Thesis Report: Application of sentimental analysis in adaptive user interfaces

By

Badruddin Kamal(08301012);
Supervisor Abu M. Hammad Ali;
Co-supervisor Dr. Mumit Khan;
BRAC University,Dhaka, Bangladesh;

Fall'2011

Badruddin Kamal	Abu M. Hammad Ali	Dr. Mumit Khan
08301012	Supervisor	Co-supervisor

Contents

Thesis: Application of sentimental analysis in adaptive user interfaces

Badruddin Kamal
Supervisor: Abu M. Hammad Ali
Co-supervisor: Dr. Mumit Khan
BRAC University, Dhaka, Bangladesh

Abstract

The objective of this project is to show how sentimental analysis can help improve the user experience over a social network or system interface. The learning algorithm will learn what our emotions are from statistical data then determine the mood. After that it will change our social interactions accordingly on our social network sites or other interfaces like desktop or system services or web-pages. Suppose you are bored or sad ,in the case of social networks one thing the computer could do is to be more suggestive of things that lighten your mood and change interactions like backgrounds color's ,icons services. The site could automatically try suggesting interactions with people and applications that would help improve the mood, while hiding others that might make it worse. The project aims to implement these in the social network community as well as services and interfaces of our systems, while making our lives better and our experience richer and efficient.

1. Introduction

"It is not God's will merely that we should be happy, but that we should make ourselves happy" Immanuel Kant.

As sentimental analysis has improved in the last few decade's so has its applications. Sentimental analysis[1][3] is now being used from specific product marketing to anti social behavior recognition. The advances in facebook ,twitter , youtube and other microbloging and social networking sites have not only contributed change to the social sites but have fundamentally changed the way we use these sites and how we share our feelings, our views with the wider audience.

Millions of messages are appearing daily in popular web-sites that provide services for microblogging. Authors of those messages write about their life, share opinions on variety of topics and discuss current issues. Because of a free format of messages and an easy accessibility of microblogging platforms, Internet users tend to shift from traditional communication tools (such as traditional blogs or mailing lists) to microblogging services. As more and more users post about products and services they use, or express their political and religious views, microblogging web- sites become valuable sources of people's opinions and sentiments. Such data can be efficiently used for adaptive user interface's.[2]

Data from these sources can be used in opinion mining and sentiment analysis tasks. For example, we may be interested about Questions about an individual:

- What do people think about this persons status post(comments) ?
- How positive (or negative) are people about it?

For our research purposes we will be targeting a specific Facebook fan page / wall post to collect corpus data, then use this data to determine a users mood so we can highlight whether they are in a positive or neutral or negative mood and determine set actions for each out come.

For example, in the case the user has a bad mood facebook can suggest an application like lifebox, games like cityville etc. or try stimulate conversation with a use who is in a better mood than our subject by suggesting their posts more frequently. It can also change the layout(Background, themes, etc.) more frequently.

We perform a linguistic analysis of our corpus and we show how to build a sentiment classifier that uses the collected corpus as training data.

Topics overview

The report is divided into sections as follows:
In 2.Background Work we discuss the work on sentimental analysis that has been done around the world. We discuss different paper's research finding's and discussion topics.

In 3. Methodology we discuss our plans for thesis progress how the project is planned to move along. What are the goals we have targeted? What results are we looking for? In 4.Implementation we discuss the overview on how the software was implemented. In 5.Training we discuss the sources how we trained our classifiers. In 6.Results we discuss the results obtained and how the accuracy of the classifiers affected out choices using different methods. In 7.Conclusion we discuss the overview and brief of the project and finale.

2. Background Work

In this digital age we are connected more and more via social media. According to Nielsen wire we spend about 6 hrs [17] or more daily being connected to them. They have become the new domains of our social interactions.

