Comparing and Evaluating the Sentiment on Newspaper Articles: A Preliminary Experiment

S Padmaja
Dept. of CSE, UCE
Osmania University
Hyderabad
bandupadmaja@gmail.com

Sasidhar Bandu
Dept. of English
Salman Bin Abdul-Aziz University
Saudi Arabia
sasibandu@gmail.com

Prof. S Sameen Fatima
Dept. of CSE, UCE
Osmania University
Hyderabad
sameenf@gmail.com

Pooja Kosala
Dept. of CSE,UCE
Osmania University
Hyderabad
kosalapooja@gmail.com

M C Abhignya
Dept. of CSE,UCE
Osmania University
Hyderabad
abhignya.1992@gmail.com

***Abstract*—Recent years have brought a symbolic growth in the volume of research in Sentiment Analysis, mostly on highly subjective text types like movie or product reviews. The main difference these texts have with news articles is that their target is apparently defined and unique across the text. Thence while dealing with news articles, we performed three subtasks namely identifying the target; separation of good and bad news content from the good and bad sentiment expressed on the target and analysis of clearly marked opinion that is expressed explicitly, not needing interpretation or the use of world knowledge. On concluding these tasks, we present our work on mining opinions about three different Indian political parties during elections in the year 2009. We built a Corpus of 689 opinion-rich instances from three different English dailies namely The Hindu, Times of India and Economic Times extracted from 02/ 01/ 2009 to 05/ 01/ 2009 (MM/ DD/ YY). In which (a) we tested the relative suitability of various sentiment analysis methods (both machine learning and lexical based) and (b) we attempted to separate positive or negative opinion from good or bad news. Evaluation includes comparison of three sentiment analysis methods (two machine learning based and one lexical based) and analyzing the choice of certain words used in political text which influence the Sentiments of public in polls. This preliminary experiment will benefit in predicting and forecasting the winning party in forthcoming Indian elections 2014.**

Keywords— Sentiment Analysis; News Articles; Public Mood

I. INTRODUCTION

The exponential increase in the volume of digitized information and difficulty in sorting and processing is the driving force behind Analysis of Data. Together with the development of technology and the growing access to information, Wiberg (2004) [1] witnessed the birth of a new type of society- that of the interaction and communication. Online Social Networks (OSNs) are one among them. Access to such popular communication platforms have given a way for the public to generate emotions, opinions, sentiments, evaluations, appraisals and attitudes towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. The volume of such information created every day is massive. Through such OSNs we share this information as part of our everyday lives for better understanding and integrating into our surrounding reality.

With the effect of such networks on a daily basis, we start recoiling our valuable decisions and actions with certain predefined notions created by others through their reviews. Thus, it is both interesting and challenging to see what and how the trends in such OSNs change from time to time. Comprehensive investigations can lead to appropriate predictions. This gives an edge to Text Mining and explores the area of Sentiment Analysis (Opinion Mining) which is designated as an automatic processing of texts to detect opinions.

Authors of such OSNs express their opinions freely, but, this is not the case with news articles. The major difference between news and product reviews is that the target of the sentiment is less concrete and is expressed much less explicitly. Another major difference is Newspaper articles is that they give an impression of objectivity to refrain from using clearly positive or negative vocabulary. They resort to other means to express their opinion, like embedding statements in a more complex discourse, omitting facts that highlight some important people. For this reason sentiment analysis on news text is rather difficult compared to others.

As we all know that politics is a practice of influencing people at individual level and in this Media plays a very important role as they form opinions based on the information and facts known to them through newspapers. Hence, in this paper, we chose the challenging job of mining opinions on news articles which are related to politics as they are the cheapest means for people to acquire awareness on political scenario and give their votes accordingly during polling.

The rest of this paper is organized as follows. In section II we illustrated various types of Sentiment Analysis Methods and their usage in literature predominantly working with news articles. Section III outlines the method of implementation and various performance evaluation measures that are to be carried out. Section IV highlights the comparison results. Finally, section V concludes and demonstrates the future work.

