Dean&Francis
ISSN 2959-6157

The Widely Use of Support Vector Machine in Many Areas

Yuxuan Zhang

No1. Middle School of Central China Normal University, Wuhan, China

*Corresponding author: 1812010725@stu.hrbust.edu.cn

Abstract:

After numerous experiments, a lot of data will be obtained, but the classification of fata is complicated and cumbersome. Hence, it is crucial for them to classifier data into different classes in a efficient way so that there is a subject called Support Vector Machine (SVM). Moreover, many research areas use SVM with other mathematical method like physics, chemistry and biology. In this article, the author will fully introduce the history, definition, composition, kernel trick and four common functions about SVM and the widely uses of it in different areas, such as image classification, text categorization and bioinformatics. After showing these typical examples, the result shows that SVM plays an significant role in helping researchers classifier data and getting the most sufficient result in their researches. In the future, there will certainly be more and more studies using this method to efficiently classify data, and scholars in different fields can make reasonable use of this method based on their types of data, thus saving research time and improving the accuracy of results.

Keywords: Support Vector Machine, image classification, text categorization and bioinformatics.

1. Introduction

This article can be divided into two parts which are the definition and uses of Support Vector Machine (SVM). According to the first part, SVM is widely known by researchers especially among mathematics students and scholars. It is a convenient and efficient method to classify linear and nonlinear data which need researchers to use kernel trick for nonlinear data. Since the data in many studies are stochastic, most of them are nonlinear, which was a difficult problem for people before SVM was proposed, but after understanding the nature and usage of SVM, this difficulty will be easily solved, which promotes the following research and lays a very solid data foundation for the conclusion. Therefore, SVM can be applied in may relevant subject researches requiring data statistics or classification.

In this paper, the role of SVM in several different fields will be listed, and its high efficiency is also illustrated by these typical examples. In the first example which is about physics, the study shows the impact of the weapon system during the launch process and this author combine CNN and SVM into the research to simulate the large amount of data. It shows the powerful nonlinear modeling capabilities of SVM and perform these data well in a high dimension. Moving to the second example that is about laser and optoelectronic. Like the first study above, the author combine two mathematical methods which are PCI and SVM. In this research, SVM classifiers are deployed to train and classify these picture data in order to achieve the most efficient classification of cavitation and crack defects. The last one is about bioinformetics. SVM is used in judging the remaining time of fingerprints which can help police handle cases in a more efficient way. Differently, the author not only use SVM, but also use SG, SNV and SG+SNV to feature wavelength extraction. Through this real examples and the detailed introduction of the SVM method, it is not difficult to see the broad application of SVM.

2. Theory of Support Vector Machine

The history of SVM can be traced back to 1963, when Vapnik and Chervonenkis proposed a classification method based on maximum interval. And in 1992, Berhard E. Boser, Isabelle M Guyon and Vladimir N Vapnik suggested a way to creat non-linear classifiers by applying the kernel trick to maximum the margin. The core idea of this method is to find a hyperplane and divide the data into two categories, so that the distance between data points of different classes and this hyperplane is as large as possible, thereby maximizing the classification interval.

SVM is a kind of machine learning method that has been widely concerned in recent years. It is based on statistical learning theory, aims at minimizing structural risk, takes kernel method as means, and takes support vector as result. It has solid theoretical foundation, simple mathematical representation, standard training algorithm, and good generalization performance [1]. Moreover, SVM are composed of three significant parts: hyperplane, support vectors and margin. The hyperplane are used for separating data into different class labels and the support vectors are the sample datasets which need to be classified but some of them are linear, the majority of them are non-linearly separable. Author will introduce these two conditions in detail later. The last one is about margin. This is a separation gap between two lines and it is calculated as the perpendicular distance between the two kinds of data. In SVMs, people should try to maximize the margin to get the most efficient result.

However, as a new machine learning method, there are still some problems in solving practical problems, such as support vector machine method can not directly deal with high-dimensional large-scale data and is sensitive to abnormal samples [2]. There are two conditions for SVM to classify datasets. On the one hand, if the data are linearly separable, in other words, these data can be separated by using a single line like the author shows in Fig. 1.

