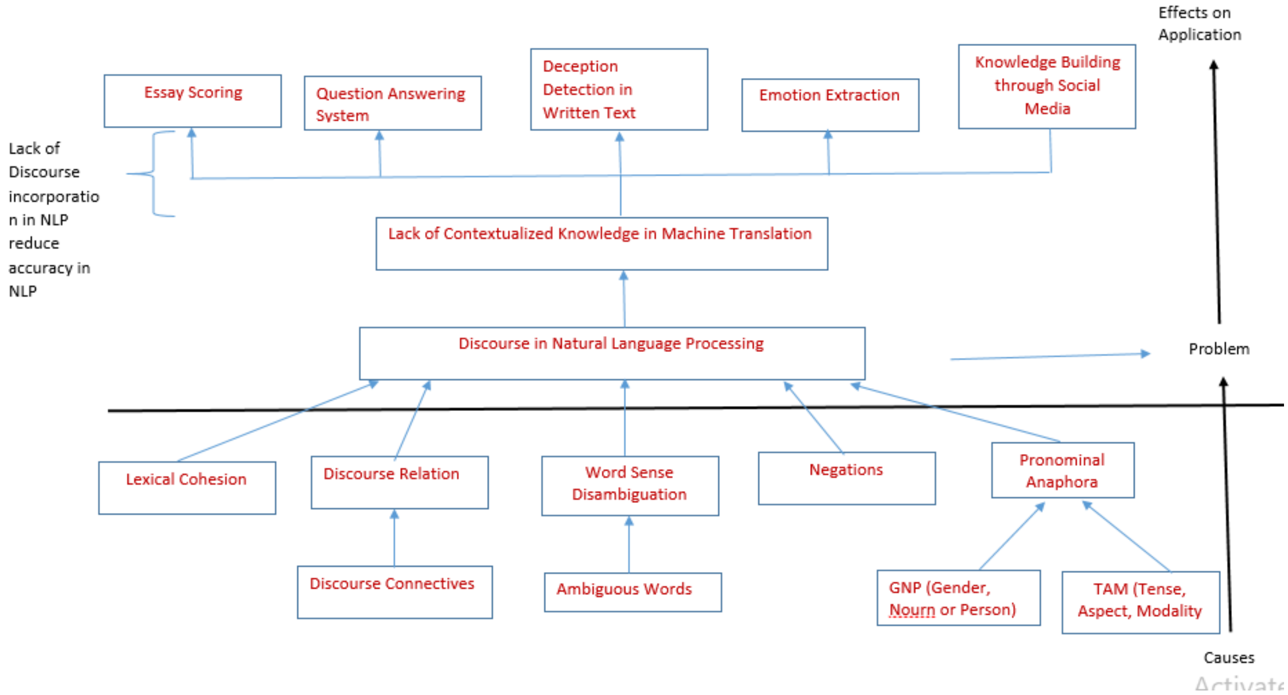
Human beings are accumulator of knowledge and knowledge is stored and organized in natural language. To automate the understanding and generation of natural language is an enduring objective of Artificial Intelligence. Well Formed representation of language was a hard problem for the computer science researcher. Initially, in the Rule Based approach (RBMT), a manual handwritten rule for almost all conditions of language was maintained and with the help of debugging tools newly arrived conditions were updated for translation. After RBMT, Example Based Machine Translation introduced to overcome the exhaustive work of RBMT. MakatoNagavoet. al. (1981), EBMT have three phases: matching, alignment and recombination. Peter Brown et. al. (1988) Later, stochastic technique have been implemented for Machine Translation and this approach of Machine Learning for Machine Translation in Natural Language Processing emerged as the incredible success , with over half of the translation acceptable. This Statistical Based Translation approach is more on statistical analysis of language for translation and highly ignore the semantic translation value. Also a lot of work requires for debugging like TAM (Tense, Aspect and Modality), GNP (Gender Noun or Person) Debugging, WSD (Word Sense Disambiguation) Debugging and Verb Agreement Debugging for the updating of the system. These lack of semantic knowledge while doing the translation becomes and obstacle in the NLP applications like sentiment analysis, text summarization, question-answering etc. Thus, the Natural Language processing required a big leap for incorporating Discourse {ie the study of analyzing relationship among two sentences or

