

Sentiment Analysis on Government Scheme Tweets using LSTM

Mrs. Jayalakshmi.V, Research Scholar, Department of Computer Science, Sathyabama Institute of Science and Technology, Chennai

Dr. M. Lakshmi, Professor, Department of Computer Science and Engineering, Savitha School of Engineering, SIMATS, Chennai.

Mr. S.V. Mahadevan, Senior Process Associate, TCS, Siruseri, Chennai.

Article Info

Volume 82

Page Number: 4939 - 4945

Publication Issue:

January-February 2020

Abstract:

The objective of this work is to study the various artificial intelligence methods used for optimizing machining parameters while machining of hybrid metal matrix composites, by conventional and unconventional methods. A hybrid composite is formed when two or more reinforcements are added to the matrix. Hybrid Composites are manufactured by stir casting method. Machining of composites is done to create holes, slots and other features that are not possible to obtain during manufacturing of the part. Various types of machining operations are done in hybrid composites with lathe, drilling, milling and EDM machine to get the desired surface roughness, tolerance. Cutting speed, feed rate and depth of cut are the machining parameters optimized while machining in lathe with desired target like surface roughness, MRR, cutting force and tool wear. Drilling of fiber-reinforced plastics (FRP's) composites facilitates assembly of several components by means of mechanical fastening. Spindle speed, feed rate, drill type are optimized for performance characteristics thrust force, surface roughness, and tool wear in drilling operation. Spindle speed, feed rate, depth of cut is optimized for desired cutting force and surface roughness in milling operation. The influence of process parameters such as pulse on time, pulse off time, spark gap voltage, peak current, wire tension and wire feed rate on response variables such as cutting speed, surface roughness and spark gap are studied in EDM of hybrid composites. ANOVA, RSM, GRA, Taguchi method are used for optimization of various machining parameters. The AI techniques used for prediction include artificial neural network (ANNs), fuzzy logic (FL), adaptive neuro-fuzzy systems (ANFIS), decision tree, genetic algorithm (GA) and genetic programming. Fuzzy rules relate the relationship between input and output variables. Expert knowledge can be built into the system through the rule base. Fuzzy models are used for predicting the thrust force and torque for drilling hybrid composite. In adaptive neuro-fuzzy systems advantages of FL and ANNs are combined for adjusting the membership functions, rule base and related parameters for training the data set. It can continuously improve the initially obtained rough model based on the daily operating data. AI techniques are used for predicting the automatic selection of inputs and predicting surface roughness, other response variables.

Keywords: Artificial Intelligence, Hybrid Metal Matrix Composites, Machining Parameters, Optimization.

Article History

Article Received: 18 May 2019

Revised: 14 July 2019

Accepted: 22 December 2019

Publication: 23 January 2020

I. INTRODUCTION

Internet based life have turned into a developing wonder due to the colossal and quick advances in data innovation. Individuals are utilizing web based life on consistent schedule to discuss their feelings

with one another about wide assortment of subjects, items and administrations, which has made it a rich asset for content mining and estimation investigation. Internet based life correspondences incorporate Facebook, twitter, and numerous others. Twitter is a standout amongst the most broadly

utilized online networking destinations. The quantity of twitter messages sent every second around the world. There is no standard strategy for mining and examining online networking business information. To dissect the tweets python has been utilized. A contextual investigation of Eight Indian Government Schemes is introduced to demonstrate the significance of dissecting client produced online suppositions from Twitter. This is useful for assessment government execution checking from Peoples point of view as opposed to making Peoples studies which are costly and tedious.

Assumption examination has been first presented by Liu, B. It is otherwise called assessment mining and subjectivity examination is the procedure to decide the demeanor or extremity of feelings given by people to specific plan. Notion examination can be connected on any printed type of sentiments, for example, websites, surveys and Microblogs. Microblogs are those little instant messages, for example, tweets, a short message that can't surpass 160 characters.

