DISCOURSE BASED OPINION MINING ON ROMAN URDU DATA

Dr. Zareen Sharf
Assistant Professor
SZABIST

Dr. Husnain Mansoor Ali
Associate Professor
SZABIST

ABSTRACT

The use of Roman Urdu (in the form of web and user-generated content) is a common mode of communication on social media. Content like comments, reviews, feedbacks and social networking posts have been generated in Roman Urdu in large volumes. But this area is not much worked on in terms of sentiment and opinion analysis. Roman Urdu (the scripting style for Urdu language) is one of the limited resource languages that brings forward the challenges and problems for performing Opinion Mining. Adequate opinion mining is not just about understanding the overall sentiment of a document or a single paragraph, but it is also important to be able to extract sentiments on a very granular level and relate each sentiment to the aspect it corresponds to. On the more advanced level, the analysis can go beyond only positive or negative attitude and identify complex attitude types. We, therefore, developed a model for performing discourse-based opinion mining, so we could also consider the impact that various discourse elements have on the overall sentiment of the text. Our work differs from the existing body of knowledge in that not much work has been carried out on processing of Roman Urdu data for opinion mining considering discourse elements. Since our work focuses on performing discourse-based opinion mining it can be considered as first attempt in this direction as none of the literature surveyed revealed discourse-based analysis of Roman Urdu text. The overall gist of this research work is to have insights of the nature of user-generated content in Roman Urdu and to build necessary resources and devise algorithms to make an advancement in Sentiment Analysis and Classification for Roman Urdu.

1. INTRODUCTION

Discourse relations play a significant role in natural language processing. This phenomenon is used to build clauses, linking relations and a coherent relation in the text. A sentence, in actual, comprises of conditionals, modals, negation and connectives. These language constructs basically change the overall sentiment at the sentence, clausal or phrasal level. For instance, “@user share 'em! I’m quite excited about Tintin, despite not really liking original comics. Probably because Joe Cornish had a hand in.” It is clear from the example sentence that the overall sentiment is positive.

Although, there is a mix use of negative and positive words, but the connective despite makes the overall sentiment positive by giving more weight to the previous discourse segment. That is the reason, it is very important to analyse discourse relations in order to extract opinions or sentiments. A bagged model (i.e. bag-of-words) would not be able to perform opinion or sentiment classification without incorporating discourse marker.

Let us take another example, *"Z10 Kaafi Intresting Set laga Lekin me bettry timing se thora dar gya hon overall set acha he Lekin 20hzaar Is set pe kharch karna Kia sahe he ya koi 20hzaar tak ka set jo apki nazar me ho jis ki ram nd storage healthy ho Ar bettry timng bi achi dy agar ap bta dein to acha hoga"*. The overall sentiment of this example comment is neutral because of the connective 'but', that gives more weigh to the following segment. That is why it is crucially important to acquire all such phenomenon in a computational model.

The main focus of this research is to perform discourse-based opinion mining on text. In this research, we are aimed at developing a model that would take into consideration different discourse features for performing opinion mining with enhanced prediction accuracy.

2. WHAT IS DISCOURSE?

Discourse (as per the concept of natural language) refers to an essential paradox in natural language processing. It greatly aids in authorising coherent relation as well as the linking of phrase and clauses within a text segment [1]. It can further be defined as a logical structured group of textual segments or units. A discourse can take any form whether a sentence, a written text, dialogue etc. [2] A connective, when appears in between the two text segments, supports the discourse relation. It gives weight to eider side of the two connected sentences or text segments, giving an overall meaning to the entire sentence, be it positive or negative.

There exist some theories that analyse discourse as well as the segments involved in discourse classification. Some of these theories are referred to as Rhetorical Structure Theory (RTS) which was proposed by Mann et al. (1988). The main focus of this theory is to idealise the two segments called nucleus and satellite, in a sentence. A lot more work has been done to establish elementary discourse units at clausal stage as well as by producing trees at sentence level. A discourse parser or (dependency parker) has been used in most of the discourse related works. [1]

3. WHAT IS DISCOURSE ANALYSIS?

Discourse analysis refers to the study of language used in text and conversation. Basically, critical discourse analysis is the branch of linguistics that takes into account about the influence of writings, means it analyses that why and how some writings put a great impact on readers or hearers as compared to the other ones. The main aim of critical discourse analysis (even by the analysis of simple grammar) is to decipher the hidden ideologies that are capable of influencing readers and hearers. The discourse analysts have thoroughly performed Discourse analysis words on a wide range of write ups and text in order to describe how writers and speakers can be influential just using their words and can become ideologically significant.

