Sentiment Polarity in Translation

Thet Thet Zin
Faculty of Computer Science
University of Computer Studies
(Thaton)
Yangon, Myanmar
ttzucsy@gmail.com,
thetthetzin@ucstt.edu.mm

Abstract

Previous year, many researchers have been sentiment analysis on many focus languages. They analyzed and categorizing opinions expressed in a text. People express their opinions and feeling on social media as a daily routine. For sentiment analysis work, data plays an important role. Thus, social media become interested platform for opinion mining. On the other hand, low resource languages face less of sentiment resources (such as sentiment lexicon, corpus) than English language. It is needed to overcome language barriers and realize a sentiment platform capable of scoring in different languages when global opinion is need to decide something. In this paper, the expectations and limitations of machine translation in sentiment polarity task for Myanmar language is presented. We experiment with comments of particular news and general news that are expressed in social media news pages. Results show that sentiment transfer can be successful through human translation. This also demonstrates that translation from Myanmar to English has a significant eﬀect on the preservation of sentiment by using translation engine. This happens primarily due to nature of Language but the results show that machine translation quality plays the important role in this work.

Keywords: *sentiment polarity, machine translation, sentiment transfer, language barriers*

I. INTRODUCTION

During these years, researchers and business industries have developed a platform for detecting opinions about the products, markets, stocks and news as expressed in social media and other media sources. They extract content from social media sites, blogs, news sites and analyze sentiment associated to each content. By doing like this, researchers can know the needs of customers, condition of specific market in particular region, public opinion on government's activities, strong and weak points of competitors and can get other benefits for industries, governments and organizations. For opinion mining work for around the world, there should not be restricted to a specific language and locale even though content appear in different languages. For successful sentiment analysis work, the system should overcome language barrier and provide sentiment analysis across languages.

At the present time in Myanmar, most people use Facebook and they read news from Facebook's news pages. They also write their opinions as comments in relative news articles using Myanmar Language. But the time of widespread information flow, we also need to listen and know opinions from other countries on our country news and events. We have already known that people from other countries use English language or their mother language to communicate the global. Therefore, at this point translation plays main role in sentiment analysis work to understand global phenomena on target work. After translating their focus language to English and then doing sentiment analysis on it that can get more opinion analysis results for particular domain from global as much as the content available for the works. Moreover, many tools developed for English such as sentiment lexicon, normalization tools or other that are not available for other languages yet. However, we should not do translation of sentiment words of one language to English without knowing how translation impact on sentiment analysis work. We should observe translation effect on sentiment work. If the machine translation system is not very accurate, there will be noise in the polarity words translated from the source language.

In this work, human translation and google machine translation are used for analysis. Human translation is very expensive and it takes time for preparation. It is not possible for many source languages. But google supports bilingual translation works. High quality machine translation from a target language to English can eliminate the necessity to develop specific sentiment analysis resources for that language.

There are two news Facebook comments datasets for Myanmar language are used in this work. In the first one comments are extracted from 21st Century Panglong Conference news articles. The second one contains comments from general news articles. Firstly, sentiment analysis on target language is done. And then target corpus is translated to English using human translation and google machine translation system. Second, translated corpus is apply for sentiment analysis on English language. And then we investigate how sentiment polarity appear in these works. The paper is organized as follows. In the next section related work for sentiment analysis in translation is briefly described. Section 3 contains motivation and contribution of this work. Data collection and preparation is presented in section 4. Section 5 provides methodology and section 6 describe experimental results of this work. The final section contains a discussion of the obtained results, some remarks and issues that remain to be addressed and that we intend to investigate in future work.