TEST Engineering & Management

May – June 2020
ISSN: 0193-4120 Page No. 94 - 101

phrases] in Machine Translation Discourse is a group of collocated and coherent sentences. Martin Jurafskyet. al. 2009. Discourse analysis/processing is the task of natural Language Processing (NLP) to evaluate the semantics from multi-sentential text at several levels which may support "downstream" language processing application. With the advancement of technology the limitation of Machine Translation i.e. either lack of semantics knowledge availability or lack of resources like lack of bilingual corpora or handling translation from High morphological language to low morphological language has been working to overcome using Deep Neural Network [4]. Machine Translation using Deep Neural Network (NMT) emerged as the standard benchmark approach for Natural Language processing   Also Neural Machine Translation approach emerges as an efficient techniques to implement discourse analysis in Machine translation.

With the increment in the computer mediated data either structured or unstructured text, a wide range of probabilistic techniques had been used to accomplish the automated learning with the emerging computational power. The idea is to show the significance of discourse incorporation in Machine Translation which in turn improved the translation quality by considering the contextual essence.A pictorial representation of the overall research is given in Fig.1, which summarizes the causes for the problem why it is difficult to incorporate Discourse in Machine Translation and also shows how the lack of Discourse knowledge reduces the accuracy of the NLP applications.



The intention of this paper is to incorporate discourse in Machine Translation will successfully lead to examine the flow of conversation, eventually lead to determine the real essence of context and also determine the social structure existing within the group participated in the conversation. Furthermore, with the advancement of technology, there is a technological shift from traditionally Statistical Machine Translation (SMT) approach to Neural Machine Translation (NMT). In the same way, most computational model for Discourse analysis also has the paradigm shift from traditional Statistical Models to Deep Neural Models.Discourse Representation Theory have been discusses in section 2; followed by a comprehend Discourse Relation and then in Section 3 Application of Discourse have been covered.

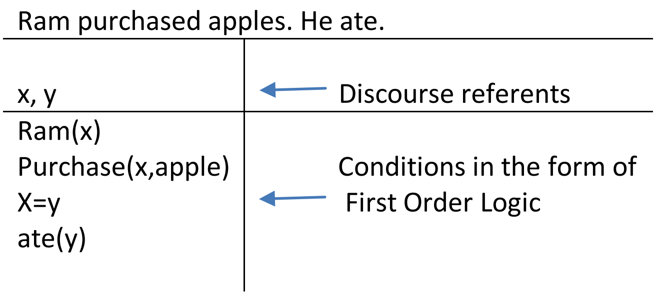
TEST Engineering & Management
May – June 2020
ISSN: 0193-4120 Page No. 94 - 101

II. Discourse Representation And Analysis

Discourse is a group of collocated and coherent sentence. Discourse Theory deals with the language phenomenon that operate beyond the single sentences. This section covers the conventional representation of discourse using First Order Logic and then the summary of Discourse relations.

A. Discourse Representation Theory (DRT)

Kamp et. al. (1981), DRT was developed to unfold and apprehend the semantic of underlying text. If formally require 2 components: (i) a formal definition of language representation consist of: (a) a recursive definition of the set of all well-defined Discourse Representation Structure (DRS). (b) a model theoretic semantics for the members of this set. (ii) a construction procedure specifying how a DRS is to be extended when new information becomes available. Sample example for Anaphora Resolution using First order logic:



B. Coherence Strucutre

Like Grammaticality is the relation and grouping between words to form a well structures sentence (Fig 2), the same way, coherent model is the grouping and relation between sentences.