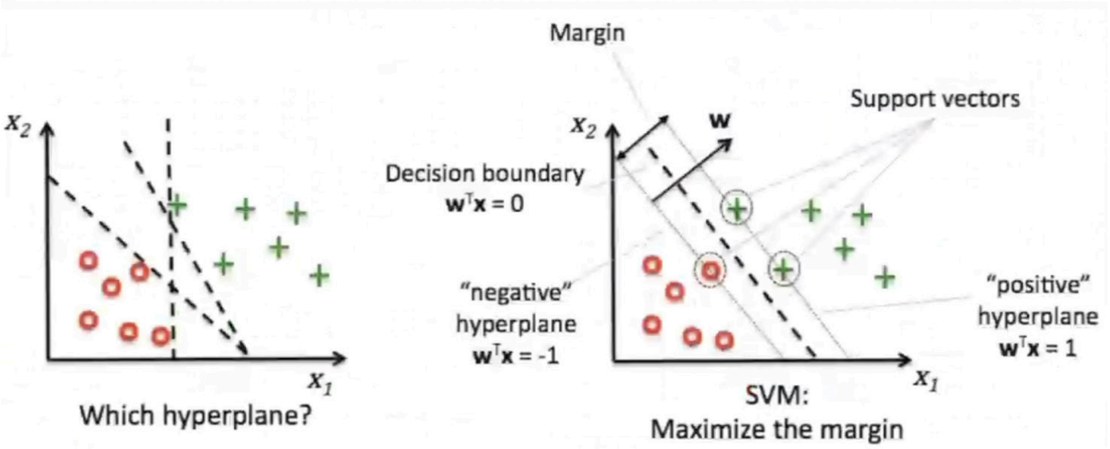

Fig. 1 SVM Classier Tutorial from Kaggle

Two different kinds of data are be classified into different groups. But there are three lines that have different slopes which can divide them. The most efficient one is in the second graph which has the biggest margin to the two different types of data. In SVM, one can try to maximize the margin so that one can get the most efficient result. Suppose the margin is 2r, and people can write the function for these two lines

$$AX_1 + BY_1 + C = r \quad (1)$$

$$AX_2 + BY_2 + C = -r \quad (2)$$

But the result can only be influenced by the margin instead of A, B or C, so let the original function be divided by r, the next function will be shown below.

$$A_1X_1 + B_1Y_1 + C = 1 \quad (3)$$

$$A_1X_1 + B_1Y_2 + C = -1 \quad (4)$$

Therefore, the classification is efficient.

As the author introduced above that the goal of SVM is to

ISSN 2959-6157

find the optimal hyperplane which separates data into different classes. And Linear SVM can only works well for linearly separable data. But in the real research, the majority of data are non-linearly like these two examples shown in Fig. 2.

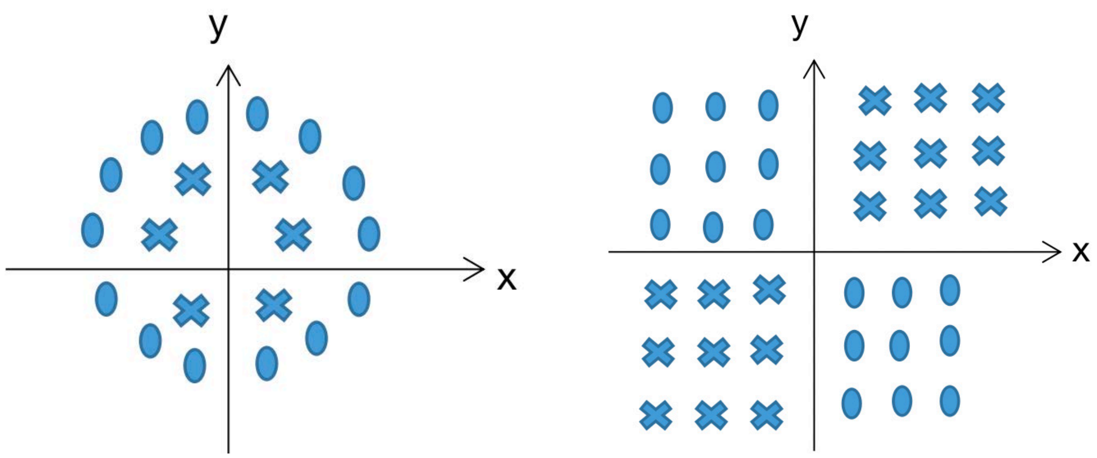

Fig. 2 Two of the non-linear examples.

Looking at the first example, one type of data like the crosses are in a circle and others are around them but just use a line cannot separate them sufficiently. Moving to the second one, one type of data are in the first and third quadrants but others are in the second and fourth quadrants. Obviously, a line cannot make them separable. There are lots of datasets like these two examples and it is a challenge of linear SVM. Then, the author will introduce the Kernel Trick which can tackle this problem.

The first thing the author want to introduce is the definition, Kernel Trick is a mathematical technique that allows people to operate in a high-dimensional space without explicitly computing the coordinates of the data in that space. Moreover, one can use Kernel Trick to transform data into a higher dimension where it becomes linearly separable.