II. Related works

As mention above, many researchers did sentiment analysis (SA) research on various languages. SA is the computational study of opinions, sentiments and emotions as they are expressed in text. A. Balahur and M.Turchi [1] expressed the impact of machine translation (MT) on sentiment analysis in French, German and Spanish. They employ three MT systems for comparison- Bing translator, Google translator and Moses. The performances of the sentiment analysis on original language and translated corpora were comparable. In the worst case, the performance difference reached 8%. The approach in [2] experiments that polarity-annotated datasets in English and Turkish for movie and products reviews domain. The authors concluded that the polarity detection task is not affected by the amount of noise introduced by MT. The publication [3] illustrate the impact of Machine Translation on sentiment analysis. The authors presented the development and evaluation of Spanish resources for the multilingual sentiment analysis tool SentiSAIL. And then they described empirically the impact of MT on sentiment analysis performance. The performance decrease in the worst case remained within negligible 5%. They drew conclusion as an outcome of the experimental setup is that substituting multilingual sentiment analysis by English sentiment analysis via MT may be an acceptable alternative. The authors of [4] present a lexicon-based sentiment analysis system in Spanish, called Sentitext, which used three feature sets- the dictionary of individual words, the dictionary of multiword expressions and the set of context rules. They concluded that multiword expressions are critical for successful sentiment analysis. Sentitext is also used in [5] to detect sentiments on Twitter messages in Spanish.

Work on multilingual sentiment analysis has mainly addressed mapping sentiment resources from English into morphologically complex languages or from target language to English language. The authors [6] translated Chinese customer reviews to English using a machine translation system. The translated reviews are classified with a rule-based system applied on English lexicons. A high accuracy is achieved by combining knowledge from Chinese and English resources. The publications [7] and [8] did sentiment analysis and normalization work on Myanmar language. In [7] researchers created domain-specific lexicon for opinion mining work. Based on experimental results that lexicon is suitable for analysis work on particular domain.

Recent research papers show that google translate has a good performance to European languages but relatively poor in Asian languages. Even thought, the effort in this paper takes a different direction as it evaluated English sentiment analysis applied to translated data from Myanmar language.

III. MOTIVATION AND CONTRIBUTION

In this golden age of social media, comments and posts on social media is a part of people’s daily routine. Social media allow people to express their feeling, opinion and debate for many users and many topics. In Myanmar, most people use Facebook for many purposes. Government and organizations post news on Facebook pages and every news journal and media groups also post news on their Facebook pages. Thus, many people read news from their favorite pages and write their opinion on news article as comments. Therefore, news from social media pages are attractive point for analysis of opinion of people. Most Myanmar people write comments using Myanmar language and some people use English. On the other hand, in the globalization age, we need to mine opinion of people from international news pages and should consider perception of people on Myanmar news articles from other countries. In the global news site, users use English or their native languages. Therefore, machine translation and sentiment analysis on English language plays a major role for capturing their opinions on news articles. In this paper sentiment analysis on news domain for Myanmar language is done. Myanmar sentiment words are translated into English language using Google translate. And then we did sentiment analysis on English language again. After these steps, we have surveyed translation effect on sentiment work. Another fact for this work is that sentiment analysis tool for English language can be more available for low-resources language such as Myanmar and other target language. Therefore, I motivated to analysis how sentiment is preserved after machine translation. In this work, sentiment annotated corpus for Myanmar language for news domain is updated. There are two datasets. One dataset, 21st century Panglong Conference, has already built. Another dataset contains comments from general news articles from popular news pages in Myanmar. Contributions in this paper are as follows: (1) Data preprocessing step and building annotated general news corpus of Myanmar language. (2) English sentiment labeled corpus is created. (3) Polarity is calculated using annotated corpus and point wise mutual information method. (4) we investigate which translations that preserve sentiment best. This is a crucial step towards multilingual sentiment platform for news domain.

IV. DATA COLLECTION AND PREPROCESSING

The internet is an active place with respect to sentiment information. From a user's view, people are able to post their own content through various social media and blogs. From a researcher's view, many social media sites release their API, allowing data collection and analysis by researchers and developers.

