



# Sentiment Analysis with Supervised Learning Techniques: A Survey

Nuha Elamin<sup>1,\*</sup>, Samani A.Talab<sup>2</sup> and Ahmed Khalid<sup>3</sup>

<sup>1</sup>Faculty of Graduate Studies, Neelian University, Khartoum, Sudan

<sup>2</sup>Faculty of Computer Science and Information Technology, Neelian University, Khartoum, Sudan

<sup>3</sup>Department of Computer Science, Najran University, USA

**Article Type:** Conference Proceedings Article

**Article Citation:** Nuha Elamin, Samani A.Talab, Ahmed Khalid. Sentiment analysis with supervised learning techniques: a survey. *Indian Journal of Science and Technology*. 2020; 13(03), 249-268. DOI: 10.17485/ijst/2020/v013i03/148900

**Received date:** November 14, 2019

**Accepted date:** December 28, 2019

**\*Author for correspondence:**

Nuha Elamin  neeahmed@nu.edu.sa  Faculty of Graduate Studies, Neelian University, Khartoum, Sudan

## Abstract

**Objectives:** This study aims two main goals; one is to provide complete notions relevant to sentiment analysis by SA mechanisms, its categorization, and its techniques. The second goal is to make a comprehensive study of supervised learning techniques used in SA classification to summarize the different works conducted in this area and track the recent developments. **Methods:** To achieve the first goal, several important survey studies, including modern and relevant works presented would be analyzed for full concepts around SA. As for the second objective of the study, the most important reports would be investigated, analyzed, and compared in the use of supervised learning techniques in SA from the previous to the recent researches till 2019. **Findings:** This study also made a comprehensive research of the supervised machine learning classifiers used in SA, its recent techniques and enhancement methods and the suggestion future works. There are still some open challenges in this area such as mining the complex reviews and implicit aspect identification. The sentiment language is also a challenge; thus, addressing each language according to its attributes is a difficult task and so the sentiment domain issue. **Application/improvements:** The information provided is used in assessing opinions and analyzing sentiment that could be used by researchers and institutions, and to identify different trends besides recommending the future research directions.

**Keywords:** Sentiment Analysis (SA), Opinion Mining (OM), Machine Learning, Supervised Learning, Sentiment Classification.

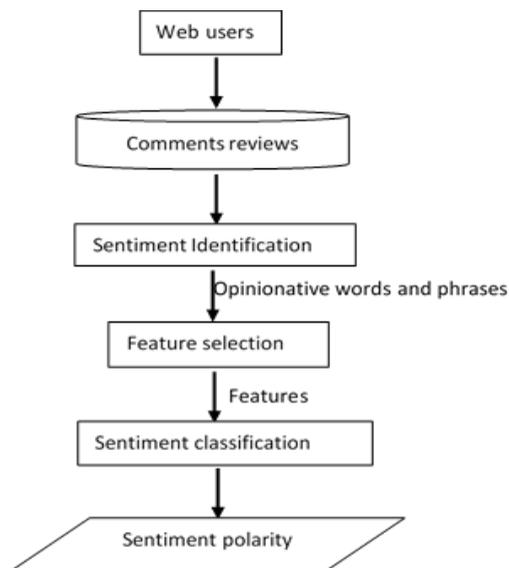
## 1. Introduction

Generally, textual information is available in two categories, target data that just contains realities, target articulation about elements or occasions and subjective information that

shows a person's sentiment or opinion about an event or object [1]. People's opinions, to a considerable degree, show their perception of reality and belief. Conclusions and ideas identified with estimations, frames of mind and feelings have become the subject of investigation of slant examination and sentiment mining [2]. Sentiment analysis applied for subjective statements. The rapidly rising popularity of various kinds of private publishing services indicates that sentiment documents will become a significant feature of the web's textual information [3]. Before sentiment, analysis discussed; the definition of "sentiment" first needs to agree on. We can represent the term sentiment as "an individual impression or private state associated with a specific goal". This perspective based on experiences, observations, concepts, and beliefs. Sentiment articulation contains every one of the individuals' administrations, assessments, examinations, mentalities, and feelings about substances. The world has recognized the important role that opinionated social media posts have played in the restructuring of companies in recent years. The public sentiment and opinions have a deep impact on our community and party-political system [4]. For instance, placements have organized the masses to make enormous party-political changes, as has occurred in some Arabic republics in 2011 and Trump's election victory in 2016. Prior to the web, assessments got physically; if an individual was to choose, he collected related opinions by asking other people around him. To obtain public opinion, companies mostly conducted surveys through focus groups and hiring external consultants; however, these methods were costly and arduous. Currently, World Wide Web (www) furnishes these data easily and almost no expense. The Internet has become an abundance of conclusions with a huge number of sites, gatherings, and tweets, social sites offering a colossal volume of refreshed data yet this web information is for the most part unstructured content that cannot be legitimately utilized for information portrayal, hence an automated opinion summarization model is needed to perform these tasks. The field breaking down a sentence of content to decode its estimations is called Sentiment examination (SA) or opinion mining (OM) [5–6], "OM is an ongoing control at the junction of data recovery, content mining and computational etymology which attempts to recognize the conclusions communicated in common language writings" [7]. The terms slant investigation and feeling mining, for the most part, utilized reciprocally in scholarly zones. Liu states that the term "sentiment analysis" investigation originally showed up in the study [8], and "opinion mining" term first appear in research [9]. Estimation investigation field characterized as a mix of data recovery, AI, natural language processing (NLP), and computational phonetics. As it were, it normally attempts to remove and assess individual's opinions from their composition. In writing, assessment examination has numerous names and somewhat various errands, for example, subjectivity investigation, conclusion mining, audit mining, feeling extraction, slant mining, evaluation extraction, influence examination, and feeling.

Subjectivity arrangement implies undertakings that order a book (terms, sentences, or report) as abstract, containing conclusion words that help decide the supposition about the substance, or target terms, the subsequent emotional sentences (opinioned term), additionally delegated communicating positive or negative sentiments, which is called extremity grouping [10].

Sentiments analyses focus on three fundamental segments, the conclusion holder or wellspring of the sentiment, the article about which the supposition is communicated and the assessment articulation [11]. The term object O, sometimes called entity, represents the intended target of the opinion expression. It is related with a couple,  $O(T, A)$ , where T is a chain of importance of segments (or parts), sub-segments, etc., and A will be a set of characteristics of O. Every part or subcomponent likewise has its own arrangement of characteristics. Liu characterizes a supposition as a quintuple  $(e_i, a_{ij}, o_{ijkl}, h_k, t_l)$ , where  $e_i$  is the name of a substance,  $a_{ij}$  is a part of  $e_i$ ,  $o_{ijkl}$  is the sentiment on the perspective  $a_{ij}$  of element  $e_i$ ,  $h_k$  is the feeling holder and the time, when the conclusion is communicated. A substance is the objective object of a supposition; it is an item, administration, subject, individual, or occasion. The perspectives speak to parts or traits of a substance (some portion of-connection). The assessment is sure, negative, unbiased, or communicated with force levels. The records i, j, k, l show that the things in the definition must compare to each other. Sentiment study considered a classification procedure as exemplified in Figure 1. It has been examined mostly at three levels; first, known as the document level which aims to find the author's general sentiment (a positive or negative) in an opinionated document, document level assumes that each document associates with a single object and opinions from a single holder [12]. The second, a sentence level, is at times a solitary archive containing various feelings even about similar substances. The assignment at this level goes to the sentences and decides if each sentence communicated a positive, negative, or nonpartisan sentiment, however, before investigating the extremity of the sentences, we should decide whether the sentences are abstract or objective and just take the emotional sentences. The two past levels are appropriate when either the entire archive or every individual sentence alludes to a solitary element. In any case, as a rule, individuals talk about substances that have numerous perspectives (characteristics) and they have an