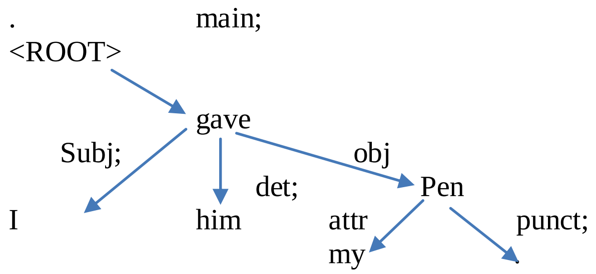


Fig 2: Groupings of Words in Grammaticality

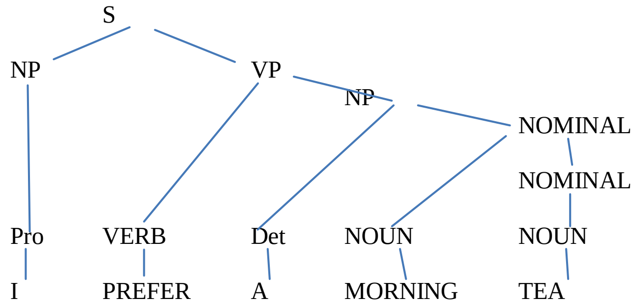


Fig 3: Relation between Words in Grammatically

C. Discourse Relation

Most of the documents have Discourse structure: a logical floe of events, states and propositions that make a coherent idea, argument or story. Discourse Relation specify the relation between sentences or clauses. Thus the two adjacent sentences looks coherent. In order to capture the semantic structure of the cohesive words and coherent sentences, the Discourse Relation analysis approach was introduced. These prior annotation marking approach also captures the information regarding attributes of discourse for meaning. In the following table various coherence relation have been summarized with key examples:

S.No.	Relation	Keywords	Examples
1.	Elaboration	And, also, then, moreover	Jack bought a gift for his wife Nancy. He bought Diamond ring.
2.	Contrast	Although, but, while, however	The man has much money. However, he isn't happy at all
3.	Cause-Effect	Because, and, so, thus, therefore	The Tin Woodman was caught in rain. His joints rusted.
4.	Equivalence	In-fact, alternativ ely, similarly	It is better to grind fresh bark, if you can. Alternatively, you can use cinnamon powder.
5.	Change Topic	By the way, incidentall y, mean-while,	Ramya spent four years studying for her law degree. Mean-while, she continued to work at

TEST Engineering & Management

May – June 2020
ISSN: 0193-4120 Page No. 94 - 101

		And now	the bank.
6.	Example	For Example, For Instance	With recent advent of technology so many application in Discourse emerges. For example, emotion extraction, sentiment analysis and so on.
7.	Narration	Next, Finally, Obviously	I got up to collect due forms and meet the coordinators. Next complete all the signatures and formalities. Finally submitted the forms to Dean office.
8.	Explanation	As, because , since	Sterlin hid Jeevan's car keys. He was drunk.
9.	Continuation	And, but	I ran home. But I was still late.
10.	Condition	If...(then), as long as , while	Madhu works consistent during semester. She scored 98%.
11.	Generalization	In general	Most of the students had scored well in Part A of End Term question paper. In general, we can say, the Part A is the scoring section.
12.	Temporal Sequence	(And)... the, first, second, before, after	First wash your hand then had your lunch.
13.	Parallel	But, so, instead of	There was no flight scheduled to Bhopal yesterday. There was several to Indore
14.	Non-Parallel	But, so , instead of	There were too few flights flying to Bhopal yesterday. We went to Indore.

There were many different ways to express types of discourse relationship. The above table broadly summarized discourse relationship. While translation the text from source language to destination language. If the discourse relationships are annotated properly, the translated result will have a semantic essence of source test in the target text. This incorporation will efficiently improves the accuracy of the Natural Language based application.