A. Data Collection

In this work there are two datasets. Data for these two datasets are extracted from eight Myanmar popular news Facebook pages according to socialbaker.com: 7Day News Journal, Eleven Media Group, BBC Burmese, MRTV-4, Mizzima-News in Burmese, The Irrawaddy-Burmese Edition, VOA Burmese News, DVB TV News. Detail data analysis and description of 21st century Panglong Conference dataset has already mentioned in [7]. It contains 27,337 comments from news pages. Average length of the comments in this dataset is 21 words and the average number of words is 238,532 words. Another dataset contains comments on general news articles from 1st September,2019 to 31st October,2019 extracted from pages. The second dataset contains 656 comments from 98 news articles and average number of words is 13,214 words. Average words of the comments are 15 words. For classification task, the text fall into one of the following three classes: positive, negative and neutral.

B.Data Collection

Nature of social media users, they write status and comments using informal text. The dataset is examined for better understanding of the nature of collected data from Facebook. According to analysis, they use abbreviating (short form or acronym), mix usages (Myanmar and English words mixing eg. today နေသာတယ်-It's sunny today), slang words(ကွီး-ကိုကြီး-brother), multiword expressions (အေးပါပါပါပါ-အေးပါ-ok), emotion icons, syntactic mistake and using myanglish (nay kg lar? -how are you? -နေကောင်းလား). There are four types of comments in collected data :(1) Comment does not contain sentiment words. (2) Neutral comments (cannot consider in classification work. In this work PMI value is zero, sentence is noted as neutral sentence) (3) Comments suitable for sentence level but it uses many conjunction words. (4) Comments contain sentiment words. Example and explanation of each type of comment is shown in Table 1.

TABLE I. TYPE OF COMMENT IN COLLECTED DATA

Comment Type	Example	Explanation
1.	ကဗျာတွေစာတွေရေးလို့နေ - writing poems and letters	Cannot extract sentiment words from comment.
2.	တကယ်-really?	Cannot conclude polarity for it.
3.	ကြိုးစားနေပါသော်လည်း အလုပ်လုပ်နေပါသော်လည်း ပြည်သူတွေရဲ့နှလုံးသားကို မရလျှင်ကြိုးစားသမျှသဲထဲရေသွန် - Even if you try hard but work, but don't get the hearts of the people, pour it into the sand	contains two positive sentiment words -ကြိုးစား(try hard), အလုပ်လုပ် (work hard),One negative words- မရ- not get But comment gives negative meaning

4.	ငြိမ်းချမ်းခြင်းရဲ့အရသာ-the taste of peace	Positive sentiment word (ငြိမ်းချမ်းခြင်း-peace)

C. Preprocessing

Online data have flaws that potentially hinder the process of sentiment analysis. Many preprocessing steps are applied on the available dataset to optimize it for further experimentations. There are three steps in preprocessing on collected data.

1) Data cleaning process: New types of words such as emoticons (☺ and <3), hashtags("#Bieber"), URLs, photos or stickers comments are presented in social media data. Data cleaning processing is done to eliminate the incomplete, noisy and inconsistent data.
2) Syllable segmentation process: Myanmar language does not have boundary word markers. Syllable segmentation is done on cleaning dataset by using syllable segmentation tool.
3) Words Extraction process: Syllable segmented dataset is tokenized applying the n-gram method by setting the minimum and maximum grams. According to our analysis result on 500 comments, bigrams is set as minimum and 5-grams as maximum gram.

In the previous publication works [7] and [8] detail process for online data preprocessing is presented.

