Çanakkale Onsekiz Mart Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 2016:2, 1, 22-35.

Çanakkale Onsekiz Mart University, Journal of Graduate School of Natural and Applied Sciences, 2016:2, 1, 22-35.

E-Nose Identification of Milk Somatic Cell Count

Melis İNALPULAT[1], Ünal KIZIL[2]*, Ertuğrul BİLGÜCÜ[3], Levent GENÇ[1]

[1]Çanakkale Onsekiz Mart University, College of Architecture and Design, Department of Urban and Regional Planning, Rural Planning and Geographic Systems Laboratory (RPGIS-LAB), 17020, Çanakkale, Turkey
[2]Canakkale Onsekiz Mart University, College of Agriculture, Department of Agricultural Structures and Irrigation, Agricultural Sensor and Remote Sensing Laboratory (ASRESEL), 17020, Çanakkale, Turkey
[3]Çanakkale Onsekiz Mart University, Biga Vocational School, Department of Milk and Milk Products, 17200, Biga, Çanakkale, Turkey

Abstract

Mastitis is a common disease among dairy animals which causes serious economic losses. It can be diagnosed via diverse clinical findings, while milk somatic cell count (SCC) is accepted as a key indicator. However, determination of SCC with traditional methods is time consuming and laborious. This paper focuses on the ability of electronic nose (e-nose) system containing 12 different metal oxide sensors (MOS) to discriminate milks with somatic cell counts (SCC) above a threshold value. Milk samples were collected from dairy farms around Biga district of Çanakkale province, Turkey. Forty-six samples were analyzed using standard protocols in laboratory, then exposed to DiagNose II e-nose system. Artificial Neural Networks (ANNs) was used to discriminate between Non-Mastitic (N-M) / Mastitic (M) samples depending on sensor responses. Results showed that 8 of 12 sensors were responded to milk samples. Thus, performances of several ANNs models with different topologies were tested using 8 sensor responses. ANNs was trained using 28 samples, and remaining 18 samples were used in validation step. Among

*Sorumlu Yazar (Corresponding Author): Ünal KIZIL,
(e-posta: unal@comu.edu.tr)

tested models, the results of the lowest overall errors for training and validation steps were found to be 35.71 % and 38.89 % respectively. To improve the performance, Principal Components Analysis (PCA) performed for dimension reduction and three components were selected to be included in ANNs model instead of 8 sensors. Performing of PCA prior to ANNs provided decreased overall errors for training (10.7 %) and validation (0 %). However, the actual performance of the system should be tested using new dataset.

Keywords: Artificial neural network, electronic nose, mastitis, milk somatic cell count, principal component analysis

Süt Somatik Hücre Sayısının E-Burun ile Belirlenmesi

Özet

Mastitis sağmal hayvanlar arasında yaygın bir hastalık olup önemli ekonomik kayıplara sebep olur. Hastalık çeşitli klinik bulgularla teşhis edilebilirken süt somatic hücre sayısı (SHS) kilit göstergelerden biri olarak kabul edilmiştir. Bununla birlikte SHS' nın geleneksel yöntemlerle belirlenmesi yoğun emek gerektirir ve zaman alıcıdır. Bu çalışma farklı metal oksit sensörler (MOS) içeren elektronik burun (e-burun) sisteminin bir eşik değerin üzerinde SHS içeren sütleri ayırt edebilme yeteneği üzerine odaklanmıştır. Süt örnekleri Çanakkale ili Biga ilçesinde bulunan çiftliklerden toplanmıştır. Kırk altı örnek laboratuarda standart protokoller ile analiz edilmiş ve ardından DiagNose-II elektronik burun sistemi ölçümüne tabi tutulmuştur. Sensör tepkilerine göre Mastitik-Olmayan (M-O) / Mastitik (M) sütlerin ayırt edilmesinde Yapay Sinir Ağları (YSA) kullanılmıştır. Sonuçlar 12 sensör içerisinden 8 sensörün süt örneklerine tepki verdiğini göstermiştir. Bu nedenle farklı topolojilere sahip çeşitli YSA modellerinin performansı 8 sensörün tepkileri kullanılarak test edilmiştir. Tüm YSAları 28 örnek kullanılarak eğitilmiş ve kalan 18 örnek ise geçerlik aşamasında kullanılmıştır. Test edilen modeller arasından eğitim ve geçerlik aşamalarına ilişkin en düşük hata sonuçları sırasıyla % 35.71 ve % 38.89 bulunmuştur. Performansın artırılması amacıyla boyutları azaltmak için Ana

