

Application of Potential Type Electronic Tongue on Milk Discrimination

Jinguo Zhang^{*1}, Chun Pan², Honghui Gao², hong Men², Yitong Li³

¹Power Distribution and Transformation Technical College, Northeast Dianli University Jilin 132012, Jilin, China

²School of Automation Engineering, Northeast Dianli University, Jilin 132012, Jinlin, China

³Beijing University of Posts and Telecommunications, School of Computer Science, Beijing 100876, China

*Corresponding author, e-mail: 505227895@qq.com

Abstract

Four brands of milk based on the potential type electronic tongue are tested by using the principal component analysis (PCA), fuzzy c-means clustering (FCM) algorithm for clustering analysis. Support vector machine (SVM) algorithm is used to forecast category for any brand of milk data, which are extracted from all the data randomly. The results show that potential type electronic tongue can distinguish four brands of milk perfectly, and the forecasting accuracy rate can reach to 100%. Potential type electronic tongue has potential application value in the identification of milk.

Keywords: milk, potential type electronic tongue, principal component analysis, fuzzy C-means clustering, support vector machine algorithm

Copyright © 2013 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

Milk is a common drink in daily life. Because milk has a high nutritional value and is easy to digest and absorb, it's becoming more and more popular. There are lots of brands and diverse forms of milk in the market, so it's very significant job to do discrimination, identification, quality identification and evaluation on milk [1].

The traditional discrimination methods mainly include physical & chemical analysis method and artificial sensory evaluation method. But these methods need tedious operation, long detection time, and have fatigue period [2]. Electronic tongue can imitate human tongue to do qualitative and quantitative analysis on liquid, and can give a rapid evaluation on the overall information of samples [3, 4]. Because of its short detection time, convenient operation, needing no sample pretreatment, lots of research has been done now on it in liquor, beverage, fermentation etc. [5-7], and certain progress has been made. The electronic tongue mainly includes volt-ampere type, potential type, surface acoustic wave type, light-addressable type and impedance spectrum type, etc. [8, 9].

The article has used potential-type electronic tongue in doing discrimination on four brands of milk—namely, Bright pure milk, Mengniu pure milk, Wandashan mellow milk and Yili High-calcium low-fat milk, and used principal component analysis (PCA), fuzzy c-means clustering (FCM) algorithm and support vector machine algorithm on experimental data for doing clustering and prediction treatment.

2. Experiment

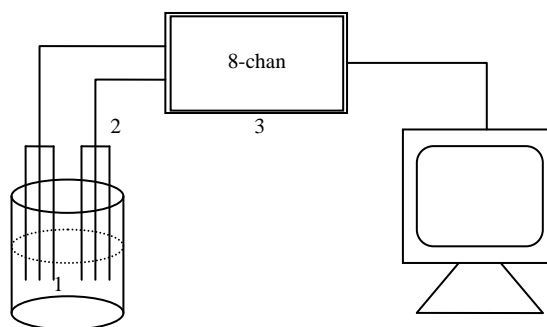
2.1. Materials and Apparatus

The four kinds of milks are all packed by carton and bought from RT Supermarket in Jilin City, each having a volume of 1L. The sample information is shown in Table 1.

The potential-type electronic tongue system (Figure 1) consists of three modules—namely sensor array, data acquisition module (8chan) and pattern recognition. The sensor array is composed of six ion-selective electrodes, among which the five ones are working electrodes K^+ , Na^+ , Ca^{2+} , H^+ , Cl^- , and the other one is LiAc reference electrode.

Table 1. Four Kinds of Milk Sample information

Name	Mengniu pure milk (MC)	Bright pure milk (GC)	Wandashan mellow milk (WZ)	Yili High-calcium low-fat milk (YG)
Date of manufacture	2012.07.20	2012.06.28	2012.08.13	2012.08.13
Guarantee period	6 months if kept tightly closed and at normal temperature	6 months	months if kept at common temperature	6 months if kept tightly closed and at normal temperature
Date of experiment	2012.09.27	2012.09.27	2012.09.27	2012.09.27



1.solution to be tested ; 2. Sensor ; 3. 8chan ion analyzer ; 4. Computer

Figure 1. Structure Diagram of Potential Type Electronic Tongue System

2.2. Experiment Process

Before experiment, beaker, alcohol, desalted water, and cotton swab are prepared. Experiment instrument adopts 8-chan ion analyzer, which is connected to computer through data serial port. The beaker is used for holding experimental samples, deionized water and alcohol are used for cleaning the beaker and ion electrode, and the electrode should be wiped dry by cotton swab after cleaning. Before experiment, the electrodes need to be activated and corrected to ensure the reliability and stability of data collecting.

