QCM based Electronic Nose for Black Tea Quality Evaluation by Different Data Analysis Techniques

Moumita Guha, BipanTudu, Pritam Singha Roy

ABSTRACT---The common and effective technique to find the quality of tea is by sensor activities of the human called Tea Tasters. But this technique is most predictable in nature. For this reason very sophisticated sensor network has been designed. Electronic Nose is the most efficiently used to calculate the behavior and quality of black tea. An Electronic Nose sensor has been developed, consisting of an array of five AT-CUT 10 MHz QCM sensors. The array has been exhibits by the aroma characteristics of different types of black tea (CTC, Orthodox) and the response has been monitored online though Data Acquisition System. The data obtained in this manner has been clustered with the help of different algorithm(PCA,LDA). A BPMLP network has also been incorporated for the purpose of classification. The paper is concluded with comparison of cluster validity indices of original data with those of clusters obtained with different algorithms and the performance of the system which is very effective, is justified.

Keywords—Black Tea; Electronic Nose; QCM; Clustering; Classification; PCI; LDA; ICA; BPMLP; Dunn's Index

I. INTRODUCTION

The quality of tea is characterized by its flavour. According to "???" few major elements are effective for the flavour of tea are listed up in Table 1 [1]. Evaluation of tea quality and the differentiation of the particle from the tea is very impossible. Traditionally this evaluation is done by a sensitive sensor made by human sense capability known as "Tea Tasters".But this may be produce error and it is very critical to measure the aroma due to different human reliability.

TABLE 1BIOCHEMICAL COMPONENT IN BLACK TEA AND ITS FLAVOUR

Components	Sense
$C_{10}H_{18}O$	Sugary
Palmarosa Oil, Citronella	Odour perfume
$C_{15}H_{26}O$, C_7H_6O	Fruit flavour
$C_6H_{12}O$, $CH_3(CH_2)OH$, 3 hexen-1-ol, Grassy, C_3H_{20O}	Clean and fine

Decayed its sensitivity due to the reason of human mental activities, brain activities . Moreover, scoring varies from taster to taster even for the same sample.

An Electronic Nose is very likely to be a solution for these problems. It may be used as a gadget which is used for fast, and reliable of the tea aroma analysis. Not only that this is a non invasive and useful for continuous monitoring monitoring of the tea aroma which helps in quality evaluation. A QCM based Electronic Nose has been devised to serve the purpose and it has been experimented with two major variants of Black Tea, viz., CTC ("Cut, Tear, Curl" or "Crush, Tear, Curl") and Orthodox. Very promising results have been obtained and post data acquisition processing has been done with different

Clustering algorithms. In this paper, those results are brought together and the utility of the E-Nose system as followed by data processing units is shown.

II. EXPERIMENTAL

A. The QCM Sensor

A QCM (Quartz Crystal Microbalance) is used to calculate the frequency characteristics when change the mass per unit area in a quartz crystal resonator. For any variation of the mass along to the surface, the resonance frequency may be changed. With some simplification and assumptions, the change in frequency can be quantified and correlated to the change in mass by Sauerbrey's Equation [2]:

$$\Delta f = \frac{-f_0^2}{A\sqrt{\rho_q \mu_q}} \Delta m$$

f_0 – Resonant Frequency, Δf – Change in Frequency, Δm – Change in Mass, A – Piezoelectrically Active Crystal Area, ρ_q – Density of Quartz, μ_q – Shear Modulus of Quartz. It is quite evident that any deposition of mass on the resonator area increased according to the of frequency of oscillation from its initial value. The very principle is quite effectively used in Gas sensing and even in Liquid phase in recent times. This also has led to the idea of devising a QCM based nose system for Tea Quality Evaluation.

B. The Digital Nose system

A setup of Electronic Nose consisting of QCM sensors has been devised for tea quality determination and classification. The schematic of the whole setup is shown in Figure 1.

Most usable adsorbent compounds for 3 significant volatile materials of different aroma flavor of the tea such Linalool, Geraniol etc. are used. Solutions of the required

Revised Manuscript Received on June 10, 2019.

MoumitaGuha,Instrumentation Dept, JadavpurUniversity , Kolkata, India. (mou.guha08@gmail.com)

BipanTudu,Instrumentation Dept, JadavpurUniversity , Kolkata,India. (bipan.tudu@gmail.com)

PritamSingha Roy, Research Scholar, Pacific University, Rajasthan, India. (prittam.pritam@gmail.com)

particles with proper solvents shown in Table 4. The required sample has been made and dropped on the surface of the Quartz crystal blanks. The all detecting transducers are located into the sample sensor chamber. Here a suction pump has been used to suck the sample aroma of the tea and the fresh air (as required by the selector button)) for purging. The most common digital IC 8084 series used to generate the oscillation in crystal material and here the DAQ system used to analysis the different response of black teas.