After the emergence of micro-blogs, a rapid influence on the society is witnessed. So, it is highly important to discuss the impact that micro-blogs have put on human perception. A micro-blog is actually a way to share content and videos among people that can be related to any specific topic. There are many theories that have been particularly proposed to analyse micro-blogs. But the implementation of such theories has raised different problems. Consider an example of Twitter, where people are not restricted to any specific kind of content or style of language. That is why, the tweets are full of slang words, abbreviations in informal tone. Even there are many grammatical errors or discontinuities as well. Due to this reason, the natural language processing tools (like taggers, parsers) get failed as they are not able to handle such unstructured data. [3]

The conventional natural language processing tools incorporate the use of connectives, modals, conditions in order to analyse the text. But in case of micro-blogs, these kinds of formalities are usually ignored as they are replaced by some specialised domain-specific characteristics like hashtags and emoticons. [4]

In order to yield trees at the sentence level, the identification of elementary discourse units was carried out in the proposed probabilistic models [5]. The proposed model incorporated syntactic and lexical information from discourse-annotated corpus. Polanyi (2004) studied about the effect of modals, connectives and negatives (that used to change the prior polarity of the said words to make out new meanings) in Contextual Valence Shifters. These models discussed about a simple weighting scheme as well as the pre-suppositional items and irony.

More work on tweets classification was performed by Joshi et al in 2011. They came up with a rule to perform classification of tweets as positive or negative that depended upon the opinion words present there.in. They utilised specific words such as emoticons, sentiment lexicons and hashtags. Due to a massive use of hashtags and emoticons on social media, it is essential to include them in any discourse analysis rule. Similarly, Gonzalez in 2011 also utilised the information provided in hashtags in order to distinguish wit, sarcasm and negative and positive tweets. According to him, hashtags are the best tool (or indicator) to detect a discourse is sarcastic or not.

A study on the basis of works done by [6], thoroughly discussed about the sentiment analysis of micro-bogs. This study had exploited particularly the works related to Twitter, to develop a bagged model (i.e. bag-of-words style model). This bagged model utilised discourse related features in order to give an enhanced accuracy. The research evaluated three sets of data making use of lexicon based classification with the help of supervised classifier. The study incorporated more than 8500 tweets in the set of labeled tweets (these were the manually used tweets) whereas the automatically annotated set contained around 15200 tweets. The datasets utilised in the study were from the domain of travel review of [7]. The major aim of the study was to demonstrate that the employed method was highly beneficial to the reviews which were structured also. [1]

4. TYPES OF DISCOURSE

There are different types of discourse. The ordinary language discourse includes the expressions of feelings, emotions or attitudes. Therefore, it can further be classified as:

Coherently Structured Discourse

One of the types of discourse is that in which a group of sentences has some relationship with each other. This refers to as coherently structured discourse. The relation in group of sentences is explained by a coherent relation and how they basically interact with each other. This coherent relation in a discourse structure actually varies with respect to two approaches. In one approach, the major target is to equate intentional structure of discourse. Here, the coherence relation emulates that how role of a segment (played with respect to interlocutor's purpose) is interfaced to the other segment's role.[8]. While the second approach works on idealising the informational structure of a discourse. Means in this approach, the coherence relation emulates how meaning transmitted by one discourse segment is divulged with the meaning conveyed by the other segment. [9] [5]. The coherence relations in a discourse are illustrated by using Conjunctions. Below is the list of those types:

- Condition – if…then; while; as long as
- Similarity – similarly; and
- Example – for instance; for example
- Cause-effect – because; due to; and so
- Generalisation – in general

- Attribution – claim that; according to; stated; maintain that
- Elaboration – more; in addition; moreover; note (more) that; for/in/with which; for/in/against/with/on whom; also; who
- Contrast – but; by contrast
- Violated Expectations – while; but; although \Temporal Sequence – before; after; (and) then; while; first; second

