

Application of Neural Network in Wine Grape Quality Evaluation

Wei Fan¹, Zhi Pan²

¹Department of Electrical and Electronic Engineering, North China Electric Power University, Beijing, 102206, China

²Department of Mathematics and Physics North China Electric Power University, Beijing, 102206, China

fanwei1218@126.com, panzhi@126.com

Abstract - The aim of this paper is to help select suitable grapes for making red wine by studying the correlation between ingredients of the grapes and the resulting wine. Two progressive models are carried out in determining the relation between physical and chemical indexes of the grapes and the resulting wine. Multiple linear regression model is used to get the function relationship between them and the related weight. According to the weight, the less relevant indexes of grape are eliminated. Single layer forward neural network is established between the rest physical and chemical indexes of the grapes and wine to make a more accurate solution of the weight. The correlated factors of each index in grape wine are obtained by analyzing the related weight. The anthocyanin, DPPH radical and soluble solid in grapes has a considerable effect on the quality of red wine.

Index Terms - Single layer forward neural network, Multiple linear regression model, Related weight, Grape wine, Ingredients.

I. Introduction

The quality of grapes would directly affect the quality of grape wine. The combination of physical and chemical indexes of the grape plays an important role in determining the quality of the grapes and ultimately the quality and the flavor of the resulting wine. There is a high correlation between physical and chemical indexes of the grapes and the resulting wine [1]. Currently, the quality of grape wine mainly depends on the evaluation of experienced wine tasters [2]. Therefore, it is urgent to solve the problem of how to effectively predict the physical and chemical indexes of grape wine on the basis of that of grape, and reduce as far as possible the grape's negative impact to grape wine.

Grape wine is made from grape after complex physical, chemical and biological reactions [3]. So there is no definite functional relation between the dependent variable and independent variables. That means we can only use data analysis to establish the connection between them [4]. Artificial neural network, which simulates human neuronal function, can deal with highly non-linear complex relationships. In this essay, we will: Firstly use multiple linear regression model to filter out high related indexes. Then, on the basis of the first step establish relation model between the physical and chemical indexes of the grape and grape wine through single layer forward neural network. We hope that this model can provide some guidance and reference for grape wine industry.

II. Data Source

The data is from a test data of National Higher Education

Cup modeling contest, in which grape wine has 27 samples, the grapes have 32 physicochemical indexes, and the resulting grape wine has 9 physicochemical indexes.

III. Multiple Linear Regression Model

The physicochemical index of wine comes after the standardization as the dependent variable Y, the physicochemical index of grape after standardization as the independent variable X. The regression model is a method of empirical model established by statistical analysis. In the essay, we get rough correlation coefficient between the independent variable and dependent variable through multiple linear regression analysis, after which the low correlation variables are eliminated.

We set $x = (x_1, \dots, x_{32})$, the dependent variable is a wine index y. Multiple linear regression model of y and x is established as follows:

We can get 27 groups of independently observed grape and wine samples $(y_i, x_{i1}, \dots, x_{i,32})$ ($i=1, \dots, 27$). The model can be expressed as:

$$\begin{cases} y_i = \beta_0 + \beta_1 x_1 + \dots + \beta_{32} x_{i32} + \varepsilon_i, \\ \varepsilon_i \sim N(0, \sigma^2), i = 1, \dots, n. \end{cases} \quad (1)$$

Set

$$X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1,32} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{27,1} & \dots & x_{27,32} \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ \vdots \\ y_{27} \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{27} \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_{32} \end{bmatrix}.$$

By the least squares estimation model we get β , which is weight matrix.

According to the correlation coefficient of weight matrix, we eliminate 15 indexes of grape whose correlation coefficient with the dependent variable tend to be zero, and choose 18 indexes as independent variables in the analysis of the neural network model including the protein, amino acid, anthocyanin, malic acid, etc.

IV. Single-layer Forward Neural Network

A. Establishment of Neural Network

In a single-layer forward neural network, there are only the input layer and the output layer. This network studies the input data of input layer and compares with the ideal output results, and then we can determine the weights. Finally we apply this neural network with determined weights to practical problems [5].

In this study, we use multiple linear regression to select 18 standardized physicochemical indexes and set them as input variables (that is to say the number of input-layer neuron is 18). We select 9 wine indexes as output variables including anthocyanins, tannins, total phenols and so on (that is to say the number of output-layer neuron is 9). Therefore, a 18-9 single layer forward neural network is established. The structure diagram is shown in Fig. 1.

This paper adopts linear network function as the activation function:

$$Y = WX + \theta \tag{2}$$

In order to make the network output and the desired output close, this paper adopts the least square criterion asⁱ the learning criterion.