A small survey taken about 25 people(Though not significant enough) suggest that

Social media's specific Usage	Percentage %
Used when even there is no objective or interest	60%
Facebook used for social interactions	85%

So for collecting data it can be considered as a good source. This data can be used later for our adaptive interfaces mood determiner.[14][23]

Chmiel, A., Sienkiewicz, J., Thelwall, M., Paltoglou, G., Buckley, K., Kappas, A. & Hołyst, J.A. (2011) [11] suggest that collective emotional states can be created and modulated via Internet communication and that emotional expressiveness is the fuel that sustains some e-communities.

Our adaptive interfaces [3] will work much like iGoogle interfaces allowing us to change our layout, but rather than being predetermined it is targeted to be more automated depending on subject sentiment.

Different people have approached the online community in different ways to solve the issue of sentimental analysis. Their work can be divided into a few parts.

- Lexical analysis the analysis of frequently used words.
- Sentimental analysis of sentiment or opinion of the subject.
- Social network analysis(SNA) study of how social network interactions and mindset affect data.

To keep our objectives simple we will only study Lexical analysis and Sentimental analysis to determine our data and exclude SNA for the time being.[5][9][10][12]

With the population of blogs and social networks, opinion mining and sentiment analysis became a field of interest for many researches. A very broad overview of the existing work was presented in [7]In their survey, the authors describe existing techniques and approaches for an opinion-oriented information retrieval.
However, not many researches in opinion mining considered blogs and even much less addressed microblogging. In [6], the authors use web-blogs to construct a corpora for sentiment analysis and use emotion icons assigned to blog posts as indicators of users' mood. The authors applied SVM and CRF learners to classify sentiments at the sentence level and then investigated several strategies to determine the overall sentiment of the document.
As the result, the winning strategy is defined by considering the sentiment of the last sentence of the document as the sentiment at the document level. J. Read in [8] used emoticons such as ":-)" and ":-(" to form a training set for the sentiment classification. For this purpose, the author collected texts containing emoticons from Usenet newsgroups. The dataset was divided into "positive" (texts with happy emoticons) and "negative" (texts with sad or angry emoticons) samples. Emoticon strained classifiers: SVM and Na¨ıve Bayes, were able to obtain up to 70% of an accuracy on the test set.

In [5], authors used Twitter to collect training data and then to perform a sentiment search. The approach is similar to [8]. The authors construct corpora by using emoticons to obtain “positive” and “negative” samples, and then use various classifiers. The best result was obtained by the Na¨ıve Bayes classifier with a mutual information measure for feature selection. The authors were able to obtain up to 81% of accuracy on their test set. However, the method showed a bad performance with three classes (“negative”, “positive” and “neutral”).[4]

From Alexander Pak, Patrick Paroubek’s[4] work we find that using bigrams yield better results, and also larger dataset yield better accuracy in POS tagging.

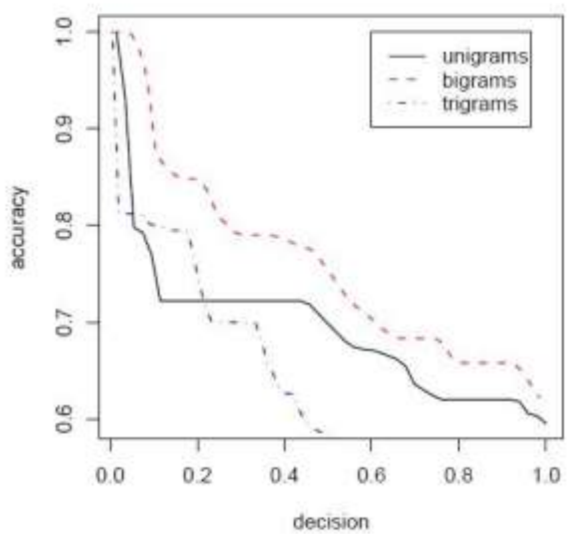


Figure : The comparison of the classification accuracy when using unigrams, bigrams, and trigrams