Bileşenler Analizi (ABA) uygulanmış ve 8 sensör yerine 3 bileşen YSA modeline dahil edilmiştir. YSA çalıştırılmadan önce ABA uygulanması eğitim (% 10.71) ve geçerlik (% 0) aşamalarındaki hataların daha düşük olmasını sağlamıştır. Ancak sistemin gerçek performansı yeni veri setleriyle test edilmelidir.

Anahtar Sözcükler: Ana bileşenler analizi, elektronik burun, mastitis, süt somatik hücre sayısı, yapay sinir ağları

1. Introduction

Milk yield and quality is strongly related to udder hygiene. Mastitis is known as one of the most common diseases affecting udder glands of dairy animals due to infections. It causes serious economic losses in Turkey, as well as many other countries. The direct cost to the public sector is estimated to be almost 30 million $ (Tekeli, 2005; Turkyilmaz et al., 2010). Therefore, it is crucial to determine its presence in early stages to take necessary precautions.

Mastitis can be diagnosed via diverse clinical tests (Eriksson et al., 2005). Milk somatic cell count (SCC) is reported to be the most valid indicator (Reneau, 1986). California Mastitis Test is most commonly used method for SCC estimation (Leach et al., 2008). Since the diagnose methods are laborious and time consuming (Reneau, 1986), rapid, reliable and relatively economic techniques are required.
Different volatile organic compounds (VOCs) are generated due to the mastitis presence. Thus, differences in odor measurements are considered as mastitis indicators (Persaud et al., 2002). Volatile headspace gases can be sensed using gas sensors which are employed in electronic nose (e-nose) systems and successfully used for odor measurements in various studies (Casalinuovo et al., 2006; Alam and Saeed, 2013; Kizil et al., 2015). These systems mimic human nose and can sense different odors even at very low concentrations. Signal processing and pattern recognition processes are other important components of an e-nose system. Several multivariate analysis techniques used with e-noses

including linear discriminant analysis (LDA), Fourier transform (FT), principal component analysis (PCA), and artificial neural networks (ANNs) (Zohora et al., 2013).

The objective of this study was to evaluate the performance of an e-nose system for estimation of milk SCC. Although there are literatures reporting e-nose identification of SCC in milk, they employed different data analysis and classification techniques. Therefore, it was aimed to evaluate performance of ANNs as a classification method. The PCA analysis was applied for improving the ANNs model performance.

2. Materials and Methods

2.1. Collection, preparation and lab analysis of milk samples

Total of 46 different milk samples were collected from randomly selected dairy barns located in Biga town of Çanakkale province, Turkey. Two portions from each sample were obtained on 24 April 2014 in 100 mL sterile jars. One portion of all samples transferred to the Raw Milk Analysis Laboratory of Çanakkale Onsekiz Mart University (ÇOMU), Biga Vocational College to be analyzed for determining the SCC. Other portions which were numbered accordingly, transferred to ÇOMU, Agricultural Sensor and Remote Sensing Laboratory (ASRESEL) for e-nose analysis. In order to maintain milk quality parameters, samples were carried in coolers filled with ice cubes. All milk samples were analyzed in the same day of sample collection in both laboratories. Milk samples with SCCs under 400,000 ml^{-1} were considered as "Non-mastitic" (N-M) and the rests are "Mastitic" (M) (Smith, 1996).

