



# A review on sentiment analysis in psychomedical diagnosis

## Abstract

As the research in sentiment analysis (SA) has reached the advanced level of maturity, scientific community is exploring possibilities of applying SA in different domains for which huge descriptive data is available. Researcher has developed tools for SA in commercial domains like feedback from customers, review rating of movie, political domains like exit poll prediction from tweets, etc. In medical science, huge data is available which can be used for predictive modelling. Psychomedical diagnosis is an interesting field which deals with exploring behavioural patterns to potential psychological problems. In the domain of medical science, little amount of work has been done in clinical SA. However, the efforts are at an elementary level in the course of research of SA for behavioural psychology. Hence, there is an opportunity of applying SA as the first step towards automatic psychomedical diagnosis of teenagers' behaviour patterns.

**Keywords:** NLP. ML. Emotion-Ontology. Mental Illness. Psychology.

**Monali Kishor Patil<sup>1</sup>,  
Nandini Chaudhari<sup>2</sup>, Ram Bhavsar<sup>3</sup>,  
BV Pawar<sup>4</sup>**

*<sup>1</sup>School of Computer Sciences, KBC-NMU, Jalgaon, Maharashtra, India, <sup>2</sup>Babaria Institute of Technology, Vadodra, Gujarat, India, <sup>3</sup>School of Computer Sciences, KBC-NMU, Jalgaon, Maharashtra, India, <sup>4</sup>School of Computer Sciences, KBC-NMU, Jalgaon, Maharashtra, India*

**Correspondence:** Monali Kishor Patil, B.E. (Computer Engineering), M.E. (Computer Engineering), Ph.D. student, School of Computer Sciences, KBC-NMU, Jalgaon, Maharashtra, India. 302/A3, Runval Garden City, Thane (west), Maharashtra, India. PIN: 400608. monalismart1981@gmail.com

**Received:** 19 January 2020

**Revised:** 3 June 2020

**Accepted:** 4 June 2020

**Epub:** 12 June 2020

## INTRODUCTION

Machine learning (ML) and natural language processing (NLP) are two streams under artificial intelligence (AI) as shown in Figure 1, while sentiment analysis (SA) is a new subfield which analyses written natural language text for the sentiment expressed in the given text. SA is a multidisciplinary field which uses both NLP and ML.

SA computationally identifies and categorises opinions which are expressed in a piece of text. It determines the writer's attitude towards topic whether it is positive, negative, or neutral. SA is used to analyse writers' sentiments, opinions, appraisals, attitudes, evaluations, and emotions, they expressed for entities. The entities can be organisations, products, services, individuals, topics, issues, and events. Nowadays, information is available online as well as in offline mode. This information can be exploited for different applications such as Customer Relationship Management (CRM), product reviews, product feedback, public voice analysis, crowd surveillance, etc.

In medical domain, research has been initiated for applications like patient opinion mining, detecting adverse drug events and effects, mining personal health information, patient knowledge retrieval, drug opinion mining, measuring healthcare quality, determining clinical outcome, etc.[1]

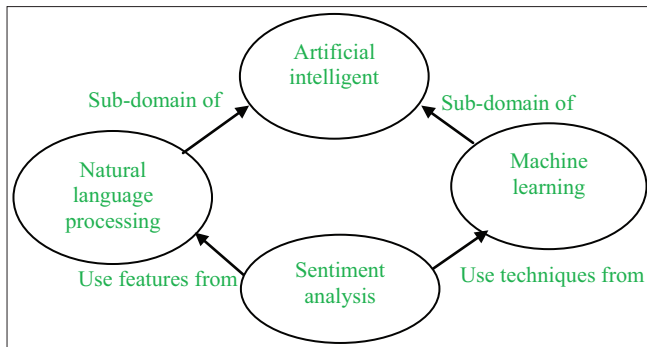
Clinical SA is concerned with patient's health status, medical conditions and treatment.[2]

In psychomedical domain, the researcher simulates psychological experiment paradigm of children's game task with an artificial emotion generating model.[3] Using NLP techniques and sentiment inference, standardised platform is developed to assist psychiatrist in patient diagnosis.[4] Behaviour pattern analysis is at an elementary level in psychomedical research.

As psychomedical research is at an elementary level, hence there is an urge to analyse behavioural patterns from written context. Youth in every country is the biggest asset for that nation. Nowadays, many psychomedical problems have been observed in teenagers. To interpret this psychomedical problem, there is a need to first understand emotion-ontology and human behaviour. So in next section, we have discussed about emotion-ontology and human behaviour followed by teenagers' behaviour patterns.

## I. SENTIMENT ANALYSIS OF TEENAGERS' BEHAVIOUR PATTERNS IN PSYCHOMEDICAL DOMAIN

From literature survey, it has been observed that sentiment is mapped according to polarity, i.e. positive, negative, and



**Figure 1:** Relationship between artificial intelligence, natural language processing, machine learning, and sentiment analysis.

neutral. There is a need to work on shades and intensity of sentiment to extract writer's emotions from written natural language text. In medical domain, clinical SA has been done. Behaviour pattern recognition is still the unfolded area in psychomedical application.

