Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 7, July 2021: 5759 - 5775

Research Article

Supervised Machine Learning Techniques for Sentiment Analysis and its Application in Image Processing and Remote Sensing

Iman M[1], Subhashish Tiwari[2], Swetha Vura[3], Ajay Sudhir Bale[3*], Bharath G[4] and Sharanabasaveshwara H B[5]

[1]Department of CSE, School of Engineering and Technology, CMR University, Bengaluru, India
[2]Department. of ECE, Vignan & Foundation for Science, Technology & Research, Guntur, India
[3]Department of ECE, School of Engineering and Technology, CMR University, Bengaluru, India
[4]Department of ME, School of Engineering and Technology, CMR University, Bengaluru, India
[5]Department of ECE, GM Institute of Technology, Davangere, India
ajaysudhirbale@gmail.com

ABSTRACT

As we see the growing social media transparency, where users want to express their opinions we also see a growing need of Sentiment analysis (SA) or Opinion mining. Sentiment Analysis is a process used to interpret user sentiments related to a specific topic. It is a field of study which is formed at the junction of Natural Language Processing (NLP), Machine learning (ML) and Information retrieval (IR). In this paper we will see various types of machine learning (ML) techniques that come under supervised learning. We will be discussing 3 techniques under the domain of Supervised ML, namely, SVM, Naive Bayes and Maximum Entropy, which are studied and established to be the most accurate. We will see how supervised ML techniques can be applied in Image processing and Remote sensing as well.

KEYWORDS: sentiment analysis, ML classifiers, Supervised ML algorithms, Image processing using Supervised ML, Remote Sensing using Supervised ML

1. INTRODUCTION

NLP is analyzing NL data and processing it with the help of ML algorithms [1]. Sentiment analysis is one of the sub discipline of NLP. SA or opinion mining is a process used to interpret the sentiments or attitude of the user giving the review. This large set of data collected of opinions, comments, reviews etc. is proven to be very useful for businesses to improve their products based on the sentiments, which in turn improves the sales [2].

There are 3 levels of categorization of SA namely Document level, Sentence level and Aspect level or Attribute level. Document level recognizes if the document expresses opinion (example: review, blogs) and if yes then what level of emotion (positive, negative or neutral). In Sentence level it recognizes if the sentence is opinionated and if yes which level of emotion does it carry (positive, negative or neutral). In the Attribute level the focus of opinion on one specific entity. A few challenges faced by SA are Subjectivity and Tone of the opinion, Irony and Sarcasm, Comparison, Defining neutral and so on.

Mainly, SA are of two types, i.e., ML based approach and lexical based approach (LBA) [3].

The process of sentiment analysis through which we can evaluate whether the given data in the text delivers a positive, negative or neutral opinion, this representation of sentiment in text or through emoticons by customers is useful in the field of business, product feedback. There are three types of approaches through which the sentiment analysis can be performed. They are:

ML approach, Lexicon based approach and Hybrid method

The types of learning methods used in Machine learning are semi supervised, supervised and unsupervised learning methods. The deep learning technique comprising of semi supervised and supervised learning gives rise to various types of neural networks like Convolutional Neural Network (CNN), Recursive Neural Network (RNN), Deep Neural Network (DNN), Recurrent Neural Network, Deep Belief Network (DBN) and Hybrid Neural Network (HNN). Using these Artificial Neural Network (NN) concepts a variety of Algorithms can be derived and implemented for Sentiment analysis [43]. Being a broad diversified learning method, supervised learning has a diversified contribution in various classifiers like deep learning technique, Decision Tree classifiers, linear classifiers, Rule - based classifiers, Probabilistic classifiers. The Support Vector Machines (SVM), that encounter the regression and classification difficulties, and the neural network are categorized under linear classifiers of supervised ML. Naives Bayes, Bayesian Network and Maximum entropy are the styles of probabilistic classifiers.

The LBA that comprises Dictionary based Approach and Corpus based Approach uses Statistical and Semantic algorithms in the analysis of social media based applications.

In this paper we will be discussing the three algorithms that are found to be the most efficient algorithms that come under supervised ML techniques.

Fig 1. Describes the basic classification of sentiment analysis algorithms using a simple informative chart. The main approaches that come under semantic analysis are ML approach and LBA, these 2 approaches when put together for more effectiveness make a hybrid approach.

The Hybrid approach deals with the collaboration of two algorithms, one from ML and the other from LBA. For example we can take SVM and SVR models, NN and EKF [4]

