IOSR Journal of Mobile Computing & Application (IOSR-JMCA)
e- ISSN: 2394-0050, P-ISSN: 2394-0042.Volume 6, Issue 2 (Mar. - Apr. 2019), PP 01-10
www.iosrjournals.org

Sentiment Summarizer for Opinionated Symbiotic Discussion Chains under Social Networks

Mohammed Tune [1 *], Abdulkadir Ahmed [2]
[1*]*Department of Information Systems, College of Computing, DebreBerhan University, DebreBerhan,Ethiopia*
[2] *Department of Information Technology, College of Computing, DebreBerhan University, DebreBerhan, Ethiopia*
Corresponding Author: Mohammed Tune

Abstract: The *sentiment discourse (i.e. responses and replay) on the social Network sites such as, Facebook, Twitter, YouTube, Forums, etc., forms the outbreaks of ample of opinionated thread chains. To get a complete message, every single opinionated text under these chains has to be seen interdependently. However, it is problematic to get a complete message from such threaded opinion chains, solely by applying the state-of-the-art computational linguistic techniques being utilized under opinion mining. In this paper, an opinion-oriented graph-based summarizing model from an opinionated discourse text of social network site is proposed. The major novelty of this paper is the use of back-trace enabled rule based opinion-oriented graph approach. Experiments are conducted and it has confirmed that the proposed model provided an encouraging result, which cannot be managed easily by the use of state of the art approaches. We put forward on the use of machine learning techniques to enhance the efficiency of the developed model.*
Keywords*: back-tracing; discourse; opinion oriented-graph; social network; thread*

Date of Submission: 08-05-2019 Date of acceptance: 24-05-2019

I. Introduction

With the emerging of the internet as an important source of information, users are able to actively express their opinions and exchange their experience on different hot-topics regarding social, political, sports, religion, businesses etc., aspects**[1] [2] [3].** Currently, Social networks (Facebook, Twitter, YouTube, Forums, etc.) are becoming a great source for discourse analysis. Facebook is one of the most popular Social Network in the world with more than 2 billion active users per month. It offers the possibility of collecting posts appearing in the form of discussions, debates, agreement-disagreement discourse **[4] [5] [6].** These chains of discussions cause information overload which can create difficulty in identifying relevant information for a political analyst, journalist and other stakeholders. As a result, it pays attention for discourse analysis in the area of opinion mining and social networking. The structure of posts on social networks like Facebook is mainly a kind of discussion thread **[7].** It typically consists of original postings (parent-node) and a plenty of additional postings (child node) that are publicly responded to original posts or responses to responses **[8].** These responses are basically conveying two types of massage between discussion posts.

The first discussion post is the one that conveys a positive message or reflects an agreement with the prior post and the second response type conveys a negative message or reflect a disagreement to the prior post. These supporting and contrasting discussion posts aimed to respond to the prior post have created a long discussion chain. Hence, as the discussion chain continuously increases making decisions is become quite complex. It is difficult for one's to tell the total tendency of a discussion thread. Moreover, it is also time consuming and boring to analyze which opinion is provided with what replay and what are the corresponding opinion relations. Understanding and back-tracing the meaning of such type of opinion thread profoundly require focused attention and more time.

This paper narrates the overall sentiment summary of the discourse posted on publicly available message about specific events. Recently a plenty of discourse analysis research in the area of social networking and opinion mining had been done and are being undertaken using machine learning and social network techniques. From the task of machine learning, Opinion Mining or Sentiment Analysis is a process of automatic extraction of opinion described in the form of positive, negative or neutral about a text in discussion topic **[9].** The opinion mining methods are explored in recent Literature mainly rooted in, natural language processing, computational linguistics, and text mining to determine the sentiment polarity of a text at a different level of granularity namely sentence, document, word or phrase and aspect **[10], [11], [12], [13].** A typical approach to these methods is to use frequencies of positive and negative words in order to determine whether a discourse is predominantly positive or negative **[14], [15].** Such an approach ignores the hierarchical structure of discourse text, whereas this hierarchical structure of discourse posts carries valuable information that tells the inter-relationship of posts. As a result, the individual opinion polarity identified by the methods does not carry out complete topic information, out of discourse text. Yet, using knowledge obtained from this hierarchical structure of discourse text is a relatively unexplored direction of discourse sentiment.

E.g. in discourse 1 "*Ethiopia has received new prime minister*". Comment 1.1"*I am happy with him, he is the right delegate.*" reply 1.1.1. "*I don't think any change from ruling coalition*" reply 1.1.2. "*You are right"*replay 1.1.2.1 "*it is bad to say this; the candidate is from reformist element*" In this discourse: the reply on the order lists 1.1.1 And 1.1.2 have shown disagreement to the initial reply1.1; However, the opinions given in reply 1.1.2.1 shows disagreement with the previous two replies, as a result, this reply shows indirectly agreement (positive) with initial post 1. But if we consider this reply message out of discourse it shows disagreement (negative) message. Here, a sequence of a positive message followed by a negative does not necessarily show a negative message between comments and vice versa.