There is a major psychological problem observed in behaviour patterns among today's teenagers. It is difficult for parents to recognise psychological problems in their teenage children. In their crucial academic years, they are found distracted from their studies. So, there is an urge to know the reasons behind their behaviour in scientific way. We are trying to resolve this teenagers' academic distraction problem by analysing the shades and intensity of sentiment. This can be done by extracting the sentiment-words (for example, *alash! Brao! Shit! Oh no! waw! Amazing!*) from self-portrayals, description of patient, and session notes written by psychiatrist. In this section, we first discussed emotion-ontology and then human behaviour patterns, followed by teenagers' psychomedical problem.

### A. Emotion ontology

Dr. Ken McGill holds a doctorate degree in clinical psychology with an emphasis on family psychology. He suggested how to use emotion-ontology chart to identify emotions and feelings expressed by writer in his written context.[5]

The emotion-ontology is developed for the emotion-areas where affective phenomena are significant. Emotion-ontology addresses affective phenomena in multidisciplinary fields like psychology, psychiatry, neuroscience, biomedicine, and the life sciences. Shades of emotions are shown in Figure 2.

In different research, capturing emotional experience by client is important scenario. The emotional experience can be self-reported emotions which are often captured to monitor mood fluctuations between clinical visits in the clinical treatment of depression and bipolar disorder.

The disturbing, shocking, or "triggering" events in someone's life can cause mental and/or emotional dysregulation within that person. These events diminish his/her ability to "see the other side of life". Self-reported emotional experiences may also be useful in the assessment of psychological problem. The emotion-ontology chart will help in this SA of written natural language text.

### B. Behavioural patterns in teenagers

Dr. Gary Chapman, has written about teenagers' problems in his "The Five Love Languages",[6] that had sold four million copies in English alone and had been translated into 36 languages including Arabic and Hindi as well.

Teenagers search for independence and to establish self-identity. There are many worries in teenagers' minds like "Which morality and ideologies should they follow?" "How to tackle sexuality issues?" "How to achieve success in life?" The excessive use of mobile phones and multimedia, nuclear families, no ethics create serious issues in today's teenagers. Verbal exploitation such as unfriendly, hurtful, strict, or demeaning words or physical abuse will damage teenagers' emotional development. A hollow love tank has a great impact on teenagers' life like an inspiration for learning is dissipated ("Why should I study in school? As nobody wants to understand what is happening to me?"). Teenagers are always susceptible to negative role models.

The above information helps in identifying the root cause of teenagers' psychological problems and predicting behavioural syndromes in teenagers' academic distractions. Now, we need to understand mental illness faced by children.

Mental illness in children is the biggest concern for today's parents. Many children who could benefit from treatment do not get the help they need at the proper time. By designing the predictive framework to address the academic distraction problems of teenagers, several psychomedical problems can be resolved at an early stage in children. The taxonomy of mental illness in children is shown in Figure 3.

Here, we have discussed the disorders considering the scope of our problem, so there may be other mental disorders that are not included in the following discussion. Mental health disorder in teenagers are-

1. Childhood behavioural disorders
 

Childhood behavioural disorders are the sixth leading cause of disease 'burden' among teenagers. Teenage can be a time where rules, limits, and boundaries are tested.

  - 1.1 Anxiety disorder
 

Obsessive-compulsive disorder, posttraumatic stress disorder, social phobia, and generalized anxiety disorder experience anxiety as a persistent problem that interferes with the daily activities of children. Many a time, heavy stress or worry causes anxiety disorder.
  - 1.2 Attention-deficit/hyperactivity disorder
 

Attention-deficit/hyperactive disorder (ADHD) typically includes symptoms like difficulty paying attention, hyperactivity, and impulsive behaviour. Children with ADHD have faced problems in all these or maybe in one only.
  - 1.3 Autism spectrum disorder
 

Autism spectrum disorder (ASD) is a serious developmental disorder that appears in early childhood, usually before age three. It was observed that ASD causes a major problem in communication and interaction of a child in society.



symptoms observed in emotional disorders as they are inter-mixed.

4. **Psychosis**  
Disorders which include symptoms of psychosis, most commonly emerge in late teenage or early adulthood. Symptoms of psychosis can contain hallucinations (such as hearing or seeing things which are not in existence) or delusions (including fixed, inappropriate beliefs). Psychosis affects teenagers' life so badly that they cannot concentrate on their studies properly.
5. **Suicide and self-harm disorder**  
According to the World Health Organization (WHO),[7] it is estimated that 62000 teenagers died in 2016 as a result of self-harm. Suicide is the third leading cause of death in older teenagers (15-19 years). Nearly 90% of the world's teenagers live in low- or middle-income countries but more than 90% of teenage suicides are among teenagers living in those countries. Suicide attempts are thoughtless and they create a feeling of hopelessness or loneliness. Many times, it was observed that excessive use of alcohol, abuse in childhood, shame using help from others cause a risk for suicide attempt. Communication through digital media about suicidal behaviour is an emerging concern for this age group.
6. **Risk-taking behaviour problem**  
Many risk-taking behaviours for health, such as substance use or sexual risk-taking starts in teenage. Risk-taking behaviours can be both an unhelpful strategy to cope with poor mental health, and can negatively contribute to and severely impact a teenager's mental and physical wellbeing.

