

# Enhanced Sentiment Analysis using Software Engineering with Machine Learning Algorithm

J.Satish Babu, B.Jahnavi, P.SaiAnusha, S.Sai Manisha, M.N Darshini, G. Krishna Mohan

**Abstract:** Sentiment analysis has been related with different software engineering (SE) tasks, for example, Studies show that sentiment analysis instruments give dangerous outcomes when utilized out-of-the-compartment, since they are not expected to process SE datasets. The silver slug for a profitable use of sentiment analysis instruments to SE datasets may be their customization to the particular use setting. To achieve our objective, we retrained on a lot of 40k physically checked sentences/words secluded from Stack Overflow, a front line thought examination instrument misusing noteworthy learning. Despite such an exertion and dull preparing system, the outcomes were negative. We changed our concentration and played out a careful examination of the precision of these gadgets on a gathering of SE datasets.

**Index Terms:** Sentiment Analysis, Software Engineering, NLP.

## I. INTRODUCTION

SA is the examination of the subjectivity (unbiased versus truly stacked) and farthest point (positive versus negative) of a substance [1]. It depends upon sentiment lexicons, that is, enormous conglomeration of words, each commented independently positive or negative introduction (i.e., earlier uttermost point). Advancing examinations propose approaches for redesigning software progress, sponsorship, and headway by applying thought examination on Stack Overflow[2], application audits[3], and tweets containing remarks on programming applications.

Further research on organizers' feelings watches out for the action of effect in social software engineering, by applying sentiment analysis to the substance open in synergistic movement conditions, such as GitHub[4],[5],[6], and Stack Over Flow.

With a striking couple of phenomenal cases, precise software engineering examinations have misused off-the-rack ,sentiment analysis tools that have been set up on non-programming building records, for example, film

surveys, or posts slithered from broadly important online frameworks organization, for example, Twitter and YouTube. Show how the decision of the estimation examination mechanical gathering may impact the end realness of positive examinations by playing out a benchmarking study on seven datasets, including trades and remarks from Stack Overflow and issue trackers[7]. By separating the evaluations of completely utilized off-the-rack sentiment analysis tools, they demonstrate that not just these mechanical congregations, don't concur with human illumination of pros' correspondence channels[8], in any case they in like way shift among themselves. Another test to address is that the negative tendency of existing sentiment analysis tools devices that are the mis-classification of reasonable explicit messages as genuinely negative. It is especially the situation of bug reports or issue depictions [10], [9]. Novielli et al. This show how sentences like "What is the most ideal approach to manage manslaughter a basic framework" or "I am feeling the death of an area yet I don't have the foggiest idea where" are wrongly relegated negative in light of the way that both 'to butcher' and 'missing' hold a negative farthest point in the SentiStrength word reference.

This affirmation is strong with the monstrosity is-use suspicion that the feeling of a verbalization is completely compelled by its setting of use[11]. From the below figure .we are taking the tweets from different users, by classifying all the tweets we will say that the given tweets are positive, negative or stressed tweets. Users can view all the details and their feeling. The input design is in between users and the information system. It specifies the manner in which data enters the system for processing. It can ensure the reliability of the system and produce reports from accurate data or it may result in output of error information. It will take input from both the software engineering and machine learning, by these two it will perform sentiment analysis and it will check whether the given tweet is positive or negative tweet.

**Revised Manuscript Received on April 25, 2019.**

**J.SatishBabu**, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, A.P., India.

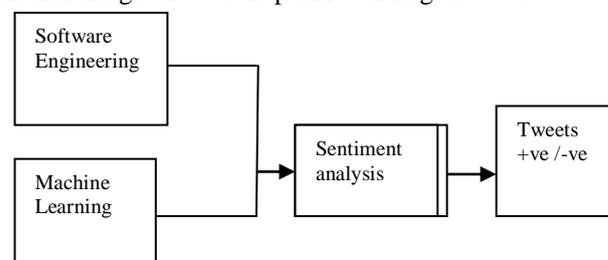
**B.Jahnavi**, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, A.P., India.