II. Related Work

Sentiment Analysis of Natural Language texts is a broad and expanding field. A text may contain both Subjective and Objective sentiments. Wiebe (1994) [2] defines Subjective text as the "linguistic expression of somebody's opinions, sentiments, emotions, evaluations, beliefs and speculations". In her definition, the author was inspired by the work of the linguist Ann Ban field (1982) [3], who defines subjective as a sentence that takes a character's point of view and that present private states (that are not open to objective observation or verification), defined by Quirk (1985) [4], of an experiencer, holding an attitude, optionally towards an object. Bing Liu (2010) [5] defines Objective text as the facts that are expressed about entities, events and their properties. Esuli and Sebastiani (2006) [6] define Sentiment Analysis as a recent discipline at the crossroads of Information Retrieval and Computational Linguistics which is concerned not with the topic a document is about, but with the opinion it expresses. While a wide range of human moods can be captured through Sentiment Analysis Hannak (2012) [7] says majority of studies focus on identifying the polarity of a given text—that is to automatically identify if a message about a certain topic is positive or negative.

Polarity analysis has numerous applications especially on news articles using several methods. Pang and Lee (2002) [8] broadly classifies Sentiment Analysis methods into machine-learning-based and lexical-based. Machine learning methods often rely on supervised classification approaches, where sentiment detection is framed as a binary (i.e., positive or negative). This approach requires labeled data to train classifiers [8]. While one advantage of learning-based methods is their ability to adapt and create trained models for specific purposes and contexts, their drawback is the availability of labeled data and hence the low applicability of the method on new data. This is because labeling data might be costly or even prohibitive for some tasks. On the other hand, lexical-based methods make use of a predefined list of words, where each word is associated with a specific sentiment. The lexical methods vary according to the context in which they were created [8].

Sentiment Analysis work has been handled heavily in subjective text types where the target is clearly defined and unique across the text as the case in movie or product reviews. But when applying Sentiment Analysis to the News domain, Alexandra Balahur (2009) [9] says it is necessary to clearly define the scope of the tasks in three levels. They are definition of the target; separation of good and bad news content from the good and bad sentiment expressed on the target and analysis of clearly marked opinion that is expressed explicitly, not needing interpretation or the use of world knowledge. Thus, on the same lines we built a corpus of 689 political instances from three different news databases namely, The Hindu, Times of India and Economic Times for three distinct Indian parties namely United Progressive Alliance (UPA), Telugu Desam Party (TDP) and Telangana Rashtra Samithi (TRS) to analyze the choice of certain words used in political texts to influence the sentiments of public in polls. Following is the glimpse of work done using news articles by various authors either by adopting machine learning based (MLB) or Lexical based (LB) methods with their respective accuracies.

TABLE I. An Overview of the Most Popular Work Done by Authors in News Articles

Title of the paper	Author	Year	SA Methods (MLB/ LB)	Data Set	Accur-acy (%)
[12]	Justin Grimmer et al.	2013	MLB	Political News articles	65
[13]	Hong Li et al.	2012	MLB	Online news articles	60
[14]	Robert P. et al	2011	MLB	news articles on finance	59
[15]	Hristo Tanev et al.	2010	LB	EMM news data	62
[16]	Alexandra Balahur et al.	2009	LB	EMM news data	50
[17]	Alexandra Balahur et al.	2009	LB	EMM news data	55
[18]	Robert P. Schumaker et al.	2008	LB	news articles on finance	51

Exclusively, our work in this paper is in comparing both machine learning and lexical based sentiment analysis methods and analyzing how best they are in extracting appropriate results.

III. Implementation

Political texts collected from three different databases as mentioned earlier are identified for subjectivity and performed preprocessing steps like tokenization, POS tagging and parsing using Natural Language Toolkit (NLTK) [20]. Our annotation framework for appropriate feature selection was on the lines of identifying target, source and text span that expresses attitudes like idioms, phrases or words.

Negation identification is done by separating noun phrase and verb phrase in each instance and then extracting their polarities. Thus, to bring a comparison between lexical based and machine learning methods, we implemented SentiWordNet (SWN) [10] which is based on English lexical dictionary called WordNet [11] and two machine learning based algorithms namely Naive Bayes (NB) and Support Vector Machine (SVM) using WEKA with 10 fold cross validation [19] respectively.

A. Comparison Measures

In order to define the metrics used to evaluate the methods, we considered the following metrics:

TABLE II. CONFUSION MATRIX

		Actual Observation	
		Positive	Negative
Predicted expectation	Positive	a	b
	Negative	c	d

Let ‘a’ represent the number of instances correctly classified as positive (i.e. true positive), ‘b’, the number of negative instances classified as positive (i.e. false positive), ‘c’, the number of positive instances classified as negative (i.e. false negative), and ‘d’, the number of instances correctly classified as negative (i.e. true negative).

