

A Hybrid System for Online Detection of Emotional Distress

Tim M.H. Li¹, Michael Chau², Paul W.C. Wong³, and Paul S.F. Yip¹

¹HKJC Center for Suicide Research and Prevention,
The University of Hong Kong, Hong Kong
{timlmh, sfpypip}@hku.hk

²School of Business, The University of Hong Kong, Hong Kong
mchau@business.hku.hk

³Department of Social Work and Social Administration, The University of Hong Kong,
Hong Kong
paulw@hku.hk

Abstract. Nowadays, people are familiar with online communication and tend to express their deeper feelings on the Web. In the light of this situation, we present a hybrid system based on affect analysis for mining emotional distress tendencies from publicly available blogs to identify needy people in order to provide timely intervention and promote better public health. We describe the system architecture with a hand-crafted model at a fine level of detail. The model, which incorporates human judgment, enables the adjustment of prediction in machine learning on blog contents. The system blending supervised and unsupervised approaches will be examined and evaluated in lab experiments and practice.

Keywords: Affect mining, blogs, emotional distress, depression, hybrid system, hand-crafted model, machine learning, public health.

1 Introduction

With the recent advent of Web 2.0 applications such as blogging and social networks, many people, especially adolescents, express their negative emotions in their blogs. Some local non-governmental organizations have started to look for such distressed and negative self-expressions in blogs to identify potential severely depressed people in order to provide help and follow-up services. Online detection is regarded as a proactive and engaging way to identify the high risk groups and is important for public health. Nevertheless, the current initiatives are very labor-intensive and ineffective because they rely on simple keyword searching and matching from the World Wide Web, theoretically, a space with infinite capacity.

We should therefore take advantage of advanced text analysis and mining technology to enhance the time-and-cost-efficiencies of the initiative. Techniques for blog

mining and affect analysis have advanced substantially in recent years [6]. Although these techniques could potentially help with online depression detection and even have life-saving applications, little has been done and measured. In this paper, we propose to use text analysis techniques to study the characteristics of bloggers with emotional distress, and design, develop, and evaluate a system that automates the detection and analysis of emotional distress expressed in blogs.

The next section presents the related work. Section 3 is the architecture of the proposed system. The technical detail of the hand-crafted model is discussed in Section 4. The evaluation plan of the system is shown in the Section 5, and our ongoing work is discussed in the final section.

2 Related Work

In the digital age, people are keen on communicating and interacting on the Internet. There is a wide range of channels on this platform where blogs have become one of the most popular communication means on the World Wide Web. Blogs are like personal websites comprising of various content that makes regular records or digital documentations of writers' lives, opinions, and feelings [16]. Blogs, therefore, have become a rich resource for extraction of useful information and analysis on the online behaviors of people with severe signs of depression.

Web content mining techniques are instrumental to the extraction and analysis of content in blogs and social networking sites [6]. For content analysis, several approaches have been proposed in locating opinions or sentiment in blogs. Many techniques have been adopted in information retrieval and opinion identification in blogs during the Blog track of the TREC conference [14]. Affect or emotion analysis, which has emerged recently, takes advantage of well-developed sentiment and opinion analysis to classify documents into multiple emotions. For example, sentiment analysis is often used to classify texts into two or three classes (e.g., positive, negative, and neutral) while affect analysis focuses on a large number of affect classes such as happiness, sadness, hate, anger, fear, surprise, and so on [21]. Texts can contain multiple affects [2][10].

Studies on detecting users' affects in blogs have been conducted. For instance, word frequencies and the usage of special characters in blogs were analyzed to estimate the mood of the general public during the London bombing incident in 2005 [15]. In addition, co-occurrence analysis has also been applied [9]. Classification methods have been used to categorize the affect intensities of Web forums and blog messages [2].