The next part is about the Kernel Function. It is a function that computes the dot product of two vectors in higher-dimensional space. The mathematical form is like this:

$$K(x,y)=\varnothing(x)\cdot\varnothing(y) \quad (5)$$

Here, $\varnothing$ is Feature mapping function and K is Kernel Function.

This is a transfer function and represents that only the intercept can influence the outcome. And it takes two input vectors X and Y and returns a scalar that represents the inner product of these vectors in a higher-dimensional feature space without working with high dimensional data. Two properties it shows is symmetry and positive semi-definite. The symmetry ensures that the similarity between two data points is consistent regardless of their order. In other words, the relationship measured by the kernel is bidirectional and does not depend on the which point is considered first. Therefore, K(x,y)=K(y,x). The other property implies that a Kernel Function K(x,y) is positive semi-definite if for any set of vectors $\{X_1, X_2, X_3, \ldots, X_n\}$ and any real numbers $\{C_1, C_2, C_3, \ldots, C_n\}$ and the following inequality holds:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j K\left(X_i, X_j\right) \geqslant 0. \quad (6)$$

The kernel trick matrix K is constructed by evaluating the kernel function on all pairs of a given set of data points. This is crucial for ensuring that the optimization problems in kernel-based algorithms which leads to unique and optimal solutions and it has stability and reliability. There are two steps of using Kernel Trick to make data separable. The first step is mapping data to a higher-dimensional space using a kernel function and the second step is applying the linear SVM algorithm in this higher-dimensional space.

According to the Kernel Function, there are four common functions, the first one is Linear Kernel: $K(x,y)=X\cdot Y$; The second one is Polynomial Kernel: $K(x,y)=(x\cdot y+c)^d$; The third one is Radial Basis Function Kernel, it is also known as the Gaussian Kernel: $K(x,y)=exp(-\frac{\|x-y\|^2}{2\sigma^2})$. The last one is Sigmoid Kernel: K(x,y)= $tanh(\alpha x\cdot y+c)$ [3].

After the introduction of kernel functions, the author will briefly introduce the choices of kernel functions. Different kernel functions contain different geometric metric characteristics, and the selection of different kernel functions leads to the difference in generalization ability of support vector machines. Because kernel function selection has an important effect on the performance of SVM model, how to effectively select kernel function is a very important research problem in the field of SVM research. Therefore, a supervised kernel function selection mechanism of SVM is constructed, which can combine the prior information of samples and take into account the geometric metric features contained in the kernel function, and effectively avoid its blindness [4].

What the author wants to add is that there is another machine method which is based on Support Vector Machine. Twin support vector machine (TWSVM) is a new machine learning method based on SVM. For classification problems, TWSVM seeks a pair of non-parallel classification hyperplane. For regression problems, TWSVM generates a pair of unparallel functions on both sides of the training data points, and determines the insensitive upper and lower bounds of the regression function respectively. TWSVM is similar in form to SVM, but its computational efficiency is 4 times that of SVM [5].

3. Examples of Support Vector Machine

Moreover, there are three kinds of classification which are image classification, text categorization and bioinformatics. The first one is about using kernels to classify images based on features and the second one is about the kernel methods in Natural Language Processing for document classification. The last one introduces the protein structure prediction and gene classification using kernel methods.

As the author introduced above, SVM are widely used in different areas like physics, economy, biology and investigation and so force based on its strong ability of classifying datasets. And the author will show examples of the widely uses of SVM in these areas to show its practicality. This study focuses on the impact of the weapon system during the launch process. The research on the separation of Adapter Structures under External Environment Factors Object, using Euler angle description method to establish multi-body dynamics during launch process. Obtain aerodynamic parameters of the adapter through learning models and wind tunnel experiments. Through imitation real experiments simulate a large number of launch conditions and construct adapters for different operating conditions. Datasets of landing point distribution, Using CNN and Support Vector Machine. The model built by Support Vector Machine(SVM) algorithm train on the obtained data and construct a fast adapter landing point prediction. Text the model and select examples of Montacalo simulaion experiments [6].

This research combined the advantages of CNN and SVM to tackle the challenging of these datasets, and the article will focus more on Support Vector Machine(SVM). SVM, as a classifier, has powerful nonlinear modeling capabilities By using SVM as the final classifier, and to use its kernel function to map the features extracted by CNN a high dimension feature space enables better handling of nonlinear problem, these enable the CNN and SVM model to perform well on large-scale data and high fidelity data model.

Here is the process in detail. The first thing id to input the feature map output of the convolutional layer into SVM classifier with preset kernel functions and regularization parameters. The second step is model training and optimization. In the training process, by BP calculation Method and gradient descent were used to optimize the parameters of the whole model updated to minimize prediction errors and loss functions. The last thing is model evaluation validation. Trained with validation set pairs. The model was evaluated and validated. The predicted results are in good agreement with actual results, the differences between the placement points can evaluate the accuracy and performance of the model.