**P.SaiAnusha**, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, A.P., India.

**S.Sai Manisha**, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, A.P., India

**M.N Darshini**, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, A.P., IndiaThird

**G. Krishna Mohan**, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, A.P., India



**Fig. 1** Architecture Diagram



## II. LITERATURE SURVEY

Sentiment analysis has been related on various software engineering ancient rarities, for example, explicit knick-knacks (e.g., issues and submit messages) and assembling made substance (e.g., talk messages and clients' surveys), to help different errands. Also, the creators found that those remarks made on Mondays will when all is said in done express continuously negative feelings. A close report was driven by on 28,466 activities inside a multiyear time span. At long last, some of the researchers isolated the relationship between makers' sentiment and gathers performed by dependable joining servers. They found that negative sentiment both aspects and is accepted by the postponed result of the make methodology.

Looking at the farthest purpose of employments' reviews is especially imperative to help the progress of versatile applications. Sentiment analysis gadget suite to in excess of 5,000 outlines seeing that supposition examination can pass on current bottlenecks to necessities organizing, yet that particular sorts of surveys will all things considered evade algorithmic examination. The existing system demonstrated a procedure subject to Aspect and Sentiment Unification Model (ASUM) to expel basic themes from applications' examinations and present clients' evaluations about those core interests. The use of SentiStrength to help a relative errand.

Naive Bayes classier to allot each sentence in clients' surveys to a "sentiment class" among negative, goal, and positive. This is one of the highlights they use to mastermind surveys subject to the data they bring (e.g., join intrigue, issue introduction, and soon.).

At long last, Garcia et al. The relationship between the assumptions and the advancement of supporters in the Open Source Software project GENTOO is broke down..They found that sponsors for example, Stack Over flow is likewise utilized by inspectors to suggest remarks on quality, for further improvement for source code or to see perilous API configuration highlights.

## III. SENTIMENT ANALYSIS TOOLS

At the time of framing, three SE-express sentiment analysis tools are uninhibitedly open, unequivocally Senti4SD, SentiStrengthSE, and SentiCR. We unite them to address RQ1 and RQ2 furthermore, give point by point data on their structure, execution. Our standard is tended to by SentiStrength, The most thoroughly utilized instrument in tendency examination considers in software engineering.

### A. SentiStrength(measure)

SentiStrength is a best level, vocabulary based classifier, which can administer short accommodating substance, including curtailed structures, intensifiers, and emoji's. In the SentiStrength vocabulary, each negative word gets an end score extending from -2 to -5, which tends to it's from as far as possible (i.e., the farthest purpose of the term out of its huge use). Positive and negative emoji's are additionally solidified into the word reference. Grounded on the supposition that a sentence can pass on blended end, SentiStrength yields both positive and negative thought scores for an information content. Given an information

sentence, SentiStrength picks its general positive and negative scores by picking the best positive and negative scores of the terms in the sentence. Intensifiers, i.e., complaint etchings or action words, for example, 'truly', are treated as supporter words and In context on their arithmetical whole, SentiStrength can report the general trinary score, for example the general positive (score = 1), negative (score = - 1) or reasonable (score = 0). With the examination, we are beginning at now reflecting, we map SentiStrengthscores utilizing this trinaryapproach.

### B. Senti4SD

Senti4SD is a point of confinement classifier unequivocally organized to help doubt examination in designers' channels. Senti4SD is uninhibitedly open for research purposes and is hovered together with. Overflow and physically commented on for sentiment uttermost point. What's more, the instrument compartment gives a plan procedure that empowers the preparation of custom classifiers starting with no outside help. Concerning the execution obtained by SentiStrength on an equivalent Stack Overflow dataset, Senti4SD diminishes the misclassifications of fair and positive posts are exactly negative.Senti4SD accomplishes an indistinguishable exhibit moreover in closeness of a superfluous set of preparing reports.