3 System Architecture

Here we present the system architecture to provide an overview of the system and implementation of our project on depressed individuals. We propose a hybrid system combining supervised machine learning approach and an unsupervised hand-crafted model in affect analysis to identify our potential targets – operationalized as expressing depression in written blogs. Affect analysis can be extended to discovery of depressed people with the publicly available blogs containing depression signs.

The system performs a retrieval of information through blog search engines like Google Blog search by a meta-search approach [5][7]. After acquisition of blog posts through RSS in XML format, the files can be read with automated processes, which also prepare the blog contents for subsequent analysis.

Blog content that is ready for analysis are separately categorized with two classifiers, a hand-crafted model and a supervised machine learning approach using Support Vector Machine (SVM). SVM is a well-known and highly effective approach yielding high accuracy in affect analysis and text mining. Training and prediction in SVM-based classification in the system are carried out with the inductive learning algorithm in Joachim's SVM^{light} package¹, which is an implementation of SVM algorithms in C. The SVM approach puts a great emphasis on document-level analysis without, however, investigating the sentence-level influence on document polarity inherently. There is possibility that the blog content containing noisy data is not sufficiently analyzed and examined. As a result, our hybrid system is aimed to combine the document-level and sentence-level analysis of affect and operating them separately to achieve an improvement of performance. The hand-crafted model proposed in this paper, which utilizes human judgment and experience in web discourse and conversation text, facilitates the sentence-level analysis. A scoring system is also incorporated in the model. The model provides a new direction for document classification from a sentence-level prospective.

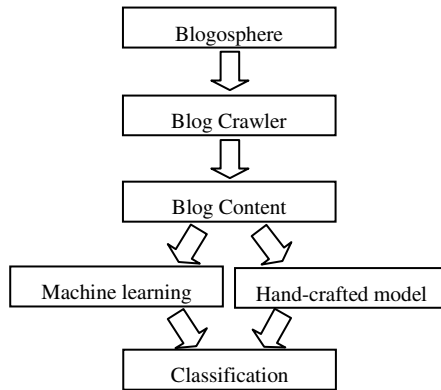


Fig. 1. System architecture

Both classifiers generate a score of prediction for each blog. The content predicted to be showing emotional distress is assigned a positive score, if no emotional distress is found a negative score is assigned. The score from the hand-crafted model is harnessed to adjust the score from SVM-based classifier to correct false decisions and generate a more precise and accurate result. The final score of content is computed with the aggregation of the scores from both classifiers and the classification decision is eventually made.

¹ <http://svmlight.joachims.org/>

4 Hand-Crafted Model

This model classifies blog content based on the sentence and document level. The sentence-level classification differentiates sentences into positive or negative emotion; as a result the model is able to determine whether the whole document shows emotional distress from the automatically annotated sentences. This approach does not require a labor-intensive annotating process of a large number of documents and sentences and is aimed to produce a high precision result.

4.1 Lexicon Creation

Since no lexicon specifically concerning emotional distress wordings is available, we develop our own lexicon in this model. The lexicon is constructed with function words and by manual inspection of blog contents by professionals familiar with web discourse terminology for emotional distress. Similar lexicon creation approaches have been used in previous studies and shown encouraging results [1][21]. The professionals read over thousands of different blog contents with emotional expressions and extract representative words of positive, negative, and neutral emotions in a macro-view. Manual lexicon creation is encompassed since blogs contain their own terminology which can be difficult to extract without human judgment and manual evaluation of conversation text. The words in the lexicon are categorized into ten groups with various functions in the hand-crafted model. All the words of which intensity or valence is neglected are treated equally in the lexicon without individual score assignment. Different groups of words are, however, used in different components in a sentence-level scoring process in the model.