These microblogs are less demanding than different types of feelings for opinion. Sentiment investigation should be possible on an archive level or a sentence level. In the main case, the entire archive is assessed to decide the conclusion extremity, where, the highlights depicting the item/administration ought to be extricated first. Though, the second one, the report is isolated into sentences everyone is assessed independently to decide the conclusion extremity.

1.1 Levels of Sentiment Analysis

a. Record level Sentiment Analysis In this Sentiment Analysis level entire report has examined and order whether an entire assessment archive communicates a positive or negative feeling [1], [2]. In one report just surveys of one item has been assessed. Furthermore, undertaking is to discover the conclusion about that item. So this errand is extensively known as archive level assumption characterization. In this dimension, communicated feeling is on single substance. This isn't material

when there is archive which contains various item audits.

b. Sentence Level Sentiment Analysis In this dimension, undertaking goes to each sentence and decide if the sentence communicates the positive, negative or unbiased feeling. This dimension mindfully identified with Subjectivity Classification [3], which recognizes target Sentences and abstract sentences. Targets sentences express truthful data about sentences where Estimation Analysis of Indian Government Schemes utilizing Twitter Datasets Subjective sentences express the abstract data about sentences. Numerous target sentences can include Opinions. This assignment is known as Sentence Level Sentiment Analysis. In the last case, the report is isolated into sentences and everyone is assessed independently to decide the feelings. To perform assumption investigation, preprocessing is a fundamental errand and preprocessing systems upgrades the exactness of the estimation order [4].

This work is completed by Python language to mine the tweets and furthermore to actualize the Opinion Mining for Government Scheme. Python is a broadly utilized universally useful, abnormal state programming language. It was at first planned by Guido van Rossum in 1991 and created by Python Software Foundation. It was fundamentally produced for accentuation on code lucidness, and its linguistic structure enables software engineers to express ideas in less lines of code. Python is a programming language that gives you a chance to work rapidly and incorporate frameworks all the more proficiently. Python takes a shot at various stages like Windows, Mac, Linux, Raspberry Pi, and so on.

II. LITERATURE REVIEW

In the present computerized world situation, the viable and proficient mining of the online surveys is generally required. In the online store, there are a huge number of suppositions thus it is troublesome for the average citizens to pursue the equivalent. In this way, online web surveys mining arrangement of

Sentiment Analysis helps everybody in knowing further. Assessment Analysis is an AI approach that distinguishes and classifies the suppositions communicated as content to discover the easiest concept of methodology and his frame of mind towards the particular space as positive or negative or impartial. Estimation Analysis sees the extremity of the content in printed position.

As it is otherwise called Opinion Mining, it infers the conclusion of the speaker about a subject. From the paper Twitter Sentiment Analysis [8] by Efthymios Kouloumpis, Theresa Wilson, Johanna Moore Sentiment investigation is a developing zone of Natural Language Processing with research extending from record level classification to learning the extremity of words and expressions. Given the character confinements on tweets, characterizing the conclusion of Twitter messages is most like sentence-level assumption examination. From the white paper Social Networks Overview: Current Trends and Research Challenges distributed [4] by European Commission, we figured out how to examine the momentum patterns of the informal organization examination and the difficulties that were available while doing the investigation of interpersonal organization investigation. Feeling examination is an exploration and building up a surge of Natural Language Processing strategy for the man-made consciousness [11].

