Sentiment Extraction From Text Using Emotion Tagged Corpus

Thasina Tabashum
Computer Science and Engineering
American International University Bangladesh
thasinatabashumabonti@gmail.com

Sheikh Abujar
Computer Science and Engineering
Daffodil International University
sheikh.cse@diu.edu.bd

Md Mohibullah
Department of CSE
Comilla University
mohib.cse.bd@gmail.com

Abdul Mutalab Shaykat Computer Science and Engineering American International University Bangladesh shaykat2057@gmail.com

Shekhor Chanda Computer Science and Engineering American International University-Bangladesh shekhor.chanda333@gmail.com

Abstract—Emotion tagging aims to tag words with appropriate emotion that a word expresses in a document or a sentence. In this paper, authors introduced an approach to tag words automatically and then analyzed in sentence and document level. The evaluation of tagged words described and the performance based on changing the tone of a document has been explained. The experiment showed that the approach could be effective if the tag words are not constant or static. Application of this approach in a contextual level analysis is briefly concluded and proposed tentative emotional tree by using emotion tagging in word level.

Keywords— Sentiment Analysis, Natural Language Processing, Emotion Tagging

I. INTRODUCTION

Emotion is a feeling from one's circumstances which is not constant and varies from place to place, time to time or person to person so determining emotion of a document or an article is not straightforward process. Sentiment classification of a text is quite difficult task. This imponderable research has lots of applicable and practical uses in the field of Natural Language Processing. Analysis of opinion, events, news articles, topics, posts on social media provide us a vast problem space. Sentiment analysis is divided into three distinct levels: document level, sentence level and aspect level[1]. Document level classification is classifying overall text for example: review of service or product. Sentence level classification is classifying a sentence opinion. In this study word tagging with emotion polarity will be approached. Assigning emotion tags to the words with Ekaman's (1993) six basic emotions- joy, sadness, anger, fear, surprise, disgust along with a distance measurement. The goal with this approach is to allow one to easily and accurately analyze these sentiments by tagging words with emotions. In this study, a tagged dataset is created. However, all the words could not be included therefore a system has been created to tag the untagged words. A dynamic system has been proposed to label the unlabeled words with a correct emotion that word expresses. The objective of this study is to tag words according to the emotion it expresses with a distance measurement from actual standard emotion level. By distance measurement the intensity of that emotion can be

concluded. In section 2, literature reviews has been discussed. Resource collection and formation is described in section 3, followed by process of data analysis and evaluation process in Section 4 and 5 respectively. Emotional tree is proposed with emotion tagged word in section 6. Finally, the last section includes a discussion of the future scope of this paper.

II. LITERATURE REVIEW

In [2] retrieved English sentiWordNet and translated it to Bangla for emotion tagging word at sentence level. For training tenacity here used CRF framework and some topographies have been nominated for six emotions classification. In [3] the author introduced co-training for the sentiment classification of multiple languages. The author used English dataset to perform classification in Chinese. Explanation of importance of adjectives and verbs for opinion mining has been introduced in [4]. In recent past, the domain of NLP for academics has been broadening. Online product recommendation system [5], determining the subjective adjectives [6], topic detection and identifying the pilot of the study [7] are focused on text processing. As the importance of opinion mining is increasing day by day in many sectors of modern life, new methods are being introduced and researchers are improving the existing methods. Researchers are mostly focused on general solution for all kind of sentences in a document. Classifications of sentences before go for sentiment analysis can improve performance and accuracy [8]. Sentiment analysis is not like other text mining researches, it depends on many parameters like context, background of the context, sentence pattern etc. and that's why accurate results from sentiment analysis is infrequent. To extract attribute from document NaÃrve Bayes classifier has been used in [9]. Authors simplified a data set model for sentiment analysis in Bangla text and Randomized Bangla text in [10], therefore many works have been completed on phrase level sentiment analysis. In NLP it is very important to identify the subjective parts of the document in order to identify the context, which is essential to determine. In [11] the author conferred the identification of subjective sentences of a document to determine its subject using SVM classifier.

