Test- Download - Java Version - Non-English - Buy! - About



Automatic sentiment analysis of up to 16,000 social web texts per second with up to human level accuracy for English - other languages available or easily added.

SentiStrength estimates the *strength* of positive and negative sentiment in *short texts*, even for informal language. It has <u>human-level accuracy</u> for short social web texts in English, except political texts. SentiStrength reports *two* sentiment strengths:

-1 (not negative) to -5 (extremely negative)

1 (not positive) to 5 (extremely positive)

Why does it use two scores? Because <u>research from psychology</u> has revealed that we process positive and negative sentiment in parallel - hence mixed emotions.

SentiStrength can also report binary (positive/negative), trinary (positive/negative/neutral) and single scale (-4 to +4) results. SentiStrength was originally developed for English and optimised for general short social web texts but can be configured for other languages and contexts by changing its input files - some variants are demonstrated below.

Quick To	ests (En	glish v	ersion)	:
----------	-----------------	---------	---------	---

Enter text: I love you but hate the current political climate.	Detect Sentiment	
Output: ● Dual, ○ binary, ○ trinary, ○ scale		
Keyword test:		
I love dogs quite a lot but cats I really hate.		
Enter keywords (comma-separated list, no spaces): dogs,frog	gs	
Detect Sentiment Around Keywords		
Topic test:		
The mayie was very frightening. My phone has a small across. That is a hanging	shuuuun Luvan tha vata but ha liad abau	+ n

The movie was very frightening. My phone has a small screen. That is a banging chuuuun. I won the vote but he lied about me.

Select domain (broad topic): Films/Movies ▼ Detect Sentiment in Domain

Other languages: Finnish, German, Dutch Spanish. Russian, Portuguese, French, Arabic, Polish, Persian, Swedish, Greek, Welsh, Italian, Turkish.

Download SentiStrength

SentiStrength is free for academic research and is <u>certified safe by Softpedia</u>. Please contact the author for the commercial Java version or a commercial license for the online version. The free version runs under Windows only and is provided without liability or guarantees for any uses. Downloading SentiStrength and/or the configuration files signifies acceptance of these conditions. This version does not contain the keyword or domain classification facilities.

- Register for the free <u>SentiStrength download</u> make sure to save the program AND the data files (this does not do the binary/trinary/scale classifications).
- if SentiStrength gives an error message when starting, please try downloading and installing the "Microsoft .NET Framework Redistributable Package" and then try running SentiStrength again.
- Remember to use Register New Location in the File menu to point SentiStrength to the location of the data files as soon as it loads, unless they are saved in the default location C:\SentStrength_Data\.

Buy SentiStrength

sentistrength.wlv.ac.uk 1/8

A commercial licence for SentiStrength is available for £1000 - please contact m.thelwall -at- wlv.ac.uk. The <u>Java version</u> of SentiStrength is normally used commercially.

SentiStrength is used by computing, language technology and market research companies in the US, Europe and Australia. Some use the default English version and others have translated it into different languages or adopted it to integrate with their existing language technology systems. Commercial users range from small start-ups to one of the world's top 10 largest corporations.

Java Version

The Java version of SentiStrength is similar to the Windows version in core functions but has additional capabilities - see the <u>SentiStrength Java manual</u> (updated February 2017) and <u>Mac users' starting instructions</u> (also helps in Linux probably) and <u>command list</u>. It can conduct binary (positive/negative), trinary (positive/neutral/negative), single-scale classifications (-4 very negative to very positive +4) in addition to the standard type, and can conduct keyword-oriented and domain-oriented classifications. It also has a special mode for binary and trinary classification on longer texts. It allows wildcards in the idiom list file. To use the Java version for research only (free), email from your academic email address. It can process about 16,000 tweets per second.

For RJB users, here is some <u>sample RJB code</u> from <u>Adam Pantanowitz</u>, University of the Witwatersrand.