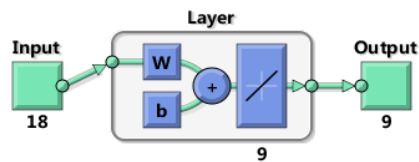


Fig. 1 Single layer forward neural network structure

TABLE I Relative Weight between Grape and Wine Physicochemical Index

Relative weight	Anthocyanin	Tannin	Total phenols	Total flavones	Resveratrol	DPPH half inhibitory volume	Color and luster(L*)	Color and luster(a*)	Color and luster(b*)
Amino acid	0.17	0.36	0.35	0.17	0.34	0.19	0.01	-0.32	-0.27
Protein	-0.24	0.06	-0.27	-0.24	-0.99	-0.27	0.09	0.28	0.01
Anthocyanin	0.55	0.03	0.47	0.55	0.95	0.58	-0.64	-0.34	0.241
Malic acid	0.14	0.10	-0.39	0.14	-1.04	0.09	0.37	0.09	-1.25
Citric acid	0.06	0.02	-0.02	0.06	0.19	0.06	-0.07	0.11	0.01
Tannin	0.21	0.35	0.29	0.22	0.01	0.22	0.01	-0.56	-0.15
DPPH radical	0.61	0.58	0.64	0.61	0.45	0.63	-0.24	-0.33	-0.57
Total phenols	-0.58	-0.37	-0.14	-0.58	0.01	-0.57	-0.37	0.43	2.07
Browning degree	0.09	-0.07	-0.10	0.09	-0.31	0.07	0.18	-0.01	-0.65
Total flavones	0.09	0.38	0.37	0.10	0.46	0.13	0.27	-0.27	-0.99
Resveratrol	0.02	0.10	0.13	0.02	0.34	0.04	-0.32	0.44	0.33
Flavonol	-0.07	-0.03	-0.24	-0.08	-0.30	-0.09	-0.12	0.22	0.34
Total sugar	-0.36	-0.34	-0.29	-0.36	0.27	-0.35	0.12	0.34	0.63
Soluble solid	0.53	0.60	0.46	0.53	-0.30	0.52	-0.21	-0.42	-0.37
PH value	-0.05	0.12	-0.28	-0.06	-0.24	-0.07	0.27	-0.17	-0.79
Titratable acid	-0.22	0.14	-0.16	-0.22	-0.01	-0.22	0.10	0.37	-0.18
Color of fruit peel a*	-0.13	-0.05	-0.20	-0.14	-0.14	-0.14	0.91	-0.89	-1.16
Color of fruit peel b*	-0.04	-0.05	0.27	-0.04	-0.07	-0.03	-0.45	-0.27	1.20

In order to improve accuracy of the weight, this paper adopts the self-adaptive linear network. In the training process, to imitate the human learning process, the self-adaptive linear network improve itself through the training of learning samples one by one, which means the system will update weight when arriving in each sample.

B. Analysis of the Result

As shown in Figure 2, in the 843rd iteration the training achieves the minimum mean square error, MSE=0.861, and from then on the mean square error remains unchanged until the learning rounds reach the limit of 3000 times and the training finishes.

It is very alike between analog output and desired output and the result is satisfactory.

We can get the following table after the output of weight:

From Table 1 we can get the connection between grape and wine physicochemical index. Using grape wine physicochemical index as dependent variables, we get the following conclusion:

TABLE II Connection between Grape and Wine Physicochemical Index

Grape wine physicochemical index	Major related factors in grape
Anthocyanin	Anthocyanin, DPPH radical, Soluble solid, Total phenols(-)
Tannin	DPPH radical, soluble solid
Total phenols	Anthocyanin, DPPH radical, Soluble solid
Total flavone of wine	Anthocyanin, DPPH radical, Soluble solid, Total phenols(-)
Resveratrol	Anthocyanin, DPPH radical(-), Total flavones of grape(-)
DPPH half inhibitory volume	Anthocyanin, DPPH radical, Soluble solid, Total phenols(-)
Color and lustre (L*)	Color of fruit peel a*(++), Anthocyanin(-)
Color and lustre (a*)	Color of fruit peel a*(--)
Color and lustre (b*)	Total phenols(++), Color of fruit peel b*(++), Malic acid(--), Color of fruit peel a*(-)

The table above roughly describes that the anthocyanin, DPPH radical and soluble solid in grape promotes generating anthocyanin in wine in the process of brewing, while total phenols in grape has a role in decomposition or destruction of generating conditions to anthocyanin in wine. Resveratrol has serious positive correlation with anthocyanin in grape, while it has serious negative correlation with protein and malic acid in grape.

Looking at the main related factors in other indicators, we can get further conclusions:

1) The anthocyanin, DPPH radical and soluble solid has close correlation with the quality of grape wine [6].

2) The indicator of color and luster of red wine primarily has correlation with the content of anthocyanin in grape and the color of grape peel.

V. Conclusion

This paper first analyses the rough relation between the grape and grape wine physicochemical index through multiple linear regression model, based on which physicochemical index whose correlation is relatively low are eliminated. Then through the single-layer forward neural network a mathematical model is established about the relation between

the physicochemical index of grape and grape wine. Combining

multiple linear regression model with single-layer forward neural network helps not only to reduce the impact of too many physical and chemical indicators on the accuracy of neural network's training, but also to simplify the structure of neural network and greatly improve the speed of network training. Finally we get the physicochemical index of grape whose correlation is relatively high, which may provide some guidance and reference for wine industry on how to select grape according to the physicochemical index of grape.

References

- [1] Wang Jinjia, Yin Tao, Li Jing, Hong Wenxue, Ma Chongxiao, Visual evaluation of wine quality from physicochemical properties, Journal of Yanshan University, 2010, 133-137.
- [2] Wang Wenjing. Development of Study on Sensory Evaluation in Wine . Liquor Making, 2007, 57-59.
- [3] Gary J. Pickering et al. Journal wine Research, 2000, 129-144.
- [4] Cortez P, Cerdeira A, Almeida F, et al. Modeling wine preferences by data mining from physicochemical properties. Decision Support Systems, 2009, 547-553.
- [5] Jang Q X, Xing W X, Xie J X, Yang D H, College Mathematics Experiment , Tsinghua University Press, 2005-2.
- [6] Boulton R. The copigmentation of anthocyanins and its role in the color of red wine: A critical review. American Journal of Enology and Viticulture, 2001, 67-84.