**FIGURE 1.** Sentiment analysis process.

alternate conclusion about every one of the angles. This regularly occurs in audits about items or in discourse discussions committed to explicit item classes (for example, autos, cameras, cell phones, and even pharmaceutical medications), so this level spotlights on their insight of all estimation articulations inside a given record and the angles to which they allude [13]. The angle level prior called the element level [14]. It will probably find opinions on substances as well as their perspectives. For instance, the sentence “this camera can take high-quality photos, but its battery life is short.” Here, the sentiment evaluates two aspects of the target or entity Camera is positive, the camera quality and the other aspect is negative, the battery life of camera. In the aspect level sentiment classification, from the already extracted aspects, opinion is determined [15]. Besides, the feeling of the content can be expressed or understood. On the off chance that express, a book legitimately gives a feeling, for example, (It is a great car) while if certain; the content infers a sort of opinion like the charger labors for multi week.

Data classification accomplish in two stages, one is the learning process where the training data are analyzed by a classification technique (a classification model is learned) and the other stage is a classification process where the test data are used to predict class labels for the given data (Figure 2). Since the class name of each preparation tuple is given, this progression is otherwise called directed learning (i.e., the learning of the classifier is “managed” in that it is advised to which class each preparation tuple has a place). It appears differently in relation to solo learning (or grouping), in which the class mark of each preparation tuple not known, and the number or set of classes to learn may not know ahead of time.

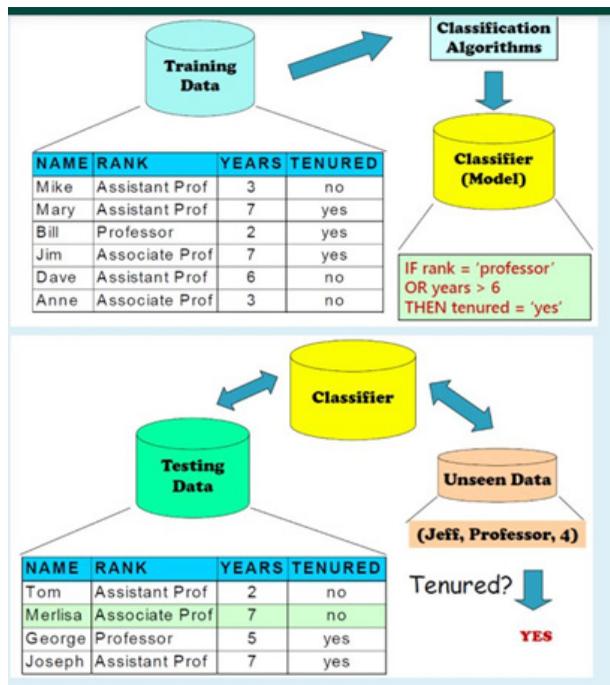


FIGURE 2. General approach for building the classification model.

Sentiment classification applied in various spaces. The most well-known spaces are motion picture surveys and client audits in a market area. Much inquire about has done in these regions [16]. News is another area examined by analysts [17–18]. The kind of information utilized in conclusion characterization contrasts starting with one space then onto the next, just as from language to language. At the end of the day, an estimation investigation framework that functions admirably for film surveys may not fill in too for client audits [19]. This issue originates from the assorted variety of the estimation starting with one space then onto the next. Subsequently, slant characterization is a very space explicit issue [18]. In [21,22] current years, a big number of surveys have conducted on SA and its related task. Most of these studies presented topics on sentiment analysis such as web data extraction and analysis, polarity and subjectivity, feature selection [23–24], sentiment analysis of comparative sentences, opinion search and retrieval exploration techniques, different classification methods and machine learning techniques for various emotion analysis tasks and finally, fake reviews or spam detection [25–28]. Schouten and Frasinca conducted comprehensive survey relevant to aspect-level sentiments [29], a survey on methods for selecting features and classifying sentiments and it outlined the depiction of the component choice techniques and posted a depth discussion on strategies of classification and related articles [30]. Fifty-four papers summarized citation the function performed, the area, the algorithm used, the polarity, the data range, the data source, and the type of language [31]. The researchers' major concern is to analyze the methods used in the articles surveyed conducted a survey of multiple emotion analysis aspects for the period 2002–2014; he noted some intelligent technologies such as random forests, evolutionary calculation, association mining, mysterious rule-based schemes, miners, and Conditional Random Field (CRF) theory. The proper idea analysis, neural network of radial foundation functions (RBFNN), and online learning algorithms have not optimized in SA. Furthermore, logic, online learning algorithms and ontology can all be very useful, especially in large data cases [32]. Ayyoub et al. provide an overview of the research on Arabic SA (ASA) so far. The research groups have published papers focused on SA-related issues they discuss and are attempting to identify the gaps for future studies in this area [33]. Research on sentiment analysis relies on the six main topics. 1- The problem of sentiment study, formalized by introducing the basic definition, concepts, and issues. 2- Sentiment and subjectivity classification, which regards supposition examination as content arrangement issue, in subjects; two sub-themes that have widely contemplated subjectivity and polarity. 3- Aspect-based sentiment analysis; this model looks for opinion targets (objects) and their components (attributes and features). 4- Sentiment analysis of near sentences. Here the assessment of an item can done in two principle ways, an immediate supposition that gives a sentiment about the article without referencing some other comparable objects (e.g., the resolution of this screen is good) and comparison which compares the objects with other similar objects (e.g., the resolution of this screen is better than that of screen-x). 5- Feeling search and recovery, conclusion search is in this way a mix of data recovery and notion investigation. 6- Opinion spam, which alludes to the counterfeit feeling that attempts to delude peruses or the robotized framework by giving under serving sentiments to some objective items to advance the articles or harm their notoriety.

The contribution of this survey offers a detailed categorization according to the methods used in a significant number of recent works. This methodology can permit scientists who know about specific systems to utilize them in the field of SA and to pick the fitting procedure for a specific application. In addition, these studies provided an overview of the importance of feature selection in refining the presentation of sentiment classification algorithms, particularly in machine-learning techniques.