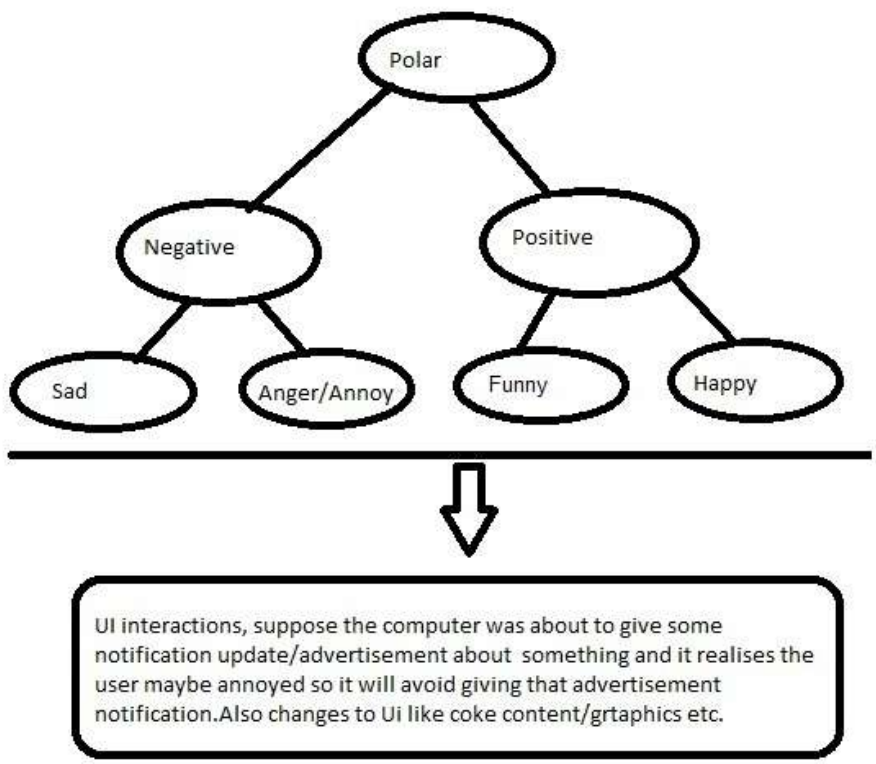
For this method to work we determine using a data set about 5000 may yield better results for out work. Datasets must be significant enough to provide accurate results. Taking smaller datasets might not provide proper accuracy.

Also form Wojciech Gryc and Karo Moilanen ,2008[10] we learn that our training dataset must be biased free so that the data set is not tampered accordingly. Eg. Taking a dataset corpus of politically biased data will make more lenient to a specific group than generalizing it.

Another aspect which should be kept enlightened to this is that a multitude of methods maybe applied to do this and each my yield different accuracy. We must cut out human data error from it and also make sure it provides better performance than 80% which is statistically close to human efficiency in detecting emotions correctly from text.[19][20]

3. Methodology

For research purposes we tried a new approach, As to just tagging positive and negative data, we have tried to tag emotions such as Happy ,Funny ,Sad & Angry. To track this data, we a new corpus will be created. We approached on a binary method to detect these emotions. We first detect positive and negative. The under positive we find happy or funny and under negative we detect sad and angry. Here positive is made of data from happy and funny and negative is also made of sad and angry.



Tools planned to be used:

- Python NLTK
- Facebook page as Corpus
- Device UI/Webpage programming

What’s planned on being done?[13][15][16][18][21][24][25]

We will use python NLTK ‘s rich library to program our analyzer. Our model will use Naive Bayes method to analyze words as groups of sections using bi-gram’s. This may not achieve much accuracy at first but, we plan to incorporate above suggested methods to achieve better accuracy.

Our model will have analyzer for POS tagging words for lexical analysis OpenNLP has shown much promise in this area.

We will also target subjectivity and objectivity for identifying polar and neutral comments. Also we will need identification of positive and negative comments.

