# STUDY ON DETECTION TECHNOLOGY OF MILK POWDER BASED ON SUPPORT VECTOR MACHINES AND NEAR INFRARED SPECTROSCOPY

Jingzhu Wu[1], Shiping Zhu[2], Yun Xu[3], Yiming Wang[3,*]

[1]、 *College of Information Engineering,Beijing Technology & Business University, Beijing, China, 100037*

[2]、 *College of Engineering and Technology, Southwest University, Chongqing, China, 400715*

[3]、 *College of Information and Electrical Engineering, China Agricultural University, Beijing, China, 100083*

* *Corresponding author, Address: P. O. Box 63, China Agricultural University, 17 Tsinghua East Road, Beijing, 100083, P. R. China, Tel:+86-10-6273759 1, Email: ym_wang@263.net*

Abstract:  This paper presents a novel classifier to identify standard and sub-standard milk powder, which is built by support vector machines (SVM) and near infrared spectroscopy (NIR). The training set is composed of 38 samples and the testing set is composed of 12samples. The correct classification ratio of the training set is up to 100%, while that of the testing set is up to 100%. The result indicates that the combination of SVM and NIR can be used as a fast, convenient, and safe technology to identify standard and sub-standard milk powder.

Key words:  Near Infrared Spectroscopy, Support Vector machines, Milk Powder

## 1.    INTRODUCTION

There appeared many kinds of sub-standard milk powder in recent years in China, which are damaged to people's life. Traditional chemical method to detect milk powder is time-consuming and complex. Near infrared spectrometry（NIR）technology is fast, green and nondestructive. NIR is gradually applied in the field of food and agriculture. Support Vector Machines (SVM) is a new machine-learning algorithm and a new technology

to data classification. This paper presents a novel classifier to identify standard and sub-standard milk powder, which is built by SVM and NIR. The study is supported by the foundation of National Advanced Technology Development Project (863 project, No. 2003AA209012) and National Natural Science Foundation of China (project No. 30671198).

## 2. EXPERIMENT

The sample set is composed of 36unit standard milk powder sample and 14unit substandard milk powder. Near infrared spectrum of sample set were collected by diffuse reflectance from 12500 to 4000 cm-1 in 16 cm-1 on MATRIX-I spectrometer. The software used for this research was implemented in MATLAB v.6.5. The software package about SVM named LIBSVM can be conveniently downloaded from the web, http://www.kernelmachines.

We selected 38 samples, including 30 standard samples and 8 sub-standard samples, as the training set. The testing set is made up of the other 6 standard samples and 6 sub-standard samples. A classifier based on SVM is built to identify the quality of milk powder in this paper. We choose the RBF kernel and the defaulted the penalty parameter C=1.

Table 1: Classifying result by SVM with original input

| $\gamma$ | Support Vectors | Recognition Ratio | Prediction Ratio |
|---|---|---|---|
| 0.5 | 18 | 89.34%(34/38) | 83.33%(10/12) |
| 1.5 | 19 | 94.73% (36/38) | 100%(12/12) |
| 2.5 | 19 | 100%(38/38) | 100%(12/12) |
| 3.5 | 20 | 100% (38/38) | 100% (12/12) |
| 4.5 | 18 | 100% (38/38) | 83.33% (10/12) |
| 5 | 19 | 100% (38/38) | 66.67% (8/12) |

The result indicates that the SVM classification has good performance to identify the stand and sub-stand milk powder. The theorem states that the less support vectors, the more powerful the model is(Zhang Xuegong, 2000). So the best parameter is γ=2.5 and the number of support vector is 19. It is obvious that while γ is increasing, the Recognition Ratio and the Prediction Ratio is increasing along with γ . After the both ratios arrived at 100%, the Prediction Ratio will drop down and Recognition Ratio will maintain 100% if γ keeps on increasing. So the kernel parameter γ influences the SVM model prediction ability.

# REFERENCES

Zhang Xuegong, Introduction to Statistical Learning Theory and Support Vector Machines, *Acta Automation Sinica*26 32-42(2000)