A preliminary solution for anomaly detection in water quality monitoring

C. Bourelly
Sensichips srl
Aprilia, Italy
carmine.bourelly@sensichips.com

A. Bria, L. Ferrigno, L. Gerevini, C. Marrocco, M. Molinara
Dept. of Electrical and Information Engineering
University of Cassino and Southern Lazio
Cassino, Italy
{a.bria; ferrigno; luca.gerevini; c.marrocco; m.molinara}@unicas.it

G. Cerro
Dept. of Medicine and Health Sciences
University of Molise
Campobasso, Italy
gianni.cerro@unimol.it

M. Cicalini, A. Ria
University of Pisa
Pisa, Italy
{mattia.cicalini@phd.; andrea.ria@ing.}unipi.it

***Abstract*—In smart city framework, the water monitoring through an efficient, low–cost, low–power and IoT–oriented sensor technology is a crucial aspect to allow, with limited resources, the analysis of contaminants eventually affecting wastewater. In this sense, common interfering substances, as detergents, cannot be classified as dangerous contaminants and should be neglected in the classification. By adopting classical machine learning approaches having a finite set of possible responses, each alteration of the sensor baseline is always classified as one out of the predetermined substances. Consequently, we developed an anomaly detection system based on one-class classifiers, able to discriminate between a recognized set of substances and an interfering source. In this way, the proposed detection system is able to provide detailed information about the water status and distinguish between harmless detergents and dangerous contaminants.**

***Index Terms*—water monitoring, sensor network, artificial neural network, machine learning**

I. Introduction

Wastewater can be defined as water polluted by human activities as a byproduct of domestic, industrial, commercial or agricultural activities. Wastewater is unsuitable for its direct use as contaminated by different types of organic and inorganic substances that represent a hazard to public health and environment. For this reason they cannot be returned directly to the environment since the final deliveries such as land, sea, rivers and lakes are not able to receive a quantity of polluting substances greater than their self-purifying capacity without compromising the normal balance of the ecosystem. In a smart city framework, it can be very interesting to map the wastewater collection system and detect sources of pollution within a metropolitan area, in order to optimize water treatment systems or identify anomalies in wastewater, which could indicate fraudulent spills.

Sensing technologies for water pollution monitoring have been widely studied in the scientific literature, e.g. see the extensive reviews in [1]–[3]. The proposed methods mainly focus on the application of electrodes of different metals [4] or covered by sensing films [5], and optical sensors [6]. Measurements from sensors are usually based on Electrochemical Impedance Spectroscopy (EIS) [7]–[9], a frequency domain technique that evaluates the response of an electrochemical cell to a low amplitude sinusoidal perturbation. After measurement acquisition, data analysis for contaminants classification is principally based on machine learning methods. For example, in [10], Artificial Neural Network and Principal Component Regression are used to estimate nitrate concentration in ground water, whereas in [4] partial least square discriminant analysis is applied to detect explosive precursors in wastewater. Also deep learning solutions have been proposed, e.g. in [11] Long Short Term Memory (LSTM) and Convolutional Neural Network (CNN) were adopted for detection and classification of chemicals in sea water.

To ensure a widespread diffusion of pollutants detection systems in wastewater, low cost, compactness and low consumption (which would make them suitable even for low maintenance) become fundamental. An ideal monitoring system should be able to detect and/or distinguish between thousands of different substances, even in very complex conditions of high water turbidity, sensor degradation, which can quickly become important, and so on. Furthermore, such a system should be able to distinguish polluting substances also on the basis of their dangerousness level and treatment processes. This work is part of a project where the ultimate goal is the creation of an end-to-end identification system (from sensing to classification) capable of detecting a predefined set of substances commonly considered as dangerous and indicative of an anomalous use of water. All the substances considered are characterized by complete solubility in water, which makes them more difficult to identify and filter. The proposed measurement system is based on a proprietary embedded IoT–ready Micro-Analytical Sensing Platform (MASP henceforth) of size 1.5×1.5 mm and 1.5 mW power absorption. The

978-1-7281-6997-2/20/$31.00 ©2020 IEEE
DOI 10.1109/SMARTCOMP50058.2020.00086

Authorized licensed use limited to: Zhejiang University of Technology. Downloaded on September 10,2024 at 07:46:40 UTC from IEEE Xplore. Restrictions apply.