V. METHODOLOGY

In this work, we are interested to study the impact of machine translation on Myanmar sentiment analysis. The whole experimental stage is outlined below:

1) Build the Myanmar general news dataset MM_data and use previous domain specific dataset MM_ Panglong_data.
2) Learn a SA system on the two datasets based on pre-learned manually annotated Myanmar SA lexicon. This is denoted as the MMSA_baseline.
3) Test the MMSA_baseline system on two test dataset MM_data and MM_ Panglong_data and compute its performance.
4) Translate MM_data and MM_ Panglong _data from Myanmar to English. The English version of these datasets are obtained.
5) Learn a SA system on the two datasets based on pre-learned manually annotated English SA lexicon. This is denoted as the MSA_MT.
6) Test the MSA_MT system on two test datasets and compute its performance.

7) Compare the performance of both systems MMSA_baseline and MSA_MT and draw and investigate some conclusions.

A. Pointwise Mutual Information (PMI) based Sentiment Classification

Pointwise mutual information (PMI) is used for sentiment classification. It is a measure of association used in information theory and statistics. In computational linguistics, PMI has been used for finding collocations and associations between words. For example, counting of occurrences and cooccurrences of words in a text corpus can be used to approximate the probabilities $p(x)$ and $p\ (x,\ y)$ respectively. This method calculates the PMI between two words to obtain numeric score. The formula is as follows.

$$\mathrm{pmi}(\mathrm{x};\mathrm{y}) = log\frac{p(x,y)}{p(x)p(y)} \quad (1)$$

$$\mathrm{pmi}(\mathrm{w1};\mathrm{w2}) = log\frac{prob(w1\&w2)}{prob(w1)*prob(w2)} \quad (2)$$

where, $prob\ (w1\&w2)$ is the probability of word1 and word2 co-occur in the comments. Sentiment orientation score is calculate using $pmi(word1, positive\ word)$ and $pmi(word1, negative\ word)$.

$$Scorepmi = pmi\ (word1, positive\ word) - pmi(word1, negative\ word \quad (3)$$

$Scorepmi$ is calculate for all phrases. If score value is positive then comment sentence is categorized as positive, if value is negative then sentence is categorized as negative and if value is zero, it categorized as neutral.

B. Creating Sentiment Labeled Data in Myanmar and English

Manual sentiment annotations were performed on the dataset. MM_ Panglong _data dataset has been annotated by previous work. In this work, only MM_data dataset is annotated sentiment label by manually. To annotate English sentiment label, we translated Myanmar words to English by human and using google translate. MM_ Panglong _data dataset is larger than MM_data dataset. Thus, we select 400 Myanmar sentiment words from MM_ Panglong _data for translation. And then we annotated sentiment label on these translated datasets. As mention above, there are three classes for annotating sentiment labels: positive, negative and neutral.

In the publication [7] has already created sentiment lexicon for 21st century panglong conference news articles. This lexicon is updated by adding some words from general news articles. English sentiment lexicons are freely available for research work. At the present time, we cannot apply these available lexicons. Thus, English sentiment lexicon is created by manually. This lexicon contains 101 positive words and 106 negative words. In the future, we will use freely available sentiment lexicons for English language and construct automatic sentiment lexicons by using the classification methods.

C. Impact of Translation on Sentiment Analysis

In this section, how to get translation words from Myanmar words to English words and how this translation is impact on sentiment analysis work is presented. There are two types of translation is done in this work. The first one is human translation and another one is by using google translation engine. Firstly, we need to know the quality of machine translation engine for Myanmar language. And then We can survey how English translation of Myanmar text alters or not in detection of sentiments. To evaluate translation quality for the translation engine, BLEU scores are calculated. Using human translation as the reference. By comparing machine translation to human translation, we can know how close it is to the human translation.

On the other hand, human translation work is very expensive and time-consuming process. Thus, human translation is done for randomly selected 400 comments for 452 words as a training dataset. At the same time, these selected comments are also translated by using google translation engine. Preprocessing for these sentences is done by segmenting the Myanmar sentence using Myanmar word segmenter which is described by [7]. Google translate is primarily designed to translate sentences; it can also provide one-word translations. Therefore, google translate is used to translate Myanmar words into English in comments.1-gram BLEU score for selected sentences is 0.801 and 5-grams BLEU score is 0.791.