Explicit Discourse

Another discourse is referred to as explicit discourse, when the text does not have any kind of explicit cues. The use of connectives (such as since, however, because, and, etc.) mark the presence of explicit discourse. [10]. These discourse relations are quite easy to be identified. The general senses including contingency, comparison, temporal and expansion are easily authorised in explicit discourse relations with accuracy of about 93%. This accuracy totally depends upon the usage of discourse connectives to signal relation. [11]. Penn Discourse Treebank is usually employed for experimentation in order to identify explicit connectors. This Treebank is a bulk and illustrated collection of discourse relations. Other samples are also used for explicit discourse relation analysis like English Gigaword Corpus that contains around four million news articles. [12]

Implicit Discourse

When there is a discourse connective between two text segments, it usually becomes easy to recognise the relationship between the segments because the connectives used are unclear. [13]. On the other hand, the discourse relation gets difficult to be identified when there are no explicit text cues are found. The implicit discourse is the one in which the discourse relation does not contain any connectives and the two text segments are almost adjacent to each other. [14]

Dataset used for Discourse Based Opinion Mining

To train a sentiment classifier, a fairly large size of training dataset of records (already labeled with sentiment) is required. We scrapped data from various blogs, discussion forums and social media sites (like Facebook) using scripts written in Python Language. The irrelevant records (i.e. records other than those in Roman Urdu) were discarded and each relevant record in the dataset was manually labeled as positive, negative or neutral sentiment. As shown in Table 3-2 and Figure 3-2 our data corpus comprises of approximately 22000 records. Out of which 4500 records are labeled positive, 4900 records are labeled negative, 13000 records were labeled neutral and approximately 5000 records were discarded for being irrelevant.

Table 1 Dataset Distribution

Total Records	Positive	Negative	Neutral
22000	4500	4900	13000

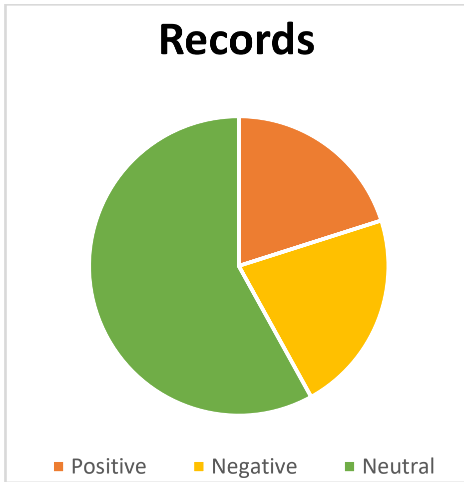

Figure 1 Dataset Distribution based on Sentiment Polarity

A list of web sources and Facebook pages used for retrieval of records is presented in Table 3-3 and Table 3-4.

Table 2 Web Sources for Data Collection

Sources	
Biographies	Womens_Fashion
UrduSafha	Express Urdu
Khabees	Reddit
Others	Comments & Reviews
Bitchy Urdu Cards	Urdu Poetry
AryNewsAsia	Sentiment Analysis Interviewer
Social Workers	Careem
Phones_tablets	2LinePoetry

Blog Khuwaar	Mens_Fashion
Beauty_Health	

Table 3 Facebook Pages for Data Collection

Source			
Expressnewspk	SHUGAL	sarcasm123	MRSMofficial
UrduAdab	TheOlaadMovement	arhamsayss	IdaraTulMustafaInternational
NaziaH	urdughazal	halalhumour	girls.pk
LolWaalay	EdaTuD	Dostonkibaateinn	english.eminem
HusseyOfficialClub	HappensInPk	Filmygyan7	jppk123
Entertainmenttrackerpage	WBloverss	PakistaniiAwaaaam	hadeed.khi
css	urdupoint	Sheikhspeareofficial	GHStrangers
ProgrammerKiBaatein	OfficialLahoreQalandars	AkbarSayss	Gulshan-E-Hadeed-Official
shahlylaB7	EmployeesUpGradation	Bfkbaatein	Hadeediansrocks
NVinez	UzairAltafPage	ChotiSiQayamat.seetv	yaradaynaalyari
TheMahiraKhanOfficial	ChupbeyTharki	Alif-on-See-TV	zindgigulzarha
ArifaSobhKhan	ItssAZee	Humorists2	Aghaz-Society
PyaareNabiKiBaatain			