For Python users, here is some <u>sample Python code</u> from Alec Larsen, University of the Witwatersrand. For Python 3+ try this <u>sample Python code</u> (or <u>text version</u>) variant instead. Unizip this <u>Jupyter Notebook</u> to use.

For GATE users, here is a <u>GATE wrapper</u> to import SentiStrength from <u>Mark</u> <u>Greenwood</u> and <u>Diana Maynard</u>, University of Sheffield. <u>Instructions for using the GATE plugin from Alex ibollit Stepanenko.</u>

For Weka users, there is a SentiStrength wrapper for the <u>AffectiveTweets Weka</u> <u>package</u> by <u>Felipe Bravo-Marquez</u> of Waikato University that can be installed via the <u>WekaPackage manager</u>.

Thank you to Sooyeon Jeong of MIT for fixing SentiStrength to work for Android apps.

About SentiStrength

SentiStrength is a sentiment analysis (opinion mining) program. It is described and evaluated in the following peer-reviewed academic articles:

- Thelwall, M., Buckley, K., Paltoglou, G. Cai, D., & Kappas, A. (2010). <u>Sentiment strength detection in short informal text</u>. *Journal of the American Society for Information Science and Technology*, 61(12), 2544–2558.
- Thelwall, M., Buckley, K., & Paltoglou, G. (2012). <u>Sentiment strength detection</u> for the social Web, Journal of the American Society for Information Science and Technology, 63(1), 163-173.
- Thelwall, M., & Buckley, K. (2013). <u>Topic-based sentiment analysis for the Social Web: The role of mood and issue-related words</u>. *Journal of the American Society for Information Science and Technology*, 64(8), 1608–1617.
- Thelwall, M., & Buckley, K., Paltoglou, G., Skowron, M., Garcia, D., Gobron, S., Ahn, J., Kappas, A., Küster, D. & Holyst, J.A. (2013). Damping Sentiment Analysis in Online Communication: Discussions, monologs and dialogs. In: A. Gelbukh (Ed.): CICLing 2013, Part II, LNCS 7817, pp. 1-12. Springer, Heidelberg.
- [Turkish version] G. Vural, B. B. Cambazoglu, P. Senkul, and O. Tokgoz (2013) <u>A framework for sentiment analysis in Turkish: Application to polarity detection of movie reviews in Turkish</u>, Computer and Information Sciences III, pp 437-445.

sentistrength.wlv.ac.uk 2/8

- Thelwall, M. (2017). <u>Heart and soul: Sentiment strength detection in the social web with SentiStrength</u> (summary book chapter). In: Holyst, J. (Ed.)
 Cyberemotions: Collective emotions in cyberspace. Berlin, Germany: Springer (pp. 119-134). doi:10.1007/978-3-319-43639-5_7
- [Spanish version] Vilares Calvo, D., Thelwall, M., & Alonso, M.A. (2015). The megaphone of the people? Spanish SentiStrength for real-time analysis of political tweets. Journal of Information Science. [Introduces an improved Spanish sentiment strength detection version of SentiStrength and shows Spanish political tweets tend to amplify news stories.]
- Thelwall, M. (2018). <u>Gender bias in sentiment analysis</u>. *Online Information Review*, 42(1), 45-57.
- Culpeper, J., Findlay, A., Cortese, B. & Thelwall, M. (2018). <u>Measuring emotional temperatures in Shakespeare's drama</u>. English Text Construction, 11(1), 10-37. [early modern English version]

It has been applied in the following research projects, amongst others.

