RESEARCH ARTICLE

SAMRIDDHI Volume 14, Issue 4, 2022 Print ISSN: 2229-7111 Online ISSN: 2454-5767

Developing a Hybrid model with Shades of Sentiment for Understanding Teenagers' Academic Distraction Problems

Monali K. Patil[1*], Nandini Chaudhari[2], B.V. Pawar[1], Ram Bhavsar[1*]

[1]School of Computer Sciences, Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon, Maharashtra, India

[2]Babaria Institute of Technology, Vadodra, Gujrat, India

Abstract

In recent years, many techniques have come up in the field of sentiment analysis. In the field of the medical domain, sentiment analysis has been used for the areas like information retrieval, feedback analysis, dialogue conversation, and review analysis. Psychological analysis through the window of sentiment analysis is still unfolded area. We have designed and developed a hybrid computational model that maps teenagers' sentiments to their behavioral patterns and academic distraction problems. We have also developed a sentiment shaded lexicon which defines the ontology for various shades of sentiment with the help of Bling Liu's positive and negative word dictionary. We have used a semantic lexicon-based approach and a rule-based classifier in our hybrid computational model. Rules are applied to user-input text, which teenagers write. We have extracted sentiments expressed in the user input text and we have also achieved to identify academic distraction problems of teenagers. We have computed the performance metrics of the model on 155 samples, randomly collected from teenagers (age group 13-19 years). We have computed accuracy at two stages, first at the sentiment extraction stage and second at the final output level. The error analysis is also presented in this paper. After working on error analysis, our model has achieved an accuracy of 90% for sentiment extraction and 87% for derived academic distraction problems.

Keywords: Sentiment analysis, Machine learning, Natural language processing, Opinion mining, Sentiment-shaded lexicon.

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology (2022); DOI: 10.18090/samriddhi.v14i04.15

Introduction

Sentiment Analysis (SA) is a relatively new area, and research is going on under artificial intelligence and natural language processing for SA in recent years. The fast growth of the information in social media over the Internet and the developments in NLP techniques present an opportunity to mine these data in different domains, such as tourism, marketing, or the political environment.[1] People participate in the information spreading over the world wide web environment with extensive social media usage.[2] In the medical domain sentiment analysis is used for summarization, feedback or review analysis, and automation of appointment scheduling. In clinical psychology, many tests are available to understand patients' psychological problems.[3,4] We want to work over students' academic distraction problems. Many times, students or their parents are unaware of different psychological problems faced by students. We want to explore these academic psycho-medical problems faced by teenagers through sentiment analysis. Because of its high social relevance, we have considered this problem a challenge.

Corresponding Author: Monali K. Patil, School of Computer Sciences, Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon, Maharashtra, India, E-mail: monalismart1981@gmail.com

How to cite this article: Patil MK, Chaudhari N, Pawar BV, Bhavsar R. (2022). Developing a Hybrid model with Shades of Sentiment for Understanding Teenagers' Academic Distraction Problems. *SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology,* 14(4), 97-104.

Source of support: Nil

Conflict of interest: None

Sentiment analysis and opinion mining are similar terms. There is a marginal difference between sentiment analysis and opinion mining.[5] In sentiment analysis, we are searching for the sentiments expressed about the object by the writer; for example, consider the sentence, "I love my country India", where the writer's positive sentiment is conveyed about an object- India. In the opinion mining opinion of the writer about an object is considered; for example, consider a sentence, "India can become a developed county if there is a