One can without much of a stretch get a point by point and graphic knowledge to foresee and examine the feelings of voters [7]. It is conceivable to follow the constant patterns and furthermore to catch any unexpected supernatural change by continue observing the discussions on the online media and get the assortment of popular sentiment well ahead of time even before the declaration and the presentation of consequences of surveys [8]. . It determines whether a bit of composing is sure, negative or impartial [5]. People's political inclinations communicated online with those distinctly seen over races with employments of Social Networking Sites [6]. A portion of the investigations express that breaking down web based

life permits a dependable and productive estimate of the last outcome [9]. Since Twitter pursues smaller scale blogging nature, it is the best hotspot for conclusion investigation. Real research work for notion examination is done utilizing tweets, yet there has been some investigation including the Amazon surveys dataset as well. The new execution for dissecting amazon surveys which involve detection of phony audits, handling the certified surveys utilizing Apache Spark lastly evaluating the items [13]. The assessment of short content is amassed by the diverse impact of connections between the modifier and the feeling word and the commitment of each sentence to the estimation computation of short message [14]. Prototype system is built to demonstrate its efficacy for automatic color palette design [15]. A sentiment-based prediction model for the financial domain which uses the combination of a clausal/phrasal sentence simplification step [16].

III. METHODOLOGY

This work has been developed by using seven phases. 1. Collection of tweets from the people, 2. Tweets Extraction, 3. Preprocessing techniques, 4. Sentiment feature selection, 5. LSTM – Long Short Term Memory, 6. Polarity identification and classification, 7. Visualizing result. Figure 1 shows the proposed methodology of this work

A. Collection of peoples tweets:

Tweets are essential for doing the Sentiment Analysis Task. Here we collected tweets from twitter API using API.get_user(), API.user_timeline() like functions.

B. Extraction of tweets:

The Twitter Application can be created with the help of twitter app. When the app is created, the user is given the access to the Consumer Key and Access Token. The Consumer Secret and the Access Secret must be confidential. The package used for authentication is tweepy.OAuthHandler(). The function used to search and extract tweets. Totally 3500 #Make in India tweets has been extracted and saved.

C. Preprocessing techniques:

Pre-Processing Data pre-handling is done to expel the deficient uproarious and conflicting information. Information must be preprocessed before utilizing in highlight choice assignment. In pre-handling following are a few assignments:

- Removing URLs, Special characters, Numbers, Punctuations and so on.
- Removing Stopwords
- Removal of Retweets (if there should be an occurrence of twitter dataset)
- Stemming
- Tokenization.

D. Sentiment feature selection

Highlight choice from pre-handled content is the troublesome assignment in feeling examination. The primary objective of the component determination is to diminish the dimensionality of the element space and along these lines computational expense. Highlight determination will lessen the overfitting of the learning plan to the preparation information. Highlight determination will lessen the overfitting of the learning plan to the preparation information. Assessment words can be characterized into positive, negative and nonpartisan words. The pre-handled tweets are the contribution to the capacities to characterize the feelings and the extremity. This should be possible with the assistance of the bundle opinion.

E. LSTM:

Long short-term memory (LSTM) is an artificial architecture used in deep learning concept. Unlike standard feed forward neural networks. LSTM Networks are appropriate to arranging, handling and making expectations dependent on time arrangement information, since there can be lags of unknown duration between important events in a time series.

F. Polarity identification and classification

The initial work in Sentiment Analysis is classify the polarity of a given tweets feature. The polarity is in three category i.e. Positive, Negative and Neutral. With the help of Bing Lui sentiment lexicon, SentiWordNet etc., lexicons used for polarity identification helps to evaluate and calculate

sentiment score, sentiment strength etc. The function classify polarity is used to classify the tweets. After being classified for each scheme, the polarity is visualized in the form of bar graphs. The polarity is classified as either positive, negative or neutral.

G. Visualizing result:

The polarities are visualized with the help of the TextBlob library in Python and matplotlib Python plot library installed to generate the graph and the word cloud with the help of WordCloud() function.

