ISSN: 2472-3703 DOI: 10.6911/WSRJ.201911_5(11).0021

Wine Classification Method Based on One-Dimensional Convolutional Neural Network

Leilei Chen[1, a], Jie Yang[1, b], Linlin Chen[1, c], Qiupu Wu[1, d] and Xingzhong Xiong[1, e, *]

[1]Artificial Intelligence Key Laboratory of Sichuan Province, Automation and Information Engineering, Sichuan University of Science & Engineering, Zigong, China, 643000;China

[a]19915493873@163.com, [b]yangjie100894@163.com, [c]1274092661@qq.com, [d]1074540172@qq.com, [e]2554438@qq.com

Abstract

A wine classification model based on one-dimensional convolutional neural network (1DCNN) is proposed for wine classification. The experimental data uses the wine dataset of the UCI machine learning library. Simulation experiments show that the wine classification model based on one-dimensional convolutional neural network has higher classification accuracy than BP neural network and extreme learning machine (ELM) classification model.

Keywords

One-dimensional convolutional neural network; BP neural network; extreme learning machine; wine classification.

1. INTRODUCTION

As people's living standards continue to increase, the demand for wine is also increasing. Because different wines have different qualities due to different materials and brewing methods, the quality of wine has always been the most concerned issue for consumers and producers. Especially for medium and high-end wines, the quality of wine is almost determined. Its value is high [1]. The traditional wine classification method mainly adopts the sensory analysis method, but this method is susceptible to the hobby and experience of the wine appraisers, and usually makes the evaluation subjective and uncertain [2]. Because wine contains many different physical and chemical components, it can be classified and identified by its physical and chemical data components. Therefore, some scholars have studied wine classification methods based on data mining. Literature [3] proposed a method for classifying wine types using fuzzy clustering. Literature [4] proposed a wine classification method based on principal component analysis (PCA) and particle swarm optimization-support vector machine (PSO-SVM), which improves the accuracy of wine classification in the case of limited training samples. . In the literature [5], for the redundancy of wine physics and chemical data components, two wine classification algorithms, Principal Component Analysis K-means and Principal Component Analysis Self-organizing Neural Network, are proposed. Both classification algorithms have high accuracy. Literature [4-5] used principal component analysis to reduce the input dimension and solved the linear redundancy between wine physicochemical composition data, but the proposed algorithm is cumbersome for the data processing steps. Existing algorithms have many shortcomings in wine classification. Therefore, this paper proposes a wine classification algorithm based on one-dimensional convolutional neural network for the accuracy of existing quality classification in wine quality, through one-dimensional convolutional nerve. The characteristics of the network wine classification model automatically extracts the physicochemical composition characteristics of wine to realize the

ISSN: 2472-3703 DOI: 10.6911/WSRJ.201911_5(11).0021

classification of wine. The algorithm has higher classification accuracy than the traditional neural network wine quality classification algorithm.

2. BASIC THEORY OF CONVOLUTIONAL NEURAL NETWORKS

As one of the classical algorithms for deep learning, convolutional neural networks have better performance in various signal and information processing tasks than standard fully connected neural networks. The main advantage of convolutional neural networks is that their local links, weight sharing and pooling operations effectively reduce the complexity of the network, reduce the number of training parameters, and make the model have a certain degree of invariance to translation, distortion, and scaling. It has strong robustness and fault tolerance, and is easy to train and optimize [6].

2.1. Convolutional Layer

A typical convolutional neural network structure consists of an input layer, a convolutional layer, a pooled layer, a fully connected layer, and an output layer. In a convolutional layer, the feature map of the upper layer is convolved by a learnable convolution kernel, and then an activation function is used to obtain the output feature map of the next layer. Each output feature map may be combined and convolved multiple times. Enter the value of the feature map. The input to the convolutional neural network is X, H^l represents the feature $maps(H^0 = X)$ of the lth layer of the convolutional neural network. Suppose H_j^l is the jth feature map of layer lth, The production process of H_j^l is as in formula (1):