The 1-gram is used to assess how much information is maintain after translation. Among other grams, 5-grams is the most correlated by human translation by human. At the present time, google performed a little poor quality in Asia language. However, BLEU is not sufficient for sentiment analysis. It only evaluates translation quality based on human translation. Therefore, we need to do other experiments for this work. Compare the sentiment labels assigned to the translated English text with manual sentiment annotations of the Myanmar text. The more similar the sentiment annotations are, the less is the impact of translation. Some extracted words are shown in table2 which are impact on the performance of sentiment analysis.

TABLE II. **SOME EXAMPLES WHICH ARE IMPACT ON THE PERFORMANCE OF SENTIMENT ANALYSIS**

comment, sentiment label	Translated by google	Translated by manually
အေးပါကွာ, *positive*	(It' s cool, *neutral)*	ok, positive
ပြောသွားတာ_က_(ကောင်းပါတယ်, positive)လက်တွေ့_ဘဲ_(လိုနေတာ, negative)	It' s good (positive) to say that It needs(negative) to be practical	good, positive need(negative) to be practical,
မလိမ့်တပတ်, *negative*	Not a week, *neural*	lie, negative
(ညာနေတာပါ, *negative*) _ကွာ	I' m right, *positive*	lie, negative
ခန့်ညားနေတာ, *positive*	It' s ugly, *negative*	Neat and tide, positive

Some Myanmar words need to normalized to improve in translation. Eg. (ဆာမ-sir (wrong translation meaning because spelling mistake in Myanmar word), if this word is normalized to the word-ဆရာမ(can get correct translation-teacher)). Out of vocabulary words in the translated text are marked as unknown words. Most of these words are slang words and abbreviating words. Sentiment class of the first comment in table2 changes positive to neutral because google translate directly by input words. But actually, it meaning is depend on content of articles. In the second comment, there are two sentiment words. Google translate gives correct translation for word by word translation for this comment. If the whole sentence is given as an input to google, translated meaning is different (missing sentiment word). (ပြောသွားတာက ကောင်းပါတယ်လက်တွေ့ဘဲလိုနေတာ-well that’s good). Therefore, in this work, we translate from Myanmar to English by segmented word by word. For the remaining three comments, translated words effect to decrease the performance of sentiment analysis. At the present time, human translation can give more accurate translated word than google from Myanmar to English language. But for the big dataset, it is very expensive, time consuming and not practical.

VI. Experimental Results

For the experimental results the two training datasets are used. The first one is 21st century Panglong Conference dataset named dataset1 which include comments from one activity, Panglong Conference article. The second one is comments from general news articles within two months which is named as dataset2. Dataset1 contains 17,337 comments and dataset2 contains 656 comments. Myanmar sentiment lexicon is constructed words extracted from 10,200 comments from both datasets. To evaluate the accuracy, we followed 3-fold and 5-fold cross validation process on these two datasets. For this experiment, we selected only 400 comments from two datasets. Because this experiment requires more manually work and human resources. To keep the time, we choose this small size of data comments.

Firstly, we investigate whether translations can maintain or not sentiment polarity from Myanmar text. For this work, sentiment lexicons for both languages are used. Firstly, sentiment analysis on Myanmar language is performed. 3-fold and 5-fold cross validation results are performed for investigation. To answer the question that human translation preserves sentiment score or not, we did sentiment analysis with human translated sentences.