All the above problems contribute to the academic distraction problem of the teenager. According to Mike Hobbiss, a Ph.D. student from the Institute of Cognitive Neuroscience at University College London (UCL),[8] teenage is a fascinating period. There are different compartments in our brains for different functions. Some areas of the brain are fully developed and working 100% while another area is not developed properly. 'Cognitive control' region of brain always develop at the end and the region where emotions and rewards get handled are developed prior. Following are ten reasons for teenagers' academic distraction-

1. **Lack of practice**  
Many young children are passive listener; they just listen the lessons taught by teacher. They never try to practice it; so, after some time, they are not able to recall the lessons and they loss interest in study.
2. **Does not understand the material**  
Sometimes students do not understand the notes which are given to them for study. This lack of understanding can lead to students stop paying attention, and consequently, fall further behind.
3. **Is not being challenged enough**  
Every student has different intellectual power. More intellectual students face problem when they do not get enough challenging task to do. If they do not receive extra difficult-task they will be bored and stop giving attention in class.
4. **Distracted by external stimuli**  
Many a time, some students are not able to concentrate because of talkative classmates or a messy lab. All children

are not capable to tackle with distracted environment and so they are not able to pay attention to a teacher.

5. **Lack of motivation**  
In some cases, the child's concentration problem may actually be a motivation problem. Child always needs motivation; else, they loose interest in studies.
6. **Mismatched learning style**  
Each student memorises his/her study in various ways. Some just read, some read and write, and some just need to be attentive in class. If teacher is working with a specific learning tool which is difficult for the student to understand the topic then the child loses his/her interest in studies.
7. **Not getting proper sleep or nutrition**  
If the child is not getting the recommended eight to ten hours of sleep each night, he or she will not have the energy needed to concentrate in class. If the child is heading to class hungry as might be he/she did not eat anything while coming to school, he or she is more apt to be distracted than learning-ready.
8. **Disorganisation problems**  
Some students do not keep their textbooks and notebooks according to timetable in their bags. Coming to class disorganised means the kid is spending time checking out the tools and material needed to find out instead of listening to what is being taught.
9. **School anxiety**  
Many students become anxious before going to school. Sometimes peer pressure or parents' pressure make child worried about their results which leads to loss of childrens' confidence.
10. **Learning difficulties**  
If the child is having severe problems in the classroom, such as constant disruptions, distractions, or poor grades and other items have been ruled out from this list then attention should be given for his/her learning difficulties. Sometimes children may face disorders like attention deficit disorder, ADHD, or dyslexia. The child may also have hearing problem such as central auditory discrimination disorder.

After a complete review of medical domain, we should focus on how SA can be used to solve this problem. In subsequent sections, we have discussed different approaches for SA classification, SA tools, and applications of SA.

## II. 360° VIEW OF SENTIMENT ANALYSIS

Here we try to cover all aspects related to SA including classification levels for SA, SA approaches, different SA tools, and applications of SA.

### A. Classification level in sentiment analysis

Levels in SA is broadly classified in three categories-

1. **Sentence level classification**  
Sentence is considered as an unit for SA. It is the most fine-grained analysis. A sentence can be categorised as subjective sentence or objective sentence. Subjective sentence has opinions, e.g. "Your saree colour is too bright". In objective sentence, facts are depicted.

Objective sentence has no judgement or opinion about object or entity, e.g. “Most of the Indian women have a huge collection of sarees”.

2. Document level classification  
The whole document is considered as an unit for SA. It is assumed that document is having opinion about single entity or object. Irrelevant sentences are removed before processing for polarity mapping of document level SA.
3. Aspect level sentiment analysis  
It classifies the sentiment with respect to specific aspect of an entity. Opinion holder may express opinions for different aspects of same entity, e.g. colour of a saree is very nice but texture quality is not up to the mark.

**B. Sentiment analysis approaches**

SA approaches broadly categorise as ‘lexicon-based approach’, ‘ML approach’, and ‘hybrid approach’. The lexicon-based approach bank on sentiment lexicon. Known and precompiled sentiment-terms group together to form a lexicon. Figure 4 gives complete view of different SA approaches.[9] Lexicon-based approach is classified as corpus-based approach and dictionary-based approach. Corpus-based approach is further divided into statistical and semantic approach. The hybrid approach makes a combination of lexicon-based and ML approaches. Following are some commonly used supervised learning ML approaches for SA.

1. Support vector machine classifiers  
Support vector machine (SVM) finds best separator for different classes in the search space. Text data is based suited for SVM as the sparse nature of text. SVM can be used for sentiment polarity classifier. In the learning

process, SVM seek for the best hyper-plane among several existing alternative hyper-planes (discrimination boundaries). In linear SVM, there are two classes which should be separated by hyper-plane, SVM should seek for the best hyper-plane to separate two classes of positive and negative.