$$H_j^l = f(\sum_{i=1}^{k} H_i^{l-1} \times W_{ij}^{(l)} + b_j^l) \tag{1}$$

Where W_{ij}^l denotes a weight matrix connecting the ith feature map of the $l-1$ layer and the jth feature map of the l layer, the operator $\times$ represents a convolution. H_i^{l-1} represents the ith feature map of the $l-1$ layer. i and j are the indices of the input and output feature maps. b_j^l represents the corresponding offset of each output feature map of layer i. Finally, the jth feature map H_j^l of the l layer is obtained by an activation function $f(\cdot)$. Commonly used activation functions are $\mathrm{Re}LU$ functions, sigmoid functions, functions, radial basis functions, etc. This paper uses a function, the form is as in formula (2):

$$y = \max(0, x) \tag{2}$$

That is, when x<0, the output value is 0, and when x>0, the output value is x .

2.2. Full Connection Layer

In a convolutional neural network, after alternating convolutional and pooling layers, one or more fully connected layers are connected. Each neuron in the fully connected layer is fully connected to all neurons in its previous layer. The fully connected layer can integrate local information with class discrimination in the convolutional layer or the pooled layer [7]. In order to improve the performance of the convolutional neural network, the activation function of each neuron in the fully connected layer generally adopts the $\mathrm{Re}LU$ function. The output value of the last layer of the fully connected layer is passed to an output layer, which can be classified

ISSN: 2472-3703 DOI: 10.6911/WSRJ.201911_5(11).0021

using the softmax function. For a specific task, it is important to choose a suitable loss function. The representative loss function is in addition to the softmax loss function, as well as the Hinge loss function, the Contrastive loss function, the Triplet loss function, and KL Divergence [8].

3. WINE CLASSIFICATION METHOD FRAMEWORK BASED ON ONE-DIMENSIONAL CONVOLUTIONAL NEURAL NETWORK

The physicochemical composition of wine can be regarded as one-dimensional sequence data, so it is more reasonable to use one-dimensional convolutional neural network. This paper proposes a wine classification model based on one-dimensional convolutional neural network. The flow chart is shown in Figure 1. The wine data set is divided into a training set and a test set. Then the training set is trained by the forward propagation and the minimized error back propagation for the one-dimensional convolutional neural network model. The test set is used to evaluate the one-dimensional convolutional neural network model. Generalization ability to get the correct classification results.

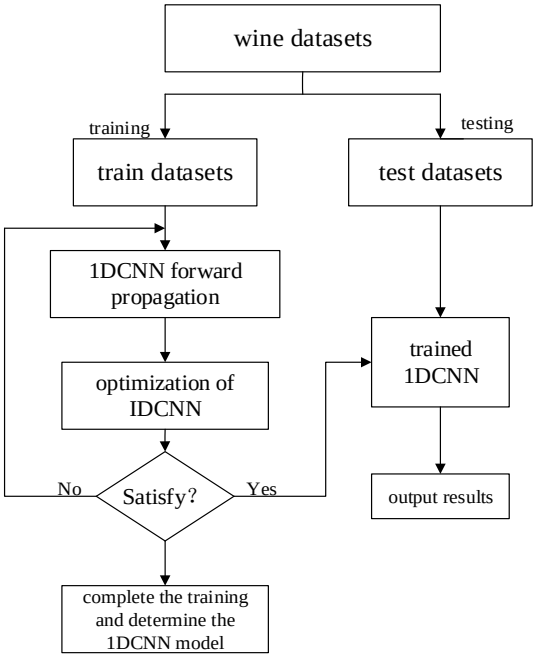

Figure 1. Flowchart of wine classification of 1DCNN

Figure 2 shows the detailed structure of the 1DCNN wine classification model. The physicochemical composition data of the wine is processed as the input matrix X of the 1DCNN model. Convolution layer C contains a convolution kernel with a size of 32×1. Next, apply the ReLU function to get a feature $maps$, Then there is the flatten layer, its role is mainly to "flatten" the obtained feature $maps$, that is, the feature is one-dimensional. In order to prevent over-fitting and improve the generalization ability of the model, the dropout technique is adopted after the flatten layer. The Dropout technique is used in supervised learning, setting the probability of dropout to p, ie each neuron is discarded with probability p, with probability $1-p$ being retained. Finally, after the two-layer fully connected layer, the output layer uses the softmax classifier, assuming that the output after the operation of the fully connected layer is $y_1, y_2, \ldots\ldots, y_n$, and the output of the softmax layer is as in formula (3):