2.2. Electronic nose measurements

Milk samples were exposed to Diagnose-II e-nose system (The eNose Company, Zutphen, Netherlands) employing 12 gas sensors. Gas generation patterns of each sensor were recorded to

monitor headspace VOCs for 280 seconds. Schematic representation of e-nose measurements is given in Figure 2.1.

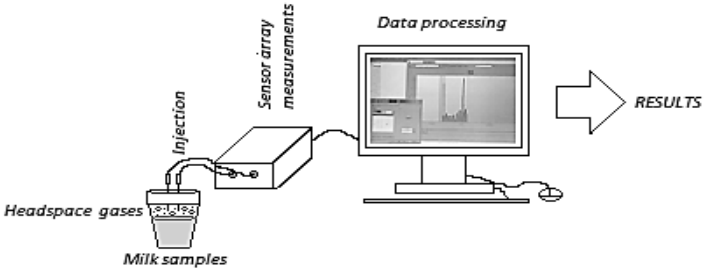

Figure 2.1. E-nose measurements

2.3. Data processing

Data were downloaded and converted to MS Excel file format after recording e-nose readings of all samples. There were 46 readings and each reading also included 8 sensor responses which lead to 368 sensor responses in total. Prior to the statistical analysis, these raw data signals were smoothed and normalized. In the smoothing, moving average algorithm was applied. A detailed data processing procedure has been described in (Kizil et al., 2015). Following the normalization process, graphical sensor responses were converted to numeric database to be used in statistical analysis. For this purpose, the area under each curve was calculated as follows. In this calculation it was assumed that the curve can be divided into rectangles.

$$A = \sum_{k=1}^{t} f(t_k)\Delta t$$

Where; *A* is the total area under curve, *t* is time (sec), and *k* is the designation number for each rectangle.

2.4. Developing ANNs model

ANNs is a non-parametric pattern recognition, and classification technique which is capable of learning. An ANNs consists of input layers, hidden layer(s) and output layers. In this study, eight sensor responses were used as inputs in the model. The outputs were in the form of string constant either "Non-mastitic" (Class I) or "Mastitic" (Class II). ANNs model was developed in two steps. Initially dataset was partitioned into two groups as training set and validation set randomly; 60 % of data was used in training step (28 readings) and remainders used in validation of ANNs model (18 readings). In order to enhance classification rate a back propagation algorithm was employed. Different network sizes were tested by changing the number of nodes, epochs and hidden layers, to identify the best model parameters. Analysis performed using 'XL Miner' add-in of MS Excel (Cytel Software Corporation, USA).

PCA is a common linear transformation technique for dimension decrement process which identifies the most efficient subgroup of original data as an input set to obtain higher classification accuracy (Perera et al., 2002; Noorsal, 2005). It is used to transform data from the many sensor patterns into a less number of principal components (Mohamed et al., 2009). Concurrently use of PCA and ANNs provides improvement in ANNs by reducing the number of inputs and disregarding insignificant sensor responses, and scaling down the computational complexity (Markom et al., 2009). Thus, it was used to enhance ANNs model performance in the study. Figure 2.2 represents the techniques followed in study.

