Original Research

Volume 3, Issue 2, pp 165-172, December 2021

ACHIEVERS JOURNAL OF SCIENTIFIC RESEARCH

Opecn Access Publications of Achievers University, Owo

Available Online at www.achieversjournalofscience.org

Adaptive Depression Diagnosis Using an Improved Support Vector Machine

A. Adegboye [1*] and M.A. Adegoke[2]

[1]Department of Mathematical Sciences, Achievers University, Owo, Ondo State, Nigeria

[2]Department of Computer Science, Bells University, Ota, Ogun State, Nigeria

*E-mail: akanbi2090@yahoo.co.uk

Submitted: October 18, 2021 Revised: December 1, 2021 Accepted: December 4, 2021 Published: December 21, 2021

ABSTRACT

The aim of this research is to modify the traditional support vector which is not adaptive to be adaptive. In the traditional SVM algorithm, the computation of the optimal plane is based only on the closet data points (support vector). In this research work, the existing support vector was modified such that instead of relying just on the closet support vector, the average distance of all the vectors (data objects) is obtained and used to compute the optimal plane. This accommodates a data point that represents a new fact or could contribute significantly which is not close to the partitioning planes to be used in construct the optimal plane. The new support vector formulation is adaptive because new facts are used to update the knowledge of the SVM training data before using it to construct a hyper-plane. The mathematical formulation showed that the new method of constructing a hyper-plane is adaptive and can accommodate new facts. Thus, this takes care of the identified lapse in traditional Support Vector. In this research work, a Support Vector that is adaptive to a new fact is formulated. This enables Support Vector to be used in an environment that is dynamic and susceptible to quick changes.

KEYWORDS: Support Vector Machine, adaptive, data points, partitioning plane, depression, and diagnosis

1.0 Introduction

The conventional support vector machine (SVM) models are not adaptive (Tan, 2004; (Bhavsar, 2017). They rely only on support vectors that are closest to the partitioning plane region for diagnosis. If the vectors of a data object that represent a new fact are not close to the partitioning planes, such new facts cannot be used to update the knowledge of the diagnosis machine. There is a need to develop a real-time diagnosis machine using a support vector machine that is capable of regularly updating its diagnosis knowledge with new facts. Such a diagnostic machine will be adaptive to the availability of new facts.

2.0 Overview of Support Vector Machine

Support vector machine (SVM) is a model that separates data objects in a feature space into two classes or two distinct groups (tan, 2014). For example, in the depression diagnosis problem, the support vector machine separates patients into either "depressed" or "not-depressed" groups. Each data object, also described as a vector, must-have features say, $x_1, \ldots ,x_p$ and a class label y_i.

The vectors from the same class fall on the same side of the separating plane as shown in Figure 1.This plane that separates the data objects into groups is called a support vector machine.

Support vector machine (SVM) thus, treats each data object as a point in the features space such that the object belongs to one class or the other. That is, a data object x_i, either belongs to a class, in which case the class label is y_i, or it does not belong to the class, in which case the class label y_i = -1. Let (x_i, y_i) be a set of training examples (i.e., data objects and their corresponding classes), the definition of the data object is therefore represented as followings:

Data object $= \{(x_i,y_i)|x_i\in R^p,y_i\in(-1,+1)\}^n_{i=1}$ (1)

Where

$x_i\in R^p$ are the data objects

$y_i\in$ (-1, + 1) are the two classes

p -is the dimension of each data object and n is the number of data objects.

The plane (SVM) which partitions data objects into groups of similar elements is a straight line that must satisfy the following equation:

$$F(x) = \text{sign}\,(w.x_i + b) \qquad (2)$$

Where:

Sign refers to the sign of ($w.x + b$)

W is a p-dimensional vector

X_i is as earlier defined, and *b* is a scalar.

The incoming data objects (vector) x_i for which f(x_i) is positive are placed in one class, while vectors for which f(x_i) is negative are placed in the other class.