Table 1: Comparative analysis of related work

Year of publication	*Title*	*Problem description*	*Computational approach (es)*	*Remark*
2011	From once upon a time to happily ever after: tracking emotions in novels and fairy tales. [6]	A system is developed which compares Fairy tales with novels based on emotional word density.	Lexicon-based sentiment analysis	Using Google books corpus, Mohammad showed how to evaluate an entity's emotional relations from co-occurring words. They made the comparison between words in fairy tales and novels to show which one was denser in emotional intensity.
2014	Sentiment analysis for opinion mining using a Cross-domain classifier. [7]	A cross-domain approach is developed to tackle the feature mismatch problem.	Unigram, bigram, feature vector	Pravin created a glossary with the help of the point-wise mutual information technique (PMI) and distributional hypothesis sentiment. A feature mismatch problem is solved by the creation of a glossary
2019	How to predict explicit recommendations in online reviews using text mining and sentiment analysis. [8]	It gives a theoretical contribution to how the negative attitude and positive feelings trigger the recommendation system.	binary logistic regression	Joao's research contributes to the literature by showing the most important predictors of an explicit recommendation in the review versus non-recommendation advice.
2020	Sentiment analysis for E-commerce product reviews in Chinese Based on Sentiment lexicon and deep learning [9]	A weighted sentiment-feature classification system is developed.	CNN and Gated Recurrent unit network	The CNN and GRU networks extracted important sentiment features and context features along with the usage of the attention technique for weights in reviews.
2020	Applying sentiment analysis to automatically classify consumer comments concerning marketing 4Cs aspects.[10]	A system is developed for aspect-based classification.	LDA topic analysis	Lin and the team collected popular comment topics of customers and used the keywords to classify the corpus with the help of lexicon comparison.
2020	SACPC: A framework based on probabilistic linguistic terms for short text sentiment analysis.[11]	Word2PLTS (probabilistic linguistic terms sets) model is developed for short text sentiment analysis. SACPC framework is developed for polarity classification.	PLTS, SVM	Chao developed a system model which considers word fuzziness and uncertainty with the help of probabilistic linguistic terms set (PLTs).

growth in its industrial domain," here the writer is expressing his opinion about India. Much work is done in opinion mining as a movie review, giving a rating to products, polling reviews before elections, and many more. Sentiment analysis can be done with text analysis from Natural Language Processing (NLP) and machine learning (ML) computational techniques. Liu[12] suggested three levels of sentiment analysis – document-based, sentence-based, and aspect-based. Our research targets document-based sentiment analysis. The focus is on teenagers' academic distraction problems based on user-input text written by teenagers. The sentiment Analysis process works as a sentiment classification problem

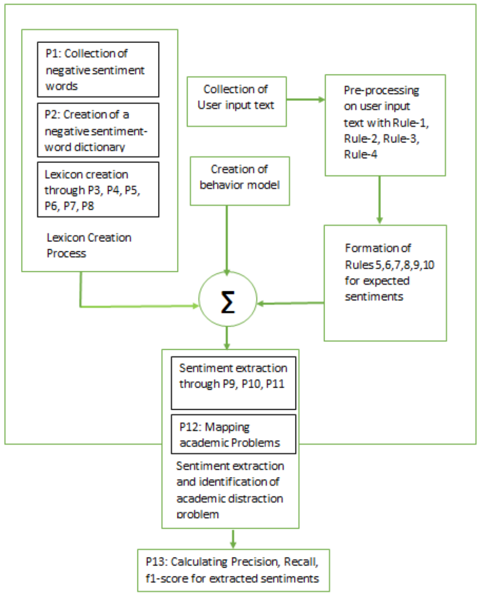

Figure 1: Hybrid Model with Lexicon and Rule-based Approach

for exploring different shades of emotions. Following are the steps in Sentiment analysis –

- A lexicon is constructed (sentiment-shaded) for the teenager's academic distraction problem.
- Identifying Parts of speech (POS)– with the help of text pre-processing.
- Identifying sentiment-related words and phrases from user input text– these are words commonly used to express sentiments, we selected noun, adjective, and verb phrases for our study.[13]
- The score is assigned to each sentiment from the sentiment-shaded-lexicon.
- With the help of a rule-based system, calculate the sentiment scores. [14,15]
- Map the extracted sentiments (from the user input text) to the maximally matching behavior model from the sentiment-behavior-lexicon.

