Journal of Computers Vol. 31 No. 4, 2020, pp. 198-210
doi:10.3966/199115992020083104015

Self-learning Consensus Voting Strategy and Its Application Based on Multiple Classification Algorithms

Yi-Qin Bao[1,4*], Qiang Zhao[2,3], Zhong-Xue Yang[1], Hao Zheng[1], Yin-tong Wang[1]

[1] College of Information Engineering of Nanjing XiaoZhuang University, Nanjing 2111711, Jiangsu, China
baoyiqin@njxzc.edu.cn

[2] Department of Information Systems Schulich School of Business, Toronto 416647, Canada
zhaoqiang@yahoo.cn

[3] Alibaba cloud university, hangzhou 310000, Zhejiang, China
392335241@qq.com

[4] Nanjing Suta Intelligent Technology co., LTD., Nanjing 210020, Jiangsu, China
358328014@qq.com

Received 10 March 2020; Revised 6 April 2020; Accepted 6 May 2020

Abstract. Machine learning classification algorithms are widely used, and each algorithm has its own characteristics. It is not straightforward to determine which classification algorithm should be used in different circumstances and how to improve the accuracy. This paper proposes a self-learning consensus voting strategy based on a variety of classification algorithms to improve the accuracy of classification. First, the classification model is trained using historical records, then the classification result is obtained by a self-learning consensus voting strategy, and finally the best classification result is attained. Through such algorithm fusion and voting, the practical problem of classification is suitably addressed. Based on a handwritten numeral dataset, this paper tests, analyzes, and compares the nine most common classification algorithms in machine learning, and makes a self-learning consensus vote on the classification result, which is verified by an intelligent handwritten numeral marking system. The system has gone through the steps of image acquisition, document upload, image preprocessing, classification algorithm learning, voting, answer comparison and scoring, etc., for examination paper autocorrection. The experimental results show that the recognition accuracy can be improved by more than 2% through the self-learning consensus voting strategy, and this voting method is universal and can be applied in practice.

Keywords: artificial intelligence, consensus voting algorithm, convolutional neural network, digit recognition, machine learning

1 Introduction

Machine learning is a subject that considers how to make computers take action under nonspecific programming conditions [1-2]. Over the past twenty years, machine learning has been used for autonomous cars, practical speech recognition, and efficient web search, and has greatly improved our ability to interpret human genes. Nowadays, machine learning technology is commonly applied. However, when to use what classification algorithm needs to be seriously considered. The main goal is to improve the result of algorithm application.

Machine learning algorithms are divided into four categories: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning [3]. Supervised learning is where the training data is labeled with the correct decision; unsupervised learning is where there is no such decision

* Corresponding Author

recorded. Semi-supervised learning is where some of the data has the decision recorded. Reinforcement learning is based on supervised learning and semi-supervised learning. The content of this paper is focused on supervised learning classification algorithms and their application.

The most commonly used machine learning classification algorithms are [4]: K-Nearest Neighbor (KNN) [5], Naïve Bayes (NB) [6], Logistic Regression (LR), Support Vector Machine (SVM) [7], Decision Tree (DT), Gradient Boost Decision Trees (GBDT) [8], Random Forests (RF), Artificial Neural Networks (ANN) [9-12], and Convolutional Neural Network (CNN) [13]. The authors in [14] gave a comparative study and application analysis of DT and NB. In [15] there is a study on DT, NB and the Minimum Distance algorithm, and [16] considers SVM, ANN, and CNN. However, these papers only compare a few algorithms, and there is no best algorithm determined. There has been much research on the application of classification algorithms to handwritten numeral recognition [17-27], but these only consider the specific application of numeral recognition, without considering the best approach.

According to the different characteristics of different classification algorithms and different applications, it is often difficult for people to choose the best approach as there are many algorithms and only a few algorithms are considered in a particular application. Common practice is as follows: 1) Choose a classification algorithm according to different application fields; 2) Compare different classification algorithms and choose the best one to make the final decision. However, although this has some utility, there is still room for improvement. AutoML (automated machine learning) is a system of learning and generalizing capabilities on given data and tasks, but this emphasizes ease of use. Due to the complexity of the learning process, the implementation efficiency needs to be improved.

Based on the above questions, this paper proposes a self-learning consensus voting strategy which integrates multiple classification algorithms. Firstly, different algorithms are trained on the same dataset, and the recognition accuracy is calculated via the historical data. Then, a self-learning consensus voting strategy is given, which integrates the results of multiple classification algorithms, and improves the recognition accuracy. Finally, the application of the system is verified, and the results are compared and analyzed. The results are satisfactory.

In Section 2, we introduce several classification algorithms related to machine learning, and compare these using datasets. In Section 3, we design a self-learning consensus voting strategy for the decision-making problem; in Section 4, we present the application of the self-learning consensus voting strategy in the intelligent handwritten numeral reading system. In Section 5, we test and compare three different classification strategies and finally we summarize the developments and findings in this paper.