- Thelwall, M., Buckley, K., & Paltoglou, G. (2011). <u>Sentiment in Twitter events</u>. Journal of the American Society for Information Science and Technology, 62(2), 406-418.
- Kucuktunc, O., Cambazoglu, B.B., Weber, I., & Ferhatosmanoglu, H. (2012).
 <u>A large-scale sentiment analysis for Yahoo! Answers</u>, Proceedings of the <u>5th ACM International Conference on Web Search and Data Mining</u>.[Used in Yahoo!]
- Weber, I, Ukkonen, A., & Gionis, A. (2012). <u>Answers, not links: extracting tips from yahoo! answers to address how-to web queries</u>, Proceedings of the fifth ACM international conference on Web search and data mining (WSDM '12). [Used in Yahoo!]
- Pfitzner, R., Garas, A., & Schweitzer, F. (2012). <u>Emotional divergence influences information spreading in Twitter</u>, ICWSM-12.
- Garas, A., Garcia, D., Skowron, M., & Schweitzer, F. (2012). <u>Emotional</u> <u>persistence in online chatting communities</u>. *Scientific Reports*, 2, article 402.
- Mihai Grigore and Christoph Rosenkranz (2011). <u>Increasing the willingness to collaborate online: an analysis of sentiment-driven interactions in peer content production</u>. ICIS 2011 Proceedings. Paper 20.
- G. Vural, B. B. Cambazoglu, and P. Senkul (2012). <u>Sentiment-focused web crawling</u>, Proceedings of the 21st ACM International Conference on Information and Knowledge Management, pp 2020-2024.
- Giorgos Giannopoulos, Ingmar Weber, Alejandro Jaimes, Timos Sellis (2012).
 <u>Diversifying User Comments on News Articles</u>, Web Information Systems
 Engineering (WISE 2012). Lecture Notes in Computer Science 2012, pp 100-113.
- From Greg Merritt: New Cities Foundation (2012), <u>Connected Commuting:</u>
 <u>Research and Analysis on the New Cities Foundation Task Force in San Jose</u>
 (SentiStrength is mentioned on page 16). See also <u>Crowdsourcing your</u>
 <u>Commute</u> (New York Times).
- Witherspoon, C. & Stone, D.(2013). Analysis and Sentiment Detection in Online Reviews of Tax Professionals: A Comparison of Three Software Packages, *Journal of Emerging Technologies in Accounting*, 10(1), 89-115.
- Zheludev, I., Smith, R., & Aste, T. (2014). When can social media lead financial markets?. Scientific Reports, 4. doi:10.1038/srep04213.
- Durahim, A. O., & Coşkun, M. (2015). <u>#iamhappybecause: Gross National Happiness through Twitter analysis and big data</u>. *Technological Forecasting and Social Change*, 99, 92-105. [Uses the Turkish version]
- Calefato, F. Lanubile, F. & Novielli, N. (2018). <u>How to ask for technical help?</u> <u>Evidence-based guidelines for writing questions on Stack Overflow</u>. *Information and Software Technology Journal, 94, 186-207.*.
- Livas, C., Delli, K., & Pandis, N. (2018). "My Invisalign experience": content, metrics and comment sentiment analysis of the most popular patient testimonials on YouTube. Progress in Orthodontics, 19(1), 3.
- Culpeper, J., Archer, D., Findlay, A. & Thelwall, M. (2018). <u>John Webster, the dark and violent playwright?</u> ANQ: A Quarterly Journal of Short Articles Notes and Reviews, 31(3). 201-210. doi:10.1080/0895769X.2018.1445515

sentistrength.wlv.ac.uk 3/8

- Orasan, C. (2018). <u>Aggressive Language Identification Using Word Embeddings</u> <u>and Sentiment Features</u>. In Proceedings of the First Workshop on Trolling, Aggression and Cyberbullying (TRAC-2018), p. 113 - 119, Santa Fe, USA, August 25. [<u>Software used</u>]
- Baviera, T., Sampietro, A., & García-Ull, F. J. (2019). <u>Political conversations on Twitter in a disruptive scenario: The role of "party evangelists" during the 2015 Spanish general elections</u>. *The Communication Review*.