In order to compare and evaluate the methods, we consider the following metrics, commonly used in information retrieval: true positive rate or recall:

$R = a/(a + c)$, false positive rate or precision: $P = a/(a + b)$, accuracy: $A = (a + d)/(a + b + c + d)$, and F-measure: $F = 2 \cdot (P \cdot R)/(P + R)$.

The true positive rate or recall can be understood as the rate at which positive instances are predicted to be positive (R), whereas the true negative rate is the rate at which negative instances are predicted to be negative. The accuracy represents the rate at which the method predicts results correctly (A). The precision rate, also called the positive predictive rate, calculates how close the measured values are to each other (P). We also use the F-measure to compare results, since it is a standard way of summarizing precision and recall (F). Ideally, a polarity identification method reaches the maximum value of the F-measure, which is 1, meaning that its polarity classification is perfect.

B. Coverage

It is a fraction of instances in a given dataset that a method is able to classify as either positive or negative. Ideally, polarity detection methods should retain high coverage to avoid bias in the results, due to the unidentified instances. In particular, suppose that a sentiment method has classified only 10% of a given set. The remaining 90% consisting of unidentified instances may completely change the result, that is, whether the context drawn from the datasets should be positive or negative. Therefore, having high coverage in data is essential in analyzing news articles.

IV. RESULTS

In order to understand the advantages, disadvantages, and limitations of the various sentiment analysis methods and analyze the choice of words from news articles in winning a particular party, we present comparison results among them which are as shown below –

A. Prediction Performance

We illustrate a comparative performance evaluation of each method in terms of correctly predicted polarity. Here we depict the results for precision, recall, accuracy, and F-measure for the three previously described methods. Table 3 shows the performance of the results obtained for each labeled dataset. For the F-measure, a score of 1 is ideal and 0 is the worst possible. Among the methods the best F-measure was SVM with the values 0.688 and 0.657 for UPA and TRS respectively and NB with the value 0.723 for TDP. On par, the accuracy of SWN, NB and SVM with the values 0.742, 0.725 and 0.666 are the highest scores in UPA, TDP and TRS respectively.

TABLE III. PREDICTED PERFORMANCE FOR LABELED DATASET

Metric	**UPA**			**TDP**			**TRS**		
	SVM	**NB**	**SWN**	**SVM**	**NB**	**SWN**	**SVM**	**NB**	**SWN**
Recall	0.719	0.652	0.58	0.700	0.725	0.50	0.667	0.636	0.54
Precision	0.690	0.725	0.86	0.689	0.721	0.54	0.658	0.623	0.63
F-Measure	0.688	0.635	0.565	0.687	0.723	0.51	0.657	0.62	0.48
Accuracy	0.712	0.652	0.742	0.70	0.725	0.601	0.666	0.636	0.601

B. Coverage

For each party as described above we computed the coverage of the three sentiment analysis methods and observed that SentiWordNet has the highest coverage with 74.2% for UPA (congress) whereas Naïve bayes with 72.5% is the next highest coverage score for TDP and SVM accounts for a fairer score of coverage with 66.7% for TRS.

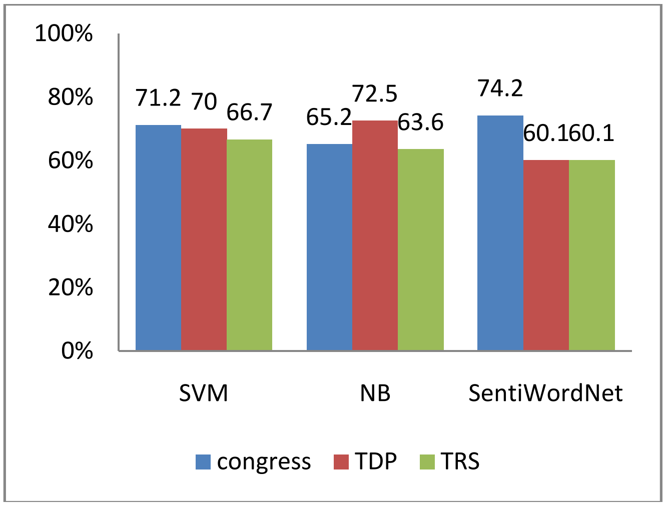

Fig. 1. Coverage of three parties

V. CONCLUSION AND FUTURE WORK

Recent efforts to analyze the moods embedded in Web 2.0 content have adapted various sentiment analysis methods originally developed in linguistics and psychology. Several of these methods became widely used in their knowledge fields and have now been applied as tools to measure polarity in the context of OSNs. In this paper, we began the comparison of three representative sentiment analysis methods like Support Vector Machine, Naïve Bayes and SentiWordNet.