2 Related Classification Algorithms

Supervised learning, where the dataset has input and output data, aims to find the corresponding relationship between the input and output, typically generating a function or model to perform this mapping. Classification is a core issue of supervised learning. If the output variable Y is a finite discrete value, the prediction problem is called a classification problem. In this case, the input variable X of the problem can be discrete or continuous. The classification model or function generated through training can also be called a classifier. The performance of a classifier is generally evaluated by the classification accuracy, that is, for a given test dataset, the ratio of the number of samples correctly classified by the classifier to the total number of samples. For classification problems, commonly used evaluation indexes include the precision and accuracy. Generally, the concerned classes are positive, and the others are negative. There are four cases in which the prediction of the classifier on the test set is correct or incorrect. They are defined as:

(1) TP - predicting the positive class as a positive class,
(2) FN - predicting the positive class as a negative class,
(3) FP - predicting the negative class as a positive class,
(4) TN - predicting the negative class as a negative class.

The definition of precision is then:

$$P = TP / (TP + FP) \tag{1}$$

The definition of accuracy is:

$$A = (TP + TN) / (TP + FN + FP + TN). \tag{2}$$

For supervised learning classification algorithms, in practical applications, the commonly used basis for selecting an algorithm is (2).

2.1 Characteristics of Classification Algorithms

(1) KNN is a mature method in theory and one of the simplest machine learning algorithms [28]. The idea of this method is: if most of the k most similar samples in the feature space belong to a certain category, then the test sample also belongs to this category. In the KNN algorithm, the selected neighbors are all correctly classified objects. In the classification decision-making, this method only depends on the category of the nearest sample or samples to determine the category of the samples to be classified.

KNN is simple, easy to understand, and easy to implement, with no need to estimate parameters or conduct training. It is suitable for rare event classification and is especially suitable for multiclassification problems where objects have multiple category labels.

(2) NB is a classification method based on Bayes theorem and independent hypothesis of characteristic conditions. It originates from classical mathematical theory and has a stable mathematical basis and classification efficiency [29]. The NB algorithm is a very simple classification algorithm. By determining the probability of each category for the given item to be classified, we can determine which category this item belongs to, and in the absence of redundant conditions, we will choose the category with the highest probability under known conditions.

NB has good performance for text classification and other domains. It is more often used in spam filtering classification.

(3) LR is a linear classifier that can calculate probabilities. Although logistic regression has “regression” in its name, it is actually a classification method, mainly used for binary classification problems [30]. Using a logistic function (or a sigmoid function), the range of an independent variable is (- INF, INF), and the range of the independent variable is (0, 1). It is often used in data mining, automatic disease diagnosis, economic prediction, and other fields. In essence, logical regression is designed for dichotomous problems, but it can also be used for multiclassification.

The LR calculation cost is not high and it is easy to understand and implement. Its disadvantage is that it is easy to underfit and the classification accuracy may not be optimal.

(4) SVM is a general linear classifier, which can minimize the empirical error and maximize the geometric edge area at the same time, so it is also called a maximum edge area classifier. At the same time, an SVM maps the vectors to a higher dimensional space, in which there is a maximum interval hyperplane, and there are two parallel hyperplanes on both sides of the hyperplane separating the data. The separating hyperplane maximizes the distance between the two parallel hyperplanes. The greater the distance or gap between the parallel hyperplanes, the smaller the total error of the classifier [31].

SVM have some problems, e.g. they are hard to train and explain, but they have excellent performance for nonlinear separable problems. SVMs are often chosen for such problems.

(5) DT is a decision-making analysis method, which is based on the known probability of occurrence of all kinds of situations. It obtains the probability that the expected value is greater than or equal to zero through the formation of a decision tree, evaluates project risk and judges its feasibility. It is a graphical method of intuitively using probability analysis [32]. DT is a prediction model, which represents a mapping relationship between object attributes and object values. It is a common method in data exploration. Each decision tree is a tree structure, which performs classification via its branches. Each decision tree can rely on the segmentation of the source database for data testing. This process can recursively prune the tree. The recursion is finished when there is no more segmentation or when a single class can be applied to a branch.

DT is easy to understand and implement. Data preparation is often simple or unnecessary. Continuous fields are difficult to predict, however. When there are too many categories, errors may increase.

(6) RF is a classifier with multiple decision trees, and the output category is determined by the mode of the output category of the individual trees [33]. RF uses bootstrap resampling to generate a new training sample set from the original training sample set n, and then generates K classification trees according to the generated sample set to form a random forest. The classification results for the new data are determined by the scores of the classification tree votes. Its essence is an improvement of the decision tree algorithm, which combines multiple decision trees together. The establishment of each tree depends on an independent sample. Each tree in the forest has the same distribution, and the classification error

depends on the classification ability of each tree and the correlation between them.

RF can use smaller samples, but also can achieve better accuracy.