### C. SentiStrengthSE

SentiStrengthSE is a SE-unequivocal thought examination instrument, in perspective on the SentiStrengthAPI. It use a physically balanced modification of the SentiStrength word reference and finishes without any preparation heuristics to address the changed according to mirror the semantics and neutral polarity of room words, for example, 'support', 'goof', or 'default'. The evaluation performed by the creators demonstrates that SentiStrengthSE beats SentiStrength on explicit organizations.

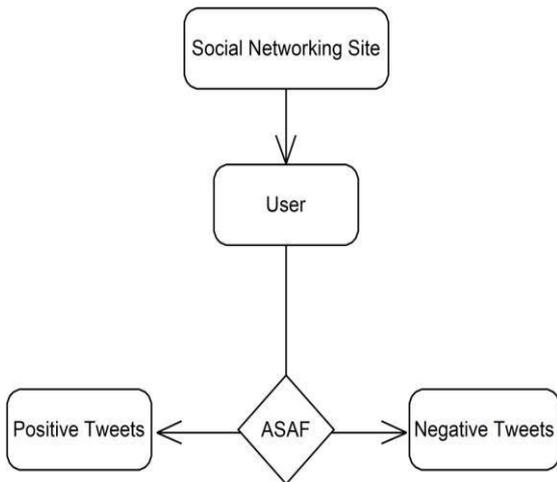
### D. SentiCR

SentiCR is an overseen sentiment analysis apparatus compartment, explicitly orchestrated and assessed for code audit remarks. SentiCR use a part vector conveyed by enrolling TF-IDF (Term Frequency - Inverse Document Frequency) for sack of-words. Before making the part vector, SentiCR. Likewise, it performs SMOTE (fabricated minority over-testing technique) to address the issue of class unevenness in the arranging information. SentiCR has been studied utilizing eight coordinated estimations in a 10-overlay cross support setting. The at present scattered change of SentiCR understands a course of action approach subject to Gradient Boosting Tree.

## IV. RESULTS

We use the proposed system focus on classification of tweets based on the ASAFs. To develop this programming language is JAVA and NETBEANS 8.0.2 IDE is used to implement and results shows the performance of the proposed system.





**Fig. 2** Flowchart for our proposed system

The flow chart explains about the integrated with social networking system to analyze the tweets given by the users. Various users are registered and are saved into the database. Every user gives there tweets and also retweets to be analyze this tweets by using ASAF.

**The ESASML algorithm as follows:**

- Step: 1 Gather all the tweets from the users.
- Step: 2 Pre-processing the tweets. (removing the noise words).
- Step: 3 Normalization of tweets-this will check the English within the given tweets.
- Step: 4 apply the ESASML algorithm.
- Step: 5 results
- Step: 6 Stop.

**A. Procedure**

Here we are creating an account by username and password.

By using their username and password users will register in that account.

Users will give different tweets.

By using ASAF the tweets are classified weather they are negative, positive or stressed.

From the below flow chart:4.1,we have given a some tweet name by using ASAF we get description and tweet date about the tweet which we game and also the temple, we also get the name of the user.

From the table 1, 2 and table 3, by using ASAF we will find weather the given tweet is positive or not.In this we will take the two user tweets and we classify weather they are positive or negative and also we can also get the time and date. From the table 4 by using ASAF we will find the given tweets are stressed tweets or not, in this also we will get the time and date.