To extract various sentiment shades, we have also prepared an emotional ontology chart for the problem under study, which includes the fundamental six emotions (love, joy, surprise, sadness, anger, and fear).[16] For our problem, we are considering only negative sentiment ontology.

The paper is organized into 6 sections, which describe the general problem definition and its challenges, related work, computational aspects and approaches, system modeling, and imperial results, followed by a conclusion.

Comparative Analysis

During our study, we came across different works related to our problem, which have been reviewed in Table 1.

System Modeling

The important techniques used in our study are lexicon creation, creation of a behavioral model, pre-processing of input text, formation of rules for expected sentiments, sentiments extraction, and identification of academic distraction problems. As the contextual meaning is considered for the words and phrases present in the given user input text, the collection of negative words and a negative sentiment dictionary formation are needed. Figure 1 shows the hybrid model with a lexicon and rule-based approach. P1, P2, and P13 are the processes used in the hybrid lexicon and rule-based model. Through the process P3 to P6, we have created a lexicon for one-gram, two-gram, three-gram, and four-gram models. P7 process creates lexicon for Booster-words. P8 process creates a lexicon for Special-words (Special words occur frequently in the input samples and imply distraction sentiments).

Rule-1 is used to convert tokens into lowercase letters. Rule-2 is used to append blank space before and after a full stop, to identify it as an end delimitator. Rule-3 removes all stop words except some negative words. With the help of Rule-4, we can select verb, noun, and adjective parts of speech (POS) from input text. Rule-5 is used to process the sentence for sentiment extraction if and only if it contains a first-person pronoun (like I, me, my, myself). Rule-6 does not allow the sentence to process if it contains a conditional clause with the word 'sometimes' and 'when'. Rule-7 and Rule-8 are formed to filter empty original sentiments and academic problems, respectively. Rule-9 checks whether a negation-word is occurred with the sentiment word with the bigram and trigram model to skip that sentiment word if the negation word is present.

Rule-10 is designed to consider words from the special-word lexicon if and only if some special-key words occur with it in the sentence. Rule-5, Rule-6, Rule-9, and Rule-10 are added to our model after error analysis. With the help of process P9, we are adding the strength of sentiments, similarly, for the final sentiment score, we have used process P10 to average the sentiment score. Process P11 is used to select the sentiments with more than and equal to a 70% sentiment score.

With the help of Google Forms, we collected input data samples from teenagers. Teenagers' information is collected as a city in which they live, gender, age, paragraphs about their sentiments when they get distracted in their studies. Paragraphs are used as user input text and city, gender, and age can be used as parameters to analyze results.

Sentiments Extraction

From the Bling Liu collection of positive and negative words,[17] a sentiment dictionary is produced, which contains the colloquial meaning of sentiment-related words. This dictionary is used to form the lexicon of negative-words and related sentiments. We have a collection of 3088 words with their sentiment-oriented meaning in the dictionary. The score is assigned to each word that is mapped to particular sentiment. The score specifies the percentage of a word related to specific sentiment. Some examples from the sentiment-shaded-lexicon are shown in Table 2. The

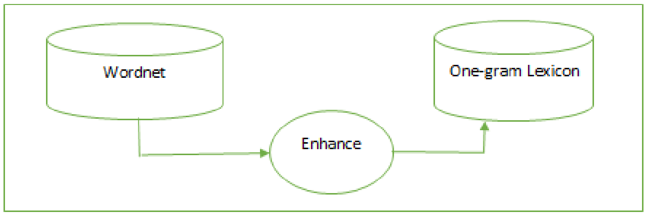

Figure 2: Enhancing One-gram lexicon with synonyms from Wordnet

Table 2: Sample sentiment-shaded-lexicon.