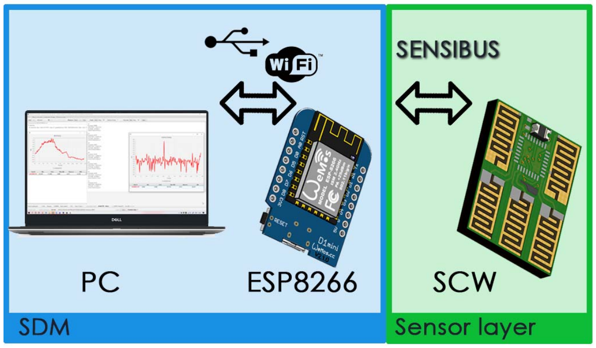

Fig. 1. The adopted acquisition and processing system

same MASP technology has been used in several studies for contaminants detection [12]–[15]. Due to time and cost issues of collecting measures for any substance that could be found in wastewater, the faced machine learning problem should be considered as an ”open-set recognition” problem, i.e., a classification problem where the training data only gives a partial view of the application domain. A first step to solve such a problem is the detection of unknown substances before the classification of the known dangerous contaminants. Thus, in this paper, stemming from our past experience in sensors and data processing [16]–[19], we present an anomaly detection system based on one class classifiers able to discriminate between a known set of substances and an interfering source so as to reduce or avoid wrong classifications or false alarms. Experiments have been conducted on a dataset realized in our laboratories to show that the proposed system is able to provide detailed information about the water status and distinguish between harmless detergents and dangerous contaminants.

II. The SENSIPLUS ecosystem

The overall system, namely SENSIPLUS, is reported in Fig. 1 and is composed of a sensor layer (also known as Smart Cable Water (SCW)) and a processing sub-system, namely Sensiplus Deep Machine (SDM). The core chip named SENSIPLUS is a proprietary technology of Sensichips developed in collaboration with the University of Pisa. It is possible to perform measurements with internal and external analog ports, allowing to work with multiple sensors. Specifically, to work with wastewater, the system has been customized by hosting it on a printed circuit board endowed with both measuring and sensing capabilities.

The physical principle adopted to achieve the goal is the electrical impedance. It has been measured through the MASP's Analog Front End (AFE). By the analysis of the single components, it is possible to see:

- the SCW, endowed with 6 Interdigitated Electrodes (IDEs) metalized through different materials;
- the processing object (SDM), composed of: (i) an ESP32 by Espressif Micro Controller Unit (MCU) running a customized software, managing the measurement layer and transmitting acquired data to a Host PC through USB or Wi–Fi technologies; and (ii) a host controller (PC equipped with Windows/Linux OSs, Android devices (Smartphone, Tablet)). All is organized through the SENSIPLUS Application programming interface (API) allowing data exchange between the MASP and external applications such as the classification module working on pre–trained machine learning models to classify substances.

The MCU works as a trans-coder bridge between USB/WiFi and a proprietary one-wire protocol, able to communicate with the SCW itself, named SENSIBUS. It is managed through the bit-banging mode GPIO pin control.

III. The measurement activity

A. Measurement Environment

An ideal solution to acquire measurements suitable to build a training set would be to operate directly in a controlled drain of a sewage network. This, however, would pose biological hazards for human health due to the presence of bacteria, viruses, parasites, and other dangers. In addition, there are also measurement issues since sewage composition is not constant due to rain and various domestic spills. These inhomogeneities could invalidate training measurements and deceive our algorithms.