This problem cannot be handled with the state of the artcomputational linguistic or text mining techniques, as the method does not consider the relationships between opinions orientations from parent to the child node, rather it determines a polarity of the individual opinion post. Thus, it is better to explore a new method which determines the total sentiment relationship in a discourse discussion thread. The new method requires adequate rules offered by experts and annotated text polarity. The focus of this article is on the presentation of an opinion-based model that facilitates the discussion analysis and not on the ways to identify opinion data. Basically, the method of opinion-oriented graph analysis requires proper data structure, storage, and representation technique. It is easy to determine the overall summary of discourse sentiment from a huge amount of collected text from Facebook reviews, prior to making a decision

II. Related works

We reviewed various related works from the concept of machine learning and social network techniques for opinion mining from discourse posts. For the task of machine learning approach, Opinion Mining or Sentiment Analysis from discourse text is defined as a process of extraction of sentiment described in the form of positive, negative or neutral about a particular topic or problem **[9].**As stated in **[13], [14], [15]** Sentiment analysis techniques can be roughly divided into the lexicon-based methods and machine-learning methods **[16]**. Lexicon-based methods rely on a sentiment lexicon, a collection of known and pre-compiled sentiment terms. Machine learning approaches make use of syntactic or linguistic features **[17], [18]** to find out sentiment statics of opinions sentence. In these, we believe that determining the meanings of discourses opinions are challenging for the method as the sentiments or opinion meanings are determined by back-tracing of each opinion post. That is when looking opinion message in the discourse shown in Table 1, some of the sentiments are ambiguous from the view of computational linguistic compared to the actual meaning throughout discourses. Because it does not mean that all positive opinions are given to imply only for a positive message and a negatives opinion are given to imply only for a negative message. In the works [19], [20], [21], [22], [23]determine the meanings of opinion in discourse posts using connectives, i.e. cue words and phrases [24], [25]discussed probabilistic models for identifying elementary discourse units at the clausal level and generating trees at the sentence level, using lexical and syntactic information from a discourse-annotated corpus. However, most of these discourse-based works narrow their scope to detect the sentiment polarity of a single unit. However, discourse analysis work concerned with the actual meaning of a message at the whole discourse remains unseen dimension. The intention of this work is to tide and compute more than one discourses unit (sentences, posts) relationship in order to determine the final sentiment meaning used for the decision. In another hand Most of existing Social Network opinion mining work deals with the analysis of the relationships between entities in a social network like who is friends with whom, who are experts and who post reaction text, what is the central node, etc. In Fisher et al. [26] analyze newsgroups by applying Social Network techniques and they interpret online communities by assigning roles to the members of the groups. This is done by observing how people relate to each other in a graph-based model of post-reply relations. In [27] the authors represent a newsgroup as a user-based graph and they base their analysis on the "reply-to" links between the users. The most related study withour work is the work of Stavrianou et al [8]. They propose a framework for discussion analysis by combining Social Network and Opinion Mining techniques and they study the structure of an online debate and analyze the user reactions, preferences and opinions on a certain subject, by combining user-based graph and opinion-based graph.

The main objective of the author was to enhance the user-based graph with additional opinion information. The authors also proposed a measure to analyze discussion thread mainly considering relationship resides between entities rather than opinion chain. They did not consider the relationship between message nodes resides in the threads. Moreover, they also did not consider the intra-relationship present between thread chains which are important to express the aggregate opinion polarity. Our work is the extension of the work of author's [8] by providing more improved opinion thread measure techniques and a well-defined set of rules that can enhance a thread discussion analysis. In general, the following aspects were considered in this study:

- We include more detail data structure concepts for the data storage and representation of opinion polarities which is important in the opinion-oriented graph model.
- An attempt was made to develop a graph-based opinion summary model from opinionated discourse posts.
- We proposed a measure and rules that consider the additional variables, the relationship between message contents (opinion polarity presented between each node).
- We consider also the intra-relationship resides between thread chains that enable us to overlook the total summary of threaded opinion polarity.

III. Methods

3.1 Dataset

The data-sets used for the experiments were crawled from the Facebook site dated between March 27, 2018, and April 3, 2018. We crawl a total of 49 threads of discussions, containing 854 comments and 280 replies, about the Current Ethiopian politics. In order to make the annotation simple for linguistic men, the irrelevant comments were removed.

3.2 Proposed Architecture

The general architecture of the model is shown in the figure. 1, the system contains different components. These components are a review of opinion data (web-crawler), linguistic annotation of opinion-reviews, design opinions-based graph, opinion measure techniques, and summary of opinion thread.