**Table. 1** This is the tweet given by the user Rakesh.

Tweet Name	Temple
Description	I am feeling very happy
Tweet Date	05/08/2019 10:59:38

**Table. 2** These are the positive tweets after the classification done by the ASAF.

Positive Tweets Of: Temple[Tweet Owner:Rakesh]			
Commented User	Comment	Date	View the Details
Omkar	So Nice	05/08/2019 13:07:42	View
Ram	Nice One	05/08/2019 16:38:36	View

**Table. 3** These are the negative tweets after the classification done by the ASAF.

Negative Tweets Of: Mountain[Tweet Owner:Rakesh]			
Commented User	Comment	Date	View The Details
Omkar	So Bad	05/08/2019 13:08:32	View
Omkar	Very Bad Day	08/08/2019 15:52:09	View

**Table. 4** These are the stressed tweets after the classification done by the ASAF.

Stressed Tweets of: Temple[Tweet Owner:Rakesh]			
Commented User	Comment	Date	View The Details
Arjun	So Stressed	05/08/2019 14:20:03	View

**V. CONCLUSION**

The proposed system in the paper is integrated with machine learning algorithm. We observed that machine learning algorithm gives best result than the semantic aims. By using sentiment analysis algorithm we classify the tweets weather there are positive or negative or stressed tweets.

**REFERENCES**

- Pang B, Lee L (2008) Opinion mining and sentiment anal- ysis. Found Trends InfRetr 2(1-2):1–135, DOI 10.1561/15000000011.
- Rahman MM, Roy CK, Keivanloo I (2015) Recommending insightful comments for source code using crowdsourcedknowledge. In: 15th IEEE International Working Conference on Source Code Analysis and Manipulation, SCAM 2015, Bremen, Germany, September 27-28, 2015, pp 81–90, DOI 10.1109/SCAM.2015.7335404.
- MaalejW, Kurtanovic Z, Nabil H, Stanik C (2016) Onthe automatic classification of app reviews. Requirements Engineering 21(3):311–331, DOI 10.1007/s00766-016-0251-9.
- Guzman E, BrueggeB (2013) towards emotional awareness in software development teams. In: Proceedings of the 2013 9thJoint Meeting on Foundations of Software Engineering, ACM, New York, NY, USA, ESEC/FSE 2013, pp 671–674, DOI 10.1145/2491411.2494578.
- Guzman E, AzocarD, Li Y (2014) Sentiment analysis of commit comments in Github: An empirical study. In: Proceedings of the 11th Working Conference on Mining Software Repositories, ACM, New York, NY, USA, MSR 2014, pp 352–355, DOI10.1145/2597073.2597118.



6. Calefato F, LanubileF, MarasciuloMC, Novielli N (2015) Mining successful answers in stack overflow. In: Proceedings of the 12th Working Conference on Mining Software Repositories, IEEE Press, Piscataway,NJ,USA, MSR '15, and pp 430–433.
7. Wittgenstein L (1965) Philosophical Investigations. The Macmillan Company, New York, NY, USA.
8. VinayakSinha, Alina Lazar, and Bonita Sharif. 2016. Analyzing developer sentiment in commit logs. In Proceedings of MSR 2016 (13th International Conference on Mining Software Repositories). ACM, 520–523.
9. Rodrigo Souza and Bruno Silva. 2017. Sentiment analysis of Travis CI builds. In Proceedings of MSR 2017 (14th International Conference on Mining Software Repositories). IEEE Press,459–462.
10. L. V. G. Carreño and K. Winbladh. 2013. Analysis of user comments: an approach for software requirements evolution. In Proceedings of ICSE 2013 (35th International Conference on Software Engineering). IEEE press,582–591.
11. Michael Goul, OliveraMarjanovic, Susan Baxley, and Karen Vizecky. 2012. Managing the Enterprise Business Intelligence App Store: Sentiment Analysis Supported Requirements Engineering. In Proceedings of HICSS 2012 (45th Hawaii International Conference on System Sciences).4168–4177.

## AUTHORS PROFILE



**Ms.S.SaiManisha** studying computer science engineering in KL University, vaddeswaram, Guntur district, Vijayawada.



**Ms.P.SaiAnusha**, studying computer science engineering in KL University, vaddeswaram, Guntur District, Vijayawada.



**Ms.M.NagaDarshini** studying computer science engineering in KL University, vaddeswaram, Guntur District, Vijayawada.



**Ms.B.Jahnavi** studying computer science engineering in KL University, vaddeswaram, Guntur District, Vijayawada.