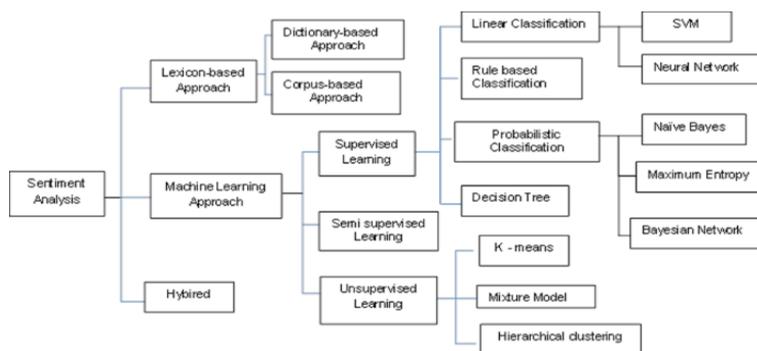
This review is structured as follows: Section 2 covers sentiment analysis approaches, defines the techniques of supervised machine learning. Section 3 describes current research gaps and key challenges of sentiment classification. Finally, the conclusion and future trend in research tackled in Section 4.

## 2. Sentiment Analysis Approaches

The approaches that manipulate SA can be categorized into three approaches; machine learning, which contains supervised and unsupervised learning, lexicon-based, which depends on the discovery and the opinion lexicon. There are two techniques in this methodology [34]. Word reference based and corpus-based methodology, in certain conditions the half breed approach, which consolidates AI with the vocabulary-based methodology and increases a moderately better presentation [35]. Figure 3 illustrates the classification approaches. The lexicon-based methodology begins with a little arrangement of seed supposition words reasonable for the current space. This arrangement of words at that point extended through the use, afterward looks the lexicon for their equivalent words and antonyms [36] whereby the corpus-based methodology starts with a seed rundown of feeling words, and afterward finds other conclusion words in a huge corpus to help discover sentiment words with setting explicit directions. This could do by utilizing factual or semantic strategies.

### 2.1. Lexicon-Based Approach

Sentiment lexicon contains arrangements of words and expressions used to express people's abstract emotions and conclusions. Vocabulary based Sentiment Analysis procedures are



**FIGURE 3.** Sentiment analysis approaches.

unaided learning in light of the fact that the order of information does not require earlier preparing. Approaches based on the lexicon use a lexicon to describe opinions by counting and measuring words related to sentiments. There are two main approaches used. One of them is the lexicon put together an approach that depends with respect to social affair an underlying arrangement of expressions of supposition and afterward looking in the word reference for their equivalent words and antonyms to grow this set. The other is the corpus-based methodology, which utilizes many expressions of supposition with known extremity, and afterward recognizes different expressions of opinion in a huge corpus to group directions.

### **2.1.1. Dictionary-Based Approach**

This technique starts by using a minor usual set of seed view words and web dictionaries, for example, WordNet. The technique first collects a minor set of opinion terms manually with established alignments and then increases it by looking for their substitutes and antonyms in the WordNet or thesaurus; the novel words added to the list of seeds until there are no more new words. Manual inspection can performed after the process has completed to remove or correct the errors. A significant shortcoming of the dictionary-based method and the view terms derived from it is it does not find words of opinion from different fields and meaning orientations, which is quite usual. For example, silent, it is typically negative for a speakerphone. However, quiet is good for a vehicle. This problem can address by the corpus-based approach.

### **2.1.2. Corpus-Based Approach**

The corpus-put together approach based with respect to the likelihood that a supposition word may happen related to a positive or negative arrangement of words via looking for enormous marked preparing information. Unlike the dictionary approach, the Corpus approach can help find domain-specific words of opinion and their orientations.

## **2.2. Hybrid Approach**

The hybrid method combines machine knowledge and lexicon-based approaches to improve performance in sentiment classification. Its main advantage is obtaining the best of both approaches, high precision from an efficient supervised learning algorithm, and lexicon-based method stability.

## **2.3. Machine Learning-Based Approach**

The machine learning method in assumption arrangement depends on the use of celebrated AI strategies on the content information to utilize the experience to make a calculation to improve the exhibition of the framework. It is functional as it completely programmed and can deal with huge assortments of information. AI-based assumption arrangement can be isolated into three primary classes: regulated, solo and semi-directed learning techniques (Figure 2).

### 2.3.1. Unsupervised Learning

Unsupervised learning is close to learning by observation; it conducts clustering. Unsupervised learning applied when there is just input information and no comparing yield factors are accessible. Its point is to show the hidden structure or appropriation in the information to become familiar with the information. Unsupervised learning issues can also group into clustering and association issues. There are some unsupervised learning algorithms commonly used in sentiment analysis (see Figure 2), such as k-means for clustering algorithms, mixture model, and hierarchical clustering.

### 2.3.2. Semi-Supervised Learning

Semi-supervised learning considers the classification problem when there are fewer corresponding labels only for a small subset of observations. These issues are of significant practical interest in a wide range of applications where unlabeled data are accessible, but it is expensive or impossible to obtain class labels for the entire data set. The key idea behindhand semi-supervised approach is unlabeled data holds a lot of class information, but it contains information on joint distribution over classification features.

### 2.3.4. Supervised Learning

Supervised learning based on the labeled dataset, so during the process the labels given to the framework. These labeled datasets trained when experienced during decision-making to produce appropriate outputs. Supervised learning classification based on the four-categories: linear classification, rule-based classification, probabilistic classification and decision tree concepts. To solve the classification problem, ML includes two stages; initial, pre-named preparing corpora utilized to gain proficiency with a “classifier” model utilizing a built up managed learning system and afterward once a classifier has been created, it tends to be applied to characterize the inconspicuous information [37]. So, one of ML's main issues is to fit a good generalization capability model to a set of training data. Traditionally, over-fitting mentions to a model that fits the training data too well but generalize poor to testing data, while under fitting refers to a model that can neither fits the training data nor generalize to testing data. A few machine learning methods have received to group the surveys in conclusion investigation. Support Vector Machine (SVM), Naïve Bayes (NB), Maximum Entropy (ME), Artificial Neural Network (ANN), and Decision Tree (DT) classifiers. Some different less generally utilized calculations are LR, K Nearest Neighbor (KNN), RF, and Bayesian Network (BN). Other less commonly used algorithms are LR, K Nearest Neighbor (KNN), RF, and Bayesian Network (BN). This article contributes to a deep understanding of machine learning in sentiment analysis especially, supervised machine learning. To accomplish this, the rest of the article offers more details about some important algorithms supervised learning approach uses in sentiment analysis.

#### 2.3.4.1. Naïve Bayes (NB)

The Naïve Bayes model is the most direct and least requesting to gather a classifier for the substance course of action framework; it subject to Bayson's theory with a supposition of

self-governance among the markers. In [38] direct terms, a Naïve Bayes classifier expects that the proximity of a particular component in a class is immaterial to the closeness of some other segment. The effortlessness of this presumption makes the calculation of Naïve Bayes classifier unquestionably increasingly productive. Bayesian classifiers have additionally shown high exactness and speed when applied to huge databases. To order as the most likely class  $c^*$  for another component  $x$ , it figures:

$$c^* = \operatorname{argmax}_c P(c|x)$$

As indicated by Bayes' hypothesis, the likelihood that we need to process  $P(c|x)$  can be communicated regarding probabilities  $P(c)$ ,  $P(x|c)$ , and  $P(x)$  as the accompanying condition:

$$P(c|x) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c)$$

Above,

- $P(c|x)$  is the back probability of class ( $c$ , target) given indicator ( $x$ , highlights).
- $P(c)$  is the earlier probability of class.
- $P(x|c)$  is the probability which is the probability of indicator given class.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

- $P(x)$  is the earlier probability of indicator

$$P_{NB}(c|x) = \frac{P(c) \prod_{i=1}^m P(f_i|c)^{n_i(x)}}{P(x)}$$

$P(x)$  assumes no job in choosing  $c^*$ . To evaluate the term  $P(x|c)$ , Naïve Bayes disintegrates it by accepting the contingent autonomy of highlights  $f_i$ 's given  $x$ 's class.