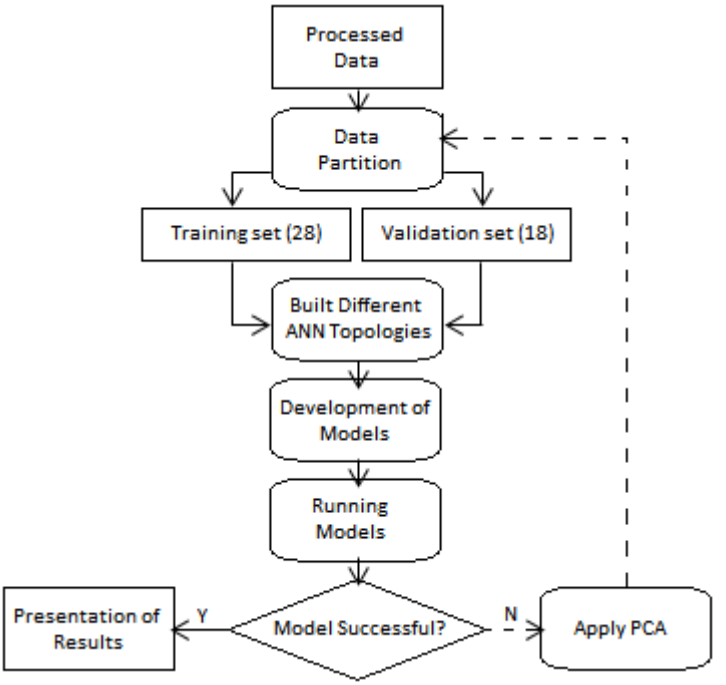

Figure 2.2. Steps of ANNs model parameters selection and model performance enhancement

3. Results and Discussion

3.1. SCC contents of milk samples

According to lab analysis 39 % of samples (18 samples) were containing somatic cells less than 400,000 ml^{-1} and the remainders (61 %) were mastitic.

3.2. Sensor responses

It was observed that 8 of 12 sensors were responded to the head-space volatiles of samples. Some polar compounds may cause malfunctioning of some sensors (Balasubramanian et al., 2004; Kizil et al., 2015), so they may not be sensitive to exposed compounds. Due to this fact, 8 sensors were considered in study as

mentioned before. Figure 3.1. shows mean responses of each sensor against samples.

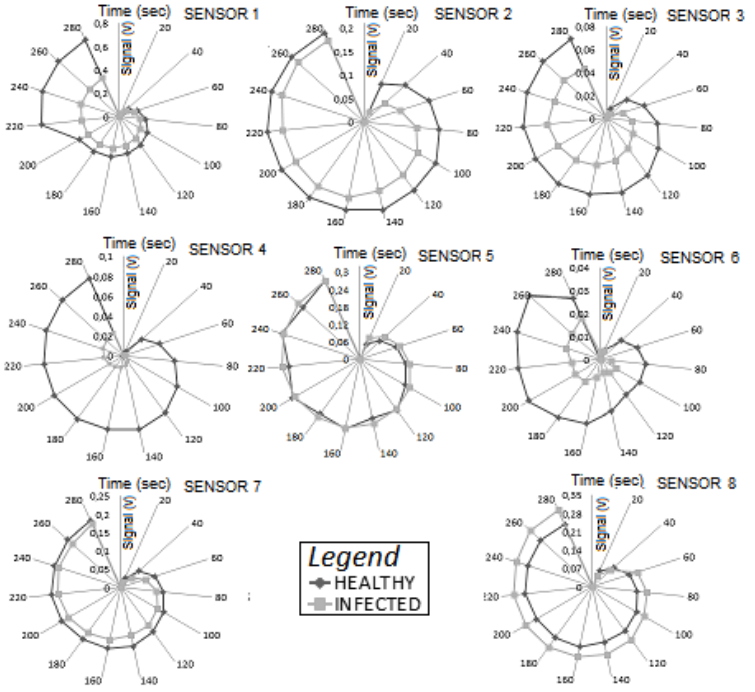

Figure 3.1. Sensor signal level changes in response to healthy and infected milk samples

3.3. Training, validation and performance improvement of ANNs model

In the study various ANNs models were developed using readings of 28 samples for training, and 18 samples of validation steps. Trial and error method was applied to select parameters that yield best classification results. Model parameters and corresponding values for the best classification are summarized below.

The best classification results were obtained using the parameters given in Table 3.1. These results showed that 9 of 13 "Non-mastitic" samples were likely to classify as "Mastitic" in training step while only 1 of "Mastitic" samples was misclassified (Table

3.2). The overall accuracy of training phase was found to be 64.29 %. On the other hand, classification accuracy of validation phase was lower than training (61.11 %), and four out of the 5 "Non-mastitic" samples were misclassified, and 3 of the 13 "Mastitic" samples were misclassified (Table 3.3).