(7) GBDT is composed of multiple decision trees and is similar to RFs. The difference is that: a) the composed tree can be generated in parallel; GBDT can only be generated in serial; b) for the final output, the random forest uses the majority of votes; c) GBDT accumulates all the results, or adds them by weight [34].

The GBDT algorithm is based on DT, which can use smaller samples and attain a better accuracy.

(8) ANN is a kind of mathematical model or calculation which imitates the structure and function of a biological neural network. It is used to estimate or approximate functions. The neural network is constructed by a large number of artificial neuron connections. In most cases, the artificial neural network can change the internal structure based on external information, resulting in an adaptive system. An artificial neural network is composed of many layers; the first layer is called the input layer, the last layer is called the output layer, and the middle layer is called the hidden layer. Each layer has many nodes, among which there are edges connected, and each edge has a weight. For text, the input value is every character, for pictures, the input value is every pixel [35].

ANN have shown excellent performance in various application domains such as voice, picture, video, games, etc., but there is a problem that a large amount of data is needed for training to improve the accuracy.

(9) CNN is one of the most representative neural networks in the field of deep learning. It has made many breakthroughs in the field of image analysis and processing. On the standard image annotation set ImageNet, commonly used in academia, CNN have had many achievements, including image feature extraction and classification, scene recognition, etc. Compared with the traditional image processing algorithms, one of the advantages of convolutional neural network is to avoid the complex pre-processing of images, especially the artificial participation in the image pre-processing process. Convolutional neural networks can directly input the original image for a series of work, which has been widely used in various image related applications [36-39].

CNN have the characteristics of sharing convolution kernels, high-dimensional data processing, with no need to manually select features or train weights, because it can achieve good feature classification performance. However, it needs parameter adjustment and a large sample size. It is best to use a GPU during training.

2.2 Accuracy of Classification Algorithms

2.2.1 Data Preparation

MNIST is a classic dataset for deep learning. It is composed of 60000 training images and 10000 test images. Each image is 28 * 28 in size. In this paper, 10000 training data and 1000 test data are used to form the total test set t, and 10% t to 100% t are taken from the T data sets to evaluate the sample set size. In data preparation, in order to achieve the universality of data training, the MNIST data is converted into a text file; the training set is saved in the train, txt file, and the test set is saved in the test.txt file. Then, the train.txt training sets are input into each classifier training model in turn; the trained model is used to test the test.txt test set, and finally the recognition accuracy of each model is output.

2.2.2 Recognition Accuracy

Accuracy is the basis for judging and selecting classification algorithms. In order to reflect the characteristics of various classification algorithms, common datasets are used for training and testing. In the experiment, 10000 training data and 1000 test data are used, and then 10% to 100% test set t are chosen from them. Nine classification algorithms are used: KNN, NB, LR, SVM, DT, RF, GBDT, ANN, CNN. The recognition accuracies are shown in Table 1.

Table 1. Comparison of the classification methods

Classification methods	Accuracy (%)									
	10%T	20%T	30%	40%T	50%T	60%T	70%T	80%T	90%T	100%T
NB	78.4	81.2	82.4	81.3	80.5	79.9	79.8	79.8	80.4	80.8
KNN	84.2	89.4	90.3	90.5	91.2	91.3	91.3	91.4	91.7	91.4
LR	90.3	93.1	92.4	90.5	90.4	88.6	88.6	88.4	89.3	89.5
RF	71.3	84.4	86.7	84.4	85.5	87.4	86.5	85.8	88.4	87.3
DT	69.4	76.1	73.9	72.9	75.4	74.8	77.1	77.6	78.9	79.2
SVM	85.5	92.4	91.8	91.4	90.4	89.5	89.2	89.5	89.4	89.6
GBDT	89.4	91.7	92.7	93.3	91.5	91.6	90.3	92.4	92.6	91.5
ANN	90.7	94.5	93.4	92.6	92.7	91.3	91.4	90.2	90.3	90.8
CNN	91.4	96.4	93.5	95.5	94.4	91.4	92.5	94.5	94.6	97.4

The data in Table 1 are graphed in Fig. 1.

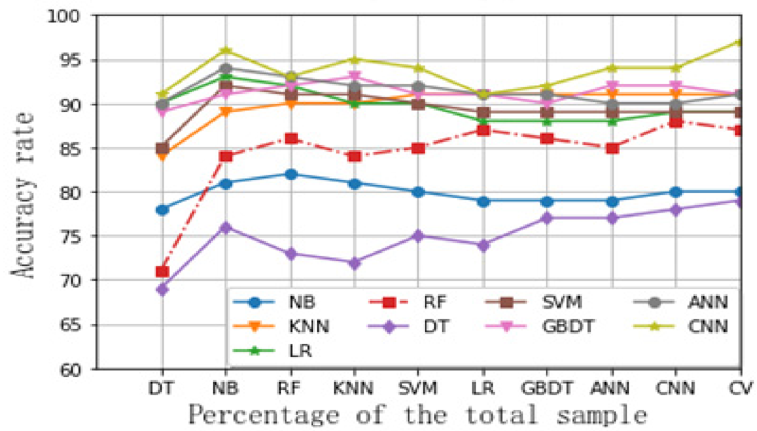

Fig. 1. Comparison of classification algorithms

From Table 1 and Fig. 1, it can be seen that the larger the dataset is, the higher the accuracy. This is determined by the characteristics of the above classification algorithms. When the test set is 100% T, the accuracy of each classification algorithm is relatively the highest.