4.2 Subjective Sentence Identification

Subjective sentence identification in this classifier is different from the previous studies that made use of subjective words in existing knowledge and sentiment databases [18][23]. Instead of finding the expressions of common affect like fear and anger which are normal expressions of feelings, the model is aimed at identifying multiple affect emotional distress which is more complex. Sentences with self-referencing pronouns are regarded as subjective sentences since those sentences directly reflect the writers' cognition. In addition, some studies in psycholinguistics reveal that people who currently have depression or suicidal ideation have a distinctive linguistic style and tend to use significantly more self-referencing words (e.g., *I, me, myself*) in their writings entailing strong self-orientation [17][19] and even withdrawal of social relationship [20]. Although this self-referencing style is difficult to identify through human judgment, those sentences with self-referencing words are believed to provide more clues on identifying disengagement behavior and hence emotional distress.

4.3 Sentence Polarity Prediction

Subjective sentence polarity of emotional distress is predicted with the positive (PW) and negative emotion words (NW). Intuitively, a sentence is classified as containing

negative emotion and hopelessness when NW is found. On the other hand, the sentence contains positive emotion when PW is found. When neither NW nor PW is found, the sentence is regarded as the same as non-subjective sentence. In the case where both NW and PW are found, the sentence is classified arbitrarily as showing negative emotion to increase the recall rate. After the polarity of a sentence is preliminarily predicted, negation is checked to determine the final polarity. The base scores of sentences are assigned as -0.1 or +0.1 for positive or negative polarity correspondingly. Negative base scores can be augmented with the occurrence and frequency of the negative events and words referring to death in negative polarity sentence. This augmentation is because detailed explanation of negative events (e.g., divorce, serious illness) provides useful information to identify emotional distress.

Non-subjective sentence polarity is not predicted with the emotion words for assignment of score. The sentence is checked for words that reference others or express thankfulness or encouragement, since these types of words do not necessarily appear in subjective sentences expressing non-negative emotion. Under disengagement theory, it is believed that people who reference other sources to offer opinion or convey information to others have a lower risk at depression [20]. Giving thankful and encouraging words to others also demonstrate a positive attitude in the writers. Negative score for marking the lack of emotional distress is thus assigned for clear identification of the absence of disengagement behavior.

4.4 Sentence Score Aggregation

The scores of automatically annotated sentences presented in the previous section 4.3 are blended with an algorithm to make the final decision of a document score. Since the emotional fluctuations throughout a document are complicated yet not meaningful to our final decision, some of the scores in the middle of the document may not be meaningful and may even be confusing. The aggregation, therefore, primarily concentrates on the scores at the beginning and the end of the document.

It is believed that the conclusion and major theme expressed by writers generally appear at the beginning and the end of documents/blogs [13]. There is difficulty, however, in defining the parameters of what constitutes the opening and ending of a document. Static positioning, however, does not yield significantly higher accuracy because of the reduced flexibility of the analysis. Furthermore, these parameters vary for documents by different writers who have diverse writing and organization styles. An algorithm that dynamically defines these parameters, therefore, is crucial for improving the analysis performance. The aggregation that takes advantage of this algorithm can capture the underlying emotion expressed in documents and provide a more accurate score.

There are many segmentation methods that have mainly been used to find sub-topics in full-length documents e.g., by grouping sentences in blocks and partitioning blog content into coherent units where the first and last blocks of subjective sentences in a document with the same polarity are taken into account for the final prediction score [11][12]. Likewise, this approach can be applied to partitioning emotional fluctuations and only capturing the main emotional state without the emotional transitions

which is sufficient and representative in determining the prediction. The final prediction score depends heavily on the emotional consistency and certainty presented in the document. In other words, the effect of some of the parts in the document is probably reduced. For example, elaborations on ideas and supplementary examples provided in the middle part of the document to make it more persuasive and interesting is not meaningful and even ambiguous to the prediction. Emotions of other people in examples or emotional transitions in writer's retrospect make the prediction less accurate. As a result, reduction of those parts is instrumental to improving the final result.

