Quality Assessment of Tea Employing Electronic Tongue (ET): Exploration of Feature Extraction Techniques and Prediction Algorithms

Synopsis of the Thesis submitted

by

Srikanta Acharya

Doctor of Philosophy (Engineering)

Department of Instrumentation and Electronics Engineering

Faculty Council of Engineering & Technology

Jadavpur University, Kolkata, India

2024

JADAVPUR UNIVERSITY

KOLKATA- 700032

INDIA

INDEX NO. 168/19/E

Title of the Thesis **Quality Assessment of Tea Employing Electronic Tongue (ET): Exploration of Feature Extraction Techniques and Prediction Algorithms**

Name, Designation& Institution of theSupervisor

Dr. Runu Banerjee Roy
Professor
Dept. of Instrumentation and Electronics Engineering Jadavpur University, Salt Lake Campus, Sector-III, Block-LB, Plot-8, Kolkata- 700106, India. Phone: 08240235742, Fax: 03323357254

List of Publications

Journals

[1] **S. Acharya**, D. Das, T. N. Chatterjee, S. Mukherjee, R. B. Roy, B. Tudu, *et al.*, "Voltammetric electrode array optimization for black tea discrimination using computational intelligence approach," *IEEE Sensors Journal*, vol. 21, pp. 20589-20595, 2021.

[2] **S. Acharya,** D. Das, S. Nag, S. Mukherjee, A. K. Hazarika, S. Sabhapondit, *et al.*, "Optimization techniques for a voltammetric signal to predict green tea quality parameters using MIP electrode," *IEEE Sensors Journal*, vol. 23, no. 17, pp. 19842-19847,2023.

[3] **S. Acharya,** S. Nag, D. Bandyopadhyay, D. Das, A. Mandal, and R. B. Roy, "A Molecular Imprinted Polymer Tethered Capacitive Sensor for Epicatechin Detection in Green Tea," *IEEE Sensors Journal*, vol. 24, no. 4, pp. 4213-4220, 2024.

Conference Proceedings

[1] **S. Acharya,** T. N. Chatterjee, S. Mukherjee, D. Das, B. Tudu, R. Bandyopadhyay, *et al.*, "Optimization of electrode array in electronic tongue for classification of black tea," in *2018 IEEE Applied Signal Processing Conference (ASPCON)*, 2018, pp. 134-137.

[2] **S. Acharya,** T. N. Chatterjee, S. Mukherjee, D. Das, R. B. Roy, B. Tudu, *et al.*, "Selection of optimum number of sensors of an electronic tongue for efficient classification of black tea: A combinatorial approach based on discrete cosine transform and artificial neural network," in *2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, 2018, pp. 108-111.

[3] **S. Acharya,** S. Nag, D. Bandyopadhyay, S. Mukherjee, D. Das, and R. B. Roy, "Unveiling the capacitive sensor performance developed for Epicatechin detection in Green tea: A Clustering Approach," in *2024 IEEE 3rd International Conference on Control, Instrumentation, Energy & Communication (CIEC)*, 2024, pp. 175-179.

[4] **S. Acharya,** D. Das,S. Nag, D.Bandyopadhyay,A. Mondal, B. Tudu, P. Pramanik, R. Bandyopadhyay, R. B. Roy, “Gallic Acid Detection Using A PMMA/PDA Coated Capacitive Sensor”, Nano-architectures for Chemical Biological and Therapeutic Applications (NACBTA)2021,November 12-14th, 2021, Department of Chemistry GLA University, Mathura, UP and Jadavpur University**,** Kolkata .