Pang and Lee [39] improve the accuracy of the Naïve Bayes classifier from 82.8 to 86.4 percent by combining it with the recognition of subjectivity. In Ref. [40], Narayanan et al. made study that suggested Adapted Naïve Bayes (ANB) to obtain information from new domain in order to solve the domain problem in SA, the study getting the farthest use of old domain and the unlabeled new domain data. Test results show that the method suggested enhances the base classifier. In Ref. [41], Dhande and Patnaik published a study showing how the suitable feature selection increase the precision of the Naïve Bayes classifier, they obtained a precision of 88.80%, but only examined a single ranker, mutual information. In Ref. [42], Bilal et al. conducted a study that combined both of NB and NN classifier for sentiment classification, experimental results show that the accuracy of sentiment analysis increased to 80.65%. A proposed model [43] applied to Roman Urdu sentiment. The training dataset used in this model contains 150 positive Roman-Urdu opinions and 150 negative that extracted from a Blogs. The study used three classification

methods NB, DT, and KNN, results show that NB classifier is the best, which compared to DT and KNN in terms of accuracy, precision, and value of F-measure.

#### 2.3.4.2. Support Vector Machine (SVM)

The support vector machine (SVM) is a factual order technique. It is dependent on the basic hazard minimization guideline from the computational learning hypothesis. SVM is a discriminative classifier; it can build a direct or non-straight choice surface to isolate the preparation information that focuses on two classes [44–45]. The key in such classifiers is to decide the ideal limits between the various classes and use them for the motivations behind the arrangement. An isolating hyperplane composed as:

$$W^*X + b = 0,$$

where  $W = \{w_1, w_2, w_3, \dots, w_n\}$ .  $w_n$  is well-defined as mass vector of  $n$  characteristics.  $b$  is well-defined as bias.

The good ways from the isolating hyperplane to any point on  $H_1$  is  $1/|W|$  and the equivalent to any point on  $H_2$

is  $1/|W|$ . In this way, the extreme edge is  $2/|W|$ . On the off chance that the hyperplane esteem  $> 0$ , at that point +ve class, if the hyperplane esteem  $< 0$ , at that point -ve classification; in the event that hyperplane esteem = 0, at that point all focuses are opposite to  $W$ . In the event that the estimation of the edge is huge, at that point an enormous punishment allotted to mistakes/edge blunders. On the off chance that the estimation of the edge is little, at that point, a few points become edge mistake and the direction of hyperplane is changed

$$W = \sum_j \alpha_j c_j d_j, \alpha_j \geq 0$$

Let  $c(1,-1)$  is class (positive, negative) for article  $d$ .

In Ref. [46], Alotaibi addressed the Political tendency classification of Twitter users by used WEKA5 exactly an SVM-based approach. In Ref. [47], Tan and Zhang created Arabic Corpus with three polarities (positive, negative, and natural) consist of 6267 documents and 33,870 sentences to use in classification problem. Different ML classifiers investigated during this task including Multinomial Naïve Bayes (MNB), Support Vector Machine (MNB) with linear kernel, and Neural Networks (NNs). SVM achieved the best results for all classification types. In Ref. [48], Manek et al. used MI, IG, CHI, and DF feature selection with a set of classification methods SVM, NB, K-nearest neighbor, winnow classifier and winnow classifier to progress a study of sentiment classification on Chinese documents. The Chinese sentiment dataset used in this study consists of 1021 documents. The result is that SVM produces the best performance for the others four classification methods. In Ref. [49], Hutto and Gilbert conducted a study based on using a suitable feature selection Gini Index to enhance the SVM classifier, the experiments emphasis that the method increases the SVM accuracy. In [50] present study that exercise TF-IDF, TF-CHI, TF-RE, and TF-OR feature selection with  $n$ -gram tokenization to improve SVM classifier, among all feature selection used, the results show that TF-IDF has the highest performance. In study [34], the

authors researched the sentiment of Movie reviews; they integrated various preprocessing strategies such as stop words erase, negation treatment and stemming, the feature election methods (FF, TF-IDF, and FP) are used to calculate three different feature matrices and then they used chi-squared technique to filter the unimportant features. Lastly, they applied SVM for classified and experimental results show that a good preprocessing leads to enhancement in classification methods.

#### 2.3.4.3. Rule-Based Classifier

In a standard-based classifier, a lot of Conditional “if ... at that point ...” style rules is normally built to decide a specific blend of examples that are in all likelihood identified with the various classes. Each standard comprises of two sections: the forerunner part and the ensuing part. The predecessor part compares to a word design and the ensuing part to a class name. Rule-based classifiers provide an advantage, as they are easily understandable by non-experts in the case of decision tree classifiers and that explanations can create easily. Arrangement controls additionally speak to each class by disjunctive ordinary structure (DNF). A k-DNF articulation is of the structure:  $(X_1 \wedge X_2 \wedge \dots \wedge X_n) \vee (X_{n+1} \wedge X_{n+2} \wedge \dots \wedge X_{2n}) \vee \dots \vee (X_{(m-1)n+1} \wedge X_{(m-1)n+2} \wedge \dots \wedge X_{mn})$ , where m is the quantity of disjunctions, n is the quantity of conjunctions in every disjunction, and  $X_n$  is characterized over the letters in order  $X_1, X_2, \dots, X_j \cup \sim X_1, \sim X_2, \dots, \sim X_j$ . The standard-based classifier intends to manufacture the littlest guideline set that is advantageous with the preparation information [51–52]. Adding extra systems to avoid over-fitting of the preparation information improved standard based arrangement. The study [53] present VADER model, a basic rule-based model for general opinion analysis, and compare its efficacy with Eleven benchmarks, including LIWC, ANEW, General Inquirer, SentiWordNet, Naïve Bayes, Maximum Entropy, and Support Vector Machine (SVM) algorithms, the results find that VADER outperforms individual human raters, F1 Classification Accuracy = 0.96 and 0.84, respectively [54] they combines the rule-based classifier and the SVM classifier To enhance the performance of SVM classifier. They used the rule-based classifier to verify the “neutral” SVM forecast, so the applied rule-based classifier for each “neutral” extracted from the of SVM classifier, although they do not obtain the best accurate, the results show that a rule-based classification can actually debug the SVM's predictions. Provided a study [55] on the use of rule-based machine learning to applied sentiment analysis on online books dataset and political reviews, they use SentiWordNet to create seven classes from strong positive to strong negative, results show this model achieved 97.4 % accuracy and minimize error average.