A chat scheme [12] is offered that customs enthusiastic text with expressive material. Here is an Arrangement to assemble information using some physiological sensors which are attached to the user's body. Context is major term in sentiment analysis and to detect contextual polarity in [13] it comes with an approach which has the capability to inevitably recognize contextual polarity. Manual footnote of contextual polarity and an inter-annotator covenant is premeditated in this paper [14]. Contextual valence is tricky thing in sentiment analysis especially in Bangla text. It is used here for sentiment analysis from Bangla text. WordNet is a strongly structured structure that contains words according to their relations with each other and contains explanations and examples in it. It is easy to get the sense of each word conferring to its parts of speech and SentiWordNet has been castoff to recognize polarity. Summarize collection of similar opinion has been presented and Graph-based technique is used for summarization and the statement is combined using sentiment analysis [15]. Data is evaluated on the collection of 250 English news text leveled with sentiment from sentiment voting system. There are some available part of speech tagger Stanford NLP parser is one of them and it is castoff to gain the part of speech tagging [16].

III. RESOURCE COLLECTION AND FORMATION

Resources are collected and separated in to six basic emotions. Six emotion categories- joy, sadness, anger, fear, disgust and surprise have been labeled. In this study, wordnet effect emotion list for tagged data has been used and all the tagged words are included in one dataset according to the tags. Total number of tagged words in dataset is 1135. Labeled words include 400 "Joy", 148 "Fear", 202 "sadness", 261 "anger", 53 " disgust" and "71" surprise In the table 1 tag names and table 2 some tagged data from dataset have been shown:

TABLE I EMOTION TAGS

Tag
JY
SD
AN
FR
DG
SP

TABLE II TAGGED DATA

Words	Tag	Words	Tag
Angry	AG	trepidation	FR
Angered	AG	timidity	FR
Enraged	AG	satisfaction	JY
Furious	AG	rejoicing	JY
Distasteful	DG	regard	JY
Disgustful	DG	amaze	SP
Unassertiveness	FR	astonish	SP

IV. DATA ANALYSIS

In this section, our objective is to label the unlabeled words. To achieve this, the words were tagged on both an

automatic and manual level. Therefore this method of emotion tagging ultimately gives the accurate result for all the level of sentiment analysis. The aim is to accurately tag words with the correct emotions using this method for opinion mining.

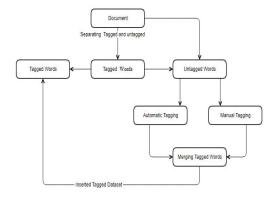


Fig. 1. Work flow of proposed emotional word tagging system

A. Automatic Word Tagging

Machine has given an article about Hospital incident [17]. Comparing with the tagged dataset only 7 words are tagged of an article of length 6531 characters, and consisting of 1246 words. Those 7 words were already existed in the tagged dataset.[('heart', 'JY'), ('bad', 'SD'), ('suffering', 'SD'), ('surprise', 'SP'), ('get', 'SP'), ('bad', 'SD'), ('contentment', 'JY')]. Since it is not possible to obtain the sentiment of that document more emotion tagged words are essential to perceive the intended result. After compiling the article through the module, it separated the untagged words. These unlabeled words needed to be labeled however they must be further analyzed in order to be accurately tagged with the right emotions. All the untagged words are compared with the tagged words. For dynamical approach Wu-Palmer metric (WUP) has been applied, WUP weights the edges based on distance in the hierarchy for the unlabeled word from all other labeled words. In this paper polarity has been assigned for each emotion. The polarity calculation is not constant. It varies with the tagged dataset. The larger the dataset, the more accurate the result is. Polarity equation: -

$$P(x = 'JY', 'SP', 'SD', 'FR', 'DG', 'AG') = \frac{Max(L)}{L(x)}$$

Max(L) = maximum length, L(x)=length of x. Here, maximum length of tagged emotion is "JY' which is 400. If, x='FR'; Length of x=148; P (x='FR') = 400/148 = 2.703

In this manner all the emotion polarities have been assigned. Six emotion percentage of a word calculated with equation (2).

$$P.E(w,x) = \frac{\sum_{i=1}^{n} S(i,w) * pix}{\sum_{i=1}^{n} S(i,w) * pj} * 100[s > = 0.5]...(2)$$

In this equation, 'j' is from joy, fear, disgust, sadness to anger. 'P' is polarity. Next,'s' is the WUP similarity between

TABLE III EMOTION TAGS WITH POLARITY

Joy	1.0
Surprise	5.633
Sad	1.980
Fear	2.703
Disgust	7.547
Anger	1.532

untagged word and the tagged data and s is greater than 0.5. n= number of total tagged data (1135) and x can be any of the six emotions. P.E of some words is shown:

P.E ("punishment", 'JY') P.E ("punishment", 'SP') P.E ("punishment", 'SD') P.E ("punishment", 'FR') P.E ("punishment", 'DG') P.E ("punishment", 'AG')	= = = =	22.026 12.428 4.114 15.117 35.700 10.616
P.E ("punishment", 'JY') P.E ("hospital", 'SP') P.E ("hospital", 'SD') P.E ("hospital", 'FR') P.E ("hospital", 'DG') P.E ("hospital", 'AG')	= = = =	22.026 34.5478 11.4684 0.0 0.0 35.78067
P.E ("punishment", 'JY') P.E ("punishment", 'SP') P.E ("punishment", 'SD') P.E ("punishment", 'FR') P.E ("punishment", 'DG') P.E ("punishment", 'AG')	= = = = =	22.026 12.428 4.114 15.117 35.700 10.616
P.E ("hospital", 'SP') P.E ("hospital", 'SD') P.E ("hospital", 'FR') P.E ("hospital", 'DG')	= = = = =	18.2031 34.5478 11.4684 0.0 0.0 35.78067

Through this process, the P.E. of 240 words was calculated. These 240 words were each tagged with the maximum P.E value that was found for a single word. Taking maximum P.E "punishment" has been tagged as "DG". The total number of words tagged summed up to 247.

TABLE IV
NUMBER OF WORDS WITH EMOTION TAG

Emotion Tag	JY	SP	SD	FR	DG	AG
Number of Words	16	33	77	12	64	43

B. Non-Automatic Word Tagging

The same article is provided to readers to manually tag the words. For this purpose a system has been developed which will provide the words and the users will select the emotion. Around 17-19 readers gave the feedback. Average values are taken then maximum valued emotion is assigned to the word.

Total number of manual tagged words is 311, which is more than dynamically tagged data because the latter searches

TABLE V
TAGGED WORDS BY READERS

Words	Joy	Surprise	Sad	Fear	Disgust	Anger
careless	0	0	25.5	18.5	15	41
childcare	38.34	1.05	2.66	42	10	5.95
parents	29.9	20	22	13.1	0	15.1
baby	27	22.5	3	2	0	7
girl	18.33	0.67	15.01	8.9	5	21.3
Ctg	13.3	0	0	0	0	0
hospital	8.2	23.9	15.8	33.8	0.3	18

TABLE VI MANUALLY TAGGED EMOTION

Emotion Tag	JY	SP	SD	FR	DG	AG
Number of words	28	65	76	26	63	53

for similarity above 0.5 and derives its own emotion. As the dataset was not substantial, it may not find many words which are above the aforementioned mark thus many words might not be taken into the calculation.

C. Merging Dynamic and Manual Tags

The automatic and manual tags were then compared. The words that had the same tag in both cases are inserted into the main dataset. The main objective of this is to create an enriched emotion tagged dataset. Manual tagging will no longer be necessary once the dataset is large enough and the system will be capable of automatically analyze accurately. Since many words were not being tagged initially, manual tagging in the small scale was important. Moreover, the manually tagged words are used for cross check the result of the dynamic tagging process. Taking set of automatically and manually tagged words, 96 words were tagged the same. These tagged data was added to the main dataset.

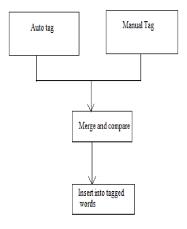


Fig. 2. Merge emotional words from Auto tagged dataset and manually tagged dataset

V. EVALUATION

Evaluation has done in word level and sentence level. Automatic tagging is compared with manual tagging for accuracy.

In this section, it has been discussed about the issues for the change in accuracy.

A. Word Level Evaluation

In order to measure the accuracy, a confusion matrix was developed comparing manual and auto tags. The accuracy was found to be at 82.75/100. Updated tagged dataset includes -406.0 "JY", 157.0 "SP", 224.0 "SD", 276.0 "FR", 76.0 "DG", 89.0 "AG". The table 7 is a confusion matrix of an article about child bus accident which gave very good results. But if this learning approach is provided with a positive article about the 'Prime minister' [18] the accuracy level was 56.67/100. Correctly tagged words were only 34 among 60. Since the system had initially been provided with a negatively toned article, it learned to tag words with negative emotions more than positive.