3 Self-learning Consensus Voting Strategy

3.1 Definitions, Assumptions, and Concepts

3.1.1 Definitions

N: The total number of voting modules with functional equivalence, each voting module is {m1, m2, …, mN}.

m: The number of voting modules that agree with the output results.

r: Output result space cardinality.

3.1.2 Assumptions

(1) All voting modules are functionally equivalent, i.e., meet the same specifications. The output of each voting module is statistically independent.

(2) For a given voting module, there is only one correct output. The output result is identified as 1,…,r. An output result of 1 is the only correct output. An output of 2,…,r is an incorrect output.

(3) A given voting module will output a result; there is no possibility of no result output.

3.1.3 Concepts of Voting and Correct Results

The result of voting is different from the correct result. The voter (also known as voting module) cannot ensure the correct identification of the correct and incorrect result, so it can only output the result of voting as the correct result. The voting process of the majority voter is shown in Fig. 2. For a majority voter, the correct output results can be divided into two situations: in the first case, the majority modules produce the correct output results and vote through, as shown in Fig. 2, correction and agreement (CAA); in the second case, the multiple modules produce incorrect results but also vote through, as shown in Fig. 2, incorrection and agreement (IAA). The same incorrect output results can be divided into two cases: the first case is the correct result because the result is not output by the majority of modules and is rejected in the voting, that is, the cotractness and failure (CAF) in Fig. 2; the second case is that the incorrect result does not pass the voting, that is, incorrect and failure (IAF).

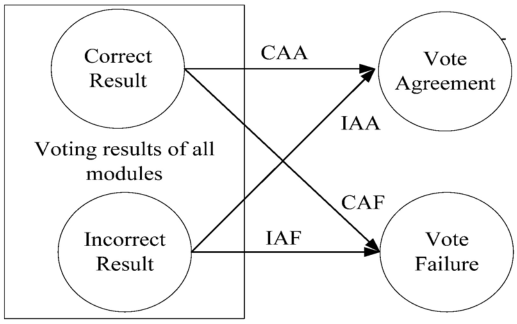

Fig. 2. Voting process of the majority voter

We try our best to make a correct result pass the voting, while an incorrect result is rejected in the voting. We also try our best to avoid the situation where a correct result fails to pass the voting and an incorrect result passes the voting.

Therefore, when the relative independent classifiers have different results, the correct results can be achieved by voting algorithm.

3.2 Consensus Voting Algorithm

A consensus voting algorithm [40] is an extension of a majority voting algorithm, which was proposed by McAllister et al in 1990. The consensus voting algorithm is described as follows:

(1) If the output result is mostly agreed, that is, m> (N + 1) / 2, N > 1, then select this result as the output.

(2) If there is a unique maximum number of endorsements, and the number of endorsements is less than (N + 1) / 2, then select this result as the output.

(3) If several outputs have the same maximum approval number, then randomly select a set of results as the output of the program.

(4) In other cases, the voting is invalid, the voting is stopped, and the voting is turned into an invalid state.

The consensus voting algorithm is highly effective for voting with a small output space, because it can make the voting automatically adapt to the change of effective output space base. The problem with the consensus voting algorithm is that it may output the wrong result as the correct result. Consider the third step of the algorithm: if several outputs have the same maximum approval number, then one of these groups is randomly selected as the output. In this way, it is possible to output incorrect results, thus reducing the accuracy of recognition.

3.3 Self-learning Consensus Voting Algorithm

The self-learning consensus voting algorithm is a kind of consensus voting algorithm based on historical information. The core step of this algorithm is to build the recent history record information of each voting module. A common way to build module history is to accumulate the number of votes passed by each module [41]. In this method, if the calculation result of the module is voted, the number of times of the module is recorded. In a certain voting cycle, the module with the largest cumulative times can be considered as the most reliable module, and the module with the smallest cumulative times can be considered as the module with the lowest reliability. Through learning from this historical information, self-learning can be achieved.

3.3.1 Self-learning State

The strategy of building the self-learning state is defined as follows:

(1) Build a set of N voting modules $\{m_1, m_2, \ldots, m_N\}$. If an output result passes the majority vote in the voting process, the module corresponding to this output is m = 1, otherwise, set the parameter m = 0. Therefore, m = 1 means that module passed the vote in this vote.

(2) The value of m is accumulated in each voting period, and H(i)=sum(m_{ij})), j=1, ..., n. H(i) is obtained in n times of system operation. This value represents the number of times module i passed the voting.