2. Naïve Bayes classifier  
It is easy to work with Naïve Bayes (NB) classifier during training phase as well as classifying phase. It is the most frequently used classifier. A NB classifier works with the principle that the existance of a specific feature in a group is not related to the existance of any other feature. For example, we can say that “this is a banana”, if it is yellow, curve shape, and about five inches long. Here, each feature of ‘banana’ seperately contributes to the probability that “this is a banana”.
3. Maximum entropy  
The maxent classifier which is also called as a conditional exponential classifier. It converts labelled feature sets to vectors using encoding. This encoded vector is then used to calculate weights for each feature followed by combining them to determine the most likely label for a feature set.[9] Maximum entropy (ME) is achieved with the help of conditional probability. It can find distribution over classes like logistic regression. The overlapping features between classes can be handled by ME. It also follows certain feature exception constraints. It follows the similar processes as NB, for finding the polarity of the sentiments.
4. Rule-based classifier  
The data space is exhibited by the set of rules. The left-side of the rule represents a condition on the feature

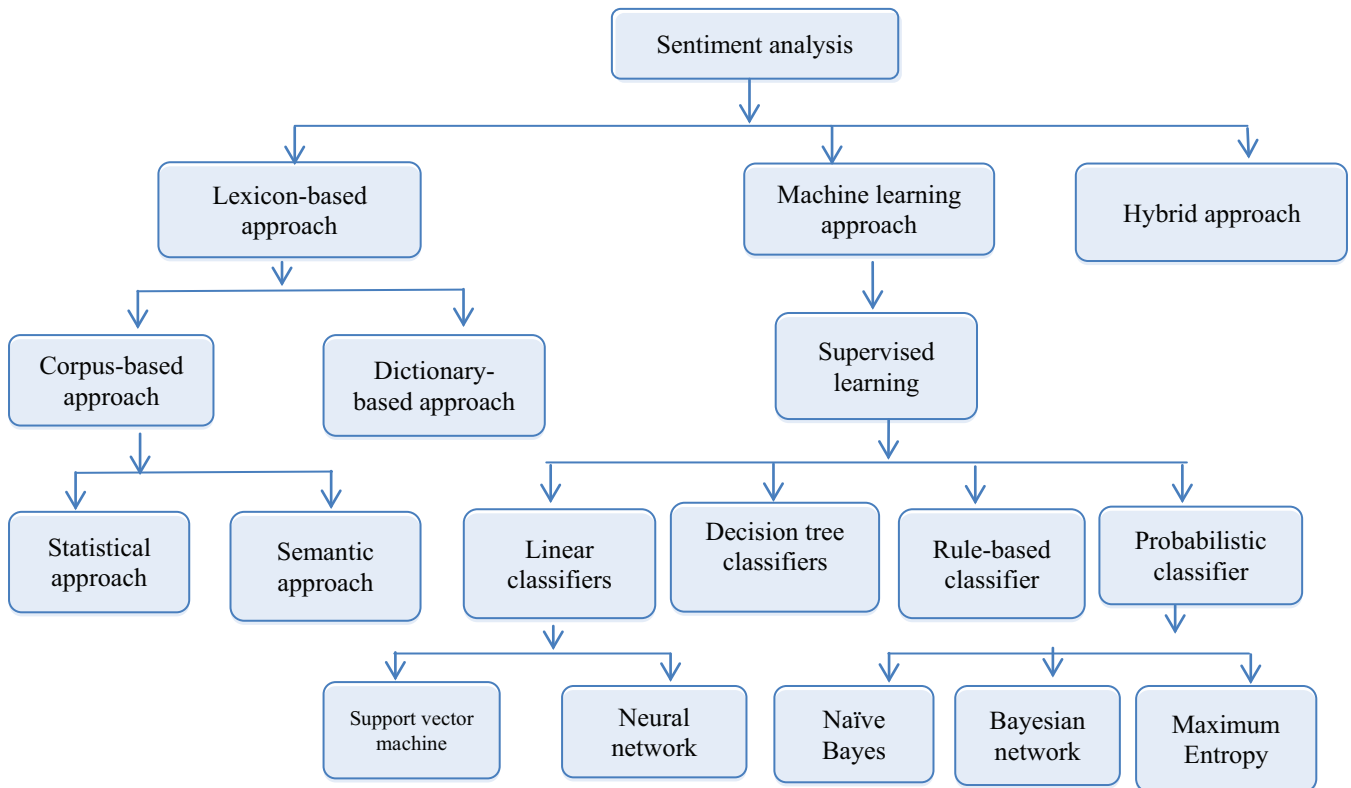


Figure 4: Classification techniques of sentiment analysis.[9]

set which is expressed in disjunctive normal form. The right-side of the rule is the class label. The conditions are based on the term presence. In training phase, all the rules are constructed depending on the criteria used. The two general criteria are support and confidence. The support is the complete number of instances in the training dataset which are appropriate to the rule. The conditional probability is denoted by confidence where the right-side part of the rule is fulfilled if the left-side part is fulfilled.

### III. SURVEY ON SENTIMENT ANALYSIS

Pang *et al.*[10] started working on syntactic attributes n-gram method. Research scholars classify opinionated text into different levels. Mullen and Collier[11] and Wiebe *et al.*[12] worked on sentiment polarity classification at document level. Riloff *et al.*[13] checked whether sentence is subjective or objective. Wilson *et al.*[14] developed the concept phrase level classification. It was observed from literature that SVM and NB classifiers from probabilistic class is used by maximum researchers as it gives better results.[15]

#### A. Survey of sentiment analysis approached

Gautam and Yadav[16] worked on three basic techniques, NB, ME, and SVM, and introduced new technique, semantic orientation with WordNet followed by basic technique. Researcher used python and natural language toolkit (NLTK) to train and classify data with NB and SVM. Total dataset of size 19340 was used; out of which, 18340 was used for training and 1000 was used or testing.