TABLE VII CONFUSION MATRIX

	JY	SP	SD	FR	DG	AG
JY	8	2	1	0	2	0
SP	4	9	2	0	2	1
SD	0	0	23	0	1	0
FR	0	0	1	15	2	0
DG	0	0	0	0	23	0
AG	1	0	0	0	2	18

Therefore, when given an article with a positive tone afterward, it was unable to accurately determine the overall sentiment of the article. This proved that word level evaluation is not sufficient, and the analysis must be done on a contextual level as well.

B. Sentence Level Evaluation

Sentence level evaluation was considered as the ultimate aim of word tagging is to determine the opinion or tone of a sentence or an article, eventually. Three tagged sentences are shown below:

"With a heavy heart, Rokhsana Akter, 21, began making preparations for the funeral of her five-day-old girl, also her first child." (s.1)

Tag	JY	SP	SD	FR	DG	AG
Words	"heart	' –	"funeral", "girl", "child"	_	-	_
Number of Words	1	0	3	0	0	0

"The authorities allegedly warned the mother and her family members not to unwrap the shroud as the baby's face was in bad shape due to excessive bleeding." (s.2)

"The authorities helped the mother with the situation, her baby is fine now" (s.3)

This approach will not always yield accurate results. It was used to accurately tag s.1 and s.2, but was unable to do the

TABLE IX Number of Words Tagged under Each tag in S2 $\,$

Tag	JY	SP	SD	FR	DG	AG
Words	_	"Unwrap",	"baby",	_	"Authorities	", "face"
		"shroud"	"shape",		"family"	
			"bad"		-	
Number	0	2	3	0	2	1
of						
Words						

Tag	JY	SP	SD	FR	DG	AG
Words	"fine"	_	"baby'	' –	"Authorities"	"situation"
Number	1	0	1	0	1	1
of						
Words						

same for s.3. This learning method gave accurate result for negative toned article because the machine learned the words from a bitter article titled "Careless Childcare" from Daily star. Thus, when happy sentences are provided to the machine, its prediction is not as accurate. Statically tagging the words with emotion cannot give the actually emotion that word expresses in a particular sentence or article, therefore a context level is needed to have an exact result.

VI. CONTEXTUAL LEVEL APPLICATION

In this section, a contextual level analysis using emotion tagged word is proposed. Contextual methodology for future study will involve the method used in this paper to create a dynamic, rather than static, tree corpus of tagged words, where the weight of the words and the distance between them is taken into consideration. Rather than static emotion tags, the dataset will learn according to the nodes around it. Tagging S.2, "authorities" tagged as "DG", but tagging manual s.3 authorities should no longer be tagged as "DG", but rather "JY". Creating sentence level tree will be very helpful to analyze sentence sentiment in the future. As an example after creating two trees, if the machine is provided with another sentence:

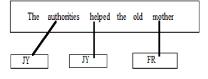


Fig. 3. Tagged words in the sentence

Here, tagging mother as "FR" is inaccurate. It should be changed. It doesn't necessarily express any emotion in this sentence. Thus there is abundant number of neutral words. If more sentences are provided about mother, for example: - mother loves her child, mother is crying, the tag for mother

is changing every time. These kinds of words, which are frequently changing shall be tagged as 'Neutral' The proposed method would create a tree where only parts of speech tags-noun, adjective, verb, adverb shall be considered. If pos =Noun that won't create another child branch under it. Like in figure 2 under "Unwrap", there are "shroud", "face", "baby". Since they are nouns, a child is not created until the word "was" is found. Proposed methodology is, initially creating a corpus of sentence level tree and emotion tagged words which will consider its branches, weight, distance and polarity in order to calculate the emotion tag. Below is a demonstration of the proposed algorithm for dynamic emotion word tagging:

2)

$$allowed_types = ["JJ", "JJR", "JJS", "NN", \\ "NNS", "RB", RBR", "RBS", \\ "VBD", "VBN", "VBS"]$$

- 3) Creating tree:
 - a) Root(first word of the sentence)
 - b) Parent=root
 - c) For p in allowed_types:
 - i) If p is not noun and is not parent: create edge from parent to p parent=p
 - ii) else if (parent-1) is not root then create edge from all(parent-1) to p
- 4) Search the created tree on the emotional tree
- 5) If found then tag the words
- 6) Else if found some branches of the tree, then compute equation (1) and (2) for untagged words and the s is WUP similarity between the words of surrounded branches and the polarity will also be calculated with those branches. Weight will be assigned with the each of neighboring words considering the distance between them. So equation (2)

$$P.E(w,x) = \frac{\sum_{i}^{n} S(i,w) * pix * dis(i,w)}{\sum_{i} \sum_{i}^{j} S(i,w) * pj * dis(i,w)} * 100$$