#### 2.3.4.4. Decision Tree (DT)

The choice tree groups the preparation information by arranging tests in the informational index contingent upon highlights. DT classifier sorted out decay progressively of the preparation information utilizing highlights. The idea is to follow the edges of a tree, starting from the root, where each non-leaf node represents a test, for example, the value

of a feature. Depending on the result of the test, one of the child nodes chosen as the next node. Nodes visited this way until a leaf node reached, which tells what class the instance should be classified as. The decision tree is useful because of its simplicity, but it is also quite powerful since the different sub-trees spanned by the children can differ depending on what tests are considered the best with the knowledge acquired on the way to the node. Despite these features, decision trees tend to have issues handling linear relationships between variables [56]. The interaction effects of variables have also problems with logistical regression. For example, the FICUS construction algorithm presented by research [57], this algorithm gets as information a set of ordered items, a set of traits, and a detail for a set of constructor capacities to deliver a set of produced highlights that can be utilized by standard idea students to make improved classifiers. ID3 is a one-approach assignment to make all conceivable choice trees that appropriately arrange the preparation set and to choose the least difficult of them. The system of this methodology repeated, it haphazardly picks a sub-set of preparing information called the window to shape a choice tree, which splendidly characterizes every one of the items in the window. In the event that the tree offers the right response for every one of these articles, at that point it is appropriate for finishing the preparation information, and the iterative closures. If not, a determination of the inaccurately arranged items added to the window and the activity proceeds [58]. The DT technique [59] used to detect the scope of the negation by dynamic determinants, which use contextual information, and static determinants, which are unambiguous words. These determinants are useful in complex sentences such that only the sentences that contain negation are considered. Experimental rules concentrate on cases where polar statements precede grammatical forms that specifically precede negative words, leading to polar expression being unique in it applied the DT technique with some other machine learning techniques [54] to classify the opinions of patients in one of the English hospitals through a point-based evaluation scale. The purpose of this analysis is to forecast the satisfaction of the patient with the hospital in terms of cleanliness and respect.

#### *2.3.4.5. A Neural Network (NN)*

NN content classifier is a system of units, where the info units speak to terms, the yield unit(s) speak to the classification or classes of intrigue and the loads on the edges associating units speak to reliance relations. For characterizing a test record  $d_j$ , its term loads  $w_{kj}$  are stacked into the information units; the enactment of these units engendered forward through the system, and the estimation of the yield unit(s) decides the classification decision(s). A study [60] presented that used the deep convolutional NN in order to extract information at the character level and then the sentence to classify short sentences of sentiments. This approach applied on two datasets and the approach demonstrated increased classification accuracy in each one based on neural networks, where they analyzed the reviews of Arab hotels using long short-term memory LSTM with aspect opinion target expressions OTEs. Results show that both applications progressed enhancement of 39 percent for the extraction of aspect-OTEs and 6 percent for the polarity classification task of aspect sentiment.

### 2.3.4.6. The Maximum Entropy Classifier

The most extreme entropy classifier applies the popular MaxEnt standard to parameter estimation. The fundamental thought is that the classifier changes over marked capabilities to vectors utilizing an encoding. This encoded vector at that point used to compute loads for each component that can then joined to decide the in all probability mark for a list of capabilities. The Max Entropy classifier can utilize to explain a huge assortment of content arrangement issues, for example, language identification, point grouping, assessment examination, and more [61]. Malouf depicted the analyses looking at the exhibition of various calculations for evaluating the parameters of a contingent ME mode [62]. The outcome shows that the standardly utilized iterative scaling calculations perform very pitifully in contrast with the others; besides, the limited memory variable metric algorithm [63] beats different calculations by a significant edge. ME classifier gauge of  $P(c | d)$  accepts the exponential structure as:

$$P_{ME}(c|d) = \frac{1}{Z(d)} \exp\left(\sum_i \lambda_i F_{i,c}(d, c)\right)$$

where,  $P_{ME}(c|d)$  is the probability of example  $d$  existence in class  $c$ ,

$\lambda_i$  denotes a feature weight parameter,

$Z(d)$  is a standardization function,

$F_{i,c}$  denotes a feature/class function for an extracted feature  $f_i$  and class  $c$ , as given in eq:

$$F_{i,c}(d, c') = \begin{cases} 1, n_i(d) > 0 \text{ and } c' = c \\ 0 \text{ otherwise} \end{cases}$$

Unlike Naïve Bayes, the ME classifier presents the best performance because it varieties no expectations about the associations between the features. A study [46] presented that uses the maximum entropy classifier for political tendency classification based on the Spanish Twitter data set this model achieved good results in the experimental test. In another study [64], the authors use a set of maximum entropy classifiers to classified the polarity for Spanish Twitter data, If all the classifiers decided on a category, the total value for the corresponding score was allocated, otherwise, the positive, negative, and objective score values are corresponding to the number of classifiers assigning the word to each category

### 2.3.4.7. Bayesian Network (BN)

The Bayesian Network (BN) assumes, unlike Naïve Bayes, that all features are completely dependent. A BN-coordinated non-cyclic chart whose hubs speak to irregular, every hub in the diagram speaks to an arbitrary variable, while the edges between the hubs speak to probabilistic conditions among the comparing irregular factors. In another depiction, BN speaks to a joint multivariate likelihood conveyance for a lot of irregular factors. Give us a chance to have a progression of sentences  $s(1), s(2), \dots, s(T)$ ; each speaks to a progression

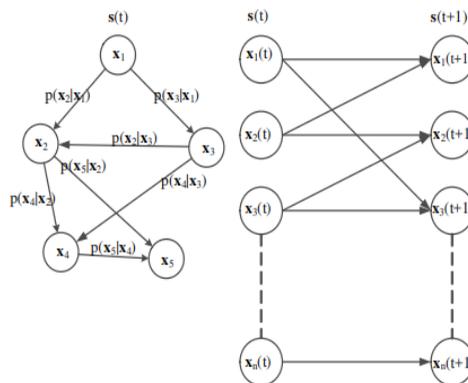
of words so that  $s(t) = (x_1(t), x_2(t), \dots, x_L(t))$ , where  $L$  is the length of sentence  $s(t)$ . Along these lines, the likelihood of a word  $p(x_i(t))$  pursues the dissemination :

$$p(x_i(t)) = P(x_i(t) | (x_1(t), x_2(t), (1) \dots, x_{i-1}(t)), (s(1), s(2), \dots, s(t-1)))$$

The BN dissects the likelihood of hub articulations into a result of restrictive probabilities by accepting the freedom of the non-relative hubs, given their folks.

$$P(X|S, \theta) = \prod_{i=1}^N p(x_i | a_i, \theta_{i,a_i})$$

where  $p(x_i | a_i, \theta_{i,a_i})$  means the restrictive likelihood of hub articulation  $x_i$  given its parent hub articulations  $a_i$ , and  $\theta_{i,a_i}$  signifies the most extreme likelihood (ML) gauge of the contingent probabilities. Figure 4 delineates the state space of a Gaussian Bayesian system (GBN) at time moment  $t$  where every hub  $x_i(t)$  is a word in the sentence  $s(t)$ . For more information review. BN occasionally used in text mining because of its computational complexity and high cost. To research a genuine issue wherein the creator’s mentality portrayed by three extraordinary (yet related) target factors. This instrument can aggregate the diverse objective factors in a similar grouping errand to profit by the conceivable measurable relations between them. Experimental results show that this approach outperforms the most common Sentiment Analysis approaches and is useful for improving the identification rates for this problem, in addition, author claim that this methodology could considered for solving future Sentiment Analysis problems. BN additionally utilized by [65] to propose a Bayesian profound convolutional conviction arrange BCDBN to Subjectivity by utilizing Bayesian systems to separate high ML ideas and word themes from the information and use them to pre-train the model grouping. This new approach accomplished very nearly 5–10% improvement in expectation precision contrasted with past approaches and it was multiple times quicker [66,67]. Bayesian networks and fuzzy recurrent NN used to enhance the extreme learning machine ELM for subjectivity detection, as the advantages of both networks used in the improvement of the ELM traditional. The results demonstrated the ability of the proposed system to detect subjectivity.



