New Media Public Relations Strategy Model Based on Support Vector Machine Algorithm

Jialin Li

790693032@qq.com

Vanderbilt University, Nashville city, USA

Keywords: The rise of new media is the result of worldwide scientific and technological development and innovation. The public relations of new media are deeply influenced by scientific and technological innovation, and the focus of public relations has gone through a process of change from management to service to relationship. In the new media era, interaction has increasingly become the focus of attention to promote and improve the quality of public relations. In this paper, the concept of support vector machine is given first, and the basic principle and research status of SVM technology are described in detail. The method of vector machine algorithm for multi-pattern classification is studied, and it is found that too many comparisons are the main reason for the large amount of calculation of this method. Through theoretical analysis and experimental results of data classification, it is shown that compared with the traditional classifier, the new method can significantly reduce the number of machine training and recognition and improve the algorithm running speed without affecting the classification accuracy. Break through the paradigm crisis that the public relations discipline is currently facing and meet its theoretical innovation needs.

Keywords: Support vector machine algorithm; New media; Public relations

1 Introduction

In recent years, with the progress of science and technology and the rapid development of various emerging communication technologies, new media has attracted more and more attention, driven the communication industry market and caused the readjustment of the industrial structure of the communication industry, which has also become a hot topic of discussion [8]. Multimedia technology is the comprehensive use of text, graphics, images, audio and video and other information communication media. It provides technical support for the input, output, transmission, storage and processing of various media [1]. New media, mainly based on the Internet and mobile phones, including blogs, podcasts, RSS, mobile text messages, mobile newspapers and other forms, have become the mainstream media in today's society [6]. At present, the traditional multimedia technology is facing new challenges. In intelligent multimedia, there are four basic information processing methods: information representation, information detection and classification, information fusion and information conversion and synchronization.

ICIDC 2022, June 17-19, Qingdao, People's Republic of China

DOI 10.4108/eai.17-6-2022.2322694

However, the theoretical research of public relations has developed rapidly in China, but the real watershed was from the late 1970s to the early 1980s. Since then, domestic scholars have made a lot of useful explorations and researches on it. Although the academic research in the field of public relations in China is rich and has made great progress, it is mostly based on western public relations theories in the 20th century [3]. As a new learning machine, support vector machine is closely related to neural network. Especially when SVM uses S-shaped function as kernel function, it realizes a single hidden layer perceptron neural network, but the weights and hidden layer nodes of the network are automatically determined by the algorithm, so it can be expected that SVM will also play an important role in the above four basic information processing methods [2]. Vector model is an effective method to digitize documents by calculating the weights of keywords in documents. With the introduction of VSM, a large number of machine learning methods, such as support vector machine, naive Bayes method and maximum entropy, are widely used in new media public relations.

2 Support vector machine algorithm for new media public relations

2.1 Algorithms Based on Support Vector Machines

The Internet has become an important channel for information exchange, which not only breaks the strict boundary between communicators and recipients, but also rewrites the way of crisis public relations. Support vector machine is a statistical machine learning method. The "one-to-one" method, a multi-classification algorithm, only considers two types of samples at a time, that is, designs a binary classification vector machine model for every two types of samples in each pattern class, so it is necessary to design k binary classifiers to distinguish $k(k-1)/2$ categories. As shown in Figure 1, the basic structure of the algorithm is shown.