Take out 300ml milk and pour into the beaker and preheat it in a water bath kettle, then connect the experimental apparatus and clean the electrodes with alcohol and deionized water, then install the electrodes adequately according to the sequence at time of calibration, and then dig the electrodes into the sample for preparation of the experiment. The experiment should be carried out in a 25 °C constant-temperature water bath kettle. During the experiment, the front end of the electrode should be at middle part of the solution, and only when the electrode enters into the solution for 5 min and the data tends to be stable, can the readings be recorded. Measure the four kinds of milk successively, and you can obtain one data every 5s, and the data will be recorded automatically. Repeat the experiment twice for every milk sample, and the average value should be taken as the original data. The four kinds of milk will receive 120 data in total for algorithm treatment.

3. Pattern Recognition Methods

The thesis mainly adopts three methods—namely, principal component analysis (PCA), K-means clustering algorithm (K-means), and support vector machine.

3.1 Principal Component Analysis Method

Principal component analysis (PCA) is a kind of statistical analysis method which reduces multiple variables into a few unrelated synthetic variables by doing dimensionality reduction treatment on data.

The algorithm process is as follows:

- (1) Standardize the data matrix;

$$\bar{x}_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (1)$$

(2) Calculate correlation coefficient matrix;

$$r_{ij} = \frac{s_{ij}}{\sqrt{s_{ii}} \times \sqrt{s_{jj}}} \quad (2)$$

(3) Compute eigenvalue;

Use Jacobi formula to calculate characteristic root $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ and corresponding eigenvector.

(4) Determine the principal component according to the cumulative contribution rate

The contribution rate of principal component refers to the percentage of No. i principal component variance in all the variances, and the value determines whether the ability of No. i principal component in reflecting the original data is strong or weak. Generally speaking, if cumulative contribution rate is between 75%-90%, it can be considered that the first i eigenvalues correspond to No.1, No.2, ..., No.m principal components. Choose first i principal components to replace the original data in analysis, so that the purpose of reducing original data dimension can be achieved.

3.2. Fuzzy C-means Method

Through optimizing objective function, Fuzzy c-means algorithm (FCM) can obtain membership grade of every sample point to all clustering centers, and can thereby determine the generic type of sample point so as to achieve automatic clustering of data samples.

The algorithm process is as follows:

- (1) Set cluster number c and parameter m;
- (2) Give initial membership grade matrix $U^{(0)}$;
- (3) Compute new clustering center V_j ;
- (4) Compute new membership grade matrix;
- (5) Use a matrix norm to compare membership grade matrix between two iterations. If

$\|U(K+1) - U(K)\| \leq e$ then iteration should be stopped.

3.3. Support Vector Machine

Support vector machine (SVM) is based on VC dimension theory and structural risk minimization theory of statistical learning theory. According to the limited sample information, best compromise should be found out between the model complexity (namely, the learning accuracy for specific training samples) and learning ability (namely, the ability of identifying any samples without errors), so as to obtain the best popularization ability.

The algorithm process is as follows:

- (1) Give training sample

$$T = \{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\}, \text{ among which } x_i \in R^n, y_i \in \{-1, +1\}, i = 1, 2, \dots, l, n$$

is dimension of sample space;

- (2) Select proper kernel function and initialize the penalty parameter $C > 0$;
- (3) Construct the programming problems corresponding to the optimum hyper plane

$$\begin{cases} \min W(\alpha) = \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) - \sum_{i=1}^l \alpha_i \\ s.t. \sum_{i=1}^l \alpha_i y_i = 0, 0 \leq \alpha_i \leq C \quad i=1, 2, \dots, l \end{cases} \quad (3)$$

Solve the above equation, the solution is: $\alpha^* = (\alpha_1^*, \alpha_2^* \dots \alpha_l^*)^T$, among which α_i^* is the support vectors of the corresponding training sample;

(4) Compute classification thresholds b^* , select a component α_i^* of α^* in open interval

$$b^* = y_j^* - \sum_{i=1}^l \alpha_i^* y_i K(x_i \cdot x_j) \quad w^* = \sum_{i=1}^l \alpha_i^* y_i x_i$$

(0,C), and compute according to

(5) Construct optimum hyperplane with w^* and b^* $(w^* \cdot \phi(x)) + b^* = 0$;

In this way, the decision function of its corresponding optimal classification face is obtained as follows:

$$f(x) = \text{sgn}[\sum \alpha_i^* y_i K(x_i \cdot x_j) + b^*] \tag{4}$$

Among which $K(x_i \cdot x_j) = \phi(x_i) \cdot \phi(x_j)$, is the kernel function satisfying Mercer Kernel theorem.