Experimental setup details

The conditions required for the experiment done are listed below:

The sensor drum vol. – 480 mlFlow rate of Air through the suction pump – 4 L/minuteAmount of dry tea sample used – 15 gramDuration taken for the sample – 33 secondTime taken for the Purging is 4 minute

C. Tea for experiments

Two major types of Black Tea are available in the market, viz., CTC and Orthodox. Two separate types of teas are collected from shop are below:

1) CTC @Rs 400 (CTC 400)
2) CTC @Rs 600 (CTC 600)
3) Orthodox @ Rs 600 (ORTH 600)
4) Orthodox @Rs 1200 (ORTH 1200)

Each type has been subjected to experimentation and 20 measurement data taken under consideration for analysis purpose.

III. DATA ANALYSIS& RESULTS

The data obtained by experimentation with Black Tea samples by the E-Nose setup, was five-dimensional and consisted of

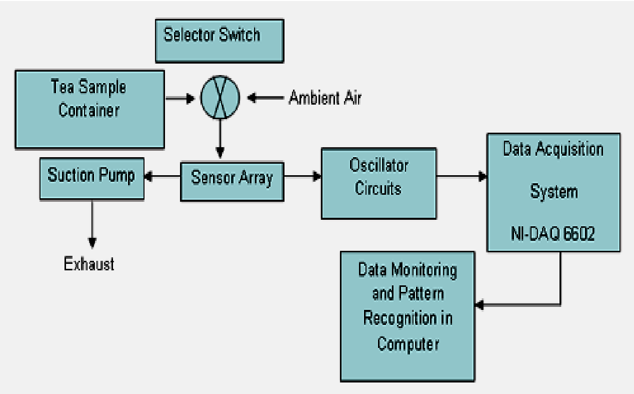

Figure 1.Functional block of QCM based Electronic Nose

TABLE II. LIST OF COATING MATERIALS ANF THE CONCENTRATION

Coating Elements	Solvents	Concentration
Glucose	Water	0.6%w/v
Adenin Phosphate	Sprit	0.65%w/v
PEG	Chloroform	0.18%w/v
$C_9H_{11}NO_2$ (D-Phy)	Hydro ethanol	0.6%w/v
$C_{20}H_{38}O_{11}$ (EC)	$(CH_2)_4O$	1.5%w/v

Over lapping clusters. As the usable presentation is not available for high dimension data, it leads to difficulty in finding patterns if dealing with the raw data. That’s why different algorithms have been used to transform the data not only for better clustering but also for dimension reduction to make the task of analysis easier. As an indication of how good the separability of classes is, a Cluster Validity Index called Dunn’s Index has been computed for each case. A BPMLP network has been used for the classification of the data.

D. Principal Component Analysis (PCA)

This is a most important method which is used to convert the number of observation sets into the values of linearly uncorrelated values called Principal Components, with the help of orthogonal transformation. The process is run in such way that the starting first data is the highest variance taken under the possible constraint. So the results have exhibits orthogonal in nature to the preceding ones. In this way the indentify the data and its pattern and that can be expressed with similarities and deviances can be highlighted [3].

Here, 80 sample values (20 from each of the four types of tea) have been transformed with PCA. The result is shown in Figure 2. Shows the 97% more similarities available in the starting two primary components .More than 97% significant values are available.

Linear Discriminate Analysis (LDA)

WhilePrincipal analysis works orthogonally and reduce the large dimension length or space .This method can be works to improved the compactness of the information by reducing the space as much as possible, Linear Discriminate method tries to reduce the dimensionality preserving the class discriminatory information. PCA does feature classification while Linear Discriminate methodhas not classified. Linear Discriminate methoddoesn’t change the location and shape of the dataset, as done by PCA, but always tri to reduce the space and enhance the efficiency with separate the different kinds of samples presents in the black tea as used in the process and quickly make decision area between the given class in the samples. [4].

The same sample set used earlier, has been transformed with LDA. The c lass independent transform has been implemented here. The result is shown in Figure 3.