[s >= 0.5](3)

- 7) Else if no match found then approach will be the one used in this paper
- 8) Create set of each word store in emotion tracking dataset:
 - a) wordEmotionTrack[key]=[[eos_tag][P.E(key, eos_tag)]]
 - b) totalOccurence[key]+=totalOccurence[key]
- 9) for w in all words

$$Wth = \frac{len(set(wordEmotionTracj[w]))}{totalOccurence[w]}$$

10) if Wth>0.5:

In step 10, Wth is word threshold which is measuring the variance of each word. If any word tag is changing frequently that word will added under neutral tag. This paper study actual purpose is for noise free sentiment analysis by tagging words with emotion which is only be possible if the analysis is in contextual level. Contextual level analysis is more accurate compared to word frequency analysis.

A. Evaluation word level tagging with dynamic tree level tagging

Initially tagged sentence "The authorities allegedly warned the mother and her family members not to unwrap the shroud as the baby's face was in bad Shape due to excessive bleeding." Another sentence "The authorities helped the mother with the situation, her baby is fine now", for word level tagging using equation (1) and (2) the result has shown table 9 where baby tagged as 'SD' which is wrong after close investigation. Contextual tree has shown in figure 4 which is built on the basis of word level tagging. But in this new sentence baby is not conveying sadness. In the new sentence "baby" should be tagged as a positive toned sentiment rather than negative toned sentiment. On the other hand, for the previous sentence sadness it is accurate for the "baby" word.

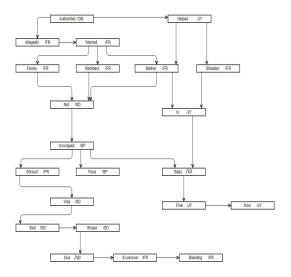


Fig. 4. Initial/Data Tree Using Word Level Tagging

Applying the contextual tree algorithm for word tagging with emotion gives more satisfying result. Using equation 4 to measure the distance to "baby" node to surround nodes has been shown in figure 5. In this tree, it will travel choosing the path of immediate parent nodes. After reaching the starting node "Authorities", it travels the existing data tree without visiting same node again. Distance is measured with below equation:

$$dis = \frac{1}{En}...(4), En = number of edge traveled.$$

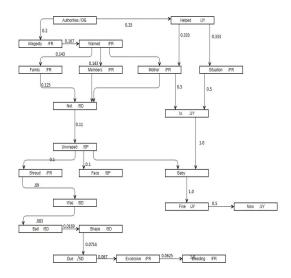


Fig. 5. Distance form a node to other nodes in contextual tree

Using Equation 3, the word "baby" tag has been updated to "SP".

P.E(Baby,SP) = 82.5263, P.E(Baby,FR)=44.23

The contextual level analysis given the better result which is dynamic, emotion tagging should be dependent on the context rather than individual words. Article [18] words extracted tags which were previously tagged incorrect after using contextual analysis tags are updated shown in table 11:

meeting	JY	DG	SD
Prince	SP	SD	SD
Victory	JY	DG	SP
criticism	SP	SD	SP

Currents tags are comparatively more accurate than the previous ones. Considering as a word how it is contributing to the sentence to expressing an emotion is giving valid result. Applying the contextual tree level learning on [18] current accurate tags number is 39 among 60. But other inaccurate tags as example "victory" tagged as "SP", which make more sense than "DG". This learning method will learn more error freely if its data tree is large enough to estimate the actual emotion of a word in a sentence or passage.

VII. CONCLUSION

The limitation of this paper was in the depth of the analysis. The word tagging also should be made dynamic instead of static so the system is better able to conclude the contextual meaning rather than go by previous learning. This way, the database will be better enriched and would not depend on its initial learning curve. Since the system was too dependent on the meaning of the initially tagged words we were unable to accurately determine the overall sentiment of an oppositely toned article supplied to it afterwards. In conclusion, as the

documents were evaluated according to word level sentiment but this alone can't give actual sentiment analysis. Thus contextual tree level approach is needed along with the tagged words to better evaluate this. Using this approach of labeling emotion on word level will give better and faster result.