4. Results and Discussion

4.1. Principal Components Analysis (PCA)

Draw 30 data from each type of milk sample, so altogether 120 data will be used for PCA algorithm treatment.

From Table 2, it can be seen that the cumulative variance contribution of the first two principal components has reached more than 90%, which represents most information of the whole. From PCA treatment result in Figure 2, it can be seen that four kinds of milk have been well differentiated, distance between classes is far. Except Yili milk, other several kinds of milk samples are near to the class center, and the distinguishing effect is ideal.

Table 2. Contribution Ratio of First Two Principle Components of PCA
Feature analysis of correlation matrix

Eigenvalue	3.7652	0.8961	0.3030	0.0355	0.0002
Ratio	0.753	0.179	0.061	0.007	0.000
Accumulation	0.753	0.932	0.993	1.000	1.000

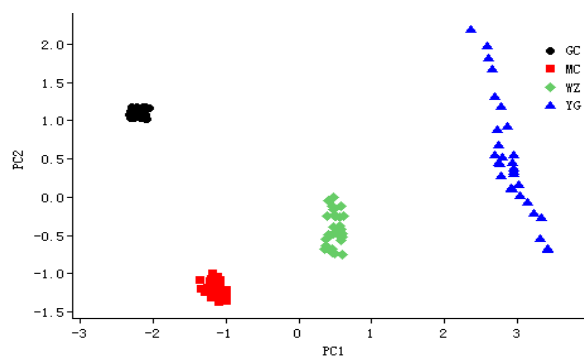


Figure 2. Treatment Results of PCA Algorithm

4.2. Fuzzy C-means Algorithm

Draw 10 data from every milk sample and altogether 40 data are used for fuzzy C-means algorithm treatment.

See Table 3 for FCM treatment results. The overall result of FCM shows that, the data of four kinds of milk samples have been successfully clustered into four groups and the membership grade are all above 90%, and the distinction effect is very obvious.

Table 3. FCM Result in Milk Brands Discrimination

FCM Classification Matrix							
Columns 1 through 8							
0.0001	0.0001	0.0001	0.0001	0.0000	0.0000	0.0000	0.0001
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0002	0.0002	0.0001	0.0001	0.0000	0.0000	0.0000	0.0001
0.9997	0.9997	0.9998	0.9998	1.0000	1.0000	1.0000	0.9998
Columns 9 through 16							
0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0001	0.0002	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.9998	0.9996	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Columns 17 through 24							
0.0000	0.0000	0.0000	0.0000	0.9967	0.9982	0.9991	0.9997
0.0000	0.0000	0.0000	0.0000	0.0016	0.0009	0.0004	0.0001
1.0000	1.0000	1.0000	1.0000	0.0007	0.0004	0.0002	0.0001
0.0000	0.0000	0.0000	0.0000	0.0010	0.0005	0.0003	0.0001
Columns 25 through 32							
0.9999	0.9999	0.9995	0.9995	0.9978	0.9963	0.0615	0.0390
0.0000	0.0000	0.0003	0.0003	0.0010	0.0017	0.9007	0.9377
0.0000	0.0000	0.0001	0.0001	0.0005	0.0008	0.0166	0.0101
0.0000	0.0000	0.0002	0.0002	0.0007	0.0012	0.0213	0.0132
Columns 33 through 40							
0.0231	0.0116	0.0003	0.0007	0.0138	0.0272	0.0451	0.0491
0.9634	0.9818	0.9995	0.9989	0.9788	0.9584	0.9309	0.9249
0.0057	0.0028	0.0001	0.0002	0.0030	0.0058	0.0095	0.0103
0.0077	0.0038	0.0001	0.0002	0.0044	0.0087	0.0145	0.0157

4.3. Support Vector Machine Algorithm

The principal component analysis algorithm and fuzzy c-means algorithm are used to cluster and distinguish the four brands of milk, and the discrimination effect is ideal. In order to classify an unknown sample data, support vector machine algorithm was used to predict and analyze a group of random-drawn data. The four kinds of milk were assigned class numbers 1, 2, 3, 4.

20 data sets are respectively drawn from four kinds of milk data for application in training samples, and then 5 samples were drawn from every remaining data to serve as prediction samples, and then use support vector machine algorithm to do prediction and classification. The data of random-drawn predicted samples is shown in Table 4.