There are usually more than one sub planed that are capable of partitioning a given data object into groups of similar elements as shown in figure 2

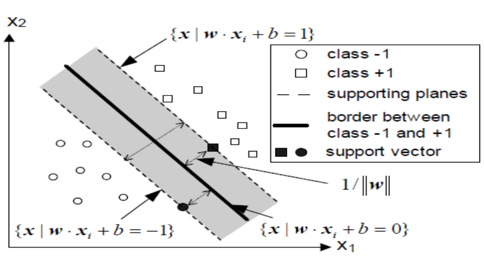

Figure 2: Linear separation of two classes -1 and +1 in two-dimensional space

With SVM classifier (Source: Burges, 1998)

However, there are only one of such planes that best segregates the data objects. The task of the SVM algorithm, therefore, is to identify that plane (line) that best separates the two classes. The optimal plane is the one with the maximum distance from the closet data objects (the data objects can belong to either class). The higher the distance of the plane from the closet data objects, the better the classifier. The objective function of the SVM, therefore, is to search for the maximum distant plane amongst candidate planes that creates the greatest margin m between the two classes. This is accomplished (done) by maximizing the margin m between the closet data object and the farthest plane. The margin width is obtained by subtracting equation (3) from equation (4) as follows:

$$w.\,x^{+} + b = 1 \;\ldots\ldots\ldots..(3)$$

$$w.\,x^{-} + b = -1 \;\ldots\ldots\ldots..(4)$$

where equation (3) and (4) represents the group of data elements that falls on the upper and lower sides of the plane respectively. Subtracting (3) from (4) produces equations (5) as follows:

$$w(x^{+} - x^{-}) + 0 = 2 \qquad (5)$$

where $(x^{+} - x^{-})$ is the margin *m*i.e

$$m = x^{+} - x^{-} = \frac{2}{|w|} \qquad (6)$$

Equation 6 can also be re-written as: m= $1/2w^{T}$. w or$m = 2|w|^{T}.w$ (Tan, 2004)

The margin is the distance of the closet data object to the classifier. Maximizing the margin requires formulating a quadratic optimization problem and solving for *w* and *b* such that :

$f(w) = \frac{1}{2}w^{T}.w$ is minimized for all $\{(x_i, y_i)\}$ $y_i\,(wx_i + b \geq 1)\forall_i$

That is,

$$\text{Minimize } f(w) = \frac{1}{2}w^{T}.w \qquad (7)$$

Subject to

$$y_i\,(w.x_i + b) \geq 1 \;\forall_i \qquad (8)$$

The solution involves constructing a dual problem where a Langrangian multiplier α_i is associated with every constraint in the primary problem. The solution has the form (tan, 2004)

$$w = \Sigma\alpha_i y_i\, x_i \qquad (9)$$

and

$$b = y_i.w^{T}.x \qquad (10)$$

for any x_k such that $\alpha_k \neq 0$ where α_i is the support vector i.e the closet data objects to the SVM,and $w^{T}$is the transpose of w. Substituting equations (9) and (10) in equation (2), that is

$$f(x) = \text{sign}\,(w.x+b) \qquad (2)$$

the classifying function, (wx + b) then becomes :

$$f(x) = \text{sign}\,(\Sigma\alpha_i y_i x_i^{T}.x + y_i - w^{T}.x_i \qquad (11)$$

it is equation (11) that partitions a data objects into either side of the plane.

The overall working of the SVM model is summarized in the following procedure:

Algorithm 1:

Step 1: A set of training data that represents both positive examples and negative examples of the problem at hand are prepared

Step 2: The closet data point called support vectors, to the dividing planes are identified from either of the training classes.

Step 3: The plane with the maximum distance from the closet support vector is determined by solving equations (9) and (10)
Step 4: Given a new data object (Input vector) the maximum hyper-plane so obtained can be used to predict the class for the new object using equation (11).