QUALITY ASSESSMENT OF TEA EMPLOYING ELECTRONIC TONGUE (ET): EXPLORATION OF FEATURE EXTRACTION TECHNIQUE AND PREDICTION ALGORITHM

1 INTRODUCTION:

Tea is most popular beverage consumed across the world and a significant cash crop in India. Tea flavor can be divided in to two categories: non-volatile component like taste and volatile component like aroma. Innumerable bio-chemical compounds present in the tea are responsible for determining taste modality [1]. Some of main bio-chemical non-volatile constituents are: polyphenol, amino acid, caffeine, theaflavins and thearubigin. These essential components of tea are responsible for astringent, brothy, bitter, ashy, and slightly astringent tastes. The gradation of tea according to the quality is a very important and critical work. Generally human panel taster are employed for the said purpose, but it suffers from inconsistency and unpredictability due to various human factors like individual variability, decrease in sensitivity due to prolonged exposure, mental and physical fatigue due to mental stress. Assessment of tea quality can be done using different analytical instruments like gas chromatography (GC), high-performance liquid chromatography (HPLC), and capillary electrophoresis (CE) techniques. However, the said techniques are expensive, time consuming and laborious. So need of instruments that are fast, repeatable, reliable and compatible with human tester score appear to top priority.

Electronic tongue (ET) or artificial tongue is an attempt to mimic the human tongue in an instrumental way. It is capable of differentiating between substances with various taste profiles and can also distinguish between different substances that produce the same taste modality. Different sensors forming the arrays in an ET system emulate the human receptors which are responsible for sensing taste. Correlation between human taste response and sensor data can be established using various pattern recognition techniques for both qualitative and quantitative analysis of test samples. The sensor array in an ET operates on liquid samples, and the combined response of the ET sensors differs from one solution to another. Three electrode voltammetry systems consist of: an Ag/AgCl reference electrode, a platinum counter electrode and a working electrode. The electrodes are actuated by the different types of voltage such as Differential pulse Voltammetry (DPV) and Cyclic Voltammetry (CV). Sensor based on Molecular Imprinted Polymer (MIP), Ploymer-Graphite Composite

Electrode are cost effective and successful to determine and distinguish the different bio-chemical constituents present in tea [2-4].

Data obtained from the electrochemical sensors are large in size and correlated. So it is very much necessary to extract the significant pattern or information from the large raw data set. Data obtained from sensor response contained the redundant data that are responsible to degrade the information signature. This redundant data should be eliminated to enhance the classification and prediction accuracy of the sensor. Several feature transformation and dimension reduction techniques exist within the pattern recognition paradigm. Among these, Principal Component Analysis (PCA) [5,6], Discrete Cosine Transform (DCT) [7,8], Singular Value Decomposition (SVD) [9], and Independent Component Analysis (ICA) [10] have been utilized in this work.

Data analysis for pattern recognition can be accomplished after pre-processing and feature extraction technique [11-13]. Various classifiers such as Artificial Neural Network (ANN), Support Vector Machine (SVM), and K-Nearest Neighbor (KNN) have been utilized for classification purposes, while feature optimization has been achieved using Genetic Algorithm (GA), Bat Algorithm (BA), and Whale Optimization Algorithm (WOA). The biochemical constituents in tea samples were assessed using standard linear regression models, such as Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR) [14]. A novel, cost-effective, and reusable capacitive sensor was fabricated for epicatechin (EC) detection in green tea. The sensor performance is then evaluated using two widely-used clustering algorithms: K-Means and Agglomerative Clustering.

2. MOTIVATION AND RESEARCH OBJECTIVES:

The literature review indicates that researchers have utilized various feature transformation techniques to extract meaningful insights from collected data, including dimension reduction and noise elimination. Additionally, both supervised and unsupervised machine learning techniques have been explored for overall quality evaluation and tea classification, leveraging biochemical component measurement through the integration of sensors with machine learning. However, the survey identified some scope of areas such as electrode arrays optimization, features optimization to enhance the performance of regression model using computational methods, developing a cognitive computing framework to

enhance prediction accuracy, and creating low-cost, durable, and sensitive sensors for detecting biochemical components in tea. The objectives of this research work are as follows:

- To investigate more meaningful and appropriate features from the sensor response employing features transformation and compression techniques.
- To optimize the number of voltammetric electrodes in an array of an electronic tongue.
- To reduce the data size, expedite computational processes, and employ various feature optimization techniques to improve the efficiency of the model.
- To predict the biochemical content present in green tea.
- To develop and characterize the low-cost, sensitive, and novel sensors for detecting specific molecules present in tea samples.