5. DISCOURSE BASED OPINION MINING MODEL

Discourse-based opinion mining has not been undertaken for many languages and this is the first time it has been done for Roman Urdu language. Therefore, we developed a hybrid model that combined the features of lexicon based and rule-based approaches. A large data corpus was collected, preprocessed and annotated to incorporate the various aspects of discourse-based opinion mining for Roman Urdu. It is also evident from the literature review that languages that have limited resources available are mostly analysed by developing models that are based on supervised or unsupervised machine learning techniques, so the same approach was undertaken for this research as well.

For this research work, we have focused on only those discourse types that have an impact on overall sentiment of the statements. These are **Condition, Cause-Effect, Elaboration, Violating Expectation, Contrast and Negation** type conjunctions as specified by [15] . Some of these types are ones that emphasise a single sentiment while others are used to refute the neighbouring discourse segment (either following or previous one). They are divided into two sub categories:

Conj_After – This corresponds to the set of all conjunctions that give more weight (or importance) to the following discourse segment.

Conj_Before – This corresponds to the set of all conjunctions that give more weight (or importance) to the previous discourse segment.

Let us elaborate it with an example: The direction was (not that great)-, but still we loved+ the movie.

When a bag-of-words model is applied, it will find the sentence comprises of one positive and one negative sentiment. So, it will finally be classified as neutral, but the overall sentiment is positive. Since, the final judgement passed through the sentence is that we loved the movie, so it means the words following the discourse 'but' should be given more weight.

Another example for explanation is here: Pakistan managed to win+ despite the initial setback -. This also presents the similar problem. And more weight is given to the segment before the discourse 'despite'. Following is the list of discourse elements recognised and processed by the developed model.

Conj_Succesor: but(laikan), however(albata), nevertheless/nonetheless (is kay bawajood), otherwise (doosri soorat mai), yet (abhi tak), still(abhi tak/phir bhi), par, pur

Conj_Predescessor: till(tak), until (jab tak), despite (halan kay), inspite(bawajood mai/ bawajood is kay), though/ although(agarchay), aur, or

The adverbs (or adverbial phrases) that strengthen the meaning of other expressions and show an emphasis are termed as Intensifiers. The words that are commonly used as intensifiers include: *extremely, absolutely, completely, likely, utterly, really, highly, rather, too, very, so, totally and at* *all*. List of **intensifiers** for Roman Urdu: '*Bohat'*, *'Kafi'*, *'Shaandar'*, *'Nihayat'*, *'bilkul'*, *'bara'*, *'bari'*.

We started off with first acquiring a list of most frequently used words in Urdu language from the "Centre for Language Engineering" website (www.cle.org.pk). This list was then transliterated into Roman Urdu and tagged with the corresponding Part of Speech using the site ijunoon.com. Once we had this list we also used the bag-of word approach to add new words from the data corpus we had collected to the existing list. This list is primarily divided into two sub-lists. One for positive words and the other for negative words for predicting the sentiment of the input string. Next we check for the presence of discourse element within the input string and if one is found the string is split into a predecessor string and a successor string. The sentiment is then calculated based on the type of discourse element and also taking into account the intensifier words and negative connotation words like 'nahi' (means No in English).

Pseudocode for detecting discourse is as follows:

1.Take a comment as input.

2.Check if discourse exists.

3.If (True)

4. Split into predecessor(pre) and successor(succ) parts.

5.Else

6. Send input to the Opinion Mining Model

7.Check for Intensifier

8.If (True)

9. Set intensifier=True

10.Else

11. Set intensifier=False

Following is the description of the proposed models working.

<u>a. Preprocessing</u>

1. Build Standardizer Routine.

It reads through the data file and uses it to find frequency of all similar sounding words, and their variations, and store it in a nested associative array. They are later use for word look ups. Using CHECK command

2. Load Standard Routine.

This reads the standard_word_list file and make a list of all standard representation of the phonetics the hashing algorithm generates. For this purpose, it builds an associative array, where each phonetic sound is mapped to its unique standard representation. e.g. ‘qitab’ is mapped to ‘Kitaab’ etc. This associative array is later used to change all input data into standard notation.