To overcome these problems, our solution is to create a synthetic wastewater (SWW) simulating the sewage composition. Our adopted recipe for SWW is a simplified version of the mixture created by [20], where all substances in trace are removed. Every batch of SWW is corrected by pH accordingly to [21] measurements of real wastewater. Correction is made with NaOH to increase pH and HCl to decrease it.

B. The adopted hardware set–up

The SCW, depicted in Fig. 2, is a MSP board engineered by Sensichips s.r.l. It is designed to integrate on the same package the sensing elements and the measurement chip. It can be flooded in water since the measuring part is protected by an acrylic resin. On SCW's front face there are five small (estimate surface 9.95 mm^2 each) interdigitated electrodes (IDEs) made by copper and metalized with gold, copper(oxide), nickel, silver, palladium. Back board has only one big (estimate surface 19.61 mm^2) IDE still copper based and metalized with platinum.

C. Measured features

Electrical circuit reported in Fig. 3 represents the equivalent circuit (Randles) of two electrodes flooded in a water solution (called bulk). Each electrode has a double layer capacitance: *Cd* and faradic resistance: *Rf*, both take into account the interface electrodes bulk and depends on electrode composition, area, geometry, roughness, bulk composition, etc. *Re* is the bulk equivalent resistance and depends on its conductivity and electrodes surface area. Measures are performed with five of the six SCW's IDEs: Gold, Platinum, Copper, Nickel,

Authorized licensed use limited to: Zhejiang University of Technology. Downloaded on September 10,2024 at 07:46:40 UTC from IEEE Xplore. Restrictions apply.

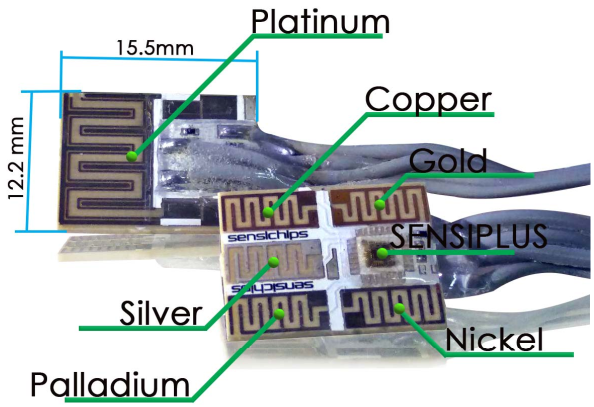

Fig. 2. Sensichips' Smart Cable Water board

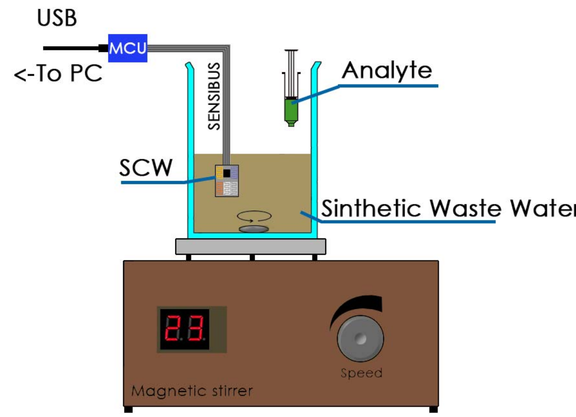

Fig. 4. Setup

Silver. For the first two sensing elements, two frequencies are evaluated: 78125 Hz, and 200.32 Hz. For the other IDEs only one frequency (200.32 Hz) is measured. According to Randles equivalent circuit, high frequency impedance is mainly defined by *Re* because in high frequency double layer capacitance has a low impedance and short circuit *Rf*. *Re* depends on bulk, therefore is useless measure it with all SCW's front sensors since all of them have similar surface area. With low frequency double layer capacitance impedance tends to be an open circuit and measurements also depend on faradic resistance over bulk resistance.