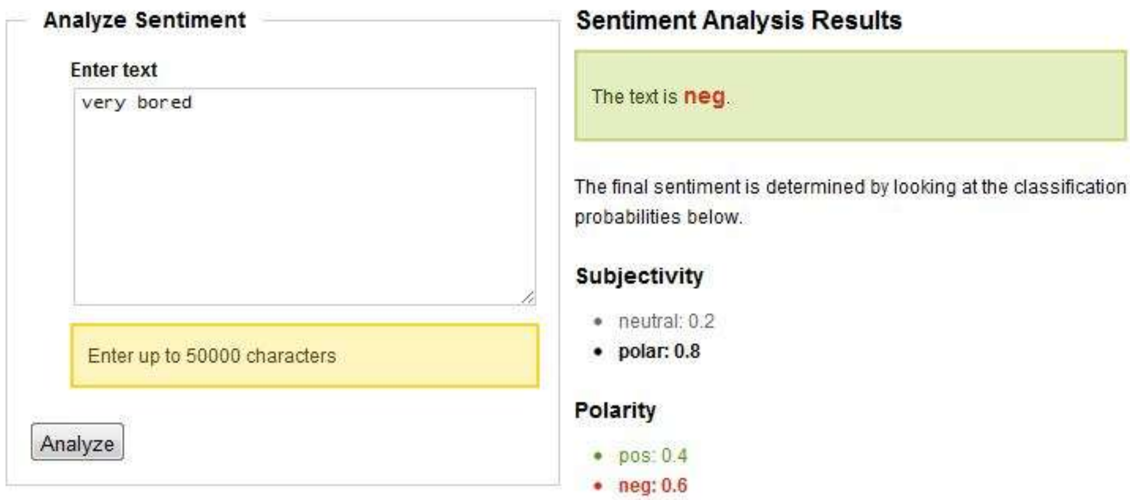
Taking datasets of only positive or only neutral data will bias the data to ensure proper dataset incorporation we must use both in our dataset training.

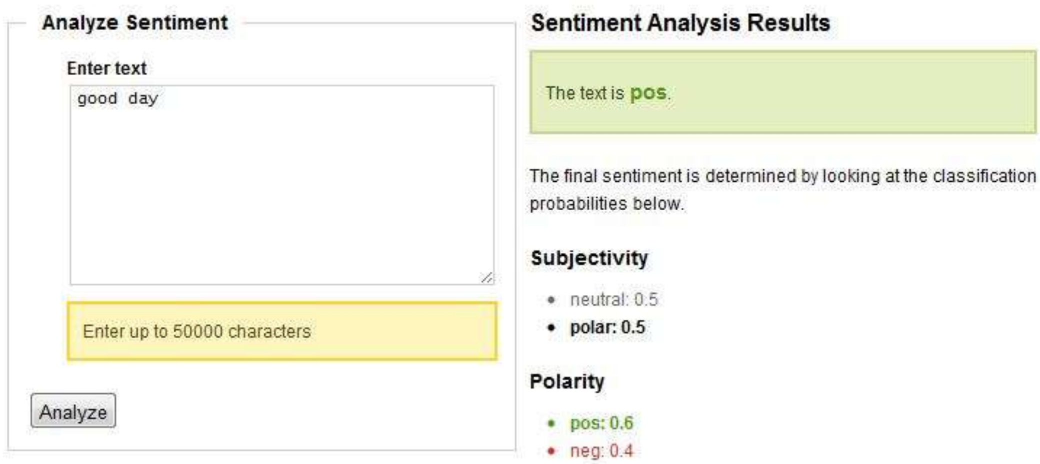
All of the NLTK classifiers work with featstructs, which can be simple dictionaries mapping a *feature name* to a *feature value*. For text, we'll use a simplified bag of words model where every word is feature name with a value of True. Here's the feature extraction method:

```
def word_feats(words):
        return dict([(word, True) for word in words])
```

We will collect corpus data using crawler from face book on targeted website topics. Of the data we will use 90% data in raining and 10% in testing data. The classifier which will be used is Naïve Bayes Classifier .[22]

86.4% is the best hierarchical performance reported by Pang and Lee using SVMs to classify polarity stacked on top of a naive Bayes subjectivity classifier. Which is better than the 85% in Naïve Bayes classifier.



