


Sentiment Analysis – SentiStrength

Για την ανάλυση συναισθήματος χρησιμοποιήθηκε το lexicon based εργαλείο SentiStrength. Το SentiStrength είναι διαθέσιμο στην ιστοσελίδα <http://sentistrength.wlv.ac.uk/>.

sentistrength.wlv.ac.uk

[Test-](#) [Download](#) - [Java Version](#) - [Non-English](#) - [Buy!](#) - [About](#)

 **SentiStrength**

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 [human-level accuracy](#) 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 [research from psychology](#) 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 Tests (English version):

Enter text:

Output: ☒ Dual, ☐ binary, ☐ trinary, ☐ scale

Keyword test:

Enter keywords (comma-separated list, no spaces):

Topic test:

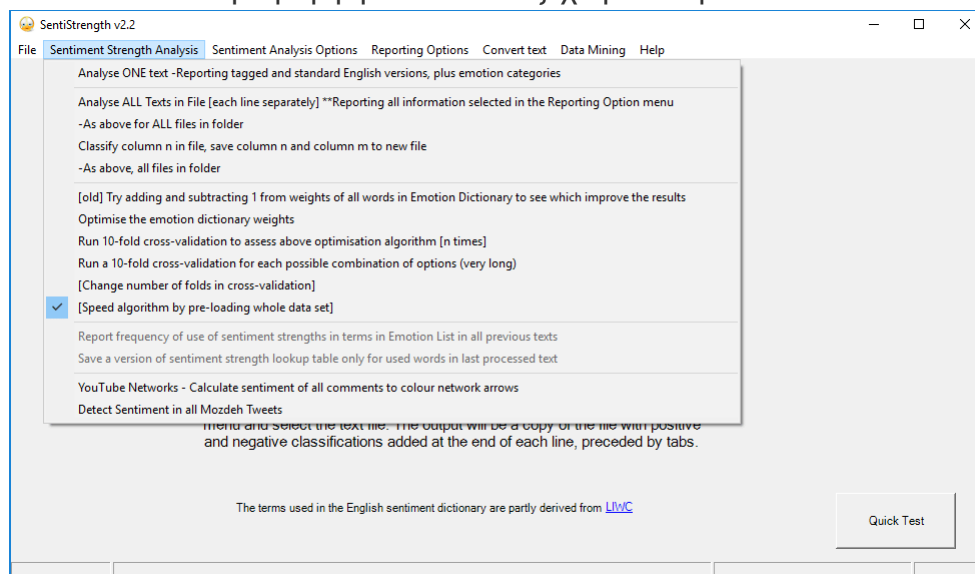
Select domain (broad topic):

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

Το εργαλείο μετράει το συναίσθημα σε απλό κείμενο με τιμές από -5 έως +5. Όπου το -1 έως -5 είναι negative Sentiment και το 1 έως 5 είναι positive Sentiment. Το 0 θεωρείται Neutral. Εκτός από την εισαγωγή data sets υπάρχει και η δυνατότητα quick test με την εξαγωγή συναισθήματος από μια μόνο γραμμή κειμένου.


```
[ItemID,Sentiment,Text]
1,0,      is so sad for my APL friend.....
2,0,      I missed the New Moon trailer...
3,1,      omg its already 7:30 :O
4,0,      .. Omigaga. Im soo0 im gunna CRY. I've been at this dentist since 11.. I was suposed 2 just get a crown put on (30mins)...
5,0,      i think mi bf is cheating on me!!!          T_T
6,0,      or i just worry too much?
7,1,      Juuuuuuuuuuuuuuuuussssst Chillin!!
8,0,      Sunny Again           Work Tomorrow :-|           TV Tonight
9,1,      handed in my uniform today . i miss you already
10,1,     hmmm.... i wonder how she my number @-)
11,0,     I must think about positive..
12,1,     thanks to all the haters up in my face all day! 112-102
13,0,     this weekend has sucked so far
14,0,     jb isnt showing in australia any more!
15,0,     ok thats it you win.
16,0,     &lt;t;----- This is the way i feel right now...
17,0,"    awihe man.... I'm completely useless rt now. Funny, all I can do is twitter. http://myloc.me/27HX"
18,1,     Feeling strangely fine. Now I'm gonna go listen to some Semisonic to celebrate
19,0,     HUGE roll of thunder just now...SO scary!!!!
20,0,     I just cut my beard off. It's only been growing for well over a year. I'm gonna start it over. @shaunamanu is happy in the meantime.
21,0,     Very sad about Iran.
22,0,     womppppp wompp
23,1,     You're the only one who can see this cause no one else is following me this is for you because you're pretty awesome
24,0,     &lt;t;---Sad level is 3. I was writing a massive blog tweet on Myspace and my comp shut down. Now it's all lost "lays in fetal position"
25,0,     ... Headed to Hospitol : Had to pull out of the Golf Tourney in 3rd place!!!!!!!!!!!!!! I Think I Re-Ripped something !!! Yeah THAT !!
26,0,     BoRinG ): whats wrong with him?? Please tell me..... :/-
27,0,     can't be bothered. i wish i could spend the rest of my life just sat here and going to gigs. seriously.
28,0,"    Feeling like shit right now. i really want to sleep, but nooo I have 3 hours of dancing and an art assignment to finish. "
29,1,"    goodbye exams, HELLO ALCOHOL TONIGHT "
30,0,     I didn't realize it was THAT deep. Geez give a girl a warning atleast!
31,0,     I hate it when any athlete appears to tear an ACL on live television.
32,0,     i miss you guys too i think i'm wearing skinny jeans a cute sweater and heels not really sure what are you doing today
33,0,     -- Meet your Meat http://bit.ly/15SSCI
```

Το αρχείο τώρα είναι έτοιμο να γίνει import στο εργαλείο του SentiStrength. Για την εξαγωγή του Sentiment χρησιμοποιήθηκε η επιλογή “Analyse All Texts in File (each line separately)”. Αυτό επιλέχθηκε γιατί το κάθε tweet στην μορφή .txt είναι ξεχωριστά μόνο του σε κάθε γραμμή.



Το SentiStrength αποτελεί ένα πολύ αξιόπιστο εργαλείο καθώς προσφέρει μια πληθώρα επιλογών για την εξαγωγή Sentiment από κείμενο. Επίσης παρέχει την δυνατότητα για εξόρυξη πληροφορίας (Data Mining). Για την παρούσα εργασία προτιμήθηκε να κρατηθούν οι default επιλογές. Κάποιες από αυτές τις επιλογές περιλαμβάνουν τα επαναλαμβανόμενα γράμματα να αυξάνουν το Positive Sentiment , λέξεις όπως το “very” να αυξάνουν το Positive Sentiment ή το “some” να μειώνουν το Negative Sentiment. Επίσης προσφέρει την δυνατότητα να διορθώνει τυχόν επαναλαμβανόμενα λάθη στην ορθογραφία.