The flow chart of the SVM procedure is as shown in figure 3

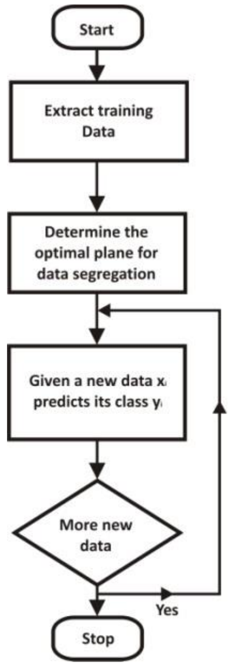

Figure 3: Flowchart of a 2- class non adaptive SVM

The above SVM equation is capable of predicting the class yi for new input data by deriving the knowledge for classification from the training logs of the previously classified and documented cases. However, the traditional SVM algorithm based the computation of the optimal plane only on the closet data points (support vector). Thus, new training data whose vector happens not to be closed to the dividing planes would not have any influence on the machine calculation of the optimal plane and hence cannot be used to update the knowledge of the support vector algorithm. New facts emerge on daily basis and these facts should be able to be used to update the knowledge of support vector machines. In this study, therefore, instead of relying just on the closet support vector, the average distance of all the vectors (data objects) is obtained and used to compute the optimal plane. In order to adjust the support vector machine to be capable of grouping data objects into more than two classes, an adaptive learning approach is adopted. In adaptive learning, the entire data objects are first partitioned into two major classes, and each of the two classes can further be partitioned into another two groups by defining the criteria for the partitioning at each stage.

3.0 Related Work on Support Vector

(Anuja and Chitra, 2013) Present Classification of Diabetes Disease Using Support Vector Machine. The experimental result shows that the Support Vector Machine can be used for disease diagnosis. A new fact (Disease symptom) that is not used in the system training dataset cannot be diagnosed. This system is not adaptive.

(Deepti and sneetal, 2013) present Classification of Heart Disease Using SVM and ANN. The research experimental result shows that SVM outperforms NN in the area of accuracy and execution time. The system does not take care of new fact detection.

In (Chitra and Anto, 2015) Diagnosis of Diabetes Using Support Vector Machine (SVM) and Ensemble Learning Approach were presented. The research experimental result shows that Support Vector Machine can be used for Diabetes disease diagnosis. A new fact (Diabetes symptom) that is not used in the system training dataset cannot be diagnosed. This system is not adaptive. (Qin *et al.,* 2016) Present automated cell selection using Support Vector Machine for application to Spectral Nan cytology. The convectional support vector machine used for diagnosing is not adaptive. A data point that represents a new fact or could contribute significantly which is not close to the partitioning planes cannot be used in construct the classification model. This could limit the model's ability to generalize relatively to new data that are not part of the training dataset.

(Zhichao *et al.,* 2017) Used a multi-kernel support vector Machine to recognize depression. The data used in the study was obtained from social media. The model was trained to recognize depression from social media data. The system had a high level of prediction accuracy but it had some major drawbacks. A data point that represents a new fact or could contribute significantly which is not close to the partitioning planes cannot be used in construct the classification model. This could limit the model's ability to generalize relatively to new data that are not part of the training dataset.

In (Hyunsik and Jeonegyeup, 2018), Automatic Task Classification via Support Vector Machine and Crowd-sourcing was presented. The research work shows that automatic task classification can be implemented for mobile devices by using the support vector machine algorithm and crowd-sourcing. The task that represents a new fact or could contribute significantly which is not close to the partitioning planes cannot be used in construct the classification model. This could limit the model's ability to generalize relatively to new data that that are not part of the training data set.