ISSN: 2472-3703 DOI: 10.6911/WSRJ.201911_5(11).0021

$$soft\max(y)_i = \frac{e^{y_i}}{\sum_{j=1}^{n} e^{y_j}} \tag{3}$$

The new output can be understood as the probability that a sample is a different class, and the distance between the predicted probability distribution and the probability distribution of the real sample is calculated by the cross entropy function. The expression of Cross Entropy Error Function (C.E.) is given by equation (4):

$$C.E. = -\frac{1}{N}\sum_{i=1}^{N}(t_i \log(y_i) + (1-t_i)\log(1-y_i)) \tag{4}$$

Where t_i represents the i-th element of the classification target vector, y_i is the ith element of the network output vector, and N is the number of network output elements.In order to minimize the loss function, the learning rate is automatically adjusted using SGD (Stochastic Gradient Descent).

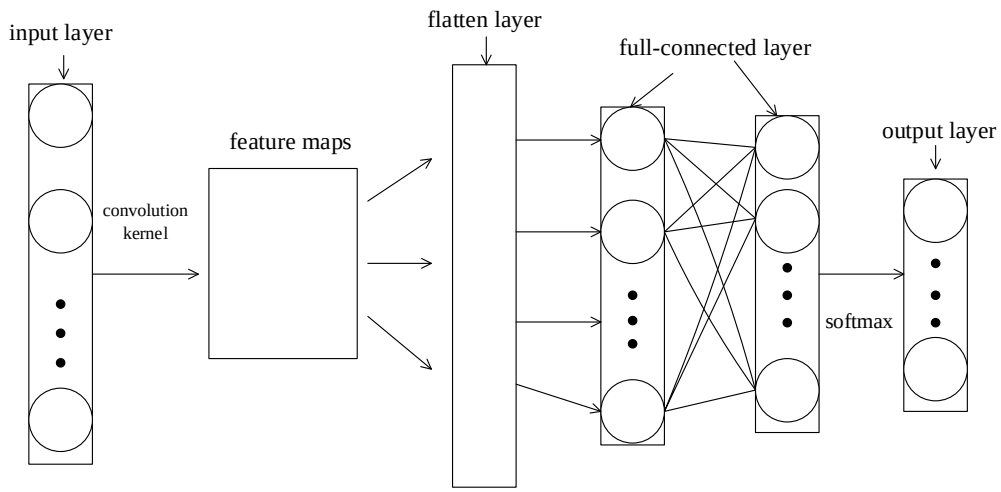

Figure 2. Architecture of 1DCNN-based wine classification model

4. EXPERIMENT

4.1. Dataset Source

The data comes from the wine dataset in the UCI database, which records the chemical composition of three different wines in the same region of Italy. The dataset has a total of 178 sample data, each sample consisting of 13 features, namely alcohol, malic acid, Ash, ash alkalinity, magnesium, phenol content, flavonoids, non-flavonoid phenol, proanthocyanidins, color intensity, color, protein of diluted wine, proline and other ingredients (represented by $X_1, X_2, \ldots\ldots, X_{13}$) Wine species system indicators. 178 samples were divided into three categories according to their quality, 59 in the first category, 71 in the second category, and 48 in the third category. Then 137 samples were selected as the training set, and 41 samples were selected as the test set.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.201911_5(11).0021

4.2. 1DCNN Model Parameter Setting

In this paper, a one-dimensional convolutional neural network model is proposed to extract and classify wine components. Considering the sample size and model calculation amount, the parameter settings for the 1DCNN model are mainly shown in Table 1.The convolution kernel size of 1DCNN is set to 32×1 , the dropout rate after the flatten layer is 0.5, the neurons of the fully connected layer are set to 128 and 64, respectively, and the output of softmax is 3 types. The number of training iterations is 100 and the learning rate is set to 0.01.