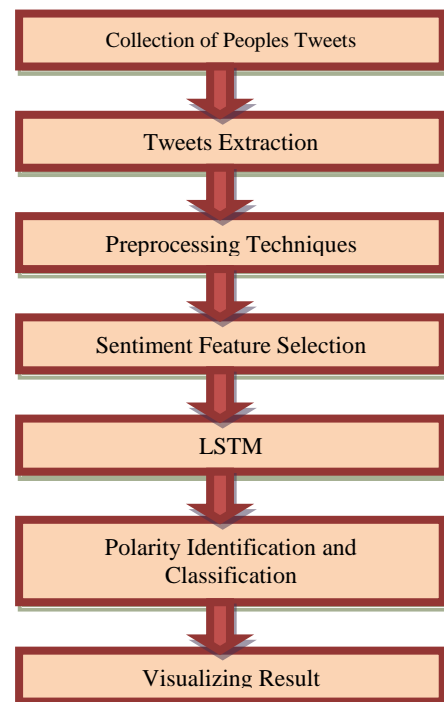


Figure 1: Proposed Methodology

IV. RESULTS AND DISCUSSION

This work has been developed using twitter dataset. Twitter is an online social network used to send and read short messages called "tweets". The tweets used to analyze and predict the future directions by public opinion. The following figures shows the result of polarity classification, which has classified the tweets into positive, negative and neutral. For the proposed work taken Make In India, Budget 2019, GST India, JammuKashmir, Swatch Bharath tweets of government scheme dataset has taken for polarity classification. Finally shown the word cloud for those government scheme tweets.