SentiStrength is part of various social media analysis software suites, including:

- Mozdeh Big data text analysis (Windows)
- COSMOS Open Data Analytics software (multi-platform)
- Chorus data harvesting and visual analytics suite for Twitter (Windows)

Press coverage and initiatives

- Reading the Riots: Investigating England's Summer of Disorder Guardian online.
- SentiStrength classified **London Olympics** tweets with the results put up in lights on the EDF Energy London Eye.YouTube videos: <u>Barge 1</u>, <u>Barge 2</u>, <u>Inside Barge</u>, <u>LondonEye lightshow</u>.
 - Time Magazine: Want to Light Up the London Eye? Just Tweet That the Olympics Are 'Totes Amazeballs', July 27, 2012.
 - Wall Street Journal: <u>Hidden Message in the London Eye</u>.
 - UK Daily Telegraph article, p. 27, 19 July 2012, "<u>Happy Olympic tweeters to light up London Eye</u>" in "the world's first social media driven light show".
 - BBC News Article: 20 July 2012, <u>London Eye Olympic Twitter positivity</u> <u>lightshow launched</u>.
 - UK Daily Mirror article, 20 July 2012, <u>The mood of the nation: Tweets to power spectacular London 2012 light show.</u>
 - Sydney Morning Herald: <u>London Eye to become giant Twitter mood ring during Olympics</u>, July 25, 2012.
 - Voice of Russia: <u>Twitter embraces Olympics</u>, <u>colors tweets in emotions and sends messages from space</u>, July 24, 2012.
 - SportPrimeur, 20 July 2012, London Eye twitter-uitlaatklep tijdens Spelen.
- SentiStrength classified tweets for the **2014 Super Bowl**, with the results transformed into a <u>lightshow on the Empire State Building</u>.
 - NBC New York, Jan 28, 2014: <u>Fans to Pick Colors in Empire State Building Super Bowl Light Shows</u>.
 - CBS Local, Jan 27, 2014: <u>Start Tweeting! Super Bowl Debates To Light Up Empire State Building</u>.
 - Forbes Jan 30, 2014: <u>Verizon's Super Bowl Scheme Is To Save \$4 Million And Light Up The Sky</u>.
 - LA Times Jan 30, 2014: <u>Empire State Building lights up for Super Bowl</u>, <u>Chinese New Year</u>.
 - Belfast Telegraph, Jan 29, 2014: <u>Empire State building lights up in Super</u> Bowl team colors.
 - Fox CT, Jan 29, 2014: <u>#WhosGonnaWin? Tweets Pick Colors Of Empire State Building</u>.

Classifying texts with SentiStrength

To get SentiStrength to classify one or more texts, put the texts into a plain text file with one text per line. Select *Analyse All Texts in File* from the *Sentiment Strength Analysis* menu and select the text file. The output will be a copy of the file with positive and negative classifications added at the end of each line, preceded by tabs. Individual texts can also be classified by selecting *Analyse One Text* from the *Sentiment Strength Analysis* menu. <u>Sets of text files</u> can also be processed.

Optimising SentiStrength term weights

The term positive and negative weights can be found in the *EmotionLookupTable.txt* file in the *SentStrength_Data* folder. These can be manually adjusted by editing the file. Alternatively, they can be automatically fine-tuned with a classified text

sentistrength.wlv.ac.uk 4/8

collection. To fine tune *EmotionLookupTable.txt* values used by SentiStrength, first create a collection of texts that have been classified by humans with positive (1-5) and negative (1-5) sentiment strengths. Put these into a plain text file in which each line has the format: positive – tab – negative – tab – text (negative sentiment can be entered as either positive or negative integers). The set should be at least 500 texts. Select *Optimise the emotion dictionary weights* from the *Sentiment Strength Analysis* menu and SentiStrength will create a new term strength list that is optimised for the sentiment in the new texts. To use the new strengths, save a copy of the original strength list and then replace it with the new list.