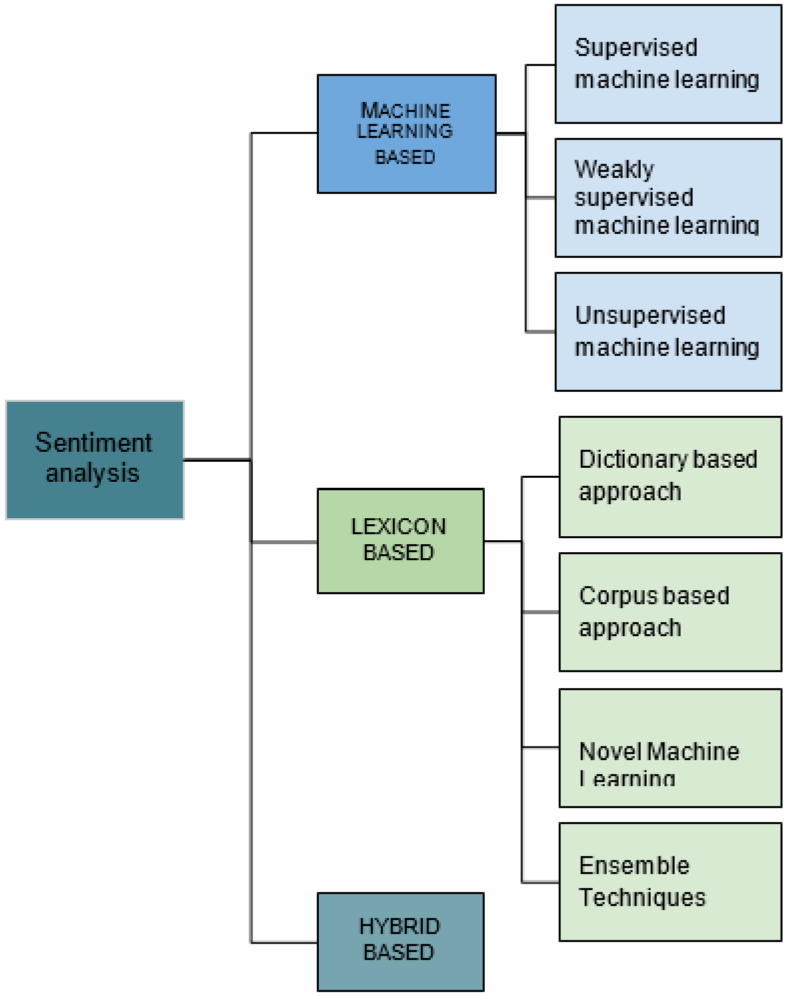

Fig 1. Basic classification of sentiment analysis

The above image is partly derived from [47]

1.1 BASIC DEFINITIONS

INTENT AND ENTITY

The intent captures the basic meaning of the utterance by the user. There can be only one main intent; for example, “I need to book a flight from LA” The intent of this statement is to book a flight.

An entity is used to give additional information about the intent. In the above statement the entity is ‘LA’. Having different entities does not change the intent of the user. Let us take another example “I want to buy a pink dress”. Here the clear intent is to buy and the entity is a pink dress.

2. MACHINE LEARNING APPROACH

ML approaches are advancing toward building algorithms that are more efficient and help the system process data. ML approaches are divided into 3 subtopics, supervised ML approach, weakly supervised ML approach and unsupervised ML approach. The ML approach treats the sentiment-classification problem as a topic-based text classification problem [5].

2.1 UNSUPERVISED MACHINE LEARNING:

Also sometimes called class discovery [6]. Unsupervised machine learning uses Artificial Intelligence algorithms to interpret data points in certain data sets that are not classified and labeled. The algorithm sorts unsorted data based on similarities and differences. Data scientist prefer using supervised and unsupervised algorithms for more

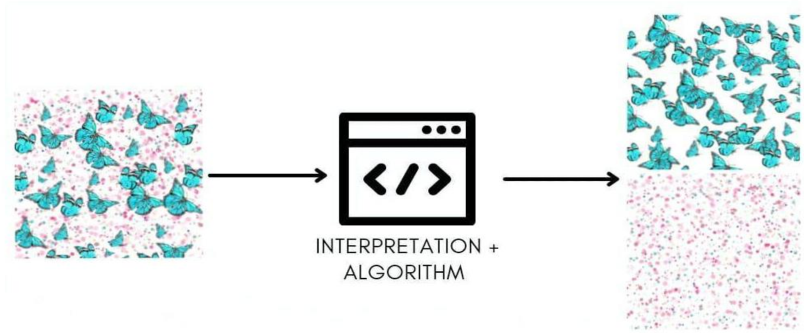

Fig 2. Unsupervised machine learning

In the figure 2 we see that the butterfly objects are learned and distinguished from the colorful objects by interpreting and using efficient algorithms.

2.2 WEAKLY SUPERVISED ML:

Weakly supervised learning is a mixture of supervised and unsupervised learning where there is a large data that is unlabeled and a small set of data that is labelled. There are 3 types of weakly supervised learning techniques:

a) Incomplete supervision: here, only a subset of training data is given with labels

b) Inexact supervision: here, the training data are given with only rough estimated labels.

c) Inaccurate supervision: the given labels here are not always correct so therefore they might have some errors due the inaccurate information [7].

2.3 SUPERVISED MACHINE LEARNING:

Supervised machine learning techniques, also known as induction classification algorithms [8] have trained data sets assigned to it containing features and labels. Every set consists of examples and every example consists of possible input and expected outputs. Supervised ML is the most common approach since the goal is to make the machine learn the data we have fed [9].

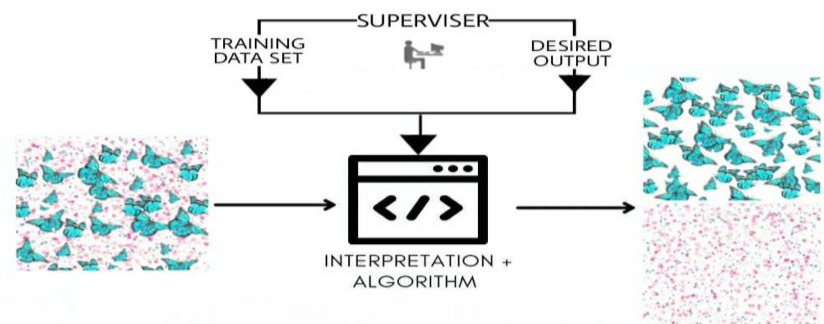

Fig 3. Supervised machine learning

From Fig 3. We can understand the importance of training data set and the role it plays in the showcasing the desired output

2.4 CLASSIFICATION TECHNIQUES OF SUPERVISED MACHINE LEARNING

According to researchers [10-11], there are multiple ways to classify supervised machine learning. Few techniques are linear classifiers, linear regression, Decision trees, neural networks, Bayesian network, Naive bayes, Maximum entropy, K means and so on