**FIGURE 4.** State space of deferent Bayesian models.

### 3. Challenges and Issues in SA

Following are a portion of the difficulties in the territory of assumption investigation, for example, nullification taking care of, area speculation, pronoun goals, language speculation, and world information. Opinion text could be in a different language, therefore, each language tackled according to its orientation, which is a formidable task. The direction of the view words could be dissimilar, from positive to negative, according to the situation, so an opinion word that is positive in one situation may consider negative in another. As the reviewer comments in free format, opinions may include abbreviations, symbols, and short words. To deal with it requires a lot of work to mine opinion. Most surveys have distinctive composition, positive and negative in same sentence, which is simple for people to see, however, increasingly hard for a PC [68]. The direction of feeling words could be diverse as indicated by their situation in the sentence; for instance, the descriptor “little” can be utilized in a positive or negative sense, in this way, to distinguish the extremity of a similar modifier words in various circumstances is likewise a difficult assignment [69]. Assessment spamming can even be terrifying as they can twist feelings and influence clients’ understanding. It protected to state that as suppositions progressively utilized practically speaking, sentiment spamming will turn out to be increasingly wild and refined, which exhibits a significant test for their recognition [47]. Since web clients are settling on choices as indicated by web audits, it is vital that the surveys be high caliber and dependable along these lines, OM experiences the nature of audits issues; in any case, just constrained work has led on supposition quality assurance. Additionally displaying an enormous test despite OM is the accessibility and availability of a standard dataset. Scarcely any information is right now accessible to encourage the arrangement, benchmarking and investigation of the inferred content. At last, some other composing styles, for example, incongruity, mockery, or nullified sentences could carry more difficulties to estimation examination [70].

### 4. Summary and Conclusion

This study reviewed the classification techniques using supervised learning machine and what tools are available for sentiment analysis. More specifically, we considered the trend of improving the classification algorithm by using appropriate feature selection methods. There are still some open challenges in this area such as mining the complex reviews and implicit aspect identification. The sentiment language is also a challenge; thus, addressing each language according to its attributes is a difficult task and so the sentiment domain issue. Concatenating the conceptual approaches with the power of machine learning, improving the feature selection methods and applying some of the intelligent techniques in sentiment analysis may lead to good and efficient solutions for the future work to enhance the performance of sentiment analysis for the challenges mentioned above.

### References

1. Kaur A, Duhan N. A survey on sentiment analysis and opinion mining. *International Journal of Innovations & Advancement in Computer Science*. 2015; 4, 107–116. <http://inpressco.com/wp-content/uploads/2017/03/Paper18444-446.pdf>

2. Cambria E, Schuller B, Xia Y, Havasi C. New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*. 2013; 28(2), 15–21. <https://ieeexplore.ieee.org/abstract/document/#>
3. Tsytsarau M, Palpanas T. Survey on mining subjective data on the web. *Data Mining and Knowledge Discovery*. 2012; 24(3), 478–514. <https://dl.acm.org/doi/10.1007/s10618-011-0238-6>
4. Wang H, Can D, Kazemzadeh A, Bar F, Narayanan S. A system for real-time twitter sentiment analysis of 2012 US presidential election cycle. Proceedings of the ACL 2012 system demonstrations. Association for Computational Linguistics, 2012 Jul 10, 115–120. <https://www.aclweb.org/anthology/P12-3020>
5. Riloff E, Patwardhan S, Wiebe J. Feature subsumption for opinion analysis. Proceedings of the 2006 conference on empirical methods in natural language processing. Association for Computational Linguistics, 2006, 440–448. <https://dl.acm.org/doi/10.5555/1610075.1610137>
6. Pang B, Lee L. Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*. 2008; 2(1–2), 1–35. <https://www.nowpublishers.com/article/Details/INR-011>
7. Nasukawa T, Yi J. Sentiment analysis: capturing favorability using natural language processing. Proceedings of the 2nd international conference on knowledge capture, 2003, 70–77. <https://dl.acm.org/doi/10.1145/945645.945658>
8. Dave K, Lawrence S, Pennock DM. Mining the peanut gallery: opinion extraction and semantic classification of product reviews. Proceedings of the 12th international conference on world wide web, 2003, 519–528. <https://dl.acm.org/doi/10.1145/775152.775226>
9. Joshi NS, Itkat SA. A survey on feature level sentiment analysis. *International Journal of Computer Science and Information Technologies*. 2014; 5(4), 5422–5425. <https://pdfs.semanticscholar.org/a65e/07deebf8a4e733ae5d9970d53e8555e948f0.pdf%5C>
10. Khan K, Baharudin B, Khan A, Ullah A. Mining opinion components from unstructured reviews: a review. *Journal of King Saud University-Computer and Information Sciences*. 2014, 26 (3), 258–275. <https://www.sciencedirect.com/science/article/pii/S131915781400010X>
11. Pang B, Lee L, Vaithyanathan S. Thumbs up? Sentiment classification using machine learning techniques. Proceedings of the ACL-02 conference on empirical methods in natural language processing. Association for Computational Linguistics. 2002, 79–86. <https://dl.acm.org/doi/10.3115/1118693.1118704>
12. Kolkur S, Dantal G, Mahe R. Study of different levels for sentiment analysis. *International Journal of Current Engineering and Technology*. 2015; 5(2), 768–770. <https://inpressco.com/wp-content/uploads/2015/03/Paper32768-770.pdf>
13. Vohra SM, Teraiya JB. A comparative study of sentiment analysis techniques. *Journal JIKRCE*. 2013; 2(2), 313–317. <https://pdfs.semanticscholar.org/3f10/b006bab60c7f363bc03e72ad405d264b8d42.pdf>
14. Turney PD. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. Proceedings of the 40th annual meeting on association for computational linguistics. Association for Computational Linguistics. 2002, 417–424. <https://dl.acm.org/doi/10.3115/1073083.1073153>
15. Abdul-Mageed M, Diab M, Kübler S. SAMAR: subjectivity and sentiment analysis for Arabic social media. *Computer Speech & Language*. 2014; 28(1), 20–37. <https://www.sciencedirect.com/science/article/pii/S088523081300017X>
16. Bautin M, Vijayarenu L, Skiena S. International sentiment analysis for news and blogs. In: ICWSM. 2008; 1–8.
17. Gamon M, Aue A. Automatic identification of sentiment vocabulary: exploiting low association with known sentiment terms. Proceedings of the ACL workshop on feature engineering for