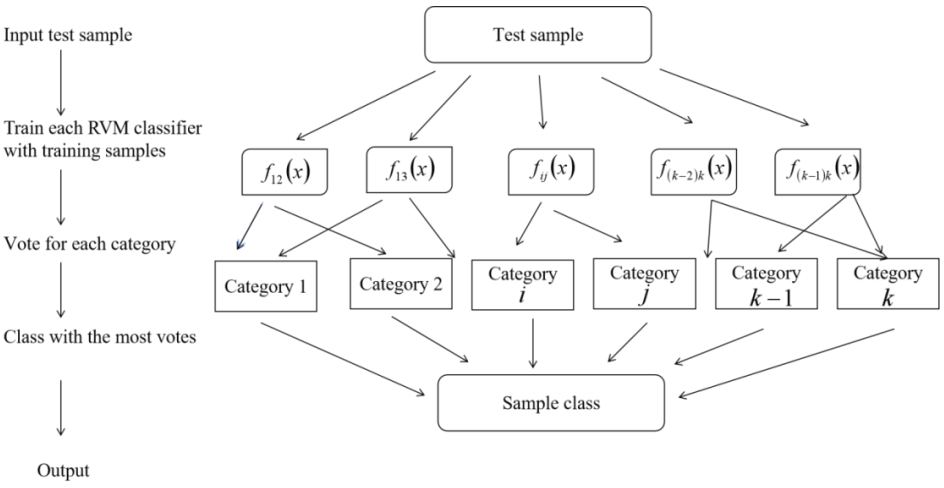

Figure 1 Basic structure of RVM algorithm

Suppose the classification function $f_{ij}(x)$ is used to discriminate i, two types of samples, if $f_{ij}(x)<0$, then the x belongs to the i class, and the i class gets one vote; otherwise, it is judged that the x belongs to the j class, and the j class gets one vote. Finally, when making a decision, Compare which category gets more votes, and classify the test samples into this category. Therefore, the classification accuracy of this method is high, and the only disadvantage is that the number of two classifiers is relatively large, and the operation speed is relatively slow. VSM converts document information into digitized vectors by calculating the weight of document feature items. For the binary classification problem, set the training sample set as the document set as $D=[D_1,D_2,\ldots,D_N]$. The feature vector of i document D_i is: $d_i=(t_{1,i},t_{2,i},\ldots,t_{n,i})$, $t_{j,i}$ represents the weight of the feature item D_i in the document t_j. The calculation methods of the four weights are:, and the definition expression of the classification function of the vector machine is:\

$$t_{ji}=\begin{cases}0, & tf_{ji}>0;\\ 1, & other.\end{cases} \tag{1}$$

If the word t_j appears in the document, its weight is 1, otherwise it is 0. For the binary classification problem, let the training sample set be $(x_n,t_n)(n=1,2,\ldots,N,x\in R^d,t\in\{0,1\}$, the classification function of the category label vector machine, and the expression is:

$$y(x,w)=\sum_{i=1}^{N}w_iK(x,x_i)+w_0, \tag{2}$$

where: $K(x,x_i)$ is the kernel function; w_i is the weight of the model. Both SVM and RVM classification methods obtain the weight w_i through a set of training samples. After multiple iterative updates, the corresponding w_i of most training samples are zero, and the training samples corresponding to non-zero w_i so are called support vectors or correlation vectors.

2.2 Improved Support Vector Algorithm for New Media Public Relations

There is a public relations department in the brand management department of an enterprise, and not all of its public relations departments have systems and personnel for crisis monitoring. An important point of corporate crisis public relations is that there are people and functional departments to solve the crisis. New media public relations should give full play to the characteristics and advantages of new media technology, avoid risks and defects, and focus on enhancing and improving "interaction" from communication channels and communication platforms [4]. And innovation, information gain is an effective selection and dimensionality reduction method. Its basic principle is to calculate the difference between the information entropy before and after the feature item appears in the text. According to the training data, calculate the information gain of each feature item and delete the information gain. Small items, the rest are sorted according to the information gain from large to small, assuming the feature item t, its information gain $IG(t)$ can be defined as:

$$IG(t) = H(D) - H(D|t) =$$

$$\sum_{d_i \in D} \left(P(d_i|t) \log \left(\frac{P(d_i|t)}{P(d_i)P(t)} \right) + P\left(d_i + \bar{t}\right) \log \left(\frac{P\left(d_i \middle| \bar{t}\right)}{P(d_i)P\left(\bar{t}\right)} \right) \right) \tag{3}$$

The larger the t information gain of the feature item, the higher the classification ability of the feature item; conversely, the smaller the classification contribution of the feature item. The classification algorithm is SVM classification algorithm, as shown in Table 1, the initial classification result table based on support vector machine.

Table 1 Initial classification results based on support vector machine

Features	Accuracy rate	Recall rate	F value
Overall situation	58.14%	58.18%	58.95%
Part	65.41%	64.23%	62.24%
Combination	72.21%	70.25%	72.31%

Multimedia information processing is not only a simple combination of text, audio and video information processing, but the mutual integration of these media information may produce better results. For example, the combination of speech and face recognition is more conducive to personal identification, which is one of the purposes of information fusion [9]. Another purpose of information fusion is to facilitate retrieval. For example, the features of audio and

video can be fused into a retrieval frame according to a certain time sequence, and the mutual index between audio and video can be realized [10]. However, the classical SVM does not output posterior probability information. At present, there are several methods to introduce posterior probability into SVM. The feature space is decomposed into sum, where orthogonal to the classification plane, the posterior probability that the test sample belongs to the first class is given by the following formula:

$$P(t=1|t,u)=\alpha_0(u)+\sum_{n=1}^{N}\alpha_n(u)\cos(nt) \tag{4}$$

$$P(t=1|x)=\frac{1}{1+\ell^{\left[k^*(-d_x+d_{sv})\right]}} \tag{5}$$

In the above formula and are the distances from the test sample and the support vector to the classification surface respectively, and a probability rejection mechanism is introduced at the same time, and only when it is greater than a certain threshold is it classified into the first category.