(3) The history H(i) is normalized, that is, P(i) = H(i) / n. Then, this value can be considered as the reference value of the accuracy of module i. The highest P(i) module i is the module most likely to produce the correct results. In addition, in any voting cycle, Pi(j) represents a state of module i from the first voting to this voting cycle j, so this value can be considered as a "self-learning state" of module i in cycle j. Hi(j) represents the history of module i from the first voting to this voting cycle j. The historical information of the module may be used for different purposes:

(1) For offline use of historical records, first count P(i) as input, and then vote;

(2) For online use of historical records, count P(i) for voting when the system is running, and then input the identification results as the next task.

The whole process of using historical data is a self-learning process.

3.3.2 Self-learning Consensus Voting

The self-learning consensus voting algorithm is improved on the basis of the consensus voting algorithm. The main idea is to use the statistical historical information to calculate the self-learning state. The algorithm is described as follows:

(1) If the output result is mostly agreed, that is, m> (N + 1) / 2, N > 1, then select this result as the output. Moreover, this output is used as the next input to recalculate P(i).

(2) If there is a unique maximum number of endorsements, and the number of endorsements is less than (N + 1) / 2, and the learning state P(i) of the corresponding module is greater than the threshold T, then this result is selected as the output. The threshold T can be set according to the actual situation of the user, for example, T = 0.95, P(i) > 0.95, indicating the voting result of the voting module i is selected. Moreover, this output is used as the next input to recalculate P(i).

(3) If several outputs have the same maximum approval number and the learning state P(i) of the corresponding module is greater than the threshold T, then this result is selected as the output. For example, m1 = m2, T= 0.95, P(2) > T, indicating the voting result of voting module 2. Moreover, this output is used as the next input to recalculate P(i).

(4) In other cases, voting is invalid, stop and transfer to the invalid state.

4 Application of the Self-learning Consensus Voting Strategy

4.1 Handwritten Digit Intelligent System

This system integrates the self-learning consensus voting strategy of nine classification algorithms and considers the specific application to an intelligent marking system. First, the original image is collected by taking photos of mobile phone, and then the collected original image is transferred to a computer for preprocessing through a wireless network. The processed data is then put into the pre-trained model for recognition. Finally, the identification results are uploaded and compared with the correct answers to judge whether this is right or wrong, so as to realize intelligent marking functionality.

The system is divided into five modules:

A. System interface module: through the front-end design, the system functions are integrated and typeset into a team website, login interface, score table interface, mobile phone interface, etc.

B. Original image acquisition module: after students answer questions, use the mobile app to extract test paper images completed by students.

C. Image preprocessing module: through a system of image processing methods, extract the numbers in the picture box and preprocess the characters in a standardized way.

D. Recognition and answer module: the processed binary matrix input is integrated with 9 classification algorithms, and the self-learning consensus voting algorithm is used for recognition. The output results are compared with the correct answers.

E. Integration data module: integrate the identified answers and corresponding student information into the database and system.

The handwritten digit intelligent marking system employs a mobile application (app), which uses the camera to obtain the image of the test paper and upload it to the PC. The mobile app interface is shown in Fig. 3, and the whole test paper is uploaded to the computer for recognition.

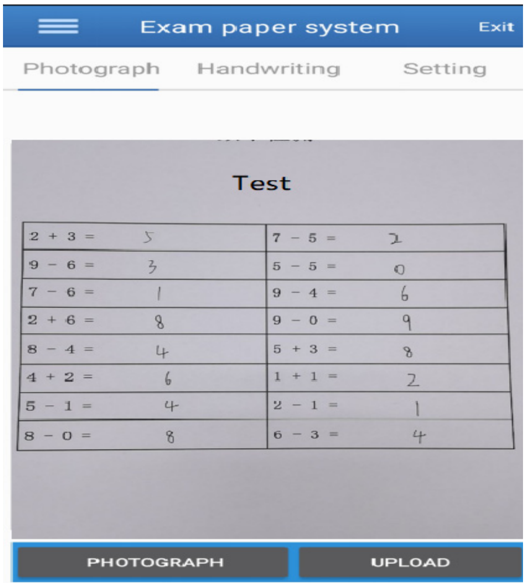

Fig. 3. Mobile app interface

4.2 Recognition Process

KNN, NB, LR, SVM, DT, RF, GBDT, ANN and CNN classification algorithms are used as voting modules in the identification process. First, the training dataset train.txt and test.txt are used as the data source, and then the recognition accuracy is calculated via machine learning as a self-learning state. Then through the mobile phone photo, the PC receives the image which is then preprocessed and the numbers are extracted and saved in the PNG file format. After adjustment, the PNG file is converted into a binary 28 * 28 matrix, and converted into a general test file. Finally, the self-learning consensus voting method is run using the nine classification algorithms. The numbers are identified and the results are output. The output file has two uses:

(1) To compare the actual classification and output;

(2) It is used for the data of the next training dataset and test dataset.