Table 3.1. Selected model parameters that gave best results

Hidden layers	2
Nodes in hidden layers	10
Hidden/output layer sigmoid	Standard/standard
Epochs	1000
Step size gradient	0.1
Momentum	0.6
Error tolerance	0.01

Table 3.2. Classification confusion and error matrices for training phase

Classification Confusion Matrix			Error Report			
Classes	Predicted Class		Classes	Cases	Error	Errors (%)
Actual Class	N-M	M	N-M	13	9	69.23
N-M	4	9	M	15	1	6.66
M	1	14	Overall	28	10	35.71

Table 3.3. Classification confusion and error matrices for validation phase

Classification Confusion Matrix			Error Report			
Classes	Predicted Class		Classes	Cases	Error	Errors (%)
Actual Class	N-M	M	N-M	5	4	80.00
N-M	1	4	M	13	3	23.08
M	3	10	Overall	18	7	38.89

As mentioned above, an overall error of 35.71 % was observed in training, and 38.89 % in validation steps for the best predictions. Since the results were not found to be satisfactory, PCA was applied to responses of an array of eight sensors to test whether ANNs model performance can be improved for prediction of Non-mastitic/Mastitic status of milk samples. The results of PCA can be seen on Table 3.4. Three principle components (PC1, PC2

and PC3) were kept which express 98.68 % of cumulative variance (CV). Individually, PC 1, PC 2 and PC 3 are accounted for 75.16 %, 17.31 %, and 6.20 % of variance respectively. Therefore, these components were used in ANNs as inputs.

Table 3.4. Variances of principal components

PCs	1	2	3	4	5	6	7	8
V	6.013	1.384	0.496	0.074	0.030	0.001	~0	~0
V(%)	75.164	17.311	6.204	0.920	0.383	0.016	~0	~0
CV	75.164	92.475	98.680	99.600	99.983	99.99	99.99	100

It was observed that satisfactory classifications was achieved after applying PCA, using given parameters in Table 3.5 even with lower numbers of hidden layers, nodes and epochs. Laboratory analysis results showed that of those 28 samples 25 were correctly classified by the ANNs model. The classification confusion matrix and error report summarizes the performance of model in training step (Table 3.6). Overall classification error is 10.71 %. The ANNs model was able to classify all milk samples correctly with a classification performance of 100 % in validation step (Table 3.7).

Table 3.5. ANNs model parameters

Hidden layers	1
Nodes in hidden layers	4
Hidden/output layer sigmoid	Standard/standard
Epochs	800
Stepsize gradient	0.1
Momentum	0.6
Error tolerance	0.01

Table 3.6. Classification confusion and error matrices for training phase

Classification Confusion Matrix			Error Report			
Classes	Predicted Class		Classes	Cases	Error	Errors (%)
Actual Class	N-M	M	N-M	13	9	69.23
N-M	5	2	M	21	1	4.76
M	1	20	Overall	28	3	10.71

Table 3.7. Classification confusion and error matrices for validation phase

Classification Confusion Matrix			Error Report			
Classes	Predicted Class		Classes	Cases	Error	Errors (%)
Actual Class	N-M	M	N-M	6	0	0
N-M	6	0	M	12	0	0
M	0	12	Overall	18	0	0

As the model performs with new datasets it updates the weight and bias values to yield better results. Therefore, one reason for the higher classification success may be sourced from the learning capability of the ANNs model.

4. Conclusions

A commercially available e-nose system, DiagNose II, was used to classify milk samples as "Non-mastitic" or "Mastitic" based on the SCC numbers. An ANNs model was used to classified samples. In ANNs model, feed-forward, back-propagation structure with sigmoid function was used. In order to improve the overall classification performance of the model, PCA was applied to dataset. A total of 46 milk samples each collected from different dairy barns were used to develop and test ANNs model.