3 Simulation experiment and result analysis

Although public relations is a service management science from the practice of modern society, since the concept of modern public relations was introduced into China, it was once misunderstood by the general public as a synonym for vulgar relations or deliberate manipulation of public opinion [7], and it was also misunderstood by professionals in the field of news dissemination as undermining the principle of truth and objectivity of news [5]. In order to verify that the improved algorithm can really improve the speed of multi-category classification, we have done many experiments with a set of sample data. First, we need to construct an n dimension function, which can represent $f(x_{1,}x_2,x_3,\ldots x_n)=x_1^2+x_2^2+x_3^2+\ldots+x_n^2$. In this way, an n dimensional space composed of n characteristic parameters is formed. In this space, 400 groups of data are randomly generated as experimental samples, and the experimental samples are arranged into f classes according to the function value k from small to large. As shown in Table 2, experimental data table of classifier based on SVM algorithm.

Table 2 Classification experimental data based on SVM algorithm

Number of categories	"One-to-One" Taxonomy		Improve taxonomy	
	Run time/s	Accuracy/%	Run time/s	Accuracy/%
K=15	0.075	94.5	0.045	94.2
K=25	0.268	93.4	0.271	93.5

K=55	0.958	95.4	0.149	95.6
K=75	2.038	95.8	0.198	95.2
K=95	4.058	95.6	0.268	95.5

We can take a look at the comparison graph of the program running time before and after the improvement of the classifier in Figure 2.

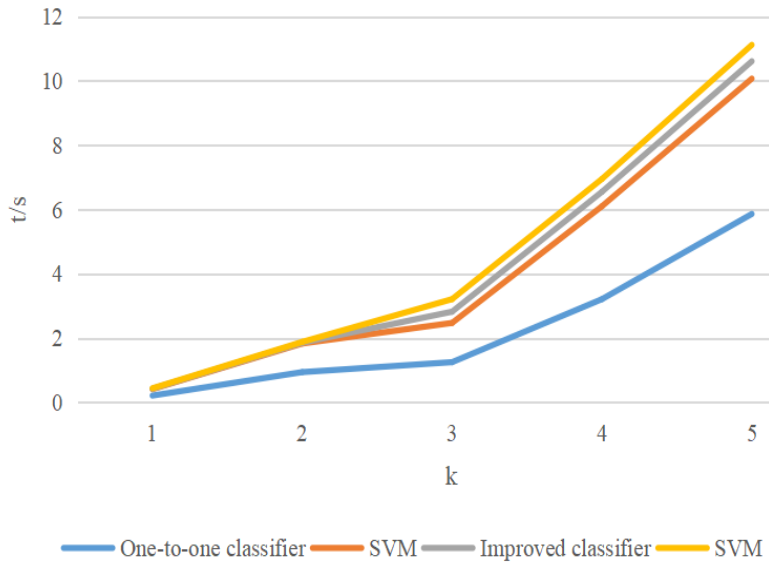

Figure 2 Comparison of program running time before and after classifier improvement

The above experimental conditions are unchanged, and 35 groups of data classification and recognition simulation experiments with different objective functions are used. In order to further verify the generalization of the improved method, several representative standard test functions are selected for classification experiments. $sqr(x)$: square root function, the range of experimental variables is $0 < x < 10;2)$ $\lg(x)$: common logarithm, the range of experimental variables is $0 < x < 10;3)$ $\exp(x)$: exponential function based on e, the range of experimental variables is $-5 < x < 5;4)$ $th(x)$ hyperbolic tangent function, and the range of experimental variables is $-1 < x < 1$. We can take a look at the comparison chart of the classification accuracy before and after the improvement of the standard test function classifier in Figure 3.