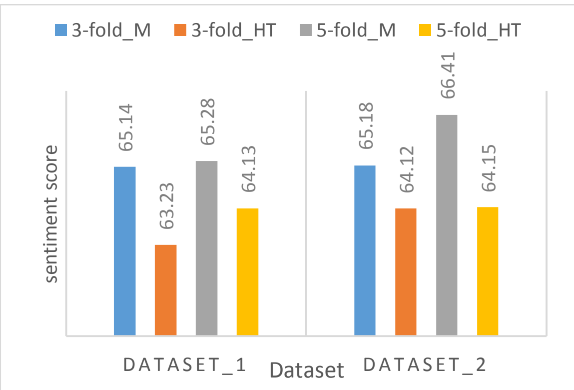

Figure 1. Sentiment Comparison: Myanmar and Human Translated English Language

If there is no significant difference between the sentiment score of Myanmar and English translation, we can conclude that the sentiment value did not change so much in translation work; if not the sentiment value is already lost in human translations. According to figure1, there is no significant difference between Myanmar and human English translation words.

According to 3-fold and 5-fold analysis results, difference between Myanmar sentiment score and human translated English sentiment score are not large. Thus, we can draw conclusion that human translation can preserve sentiment from original text. Second, we need to know the performance of machine translation engines on sentiment preservation. We compared the sentiment of the Myanmar with the sentiment of machine translation. We found that there were significant differences between this pairs. The results are shown in figure2.

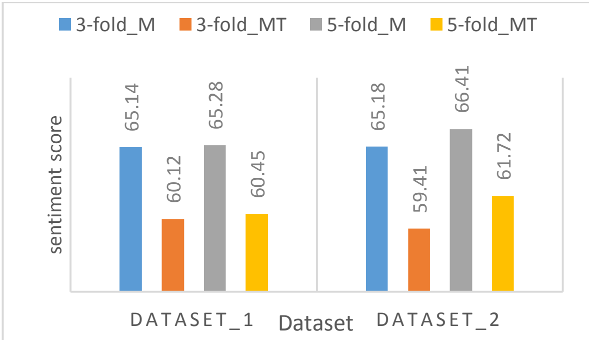

Figure 2. Sentiment Comparison: Myanmar and Machine Translated by google

In this experiment, the performance decrease in the worst case remain between 5% and 6%. We notice that both human and machine translation have lower sentiment score than original text. But we can ignore the difference between human translation and original text. One of the reasons for performance decreasing is that translations changes some words' sentiment to neutral value. Actually, these words have positive or negative sentiment values. Because translation lose content information of the articles. Other fact that complex nature of Myanmar language and most social media users use many slang words and grammar mistakes in comments. Moreover, other native languages have more than one machine translation engine such as Google, Microsoft. But Myanmar language has only one translate engine, Google. At the present time, the results indicate that translations generated by machine engine are not of the desired high quality and losing sentiment value. We can conclude that human translation is successful to preserve sentiment value. Thus, later we can use human translation as benchmark to compare other translation system.

VII. CONCLUSION AND FUTURE WORK

In this paper, we presented experimental results to study the impact of English machine translation on sentiment analysis of Myanmar comments on social media data. The results based on these datasets show that sentiment analysis on English translation by google translation engine is not good as sentiment analysis of Myanmar language at the present time. But English translation reach competitive results by using human translated English language. Thus, we observed that the quality of translation system is very important for sentiment analysis work on across languages. The important fact that this approach fails to take into account the divergence in the expression of sentiments across languages and content of articles. Moreover, according to this analysis 3-fold and 5-fold, accuracy is not completely depending on the size of training and testing data, but somewhere by increasing training data to get higher testing data accuracy. Other perspective is to investigate the use of deep learning classifiers, CNN in sentiment analysis work for better performance. In the future we will integrate other machine translation system and normalization process to enhance the performance and will evaluate the performance on data from multilingual social media platform. Moreover, we will use freely available English sentiment resources for comparative performances.

References

[1] A. Balahur and M. Turchi, "Multilingual Sentiment Analysis Using Machine Translation?" in Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis, ser. WASSA '12. Stroudsburg, PA, USA:Association for Computational Linguistics, 2012, pp. 52–60.