(Onwuegbuche *et al.,* 2019) Present Support Vector Machine for Sentiment Analysis of Nigerian Banks Financial Tweets. The research assists the Nigerian banks in understanding their customers better through their opinion. However, data points (opinion about bank) that contribute meaningfully which is not close to the partitioning planes cannot be used in construct the classification model. This can create an information gap in the above understanding, thus limit decision-making concern banks in Nigeria using social media data.

In (Muhammad *et al.,* 2020) Support Vector Machine-Based Classification of Malicious Users in Cognitive Radio Networks was presented. In this research work, a machine learning algorithm based on support vector machine (SVM) is used to classify legitimate secondary users (SUs) and malicious users (MUs) in the cognitive radio network (CRN). Legitimate secondary users which are not close to the partitioning planes cannot be used in construct the classification model. Thus, such users may prevent access to the radio spectrum.

In (Lumbanraju *et al.*, 2021) Abstract Classification Using Support Vector Machine Algorithm (Case Study: Abstract in a Computer Science Journal) was presented. The conclusion based on this research is that there are two factors that affect classification accuracy that is the number of members between scientific classes that are not balanced and the number of features generated from text mining. The convectional support vectors used in this work cannot make use of the scientific classes that are not balanced and the number of features generated from text mining, that are not close to the hyper, even though they can contribute to classification accuracy

4.0 Description of the adaptive depression diagnosis system

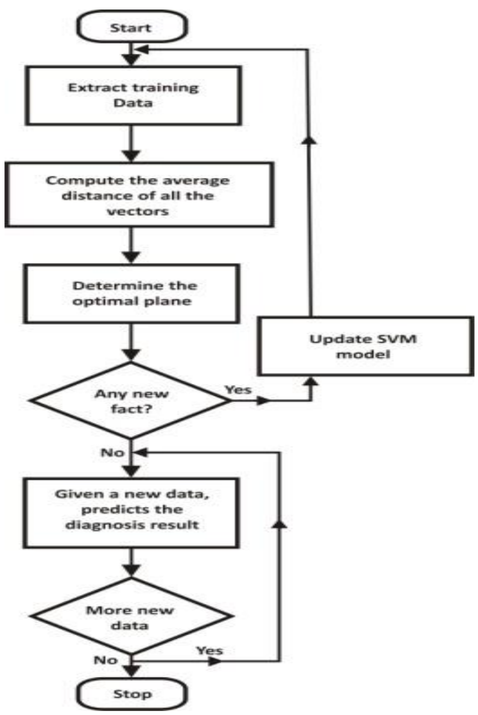

Figure 4: flowchart of the adaptive depression diagnosis system

Figure 4 depicts the flowchart of a multi-classed adaptive depression diagnosis system. It is an improvement in figure 3. The difference between the non-adaptive support vector machine (The traditional SVM model) and the work presented in this section is the module that computes the average distance of all the vectors to the partitioning region. In the conventional SVM model, only the vectors that are closest to the partitioning region are relevant for diagnosis decisions. Thus new information (data object) that is not among the subset of the training data cannot be used to update the knowledge of the diagnosis machine.

Figure 4 consists of seven sub-modules vis-a-vis: training data extraction module, the module that computes the average distance of all the vectors, the module that determines the optimal plane, the module that regularly checks for new facts, the update module, the module that predicts the cause of depression symptoms and the iterative module. The module that checks for new facts is connected to the update module. This model is adaptive because new facts are used to update the knowledge of the SVM training data before using the model to predict the cause of depression symptoms. In order to diagnose symptoms into “Depressed”, “might be depressed” and “not depressed” modes, (which are more than two classes), an installment prediction that uses one – versus all approach is adopted (Bhavsar and Ganatra, 2012). In the installment diagnosis approach, a patient, based on the symptoms he or she manifests, is first diagnosed as “not depressed” or “depressed”. The depressed patient can further diagnose as “mildly depressed” or “critically depressed” depending on the intensity of the symptoms as compared to the available training data.