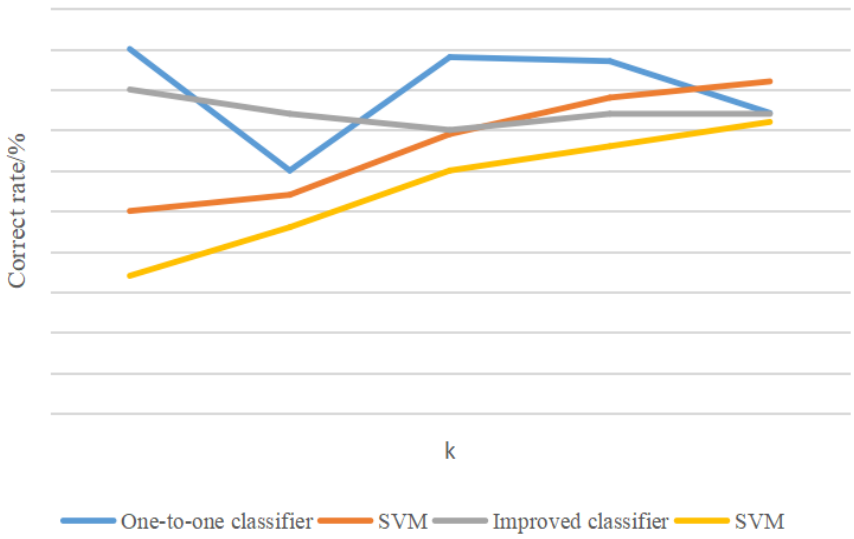

Figure 3 Comparison of classification accuracy before and after the improvement of the standard test function classifier

From the above figure, it can be seen that the running time of the improved algorithm is significantly shortened, and the more categories, the more obvious the improvement of this efficiency. After the improvement, the running time and the number of categories increased approximately linearly, and the increasing rate was much higher than before. When $k > 35$ is used, the difference between the improved new method and "one-to-one" classification accuracy is less than 1%. When the number of categories is large enough, the recognition accuracy of the two methods is almost the same. Therefore, this method can be popularized and can meet the needs of improving the real-time performance of pattern recognition software in different applications. At ordinary times, enterprises should conduct crisis public relations exercises to deepen employees' awareness of taking precautions, and they can respond in an orderly manner after the crisis, so as to enhance their crisis public relations ability. Through the routine emergency training of emergency handling, the abnormal change of enterprise leadership, the simulation of product quality injury crisis and other situations, the public relations team and employees of the enterprise can be urged to master the crisis response procedures, and the ability of employees to solve problems in crisis situations can be cultivated.

4 Conclusions

In the era of the rise of new media, under the impact and influence of technological change, modern public relations is facing a paradigm crisis in the theory of this discipline. The application of SVM is very extensive. Anything that can be transformed into classification, fitting and prediction problems can be solved by SVM. With the in-depth study of SVM, its application will be more extensive. This paper proposes a multi-class pattern recognition method that can be used for SVM or machine learning algorithms. This method makes new improvements on the basis of the existing generalization performance, classification accuracy

"one-to-one" voting method, and follows the "minimum" voting method. The method of eliminating votes by rounds" gradually approaches the correct solution. It is not only conducive to the improvement of the training speed of SVM, but also the classification performance of the system is also significantly improved. Simulation results verify the effectiveness and feasibility of the proposed algorithm. Under the premise that this method does not reduce the classification accuracy, the amount of calculation is significantly reduced, especially in the pattern recognition problem with many categories, this method significantly saves the running time of the program and improves the multi-class recognition. work efficiency. In practical applications with a large number of pattern categories, this method can greatly improve the efficiency of classification and recognition, and has certain application value.

References

[1] Fang Hanming, Shui Aishe, Wang Hui, et al. Dynamic Multi-Classification Method for Support Vector Machines [J]. Journal of the School of Logistics Engineering, 2017, 33(2):7.

[2] Huang Jing, Li Menghan, Kang Jinle, et al. Real-time mining and analysis of rainstorm disaster information based on social media—taking the "4.11 Shenzhen rainstorm" in 2019 as an example [J]. Water Conservancy Economy, 2021, 39(2): 10.

[3] Liu Zhe. The exploration and attempt of new media public relations in the new network environment [J]. Fenghui, 2019(11):2.

[4] Li Juwen. Research on data acquisition and analysis technology of convergent media platform based on ETL and SVM [J]. Electronic Design Engineering, 2021, 29(15):5.

[5] Liang Yiming, Zhao Yongyi. Research on fake news detection based on SVM [J]. Information and Computer, 2022, 34(2):3.

[6] Qiang Annie, Tao Xin. Technology Dependence, Tuning Strategies and Research Reflections on New Media Public Relations [J]. Communication and Copyright, 2021(9):45-48.

[7] Wang Yan, Wang Xuteng, Chang Xiaoyu, et al. Sentiment Analysis of Weibo Movie Reviews Based on Support Vector Machines: Taking "The Wandering Earth" as an Example [J]. Modern Film Technology, 2019(8): 5.

[8] Yin Xueting. Construction of new media news event recognition and extraction algorithm structure [J]. Electronic Technology and Software Engineering, 2018(8):1.

[9] Zhang Di, Luo Qun, Lei Min. A Copyright Protection Method Based on Image Texture Features and Support Vector Machines [J]. Network Security Technology and Application, 2017(8):3.

[10] Zheng Zhihao, Wu Wenbing, Chen Xin, et al. Traffic perception analysis system based on social media big data [J]. Journal of Automation, 2018, 44(4):11.