Word	*Sentiment*	*Strength-out-of-5*	*Percentage-Strength*
Guilt	sad	5	100
Steal	awful	4.5	90
Lie	Guilty	4	80
absurd	sad	4	80
accuse	irritable	4	80
askance	confused	3	60
Belie	Guilty	4	80

Table 3: Sample Bigram word Lexicon

Word	*sentiment*	*sentiment strength out of 5*	*percentage sentiment strength*
fed up	frustrated	4	80
not happy	Sad	5	100
back logged	Fearful	3.5	70
screwed up	Hate	4	80
feeling low	Upset	4.5	90
can't study	Weak	4	80
don't want	Crazy	4	80
deep sorrow	Sad	4.5	90

maximum strength assigned is 5 and the minimum is 1. Figure 2 shows that to enrich the lexicon, we have developed the code to add words from Wordnet synonymous with those in the lexicon. Altogether, we have collected 5979 words in the lexicon and all these words are mapped to different negative sentiment shades.

The user input text is taken into csv file. First tokenization is applied with the help of the spacy library of python. The verb phrases, the noun phrases, and adjective phrases are separated with the help of token tags by selecting VB, VBP, VBN, VBD, VBG, as a verb phrases, JJ, JJR, JJS as an adjective phrases, and NN as a noun phrase.

Bigram and Trigram Model

To address the problem of conjugative-words impact, we have to consider the bigram and trigram model. In the bigram model, two conjugative words are tokenized, while three

Table 4: Sample Trigram word Lexicon

Word	*sentiment*	*sentiment strength out of 5*	*percentage sentiment strength*
unable to concentrate	irritable	3.5	70
unable to remember	forgetful	4	80
unable to study	anxious	4	80
down the drain	lost	4	80
lack of concentration	distracted	4	80
lack of motivation	upset	4	80

Table 5: Sample Booster words

Word	*Affect on sentiment strength*
extremely	10
Quite	5
Just	5
Almost	5
Very	10

Table 6: Sample special words

Word	*sentiment*	*sentiment strength out of 5*	*percentage sentiment strength*
mobile	distracted	4.5	90
Youtube	distracted	4.5	90
T.V	distracted	4.5	90
Tv	distracted	4.5	90
Net	distracted	4.5	90
Internet	distracted	4.5	90

conjugative words are tokenized in the trigram model. The sample bigram model is shown in Table 3 and the trigram model is in Table 4. Negation of positive words frequently appearing in the context can be considered with the help of the bigram model.

Booster words and special words

Booster words are the words that affect the strength of the sentiment. It is also called adverbs of degree, e.g., very, extremely, etc, as shown in Table 5. Special words are the noun phrases that appear in the paragraph. These are not sentiment-related words but have an impact on sentiment when used in the paragraph, as shown in Table 6. the political environment [1].

Identify academic distraction problem

Table 7: Sample Sentiment-behavior lexicon.

Sentiment	*Behavior-pattern*	*Academic-problem*
Tired	Lethargic	poor academic performance
Tired	Inactive	poor academic performance
Careless	Inattentive	silly mistakes in writing or oral
Careless	Inattentive	poor handwriting
Careless	Restless	incomplete class work
Fearful	Nervous	lack of practice
frustrated	bad-tempered	poor performance
frustrated	bad-tempered	learning disorder
frustrated	Hyper	Attention-deficit/ hyperactivity disorder (ADHD)
frustrated	Restless	Attention-deficit/ hyperactivity disorder (ADHD)

After the extraction of sentiment from the paragraph, the average sentiment for each sentiment type is calculated. Many distraction problems arise in teenagers. In the current research, we have focussed on the academic distraction problem. It was observed that for each academic distraction problem, there is a particular behavioral pattern and each behavior pattern is related to the sentiments of teenagers. So, the sentiment, behavior pattern, and academic distraction problem mapping are prepared in the sentiment-behavior lexicon. Table 7 shows some examples of this mapping.