E. Independent Component Analysis (ICA)

In a simple notion of experimentation it was assumed that a particular compound in the tea vapour will affect a single and specific sensor out of the five sensor array. However, this is

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

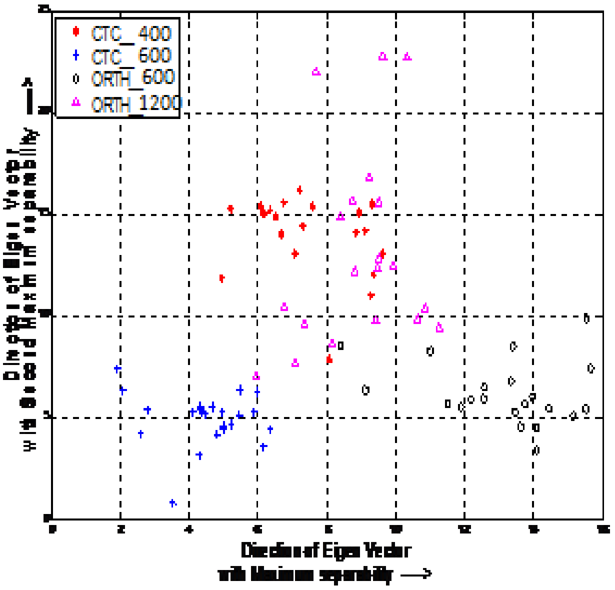

Figure 2 Transformantion results by LDA

not the case in reality. Here lied the necessity of extracting the independent sources individually from what we obtained as the raw data.

Independent Component techniques for sample analysis is based on the fully software computational technique which is quite effectively used to differentiate a multivariable data signal into different components assuming the mutual independence of source signals. It may well be defined as a most highly efficient type of sample distingusition. This methods generally determine the independent pariclesby doing maximum of statistical seperated of the calculated components.

In this paper the method of Maximisation of Non-Gaussianity has been adopted for the necessary data analysis. The Central Limitation of the Probability states that the exhibition of a addition of independent variable particles tends toward a Gaussian distribution by taking some boundary condition. Consider a sample data x is let to be a mixture of independent components as it modeled by Principal Component Thecnique, it may be states as $x = As$ where s define as sourse vector and A is the mixing coefficients which is maximum. Evidently the source s may be expressed as $s = A^{-1}x$. Now this sourse and coefficient data matrix can be find out by neural network iteration .It must be remember that the resultant summation of independent variable is more Gaussian than the original variables and it becomes least Gaussian when the estimated signal equals the original source [5].

In this attempt, the targeted chemical constituents of Black Tea are considered as the source whereas the sensors are considered to be the mixers.

F. ICA in Conjunction with LDA

Though improvement has been achieved in terms of separabilty with application of ICA (viz. Table III), the extracted data is of high dimension (five dimensional). For the reduction of dimension, the extracted data has been subjected to further transformation through LDA. As the result, even better clustering has been achieved. The result is shown in Figure 4.

G. Dunn's Index

Validity of clusters is worth being evaluated after clustering algorithms are applied, to find whether the result is appreciable or not. Different cluster validity algorithms are available for this purpose. In this paper Dunn's Index has been adopted. The main objective is to separate the clusters. The index is the distance to maximal intra-cluster distance. It can be depicted by the following expression:

$$D = \min_{1\le i\le n}\left\{ \min_{1\le j\le n, i\ne j} \left\{ \frac{d(i,j)}{\max_{1\le k\le n} d'(k)} \right\} \right\}$$

when d(i,j) defined as the path between two clusters i and j and d'(k) is the intra difference of cluster k. The inner cluster distances have been measured as the distance of centroids of the corresponding clusters. The high and low cluster similarity found in an internal cluster validation. Dunn's Indices are considered to be better and desirable.

TABLE III.DUNN CLUSTERS OF RAW DATA AND THOSE OBTAINED FROM TRANSFORMATIONS OF DIFFERENT ALGORITHMS

Dataset	Dunn's Index (approximate values)
Raw data obtained from QCM sensor response	0.0873
Data transformed by PCA	0.2037
Data transformed by LDA	0.2126
Data transformed by ICA	0.5030
Data transformed by ICA in conjunction with LDA	0.5338