Hamzah and Widyastuti[17] checked performance of ME and k-mean clustering (KMC). Result showed that KMC gave better performance than ME on an average three per cent precision. KMC was faster than ME by 25 msec in analysing 2000 text opinions.[17]

Ficamos *et al.*[18] worked on ME approach to SA which captures domain specific data in Webio. Complexity of n-gram feature extraction was more than bi-gram method.

A multi-criteria decision making approach was created by Kumar[19] for endorsing a product with the help of SA. The list of top essential features which influenced customer was identified by decision making process.

Rani and Kumar[20] analysed rule based SA system for tweets.

Rane and Kumar[21] worked on multiclass SA with preprocessing on data to clean the tweets.

Kurniawat and Pardede[22] used particle swarm optimization (PSO), information gain techniques, and SVM classifier to choose appropriate characteristics from input documents.

Idrus *et al.*[23] worked on text mining which is a technology capable of analysing semi-structured and unstructured text data where data mining processes structured data only. The accuracy for NB classifier algorithm was 73.33% while combing PSO with NB improved its accuracy to 76%.

Kristiyanti *et al.*[24] compared SVM and NB algorithms. They took public opinion from Twitter as input data for the topic “West Java governor candidate” from the period of 2018 to 2023. Datasets collected in this study was a separate set of texts in the form of documents collected from Twitter with crawling methods from social media of Twitter only for Indonesian language.[24] SVM and NB is popular algorithm in SA classification because of their good performance and high accuracy. 2-gram method is used before applying SVM and NB in this study. NB exceled with highest accuracy of 94% as SVM produced highest of 75.5 % accuracy.

Ruskanda *et al.*[25] focused on language study with the help of rule-based approach. Here, researchers concentrated on content of dataset as subjective sentences, annotated aspects per sentence, and the value of polarity per aspect as the first step in their research.

Birbeck and Cliff[26] used Bayesian classifier to generate stock predictions, which was input to an automated trading algorithm. This approach performed better than random chance.

Neural networks are always applicable to transfer features for adapting text classification models from source domain to a target domain. A capsule-based hybrid neural network which mined richer textual information to improve expression capacity was created by Du *et al.*[27] Attention-based model uses self-attention mechanism and convolutional neural networks (CNN). The hybrid capsule-based model results into better performance as it uses less training time and simple network structure. This model was tested on two short text review datasets. On movie review data, it resulted in highest accuracy of 82.55%.[27]

Ding *et al.*[28] developed multi-domain adversarial neural network for text classification. Multi-domain adversarial training strategies along with orthogonal constraints were used to separate private and shared feature with each other. In this way performance of source and target domain was improved.[28]

Manshu and Bing[29] worked on cross-domain sentiment classification (CDSC). Sometimes domain has divergence, so a sentiment classifier performs less well when directly applied to these domains. Domain independent features were learned by adversarial neural networks. For example ‘pivots’, these are words with same sentiment polarity in different domains.[29] A hierarchical attention network with prior knowledge information (HANP) for CDSC task was proposed. Both domain independent and domain specific features were obtained. Considering Amazon reviews as dataset, HANP showed effective accuracy in experimental results. Table 1 concludes the survey of SA approaches.

The summary of above research work for SA is mapped in terms of advantages and disadvantages in Table 2. Each SA techniques has its pros and cons. Dictionary- and corpus-based approach should be combined with probabilistic or linear classifier to get good result. From literature, it is observed that SVM and NB techniques were used by maximum researchers and they got comparatively good results in terms of accuracy and efficiency.

## B. Sentiment analysis in medical domain

Exploring information in health-related, social media services or in terms of hard bond reports is of great interest for patients, researchers, and medical companies. Patient's health status, medical conditions, and treatment for his/her illness is the main concern in medical SA. The challenge is, however, to provide easy, quick, and relevant analysis of the vast amounts of information available online and offline.

Internet users searching for health information and support are increasing day by day. According to Google, an

estimated seven per cent of Google's daily searches are health-related; Google's total daily health-related searches amount to 70,000 each minute.[30] SA for the medical domain has a broader range as compared to SA for the business domain. Facets of sentiment in health-related texts concern patient status, medical condition or treatment.

### Survey of psychomedical sentiment analysis

SA is a multidisciplinary problem. Research work has been done for polarity mapping as positive, negative, or neutral. Some devices are also developed to recognise emotions from

**Table 1: Survey of sentiment analysis approaches**

Author	Year	Technique used	Result
Gautam and Yadav[16]	2014	ME, NB, and SVM	Accuracy of NB is 88.2%, ME is 82.8%, and SVM is 85.5%; with the help of semantic analysis, WordNet accuracy is improved to 89.9%.
Hamzah and Widyastuti[17]	2016	ME and KMC	KMC gave better performance than ME on 3% precision.
Ficamos et al.[18]	2017	ME and POS to extract uni-gram and bi-gram features	Complexity of n-gram feature extraction was more than bi-gram method.
Kumar[19]	2017	Multi-criteria decision making approach for product recommendation system	Developed multi-criteria decision making system.
Rani and Kumar[20]	2017	Rule-based SA	86% accuracy is achieved for 500 tweet records and 94% for 200 English sentences.
Rane and Kumar[21]	2018	Multicast SA with 7 techniques -Decision tree, Random forest, SVM, K-nearest neighbour, Logistic regression, Gaussian NB, and AdaBoost	84.5% accuracy of AdaBoost classifier.
Kurniawat and Pardede[22]	2018	PSO with SVM classifier	Accuracy of SVM was improved by 94.8%.
Idrus et al.[23]	2018	NB	Accuracy was improved to 76% of NB with PSO.
Kristiyanti et al.[24]	2018	SVM and NB	2-gram method is used before applying SVM and NB. NB exceeded with highest accuracy of 94% as SVM produced highest of 75.5 % accuracy.
Ruskanda et al.[25]	2018	Rule-based approach	Aspect extraction in SA is achieved.
Birbeck and Cliff[26]	2018	Bayesian classifier	Stock prediction system is developed.
Du et al.[27]	2019	Capsule-based hybrid neural network	On movie review, 82.55 % accuracy is achieved.
Ding et al.[28]	2019	Multi-domain adversarial neural network	Performance of source and target domain is improved.