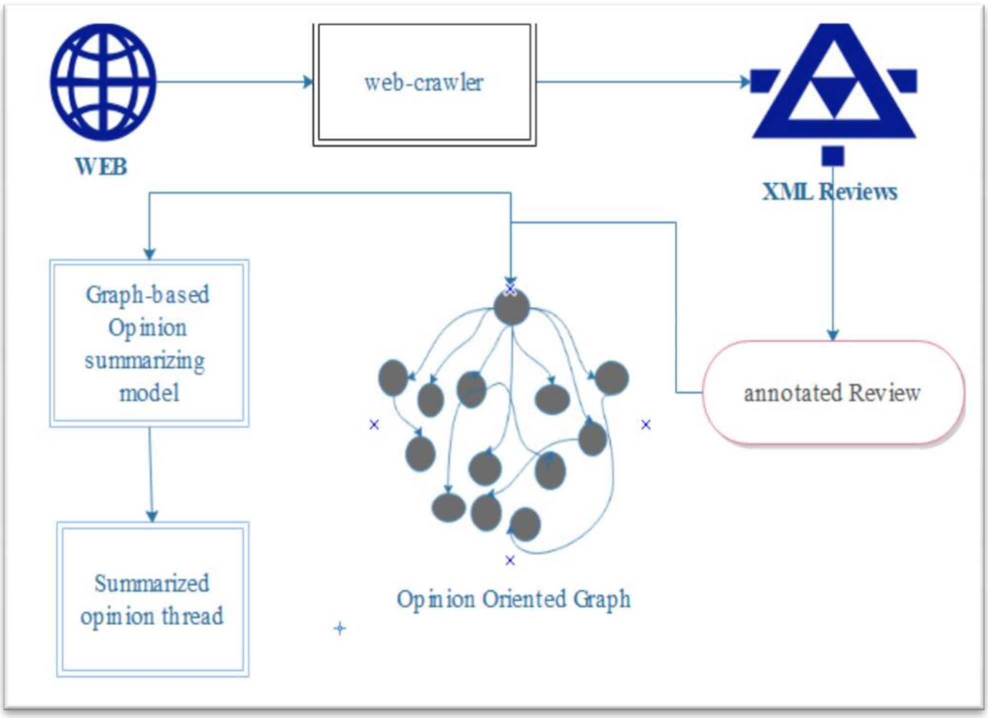

Figure 1: Proposed architecture of Graph-based opinion summarizing model (*Source: Own)*

3.3. Structuring opinion information

Different Facebook users can post thousands of messages; a design scheme is required to make sense of all the information exchanged on the Facebook page. The basic organizational unit of the Facebook post is the Thread of discussion. A thread is a collection of messages which address discussion topic declared in the first message called thread head [28]. In order to represent these thread chains, data structure tree node is used. This method is appropriate to show whether each message is a follow-up to the original post or to one of the replies arises from it. We used the Facebook Graph API to crawl public posts and its' associated public comments and replies. The Graph API allows us to navigate through the graph of the social network, which is organized into tree nodes. The crawled datasets are labels of text available in XML file format. The labeled XML file is converted into data structure nodes to indicate parent-child relationships. This node is directly mapped to a number list to refer to the parent of the reply (child node). The parent-child node has created a forest tree structure, for this we applied the data structure storage technique linked list, in order to summarize of threaded opinion polarity. However, before opinion summarization by the use of graph theory (tree structure), for crawled labels of opinion text, a very basic preprocessing phase has been applied to the corpus before linguistic annotation. The ambiguous text expressed in idioms, slang, Misspellings, Laughter, neutral opinions are ignored and only the structured text of thread opinion has been given to linguistic expertise to annotate the thread texts as positive, negative and neutral. However, after annotation, we exempt neutral opinions, because, it has no any influence on decision going to be made so that it is not considered in our model as input. But the neutral opinion is seen only from an equal number of positive and negative opinions extracted in opinion chain while summarization. To illustrate this, Table.1 represents the structure of the annotated opinion thread.

Table 1: Structure of opinion thread

Opinions	Nodes	Opinion text	Polarity
Post	1	"Ethiopia gets new prime minister"?	Neutral
Comment	1.1	"I am happy with him, he is the right delegate"	Positive
Reply	1.1.1	"I don't think any change from ruling coalition"	Negative
Reply	1.1.2	"you are right"	Positive
Reply	1.1.2.1	"it is bad to say this; the candidate is from reformist element"	Negative

To structure the corpus of opinions thread, the following three-fold hierarchical structure of Facebook discussion were considered. Post: it is the main issue about which user's posts or it is the topic/issue that requires different comments from different users to make decisions. It is clear that the suspense content of the post has been often neutral, but this post can inspire the user to create long discourse. Comment: is a reaction text written on the front page of a public Facebook page for the initial message post. Replies is a hierarchy of sub-comments written to comment about a given message post. It is a special mention from one user in response to another user's comments or replies. Opinion: is the sentiment behavior towards a previous post or parent post either as positive or negative. Another issue related to this concept such as share, like, unlike, friendship and tags are points that will be considered in our new dimension of future research work.

1.3. Opinion-oriented model

Most graph-based existing works, consider users to be the vertices of the graph [8]. In this study, we suggest using polarity of message objects as the vertices. Because in Data structure, vertices are representing the nodes, these nodes are used to store opinion polarity. We represent this framework as "opinion-oriented graph", whose definition is as follows.

***Definition 1**:* Opinion-oriented graph (OOG) is a graph $G = (T_n; R_n)$ in which T_n indicates an initial node and a set of T_n reply node. The thread node T_{ni} represents a "message object" or thread head and its weight values are (0). Each reply$R_{nij} = (r_{ni}; r_{nj})$points out direction from r_{ni} tor_{nj} , and it is weighted by a value that represents the opinion polarities expressed in the message objectr_{ni} as a reply to what it has been posted in the message objectr_{nj} .