Illustration

The following paragraph is given as input to the system:

"My study speed is too slow due to which I feel deep sorrow. I can't study for more than 3 hrs that is also the reason for my depression, the bothers like television sound, communication of my family members and also vehicles sound in front of my house totally distract me from my study. I never feel pressure about my study but it is true that sometimes I feel that I am doing very less study comparatively to my classmates. sometime I feel confuse due to distict heory of same topic given by two different teachers. Sometime I hesitate to ask questions in class because I am thinking so my questions are silliest ones".

The above paragraph is one of the samples from user input text collected through Google form. As it is written by the student, it may contain a typo error. The resultant sentiments, behavior patterns, and academic distraction problems for the above user input text:

the final fine-grained sentiments are calculated by the model:

{'Confused', 'irritable', 'Depressed', 'tired', 'shy', 'anxious', 'low confidence', 'nervous'}

The above fine-grained sentiments yield the following behavior patterns:

[apprehensive, uneasy, withdrawn, low mood, lethargic]

The above behavioral patterns conclude the following academic distraction problems:

[unable to understand the material, lack of practice, incomplete classwork, mental health issues affecting academics, poor academic performance]

Experimentation

We have collected 155 sample user inputs that are written by the target teenage group (13 to 19 years). For calculating the accuracy of the computational model, we have consulted the domain experts (psychologists and psychiatrists). Their valuable inputs are incorporated at different levels in our computational model. We have used precision, recall, and f1 measure metrics to measure the performance of our model, for which the following assumptions in the context of the contingency matrix are being made.

Objective: The objective is to identify academic distraction problems from user input text. So, a positive outcome here means academic distraction-related sentiments are observed and a negative outcome means academic distraction-related sentiments are not observed. Similarly,

True positive: Sentiments are observed in the gold standard as well as in the predicted outcomes.

***True negative*:** Sentiments are not observed in the gold standard nor predicted outcomes.

***False-positive*:** Sentiments are not observed in the gold-standard outcome but sentiments are observed in the predicted outcomes.

False-negative: Sentiments are observed in gold-standard outcomes, but sentiments are not observed in predicted outcomes.

Precision emphasizes model accuracy considering true positives and total predicted positives. Precision is always recommended when the cost of false positives is important for the model. The recall is recommended when the false negative is important for the model.

For the F-score, we have considered β =1, as precision and

$$\text{Accuracy} = \frac{\text{True positive} + \text{True negative}}{\text{Total Sample space}}$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{Total predicated positive}}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{Total gold standard Positive}}$$

$$\text{F1 Score} = (1+\beta^2) * \frac{(\text{Precision} * \text{Recall})}{(\beta^2 * \text{precision}) + \text{Recall}}$$

Table 8: Sample result calculations for extracted sentiments.

Sample No	*Precision*	*Recall*	*f1 Score*	*Sample No*	*Precision*	*Recall*	*f1 Score*	*Sample No*	*Precision*	*Recall*	*f1 Score*
1	1.00	1.00	1.00	14.	0.33	1.00	0.50	27.	0.60	1.00	0.75
2	1.00	1.00	1.00	15.	1.00	1.00	1.00	28.	0.00	0.00	0.00
3	0.50	0.80	0.62	16.	0.50	1.00	0.67	29.	1.00	1.00	1.00
4	1.00	0.67	0.80	17.	0.00	0.00	0.00	30.	0.50	0.50	0.50
5	0.50	0.80	0.62	18.	1.00	1.00	1.00	31.	1.00	1.00	1.00
6	0.50	0.67	0.57	19.	1.00	1.00	1.00	32.	0.00	0.00	0.00
7	0.00	0.00	0.00	20.	0.50	0.67	0.57	…..	……	…..	……
8	1.00	1.00	1.00	21.	1.00	1.00	1.00	…..	……	…..	……
9	0.20	0.33	0.25	22.	1.00	1.00	1.00	152.	1.00	1.00	1.00
10	0.50	1.00	0.67	23.	1.00	1.00	1.00	153.	0.25	1.00	0.40
11	0.25	1.00	0.40	24.	0.00	0.00	0.00	154.	0.00	0.00	0.00
12	1.00	1.00	1.00	25.	0.25	0.50	0.33	155.	0.00	0.00	0.00
13	0.00	0.00	0.00	26.	1.00	1.00	1.00	AVG	0.51	0.65	0.55

Table 9: Sample result calculations for derived academic distraction problem.