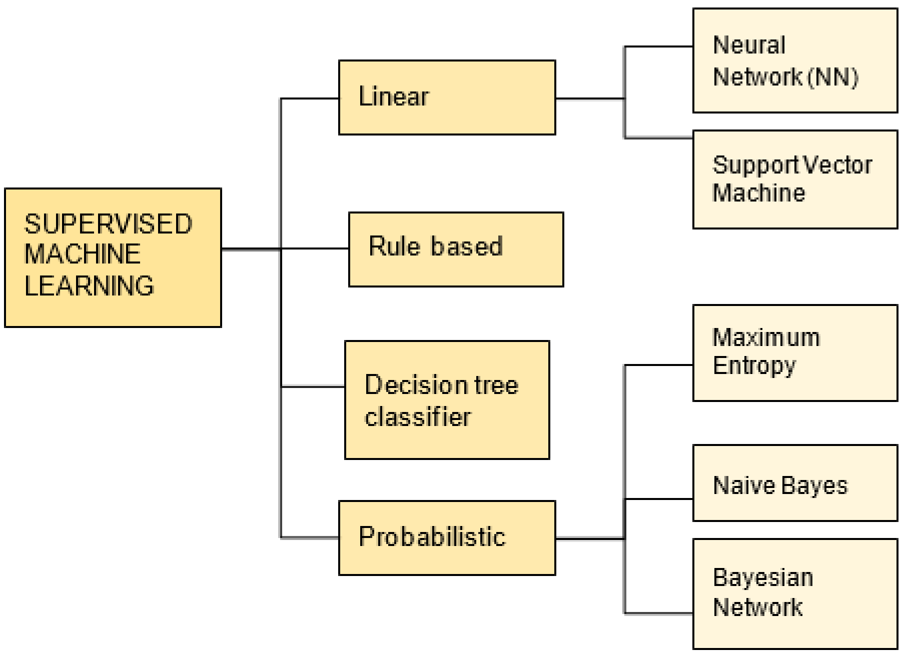

Fig 4. Basic Supervised machine learning classification

The above image is derived from [47]

From the above classification in Fig 4. We observe supervised machine learning has 4 basic classifiers. There are more algorithms that are found to be under supervised machine learning but the above mentioned are specifically found to be most useful and accurate when it comes to working for sentiment evaluation.

We will discuss only four of the most accurate supervised machine learning techniques:

2.4.1 SVM (Support Vector Machines): They are known to be the most recent algorithms and are also the most accurate algorithms when it comes to speech classification [12-13].

SVM is known to be highly effective in traditional text classification techniques [14]. It also belongs to the linear classifiers. Linear classifiers group similar feature values.

In SVM algorithm, data set is drawn on n dimensional space, n being the number of features. A hyperplane is drawn to separate two different classes. The SVM kernel is a function that helps separate the features when they can't be separated by a linear hyperplane. The support vectors are less in number in comparison to the entire data set, therefore making the algorithm good at execution time [2].

SVM framework consists of 4 stages: Dataset, Pre-processing, Classification and Results.

1. Dataset: The input data is inserted into the WEKA (Waikato Environment for Knowledge Analysis) environment [15], which is a widely used java language based software used for various applications such as visualization, predictive modelling etc. it deals with the arrangement of appropriate data and its conversion to CSV/ARFF format to use in WEKA Workbench.
2. Preprocessing: it deals with:

   Multi stop words: removal of multi stop words such as ‘A’, ‘the’, ‘is’, ‘of’ etc. may not be very useful to texts.

   Term Frequency-Inverse Document Frequency: which detects the frequency of used words. Identification of the frequency and the word might be useful in understanding the opinion in the text.

   Stemming: Is a very useful process, it removes the prefix and suffix from a word to extract the original word [16]. For example the word ‘starting’ is broken down to start, the ‘ing’ suffix is removed.
3. Classification: Here, the input data (test data) is classified on the bases of the rules that are made using the training data which is pre classified. For classification, SVM using grid search can be used for more accurate results. Overfitting problems can also be solved using K-folds cross validation.
4. Results with dataset: the below data about tweets on Apple, Google, Microsoft, Twitter is obtained from [17]. The average precision was found to be 0.745, average recall was 0.752 and average F- measure was found to be 0.747. The precision, recall and F-measure can be found with the following formulae.

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \qquad \text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})}$$

$$\text{F-measure} = \frac{\text{Precision} * \text{Recall} * 2}{(\text{Precision} + \text{Recall})}$$

Where TP - correctly classified sentences; FP- wrongly classified sentences.

Detailed results for the data set are proposed by [18] [19] as in the following table 1 and figure 5

TABLE 1. [18][19]

Class	Precision	Recall	F-Measure
Negative	0.583	0.539	0.561
Positive	0.504	0.378	0.432
Neutral	0.745	0.804	0.773
Irrelevant	0.872	0.867	0.87
Average	0.745	0.752	0.747

Graphically described data is shown as follows:

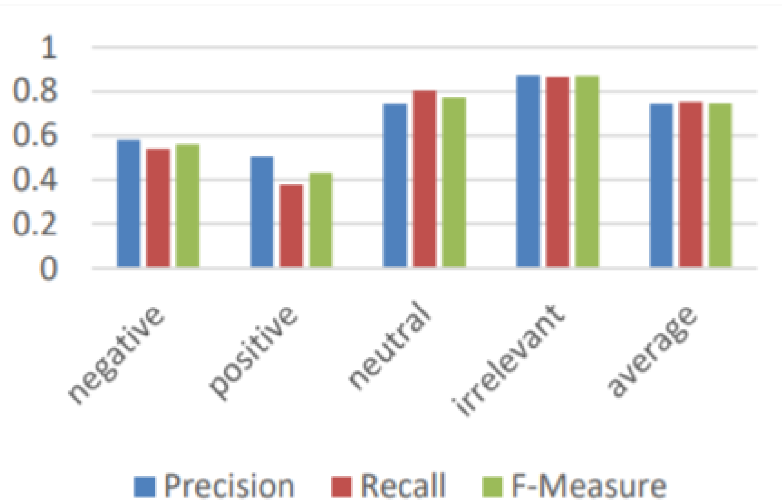

Fig 5. Results of data set in TABLE 1. [18]

Few advantages of SVM are that it is capable of handling large feature sets and is a sturdy algorithm in the presence of very few example sets [15]. In [20], SVM feature extraction is very neatly discussed in [20], the researchers have compared SVM to FDO. SVM more suitable for text categorization because of high dimensional input space, removal of few irrelevant features, sparse instances [21].