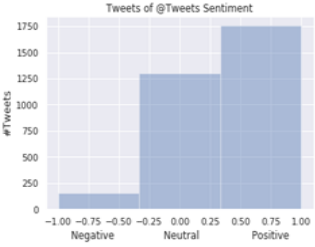
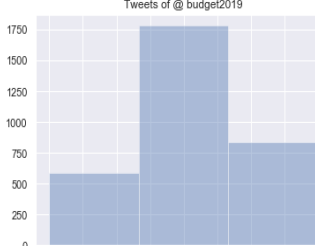
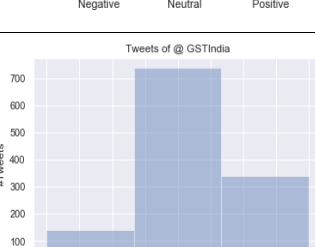
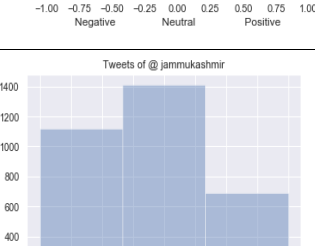
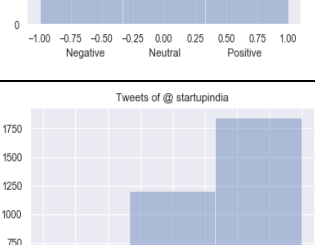
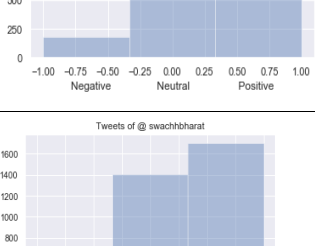
Government Schemes	Polarity Classification																				
Make In India	 <table border="1"> <caption>Tweets of @Tweets Sentiment</caption> <thead> <tr> <th>Sentiment</th> <th>#Tweets</th> </tr> </thead> <tbody> <tr><td>-1.00</td><td>100</td></tr> <tr><td>-0.75</td><td>150</td></tr> <tr><td>-0.50</td><td>150</td></tr> <tr><td>-0.25</td><td>1300</td></tr> <tr><td>0.00</td><td>1300</td></tr> <tr><td>0.25</td><td>1300</td></tr> <tr><td>0.50</td><td>1650</td></tr> <tr><td>0.75</td><td>1650</td></tr> <tr><td>1.00</td><td>1650</td></tr> </tbody> </table>	Sentiment	#Tweets	-1.00	100	-0.75	150	-0.50	150	-0.25	1300	0.00	1300	0.25	1300	0.50	1650	0.75	1650	1.00	1650
Sentiment	#Tweets																				
-1.00	100																				
-0.75	150																				
-0.50	150																				
-0.25	1300																				
0.00	1300																				
0.25	1300																				
0.50	1650																				
0.75	1650																				
1.00	1650																				
Budget 2019	 <table border="1"> <caption>Tweets of @ budget2019</caption> <thead> <tr> <th>Sentiment</th> <th>#Tweets</th> </tr> </thead> <tbody> <tr><td>-1.00</td><td>600</td></tr> <tr><td>-0.75</td><td>600</td></tr> <tr><td>-0.50</td><td>600</td></tr> <tr><td>-0.25</td><td>1750</td></tr> <tr><td>0.00</td><td>1750</td></tr> <tr><td>0.25</td><td>1750</td></tr> <tr><td>0.50</td><td>800</td></tr> <tr><td>0.75</td><td>800</td></tr> <tr><td>1.00</td><td>800</td></tr> </tbody> </table>	Sentiment	#Tweets	-1.00	600	-0.75	600	-0.50	600	-0.25	1750	0.00	1750	0.25	1750	0.50	800	0.75	800	1.00	800
Sentiment	#Tweets																				
-1.00	600																				
-0.75	600																				
-0.50	600																				
-0.25	1750																				
0.00	1750																				
0.25	1750																				
0.50	800																				
0.75	800																				
1.00	800																				
GST India	 <table border="1"> <caption>Tweets of @ GSTIndia</caption> <thead> <tr> <th>Sentiment</th> <th>#Tweets</th> </tr> </thead> <tbody> <tr><td>-1.00</td><td>150</td></tr> <tr><td>-0.75</td><td>150</td></tr> <tr><td>-0.50</td><td>150</td></tr> <tr><td>-0.25</td><td>700</td></tr> <tr><td>0.00</td><td>700</td></tr> <tr><td>0.25</td><td>700</td></tr> <tr><td>0.50</td><td>350</td></tr> <tr><td>0.75</td><td>350</td></tr> <tr><td>1.00</td><td>350</td></tr> </tbody> </table>	Sentiment	#Tweets	-1.00	150	-0.75	150	-0.50	150	-0.25	700	0.00	700	0.25	700	0.50	350	0.75	350	1.00	350
Sentiment	#Tweets																				
-1.00	150																				
-0.75	150																				
-0.50	150																				
-0.25	700																				
0.00	700																				
0.25	700																				
0.50	350																				
0.75	350																				
1.00	350																				
Jammu Kashmir	 <table border="1"> <caption>Tweets of @ jammukashmir</caption> <thead> <tr> <th>Sentiment</th> <th>#Tweets</th> </tr> </thead> <tbody> <tr><td>-1.00</td><td>1100</td></tr> <tr><td>-0.75</td><td>1100</td></tr> <tr><td>-0.50</td><td>1100</td></tr> <tr><td>-0.25</td><td>1400</td></tr> <tr><td>0.00</td><td>1400</td></tr> <tr><td>0.25</td><td>1400</td></tr> <tr><td>0.50</td><td>700</td></tr> <tr><td>0.75</td><td>700</td></tr> <tr><td>1.00</td><td>700</td></tr> </tbody> </table>	Sentiment	#Tweets	-1.00	1100	-0.75	1100	-0.50	1100	-0.25	1400	0.00	1400	0.25	1400	0.50	700	0.75	700	1.00	700
Sentiment	#Tweets																				
-1.00	1100																				
-0.75	1100																				
-0.50	1100																				
-0.25	1400																				
0.00	1400																				
0.25	1400																				
0.50	700																				
0.75	700																				
1.00	700																				
Startup India	 <table border="1"> <caption>Tweets of @ startupindia</caption> <thead> <tr> <th>Sentiment</th> <th>#Tweets</th> </tr> </thead> <tbody> <tr><td>-1.00</td><td>200</td></tr> <tr><td>-0.75</td><td>200</td></tr> <tr><td>-0.50</td><td>200</td></tr> <tr><td>-0.25</td><td>1200</td></tr> <tr><td>0.00</td><td>1200</td></tr> <tr><td>0.25</td><td>1200</td></tr> <tr><td>0.50</td><td>1750</td></tr> <tr><td>0.75</td><td>1750</td></tr> <tr><td>1.00</td><td>1750</td></tr> </tbody> </table>	Sentiment	#Tweets	-1.00	200	-0.75	200	-0.50	200	-0.25	1200	0.00	1200	0.25	1200	0.50	1750	0.75	1750	1.00	1750
Sentiment	#Tweets																				
-1.00	200																				
-0.75	200																				
-0.50	200																				
-0.25	1200																				
0.00	1200																				
0.25	1200																				
0.50	1750																				
0.75	1750																				
1.00	1750																				
Swachh Bharath	 <table border="1"> <caption>Tweets of @ swachhbharat</caption> <thead> <tr> <th>Sentiment</th> <th>#Tweets</th> </tr> </thead> <tbody> <tr><td>-1.00</td><td>100</td></tr> <tr><td>-0.75</td><td>100</td></tr> <tr><td>-0.50</td><td>100</td></tr> <tr><td>-0.25</td><td>1400</td></tr> <tr><td>0.00</td><td>1400</td></tr> <tr><td>0.25</td><td>1400</td></tr> <tr><td>0.50</td><td>1600</td></tr> <tr><td>0.75</td><td>1600</td></tr> <tr><td>1.00</td><td>1600</td></tr> </tbody> </table>	Sentiment	#Tweets	-1.00	100	-0.75	100	-0.50	100	-0.25	1400	0.00	1400	0.25	1400	0.50	1600	0.75	1600	1.00	1600
Sentiment	#Tweets																				
-1.00	100																				
-0.75	100																				
-0.50	100																				
-0.25	1400																				
0.00	1400																				
0.25	1400																				
0.50	1600																				
0.75	1600																				
1.00	1600																				