ME: Maximum entropy; NB: Naive Bayes; SVM: Support vector machine; KMC: K-mean clustering; POS: Part of speech; SA: Sentiment analysis; PSO: Particle swarm optimization.

**Table 2: Comparison of commonly used machine learning approaches for sentiment analysis**

	Approaches				
	Rule-based system	SVM	NB	ME	n-gram
Advantages	Writing and implementing rules is easy.	1. High dimensional input space. 2. Document vectors are sparse.	1. Simple and intuitive technique. 2. It combines efficiency with reasonable accuracy.	This technique does not assume the independent features like NB technique.	Usage of 2-gram and more improve accuracy of SA instead of single word SA.
Disadvantages	It loses track when the data and scenario change faster than updates in rules.	1. More datasets are required in training phase. 2. Data collection is not easy.	1. Machine can be trained with less datasets. 2. It works with conditional independence between the linguistic features.	It is tedious.	1. n-gram dependencies can not be handled if n value is large. 2. Corpus data required to train machine.

SVM: Support vector machine; NB: Naive Bayes; ME: Maximum entropy; SA: Sentiment analysis.

human-activities, mostly to help old age people. Recent trends in the field of SA are finding strength of sentiment in symbolic texts, making categories of reviews, detecting attitude from the tweet data, detecting sentiment in clinical texts, and the moods from music.

Many researchers are working on automated text analysis methods to find out how to measure psychological and demographic properties. An enormous amount of human-generated text needed to be analysed. So, there is a need to search a suitable ML method which handles this psychologically relevant data (Table 3).

Iliev *et al.*[31] reviewed the widespread methods of automated text analysis from the viewpoint of social scientists. User-defined dictionary (UDD), feature extraction, and word co-occurrences are the three approaches discussed along with semantic role labelling, cohesion, hybrid methods by Iliev *et al.*[31] to encourage psychologists to add automated text analysis to their methodological toolbox.

Tighe *et al.*[32] specified techniques of feature reduction to categorise the writer's personality characters. Principal components analysis (PCA) is used to identify patterns and highlights the similarities and differences in data. Learning algorithms used were two SVM (library for SVM [libSVM] and sequential minimal optimization [SMO]), linear logistic regression (simple logistic), one k-nearest neighbour (instance-bases learning with parameter k [IBk], where k equaled one and five), C4.5 decision tree (J48), NB, Random forest.[32] www.MyPersonality.org website was used to retrieve data consisting of 2468 essays or daily writing submissions from 34 psychology students. Waikato environment for knowledge analysis or Weka, a machine learning algorithm and data pre-processing tool was used.

Ronad *et al.*[33] wrote theoretically about child and adolescent mental health in the Indian context. They specified that many people suffering from psychiatric illnesses remain untreated, even though treatment exists, as stated by WHO. WHO reported that one in every five children has mental health issues. If the problems get identified in early-stage, it will prevent expensive rehabilitation programmes and adult treatment. It is possible to prevent the majority of behaviour disorders in the school environment itself.[33] Mental health problems in children are associated with educational failure, family disruption, disability, offending and antisocial behaviour, substance abuse, and delinquency.

Kamath *et al.*[4] have provided an intelligent digital assistance system for psychiatrists in the treatment of critical psychological disorders like schizophrenia and bipolar disorder. They developed the semantics-based gradient processing algorithm. They used a NB model, term frequency-inverse document frequency (TF-IDF) method, and NLTK. They used web-based application, Apache Mahout, Java code, and for pre-processing, python language was used.[4]

Garcia-Ceja *et al.*[34] used sensor data in their work. The main concern in this survey was mental disorders/conditions like depression, anxiety, bipolar disorder, and stress. Wearable devices like smartphones, smartwatches, and fitness bands have embedded sensors in them, as well as they have communication devices (Wi-Fi, Bluetooth), inertial sensors (accelerometer, gyroscope), physiological sensors (heart rate, dermal activity), and ambient sensors (ambient pressure, temperature).[34] Using this sensor data along with ML methods can give a piece of meaningful information about a person's current health state. Researchers surveyed on the basis of three taxonomy: type of study (association, detection,