2.4.2 NAIVE-BAYES: This classification method is a statistical or probabilistic method of classification under supervised machine learning. Naive-Bayes algorithm is known for its computational efficiency for real- world problems. The word "Naive" is used since the algorithm assumes all attributes are completely independent, even though this seems less realistic, It works very fine for normal distributed data in actual world problems [22]. Basically the algorithm assumes that any feature of the class is independent of any other feature of the same class.

Here, the starting point is Bayes theorem of conditional probabilities

$$P(x|A) = (P(A|x).\ P(x))/\ P(A)$$

Here x is the data points and A is the class. This algorithm can be used for prediction solving. The lower computational power of this algorithm is one of the advantages. The independent assumptions sometimes cause inaccurate conclusions. This method assumes an underlying probabilistic scheme by finding the uncertainty of the model. For example, Trees have features such as leaves, branches and trunk and Naive Bayes algorithm considers all these features to be independently contributing to the feature of a tree.

The Naive bayes formula is as follows [23]:

$$P(Y = K|\mathrm{X}) = \frac{P(X|\mathrm{Y} = K) * P(Y = K)}{P(X)}$$

In the above equation K represents the class of Y.

From the definition of naive bayes, if we consider all features of X to be independent then the Bayes equation becomes

$$P(Y = K|X1, \ldots\ldots Xn) = \frac{P(X1|\mathrm{Y} = K) * P(X2|\mathrm{Y} = K) \ldots\ldots.* P(Xn|\mathrm{Y} = K) * P(Y = K)}{P(X1) * P(X2) \ldots.* P(Xn)}$$

The first term defines the conditional probability of all X's with respect to Y

Few advantages of Naive Bayes algorithm includes that it is easy to implement and it performs well even if it doesn't hold conditional independence. There are 2 processes you can use Naive Bayes algorithm for sentiment analysis [11]

Classifying text: uses 3 inputs, input data to training classifiers, data to test classifiers and text that is used to classify. The text is manually classified based on sentiment into positive and negative sentiments.

[text]

[sentiment]

Example for positive sentiment

[text] - [The show was amazing]

[sentiment] - [positive]

Example for negative sentiment

[text]- [The features of the car are bad]

[sentiment] - [negative]

Extraction: in extraction all the prepositions or stop words such as is, the, if, and etc is removed. The words are later extracted and weighed through Bag of word model therefore the order of the words does not matter

1) Extract words with length greater than 3.
2) Combine extracted words into one file.
3) Get word sentiment
4) To get frequency distribution for words, Count each word occurrence with corresponding sentiment
5) Create Frequency distribution.
6) Using frequency distribution, create a training set.

2.4.2 MAXIMUM ENTROPY: It uses probabilistic latent semantic analysis and is very convenient for NLP [24]. The MaxEnt algorithm uses probability to find which class the particular sentence belongs to, for this we would require the distribution to be uniform [26]. Maximum entropy is based on a principle called ‘Principle of Insufficient reasoning’ when one has insufficient information to distinguish between the probability of two events, the best strategy is to consider them equally likely.

Consider y to be the output from the finite set Y. x maybe an information (contextual) that might influence y, x belongs to the finite set X. we need to construct a model that precisely denotes the behaviour of random processes. P(y|x) will be the probability, x being the context and y being the output [27]. Entire algorithm can be found in [26-28].

Maximum entropy uses important classification feature like the relationship between the degree of relevance with degree of adverbs [44] relationship between words and context of words etc. [29]

A lot of researchers like [30, 31, 25, 5] and more have found that MaxEnt surpasses the Naive Bayes most of the time but not always; In [32] the Maximum Entropy is found to be more accurate than Naive Bayes by 1.04% in accuracy.

3. SENTIMENT ANALYSIS IN IMAGES

Images serve as a medium through which everyone can express their feelings on a variety of topics. It is said that a picture is worth a thousand words. When it comes to expressing human feelings and sentiments, it is unquestionably more valuable. Predicting such emotions in Images is a very challenging task because emotions very from person to person, for example people in India light candles to celebrate Diwali where as people from other countries light candles during death of a person. An image with such a notion can denote 2 different things in 2 different places. Hence emotions and sentiments are considered to be very challenging predicting them. Sentiment Analysis in Images is used in many places on the internet. Google is one such example. When we search for an image in Google Images we get the results based on that particular emotion. Figure 6 shows the Results for Google Image search for the word “Anger”. We observe that both the images are different but they share a common emotion which is anger.

Figure 6 Google’s Image Sentiment Algorithm showing images of Anger

Two techniques of neural networks are very commonly used

1. Convolutional Neural Networks (CNN) for image processing
2. Recurrent Neural Networks (RNN) for natural language processing.

Unlike Statement Sentimental Analysis, Image Sentiment Analysis requires Image Processing. [40] Proposed a method for classification of sentiments from images to 6 labels i.e. Happiness, Sadness, Fear, Disgust, Anger and Surprise. The data set consists of 9854 images and approximately 1900 images/label. 75% of the total data set was used for training and 25% is used for testing. The ability of software to determine persons, places, objects, actions, and text in photos is known as image recognition. Image recognition is a very important process in the Sentiment Analysis. A convolutional neural network (CNN) is a type of deep and feed-forward artificial neural network that has been used to view pictures successfully in machine learning. CNNs employ a multilayer perceptron variant that requires very little pre-processing. In comparison to other image classification techniques. CNN utilizes deep learning to perform image recognition tasks. The positions and companions of each picture pixel are taken into account since they are crucial for collecting relevant characteristics of an image [41]. CNN allows us to extract certain sophisticated characteristics from the face, making it highly useful for picture categorization based on expressions. Standard data augmentation techniques such as mirroring and random cropping of pictures have been shown to improve accuracy. We can further improve the accuracy by using Stochastic Gradient Descent optimization algorithm. The Modal classifies the Sentiment based on the features/objects that it learns from set of images. Modal in [40] learns sunset, hearts and flowers in order to categorize an image as Love, faces as happiness. This work also demonstrated the increment of accuracy by adjusting base learning rate at a start value of 0.0001.