The whole identification process is shown in Fig. 4.

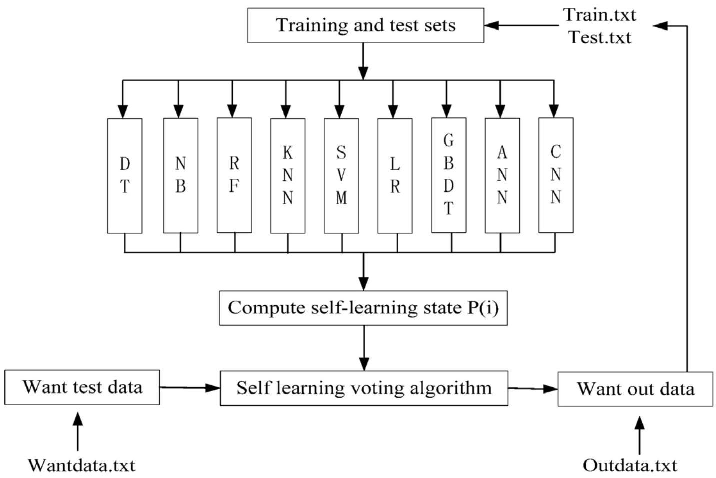

Fig. 4. Handwritten digit recognition process

4.3 Implementation of Self-learning Voting

The intelligent paper marking system for handwritten numerals is initialized and defined before recognition. The process is:

(1) Initialization: N = 9, r = 10;

Voting modules: KNN, NB, LR, SVM, DT, RF, GBDT, ANN, CNN. Defined as {m1, m2, …, m9}.

Initially, take the accuracy of the classification algorithm from Table 1 (using the accuracy of 100% T), round it, and then determine the self-learning state, as shown in Table 2.

Table 2. Self-learning state

Classification	Voting module	Self-learning state
DT	P(1)	0.79
NB	P(2)	0.82
RF	P(3)	0.88
KNN	P(4)	0.91
SVM	P(5)	0.92
LR	P(6)	0.93
GBDT	P(7)	0.93
ANN	P(8)	0.94
CNN	P(9)	0.97

(2) If the output result is mostly agreed, that is, m> 5, then select this result as the output.

(3) If there is a unique maximum number of endorsements, and the number of endorsements is less than 5, and the learning state P(i) of the corresponding module is greater than the threshold value T = 0.93, select this result as the output.

(4) If several outputs have the same maximum approval number, and the learning state P(i) of the corresponding module is greater than the threshold T = 0.93, then select this result as the output.

(5) In other cases, P(9) was taken as the result of voting (some adjustments were made in this area according to the actual situation).

Implementation result: for the test paper as shown in Fig. 3, run the identification process, and the interface is as shown in Fig. 5.

```
connected from ("192.168.1.123",43467)
          《Test Results》
====================================
2+3=5     OK          7-5=2     OK
9-6=3     OK          6-5=0     ERR
7-6=1     OK          9-4=6     ERR
2+6=8     OK          9-0=9     OK
9-4=4     ERR         6+3=8     ERR
4+2=6     OK          1+1=2     OK
5-1=4     OK          2-1=1     OK
8-0=8     OK          6-3=4     ERR
====================================
OK=11     ERR=5
waiting for connection ...
```

Fig. 5. Handwritten test paper recognition results

As can be seen from Fig. 5, the recognition result is ideal, with a recognition rate of 100%. Of course, this is just a simple example.

5 Testing and Comparison

The most commonly-used strategy for digit recognition is to select the most accurate classification algorithm. This method is defined as maximum voting (MV). We compare MV with consensus voting (CV) and self-learning consensus voting (SCV) in the intelligent paper marking system for the purpose of ensuring the availability of the system in practical applications. In the test, 100 real test papers were collected and photographed by 10 people. The numbers in each paper were identified by the humans first. After preprocessing, the image data is finally converted into the wantdata.txt file. After the system identification process such as algorithm operation, the test results in outdata.txt are finally obtained.

The comparison results regarding recognition accuracy are shown in Table 3. Here, MV is CNN, MV = 97.4, CV = 98.2%, and SCV = 99.5%. SCV shows a good improvement, two percentage points higher than CNN. The results show that SCV has better performance.

Table 3. Comparison of accuracy

Method	Accuracy (%)
Maximum voting (MV)	97.4
Consensus voting (CV)	98.2
Self-learning consensus voting (SCV)	99.5

In this experiment, all algorithms use the same training samples and test samples,This is so that the algorithms can be compared fairly. The self-learning consensus voting method has the highest recognition accuracy. Combined with all classification algorithms, the recognition accuracy line chart is shown in Fig. 6.