Assessing the accuracy of SentiStrength

To assess the accuracy of SentiStrength on a set of texts, a sample must first be classified and formatted as above. The human classifications can then be compared with the SentiStrength classifications on the same sample.

Alternatively, if one data set is available to optimise the word strength list and the same set is to be used for validation then the 10-fold cross-validation procedure can be used. This uses 90% of the data to train the term weights and the remaining 10% to assess the accuracy of the adjusted weights. This is repeated 10 times with a different 10% left out and the total results are reported. To run a 10-fold cross-validation, create the classified text as above and select *Run a 10-fold cross-validation to assess the above algorithm* from the *Sentiment Strength Analysis* menu.

Extra resources:

- <u>Six sets of at least 1000 human coded texts</u>, each coded by three independent coders.
- <u>Coding manual for sentiment in texts</u> (<u>pdf version</u>) (for Tweets, but easily changed for other texts)
- SentiStrength based 6-hour sentiment analysis course.

Support files

The various files with SentiStrength contain information used in the algorithm and may be customised.

- The EmotionLookUpTable is just a list of emotion-bearing words, each one with the word then a tab, then an integer 1 to 5 or -1 to -5. This can be edited and extended. Note that strengths of +1 and -1 have no effect on the program. There are some in the list, just to indicate that they words have been considered but not used. Each word can end with a wild card * but this can only go at the end.
- The EmoticonLookUpTable is as above but for a list of emoticons.
- EnglishWordList.txt is just a list of English words it is used for the part of the algorithm that tries to correct words with non-standard spellings.
- NegatingWordList.txt reverses the polarity of subsequent words -e.g., not happy is negative.
- BoosterWordList.txt increases sentiment intensity -e.g., very happy is more positive than happy.
- SlangLookupTable.txt replaces common slang with equivalent words or expressions
- IdiomLookupTable.txt-overrides the sentiment strength of the individual words in the phrase

Language customisation

SentiStrength can be adjusted for other languages by translating the term list *EmotionLookupTable.txt* and adding any other sentiment-bearing words that have been omitted. *Note that the sentiment scores for terms should be in the range 2 to 5 (positive) or -2 to -5 (negative). A score of +1 or -1 means neutral and neutral terms are ignored.* A training corpus in the new language is recommended to help adjust the term weight strengths (see Optimising SentiStrength term weights). You will need the Java version or Windows version **2.2** of SentiStrength to cope with accented

sentistrength.wlv.ac.uk 5/8

characters or characters not found in English as well as some additional linguistic features.

The following files will also need to be translated or replaced with a local equivalent (see the <u>extra instructions</u>):

- Emot**icon**LookupTable.txt check the strengths are appropriate and add any common new national variations
- SlangLookupTable.txt replace with a list of common slang in the new language
- EnglishWordList.txt replace with a word list of correct spellings in the new language (many such lists are on the web, but this step is optional)
- NegatingWordList.txt translate/replace with a list of negating words in the new language
- IdiomLookupTable.txt-replace with a list of common idioms in the new language
- BoosterWordList.txt translate/replace with a list of booster words in the new language – words that emphasise the strength of emotion in any subsequent words
- QuestionWords.txt- translate/replace with a list of words in the new language that reliably indicate that a question is being asked

You will also need to register a list of non-English common multiple letters (e.g., ii is common in some languages but not English). For the Java version please see the manual for this option. For the Windows version, please check the options menu for this customisation. Spell-checking can also be completely disabled in both versions, if needed.

Negating words occurring after sentiment words (e.g., "I am happy not" is OK in German but not English) can be customised in the Java version of SentiStrength but not the Windows version, sorry. The Java version may need the utf8 option to read the input files, if in UTF8 rather than ASCII format (note that utf8 does not always work on ANSI text files so it should not be used as the default).

SentiStrength versions for other languages

Would you like to help? If you are a linguist with knowledge of any of these languages then you could help by:

- Checking the dictionaries for accuracy and missing sentiment words
- Reporting any badly classified texts.