4. Supervised Learning in Remote Sensing

Deep learning faces time and model complexity issues handling big data used during remote sensing [42]. Handling this requires algorithms with a distributed framework and parallelized machines. In order to achieve this, a 2 stage approach is proposed. These stages include storage as well as processing of BRSD, as depicted in figure 7 [42].

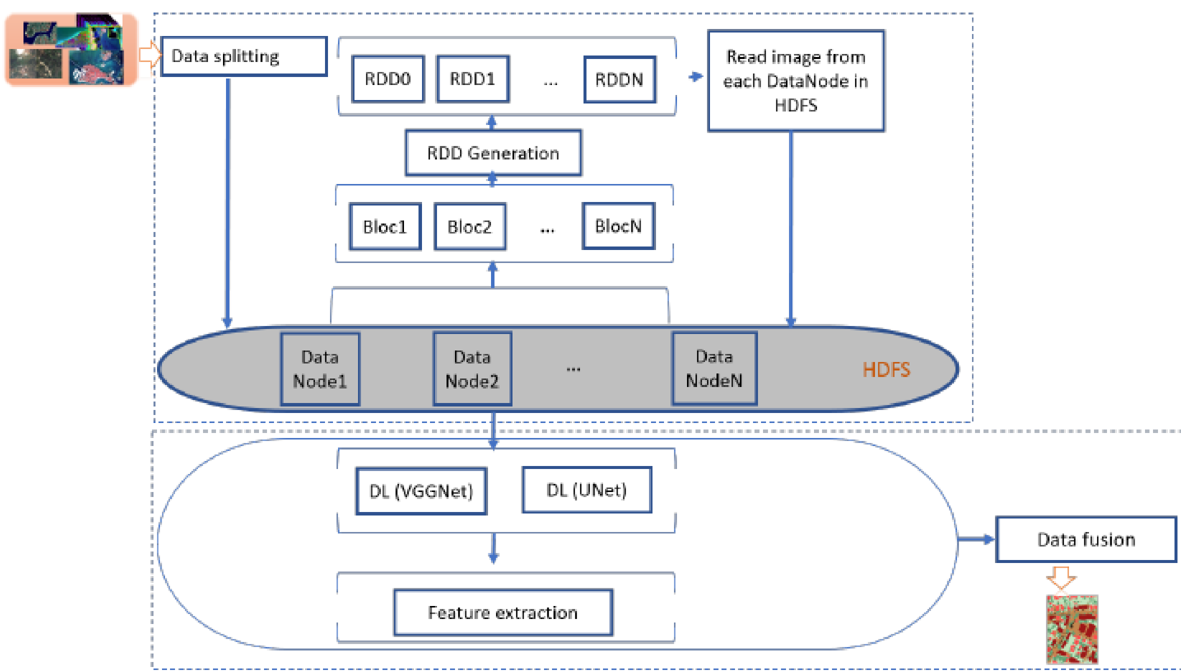

Figure 7: Proposed method in [42]. Reproduced from open access article.

The storage phase is built on Spark-resilient set of data and HDPS. First, a huge amount of RS images are loaded into the HDPS data frame, which are taken as BRSD and classified based on Volume, Velocity, Veracity and Variety. Once the images are loaded, they are split and duplicated using an HDFS data node. A master-slave architecture is used for this, and a clustered computer architecture is used – specifically Hadoop and Spark clusters for their benefits offered. Hadoop clusters consist of master nodes which are responsible for storing large data quantities using HDFS and making parallel calculations and slave nodes using MapReduce and slave nodes which are responsible for data storing task performance and calculations using a Task tracker daemon. Figure 8 depicts these two types of nodes in detail.