3. ORGANIZION OF THESIS

The thesis has six chapters. Chapter 1 introduces the topic, literature review of research along with the objective and motivations behind the work. It outlines the manufacturing processes for different tea grades, explores tea quality assessment, and discusses ET along with its sensor types. A concise overview of ET and its various sensing technologies has been presented.

Chapter 2 of this dissertation explores a variety of techniques for pattern recognition and feature transformation, such as Principal Component Analysis (PCA), Discrete Cosine Transform (DCT), Singular Value Decomposition (SVD), and Independent Component Analysis (ICA) concisely. The section also presents a concise overview of unsupervised machine learning methodologies like K-means clustering and agglomerative clustering, as well as supervised learning approaches such as Support Vector Machines (SVM), k-Nearest Neighbors (KNN), ensemble bagging trees, decision trees, and artificial neural networks (ANN). Furthermore, it examines methods for feature optimization, including Genetic Algorithms (GA), Bat Algorithm (BA), and Whale Optimization Algorithm (WOA), in addition to regression models like Partial Least Squares Regression (PLSR) and Principal Component Regression (PCR).

Chapter 3 of this thesis introduces voltammetric electrode array optimization technique using computational approach to distinguish between different varieties of black tea. Polymer graphite composite electrodes (PGE) were developed using varying ratios of

three monomers: acrylamide, aniline, and pyrrole, combined with graphite. Electrode array optimization was performed using four feature extraction techniques Principal Component Analysis (PCA), Discrete Cosine Transform (DCT), Singular Value Decomposition (SVD), and Independent Component Analysis (ICA) and five classification algorithms: Support Vector Machine (SVM), k-Nearest Neighbor (KNN), ensemble, decision tree, and discriminant analysis, followed by a polling process.

Chapter 4 introduces a machine learning model that has been formulated to improve the precision of predicting green tea quality in electrochemical systems by employing sensor data obtained from electrodes made of molecularly imprinted polymers (MIPs). The sensor data from two MIP electrodes, namely MIP-GAL and Q-IPG, undergo transformation through the Discrete Cosine Transform (DCT) method. The process of data optimization is carried out utilizing bio-inspired metaheuristic algorithms such as the genetic algorithm (GA), bat algorithm (BA), and whale optimization algorithm (WOA). Subsequently, these refined feature sets are utilized in constructing prediction models using partial least squares regression (PLSR) and principal component regression (PCR) to enhance the accuracy of predictions.

Chapter 5 explores the fabrication of a robust, discriminating, reuseable and economical capacitive sensor employing molecularly imprinted polymer (MIP) methodology to identify epicatechin (EC) in green tea. Subsequently, the efficacy of the sensor is evaluated through the application of two commonly utilized clustering techniques: K-Means and Agglomerative Clustering.

Finally, Chapter 6 presents the conclusion, summery of findings and future scope of the research work.

4. CONTRIBUTION OF THE THESIS

Based on the research objectives, the contributions of this thesis are described chapter wise.

Chapter 3 is designed to work upon the objectives 1 and 2. This chapter contributed the study and analysis of pattern recognition and feature transformation techniques along with the novel polling method to optimize the electrodes in an array.

Chapter 4 is designed to work upon the objectives 3 and 4. This chapter contributed the optimization of features of MIP electrodes to enhance the prediction performance of the regression model for measuring the biochemical content present in green tea.

Chapter 5 is designed to work upon the objective 5. This chapter contributed the fabrication of novel, cost-effective, reusable and robust, MIP tethered capacitive sensor for epicatechin detection in green tea. The performance of the synthesized sensor was evaluated using two popular clustering algorithm.