The handcraft model is capable of individually providing orientation predictions for documents like other unsupervised rule-based classifiers and the resulting scores can also help adjust predictions from the SVM-based classification. To attain the adjustment, the number of emotional state changes is considered for computation of scores since some of the parts like examples and retrospect mentioned are regarded as noise to the SVM-based document-level classifier. The classifier capturing unimportant features from emotional transition areas, which are not consistent and representative to the main feeling of the writers, lead to a deteriorated performance. The model reduces noise by controlling the weighting of the document's opening and ending. The weighting depends on the number of emotional transitions which is measured by the number of partitions between coherent emotion units in a document.

5 Evaluation Plan

The system will be comprehensively evaluated in both a laboratory setting and in real life situations in non-governmental organizations in Hong Kong by their staff.

In the laboratory setting, three experiments will be conducted to investigate the performance of SVM-based classification, hand-crafted model, and the hybrid system in distinguishing blog content showing emotional distress. We are continuously collecting more blog content with emotional expression from different sources on the Internet in order to enlarge the database to reflect the real life situation as closely as possible. The blog contents will be reviewed and rated systemically by professionals, such as clinical psychologists, to construct a better test bed for the evaluation of the system.

Standard evaluation metrics for classification namely precision, recall, and F-measures are used to evaluate the performance of affect analysis approaches. Precision is the fraction of retrieved documents relevant for indicating correctness while recall is the fraction of the relevant documents retrieved for indicating completeness. F-measure is the harmonic mean of the two measures.

$$precision = tp / (tp + fp)$$

$$recall = tp / (tp + fn)$$

where tp is true positive, fp is false positive, and fn is false negative.

$$F = 2 \times (precision \times recall) / (precision + recall)$$

After sufficient testing, the system will be delivered to various end-users who are interested in using this practical tool in searching for needy people on the Internet. Feedback in system satisfaction and usability will be obtained and the effectiveness of the system in real-life application will also be examined. We will study the number of cases correctly reported by the system and compare with the number of cases found by social workers of interested non-governmental organizations who have been using their current manual search method.

6 Ongoing Work

A considerable number of studies have been conducted on machine learning and on unsupervised learning on affect analysis but only separately and there is a paucity of research on the combined approach in a hybrid system with an appropriate evaluation. In the future, we will evaluate the hybrid system in order to show its practicality and effectiveness. We also plan to investigate the use of social network analysis [8][22] in studying online communities who appear to be emotionally distressed and to apply the analyses to multiple languages [3][4][24]. It is believed that a system utilizing advanced techniques will facilitate the identification of individuals with emotional distress, in terms of time and cost efficiencies. It is hoped that limited and scarce resources can be shifted from the labor-intensive searching job to the implementation of interventions so that more people in need can benefit.

References

1. Abbasi, A., Chen, H.: Affect intensity analysis of dark web forums. In: Proceedings of IEEE International Conference on Intelligence and Security Informatics, pp. 282–288 (2007)
2. Abbasi, A., Chen, H., Thoms, S., Fu, T.: Affect Analysis of Web Forums and Blogs Using Correlation Ensembles. *IEEE Transactions on Knowledge and Data Engineering* 20(9), 1168–1180 (2008)
3. Chau, M., Qin, J., Zhou, Y., Tseng, C., Chen, H.: SpidersRUs: Automated Development of Vertical Search Engines in Different Domains and Languages. In: Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2005), Denver, Colorado, USA, June 7–11, pp. 110–111 (2005)
4. Chau, M., Qin, J., Zhou, Y., Tseng, C., Chen, H.: SpidersRUs: Creating Specialized Search Engines in Multiple Languages. *Decision Support Systems* 45(3), 621–640 (2008)
5. Chau, M., Shiu, B., Chan, I., Chen, H.: Redips: Backlink Search and Analysis on the Web for Business Intelligence Analysis. *Journal of the American Society for Information Science and Technology* 58(3), 351–365 (2007)
6. Chau, M., Xu, J., Cao, J., Lam, P., Shiu, B.: A Blog Mining Framework. *IEEE IT Professional* 11(1), 36–41 (2009)
7. Chen, H., Fan, H., Chau, M., Zeng, D.: Testing a Cancer Meta Spider. *International Journal of Human-Computer Studies* 59(5), 755–776 (2003)