The weight is a function w: $(R \rightarrow Z)$ and it takes negative values when the opinion polarity is negative (-1), and a positive (1) value when a positive opinion is expressed. An opinion-oriented graph consists of opinion orientation, which allows us to define measures in order to extract useful information from such graphs. In this study, we discuss three basic measures at; thread level, at thread chains, and at the node level.

A Message Object $Tn_{x,} \in N$ may be replied during the discussion chain. These posts may contain the opinions of the respondent expressed as positive or negative. Opinion measure per node is determined by the level of opinion chain. This level is defined from *($L_1 \rightarrow L_n$)* which implies the flow of opinion polarity from root to leaf node. Level one(L_1)Opinions are a direct post for root node (discussion point). It does not have an opinion chain. It is measure is straightforward it can be solved by the means of computational linguistics. But the replies posted starting from level tow or second node of opinion chains to L_n level node is indirect opinion posted about discussion point. Measure it is done only for nodes, the relation is going to be determined at chin level. These level opinions measures identify individual opinion polarity; it is objective to determine only sentiment polarities of the nodes.

1.3.1. Opinion Relations measures between chains

The second opinion measure task is done the through sway of opinion chain. The objective of this measure is to determine the sentiment relations throughout discussion chains. This measure is identified starting from level two ($L_1 - L_n$). A discussion chain $G_{C,}$= (N_C;R_C), $R_C \in E_C$ in the graph G is a path whose starting node is a root and ending node is a leaf. In this we defined the opinion received by a message object $N_{x,}$ from the root node to the leaf node as:

Reply:($N_{X,}r$) =$r_1 * r_2 \ldots\ldots r_i$, where, $r_i \in R_C$ and $R_C \in E_C$ (1)

In general, to determine the final summary of opinion chain computed from these relations, we identified the following Rules. Consider from figure 2.

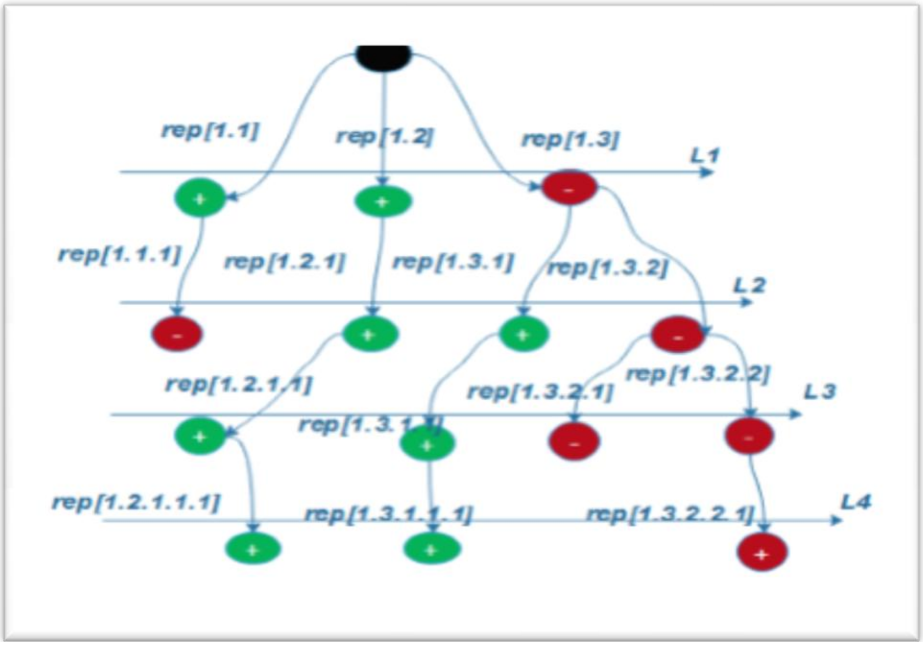

Figure 2: opinion-oriented graph (source own).

Rule 1: if all reply in the opinion chain is positive, the product of reply in a sequence of opinions chain will give positive opinion summary (+). It points out an agreement between the reply's node regarding a discussion topic. E.g.: from the above opinion thread, opinion chain $\{rep_{[1.2]}, rep_{[1.2.1]}, rep_{[1.2.1.1]}$ and $rep_{[2.1.1.1]}\}$ indicates positive reply, which means there is no disagreement between each reply post throughout the chain.

Rule 2: if the number of replies in the opinion chain is odd and their annotated polarity is negative the product of reply will be negative.

For all Replies $R_{i=-ve,}$ the, $\prod_{i=1}^{n} R_i = \begin{cases} -ve for n is odd node \\ +ve for n is even node \end{cases}$ (2)

From Thread above the opinion chain $\{rep_{[1.3]}, rep_{[1.3.2]}$ and $rep_{[1.3.2.1]}\}$ are the three replies annotated as negative. But it does not mean that all reply implies negative opinion about the message object, $rep_{[1.3]}$is a reply for negative, $rep_{[1.3.2]}$is a reply for positive, it indirectly supports the message object, $rep_{[1.3.2.1]}$ is a reply for negatives. It indirectly antagonized the message object. In general, the number of reply message that indicates the negative idea about the topic is greater than that of a positive one. The product between these opinions also gives a negative result. Therefore, the assumption made by this Rule will provide a negative conclusion.

Rule 3: If the first reply is negative and the remaining reply in the chain is annotated as positive, the product of all replies throughout the chain will be negative.