Table 1. Characteristics of the 1DCNN used in wine classification

Deep learning toolkit	Keras (TensorFlow backend)
non-linear function for convolution layer and full-connected layers	rectified linear unit (Relu)
non-linear function for output layer	softmax
number of convolution layer	1
number of full-connected layer	2
size of full-connected layer	128 and 64
size of output layer	3
filter size	32
filter length	1
loss function	C.E.
optimiser	SGD
learning rate	0.01
momentum	0.9
epoch	100
dropout rate	0.5

4.3. Experimental Results and Discussion

In order to verify the feasibility of the proposed one-dimensional convolutional neural network model for wine classification, this paper uses the Keras framework, its back-end Tensorflow, deployed in a processor for Intel(R) i5-3470, clocked at 3.2, and memory 16G. On a 64-bit Windows operating system.

At the input level, 137 samples randomly selected were input as training samples. After forward propagation, the model was tuned to update the model to determine the appropriate weight. The other 41 samples were used as test samples to test the wine classification performance. After the model is iterated 100 times, the classification accuracy of the model can be judged by observing the Loss and Accuracy curves of the training process and the test process as shown in Figure 3 and Figure 4.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.201911_5(11).0021

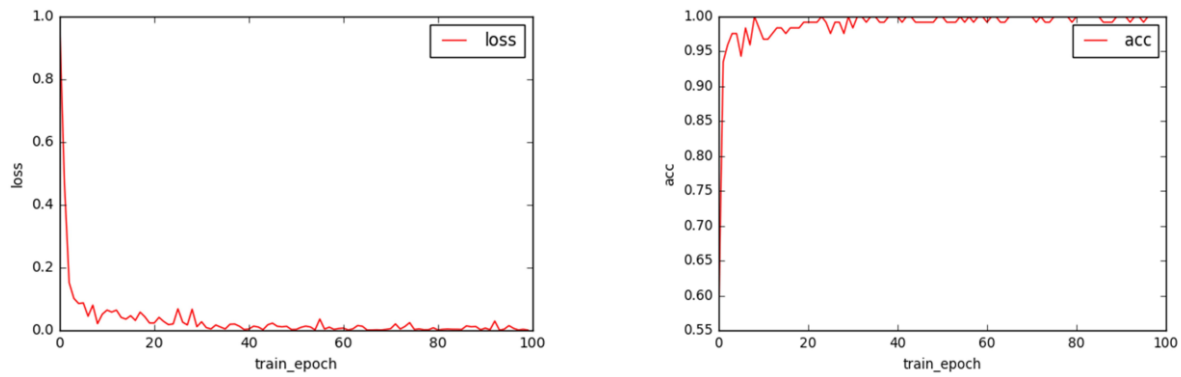

Figure 1. Loss and accuracy curves of the training process

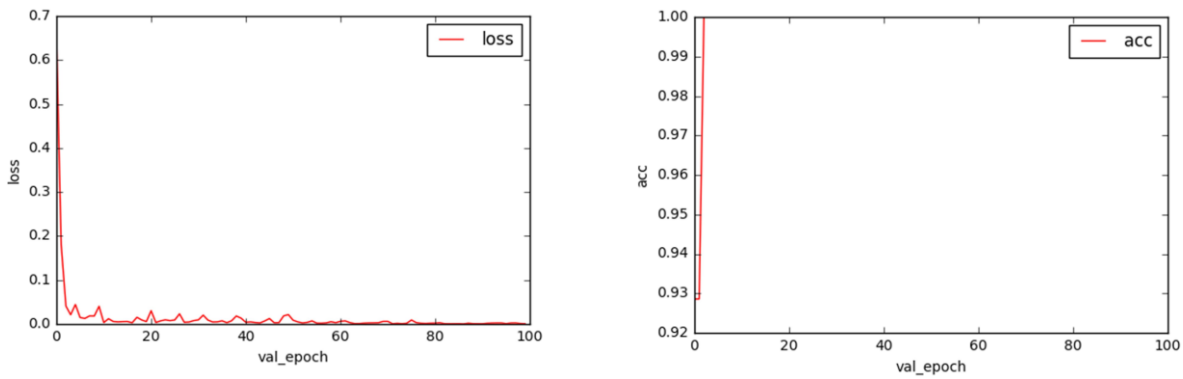

Figure 2. Loss and accuracy curves of the training proce

It can be seen from Fig. 3 that during the training process, the Loss curve converges rapidly and the Accuracy curve converges to 100%. It can be seen from Fig. 4 that during the test set verification process, the Loss curve converges rapidly, and the Accuracy curve tends to be stable after converging to 100%.

To explore the efficiency of the 1DCNN wine classification model, the contrast method used the BP (back propagation) neural network model proposed by Rumelhart and McClelland scientists and the ELM (Extreme Learning Machines) model proposed and summarized by Professor Huang Guangbin. Using the same training set and test set, the results obtained by comparison are shown in Table 2. As can be seen from Table 2, the training accuracy and verification accuracy of the BP and ELM models are smaller than those of 1DCNN.

Table 2. Comparison diagram of different network models