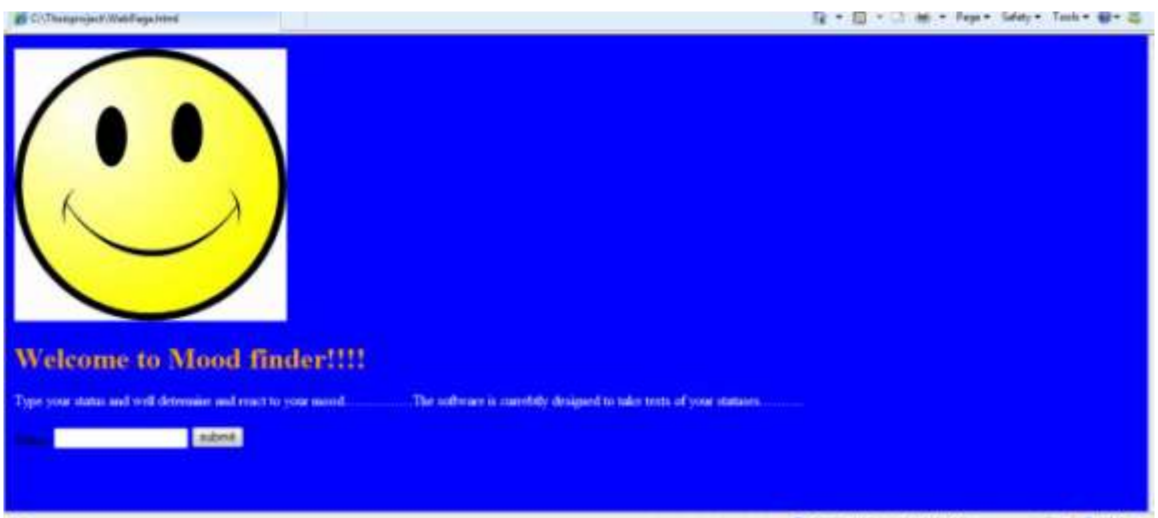


After Training our NLTK we will send User data via <XML> tags to our NLTK and check for responses to our polarity questions. The polarity will be stored in a file from which its will be read by a daemon program. We will also assign a set of action to auto generated text values to identify the values of our actions in the UI.

After reading the file the daemon program will decide on a set of actions to perform. Some of the basic ones will be to change our layouts. Suppose our background to blue when we are sad and White when we are happy. Other's maybe to open web sited change theme’s and maybe even back ground images, Text color as we aren’t able to directly tap into facebook or any other websites actions, we can only suggest predetermined sets of web links from these websites.