Sample No	*Precision*	*Recall*	*f1 Score*	*Sample No*	*Precision*	*Recall*	*f1 Score*	*Sample No*	*Precision*	*Recall*	*f1 Score*
	1.00	1.00	1.00	14.	0.43	1.00	0.60	27.	0.67	0.80	0.73
	1.00	1.00	1.00	15.	1.00	1.00	1.00	28.	0.00	0.00	0.00
	0.67	0.67	0.67	16.	0.80	1.00	0.89	29.	1.00	0.83	0.91
	1.00	1.00	1.00	17.	0.17	0.25	0.20	30.	0.67	0.40	0.50
	0.63	0.63	0.63	18.	1.00	1.00	1.00	…..	……	…..	……
	1.00	1.00	1.00	19.	1.00	1.00	1.00	…..	……	…..	……
	0.00	0.00	0.00	20.	1.00	0.75	0.86	150.	1.00	1.00	1.00
	1.00	1.00	1.00	21.	1.00	1.00	1.00	151.	1.00	0.40	0.57
	0.80	0.80	0.80	22.	1.00	1.00	1.00	152.	1.00	1.00	1.00
	0.50	1.00	0.67	23.	1.00	1.00	1.00	153.	0.17	1.00	0.29
	0.33	0.50	0.40	24.	0.00	0.00	0.00	154.	0.00	0.00	0.00
	1.00	1.00	1.00	25.	0.67	1.00	0.80	155.	0.00	0.00	0.00
	0.25	0.25	0.25	26.	1.00	1.00	1.00	AVG	0.58	0.65	0.60

recall are equally important in our problem. False-positive and false-negative are equally important for our problem.

Result

We calculated precision, recall, and f1 score separately for individual input of the model. In the model, for each input paragraph, there may be a different outcome of different combinations from 29 core sentiments. So, we must process each outcome separately for precision and recall calculation. We have computed the results based on extracted sentiments and derived academic distraction problems that are highly correlated. The results are depicted in the following tables, Table 8 and Table 9. The calculated accuracy for extracted sentiment is 71 %

The calculated accuracy for derived academic distraction problem is 68 %

Error Analysis

During the discussion with experts, it came out that there are four types of errors, in this case, processing generalized vs personalized statements, conditional vs non-conditional statements, handling negation, and processing special words.

Error 1

There are two types of statements in the user input text, generalized statement, and personalized statement. In a generalized statement, the student writes a general perspective about teenagers' academic distraction problems whereas in a personalized statement student writes about his/her problems that distract him/her from academics. So, processing only those statements which include first-person pronouns (I, me, my), improves the cases where the predicted outcome is a false positive.

Illustration

Fear of failure is a thing to tackle amidst studies. After a hectic college day and HW studying becomes a difficult task.Pressure is automatically generated before examination period which is also good thing as it helps boost the confidence and in turn leads to more focused study tho it can be stressful at times.
Gold standard result- NIL (no sentiment found)
Predicted outcomes:
Sentiments: {'Fearful', 'anxious', 'tired', 'frustrated'}

Error 2

There are conditional statements written by teenagers in user input text where they want to specify sentiments that occur sometimes or with particular instances. In the gold standard, certain conditional sentiments are not considered for sentiment extraction. We have to process non-conditional statements while extracting sentiments.