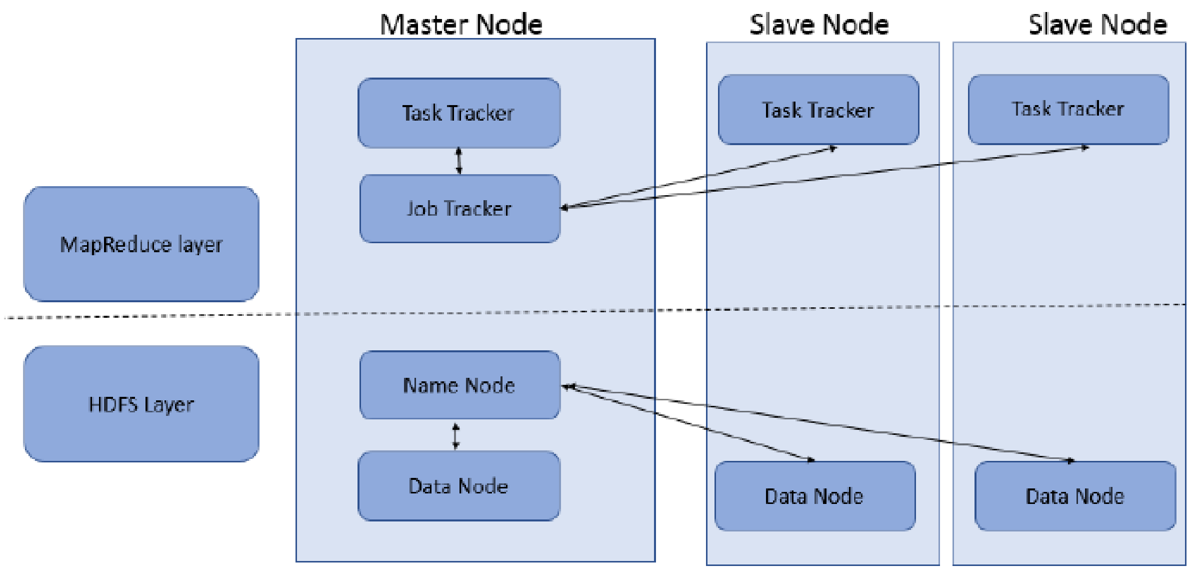

Figure 8: MS architecture [42]. Reproduced from open access article

The concept of the RDD Spark and MapReduce is used for efficiency and preparation for Phase 2 of the process. A distributed computer platform, the Hadoop Spark platform is set up to test these methods. The prerequisites for this platform are data in minute independent blocks, placement of data and processes together and the availability of data in RDD's divided into parts logically. RDD stores memory as objects which can be shared by jobs, hence increasing speed exponentially. Therefore, the matching job returns processed key-value-paired data as output for the next phase.

Figure 4 shows the extraction and classification features of the deep learning processing step.

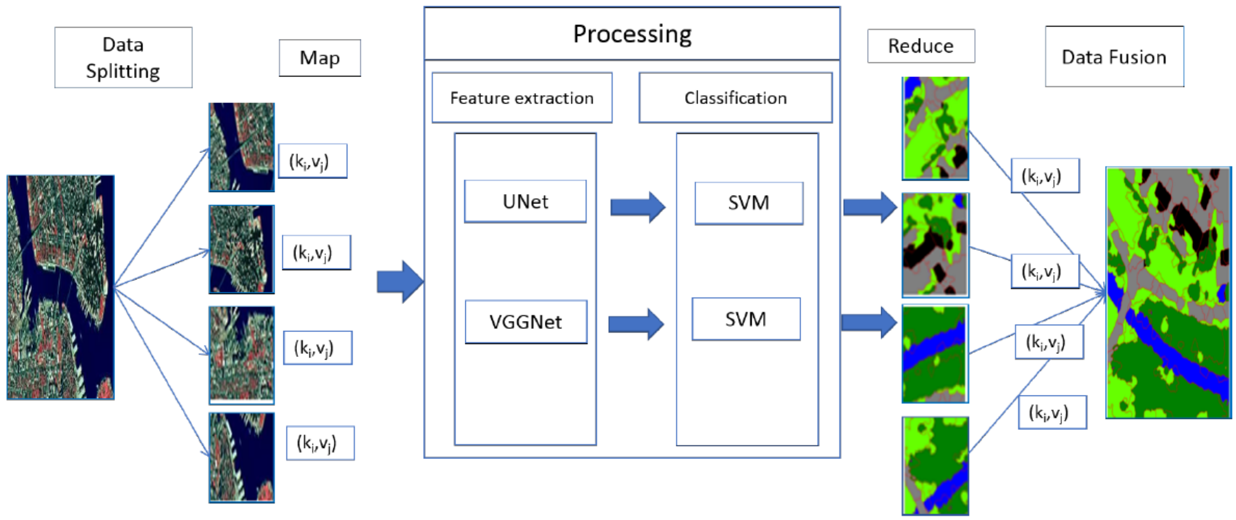

Figure 9: Processing phase [42]. Reproduced from open access article

In Feature extraction as shown in figure 9, VGGNet and UNet pre-trained models are used to process the dataset of satellite images. The VGGNet model contains 2 parts – the extraction part and the classification part. The output layer allows recognition of several physical objects. It is used for RGB images. The UNet model contains an encoder and decoder. This model can take all types of images as input, hence it is used for multispectral images which are fed along with reactance index, and the output is extracted from the bottleneck layer.

The output layers of these models are trained to include predefined labels after which they are disconnected from deeper layers. For validation of this, SVM algorithm is used to plan large margin and large distance between class data points.

After extraction and classification, reconstruction is performed using MapReduce's reducer that includes the key-value data along with intermediate tuple data to produce the final data. At this point, data fusion takes place [42].

5. DISCUSSION:

In the below table we see how vividly the researchers have used various data sets to discuss and convey the accuracy of SVM, MaxEnt and NB. The table has partly been derived from [5].

In [33] they have used features such as Stanford POS tagger, SentiWordNet and other features. They have considered Feature sets F1, F2 and F3. F1 includes content free features and is domain independent, F2 consists of domain dependent data, F1+F2 consists of content-free and content-specific features. They're accuracy is comparably good because of evaluation procedure and use of stop words and the feature filtering criteria.

In [34], the researchers have not shown accuracy specifically but the discussion was mainly on SVM algorithm and the different examples. They have taken different input classes for comparison and inputs in different dimensions as well. In [14] the sentiment of tweets are examined. Unigrams, Bigrams, Unigrams+Bigrams; they have also used post processing features on training data sets which helps determine the accuracy of algorithms. They have also collected their own data set from Twitter. [35] Has not only performed 3-fold and 10-fold cross validation for each corpus but also presented comparative study of SVM, NB and MaxEnt with graphs and tables very elegantly; they prove that in most or all of the data set input comparison SVM outperforms the other 2 algorithms. [36] Illustrates the performance of the most used algorithm such as K-NN, Naive Bayes and SVM. The researchers compare the accuracy of the algorithm normally and accuracy of the algorithm with FS-INS; they further on discussed that accuracy with FS-INS is more. [38] Shows the comparative study of algorithms for sentiment analysis for online messages. 5 data sets have been made use of by the researchers in this paper, such as E1: Opinion mining using raw data; E2: Opinion mining using raw data and pre-process activities; E3: Opinion mining using raw data and FS-INS; and so on. No evaluation has been proposed in [39] but Naive Bayes algorithm has been used

Table 2. Below shows a combined comparative study result from different papers that have taken into account different datasets and input data.