4. Implementation

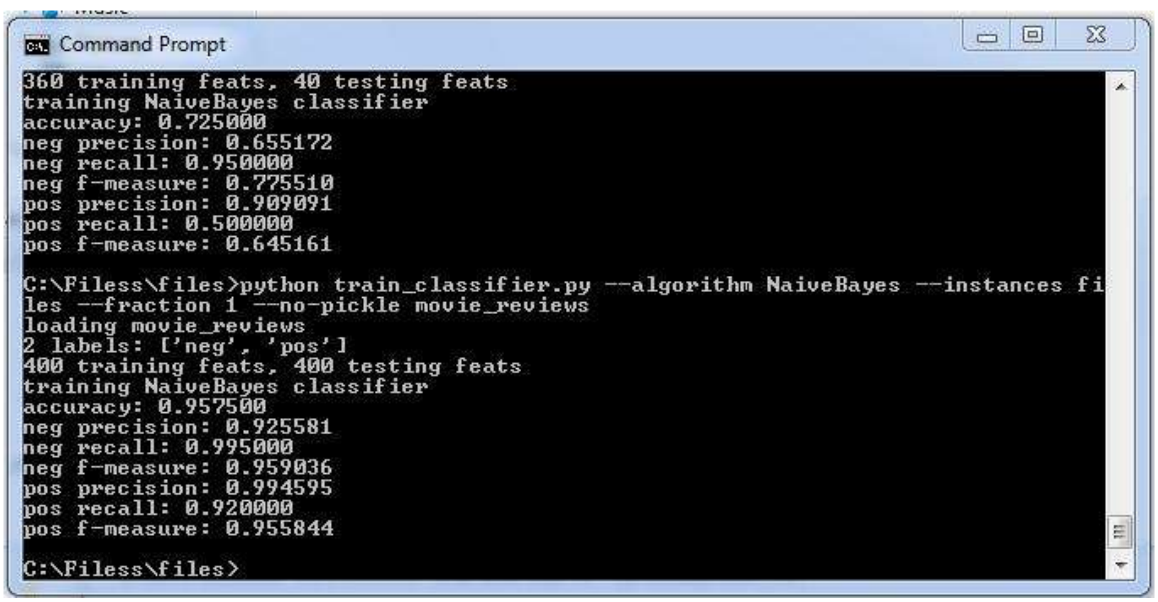
To implement the software, we had installed python nltk 2.0 with supporting packages from nltk.org. The version of python used in the experiment was 2.7, as the classifier software’s we used trained worked specific to that version of python. We first installed the system on a windows 7 machine. Tests on the Movie reviews database ran smoothly for positive and negative sentiments for a Bayesian classifier. After that we tried to move our experiment to a maximum entropy algorithm but due lack of resources it wasn’t possible to use it. However, using bigram featuresets we were able greatly increase accuracy to the point that we did not require it. We collected data online from tweets of the different sentiments. We collected 100 datasets as a base for each of the emotions. We trained classifiers for Positive Vs Negative and trained it for the sub classes Happy Vs Funny and Sad Vs angry.[18][30][31][32]



We used the word punctuation analyzer, which uses mathematical regression to identify the featuresets of the sentences, and compares and find a mood. We created a demo webpage that takes the microblog data and writes to a file using java script. The data is read by the analyzer and then the mood is determined and written to a output file. The ouptput file is read by the webpage to determine which determines which features it should use or not. in our case the webpage options like color, images, options/links etc. The reason we are using file message passing using javascripts now is to make debugging simpler, in practical application we can use php ,<xml> and port commands to help increase functionality.

5. Training Classifier

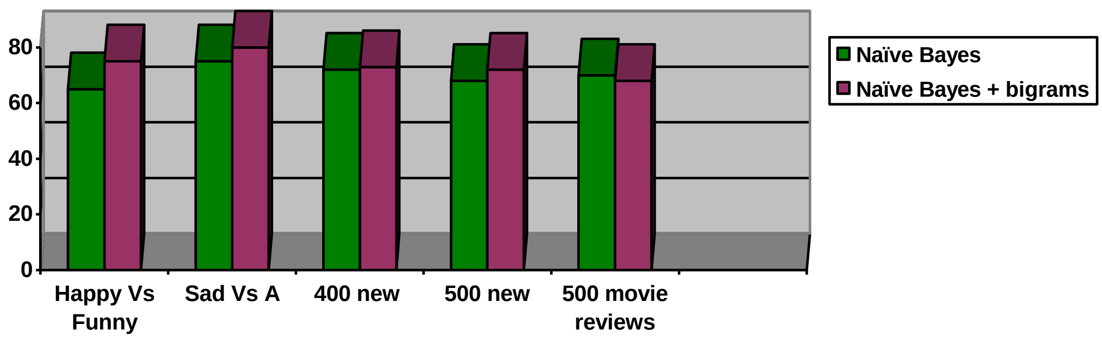
To train the classifier we collected tweets from the internet 100 tweets for happiness, 100 for funny, 100 for sad and 100 for angry. From these we constructed 200 for datasets positive and 200 for negative. We also used very informal data with strong words and smileys such as ☺, ☹, LOL, kill, happy etc. Opposed to the previous corpra used in the movie reviews dataset initial tests showed a lot of promise.[26][27][28][29]