Illustration

When my mom scolds me ,i get sad and am unable to study properly. I get irritated when i'm not understanding a concept .When I try reading it over and over and i still don't understand it.What bothers me is when i study History or civics i don't like these subjects.
Gold standard result- NIL (no sentiment found)

Predicted outcomes

Sentiments: {'Confused', 'irritable', 'Sad', 'anxious', 'upset', 'hate'}

Error 3

Sometimes while writing the text teenagers used to write sentiment words along with negation words, if it is so we have to skip including that sentiment in the output.

Illustration

No, sorrow does not distract me from studying.
Gold standard result- NIL (no sentiment found)
Predicted outcomes:
Sentiments: {'sad', 'distracted'}

Error 4

Special keywords like 'whatsapp', 'mobile' are used by teenagers with different context. We have to consider distracted sentiment only if it is written with that intension.

Illustration

I always forgot to see the timing of class in the WhatsApp.
Gold standard result-
Sentiments: {'forgetful'}
Predicted outcomes:
Sentiments: {'forgetful', 'distracted'}

Addition of Rules to improve accuracy

There are some lists that are used in the following Ruleset. The pronouns list is used for processing personalized sentences, one_time_words list is used to check conditional clauses. Negation_words list is prepared not to process sentiments when accompanied by negation_words, list_special_key is used to process special_words only when accompanied by these keys in a sentence.

- pronouns=['i','my','myself','me',"i'm"]
- one_time_words = ["sometimes","sometime","when"]
- negation_words = ['not','never','rarely','hardly','nothing', 'no','sometimes']
- list_special_key = ['attracted', 'attract', 'like', 'liked', 'distracted', 'distract', 'habit']

Rule-5: Processing personalized sentences only

If ((len(set(token).intersection(pronouns))>0):
then add sentiment record matching with one from lexicon to output-sentiment

Rule-6: Not processing sentences if conditional clause is present

If (len(set(token).intersection(one_time_words))==0)):
then add sentiment record matching with one from lexicon to output-sentiment

Rule-9: Not processing sentences if negative-words are present along with sentiment-word

If len(set(token).intersection(negation_list)) == 0:
then add sentiment record matching with one from lexicon to output-sentiment

Rule-10: Special words are considered when special-key words are present in sentence

If len(set(token).intersection(list_special_key)) > 0:
then add sentiment record matching with one from lexicon to output-sentiment

We have added the rules to overcome specified errors and validate the accuracy of experiments with precision, recall, and f1-score. We have improved the results with

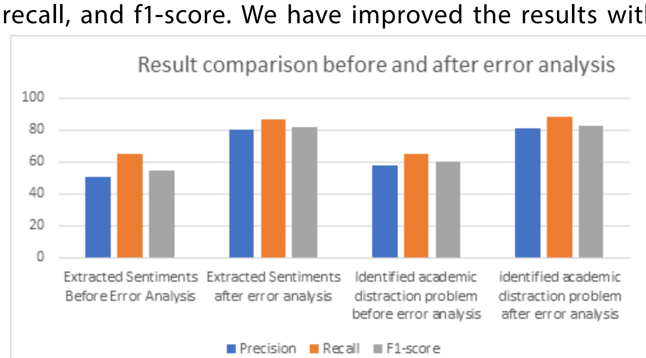

Figure 3: Result comparison before and after error analysis

80% precision, 87% recall, and 82% f1-score for extracted sentiments and 81% precision, 88% recall, and 83% for derived academic distraction problems. The calculated accuracy for extracted sentiment is 90%, and the calculated accuracy for derived academic distraction problem is 87%, after working on error analysis. Figure 3 shows the result comparison.