Note: the idea is derived from [5].

Table 2: Comparative study [5]

Type of Machine Learning Classifier	Technique	Dataset	Accuracy (%)	Domain	Ref
Supervised Machine Learning	Support Vector Machine (SVM)	reviews	82.7	Marketing	[33]
		reviews	No evaluation	General	[34]
		reviews	91.5		[35]
		reviews	82.9		[36]
		tweets	81.6		[14]
		reviews	82.7		[1]
	Maximum Entropy (Maxent)	tweets	83.0	General	[14]
		reviews	77.1		[24]
		reviews	80.8		[1]

	Naïve Bayes (NB)	reviews	84.5	Marketing	[37]
		tweets	82.7	General	[22]
		tweets	No evaluation		[22]
		reviews	84.9		[36]
		reviews	81.4		[24]
		Online messages	91.4		[38]
		reviews	No evaluation		[39]
		reviews	80.6		[1]

CONCLUSION

This paper on the whole contours discussed research and puts forth reviewed accurate supervised algorithms. It puts together all studied information and firmly stands by the point that Support Vector Machines (SVM) is accurate most of the time. Naive Bayes chiefly being better than Maximum Entropy. Therefore these algorithms have not only been beneficial but also given growth to hybrid algorithms paving even more lethal algorithms in Natural language understanding for sentiment analysis. We have also discussed basic definitions, how the algorithm works and how it outperforms the others. We have reviewed that Sentiment analysis can also be used for Images processing using CNN and remote sensing using SVM. Therefore sentiment analysis has various applications and so do the supervised machine learning algorithms. They can be put in innumerable fields of sciences. The future studies could have enhancements concentrating on building algorithms or renovating the same so that actual meaning and contextual meaning could be understood; it should be able to differentiate one from the other. For example words like, *bit, play, watch, bear* etc. Accordingly, innovative techniques that overcome such problems must be encouraged for technological growth.

REFERENCES

1. Turney P., “Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews”, Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), pp. 417-424, 2002.

2. P. Kalaivani, “Sentiment classification of movie reviews by supervised machine learning approaches,” Ijcse.com. [Online]. Available: http://ijcse.com/docs/INDJCSE13-04-04-034.pdf. [Accessed: 05-Aug-2021].
3. Dang, N.C.; Moreno-García, M.N.; De la Prieta, F. Sentiment Analysis Based on Deep Learning: A Comparative Study. Electronics 2020, 9, 483. https://doi.org/10.3390/electronics9030483
4. H. Khayyam et al., "A Novel Hybrid Machine Learning Algorithm for Limited and Big Data Modeling With Application in Industry 4.0," in IEEE Access, vol. 8, pp. 111381-111393, 2020, doi: 10.1109/ACCESS.2020.2999898.
5. Othman, Mahmoud & Hassan, Hesham & Moawad, Ramadan & Elkorany, Abeer. (2014). Opinion mining and sentimental analysis approaches: A survey. Life Science Journal. 11. 321-326.
6. Gentleman, R., & Carey, V. J. (2008). Unsupervised Machine Learning. Bioconductor Case Studies, 137–157. doi:10.1007/978-0-387-77240-0_10
7. Zhi-Hua Zhou, A brief introduction to weakly supervised learning, National Science Review, Volume 5, Issue 1, January 2018, Pages 44–53, https://doi.org/10.1093/nsr/nwx106
8. I. G. Maglogiannis, Emerging artificial intelligence applications in computer engineering: Real word AI systems with applications in eHealth, HCI, information retrieval and pervasive technologies. Amsterdam, NY: IOS Press, 2007.
9. Nasteski, Vladimir. “An overview of the supervised machine learning methods.” Horizons 4 (2017): 51-62.
10. T. Oladipupo, “Types of machine learning algorithms,” in New Advances in Machine Learning, InTech, 2010.
11. B. K. Bhavitha, A. P. Rodrigues and N. N. Chiplunkar, "Comparative study of machine learning techniques in sentimental analysis," 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT), 2017, pp. 216-221, doi: 10.1109/ICICCT.2017.7975191.
12. Osisanwo F.Y., Akinsola J.E.T., Awodele O., Hinmikaiye J. O., Olakanmi O., Akinjobi J. "Supervised Machine Learning Algorithms: Classification and Comparison". International Journal of Computer Trends and Technology (IJCTT) V48(3):128-138, June 2017. ISSN:2231-2803. www.ijcttjournal.org. Published by Seventh Sense Research Group.
13. Vapnik, V. N. (1995). The Nature of Statistical Learning Theory. (2nd ed.). Springer Verlag. Pp. 1 – 20. Retrieved from website: https://www.andrew.cmu.edu/user/kk3n/simplicity/vapnik2 000.pdf
14. Go A., R. Bhayani, L. Huang, “Twitter sentiment classification using distant supervision”, CS224N Project Report , 2009
15. E. Frank, M. A. Hall, and I. H. Witten, “The WEKA Workbench,” in Morgan Kaufmann, Fourth Edition, 2016, pp. 553–571.
16. C. C. Aggarwal and C. X. Zhai, Mining text data. 2013
17. N. J. Sanders, “Sanders-twitter sentiment corpus,” Sanders Anal. LLC., 2011
18. M. Ahmad, S. Aftab, M. Salman, N. Hameed, I. Ali, and Z. Nawaz, “SVM Optimization for Sentiment Analysis,” Int. J. Adv. Comput. Sci. Appl., vol. 9, no. 4, 2018.
19. Bhavitha, B. & Rodrigues, Anisha & Chiplunkar, Niranjan. (2017). Comparative study of ML techniques in sentimental analysis. 216-221. 10.1109/ICICCT.2017.7975191.
20. Ali, F., Kwak, K.-S., & Kim, Y.-G. (2016). Opinion mining based on fuzzy domain ontology and Support Vector Machine: A proposal to automate online review classification. Applied Soft Computing, 47, 235–250. doi:10.1016/j.asoc.2016.06.003