However due to the small size its effectiveness can not be properly measured. But initial reports show stability in the performance of the program. Also when microbloging users don’t always use strong so its would be good to add other data so the range of the analyzer can be tested and prepared for these scenario’s and for this reason with the previous 200 positive datasets we added 50 datasets of positive data from movie reviews corpra and with 200 negative datasets we added 50 datasets of negative data from movie reviews corpra. We also trained a classifier of 500 datasets from original movie reviews corpus to compare results.

6. Results

With a fraction of 0.9 in the training sets the accuracy's are shown below for the different classifications.



The accuracy for Bayesian based bigram classifier for 400 datasets had a distinct advantage over the other 500 datasets new+old and 500 datasets movie reviews classifier. It showed 73% accuracy where as the others showed 72% and 68% of accuracy respectively. As such we used it to build our software as it showed promise. Also the accuracy rate of the happy vs funny and sad vs angry classifiers 75% and 80%

7. Conclusion

The work being done on the topic is vastly narrow and only addresses the issue of USER sentimental analysis and not SNA. Incorporating this will be the next step in achieving better results. Also better incorporation with social Networking sites and other Facilities and supposed android devices can help our program to achieve a more far-reaching experience.

City's, Govt's and corporations are all opting towards creating happier, healthier lives for their citizen's and customer's, And sentimental analysis over social networks and user interfaces are going to start playing crucial parts in this goal. And a very important and unavoidable part for this improvement will be the interface improvements that can be achieved via sentimental analysis of user’s.

Special thanks to BRACU SECS for their help.

My Supervisor, Co-supervisor, Mahafuz Aziz Aveek, Imran Ahmed, Shiblee Imtiaz Hasan, Tasbeer Ahsan, Shammur Absar Chowdhury and repected teachers.

Bibliography

[1] CyberEmotions, http://www.cyberemotions.eu/
[2] Adaptive Websties Wikipedia ,http://en.wikipedia.org/wiki/Adaptive_website
[3] Sentimental analysis Wikipedia,
http://en.wikipedia.org/wiki/Sentiment_analysis

[4] Alexander Pak, Patrick Paroubek , "*Twitter as a Corpus for Sentiment Analysis and Opinion Mining*" , Universit´e de Paris-Sud, Laboratoire LIMSI-CNRS, Bˆatiment 508, F-91405 Orsay Cedex, France alexpak@limsi.fr, pap@limsi.fr

[5] Alec Go, Lei Huang, and Richa Bhayani. 2009. *Twitter sentiment analysis*. Final Projects from CS224N for Spring 2008/2009 at The Stanford Natural Language Processing Group.

[6] Changhua Yang, Kevin Hsin-Yih Lin, and Hsin-Hsi Chen.
2007. *Emotion classification using web blog corpora.*
In WI '07: Proceedings of the IEEE/WIC/ACM International
Conference on Web Intelligence, pages 275–278,
Washington, DC, USA. IEEE Computer Society.

[7] Bo Pang and Lillian Lee. 2008. *Opinion mining and sentiment analysis*. Found. Trends Inf. Retr., 2(1-2):1–135.

[8] Jonathon Read. 2005. *Using emoticons to reduce dependency in machine learning techniques for sentiment classification.*
In ACL. The Association for Computer Linguistics.