In conclusion, the CNN-SVM network model is not only accurate, the perfoemance is excellent, and its efficient solution speed makes it a launch ever ideal for process adapter drop point prediction.

Here is an another example on laser and optoelectronics. This research based on Support Vector Machine(SVM) algorithm, the PCI-SVM system was built. Through parameter optimization and model adjustment , the automatic classification of cavitation and crack defects was realized, and the classification accuracy reached 94.6%. SVM classification is achieved by converting raw training data into multi-dimensional space and constructing hyperplanes on high dimensions, and striving to be in high dimension. The eigenspace finds such a hyperlane that maximizes the spacing between different samples. The author uses the PCI-SVM system to complete the research. It was firstly uses the PCI algorithm to increase the weight of phase information, so as to realize high resolution identification and location of defects in pipes and plates, and makes the obtaining imaging results into picture datasets. Subsequently, SVM classifiers are deployed to train and classify these picture data, and the SVM model is constantly adjusted and optimized, so as to achieve efficient classification of cavitation and crack defects [7].

Moreover, SVM can also be used in judging the remaining time of fingerprints greatly improves the efficiency of police handling cases. In this research, SVM models were constructed by SG, SNV, SG+SNV preprocessing methods and feature wavelength extraction respectively, and the applicability and prediction effect of the models were

ISSN 2959-6157

compared under different conditions. The results show that the SVM mode, after hyperspectral imaging combined with proper preprocessing and feature extraction can be used to predict the residual time of latent sweat fingerprints on common indoor objects [8].

There is an another use of SVM which is about the land use, classification extraction and changes. This SVM algorithm mainly seekd a hyperplane between classes, which is the hyperplane that makes the maximun geometric interval between classes, so as to achieve more accurate classification purposes. The optimal classification function is obtained by solving the SVM mathematical formula. As a result, the land use types of this city were divided into cultivated land,m forest land, grassland, water area, construction land and unused land. Accuracy and selection of classification results. The selected classified samples are closely related to the test samples. In the selection of samples, high-resolution images are referred to, and the principles of uniform sample selection, obvious feature features and moderate quantity are followed. The separability of all sample is above 1.8 and the sample are qualified.

4. Conclusion

All research requires the support of a large amount of data. If the data collected through a series of experiments, questionnaires or other forms are not sorted out scientifically and efficiently, it will make the following experiments extremely difficult. Therefore, sorting and analyzing data has always been an inescapable problem for scholars. But fortunately, SVM has put forward an effective and powerful solution to this matter. Therefore, the author introduces the definition of SVM and its application methods as well as its application in practical research. It can be seen that this mathematical method is highly efficient in data processing. SVM can be used in any experiment that requires data classification and sorting, and even in several examples listed by the author in the paper. According to the data itself and the demand of the research results, it can be properly combined with other mathematical methods, which can not only greatly improve the efficiency of the research, but also lay a solid foundation for the authenticity of the conclusion. However, due to time reasons, there is no more time to collect more examples of SVM application in practical research such as chemistry, finance, art and so force. But examples from these three different fields can also reflect the convenience and efficiency of SVM. In future research, the author believe that SVM will be widely used and scholars will develop more efficient methods to process data. Let's look forward to the upgrade of SVM.

References

[1] Wang Chong, et al. Load forecasting method based on CNN and extended LSTM. Energy Reports, 2024, 12: 2452–61.
[2] Han hu. Support Vector Machine based on rough Sets and its application research. Lanzhou Jiaotong University, 2011.
[3] Bai, Yilan, et al. Parachute Deployment Fault Identification and Landing Point Prediction of Spacecraft Recovery System Based on Swin Transformer and Edge Detection. 2022 5th International Conference on Mechatronics, Robotics and Automation (ICMRA), IEEE, 2022: 90–95.
[4] Zeng jingjing. Generalization performance of support vector machine in big data environment. Hubei University, 2021.
[5] Liang Liming, et al. Sparse Representation for Kernel Function Selection of Support Vector Machine. International Journal of Autonomous and Adaptive Communications Systems, 2018, 11(1): 83.
[6] Huang huajuan. Research on Key problems of twin support vector machine. China University of Mining and Technology, 2014.
[7] Li ling and Wu mingdong. Research on classification extraction and change of land use in Yanting County based on support vector machine model. Science and Innovation. 2024, 16:148-150.
[8] Do, Thanh-Nghi. Training Neural Networks on Top of Support Vector Machine Models for Classifying Fingerprint Images. SN Computer Science, 2021, 2(5): 355.