21. T. Joachims, “Text categorization with Support Vector Machines: Learning with many relevant features,” in Machine Learning: ECML-98, Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 137–142.
22. Mahendran A., A. Duraiswamy, A. Reddy, C. Gonsalves, “Opinion Mining for Text Classification”, International Journal of Scientific Engineering and Technology, Vol. 2, 2013
23. Kevin P. Murphy. Ubc.ca. [Online]. Available: https://www.cs.ubc.ca/~murphyk/Teaching/CS340-Fall06/reading/notation.pdf. [Accessed: 05-Aug-2021].
24. Smeureanu, Ion & Cristian, Bucur. (2012). Applying Supervised Opinion Mining Techniques on Online User Reviews. Informatica Economica Journal. 16.
25. H. Y. Lee and H. Renganathan, “Chinese Sentiment Analysis Using Maximum Entropy,” Aclanthology.org. [Online]. Available: https://aclanthology.org/W11-3713.pdf. [Accessed: 05-Aug-2021].
26. A. L. Berger, S. A. Della Pietra, and V. J. Della Pietra, “A maximum entropy approach to natural language processing,” Msu.edu. [Online]. Available: http://www.cse.msu.edu/~jchai/CSE891Spring04/MEforNLP.pdf. [Accessed: 05-Aug-2021].
27. A. Ratnaparkhi, “A simple IntrA simple introduction toduction to maximum Entro maximum entropy models for naturopy models for natural al language PrLanguage processing ocessing,” Upenn.edu. [Online].Available:https://repository.upenn.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1083&context=ircs_reports. [Accessed: 05-Aug-2021].
28. K. Nigam, “Using maximum entropy for text classification,” in In IJCAI-99 Workshop on Machine Learning for Information Filtering, 1999.
29. Xie, X., Ge, S., Hu, F., Xie, M., & Jiang, N. (2017). An improved algorithm for sentiment analysis based on maximum entropy. Soft Computing. doi:10.1007/s00500-017-2904-0
30. Raychaudhuri, S. (2002). Associating Genes with Gene Ontology Codes Using a Maximum Entropy Analysis of Biomedical Literature. Genome Research, 12(1), 203–214. doi:10.1101/gr.199701
31. Mehra, Nipun, S. Khandelwal and Priya Patel. “Sentiment Identification Using Maximum Entropy Analysis of Movie Reviews.” (2002).
32. Saif, H., He, Y., Fernandez, M., & Alani, H. (2014). Semantic Patterns for Sentiment Analysis of Twitter. Lecture Notes in Computer Science, 324–340. doi:10.1007/978-3-319-11915-1_21
33. Y. Dang, Y. Zhang and H. Chen, "A Lexicon-Enhanced Method for Sentiment Classification: An Experiment on Online Product Reviews," in IEEE Intelligent Systems, vol. 25, no. 4, pp. 46-53, July-Aug. 2010, doi: 10.1109/MIS.2009.105.
34. Khairnar J., M. Kinikar, “Machine Learning Algorithms for Opinion Mining and Sentiment Classification”, International Journal of Scientific and Research Publications, Vol. 3, pp. 153-161, 2013
35. Saleh M., M. Valdivia, A. Ráez, L. López, “Experiments with SVM to Classify Opinions in Different Domains”, Expert Systems with Applications Journal, Vol. 38, pp. 90-105, 2011
36. Samsudin N., M. Puteh, A. Hamdan, M. Zakree , A. Nazri," Immune Based Feature Selection for Opinion Mining", in: Proceedings of the World Congress on Engineering, Vol. 3, pp. 120-127, 2013

37. Zhang, Ziqiong, Q. Ye, Zili Zhang and Yijun Li. “Sentiment classification of Internet restaurant reviews written in Cantonese.” Expert Syst. Appl. 38 (2011): 7674-7682.
38. Samsudin N., A. Hamdan, M. Puteh, M. Zakree, A. Nazri , “Mining Opinion in Online Messages”, International Journal of Advanced Computer Science and Applications, Vol. 4, pp. 52-59, 2013
39. Kishore P., K. Naidu, "Sentiment Analysis Using Semi-Supervised Naïve Bayes Classifier ", International Journal of Innovative Technology and Research, Vol. 1, 2013
40. V. Gajarla and A. Gupta, “Emotion detection and sentiment analysis of images,” Gatech.edu. [Online]. Available: https://www.cc.gatech.edu/~hays/7476/projects/Aditi_Vasavi.pdf. [Accessed: 09-Aug-2021].
41. G.Sirisha, Sai Jyothi Bolla, V.Jyothsna, R.Mounica, “Sentiment Analysis Using Image Processing:A Survey,” International Journal of Grid and Distributed Computing, vol. 13, no. 1, pp. 2955–2962, 2020.
    Sentiment Analysis Using Image Processing: A Survey
42. Chebbi, I.; Mellouli, N.; Farah, I.R.; Lamolle, M. Big Remote Sensing Image Classification Based on Deep Learning Extraction Features and Distributed Spark Frameworks. Big Data Cogn. Comput. 2021, 5, 21. https://doi.org/10.3390/bdcc5020021
43. RAJASEKARAN, Abirami; R. VARALAKSHMI, Dr. Review on automatic text summarization. International Journal of Engineering & Technology, [S.l.], v. 7, n. 2.33, p. 456-460, june 2018. ISSN 2227-524X. Available at: <https://www.sciencepubco.com/index.php/ijet/article/view/14210>. Date accessed: 09 aug. 2021. doi:http://dx.doi.org/10.14419/ijet.v7i2.33.14210.
44. Xie, X., Ge, S., Hu, F. et al. An improved algorithm for sentiment analysis based on maximum entropy. Soft Comput 23, 599–611 (2019). https://doi.org/10.1007/s00500-017-2904-0