[9] Adam Bermingham1, Maura Conway2, Lisa McInerney2,
Neil O'Hare1, Alan F. Smeaton1
1CLARITY: Centre for Sensor Web Technologies
and 2School of Law and Government,
Sentiment Analysis to Explore the Potential for Online Radicalisation .
Dublin City University, Glasnevin, Dublin 9, Ireland, *Combining Social Network Analysis and*

[10] Wojciech Gryc and Karo Moilanen,
Sentiment Analysis of Political Blogs in the 2008 U.S. Presidential Election ,
Oxford Internet Institute, University of Oxford Oxford University.

[11] Chmiel, A., Sienkiewicz, J., Thelwall, M., Paltoglou, G., Buckley, K., Kappas, A. & Hołyst, J.A. (2011). *Collective emotions online and their influence on community life*, PLoS ONE, 6(7): e22207.

[12] Thelwall, M., Buckley, K., & Paltoglou, G. (in press). *Sentiment in Twitter events. Journal of the American Society for Information Science and Technology.*

[13] OpinionFinder: Open Source Sentiment Analysis Toolkit
, http://blog.typeslashcode.com/voxpop/2010/02/opinionfinder-open-source-sentiment-analysis-toolkit/

[14] Is it possible to guess a user's mood based on the structure of text?,
http://stackoverflow.com/questions/933212/is-it-possible-to-guess-a-users-mood-based-on-the-structure-of-text

[15] Sentiment Tutorial,
http://alias-i.com/lingpipe/demos/tutorial/sentiment/read-me.html

[16] The Stanford Parser: A statistical parser,
http://www-nlp.stanford.edu/downloads/lex-parser.shtml

---

[17] Led by Facebook, Twitter, Global Time Spent on Social Media Sites up 82% Year over Year
, nielsenwire , http://blog.nielsen.com/nielsenwire/global/led-by-facebook-twitter-global-time-spent-on-social-media-sites-up-82-year-over-year/

[18] Sentiment Analysis with Python NLTK Text Classification,

http://text-processing.com/demo/sentiment/, Trainer ,
https://github.com/japerk/nltk-trainer

[19] How Companies Can Use Sentiment Analysis to Improve Their Business
, Maria Ogneva , http://mashable.com/2010/04/19/sentiment-analysis/

[20] Applications in Social Media: Sentiment Analysis
, Neil Glassman ,http://socialtimes.com/sentiment-analysis-socialmedia-marketing3-2_b59924

[21] Sentiment analysis using the Dutch Netlog Corpus
, http://mollom.com/sentiment-analysis-dutch-netlog-corpus

[22] Text classification for sentimental Analysis –Naïve Bays Classifier,
http://streamhacker.com/2010/05/10/text-classification-sentiment-analysis-naive-bayes-classifier/

[23] Yuchul Jung, Hogun Park and Sung Hyon Myaeng,
A Hybrid Mood Classification Approach for Blog Text,
http://www.springerlink.com/content/h4371783u704n93x/

[24] Part of speech Tagging ,
http://www1.cs.columbia.edu/~kathy/NLP/ClassSlides/Class5-POS/mypos.ppt

[25] Comparing Data-Driven Learning Algorithms for PoS Tagging of Swedish (2001), http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.12.9840

[26] Facebook Statuses,
http://facebook-statuses.com/

[27] Sarah Palin Death Threats on Tweeter,
http://markhumphrys.com/twitter.palin.html

[28] 100 happy tweets 100 sad tweets from16.44 on Thursday 14/04/2011,
http://www.scribd.com/doc/56012075/100-sample-happy-tweets

[29] The 100 Funniest Tweets of 2010,
http://chirpstory.com/li/143

[30] japerk / nltk-trainer,
https://github.com/japerk/nltk-trainer

[31] Download NLTK 2.0,
http://www.nltk.org/download

[32] Python 2.7,
http://www.python.org/ftp/python/2.7.2/python-2.7.2.msi