3. The negative and positive words are loaded into separate list data structure from their files, neg.txt and pos.txt respectively.

4. The program takes sentence as an input. And it checks

i. If it is ‘CHECK’ then look up the following word into the nested associative array it built in Step 1.

ii. If it is q, it exits.

iii. Otherwise it processes the sentence as in the steps mentioned below.

5. Firstly it checks whether the given sentences are inherently more than one sentence long. It checks it by whether:

i. It has fullstops

ii. It has multiple conjunctions

The function returns, a Boolean, and a list of sentences. Single sentence, if there are no multiple sentences.

6. It cleans every word of the sentence via cleaner routine. And prints

7. It hashes every cleaned word via hasher routine. And prints

8. It finds the standard notation of the cleaned, hashed word. And prints.

9. It sends the sentence to find_sentiment function, which would return a number for the sentiment.

b. Intensifier Identification

Pseudocode for performing discourse-based opinion mining is presented as follows:

1.Calculate Sentiment using Lexicon Based Approach.

2. If (intensifier=True)

3. Set Over All Sentiment to the sentiment of the Part containing intensifier

4. Else if (Pre= True)

5. Set Over All Sentiment to the sentiment of the Predecessor

6. Else if (Succ= True)

7. Set Over All Sentiment to the sentiment of the Successor

8. Return (Over All Sentiment)

9. Check if discourse exists.

10 .If (True)

11. Split into predecessor and successor parts.

12. Else

13. Send input to the Opinion Mining Model

c. Discourse Identification

1. It first checks for discourse word in the sentence, using function check_discourse. That function returns the discourse word, along with its type (pre,succ).

2. If discourse is present, the program splits the sentence on the conjunction to get two clauses.

3. It then checks for presence of Intensifier in either clause.

If Both clauses have intensifier, then intensifier affect is cancelled

Else if one clause has intensifier than the sentiment of that clause is use to determine the sentiment of the sentence.

4. If intensifier is not present, it uses the first or second clause based on the type of conjunction present (pre or succ)

5. If no conjunction was present in Step 1 here, the sentiment of the whole sentence is calculated and returned.

d. Negation (nahi) Handling:

1. Sentiment is calculated via get sentiment function. And it is supposed to receive a list of words.

2. Sentiment is initialised with a 0.

3. Boolean ‘nahi_occured’ is initialised with false.

4. For every word, it first checks whether it is a ‘nahi’. If it is, the boolean nahi_occured is made TRUE.

5. It then checks for 2-gram (and also 1-gram) occurrence of that phrase (or word) in the neg and pos data structure it initially made from those files. The value of sentiment is added or subtracted based on that.

6. At the end sentiment value is changed due to addition and/or subtraction by the words.

7. If nahi_occured is TRUE, the polarity is inverted by multiplying the sentiment by -1.

8. This sentiment is returned to whatever function called it.

A detailed discussion of all the results and findings is presented in the following section.

6. RESULTS

In order to carry out our research, we employed a dataset consisting of more than 22,604 statements in Roman Urdu. Out of these statements, around 5083 records were the ones that contained discourse element, whereas 469 records contained intensifiers. We have used F1 score to evaluate the accuracy of results. F1 score (also referred to as F-score or F-measure) is used in statistical analysis of binary classification, to measure a test's accuracy. This measure considers precision 'p' and recall 'r' of the test to calculate the score, where, p is the number off correct positive results divided by all positive results and r is the number of correct positive results divided by the number of positive results that should have been returned. The F1 score is actually a weighted average of precision and recall, when the score reaches to 1, it is the best value whereas the worst score is at 0. The formulas used for the calculation of Precision and Recall are as follows and the values for these attributes are presented in Table 4 (where Pos=positive, Net=neutral and Neg=negative).