ACKNOWLEDGMENT

We thank Daffodil International University Machine Learning and NLP lab for their support through-out the research.

REFERENCES

- Liu, B. (2014). Sentiment Analysis and Opinion Mining. Dissertation Abstracts International, B: Sciences and Engineering (Vol. 70). https://doi.org/10.1162/COLI
- [2] Das, D. (2009). Word to Sentence Level Emotion Tagging for Bengali Blogs. American Psychologist, (August), 149âÅŞ152.
- [3] Wan, X. (2009). Co-Training for Cross-Lingual Sentiment Classification, (August), 235âĂŞ243.
- [4] Chesley, P., Vincent, B., Xu, L., Srihari, R. (2006). Using verbs and adjectives to automatically classify blog sentiment. Training, 9(4), 8. https://doi.org/10.1016/0020-0271(73)90092-2
- [5] Chai, J., Horvath, V., Nicolov, N., Stys, M., Kambhatla, N., Zadrozny, W., Melville, P. (2002). Natural Language Assistant: A Dialog System for Online Product Recommendation. AI Magazine, 23(2), 63. https://doi.org/10.1609/aimag.v23i2.1641
- [6] Baroni, M., Vegnaduzzo, S. (2004). Identifying subjective adjectives through web-based mutual information. Proceedings of KON-VENS, (2000), 17âĂŞ24. Retrieved from http://sslmit.unibo.it/baroni/publications/konvens2004/wmiKONV.pdf
- [7] Allan, J., Carbonell, J., Doddington, G., Yamron, J., Yang, Y. (1998). Topic Detection and Tracking Pilot Study: Final Report. Proceedings of the DARPA Broadcast News Transcription and Understanding Workshop, 194âÅŞ218.
- [8] Chen, T., Xu, R., He, Y., Wang, X. (2017). Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. Expert Systems with Applications, 72, 221åÅŞ230. https://doi.org/10.1016/j.eswa.2016.10.065
- [9] Hasan, K. M. A., Sabuj, M. S., and Afrin, Z. (2015). Opinion Mining using Naive Bayes. 2015 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), 511âAŞ514. https://doi.org/10.1109/WIECON-ECE.2015.7443981
- [10] Wilson, T., Wiebe, J., and Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing HLT '05, (October), 347âÄŞ354. https://doi.org/10.3115/1220575.1220619
- [11] Das, a., and Bandyopadhyay, S. (2010). Opinion-Polarity Identification in Bengali. In ICCPOL 2010, (October). Retrieved from http://www.amitavadas.com/Pub/ICCPOL 2010.pdf
- [12] Hasan, K. M. A., Rahman, M., and Badiuzzaman. (2014). Sentiment detection from Bangla text using contextual valency analysis. 2014 17th International Conference on Computer and Information Technology, IC-CIT 2014, 292âAŞ295. https://doi.org/10.1109/ICCITechn.2014.7073151
- [13] Asif Hasan, Mohammad Rashedul Amin, Abul Kalam Al Azad, Nabeel Mohammad, Sentiment Analysis on Bangla and Romanized Bangla Text using Deep Recurrent Models. 2016 International Workshop on Computational Intelligence (IWCI)12-13 December 2016, Dhaka, Bangladesh
- [14] Wang, H., Prendinger, H., Igarashi, T. (2004). Communicating emotions in online chat using physiological sensors and animated text. Extended Abstracts of the 2004 Conference on Human Factors and Computing Systems - CHI '04, 1171. https://doi.org/10.1145/985921.986016
- [15] Bhargava, R., Sharma, Y., Sharma, G. (2016). ATSSI: Abstractive Text Summarization Using Sentiment Infusion. Procedia Computer Science, 89, 404åÄŞ411. https://doi.org/10.1016/j.procs.2016.06.088
- [16] Hui, J. L. O., Hoon, G. K., Zainon, W. M. N. W. (2017). Effects of Word Class and Text Position in Sentiment-based News Classification. Procedia Computer Science, 124, 77âĂŞ85. https://doi.org/10.1016/j.procs.2017.12.132
- [17] https://www.thedailystar.net/frontpage/careless-childcare-1564540
- [18] https://www.thedailystar.net/backpage/hasina-makes-times-list-100-most-influential-people-1565233