Example: in the opinion chain $\{\mathrm{rep}_{[1.3]}, \mathrm{rep}_{[1.3.1]}, \mathrm{rep}_{[1.3.1.1]}$and $\mathrm{rep}_{[1.3.1.1.1]}\}$. The first reply$\mathrm{rep}_{[1.3]}$, is posted as having a negative opinion to the root node. The remaining replies are annotated as positive opinion, which implies that they are the supporters of the first reply. Hence, they convey indirectly negative message to the root message. Thus, the product throughout the opinion chain in this assumption will give negative result so this assumption is convenient in deciding the final conclusion as negative.

Rule 4: if the number of replies in the opinion chain is even and their annotation polarity is negative the product of all reply throughout the chain will be positive.

Example: The opinion chain $\{\mathrm{rep}_{[1.3]}, \mathrm{rep}_{[1.3.2]}, \mathrm{rep}_{[1.3.2.2]}$and $\mathrm{rep}_{[1.32.2.1]}\}$ are the four replies annotated as negative. In this even though they annotate as negative the opinion idea they express is different. Here the $\mathrm{rep}_{[1,.3]}$and $\mathrm{rep}_{[1.3.2.2]}$ are show negative idea about the root message but the$\mathrm{rep}_{[1.3.2]}$ and $\mathrm{rep}_{[1.32.2.1]}$ show positive idea. In this rua le, the product of all replies throughout the chain will give positive conclusion but the numbers of replies that convey positive and negative opinion are the same so it is inconvenient to take it as a positive conclusion. Due to this rule and other combinations of opinion chains the opinion measurement techniques is improved from the product to the average sum of the product. In this new measure the final conclusion of Rule 4: is taken as a neutral reply. For this, we defined a new model that measures the average opinion received by a message object $v_{x,}$ from the root node to the leaf node as:

$$\text{AvgMsgOpinion}(N_{x,}) = \frac{\sum_{i=1}^{n} (N_{x,(}\pi R_{i,})}{n} \quad (3)$$

Where $(\pi R_{i,})$ indicates the Reply of the node $N_{x,}$and it shows the products of replies across a series of opinion chain from first reply-to last reply (leaf node), n is the total number of replies travels from the root node to leaf ntheode. The average Message opinion towards a message object is an indication of the polarity of the discussion chain toward the specific post.

$$\text{AvgMsgOpinion}(N_{x,}) = \begin{Bmatrix} -1 \text{ if AvgMsgOpinion} < 0 \\ 0 \text{ ifAvgMsgOpinion} = 0 \\ 1 \text{ if AvgMsgOpinion} > 0 \end{Bmatrix}$$

For instance, in the opinion-oriented graph of Figure2, we can observe that the thread has 5 opinion chains. The average message opinion and its polarities are shown in Table 2

Table 3: AvgMsgOpinion short discussion

Message no	Chain no	AvgMsgOpinion	Polarities
1	1	0	0
2	2	1	+
3	3	-1	-
4	4	-0.33	-
5	5	0	0

1.3.2. Opinion measure at the thread level

The last opinion measure in the model is the measure of the whole discussion chain. It involves a measure at the thread level. Here as the number of discussion threads, discussion chains, and replay-posts with mixed opinion polarities increases; determining the most interesting opinion information in the opinion-based graph has become quite complex. To simplify this complexity some global analysis is important. This measure is used to identify interesting opinion information at the whole. This opinion information is measured by a disorder of opinion polarity or entropy $Hn_{x,}$ and it defines the amount of opinion information held by a node n_x ∈N (that has been replied to), as:

$$H(n_{x,}) = -\sum_{i=-1,,1}^{n} \left(\frac{AvgMsgOpinion\ (n_{x,i})}{|replay\ (n_{x,})|} log \frac{AvgMsgOpinion\ (n_{x,i})}{|replay\ (n_x)|}\right) \quad (4)$$

The opinion information is an indication of the replies of opinions received by a node. If, for instance, a node has received reply posts that are all of the same opinion polarity, then the entropy will be 0. This information is implying that: there is common opinion regarding the message object or post expressed under particular thread.Applying the model to bigger discussions with hundreds of Reply-posts is important. The table below depicts the real statics of sample data set collected fromtopic concerning Ethiopian politics. From this, the Entropy ($H(n_{x,})$) is computed to know the most popular messages of the discussion at a global level or for the entire discussion. Consider Table 4.

Table 4: sample statics of the data set and its entropy measure

Message post	**Reply-posts**	**Opinion Chains**	**AvgMsgOpinion**			**Entropy**
			(+)	(-)	(0)	$\boldsymbol{H(n_{x,})}$
$MsgPost_{[1]}$	8	8	6	2	0	0.244
$MsgPost_{[2]}$	9	5	4	1	0	0.217
$MsgPost_{[3]}$	32	28	18	8	2	0.278
$MsgPost_{[4]}$	29	24	18	5	1	0.235
$MsgPost_{[5]}$	10	9	2	7	0	0.230
$MsgPost_{[6]}$	13	12	8	3	1	0.267
$MsgPost_{[7]}$	22	17	9	2	6	0.255
$MsgPost_{[8]}$	35	34	29	5	0	0.181
$MsgPost_{[9]}$	16	14	8	6	0	0.296
$MsgPost_{[10]}$	8	6	6	0	0	0
$MsgPost_{[11]}$	24	23	23	0	0	0
$MsgPost_{[12]}$	2	2	1	1	0	0.301
$MsgPost_{[13]}$	2	1	0	1	0	0
$MsgPost_{[14]}$	4	4	3	1	0	0.244