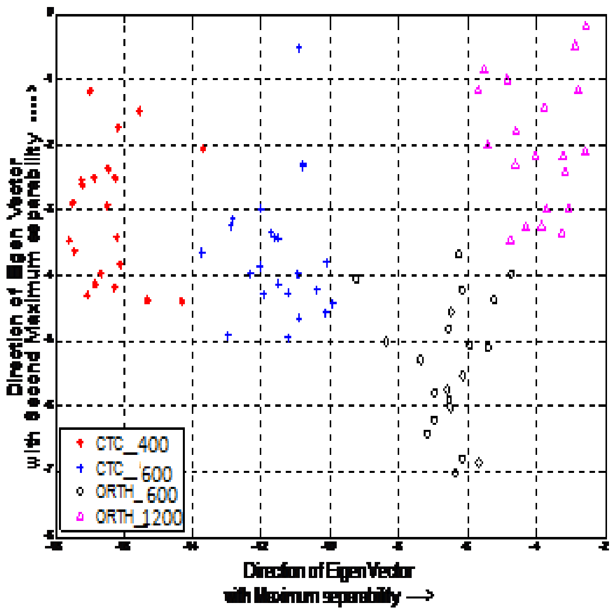

Figure 3.Results of different teasaafter Independent Component Analysis

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

H. BPMLP Model

The BPMLP network consisted of three sensory units that are input layer, hidden layer and output layer. Here input and output layer consists of four neurons each. The working principal of neuron is based on the minimize the weight of the neurons and updated by the new weight and continuously monitoring and minimized the errors. When input is applied and passes through the network then a output appears in the output layers. Then check the error and updated the wight of the neurons. The main target of this work is to minimize the error between input and output neurons with adjustment the weights of the neurons. Here the rule base techniques used to minimize the errors and use a algorithms to reduce the errors between input and output layers.

The main thing is that in back propagation method the variation of the synaptic weight adjusted in a iterative way to reduce the changes the synaptic weights in an iterative manner, so found the output $y_j(n)$ is very closer to the target $d_j(n)$. The error signal of the neuron and target can be defined as

$$e_j(n) = d_j(n) - y_j(n)$$

The weights are changed till the error cannot be minimized any further.

TABLE IV.VALIDATION FOR FOLD AND CLASSIFICATION RATE

Fold	**Number of Data: Train/Test**	**Data Misclassified**	**Classification rate**
1	63/7	4	74
2	63/7	2	73
3	63/7	4	74
4	63/7	2	73
5	63/7	0	84
6	63/7	0	88
7	63/7	0	96
8	63/7	0	100
9	63/7	0	100
10	63/7	0	100
Misclassified Data		12	
Classification Rate Average		71.83%	
Standard deviation(SD)		11.62%	

IV. CONCLUSION

Most of the essential consumer products are tea in the daily life of the people. It is very important to find the best quality of the tea .Because of the presence of the different particles and compounds and aroma, flavor etc., it is very critical and complex work. Black tea is generally classified with its aroma and taste .The teste is sensed by the human nose tester and the classified by the results obtained by the human tester after the experiment the compounds of teas and finally decide the rate of the tea and placed in the market with market price. There are many problems are arise in the results found from the human tester due to inaccuracy, local atmosphere effect, undesired particle present in the sample and is laborious .Also a large required for analysis the different human factors such as different human variability. The main disadvantage is that the sensitivity may be decreased due to the uncelebrated of tools used and human mistake due to the long time involvement and trace. This problem can be overcomethen demand a new technology like Electronic Nose by this we overcome the mentioned problems. Electronic nose technologies have been used in various food product and aro product and to improve the teste of the food products, viz., wine, cola. meat, fish etc. Here the Quartz Crystal Microbalance type sensor id developed to analysis the aroma of the different black teas ,food products and obtained the resuts very satisfactory as well as very promisticThe system of electronic nose have been able to challenge the apprehensions that are attached to fresh concept and are able to classify the various black tea qualities when tagged with a suitable pattern recognition engine. Moreover the samples are to be collected from geographically dispersed tea gardens to bulid up the knowledge base. This criterion is challenging as samples collected from the same gardens in different sensors of flush exhibit varying characteristics.

REFERENCES

1. S.Kar,P.Sharma,P.J.Soren, "QCM anlysis for Black tea discriminations," IEEE Trans.Inst. Meas,.vol.58,pp.1318-1322,2010.
2. B. Bhatia,A.Paul, Tea manufacturing palant. New Delhi, India: Oxford and Kanara Publications., 1996.
3. S.Choudhury: A Comprehensive study on Neural network Englewood Cliffs, NJ: Prentice-Hall, 2001.
4. D.H. Donald, G.E. Hinton, and T.J. Sejnowski, "A leanring algorithms on neural network," Cognitive Science, 1995
5. J.N. Foreman and D.S. Balaswami, "Neural Networks: Algorithms, Applications, and Matlab ," Addison Wesley, 1999

Published By: Blue Eyes Intelligence Engineering & Sciences Publication