For Recall:

recall_pos = (PosPos/(PosPos+PosNet+PosNeg))

recall_net = (NetNet/(NetPos+NetNet+NetNeg))

recall_neg = (NegNeg/(NegPos+NegNet+NegNeg))

recall = (recall_pos + recall_net + recall_neg)/3

For Precision:

precision_pos = (PosPos/(PosPos+NetPos+NegPos))

precision_net = (NetNet/(PosNet+NetNet+NegNet))

precision_neg = (NegNeg/(PosNeg+NetNeg+NegNeg))

precision = (precision_pos + precision_net + precision_neg)/3

Table 4 F1-Scores for Discourse Based Sentiment Tagging

Sentiment	Precision	Recall	F1-Score
Positive	0.92	0.73	0.81
Neutral	0.88	0.96	0.92
Negative	0.9	0.86	0.88

Table 5 presents the confusion matrix for analysing false positives and false negative cases.

Table 5 Confusion Matrix for Discourse Based Sentiment Tagging

	Positive	Neutral	Negative
Positive	3626	1173	162
Neutral	237	12641	261
Negative	75	566	3863

The developed model gives quite satisfactory results. Table 5 shows significant improvement in the F1-scores of all three classes of sentiment giving an overall Precision of 90 percent, Recall of 85 percent and F1-Score of 87 percent as compared to the results given by the baseline classifiers. Table 6 shows examples of instances where the human annotated value differs from the model predicted values. If we consider the second instance of Table 6 it can be observed that a human would consider this a positive statement based on his belief and perception, but the model does not find any term in the sentence that leads to a positive sentiment interpretation and thus predicts it as neutral. Similarly, in the fifth instance although the tone of the sentiment is negative and that’s understood by the human but use of sarcasm is not detected by the model that predicts the sentence to be positive. Factors like perception, experience, and sarcasm can influence human annotation whereas the machine dependent models do not take these into considerations while predicting.

Table 6 Sample of Contradicting Actual and Predicted Sentiment

S#	Comment	Human Annotated	Predicted
1	Iske monitor ke peche belt hay jo aap apnay hath pe lagaa kar ba-asaani apna blood pressure note kar saktay hai	Positive	Neutral
2	Quran kehta hy k shaeed Allah k pas zinda he	Positive	Neutral
3	ye ke humari hukumat ke khilaf aala satah par corruption ka koi case samne nahi aya	Neutral	Negative
4	Wo bunyadi taur par roshan khayal thay	Neutral	Positive
5	agr ap bhi Khursheed shah ki comittee k zariy regular hooy hoty to????? ap apni promotion ure ziada noto/peso ki khatir doosro ka rozgar hee khatm krna chahty hain wah kia bat hai.	Negative	Positive
6	first time me hi silai khul gai or kapra bilkul dheela hogaya medium se large hogai	Negative	Neutral

Sentiment analysis or opinion mining is the automated extraction of writer's attitude from the text and is one of the major challenges in natural language processing. Adequate sentiment analysis is not just about understanding the overall sentiment of a document or a single paragraph, but it is also important to be able to extract sentiments on a very granular level and relate each sentiment to the aspect it corresponds to. On the more advanced level, the analysis can go beyond only positive or negative attitude and identify complex attitude types. We, therefore, developed a model for performing discourse-based opinion mining, so we could also consider the impact that various discourse elements have on the overall sentiment of the text. Our study is specifically focused on datasets in Roman Urdu that are extracted from different social media websites. The results revealed a high percentage of success rate based on F1 score. ***This was established on the basis of comparison done with baseline classifiers as no such comparison could be established from existing literature on Roman Urdu.*** Better results can be achieved by introducing more enhanced rules and other machine learning techniques.

7. CONCLUSION

Sentiment analysis or opinion mining is the automated extraction of writer's attitude from the text and is one of the major challenges in natural language processing. Adequate sentiment analysis is

not just about understanding the overall sentiment of a document or a single paragraph, but it is also important to be able to extract sentiments on a very granular level and relate each sentiment to the aspect it corresponds to. On the more advanced level, the analysis can go beyond only positive or negative attitude and identify complex attitude types. We, therefore, developed a model for performing discourse-based opinion mining, so we could also consider the impact that various discourse elements have on the overall sentiment of the text. Our study is specifically focused on datasets in Roman Urdu that are extracted from different social media websites. The results revealed a high percentage of success rate based on F1 score. Better results can be achieved by introducing more enhanced rules and other machine learning techniques.