III. Discourse Application

As already discussed in above sections, Discourse analysis is the study of languages in either spoken or written. Text pattern, coherence sentences, cohesive words and clause relations are all part of written discourse. With the increment of Computer mediated data many innovative application emerged which have a deep impact of discourse analysis. In the discussed application if the Discourse Relations are annotated in the source text analysis level, the accuracy of the application will increase significantly. Following are the prominent application of discourse analysis in computer mediated:

A. Essay Scoring

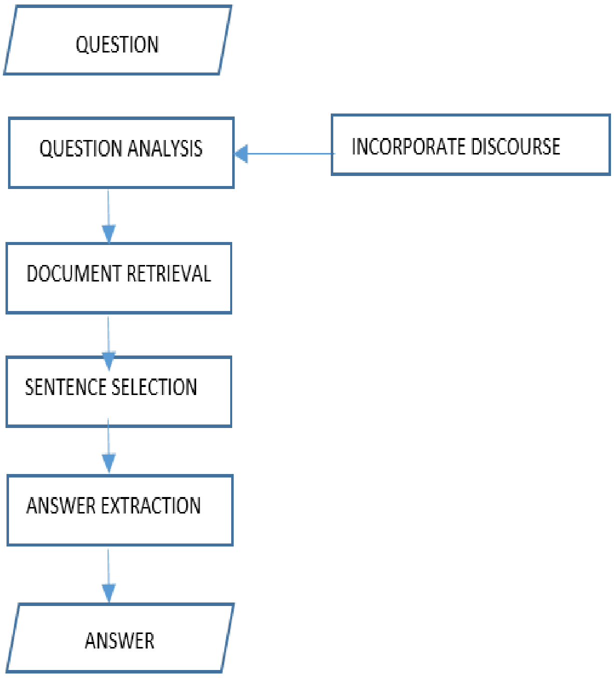
To grade manually Essay is time consuming and expensive task. [17] The automated grading system will reduce the cost significantly. [18]The methodology used is Linear Regression to learn the parameters based on the features. Some of the main feature extracted are Bag of Words, Numerical features like sentence count, length of word, word count, Part-of-Speech, Orthography and Structure and Organization. But in the entire process, extraction of features to analyze semantic knowledge per essay is lacking. Thus the performance of learning system will significantly vary from Human Evaluation. Thus if the connective relations of the coherent sentences and cohesive words annotated at the time of analyzing the essay text, the accuracy of grading will enhance considerably. Thus, discourse incorporation will improve the electronic essay evaluator's performance in comparison to the Human evaluator which led to saving a significant amount of time and cost.

B. Question-Answering System

The question answering system addressed to User's Information need. Chatbots are the typical example of question answering system. Chatbots are hugely

May – June 2020
ISSN: 0193-4120 Page No. 94 - 101

used as the Customer Care Services for worldwide many application which saves time and reduced significant cost of hiring human resources for the repetitive jobs. If the user asked for the summarization for a particular task, this automated task of Text Summarization is also consider as an application of Question-Answering. Following is the



Typical Question Answering pipeline Architecture:

Fig 4: Question Answering Pipeline with Incorporating Discourse at Question Analysis Level

Tree Recurrent Neural Network used to reproduce logical inference behavior from reasonably sized training sets. But as the complexity and size of expressions increase, then in a steady falling in performance. Thus, discourse knowledge need to incorporate while analyzing the question which lead to relevant selection of sentences for answers.

C. Deception Detection in Written Text

Fake news is any news that either factually wrong, misrepresentation of the facts; alternatively can be known as Junk news, Pseudo News, Hoax News. [7] 2019 et.al. Zhang It spreads virally on to targeted audience through social media or conventional News Channel. Fake news detection considered as the classic example of Text Classification problem. Conventionally, for fake news detection model, n-gram words and shallow Parts of Speech (POS) tagging model was used but it had proven insufficient. [16] Deep syntax analysis using Probabilistic Context Free grammar improves accuracy still context level interpretation is lacking. This leads to the discourse incorporation while doing text Classification. Thus, it has been observed here that prior discourse relations annotated text improves the application accuracy.