Please email **m** dot **thelwall** at **wlv.ac.uk** if you would like to help. This makes a good student project.

Classifiers with some testing (8)

Enter text: Minä rakastan sinua, mutta inhoan nykyistä poliittista ilmapiiriä.

Enter text: Ich liebe dich aber hasse das derzeitige politische Klima.

Created by Hannes Pirker, Interaction Technologies Group at the Austrian Research Institute for Artificial Intelligence (OFAI), with additions from Elias Kyewski of the University of Duisburg. The German version can be obtained via this download form Enter text: Ik hou van je, maar haat het huidige politieke klimaat.

Enter text: Te quiero, pero odio el clima político actual.

Detect Sentiment in Finnish Detect Sentiment in German

Detect Sentiment in Dutch Detect Sentiment in Dutch

Created by <u>David Vilares</u> from previous versions made by <u>David Garcia</u>. Incorporates terms from <u>Ancora</u> and <u>SOCAL</u>. Human-coded <u>Spanish tweets sentiment strength test and development sets</u>. See also an English-Spanish <u>code-switching corpus</u> from this <u>paper</u>.

Enter text: lo ti amo ma odio l'attuale clima politico.

Made with pride in the European Union by <u>Dr. Eng. Cosmo "Pyo" DI MILLE and his team.</u>

sentistrength.wlv.ac.uk 6/8

Enter text: Нашла дневник, который я вела в 8 классе) ужас, что творил Detect Sentiment in Russian Enter text: Нашла дневник, который я вела в 8 классе) ужас, что творил Detect Sentiment in Russian2

Thank you to Eismont Polina, Efanova Iuliia, Konovalova Svetlana, Losev Viktor and Velichko Alena of Saint Petersburg State University of Aerospace Instrumentation, Department of Applied Linguistics for help with the first Russian version. (+ve correl. 0.28-0.47, -ve correl. 0.31-0.46 on tweets - the second number is overfitted due to testing on the evaluation data set, so the real correlation is probably about 0.35 for both). 3972 human-classified Russian tweets.

Thank you to Юлия Павлова, Olessia Koltsova and Sergei Koltsov for the second Russian version. It was developed by the <u>Laboratory for Internet Studies</u>, National Research University Higher School of Economics (NRU HSE), and supported by the Russian Humanitarian Research Foundation and NRU HSE.

There is also a Turkish sentiment strength classifier that is a variant of SentiStrength created by Gural VURAL, METU Computer Eng. Dept. This is available on the same basis as the <u>Java</u> version.

Enter text: Seni seviyorum fakat mevcut politik havadan nefret ediyorum. Detect Sentiment in Turkish (this version includes only part of Gural VURAL's system so its results are not as good as Gural VURAL's full version.)

Enter text: Cowards die many times before their deaths; the valiant never ta Detect Sentiment in Early Modern English

Created by <u>Jonathan Culpeper</u>, <u>Alison Findlay</u> and <u>Beth Cortese</u> of Lancaster University. See <u>paper</u>.

Citation context classifiers are also available.

Hindi. SentiStrength cannot cope with the unicode marks in Hindi but Ashutosh Khanna has created <u>Hindi resources and code</u> that can be used for a similar sentiment analysis. He welcomes additions, comments or suggestions (ashutosh.khanna26 [at] gmail.com).

Completely untested classifiers (9) [just for fun: please email m dot thelwall at wlv.ac.uk if you would like to help improve them - this makes a good student project for linguists or computer scientists, together with testing the results, and making a small corpus of sentiment-classified texts! Here are the <u>language files</u> for 9 of these languages - please improve them and send back if you like. Please also send us a few positive and negative words from any languages not listed here and we will make a new version for your language!]