$MsgPost_{[15]}$	66	40	21	15	4	0.306
Total	**280**	**227**	**156**	**57**	**14**	------------

The summary result shows that from 280 comments156 were summarized as positive, 57 as negative and, 14 as neutral, where 69% were "positive" 25% were "negative" and 6% were neutral. Here the neutral is computed from the occurrence of an equal number of positive and negative opinions. Here we can observe that about 19% of opinions are summarized between the three classes, it reduces the complexity of opinion chains into a minimum manageable level. The model summarizes more as opinion chain increased or become more complex. We also noticed that the$MsgPost_{[15]}$is the most popular message post; having 40 replies of which 21 were positive, 15 were negative. We notice that the $MsgPost_{[15]}$has the highest entropy of all, Indeed, this is the message post that has received replies with the highest variety of opinions. We also notice that the average opinion of all messages is positive which indicates the general tendency of discussion. In contrast to this we realize that the $MsgPost_{[8]}$is the least popular one having 34 replies of which 29 were positive, 5 was negative. We notice that the $MsgPost_{[8]}$has the least entropy of all. Here this is the message that has received a plenty of positive replies that shows the general tendency of the discussion as positive. In general, we conclude that the measure of entropies in the above result resides between 0 and 1. Hence,0 Entropy indicates all of the opinion posts are from the same polarity either positive or negative. This shows that the opinion of the discussion thread is not debating issue. However, thethemaximum entropy 1 indicates that the balance number of opinion polarities, this measure indicates that the issue under discussion is a hot topic or debating point. It is very important for the decision maker to give more emphasis on the issue. In this analysis, we considered only two polarities of opinions, positive and negative for calculating entropy measure because the neutral opinions are not important for this measure as it is already used for the equal number of positive and negative polarity in the model. In general, the average entropy of the above message objects is below 0.5, it indicates that the opinions are more of the same type, which is more of a positive opinion. As we observe in the above table, the entropy measures of the opinion posted for the political domain indicates users are posted more of positive opinion about the given topic.

1.4. Proposed Algorithm

Given annotated opinion thread, the proposed graph-based opinion mining mode operates in three steps. First, it reads the polarity of annotated opinion thread from the text file. Then create a tree structure that contains parent-child relationships. Then -Insert the polarity of a text into a created tree structure for the opinion thread. Next opinions measure is applied to the automatically created opinion-oriented graph. This is done in two ways. First is done at discussion chain and the second is the measure at the whole thread. Finally, all the polarity of annotated opinion thread is summarized into predefined categories: positive (+), negative (-) or neutral. The following algorithm is the high-level view of algorithms which describe, how tree structures are created, how opinion polarities are stored and how sentiment polarity values are summarized into its pre-defined class.

Algorithm: Back Tracing Algorithm (Source: own)

1. *for every annotated opinion discourse Thread* $\boldsymbol{D_T}$
2. *for every opinion polarity* O_P *of discourse thread* $\boldsymbol{O_p}$
3. *Read Its* O_P *of* $\boldsymbol{D_T}$
4. *If the* O_P *of a text is found in* $\boldsymbol{D_T}$
4.1 *create parent node*$\boldsymbol{p_n}$
4.2 *If* p_n *has a child node* $\boldsymbol{C_n}$
4.3 *Create a child node* $\boldsymbol{C_n}$
4.4 *Read its* O_P*then*
4.4.1 *Insert* O_P *to* $\boldsymbol{C_n}$
4.5 *Repeat from step 4.3 to create new* $\boldsymbol{C_n}$
4.6 *if new* $\boldsymbol{C_n}$ *is created*
4.6.1 ***then*** *Insert* O_P *to new* $\boldsymbol{C_n}$
4.7 *If the* $\boldsymbol{C_n}$*is leaf node*
4.7.1 *Computes the AvgMsg* O_P
4.7.2 *If AvgMsg*O_P *is > 0*
4.7.2.1 *Assign* O_P*to class (positive)*
4.7.3 *IfAvgMsg*O_P *is <0*
4.7.3.1 *Assign* O_P *to class(negative)*
4.7.4 *IfAvgMsg*O_P *is=0*

4.7.4.1 Assign AvgMsgO_P to class (neutral)
4.8 *else repeat from repeat step 4.3*
4.9. End
Finally, the concept in the above algorithm implements and summarizes the threaded opinion reviews as depicted in Figure 3.