D. Emotional Extraction

In this application, the task is to extract potential reasons for certain emotions in written text is evaluated. Emotion Cause Extraction (ECE) [23], word level sequence labelling problem have the short coming that before the cause is extracted, there should be a prior marking of the emotion should be done. [6] Emotion Cause Pair Extraction (ECPE) is a 2 steps approach which proved to be a benchmark for Emotion Extraction. The shortcoming for this approach is if error happens in first step then it will influence the analysis of next step.. [22] Also for the real world application where enormous amount of unstructured data is available; it's difficult even to evaluate the polarity of the sentences. Thus the introduction of discourse particles in Emotion Extraction plays a significant role to overcome these short comings. Discourse relations like connectives, conditions, contrast, elaboration etc. can be incorporated in existing model to improve the accuracy of Emotion-Cause Extraction Pair.

E. Knowledge Building Trhough Social Netwrok

Social Media data is different from traditional documents such as Newspaper, Articles, Books and Research Publications. These data come from authors who are not professional writers; data can be found in many formats, languages and styles. [24] Many task for evaluation of new datasets collected from Social Media are organized by Association of Computational Linguistic (such as SemEval Task) or by the National Institute of Standards and Technology via Text Retrieval Conference (TREC) and the Text Analysis Conference. While building the relevant knowledge data base from the Social Media, the NLP approaches and algorithms need to analyze text which not only deal with words, grammar and syntax but also semantics and meanings. To analyze the meaning we need to incorporate discourse relations while extracting data for knowledge building.

IV. Conclusion and Future Work

Discourse is a group of collocated and coherent sentences (i.e. it deals with the relation between the sentences for the given text). In this paper, we have done an exhaustive research to present the summarized table of the popular discourse connectives with key example. Also discussed the significance of Discourse incorporation to Machine Translation irrespective of the underlying technology used. With the advent of the technology, it is observed that there is a rapid increment of application where there is direct impact of discourse relations. Discourse analysis is the State-of-the-art problem in Association for Computational Linguistics (ACL) community, where automatically model language tools are develop that go beyond individual sentence interpretation. These models can be used further solve the program of essay scoring, text summarization (sentence selection and ordering) and so on. In fig. 4, it is shown that to integrate discourse in Machine Translation, the discourse problem can be reduced to various categories, few of which are showed in fig. 5.