**Table 3: Survey of psychomedical sentiment analysis**

Author	Year	Techniques and tools used	Result	Review
Iliev <i>et al.</i> [31]	2014	Survey paper	-	Survey has been done under three groups – user-defined dictionaries, feature extraction, and word co-occurrence.
Tighe <i>et al.</i> [32]	2016	LibSVM, simple logistic, SMO, ZeroR	Accuracy achieved was 57.54%, 56.04%, 55.75%, 58.31%, and 61.95% for agreeableness, conscientiousness, extraversion, neuroticism, and openness to experience respectively using feature reduced datasets.	The research is beneficial to culture-specific and cross-culture psychology.
Ronad <i>et al.</i> [33]	2017	-	Theoretically explained child and adolescent mental health in Indian context.	Mental health problem in children are associated with educational failure, family disruption, disability, offending and antisocial behaviour, substance abuse, delinquency.
Kamath <i>et al.</i> [4]	2018	NB, TF-IDF, NLTK	It provided an intelligent digital assistant system for psychiatrist.	Predictive model for behaviour patterns needs to develop.
Shi and Li[3]	2018	Quality simulation (QSIM) reasoning and analytic hierarchy process, BP (back propagation) neural network, mood emotion model, association memory model	Emotion decision and associative memory model of humanoid robot in the process of facial expression is developed.	The computer emotion decision based on artificial intelligence can be applied in different fields.

LibSVM: Library for support vector machine; SMO: Sequential minimal optimization; ZeroR: Zero rule; NB: Naïve Bayes; TF-IDF: Term frequency-inverse document frequency; NLTK: Natural language toolkit



forecasting), the timeline of study (short term or long term), and type of sensing devices (wearable, external, and software/social media).

Shi and Li[3] contributed in the AI domain by giving the psychological experiment paradigm of children's game task, test process of emotional spontaneous transfer and stimulus transfer model, reasoning method and analytical hierarchy process (AHP) is multi-system. They established a mood state regulation algorithm, based on emotion energy theory and a combination of Hidden Markov Model (HMM)-based spontaneous transfer and stimulus transfer.[3] They developed an emotion decision and associative memory model of a humanoid robot in the process of facial expression interaction in 2018.

## Conclusion

There is a need to analyse the behaviour patterns of teenagers who are facing psychological problem and taking treatment under psychologist or psychiatrist. The written report about patient by doctor, the self-portrait of teenager patient, the description about teenager by parents about his behaviour will be considered as input context. The behaviour problem will be extracted from the text by analysing shades of sentiment observed in the input context. This will help to link the behaviour problem at teenage with some symptoms in early childhood. With the help of this prediction, we will be able to avoid the brutality of psychomedical problems at teenage.

## REFERENCES

- Denecke K, Deng Y. Sentiment analysis in medical settings: new opportunities and challenges. *Artif Intell Med.* 2015;64:17-27.
- Deng Y, Declerck T, Lendvai P, Denecke K. The generation of a corpus for clinical sentiment analysis. In: Sack H, Rizzo G, Steinmetz N, Mladenici D, Auer S, Lange C, editors. *The Semantic Web -- ESWC 2016 Satellite Events. Lecture Notes in Computer Science (LNCS) 9989.* Springer; 2016:311-24.
- Shi Y, Li C. Exploration of computer emotion decision based on artificial intelligence [Internet]. In: 2018 International Conference on Virtual Reality and Intelligent Systems (ICVRIS) [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/8531405>
- Kamath A, Raje N, Konduri S, Shah H, Naik V, Bhattacharjee K, et al. Intelligent AI assisted psychological disorder analysis using sentiment inference [Internet]. In: 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI) [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/8554493>
- McGill K. Emotions and feelings charts [Internet]. 2015 Aug 13 [cited 2020 Jan 19]. Available from: <https://325424.com/2015/08/13/emotions-and-feelings-charts/>
- Chapman G. *The five love languages.* Northfield Publishing; 1992.
- World Health Organization. Adolescent mental health [Internet]. 2019 Oct 23 [cited 2020 Jan 19]. Available from: <https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health>
- Brookman-Byrne A. Educational implications of attention and distraction in teenagers [Internet]. In: *npj Science of Learning.* 2017 Feb 6 [cited 2020 Jan 19]. Available from: <https://npjscilearncommunity.nature.com/users/20788-dr-annie-brookman-byrne/posts/14855-educational-implications-of-attention-and-distraction-in-teenagers>
- Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: a survey. *Ain Shams Eng J.* 2014;5:1093-113.
- Pang B, Lee L, Vaithyanathan S. Thumbs up?: sentiment classification using machine learning techniques [Internet]. Philadelphia: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP); 2002 Jul [cited 2020 Jan 19]; 79-86. Available from: <https://www.aclweb.org/anthology/W02-1011.pdf>
- Mullen T, Collier N. Sentiment analysis using support vector machines with diverse information sources [Internet]. Barcelona: Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing; 2004 Jul [cited 2020 Jan 19]; 412-8. Available from: <https://www.aclweb.org/anthology/W04-3253.pdf>
- Wiebe J, Wilson T, Cardie C. Annotating expressions of opinions and emotions in language. *Language Res Eval.* 2005;39:165-210.
- Riloff E, Wiebe J, Wilson T. Learning subjective nouns using extraction pattern bootstrapping [Internet]. Edmonton: Proceedings of the Seventh CoNLL conference at HLT-NAACL; 2003 May-Jun [cited 2020 Jan 19]; 25-32. Available from: <https://www.aclweb.org/anthology/W03-0404.pdf>
- Wilson T, Wiebe J, Hoffmann P. Recognizing contextual polarity in phrase-level sentiment analysis [Internet]. Vancouver: Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP); 2005 Oct [cited 2020 Jan 19]; 347-54. Available from: <https://www.aclweb.org/anthology/H05-1044.pdf>
- Sharef NM, Zin HM, Nadali S. Overview and future opportunities of sentiment analysis approaches for big data. *J Comput Sci.* 2016;12:153-68.
- Gautam G, Yadav D. Sentiment analysis of twitter data using machine learning approaches and semantic analysis [Internet]. In: 2014 Seventh International Conference on Contemporary Computing (IC3) [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/6897213>
- Hamzah A, Widyastuti N. Opinion classification using maximum entropy and k-means clustering [Internet]. In: 2016 International Conference on Information & Communication Technology and Systems (ICTS) [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/7910292>
- Ficamos P, Liu Y, Chen W. A naive Bayes and maximum entropy approach to sentiment analysis: capturing domain-specific data in Weibo [Internet]. In: 2017 IEEE International Conference on Big Data and Smart Computing (BigComp) [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/7881689>
- Kumar G. A multi-criteria decision making approach for recommending a product using sentiment analysis [Internet]. In: 2018 12th International Conference on Research Challenges in Information Science (RCIS) [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/8406679>
- Rani S, Kumar P. Rule based sentiment analysis system for analyzing tweets [Internet]. In: 2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS) [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/8286061>
- Rane A, Kumar A. Sentiment classification system of Twitter data for US Airline service analysis [Internet]. In: 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC) [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/8377739>
- Kurniawat I, Pardede HF. Hybrid method of information gain and particle swarm optimization for selection of features of SVM-based sentiment analysis [Internet]. In: 2018 International Conference on Information Technology Systems and Innovation (ICITSI) [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/8695953>
- Idrus A, Brawijaya H, Maruloh. Sentiment analysis of state officials news on online media based on public opinion using Naive Bayes classifier algorithm and particle swarm optimization [Internet]. Published in: 2018 6th International Conference on Cyber and IT Service Management (CITSM) [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/8674331/>
- Kristiyanti DA, Umam AH, Wahyudi M, Amin R, Marlinda L. Comparison of SVM & Naive Bayes algorithm for sentiment analysis toward West Java governor candidate period 2018-2023 based on public opinion on twitter [Internet]. In: 2018 6th International Conference on Cyber and IT Service Management