[2] E. Demirtas and M. Pechenizkiy, "Cross-lingual Polarity Detection with Machine Translation," in Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining, ser. WISDOM '13. New York, NY, USA: ACM, 2013, pp. 9:1–9:8.

[3] Commeignes, "The Impact of Machine Translation on Sentiment Analysis", The fifth international conference on data analytics, 2016, ISBN:978-1-61208-510-4.

[4] A. Moreno-Ortiz and C. P. Hern´andez, "Lexicon-based sentiment analysis of twitter messages in spanish," Procesamiento del Lenguaje Natural, vol. 50, 2013, pp. 93–100.

[5] E. Boldrini, A. Balahur, P. Martnez-Barco, and A. Montoyo, "Emotiblog: an annotation scheme for emotion detection and analysis in non-traditional textual genres," in Proceedings of the 5th International Conference on Data Mining (DMIN). Las-Vegas, USA: CSREA Press, 2009, pp. 491–497.

[6] Xiaojun Wan. "Using bilingual knowledge and ensemble techniques for unsupervised Chinese sentiment analysis", In Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP '08, pages 553–561, Stroudsburg, PA, USA. Association for Computational Linguistics,2008.

[7] T.T.Zin, K.T.Yar, S.S.Htay, K.K.Htwe, N.T.T.Aung. W.W.Thant, "Domain-Specific Sentiment Lexicon for Classification", 1st International Conference on Advanced Information Technologies (ICAIT), Nov.1-2, 2017, Yangon, Myanmar.

[8] T.T.Zin, "Social Media Text Normalization", 2nd International Conference on Advanced Information Technologies (ICAIT), Nov.1-2, 2018, Yangon, Myanmar.

[9] S.M.Mohammad, M.Salameh, S.Kiritchenko, "How Translation Alters Sentiment", Journal of Artificial Intelligence Research 55, pp 95-130,2016.

[10] C.Zhang, M.Capelletti, A.Poulis, T.Stemann, J.Nemcova, "A Case Study of Machine Translation in Financial Sentiment Analysis", project co-financed by the European Union, Connecting Europe Facility.

[11] A.Balahur and M.Turchi, "Multilingual Sentiment Analysis using Machine Translation?", Proceedings of the 3rd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, pp52-60, Jeju, Republic of Korea, 12 July 2012.

[12] E.Demirtas, M.Pechenizkiy, "Cross-lingual Polarity Detection with Machine Translation", WISDOM'13, August 11, 2013. Chicago, USA. ACM 978-1-4503-2332-1/13/08.

[13] A. Balahur and M.Turchi, "Comparative Experiments for Multilingual Sentiment Analysis Using Machine Translation" ,published in SDAD@ECML/PKDD 2012.

[14] H.Aflli, S.McGuire and A.way, "Sentiment Translation for low resourced languages: Experiments on Irish General Election Tweets ", April 2017.

[15] http://translate.google.com

[16] T.Zagibalov and J. Carroll, "Automatic Seed Word Selection for Unsupervised Sentiment Classification of Chinese Text", Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008), pp1073-1080, Manchester, August,2008

[17] A.Shirbhate and S.Deshmukh, "Feature Extraction for Sentiment Classification on Twitter Data", International Journal of Science and Research (IJSR), volume 5 Issue2, February 2016.

[18] Q.Zhao, H.Wang and Pin Lv, "Joint Propagation and Refinement for Mining Opinion Words and Targets", IEEE International Conference on Data Mining Workshop, 2014.

[19] Lun.Ku, Yu.Liang and H.Chen, “Opinion Extraction, Summarization and Tracking in News and Blog Corpora”, American Association for Artificial Intelligence, 2006.

[20] P.B.Filho, L.Avanco, T.S.Pardo, M.G.V.Nunes, “NILC_USP: An Improved Hybrid System for Sentiment Analysis in Twitter Messages”, International Workshop on Semantic Evaluation, 8^{th} ,2014, Dublin.