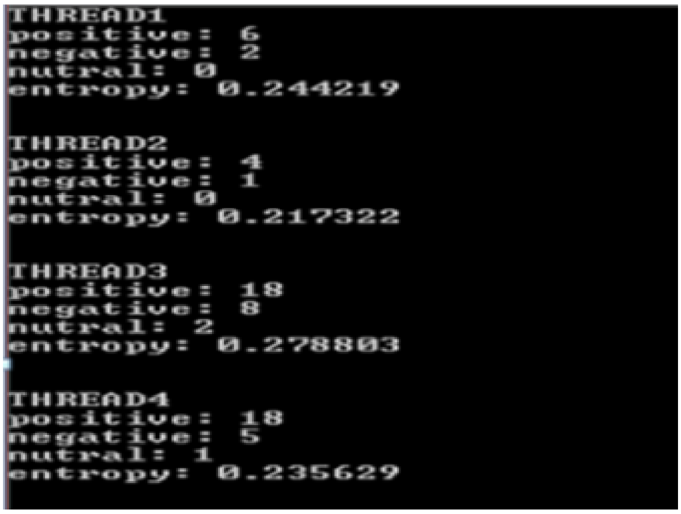

Figure 3: summary of opinion polarity (source own).

Here, as depicted in figure 3 the model generates the summary of replies in each opinion thread. In fact, in the absence of such a model, determining the final summary of even a single thread is not easy. Hence, the summarized opinion information can minimize the problem of information overload. This can support decision makers to identify quickly, the main interesting topic that they have to give emphasis.

IV. Results

In this section, we evaluate the proposed model. Typically, the evaluation is done manually, by comparing expert's annotated results in the presence of discussion threads. The following table 5 shows that the performance evaluation is done on 15 opinions thread provided by five annotators using Cohen's Kappa: to determine the agreement disagreement between annotators.

Table 5: annotation of opinion by experts

System	Reviews	Class	Number
In discussion chain	Annotator 1	Positive	160
		Negative	120
	Annotator 2	Positive	173
		Negative	107
	Annotator 3	Positive	165
		Negative	115
	Annotator 4	Positive	166
		Negative	114
	Annotator 5	Positive	180
		Negative	120
	Cohen's kappa	Positive	0.5
		Negative	0.58

The average annotation result done by all annotators from the total discussion is 60%, 40% positive and negative respectively. From the total dataset, 95% agreement is shown between the annotators.
These results were also tested by using machine learning developed online available tool that has better performance. Here, we used the known performance evaluation matrixes, precision, recall, and F-measure as depicted table 6.

Table 6: the summary of experimental result for opinionated comment from review site

Reviews	Confusion matrix			Performance Evaluation		
	Class	Positive	Negative	Precision	Recall	F-measure
Using machine learning (SVM)	Positive	142	24	0.84	0.85	0.84
	Negative	27	67	0.73	0.71	0.71

V. Discussion

The experiment shows that 74% accuracy when we compared these results with linguistics annotation the overall accuracy of the system is reduced result. The model shows about 67.5% of opinions are annotated as positive and 33.5% is annotated as negative. Whereas experts annotate 60% of it as positive and 40% as negative. The experimental results show that there is a thread off among positive and negative polarities compared to linguistic annotation. This is due to some of the opinions given to the initial posts are indirect or depend on previous one, the order in which these opinions text comes have an impact on the message it tends to expresses.

E.g.: if a reply text "*it is a right comment*" is followed by negative reply, it conveys negative message to the initial post, even though the sentiment polarity is positive, in another hand if a negative reply" *it is not a right comment*" is followed by negative comment it conveys positive message for the initial post, despite its' sentiment polarity which is negative. Due to such opinion text exist in positive but express negative and vice versa the total accuracy of the model is reduced. From these, we observe that the meaning of opinion and its' relationships in discussion threads are determined through back word look of the previous post. This cannot be resolved from the use of solely straightforward text mining or Computational linguistics in the absence of the hybrid graph model. The other challenge that affects the annotation result in both experiments is the use of an informal expression like proverb, pragmatics, idiomatic, slang, semantic, Misspellings, Laughter etc. as annotation result may vary from expert to expert in such ambiguous text.

VI. Conclusion and Recommendation

The web has dramatically changed the way that people express their views and opinions. One can now post reviews of the event at social Network sites and express their view on almost everything on tweeter, Facebook, forums, and blogs etc. on current hot topics like politics and sport, knowing what other people think is a determinant factor in decision making. In politics Journalist, politician and analysisist can find out the opinions of people, and political parties to analyze public view of government policies, elections, debates etc. Hence analyzing data from it can help in determining public view. However, it is difficult for one to make sense of all data in such domain for the required objective. As a result, automated graph-based opinion measure and summary systems are important. Thus, this article work employed the method of graph-based opinion summarizing model from the corpus of the opinion discussion thread that was selected from a different Facebook page about political discussion. The developed opinion measures model determines the sentiment orientation and generates aggregate relationships between discussion chains. Furthermore graph-based opinion-oriented model presents measure that cannot be solved by the use of computational linguistics from threaded opinions. And the model summarizes more, for a more complex opinion chain which is difficult to decide the final decision than simple opinion chain. The evaluation of the model shows that graph-based opinion summarizing model works hundred percent for any annotated opinion text done either by text mining or linguistic men. However, improving the annotation result prior to graph-based summary need further study due to the complexity of social network text.We put forward on the use of machine learning techniques to enhance the efficiency of the developed model.

Acknowledgements

We would like to extend our gratitude to Debre Berhan University staff for their invaluable support and motivation they have provided us to commence the study. We would also like to express our appreciations for the study participants and data collectors.