It was observed that DiagNose II e-nose system can be used to classify milk samples based on their headspace volatile measurements. In the model development step ~36 % of classification error was observed while it was found to be ~39 %. However, applying PCA improved the classification results and errors of training phase were ~11 %, and the e-nose system was able to correctly classify all samples in validation. Our current effort focuses on developing a metal-oxide sensor based e-nose system that is capable of acquiring and processing sensor signals via built-in software. Thus, it will be possible to evaluate the quality parameters of biological materials in shorter times. It will also enable us to test this system in other disciplines such as waste management, air quality, and etc.

Kaynakça

Alam H., Saeed H., 2013. Modern Applications of Electronic Nose: A Review. International Journal of Electrical and Computer Engineering 3(1): 52–63.

Balasubramanian S., Panigrahi S., Logue C.M., Marchello M., Doetkott C., Gu H., Sherwood J., Nolan L., 2004. Spoilage Identification of Beef Using An Electronic Nose System. Transections of ASAE 47(5): 1625–1633.

Casalinuovo I.A., Pierro D.D., Coletta M., Francesco, P.D., 2006. Application of Electronic Noses For Disease Diagnosis and Food Spoilage Detection. Sensors 2006 (6): 1428–1439.

Eriksson A, Waller K.P., Svennersten-Sjaunja K., Haugen J-E., Lundby F, Lind O., 2005. Detection of Mastitic Milk Using A Gas-Sensor Array System (Electronic Nose). International Dairy Journal 15: 1193–1201.

Kizil U., Genc L., Genc T.T., Rahman S., Khaitsa M.L., 2015. E-Nose Identification of Salmonella enterica in Poultry Manure. British Poultry Science 56(2): 149–156.

Leach K.A., Green M.J., Breen J.E., Huxley J.N., Macaulay R., Newton H.T., Bradley A.J., 2008. Use of Domestic Detergents in The California Mastitis Test For High

Somatic Cell Counts in Milk. Veterinary Record 163: 566–570.

Markom M.A., Shakaff A.Y.M., Adom A.H., Ahmad M.N., Hidayat W., Abdullah A.H., Fikri, N.A., 2009. Intelligent Electronic Nose System for Basal Stem Rot Disease Detection. Computers and Electronics in Agriculture 66: 140–146.

Mohamed E.I., Maher H.M., Shaban A.M., Hussein, F.M., 2009. Electronic Nose Technology For The Accurate Detection of Estrus By Monitoring Changes in Perineal Odors in Dairy Cows. Journal of Biophysics and Biomedical Sciences 2(1): 128–133.

Noorsal E., 2005. Development of Electronic Nose System Using Quartz Crystal Microbalance Odour Sensor Array. MSc. Thesis. Universiti Sains, Malaysia.

Perera A., Gomez-Baena A., Sundic T., Pardo T., Marco S., 2002. Machine Olfaction: Pattern Recognition For The Identification of Aromas. Proceedings of 16th International Conference on Pattern Recognition 2: 410–413.

Persaud K.C., Lee D.H., Byun H.G., 2002. Objective Odour Measurements. Biometrics, Springer, Netherlands. 251–272.

Reneau J.K., 1986. Effective Use of Dairy Herd Improvement Somatic Cell Counts in Mastitis Control. Journal of Dairy Science 69: 1708–1720.

Smith K.L., 1996. Standards For Somatic Cells in Milk: Physiological and Regulatory. Mastitis Newsletter, Newsletters of the IDF 144: 7–9.

Tekeli T., 2005. Mastitis. AB Sürecinde Kaliteli Süt Üretimi ve Somatik Hücre Sayısı. Güzeliş Matbaası, Konya. 19-35.

Turkyilmaz S., Yildiz O., Oryasin E., Kaynarca S., Bozdogan B., 2010. Molecular Identification of Bacteria Isolated from Dairy Herds with Mastitis. Kafkas Univ.Vet. Fak. Derg. 16(6): 1025–1032.

Zohora S.E., Khan A.M., Srivastava A.K., Hundewale N., 2013. Electronic Noses Application to Food Analysis Using Metal Oxide Sensors: A Review. International Journal of Soft Computing and Engineering 3(5): 199–205.