- machine learning in natural language processing. Association for Computational Linguistics. 2005, 57–64. <https://dl.acm.org/doi/10.5555/1610230.1610241>
18. Dey S. Performance investigation of feature selection methods. arXiv preprint arXiv:1309.3949. 2013 Sep 16. <https://arxiv.org/abs/1309.3949>
  19. Thet TT, Na JC, Khoo CS. Aspect-based sentiment analysis of movie reviews on discussion boards. *Journal of Information Science*. 2010; 36(6), 823–848. [https://journals.sagepub.com/doi/abs/10.1177/0165551510388123?casa\\_token=Xx8cTUOAxYkAAAAA:UPFCT2D16HadJWy4f2rPYCTzhI-t22DKzk4ZmP46udEEwcAgzKklJQbdEh1qb7arJ\\_4wf44G58-1Sw](https://journals.sagepub.com/doi/abs/10.1177/0165551510388123?casa_token=Xx8cTUOAxYkAAAAA:UPFCT2D16HadJWy4f2rPYCTzhI-t22DKzk4ZmP46udEEwcAgzKklJQbdEh1qb7arJ_4wf44G58-1Sw)
  20. Duric A, Song F. Feature selection for sentiment analysis based on content and syntax models. *Decision Support Systems*. 2012; 53(4), 704–711. <https://www.sciencedirect.com/science/article/pii/S0167923612001340>
  21. O'Keefe T, Koprinska I. Feature selection and weighting methods in sentiment analysis. Proceedings of the 14th Australasian document-computing symposium, Sydney. 2009, 67–74. <https://pdfs.semanticscholar.org/4bf9/6ed718c89af435929380a1a529f6721ef2ee.pdf>
  22. Savoy OK. Feature selection in sentiment analysis. Proc. CORIA. 2012, 273–84. <http://www.asso-aria.org/coria/2012/273.pdf>
  23. Mohod SW, Dhote CA. Feature selection technique for text document classification: an alternative approach. *International Journal on Recent and Innovation Trends in Computing and Communication*. 2014; 2(9), 2914–2917. <https://pdfs.semanticscholar.org/c366/9a107cf135b10e716075a4d40772445824e9.pdf>
  24. Aydoğan E, Akcayol MA. A comprehensive survey for sentiment analysis tasks using machine learning techniques. In: 2016 international symposium on innovations in intelligent systems and applications. 2016; 1–7. <https://ieeexplore.ieee.org/abstract/document/7571856>
  25. Vasantharaj S, Martin A, Meiappane A, Madhusudnan J. A survey on sentiment analysis applied in opinion mining. *Journal of Network Communications and Emerging Technologies*. 2015; 1(1), 1–7. <http://www.jncet.org/Manuscripts/Volume-1/Issue-1/Vol-1-issue-1-M-04.pdf>
  26. Hussein DM. A survey on sentiment analysis challenges. *Journal of King Saud University-Engineering Sciences*. 2018; 30(4), 330–338. <https://www.sciencedirect.com/science/article/pii/S1018363916300071>
  27. Nikhil R, Tikoo N, Kurlle S, Pisupati HS, Prasad GR. A survey on text mining and sentiment analysis for unstructured web data. *Journal of Emerging Technologies and Innovative Research*. 2015, 2(4), 1292–1296. <https://pdfs.semanticscholar.org/acdb/d896f7bbde0858fde337df657edeba71fff2.pdf>
  28. Schouten K, Frasinca F. Survey on aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering*. 2015; 28(3), 813–830. <https://ieeexplore.ieee.org/abstract/document/7286808>
  29. Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: a survey. *Ain Shams Engineering Journal*. 2014; 5(4), 1093–1113. <https://www.sciencedirect.com/science/article/pii/S2090447914000550>
  30. Ravi K, Ravi V. A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-Based Systems*. 2015; 89, 14–46. <https://www.sciencedirect.com/science/article/abs/pii/S0950705115002336>
  31. Al-Ayyoub M, Khamaiseh AA, Jararweh Y, Al-Kabi MN. A comprehensive survey of Arabic sentiment analysis. *Information Processing & Management*. 2019; 56(2), 320–342. <https://www.sciencedirect.com/science/article/abs/pii/S0306457316306689>
  32. Petz G, Karpowicz M, Fürschuß H, Auinger A, Striteský V, Holzinger A. Reprint of: computational approaches for mining user's opinions on the web 2.0. *Information Processing*

- Management. 2015; 51(4), 510–519. <https://www.sciencedirect.com/science/article/abs/pii/S0306457315000655>
33. Aggarwal CC, Zhai C. A survey of text classification algorithms. In: Mining text data, Springer: Boston, MA. 2012; 163–222. [https://link.springer.com/chapter/10.1007/978-1-4614-3223-4\\_6](https://link.springer.com/chapter/10.1007/978-1-4614-3223-4_6)
34. Feldman R. Techniques and applications for sentiment analysis. *Communications of the ACM*. 2013; 56(4), 82–89. <https://airtonbjunior.github.io/mestrado/seminars/presentations/2/Presentation2AirtonV1.pdf>
35. Abbasi A, Chen H, Salem A. Sentiment analysis in multiple languages: feature selection for opinion classification in web forums. *ACM Transactions on Information Systems (TOIS)*. 2008; 26(3), 12. <https://dl.acm.org/doi/10.1145/1361684.1361685>
36. Vapnik V. Statistical learning theory. A Wiley-inter science publication, John Wiley & Sons, Inc. 1998, 1–740. <https://www.wiley.com/en-us/Statistical+Learning+Theory-p-9780471030034>
37. Dey Sarkar S, Goswami S, Agarwal A, Aktar J. A novel feature selection technique for text classification using naive bayes. *International Scholarly Research Notices*. 2014; 10. <https://www.hindawi.com/journals/isrn/2014/717092/abs/>
38. Pang B, Lee L. A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. Proceedings of the 42nd annual meeting on Association for Computational Linguistics. 2004 Jul 21, 271. Association for Computational Linguistics. <https://dl.acm.org/doi/10.3115/1218955.1218990>
39. Tan S, Cheng X, Wang Y, Xu H. Adapting naive bayes to domain adaptation for sentiment analysis. In: European conference on information retrieval. 2009; 337–349. [https://link.springer.com/chapter/10.1007/978-3-642-00958-7\\_31](https://link.springer.com/chapter/10.1007/978-3-642-00958-7_31)
40. Narayanan V, Arora I, Bhatia A. Fast and accurate sentiment classification using an enhanced Naive Bayes model. In: International conference on intelligent data engineering and automated learning. 2013; 194–201. [https://link.springer.com/chapter/10.1007/978-3-642-41278-3\\_24](https://link.springer.com/chapter/10.1007/978-3-642-41278-3_24)
41. Dhande LL, Patnaik GK. Analyzing sentiment of movie review data using Naive Bayes neural classifier. *International Journal of Emerging Trends & Technology in Computer Science*. 2014; 3(4), 313–320. <https://pdfs.semanticscholar.org/a030/b72c0546a3c832627855d42bcc3f6819e75.pdf>
42. Bilal M, Israr H, Shahid M, Khan A. Sentiment classification of Roman-Urdu opinions using Naïve Bayesian, decision tree and KNN classification techniques. *Journal of King Saud University-Computer and Information Sciences*. 2016; 28(3), 330–44. <https://www.sciencedirect.com/science/article/pii/S1319157815001330>
43. Vinodhini G, Chandrasekaran RM. Sentiment analysis and opinion mining: a survey. *International Journal*. 2012; 2(6), 282–292. <https://pdfs.semanticscholar.org/261e/26ae134b8f63270dbcac2d07fa700fdf593.pdf>
44. Sebastiani F. Machine learning in automated text categorization. *ACM Computing Surveys (CSUR)*. 2002; 34(1), 1–47. <https://dl.acm.org/doi/10.1145/505282.505283>
45. Pla F, Hurtado LF. Political tendency identification in twitter using sentiment analysis techniques. Proceedings of COLING 2014, the 25th international conference on computational linguistics: technical papers. 2014, 183–192. <https://www.aclweb.org/anthology/C14-1019>
46. Alotaibi SS. Sentiment analysis in the Arabic language using machine learning. Doctoral dissertation. Colorado State University. 2015, 1–203. <https://pdfs.semanticscholar.org/054e/8331d7bea8fba2916fc57d6d8781edd5a3e3.pdf>
47. Tan S, Zhang J. An empirical study of sentiment analysis for Chinese documents. *Expert Systems with Applications*. 2008; 34(4), 2622–2629. <https://www.sciencedirect.com/science/article/pii/S0957417407001534>