Conclusion

The system employs a hybrid model with a rule-based classifier and a sentiment-semantic lexicon-based approach to address the academic distraction problems of teenagers. We have gathered data from teenagers ranging in age from 13 to 19 years old. We have retrieved sentiments from user input text using a word-to-sentiment lexicon. For the supplied user input text, sentiment-to-behavior mapping is employed to derive behavior patterns and academic distraction problems. We have tested the model for extracting sentiment and academic distraction problems and found that it was accurate to 90 and 87%, respectively. Error analysis has pointed out clues for improving performance by processing non-conditional statements and personalized statements. We have created rules for non-conditional statements and personalized statements. With the use of shaded sentiment ontology, we will strive to treat various psychological difficulties in people of different ages in the future. To solve the numerous multi-class challenge, we are aiming to create a machine learning model. This model can also be extended to other domains where negative shaded sentiments are used, such as analyzing the impact of COVID-19 on human psychology, understanding employee sentiments to create a healthier company-work culture, and analyzing suicide instances from suicide notes.

Source(s) of support: Nil
Conflict of Interest: None

References

[1] Walaa Medhat a, Ahmed Hassan b, Hoda Korashy (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal 5, 1093–1113.
[2] Nurfadhlina Mohd Sharef, Harnani Mat Zin and Samaneh Nadali (2016). Overview and Future Opportunities of Sentiment Analysis Approaches for Big Data. Journal of Computer Sciences.
[3] Kerstin Denecke, Yihan Deng (2015). Sentiment analysis in medical settings: New opportunities and challenges. Artificial Intelligence in Medicine, Elsevier B.V,64,17–27.
[4] Yihan Deng, Thierry Declerck and Piroska Lendvai, Kerstin Denecke (2016). The Generation of a Corpus for Clinical Sentiment Analysis. ESWC: The Semantic Web pp 311-324, Springer.
[5] Geetika Gautam, Divakar Yadav (2014). Sentiment Analysis of Twitter Data Using Machine Learning Approaches and Semantic Analysis. IEEE-978-1-4799-5173-4/14.
[6] Saif Mohammad (2011). From once upon a time to happily ever after: tracking emotions in novels and fairy tales. Proceeding of the 5th ACL-HLT workshop.
[7] Pravin Jambhulkar et. Al (2014). Sentiment analysis for opinion mining using Cross-domain Classifier. International Journal on Recent and Innovation Trends in computing and communication.
[8] Joao Guerreiro et. Al (2019). How to predict explicit recommendations in online reviews using text mining and sentiment analysis. Elsevier.
[9] Li Yang et al (2020). Sentiment analysis for E-commerce product reviews in Chinese Based on Sentiment lexicon and deep learning. IEEE Access -Volume 8.
[10] H.-C.K.Lin et al (2020). Applying sentiment analysis to automatically classify consumer comments concerning marketing 4Cs aspects. Applied soft computing journal.
[11] Chao Song et. al (2020). SACPC: A framework based on probabilistic linguistic terms for short text sentiment analysis. ScienceDirect, knowledge-based systems 194.
[12] B. Liu (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies 5(1) 1-167.
[13] Ankita Rane, Dr. Anand Kumar (2018). Sentiment Classification System of Twitter Data for US Airline Service Analysis. 42nd IEEE International Conference on Computer Software & Applications.
[14] Fariska Zakhralativa Ruskanda, Dwi Hendratmo Widyantoro, Ayu Purwarianti (2018). Comparative Study on Language Rule-Based Methods for Aspect Extraction in Sentiment Analysis. IEEE.
[15] Sujata Rani, Parteek Kumar (2017). Rule-based Sentiment Analysis System for Analyzing Tweets. IEEE.
[16] Monali Kishor Patil, Nandini Chaudhari, Ram Bhavsar, BV Pawar (2020). A review on sentiment analysis in psychomedical diagnosis. REVIEW ARTICLE, Open Journal of Psychiatry & Allied Sciences, Epub: 12.
[17] Minqing Hu and Bing Liu (2004). Mining and Summarizing Customer Reviews. Proceedings of the ACM SIGKDD International Conference on Knowledge, Discovery and Data Mining (KDD2004), Aug 22-25, Seattle, Washington, USA.