Table 1: Polarity Classification of various government scheme tweets

[illegible]

Table 2: Word Cloud of various government schemes

Trend name	Positive Sentiments	Negative Sentiments	Neutral Sentiments	Tweets used
MakeIn India	1750	200	1300	3250
Budget 2019	841	584	1783	3208
GST India	338	141	737	1216

Jammu Kashmir	692	1118	1411	3221
Start up India	1841	182	1202	3225
Swatch Bharath	1710	123	1400	3233

Table 3: Government Scheme tweets with Sentiment score

As the tweets were classified with the polarity based on Subjectivity Lexicon, the tweets which have more number of positive words are classified as positive and with more negative words are negative and the tweets which neither falls under positive nor negative are classified as neutral tweets. After being classified separately and found the tweets has more positive and neutral than negative in their sentiment score. Table 3 represents the result of sentiment score.

V. CONCLUSION

Sentiment Analysis is one of the effective and efficient techniques for analyzing the behavior of the people. The twitter data, commonly known as tweets are the samples for the different opinions of different people. Sentiment Analysis helps to find the different perspective and aspects of the people's mind and thoughts with which the polarity of the schemes can be found and it helps in predicting the implementation of the various Government schemes effectively and also to take decisions for the future schemes.

In this paper, Polarity Classification is performed such as obtaining positive, negative and neutral tweets for particular Government Schemes such as Make in India. From this people and Government can find out the different opinion behind the declared scheme. The manual survey from people is tiresome as it is costly and time consuming. So, this Sentiment Analysis is very handy and effective for the evaluation of the Government scheme and also to monitor them on a regular basis of its growth among the people and also to know the people's perspective.