Our comparison study focused on detecting the polarity of content (i.e., positive and negative affects) from good or bad news for three different Indian political parties. Thus by extracting the average predicted performance we observed that the choice of certain words used in political text was influencing the Sentiments in favor of UPA which might be one of the causes for them be the winners in Elections 2009.

We also adopted coverage (measuring the fraction of instances whose sentiments are detected) as a measure of efficacy. We found that all the three methods have varying degrees of coverage and no single method is always best across different text sources. T

his work has demonstrated a framework with which various sentiment analysis methods can be compared. As a natural extension of this work, we would like to continue to add more existing methods for comparison, such as Happiness Index, SentiWordNet, SASA, PANAS-t, Emoticons and SenticNet on two datasets created by formal and informal texts with appropriate feature extraction on the topic related to present political scenario for various leading parties in India.

ACKNOWLEDGMENT

I am deeply indebted to EMR Division under HRD Group of Council of Scientific & Industrial Research (CSIR) for providing me CSIR Research Fellowship (SRF) and also to Technical Education Quality Improvement Programme (TEQIP) phase II for providing me with travel funds.

REFERENCES

[1] Wiberg, M., "The interaction Society: Theories Practice and Supportive Technologies." Idea Group Inc. 2004.

[2] Wiebe, J., "Tracking point of view in narrative." Computational Linguistics, 20, 1994.

[3] Banfield, A. "Unspeakable sentences: Narration and Representation in the Language of Fiction." Routledge and Kegan Paul, 1982.

[4] Quirk, R., " A Comprehensive Grammar of the English Language", Longman Publishing House, 1985.

[5] Bing Liu, "Sentiment Analysis and Subjectivity", Handbook of Natural Language Processing, Second Edition, 2010.

[6] Esuli, A. and F. Sebastiani, "SentiWordNet: A Publicly Available Resource for Opinion Mining." In Proceedings of the 6th International Conference on Language Resources and Evaluation, LREC 2006, Italy. 2006.

[7] A. Hannak, E. Anderson, L. F. Barrett, S. Lehmann, A. Mislove, and M. Riedewald. Tweetin'in the rain: Exploring societal-scale effects of weather on mood. In Int'l AAAI Conference on Weblogs and Social Media (ICWSM), 2012.

[8] B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up?: sentiment classification using machine Learning techniques. In ACL Conference on Empirical Methods in Natural Language Processing, pages 79–86, 2002.

[9] Alexandra Balahur, Ralf Steinberger, "Rethinking Sentiment Analysis in the News: from Theory to Practice and back", European Commission-Joint Research Centre, University of Alicante, Department of Software And Computing Systems.WOMSA, pp.1-12, 2009.

[10] Esuli and Sebastiani. Sentwordnet: A publicly available lexical resource for opinion mining. In International Conference on Language Resources and Evaluation (LREC), pages 417– 422, 2006.

[11] G. A. Miller. Wordnet: a lexical database for english. Communications of the ACM, 38(11):39–41, 1995.

[12] Justin Grimmer,Brandon M. Stewart, "Text as Data: The Promise and Pitfalls of Automatic ContentAnalysis Methods for Political Texts", 2013.

[13] Hong Li, Xiwen Cheng, Kristina Adson, Tal Kirshboim, Feiyu Xu, "Annotating Opinions in German Political News", 2012.

[14] Robert P. Schumaker, Yulei Zhang, , Chun-Neng Huang and Hsinchun Chen, "Evaluating Sentiment in Financial News Articles", 2011.

[15] Hristo Tanev, Bruno Pouliquen, Vanni Zavarella, and Ralf Steinberger, "Automatic Expansion of a Social Network Using Sentiment Analysis", Annals of Information Systems 12, DOI 10.1007/ 978-1-4419-6287-4_2, Springer Science+Business Media, LLC 2010.

[16] Alexandra Balahur, Ralf Steinberger, "Rethinking Sentiment Analysis in the News: from Theory to practice and back", WOMSA, pp. 1-12, 2009.

[17] Alexandra Balahur, Erik van der Goot, Ralf Steinberger, Matina Halkia, "Sentiment Analysis in the News", OPTIMA, 2009.

[18] Robert P., Schumaker, Yulei Zhang and Chun-Neng Huang, "Sentiment Analysis of Financial News Articles", 2008.

[19] Weka User Manual, http://iasri.res.in/ebook/win_school_aa/notes/WEKA.pdf

[20] NLTK 3.0 documentation, http://nltk.org/.