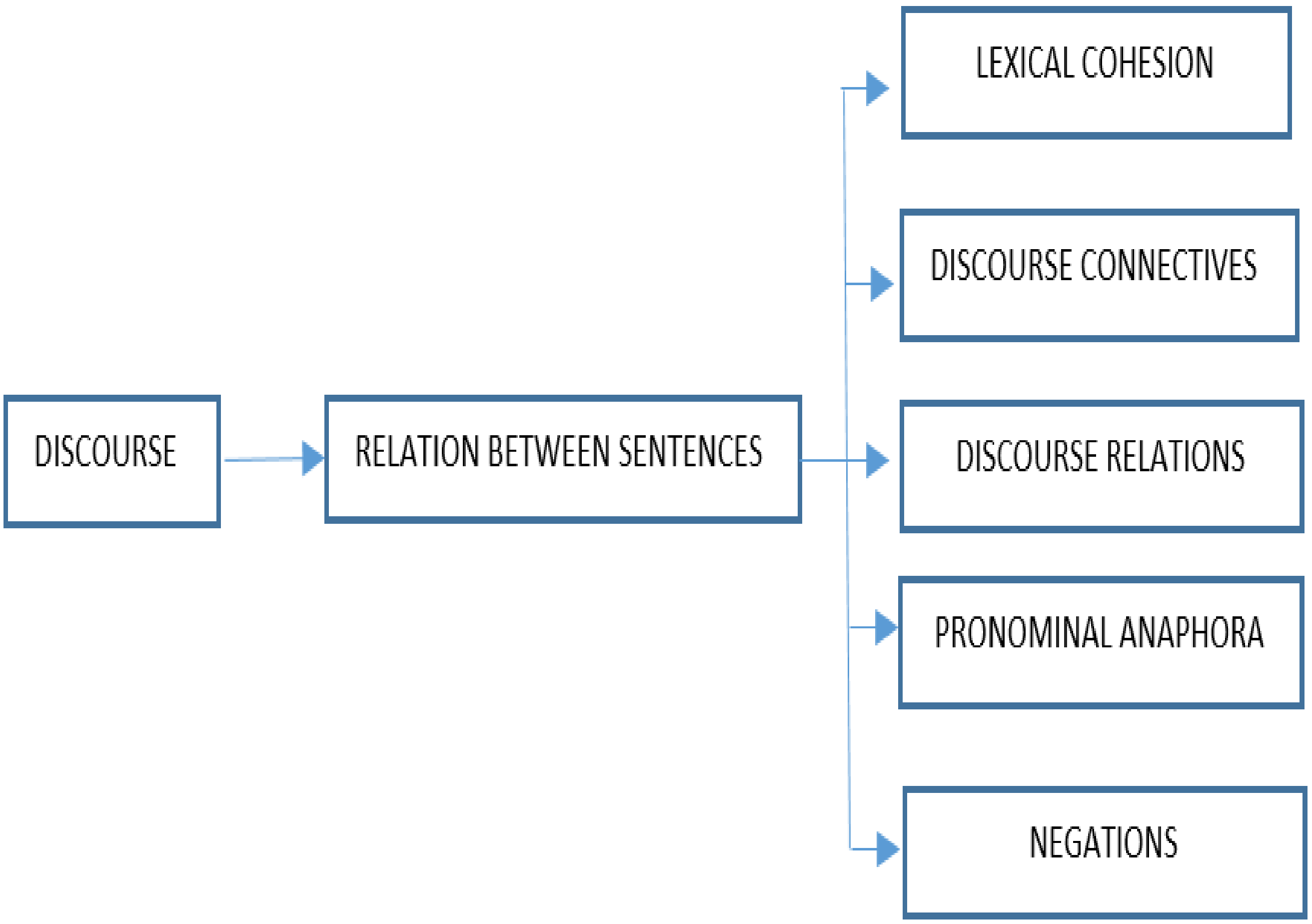


Fig 5: Reductionist approach for Discourse Incorporation in Machine Translation

Further, considering the subcategories of Discourse and the detailed analysis of paradigm involved in each reduced approach of discourse could incorporate in Machine Translation to improve the accuracy of the translated result. It has been observed that though there is a consistent attempt to analyze and process discourse relation while translation still discourse analysis is in the verge of in-depth analysis. So much work has yet to be done in the fields of anaphora resolutions, discourse connectives, negation, cohesive words and word sense disambiguation. Furthermore, we can discuss the problem of Discourse Handling in SMT, providing a more detailed survey of recent research activities and analyzing possible reasons for past failures in this task. Research in this area would result in the improvement of all the Natural Language processing application specially which dealt with written text.