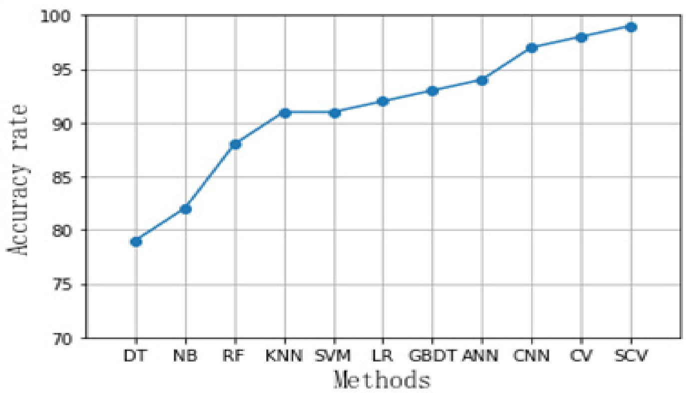

Fig. 6. Comparison of accuracy

The following conclusions can be drawn from Fig. 6:

(1) The algorithms are arranged from low to high according to recognition accuracy: DT, NB, RF, KNN, SVM, LR, GBDT, ANN, CNN. CNN is considered to be the best algorithm for handwritten digit recognition among the commonly-used algorithms.

(2) On the basis of consensus voting, by adding historical information, the accuracy of recognition is further improved by the self-learning consensus voting method.

6 Conclusions

This paper analyzes the principles of the classification algorithms in machine learning, compares the recognition accuracy of the main algorithms for handwritten data, and finally proposes the self-learning consensus voting algorithm which is applied in the handwritten digit intelligent marking system. The self-learning consensus voting algorithm can also be used in other classification scenarios. It is universal, intelligent, and shows a significant improvement in recognition accuracy. Moreover, it reduces the amount of time required to select and understand appropriate algorithms. At present, in the application of digit recognition, Paperless examination is not popular, so a handwritten digit examination system is needed. There are still many problems to be solved in the practical application of a perfect digit recognition system. For example, the efficiency can be improved through distributed operation, and the self-learning consensus voting algorithm of the general fusion classification method only addresses one problem.

Acknowledgements

This work is supported by the Jiangsu natural science foundation project (BK20180142) and Jiangsu “six talent peaks” project (rjfw-038).

References

[1] P. Harrington, Machine Learning in Action, Manning Publications Co., Shelter Island, NY, 2012.

[2] T.M Mitchell, J.G. Carbonell, R.S. Michalski, Machine Learning, McGraw-Hill, New York, NY, 2003.

[3] W. Jue, C.-Y. Shi, Machine learning research, Journal of Guangxi Normal University Natural Science 21(2)(2003) 1-15.

[4] X.-X. Xu, F.-Z. Li, L. Zhang, Z. Zhang, The category representation of machine learning algorithm. Journal of Computer Research and Development 54(11)(2017) 2567-2575.

[5] H.-Y. Meng, Z. Liu, F. Wang, J.-D. Xu, G.-Q. Zhang, An efficient collaborative filtering algorithm based on graph model and improved KNN, Journal of Computer Research and Development 54(7)(2017) 1426-1438.

[6] J.-H. Wang, W.-L. Zuo, Z. Yan, Word semantic similarity measurement based on Naïve Bayes model, Journal of Computer Research and Development 52(7)(2015) 1499-1509.

[7] K. Xiao, S. Dai, Y.-H. He, L.-M. Sun, Chinese text extraction method of natural scene images based on city monitoring, Journal of Computer Research and Development 56(7)(2019) 1525-1533.

[8] J.-S. Liu, B.-W. Liu, M. Zhang, Q. Liu, Fault prediction of power metering equipment based on GBDT, Computer Science 46(6A)(2019) 392-396.

[9] G.-L. Chen, S. Ma, Y. Guo, Survey on Accelerating Neural Network with Hardware, Journal of Computer Research and Development 56(2)(2019) 240-253.

[10] D. Han, S.-Y. Zhou, T. Zhi, Y.-J. Chen, T.-S. Chen, A survey of artificial intelligence chip, Journal of Computer Research and Development 56(1)(2019) 7-22.

[11] J. Guo, Method of predicting performance of storage system based on improved artificial neural network, Computer Science 46(6A)(2019) 52-55.

[12] S.-W. Liu, J.-M. Sun, X. Gao, M. Wang, Analysis and establishment of drilling speed prediction model for drilling machinery based on artificial neural networks, Computer Science 46(6A)(2019) 605-608.

[13] W.-H. Shi, J. Meng, P. Zhang, C.-J. Liu, Prediction of miRNA-lncRNA interaction by combining CNN and Bi-LSTM, Journal of Computer Research and Development 56(8)(2019) 1652-1660.

[14] B.N. Lakshmi, T.S. Indumathi, R. Nandini, A comparative study of classification algorithms for risk prediction in pregnancy, in: Proc. TENCON 2015-2015 IEEE Region 10 Conference, 2015.

[15] L.-J. Zhu; S. Liu, An experimental comparative study on three classification algorithms on unknown malicious code identification, in: Proc. 2011 International Conference on Multimedia Technology, 2011.