Table 4. Data and Actual Classification of Predicted Samples

	Actual class		Prediction data		
1	429	353	150.6	9.7	74.2
4	437.9	363.9	145.7	18.4	89.5
2	421.4	347	144.4	8	74.2
1	429	353.1	150.7	9.9	74.4
4	437.7	363.6	145.5	18.1	89.2
1	429	353.2	150.8	9.9	74.4
2	421.4	347	144.4	8	74.2
3	429.4	354.2	148.6	12.1	87.7
2	421.4	347	144.4	8	74.2
4	437	363	144.6	17.4	88.6
2	421.4	347	144.4	8	74.2
3	429.4	354.1	148.5	12	87.7
2	421.4	347	144.4	8	74.2
3	429.4	354.1	148.5	12	87.7
1	429	353.1	150.7	9.8	74.3
3	429.4	354.2	148.8	12.1	87.7
1	429	353.1	150.7	9.8	74.3
4	436.9	362.8	144.3	17.2	88.4
3	429.4	354.1	148.7	12.1	87.7
4	437.4	363.4	145.2	17.9	89

SVM prediction results of the 20 sets of random-drawn data are shown in Figure 3. Compared with the actual class in Table 4, the data waiting to be predicted were successfully classified correctly, and the success rate reached 100%, which shows that support vector machine algorithm is very suitable for discrimination research of milk brands.

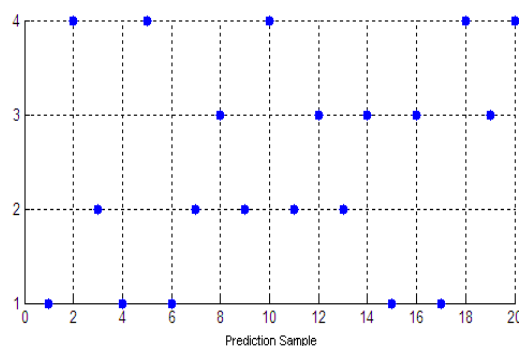


Figure 3. Predicted Result of SVM

5. Conclusion

The potential-type electronic tongue system has been used in discrimination research on four brands of milk in the market, and the two methods--principal component analysis and fuzzy c-means methods have been used in distinguishing the four kinds of milk. We can see from the result that the effect is good. The distance between each class is far, and within-class samples are relatively concentrated. Based on it, support vector machine has been used in training the four milk classes which are already known. In the prediction of unknown randomly-drawn samples, the predicted samples have been successfully classified, and the success rate reaches 100%. It shows that potential-type electronic tongue is quite suitable for the discrimination of milk brands, and has the advantages such as simple operation, rapid reaction, less damage on samples, etc., and therefore has a broad application prospect in food detection field and other fields.

References

- [1] Congyuan Wu, Jun Wang, Sheng Ye. Canonical Correlation Analysis of Relationship between Response Signals of E-tongue and Physicochemical Index of Milk. *Chinese Journal of Sensors and Actuators*. 2010; 23(1): 5-9.
- [2] Shaoping Deng, Shiyi Tian. Progress and its Technique Background of the Electronic Tongue. *Journal of Food Science and Biotechnology*. 2007; 26(4): 110-116.
- [3] Hong Men, Shaofang Zou, Yi Li. A novel electronic tongue combined MLAPS with stripping voltammetry for environmental detection. *Sensors and Actuators*. 2005; 110(2): 350-357.
- [4] Dongmei Chen, Yuan Zhou. *Electronic Tongue and Application of It in Food Industry*. Modern Agricultural Science and Technolog. 2010; (7): 26-31.
- [5] Gallardo J, Alegret S, Valle MD. Application of a potentiometric electronic tongue as a classification tool in food analysis. *Talanta*. 2005; 66(5): 1303-1309.
- [6] Hong Men, Yanping Zhang, Caiwa Zhang. The Development of the Electronic Tongue. *Journal of Northeast Dianli University*. 2012; 32(3): 1-5.
- [7] TRLC Paixão, M Bertotti. Fabrication of disposable voltammetric electronic tongues by using Prussian Blue films electrodeposited onto CD-R gold surfaces and recognition of milk adulteration. *Sensor and Actuators B*. 2009; 137(2): 266-273.
- [8] S Holmin, P Spangeus, F Winquist. Compression electronic tongue data based on voltammetry-a comparative study. *Sensor and Actuators B*. 2011; 176(1): 455-464.
- [9] DB Kantor, G Hitka, A Fekete, C Balla. Electronic tongue for sensing taste changes with apricots during storage. *Sensors and Actuators B*. 2008; 131(1): 43-47.