REFERENCES

1. YanyanJia, Yuan Ye, Yansong Feng, Yuxuan Lai, Rui Yan and Dongyan Zhao. Modeling Discourse Cohesion for Discourse Parsing via Memory Network. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Short Papers), pages 438–443 Melbourne, Australia, July 15 - 20, 2018. c2018 Association for Computational Linguistics
2. Zihan Liu, Yan Xu, GentaIndraWinata, Pascale Fung. Incorporating Word and Subword Units in Unsupervised Machine Translation Using Language Model Rescoring. arXiv:1908.05925v1 [cs.CL] 16 Aug 2019
3. Mikel Artetxe, GorkaLabaka, EnekoAgirre, and Kyunghyun Cho. 2017b. Unsupervised neural machine translation. arXiv preprint arXiv:1710.11041
4. ShuoRen, Zhirui Zhang, Shujie Liu, Ming Zhou, Shuai Ma. Unsupervised Neural Machine Translation with SMT as Posterior Regularization. Presented at AAAI 2019, pages, 4. Computation and language (cs. CL) arXiv: 1901.04112 [cs.CL]
5. Xiaomian Kang, ChengqingZong, and NianwenXue. 2019. A Survey of Discourse Representations for Chinese Discourse Annotation. ACM Trans. Asian Low-Resour. Lang. Inf. Process. 18, 3, Article 26 (January 2019), 25 pages. https://doi.org/10.1145/3293442
6. Xia, Rui& Ding, Zixiang. (2019). Emotion-Cause Pair Extraction: A New Task to Emotion Analysis in Texts. arXiv:1906.01267v1 [cs.CL] 4 Jun 2019
7. Jiawei Zhang, Bowen Dong, Philip S. Yu. Fake Detector: Effective Fake News Detection with Deep Diffusion Neural Network arXiv:1805.08751v2 [cs.SI] 10 Aug 2019
8. Mikel Artetxe, GorkaLabaka, EnekoAgirre, and Kyunghyun Cho. 2017b. Unsupervised neural machine translation. arXiv preprint arXiv:1710.11041
9. MikelArtetxe, GorkaLabaka, EnekoAgirre. An effective approach to Unsupervised Machine Translation. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 194–203 Florence, Italy, July 28 - August 2, 2019. c2019 Association for Computational Linguistics
10. Mikel Artetxe, GorkaLabaka, EnekoAgirre. Learning bilingual word embeddings with (almost) no bilingual data. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, pages 451–462 Vancouver, Canada, July 30 - August 4, 2017. C 2017 Association for Computational Linguistics https://doi.org/10.18653/v1/P17-1042
11. Zeng Li He GaoLyu King. 2019. "What you Say and How You Say it: Joint Modelling of Topics and Discourse in Microblog". In Transactions of Association of Computational Linguistic, Volume 7
12. Sim Smith Karin.2017 "On Integrating Discourse in Machine Translation", In

May – June 2020
ISSN: 0193-4120 Page No. 94 - 101

proceedings of the third Workshop on Discourse in Machine Translation, Copenhagen, Denmark, publisher ="Association for Computational Linguistic",url=https://www.aclweb.org/anthropology/W17-4814
13. Christian Hardmeier. 2015. On Statistical Machine Translation and Translation Theory. In Proceedings of the Second Workshop on Discourse in Machine Translation. Association for ComputationalLinguistics ,Lisbon, Portugal, pages168–172. http://aclweb.org/anthology/W15-2522.
14. Christian Hardmeier. 2012. Discourse in Statistical Machine Translation. Discours 11-2012
15. Song Feng, Ritwik Banerjee, yejin Choi, Syntactic Stylometry for Deception Detection. 2012 Proceedings of the 50th Annual Meeting of the Association for Computational Linguistic.
16. ManviMahana, Mishel Johns, AshwinApte. 2012. Final Report of Project submitted in Stanford University
17. S. Valenti ,NeriCucchiiarlli. 2003. An overview of Current Research on Automates Essay Grading.
18. Yashvardhan Sharma, Sahil Gupta. 1877-0509 © 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/3.0/) Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2018). 10.1016/j.procs.2018.05.090
19. Tuan Manh Lai, Trung Bui, Sheng Li.Proceedings of the 27th International Conference on Computational Linguistics, pages 2132–2144 Santa Fe, New Mexico, USA, August 20-26, 2018.
20. Samuel R. Bowman, Christopher Potts, and Christopher D. Manning. (2015)"Recursive neural networks can learn logical semantics". Proc. of the 3rd Workshop on Continuous Vector Space Models and their Compositionality.
21. RoyalSequiera, GauravBaruah 2017. ExploringtheEffectivenessofConvolutionalNeuralNetworks forAnswerSelection inEnd-to-EndQuestionAnswering. arXiv:1707.07804v1 [cs.IR].
22. Subhabrata Mukherjee, PushpakBhattacharjee.[2012], Sentiment Analysis in Twitter with Lightweight Discourse Analysis. International Conference of Computational Linguistic. Association of Computational Linguistic.
23. Sophia Yat Mei Lee, Ying Chen, and Chu-Ren Huang. 2010. A text-driven rule-based system for emotion cause detection. In Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, pages 45–53.
24. AtefehFarzindar, Diana Inkpen, Graeme Hirst. 2017. Natural Language Processing for Social Media, ISBN: 9781681736136, Publisher : Morgan and Claypool: IEEExplore.