References

[1]. A. Ahmad. "Is Twitter a useful tool for journalists?" Journal of Media Practice, vol. 11, no.2, pp. 145–155,2010
[2]. A. Ahmad. "Whats in a tweet?" foreign correspondents use of social media. Journalism Practice, vol. 7, no. 1 pp. 33–46,2013
[3]. M. L. Sheffer and B. Schultz. Paradigm shift or passing fad? twitter and sports journalism. International Journal of Sports Communication, vol.3, no. 4, pp. 472–484,2010
[4]. Lucia C. Passaro, Alessandro Bondielli and Alessandro Lenci, "A Topic-Annotated Facebook Corpus for Emotion Detection and Sentiment Analysis",2015
[5]. Márton, "Beyond Sentiment: Social Psychological Analysis of Political Facebook Comments in Hungary", 2015.
[6]. A. Jacob, R. Sara, and M. Kathy, "Annotating Agreement and Disagreement in Threaded Discussion," Columbia University, 2005.
[7]. Somasundaran, Swapna, N. Galileo, G. Lise, and J. Wiebe, "Opinion graphs for polarity and discourse classification," in Proceedings of the 2009 Workshop on Graph-based Methods for Natural Language Processing, 2011.
[8]. A. Stavrianou, J. Velcin, and J.-H. Chauchat, "A combination of opinion mining and social network techniques for discussion analysis," ERIC Laboratoire, Université Lumière Lyon 2, 2012.
[9]. M. K. Jayashri Khairnar, "Machine Learning Algorithms for Opinion Mining and Sentiment Classification," International Journal of Scientific and Research Publications, vol. 3, no.6, 2013.
[10]. B. Liu, "Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies," in Morgan & Claypool, 2012.
[11]. B. Pang, L. Lee and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques", in In Proceeding of the conference on empirical methods in natural language, pp. 79-86, Philadelphia, US, 2002.

[12]. SwapnaSomasundaran, JanyceWiebe, and Josef Ruppenhofer“Discourse level opinion interpretation”, In Proceedings of COLING, 2008.
[13]. M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede,” Lexicon-based methods for sentiment analysis.Computational Linguistics”, Vol. 1, pp.1–41,2010.
[14]. Popescu and O. Etzioni. “Extracting product features and opinions from reviews”, In Proc. of the Conference on Empirical Methods for Natural Language Processing (EMNLP), pp. 339–346, Vancouver, Canada, 2004.
[15]. Scharl and A. Weichselbraun. “An automated approach to investigating the online media coverage of US presidential elections”, Journal of Information Technology and Politics, Vol. 5, no. 1, pp. 121–132, 2008.
[16]. Boy and M-F. Moens. “A machine learning approach to sentiment analysis in multilingual web texts. Information Retrieval”, Vol. 12, no. 5, pp.526–558, 2009
[17]. Pak and P. Paroubek. “Twitter Based System:Using Twitter for Disambiguating Sentiment Ambiguous Adjectives” In Proceedings of the 5th International Workshop on Semantic Evaluation, Association for Computational Linguistics, pp.436-439, 2010b.
[18]. Alec Go, RichaBhayani, and Lei Huang,“Twitter sentiment classification using distant supervision”, CS224N Project Report, Stanford,2009.
[19]. Carley, K.M., Diesner, J.: “AutoMap: Software for Network Text Analysis”, CASOS (Center for Computational Analysis of Social and Organizational Systems), ISRI, CMU,2005)
[20]. Van Atteveldt, W.H.: “Semantic Network Analysis. Techniques for Extracting, Representing, and Querying Media Content.” VrijeUniversiteit, Amsterdam, 2008)
[21]. Pitler, E., Raghupathy, M., Mehta, H., Nenkova, A., Lee, A., and A. Joshi. “Easily identifiable discourse relations.” In Proceedings of COLING, pages 87-90, 2008
[22]. Somasundaran, Swapna. “Discourse-level relations for Opinion Analysis.” PhD Thesis, University of Pittsburgh. 2010
[23]. Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A. K., and Webber, B. L. “The Penn discourse treebank 2.0.” In LREC. European Language Resources Association.,2008
[24]. Marcu, Daniel. “The Theory and Practice of Discourse Parsing and Summarization”, MIT Press, Cambridge, MA. 2000
[25]. Asher, Nicholas and Benamara, Farah and Mathieu, Yvette Yannick. “Distilling opinion in discourse”: A preliminary study. In Proceedings of Computational Linguistics (CoLing). 2008
[26]. Fisher, D., M. Smith, and H.Welser, “You are who you talk to Detecting roles in Usenet newsgroups”, Proceedings of the 39th Annual HICSS. IEEE Computer Society, 2006
[27]. Jindal and L. Nitin and Bing, "Identifying comparative sentences in text documents: in Proceedings of ACM SIGIR Conf," in on Research and Development in Information Retrieval, SIGIR, 2006,2006a.
[28]. Zhang, M. Ackerman and L. Adamic, "Expertise networks in online communities: Structure and algorithms," in In Proc. of the 16th International Conference on World Wide Web, 2007.

Mohammed Tune. " Sentiment Summarizer for Opinionated Symbiotic Discussion Chains under Social Networks." IOSR Journal of Mobile Computing & Application (IOSR-JMCA) 6.2 (2019): 01-10.