REFERENCES:

- [1] Rajkumar S. Jagdale, Vishal S. Shirsat, Sachin N. Deshmukh," Sentiment Analysis of Events from Twitter Using Open Source Tool International Journal of Computer Science and Mobile Computing ISSN 2320-088X IMPACT FACTOR: 5.258 IJCSMC, Vol. 5, Issue. 4, April 2016, pg.475 – 485
- [2] Xing Fang and Justin Zhan."Sentiment analysis using product review data". Journal of Big Data 2015. DOI: 10.1186/s40537-015-0015-2
- [3] Seyed-Ali Bahrainian, Andreas Dengel, "Sentiment Analysis and Summarization of Twitter Data", CSE, 2013, 2013 IEEE 16th International Conference on Computational Science and Engineering (CSE), 2013 IEEE 16th International Conference on Computational Science and Engineering (CSE) 2013, pp. 227-234, doi:10.1109/CSE.2013.44
- [4] M.Bhuvaneswari, V.Srividhya, "Enhancing The Sentiment Classification Accuracy of Twitter Data Using Preprocessing Technique"International Journal of Engineering Research, Vol. 4 , Issue 5, 2016
- [5] EfthymiosKouloumpis, Theresa Wilson, Johanna Moore "Twitter Sentiment Analysis: The Good the Badandthe OMG!" Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media.
- [6] An Approach to Harvesting, Cleaning, and Analyzing Data from Twitter Using R , Stephen Hill hills@uncw.edu , Rebecca Scott scotra@uncw.edu ,2016 Proceedings of the EDSIG Conference ISSN: 2473-3857 Las Vegas, Nevada USA
- [7] H.Yu. V.Hatzivassiloglou, "Towards answering opinion questions: separating facts from opinions and identifying the polarity of opinion sentences" In the Proceedings of the Conference on Empirical Methods in Natural Language Processing,pp. 129-136, 2003
- [8] Xin J, Gallagher A, Cao L."The wisdom of social multimedia: Using flickr for prediction and forecast",Proceedings of the 18th ACM international conference on Multimedia, New York, pp.1235-1244, 2010.
- [9] MJ.Jensen, L.Jorba E. Anduiza "Digitalmedia and political engagement worldwide: A comparative

- study”, Cambridge University Press, New York, NY, pp1-15, 2012
- [10] BholaneSavitaDattu, Prof.Deipali V. Gore “A Survey on Sentiment Analysis on Twitter Data Using DifferentTechniques”, International Journal of Computer Science and Information Technologies, Vol. 6, Issue 6, pp. 5358-5362, 2015.
- [11] O’Connor B, Balasubramanyan R, Routledge BR, Smith NA. “From tweets to polls: Linking text sentiment to public opinion time series” Proceedings of the fourth International AAAI Conference on Weblogs and Social Media, Washington, pp. 122-129, 2010.
- [12] Tjong Kim Sang E, J.Bos, “Predicting the 2011 Dutch senate election results with twitter”, Proceedings of SASN 2012, Avignon, France, pp.1-8, 2012.
- [13] S.R.Chheda, A.K.Singh, P.S.Singh, A.S.Bhole, “Automated Trading of Cryptocurrency Using Twitter Sentimental Analysis”, International Journal of Computer Sciences and Engineering, Vol 6, Issue 5, pp.209-214.
- [14] JieLi ;LirongQiu , “A Sentiment Analysis Method of Short Texts in Microblog”, IEEE 2017, ISBN: 17098775.
- [15] Ya-Fang Peng ;Tzren-Ru Chou , Automatic Color Palette Design Using Color Image and Sentiment Analysis, 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA).
- [16] Macedo Maia ; Andr Freitas ; Siegfried Handschuh , FinSSLx: A Sentiment Analysis Model for the Financial Domain Using Text Simplification, 2018 IEEE 12th International Conference on Semantic Computing (ICSC), Electronic ISBN: 978-1-5386-4408-9 Print on Demand(PoD) ISBN: 978-1-5386-4409-6