[16] K. Xiao, S. Dai, Y.-H. He, L.-M. Sun, Chinese text extraction method of natural scene images based on city monitoring, Journal of Computer Research and Development 56(7)(2019) 1525-1533.

[17] D. Cao, C.-X. Yang, W. Zhang, Research progress of handwritten digit recognition, Computer Knowledge and Technology: Academic Exchange 5(3)(2009) 688-688.

[18] X. Han, Q. Liu, J. Xu, H.-Y. Chen, Handwritten numeral recognition algorithm based on similar principal component analysis, Computer Science 45(11A)(2018) 278-281, 307.

[19] W.-L. Jiang, W. Wang, Z.-X. Sun, Fast online handwritten digit recognition technology based on decision tree, Computer Science 33(9)(2006) 207-210.

[20] E.-Z, Shi. KNN fast handwritten Arabic digit recognition based on vector projection, Bulletin of Science and Technology (12)(2013) 127-129.

[21] F.-F. Wang, Research and application based on improved convolutional neural network algorithm, [dissertation] Nanjing, China: Nanjing University of Posts and Telecommunications, 2016.

[22] L. Wang, Review of offline handwritten digit recognition, Journal of Yunnan University For Nationalities Natural Science 22(S1)(2013) 67-71.

[23] Z. Wang, M.-T. Gao, Design and implementation of image recognition algorithm based on convolutional neural network, Modern Computer: Universal Version 20(2015) 61-66.

[24] L.-L. Wu, Research on handwritten digit recognition based on SVM, [dissertation] Shandong, China: Shandong Normal University, 2006.

[25] Z. Wu, G.-L. Zhu, G.-F. Huang, Handwritten digit recognition method based on image recognition technology, Computer Technology and Development 21(12)(2011) 48-51.

[26] K. Xu, Research on the application of convolution neural network in image recognition, [dissertation] Zhejiang, China: Zhejiang University, 2012.

[27] X. Zhang, The prospects and difficulties of handwritten digit recognition, Digital World (1)(2016) 7-7.

[28] Z.-M. Bao, S.-R. Gong, S. Zhong, R. Yan, X.-H. Dai, Person re-identification algorithm based on bidirectional KNN ranking optimization, Computer Science 46(11)(2019) 267-271.

[29] Q.-S. Wang, R.-Y. Wei, Bayesian Chinese spam filtering method based on phrases, Computer Science 43(4)(2016) 256-259, 269.

[30] S.-X. Zhang, W.-G. Li, T. Li, N.-N. Zhu, Improved CoreSets construction algorithm for Bayesian logistic regression, Computer Science 46(11A)(2019) 98-102.

[31] T. Li, X.-C. Liu, S. Zhang, K. Wang, Y.-L. Yang, Parallel support vector machine training with hybrid programming model, Journal of Computer Research and Development 52(5)(2015) 1098-1108.

[32] Y.-M. Ji, Y.-P. Zhang, X.-B. Lang, D.-C. Zhang, R.-H. Wang, Parallel of decision tree classification algorithm for stream data, Journal of Computer Research and Development 54(9)(2017) 1945-1957.

[33] B.-B. Zhang, J. Wang, K. Yue, H. Wu, J. Hao, Performance prediction and configuration optimization of virtual machines based on random forest , Computer Science 46(9)(2019) 85-92.

[34] H.-Y. Li, K. Qin, J.-Z. Li, W. Luo, K.-M. Zhao, A target threat assessment model based on improved iterative decision tree algorithm, Naval Electronic Engineering 37(10)(2017) 25-29.

[35] S.-W. Liu, J.-M. Sun, X. Gao, M. Wang, Analysis and establishment of drilling speed prediction model for drilling machinery based on artificial neural networks, Computer Science 46(6A)(2019) 605-608.

[36] W.-H. Shi, J. Meng, P. Zhang, C.-J. Liu, Prediction of miRNA-lncRNA interaction by combining CNN and Bi-LSTM, Journal of Computer Research and Development 56(8)(2019) 1652-1660.

[37] C.-W. Tian, Y. Xu, W.-M. Zuo, Image denoising using deep CNN with batch renormalization, Neural Networks 121(2020) 461-473.

[38] D. Yuan, N.-N. Fan, Z.-Y. He, Learning target-focusing convolutional regression model for visual object tracking, Knowledge-Based Systems 194(2020) 105526.

[39] D. Yuan, X. Li, Z. He, Q. Liu, S. Lu, Visual object tracking with adaptive structural convolutional network, Knowledge Based Systems 194(2020) 105554.

[40] D.F. McAllister, C.-E. Sun, M.A. Vouk. Reliability of voting in fault-tolerant software systems for small output-spaces, IEEE Log 39(5)(1990) 524-534.

[41] C. Latif-Shabgahi, S. Bennett, Adaptive majority voter: a novel voting algorithm for real-time fault-tolerant control systems, in: Proc. 25th EUROMICRO Conference, 1999.