Enter text:	Je t'aime mais je déteste le climat politique actuel.	Detect Sentiment in French
Enter text:	Kocham cię, ale nienawidzę obecny klimat polityczny.	Detect Sentiment in Polish
Enter text:	Σ 'αγαπώ, αλλά το μίσος το σημερινό πολιτικό κλίμα.	Detect Sentiment in Greek
Enter text:	Ina son ki amma kin ki jini na yanzu siyasa sauyin yanayi.	Detect Sentiment in Hausa
Enter text:	Jag älskar dig men hatar det rådande politiska klimatet.	Detect Sentiment in Swedish
Enter text:	Eu te amo, mas odeio o clima político atual.	Detect Sentiment in Portuguese
Enter text:	من شما را خیلی دوست دارم اما از وضعیت سیاسی فعلی متنفر هستم	Detect Sentiment in Persian
Enter text:	أنا أحبك ولكن أكره المناخ السياسي الحالي	Detect Sentiment in Arabic
Enter text:	Rwy'n dy garu di, ond dwi'n casu'r sefyllfa wleidyddol bresennol.	Detect Sentiment in Welsh
Enter text:	Is breá liom tú, ach is fuath leis an aeráid reatha polaitiúla.	Detect Sentiment in Irish

Thank you to Usman A. Adam (Kaduna State University) for the Hausa version.

Basic classifiers (10) that recognise only a few sentiment words. Please email m dot thelwall at wlv.ac.uk if you would like to help improve them or to send a list of at least 10 common sentiment words for any language. We can't get Hindi, Punjabi and Bengali to work at the moment, sorry (see above for a partial <u>Hindi</u> solution). This is due to word segmentation issues related to these languages containing characters

sentistrength.wlv.ac.uk 7/8

classified as 'marks' in unicode. Also, the Chinese simplified and traditional and the Japanese are artificial versions that add spaces between words or phrases into the language.

Enter text: Wǒ ài nǐ, dàn wǒ bù xǐhuān dāngqián de zhèngzhì qìhòu. Detect Sent. in Chinese-Pinyin Enter text: 我爱你,但我不喜欢 当前的政治气候。 Detect Sent. in Chinese-Simplified Enter text: 我愛你,但討厭目前的政治氣候。 Detect Sent. in Chinese-Traditional Enter text: Mahal kita pero ayaw mo sa kasalukuyang pampulitika klima. Detect Sentiment in Filipino Enter text: Aku mencintaimu tapi benci iklim politik saat ini. Detect Sentiment in Indonesian Enter text: 私はあなた を愛して いるけれど、現在の政治情勢を 憎む。 Detect Sentiment in Japanese Enter text: 내가 당신을 사랑하지만 현재의 정치 사정 싫어. Detect Sentiment in Korean Enter text: Ndino da iwe asi venga iripo zvenyika. Detect Sentiment in Shona Enter text: Nakupenda lakini chuki ya hali ya hewa ya sasa ya kisiasa. Detect Sentiment in Swahili

Domain customisation

SentiStrength can be adjusted for other domains (e.g., Twitter, product reviews) by adding new relevant words and sentiment strengths to the term list <code>EmotionLookupTable.txt</code> and adjusting any relevant existing term strengths. The other files can also be adjusted, as for language customisation. For example, the file <code>EmotionLookupTableGeneral.txt</code> in the download zipfile contains a slightly adjusted set of term weights to cope with more impersonal communication than MySpace. In this alternative file, the word "love" has a higher strength because it is less likely to be used in formulaic message endings, such as "love from" or "love u" or "love x".

Stress and relaxation

SentiStrength's sister program <u>TensiStrength</u> detects the strength of stress and relaxation in social media texts.

Data mining

The data mining menu and ARFF menu items are not part of the main SentiStrength functionality nor documented. Please ignore them unless they make sense to you.

Other

For further issues, please see the <u>Frequently Asked Questions</u>.

SentiStrength was produced as part of the <u>CyberEmotions</u> project, funded by the EU (FP7).

sentistrength.wlv.ac.uk 8/8