8. Chen, H., Chung, W., Qin, Y., Chau, M., Xu, J.J., Wang, G., Zheng, R., Atabakhsh, H.: Crime Data Mining: An Overview and Case Studies. In: Proceedings of the National Conference for Digital Government Research (dg.o 2003), pp. 45–48 (2003)
9. Gill, A.J., French, R.M., Gergle, D., Oberlander, J.: The Language of Emotion in Short Blog Texts. In: Proceedings of ACM Conference on Computer-Supported Collaborative Work (CSCW), San Diego, California, USA, November 8-12 (2008)
10. Grefenstette, G., Qu, Y., Evans, D.A., Shanahan, J.G.: Validating the Coverage of Lexical Resources for Affect Analysis and Automatically Classifying New Words Along Semantic Axes. In: Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications, AAAI-EAAT 2004 (2004)
11. Hearst, M.A., Plaunt, C.: Subtopic structuring for full-length document access. In: Proceedings of the 16th Annual International ACM/SIGIR Conference, pp. 59–68 (1993)
12. Hearst, M.A.: TextTiling: Segmenting text into multi-paragraph subtopic passages. *Computational Linguistics* 23, 33–64 (1997)
13. Lee, L., Pang, B., Vaithyanathan, S.: Thumbs up? Sentiment Classification using machine learning techniques. In: EMNLP, pp. 79–86 (2002)
14. Macdonald, C., Santos, R.L.T., Ounis, I., Soboroff, I.: Blog track research at TREC. *SIGIR Forum* 44(1), 58–75 (2010)
15. Mishne, G., de Rijke, M.: Capturing Global Mood Levels Using Blog Posts. In: Proceedings of the AAAI Spring Symposium on Computational Approaches to Analysing Weblogs, AAAI-CAAW (2006)
16. Nardi, B.A., Schiano, D.J., Gumbrecht, M., Swartz, L.: Why we blog. *Communications of the ACM* 47(12), 41–46 (2004)
17. Ramirez-Esparza, N., Chung, C.K., Kacewicz, E., Pennebaker, J.W.: The psychology of word use in depression forums in English and in Spanish: Testing two text analytic approaches. Paper Presented at the International Conference on Weblogs and Social Media, Seattle, WA (2008)
18. Riloff, E., Wiebe, J.: Learning extraction patterns for subjective expressions. In: EMNLP 2003 (2003)
19. Rude, S.S., Gortner, E.M., Pennebaker, J.W.: Language use of depressed and depression-vulnerable college students. *Cognition and Emotion* 18, 1121–1133 (2004)
20. Stirman, S.W., Pennebaker, J.W.: Word use in the poetry of suicidal and non-suicidal poets. *Psychosomatic Medicine* 63, 517–522 (2001)
21. Subasic, P., Huettner, A.: Affect analysis of text using fuzzy semantic typing. In: Proceedings of the Ninth IEEE International Conference on Fuzzy Systems, San Antonio, TX, USA, pp. 647–652 (May 2000)
22. Xu, J., Chau, M.: The social identity of IS: analyzing the collaboration network of the ICIS conferences (1980-2005). In: Proceedings of the International Conference on Information Systems, Milwaukee, Wisconsin, USA, December 10-13 (2006)
23. Zhang, C., Zeng, D., Li, J., Wang, F.Y., Zuo, W.: Sentiment analysis of Chinese documents: from sentence to document level. *Journal of the American Society for Information Science and Technology* 60(12), 2474–2487 (2009)
24. Zeng, D., Wei, D., Chau, M., Wang, F.: Domain-Specific Chinese Word Segmentation Using Suffix Tree and Mutual Information. *Information Systems Frontiers* 13(1), 115–125 (2011)