5.0 Adaptation Algorithm

If the average vectors distance in the training data is denoted by p(n) and the distance of a new fact to the partitioning region is denoted by y(n), then the updated average distance of the support vectors to the partitioning region is given by :

$$g(n) = \frac{p(n) + y(n)}{n} \qquad 12$$

where n is the number of data objects in each training instance. The average distance is used to replace $\alpha_i$in equation (9) as follows:

$$w = \Sigma\, g(n) y_i x_i \qquad 13$$

The adaptation algorithm is therefore given as follows:

Algorithm 2

Step1: Extract training data for depression cases from previously documented cases

Step 2: Compute the average distance of all the support vectors using equation (12)

Step 3: Determine the optimal plane using equations (9) and (10) as modified by equation (12) obtained in step 2

Step 4: Check for a new fact about depression diagnosis

If new fact exist

Update the SVM database

Using step 2

Else

Step 5: Diagnoses the patient using the adaptation diagnosis algorithm

Step 6: Output result

Step 7: end.

The adaptation algorithm (step 5) is given as follows:

1. Given input x_i ; i = 1, 2, ……,n
2. diagnose x_i into y_i ; j = -1, + 1

   Where: -1: represent not depressed

   +1: represent depressed
3. If y_j = +1

   re-diagnose x_i into y_k ; k = +1, 0

   where: 0 = mildly depressed

   +1 = critically depressed
4. Output result
5. end.

6.0 Conclusion

A major advantage of SVMs is that a relatively small sample might be sufficient to build an effective model for classification. A data point that represents a new fact or could contribute significantly which is not close to the partitioning planes cannot be used to construct the classification model. This could limit the model's ability to generalize relatively to new data that are not part of the training dataset. In this research work, a modified support vector machine algorithm that takes care of this identified lapse is proposed. The algorithm can be used to diagnosis new facts in real-time. It is recommended in application areas such as classification, regression, and many practical problems.

References

1. Anuja, B. and Chitra, A. (2013): Classification of Diabetes Disease Using Support Vector Machine. International Journal of Emerging Research and Applications.Volume 3, Issue 2..1797-1801.
2. Bhavsar and Ganatra (2012): A Comparative study of training Algorithms for Supervised Machine Learning. International Journal of

Soft Computing and Engineeringvol. 6, Issue 3. Pp.74-81
3. Bridgeall Raj (2017): Lecture Notes on Support Vector Machines
4. Burges, (1998): A Tutorial on Support Vector Machine for pattern recognition: Data Mining and Knowledge Discovery Vol.1 121-167.
5. Chitra and Anto, (2015): Diagnosis of Diabetes Using Support Vector Machine and Ensemble learning Approach. International Journal of Engineering and Applied Sciences Volume 2, Issue 11.Pp.68-72.
6. Deepti and Sheetal, (2013): Classification of Heart disease Using SVM and ANN. International Journal of research in Computer and Communication Technology, Vol 2, Issue 9, 694-701. https://documents.pub/document/introduction-to-support-vector-machines-lecture-notes-introduction-to-support-vector.html (Accessed March, 2021).
7. Hyungsik, S. and Jeongyeup, P.(2018): "Automatic Task Classification via Support Vector Machine and Crowdsourcing", *Mobile Information Systems*, vol. 7, Article ID 6920679, 9 pages. https://doi.org/10.1155/2018/6920679
8. Maltarollo, V., Kronenberger, T., Espinoza G., Oliveira P., Honorio K. (2019): Advances with support vector machines for novel drug discovery. Expert Opin Drug Discov. Jan;14(1):23-33. doi:
9. Muhammad.S, Liagat.K, Noor.G, Amir. M, Junsu.M, and Su.K (2020): Support Vector Machine-Based Classification of Malicious Users in Cognitive Radio Networks. Hindawi Wireless Communications and Mobile Computing Volume 2020, Article ID 8846948, 11 pages https://doi.org/10.1155/2020/8846948