- (CITSM) [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/8674352>
25. Ruskanda FZ, Widyantoro DH, Purwarianti A. Comparative study on language rule based methods for aspect extraction in sentiment analysis [Internet]. In: 2018 International Conference on Asian Language Processing (IALP). Available from: <https://ieeexplore.ieee.org/document/8629140>
  26. Birbeck E, Cliff D. Using stock prices as ground truth in sentiment analysis to generate profitable trading signals [Internet]. In: 2018 IEEE Symposium Series on Computational Intelligence (SSCI) [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/8628841>
  27. Du Y, Zhao X, He M, Guo W. A novel capsule based hybrid neural network for sentiment classification [Internet]. In: IEEE Access 2019 Mar 20 [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/8672127>
  28. Ding X, Shi Q, Cai B, Liu T, Zhao Y, Ye Q. Learning multi-domain adversarial neural networks for text classification [Internet]. In: IEEE Access 2019 Mar 13 [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/8666710>
  29. Manshu T, Bing W. Adding prior knowledge in hierarchical attention neural network for cross domain sentiment classification [Internet]. In: IEEE Access 2019 Mar 5 [cited 2020 Jan 19]. Available from: <https://ieeexplore.ieee.org/document/8658231>
  30. Murphy M. Dr Google will see you now: Search giant wants to cash in on your medical queries [Internet]. In: The Telegraph. 2019 Mar 10 [cited 2020 Jan 19]. Available from: <https://www.telegraph.co.uk/technology/2019/03/10/google-sifting-one-billion-health-questions-day/>
  31. Iliev R, Dehghani M, Sagi E. Automated text analysis in psychology: methods, applications, and future developments. *Lang Cogn*. 2015;7:265-90.
  32. Tighe EP, Ureta JC, Pollo BAL, Cheng CK, Bulos RDD. Personality trait classification of essays with the application of feature reduction [Internet]. In: Proceedings of the 4th Workshop on Sentiment Analysis where AI meets Psychology (SAAIP) co-located with 25th International Joint Conference on Artificial Intelligence (IJCAI 2016); New York City, USA. 2016 Jul 10 [cited 2020 Jan 19]. Available from: <http://ceur-ws.org/Vol-1619/paper4.pdf>
  33. Ronad SV, Babu RB, Patali CS, K S, Ugargol SS, Gondabal SH, *et al.* Child and adolescent mental health in Indian context. *Res Med Eng Sci*. 2017;2:172-7.
  34. Garcia-Ceja E, Riegler M, Nordgreen T, Jakobsen P, Oedegaard KJ, Tørresen J. Mental health monitoring with multimodal sensing and machine learning: a survey. *Pervasive Mob Comput*. 2018;51:1-26.

Patil MK, Chaudhari N, Bhavsar R, Pawar BV. A review on sentiment analysis in psychomedical diagnosis. *Open J Psychiatry Allied Sci*. 2020 Jun 12. [Epub ahead of print]

**Source of support:** Nil. **Declaration of interest:** None.