48. Manek AS, Shenoy PD, Mohan MC, Venugopal KR. Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier. *World Wide Web*. 2017; 20(2), 135–154. <https://link.springer.com/article/10.1007/s11280-015-0381-x>
49. Hutto CJ, Gilbert E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In: Eighth international AAAI conference on weblogs and social media. 2014. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewPaper/8109>
50. Fürnkranz J. Pruning algorithms for rule learning. *Machine Learning*. 1997; 27(2), 139–172. <https://link.springer.com/article/10.1023/A:1007329424533>
51. Quinlan JR. Induction of decision trees. *Machine Learning*. 1986; 1(1), 81–106. <http://hunch.net/~coms-4771/quinlan.pdf>
52. Chikersal P, Poria S, Cambria E. SeNTU: sentiment analysis of tweets by combining a rule-based classifier with supervised learning. Proceedings of the 9th international workshop on semantic evaluation SemEval2015, 2015, 647–651). <https://www.aclweb.org/anthology/S15-2108>
53. Jia L, Yu C, Meng W. The effect of negation on sentiment analysis and retrieval effectiveness. Proceedings of the 18th ACM conference on Information and knowledge management, 2009, 1827–1830. <https://dl.acm.org/doi/10.1145/1645953.1646241>
54. Chaturvedi I, Ragusa E, Gastaldo P, Zunino R, Cambria E. Bayesian network based extreme learning machine for subjectivity detection. *Journal of the Franklin Institute*. 2018; 355(4), 1780–1797. <https://www.sciencedirect.com/science/article/abs/pii/S0016003217303009>
55. De Caigny A, Coussement K, De Bock KW. A new hybrid classification algorithm for customer churns prediction based on logistic regression and decision trees. *European Journal of Operational Research*. 2018; 269(2), 760–772. <https://www.sciencedirect.com/science/article/abs/pii/S0377221718301243>
56. Markovitch S, Rosenstein D. Feature generation using general constructor functions. *Machine Learning*. 2002, 49(1), 59–98. <https://link.springer.com/article/10.1023/A:1014046307775>
57. Nigam K, Lafferty J, McCallum A. Using maximum entropy for text classification. In: IJCAI-99 workshop on machine learning for information filtering. 1999; 1(1), 61–67. <http://www.kamalnigam.com/papers/maxent-ijcaiws99.pdf>
58. Tepper JE, O'Connell MJ, Niedzwiecki D, Hollis D, Compton C, Benson III AB, Cummings B, Gunderson L, Macdonald JS, Mayer RJ. Impact of number of nodes retrieved on outcome in patients with rectal cancer. *Journal of Clinical Oncology*. 2001; 19(1), 157–163. <https://pdfs.semanticscholar.org/77b7/7fdef509b2c91392637fd2544013893a1325.pdf>
59. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Use of sentiment analysis for capturing patient experience from free-text comments posted online. *Journal of Medical Internet Research*. 2013; 15(11), e239. <https://www.jmir.org/2013/11/e239>
60. Chaturvedi I, Cambria E, Poria S, Bajpai R. Bayesian deep convolution belief networks for subjectivity detection. In: 2016 IEEE 16th international conference on data mining workshops. 2016; 916–923. <https://ieeexplore.ieee.org/abstract/document/7836765>
61. Dave K, Lawrence S, Pennock DM. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. Proceedings of the 12th international conference on world wide web, 2003, 519–528. <https://dl.acm.org/doi/10.1145/775152.775226>
62. Tepper JE, O'Connell MJ, Niedzwiecki D, Hollis D, Compton C, Benson III AB, Cummings B, Gunderson L, Macdonald JS, Mayer RJ. Impact of number of nodes retrieved on outcome in patients with rectal cancer. *Journal of Clinical Oncology*. 2001; 19(1), 157–163.
63. Ahmed S, Danti A. Effective sentimental analysis and opinion mining of web reviews using rule based classifiers. *Computational Intelligence in Data Mining*. 2016; 1, 171–179. [https://link.springer.com/chapter/10.1007/978-81-322-2734-2\\_18](https://link.springer.com/chapter/10.1007/978-81-322-2734-2_18)

64. Harb A, Plantié M, Dray G, Roche M, Troussel F, Poncelet P. Web Opinion Mining: how to extract opinions from blogs? Proceedings of the 5th international conference on soft computing as transdisciplinary science and technology, 2008, 211–217. <https://dl.acm.org/doi/10.1145/1456223.1456269>
65. Danyang L, Huimin F. Research and improvement of CHI feature selection in sentiment analysis. *Journal of Physics: Conference Series*. 2019; 1187(5), 052022. <https://iopscience.iop.org/article/10.1088/1742-6596/1187/5/052022/meta>
66. Ortigosa-Hernández J, Rodríguez JD, Alzate L, Lucania M, Inza I, Lozano JA. Approaching sentiment analysis by using semi-supervised learning of multi-dimensional classifiers. *Neurocomputing*. 2012; 92, 98–115. <https://www.sciencedirect.com/science/article/abs/pii/S0925231212001282>
67. Al-Smadi M, Talafha B, Al-Ayyoub M, Jararweh Y. Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews. *International Journal of Machine Learning and Cybernetics*. 2019; 10(8), 2163–2175. <https://link.springer.com/article/10.1007/s13042-018-0799-4>
68. Yao T, Li L. A kernel-based sentiment classification approach for Chinese sentences. In: 2009 WRI world congress on computer science and information engineering. 2009; 5, 513–518. <https://ieeexplore.ieee.org/abstract/document/5170588>
69. Osimo D, Mureddu F. Research challenge on opinion mining and sentiment analysis. Université de Paris-Sud, Laboratoire LIMSI-CNRS, Bâtiment. 2012, 508. <https://www.w3.org/2012/06/pmod/opinionmining.pdf>
70. M, Liu B. Mining and summarizing customer reviews. Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining, 2004, 168–177. <https://dl.acm.org/doi/10.1145/1014052.1014073>