References:

[1] B. Banerjee, *Tea Production and Processing*. New Delhi, India: Oxford & IBH Publishing Co., 1993.

[2] T.N. Chatterjee, D.Das, R.B.Roy,B.Tudu, S.Sabhapondit,P.Tamuly,P.Parmanik,R. Bandyopadhyay, “Molecular imprinted polymer based electrode for sensing catechin(+C) in green tea,” IEEE Sensor Journal,18(2018)2236-2244.

[3] T.N. Chatterjee, D.Das, R.B.Roy, B. Tudu, A.K.Hazarika, S.Sabhapondit, P.Tamuly ,R. Bandyopadhyay, “Development of nickel hydroxide nanopetal decorated molecular imprinted polymer based electrode for sensitive detection of epigallocatechin-3-gallate in green tea,” Sensor and Actuators B: Chemical, (2019) 69-78.

[4] S.Acharya, S.Nag, D.Bandyopadhyay, D.Das, A.Mandal, and R.Banerjee Roy, “A Molecular Imprinted Polymer Tethered Capacitive Sensor for Epicatechin Detection in Green Tea ", IEEE SensorJournal,pp.4213-4220,vol.24,no.4,2024,DOI:10.1109/JSEN.2023.3344668.

[5] P. Ciosek, K. Brudzewski, and W. Wróblewski, "Milk classification by means of an electronic tongue and Support Vector Machine neural network," Measurement Science and Technology, vol. 17, p. 1379, 2006.

[6] L. Lvova, A. Legin, Y. Vlasov, G. S. Cha, and H. Nam, "Multicomponent analysis of Korean green tea by means of disposable all-solid-state potentiometric electronic tongue microsystem," Sensors and Actuators B: Chemical, vol. 95, pp. 391-399, 2003.

[7] I. Batal and M. Hauskrecht, "A supervised time series feature extraction technique using dct and dwt," in 2009 international conference on machine learning and applications, 2009, pp. 735-739.

[8] B. Schwerin and K. Paliwal, "Local-DCT features for facial recognition," in 2008 2nd International Conference on Signal Processing and Communication Systems, 2008, pp. 1-6.

[9] M. Narwaria and W. Lin, "SVD-based quality metric for image and video using machine learning," IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 42, pp. 347-364, 2011.

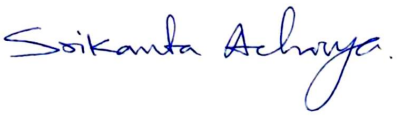
[10] X. Zhang, V. Ramani, Z. Long, Y. Zeng, A. Ganapathiraju, and J. Picone, "Scenic beauty estimation using independent component analysis and support vector machines," in Proceedings IEEE Southeastcon'99. Technology on the Brink of 2000 (Cat. No. 99CH36300), 1999, pp. 274-277.

[11] H. Xiao and J. Wang, "Discrimination of Xihulongjing tea grade using an electronic tongue," *African Journal of Biotechnology*, vol. 8, 2009.

[12] C. Söderström, F. Winquist, and C. Krantz-Rülcker, "Recognition of six microbial species with an electronic tongue," *Sensors and Actuators B: Chemical*, vol. 89, pp. 248-255, 2003.

[13] M. Palit, B. Tudu, N. Bhattacharyya, A. Dutta, P. K. Dutta, A. Jana, *et al.*, "Comparison of multivariate preprocessing techniques as applied to electronic tongue based pattern classification for black tea," *Analytica chimica acta*, vol. 675, pp. 8-15, 2010.

[14] S. V. Suryakala and S. Prince, "Investigation of goodness of model data fit using PLSR and PCR regression models to determine informative wavelength band in NIR region for non-invasive blood glucose prediction," Optical and Quantum Electronics, vol. 51, pp. 1-20, 2019.



SIGNATURE OF CANDIDATE **SIGNATURE OF SUPERVISOR**

Srikanta Acharya ***Dr. Runu Banerjee Roy***