network model	training accuracy (%)	validation accuracy (%)
BP	83.16±0.018	84.12±0.04
ELM	70.073±0.05	75.601±0.015
1DCNN	99.8±0.02	100

5. SUMMARY

In this paper, based on the classification problem of wine, a one-dimensional convolutional neural network is used to analyze the nonlinear relationship between physicochemical components in wine to realize the classification of wine. The experimental data is based on the wine dataset in the UCI machine learning library. The classification accuracy of the model is

ISSN: 2472-3703 DOI: 10.6911/WSRJ.201911_5(11).0021

100% through the established one-dimensional convolutional neural network wine classification model. The classification results obtained by inputting the same dataset into BP neural network and extreme learning machine classification model are worse than the classification effect of wine classification model based on one-dimensional convolutional neural network. The above experiments show that the one-dimensional convolutional neural network wine classification model proposed in this paper has better feature learning ability, and its classification accuracy is also higher.

ACKNOWLEDGEMENTS

Sichuan University of Science and Engineering Talent Introduction Project (No.2017RCL11). Sichuan University of Science and Engineering Postgraduate Innovation Fund Project (y2019011).

REFERENCES

[1] R.Boyer, J. Moore. A fast string searching algorithm[J]. Communications of the ACM,1977, 20(10): 762-772.

[2] R. N. Horspool. Practical fast searching in strings[J].Software Practice &Experience, 1980, 10(6): 501-506.

[3] SP Yuan, JH Wan. Simple Implementation of The Wine Classification with Fuzzy Clustering [J]. Technological horizon,2012(19):96-98+256.

[4] XH Xu, XS Quan, ZF Zhang, XF Hu. A study on the application of PCA and PSO-SVM to the classification of grape wines[J]. Journal of Yunnan University of Nationalities (Natural Sciences Edition),2014,23(06):456-459.

[5] SG Huo, HP Hu, YP Bai, JZ Wang. Wine classification based on PCA-K-means and PCA-SOM neural network [J]. Practice and understanding of mathematics,2016,46(17):168-173.

[6] FY Zhou, LP Jin, J Dong. A review of convolutional neural networks [J]. Journal of Computer, 2017(6).

[7] Sainath T N, Mohamed A R, Kingsbury B, et al. Deep convolutional neural networks for LVCSR[C]// IEEE International Conference on Acoustics. 2013.

[8] Gu J , Wang Z , Kuen J , et al. Recent Advances in Convolutional Neural Networks[J]. Computer Science, 2015.