There are also certain limitations in the paper. First being the size of the data corpus. This is a problem to solve for doing scale sentiment classification by applying big data techniques. Besides that, the data collected was very unstructured and noisy and it contained a lot of typos and abbreviations. It is desirable to auto correct all the typos and to extend the abbreviations to regular words, which requires high level Natural Language Processing techniques. The normalisation of text data can be done using one of the other techniques discussed in detail in the literature review section. Formulation of rules needs to be improved. Analysis can be made more accurate by processing multilingual text present in the dataset. Taking emotion icons and sarcasm into account can also significantly improve the accuracy of results. Finally, there are many types of discourse elements of which we have considered only a few.

BIBLIOGRAPHY

[1] Mukherjee, Subhabrata, and Pushpak Bhattacharyya. , "Sentiment analysis in twitter with lightweight discourse analysis.," in *Proceedings of COLING 2012*, 2012.

[2] "Computational Discourse," [Online]. Available: http://www3.cs.stonybrook.edu/~ychoi/cse507/slides/06-discourse.pdf.

[3] Dey, Lipika and Haque, Sk., " Opinion Mining from Noisy Text Data," *International Journal on Document Analysis and Recognition 12(3).* , pp. 205-226, 2009.

[4] Go, Alec, et al., "Twitter sentiment classification using distant supervision," *CS224N Project Report.*, 2009.

[5] Marcu, Daniel, "The Theory and Practice of Discourse, Parsing and Summarisation," MIT Press, Cambridge, M.A., 2000.

[6] Wolf, Florian and Gibson, Edward, "Representing discourse coherence: A corpus-based study.," *Computational Linguistics,* p. 249–287, 2005.

[7] B. a. J. A. a. B. P. AR, "Harnessing WordNet Senses for Supervised Sentiment Classification.," in *In Proceedings of Empirical Methods in Natural Language Processing (EMNLP)*, 2011.

[8] Grosz, Barbara J. and Candace L. Sidner, "Attention, intentions, and the structure of discourse.," *Computational Linguistics,* p. 175–204, 1986.

[9] Hobbs, Jerry R., "On the coherence and structure of discourse.," Center for the Study of Language and Information (CSLI), Stanford, C.A., 1985.

[10] Syeed Ibn Faiz and Robert E. Mercer, "Identifying Explicit Discourse Connectives in Text," in *Canadian Conference on Artificial Intelligence.*, London, ON, Canada, 2013.

[11] Pitler, Emily, et al., "Easily identifiable discourse relations.," *{Technical Reports (CIS),* p. 884, 2008.

[12] Graff, " English gigaword corpus," 2003.

[13] Miltsakaki, Eleni, et al., "Experiments on sense annotation and sense disambiguation of discourse connectives.," in *Proceedings of the Fourth Workshop on Treebanks and Linguistic Theories (TLT2005),* 2005.

[14] Ziheng Lin, et al., "Recognizing Implicit Discourse Relations in the Penn Discourse Treebank," in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1. Association for Computational Linguistics.*, Singapore, 2009.

[15] Mukherjee, Subhabrata, and Pushpak Bhattacharyya., "Sentiment analysis in twitter with lightweight discourse analysis.," in *Proceedings of COLING*, 2012.

[16] Al-Moslmi, Tareq, "Enhanced Malay sentiment analysis with an ensemble classification machine learning approach.," *Journal of Engineering and Applied Sciences,* pp. 5226-5232, 2017.

[17] T. Z. Zhao, "Learning discourse-level diversity for neural dialog models using conditional variational autoencoders.," arXiv, 2017.

[18] B. Pan, " Discourse marker augmented network with reinforcement learning for natural language inference.," in *Proceedings of the 56thAnnual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers).*, 2018.

[19] C. Bothe, "Discourse-wizard: discovering deep discourse structure in your conversation with RNNs.," arXiv preprint arXiv:1806.11420., 2018.

[20] RAFIQUE, Ayesha et al. Sentiment Analysis for Roman Urdu. **Mehran University Research Journal of Engineering and Technology**, [S.l.], v. 38, n. 2, p. 463-470, apr. 2019. ISSN 2413-7219. Available at: <https://publications.muet.edu.pk/index.php/muetrj/article/view/977>. Date accessed: 06 july 2020. doi: http://dx.doi.org/10.22581/muet1982.1902.20.