D. Measurement Setup and Measurement Protocol

Measurement setup (Fig. 4) is composed of the measurement chain (SCW, MCU, PC), a magnetic stirrer, 300 ml beaker, 8 Analytes (9, including background), 3 Interferants. To avoid false positives classification due to spills of non-dangerous substances, that could fool the classification algorithm, a series of measurements with only most common interferants were carried out. A magnetic stirrer equipped with 25 mm anchor simulates the water movement: all measurements are performed under the same water stirring conditions. Anchor rotation is set to 50 rpm. At this speed, 100 ml of water do not incorporate air bubbles, otherwise contributing to obtain a noisy measurement. In order to have any measure independent from the others, cleaning all glasses is fundamental. Beaker is cleaned with soap and fresh water or sometimes with concentrated 95% sulphuric acid.

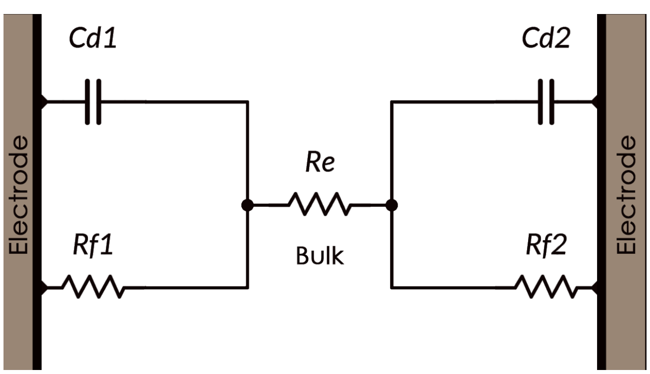

Fig. 3. Randles equivalent circuit

The measurement protocol is divided into three sections:

- First 300 samples are acquired in warm–up mode, thus not considered as useful data and consequently discarded;
- 300 samples of baseline: in this period of time, sensors are exposed to SWW only;
- 1000 samples of sensor evolution after analyte injection.

IV. The Classification Activity

A. Baseline and Normalization

When working with sensor data analysis, data pre–processing is a very important task in order to make the feature data easily understandable by the algorithm. Typically, the drifting behaviors that affect sensor measurements as environmental effects, sensor poisoning, sensor drift etc. cause the raw measurements to be unsuitable for classification purposes. To deal with them, considering their slow dynamics, we implemented an *Exponential Moving Average* (EMA), which applies exponentially decreasing weights to the raw data, so that older data are progressively neglected with respect to the new ones. In detail, the EMA analytical formulation is reported in Eq. 1:

$$\boldsymbol{s}_t = \begin{cases} \boldsymbol{f}_t, & t = 1 \\ \alpha \boldsymbol{f}_t + (1-\alpha)\cdot \boldsymbol{s}_{t-1}, & t > 1 \end{cases} \tag{1}$$

where $\boldsymbol{f}_t$ represents the feature vector, $\boldsymbol{s}_t$ is the value of the EMA at time t and α is the degree of weighting decrease

Authorized licensed use limited to: Zhejiang University of Technology. Downloaded on September 10,2024 at 07:46:40 UTC from IEEE Xplore. Restrictions apply.

[1]. To compensate the slow dynamics of the before mentioned factors, we choose a decaying coefficient α that evolves until reaching a threshold value, as follows:

$$\begin{cases} \beta(1) = 1 \\ \vdots \\ \beta(n) = \beta(n-1) + n \end{cases} \quad \text{with} \quad \begin{cases} \alpha = \dfrac{1}{\beta(n)} > \mathrm{A} \\ n \in \mathbb{N} \end{cases} \tag{2}$$

where A represents the lower bound value until α keeps evolving. We chose a decaying α with an empirically chosen lower bound A equals to 10^{-4}, in order to build a robust baseline that is able to compensate the drift effect caused by environment, sensor drift etc. From now on, all new data will be normalized with the respect to the baseline according to the following equation:

$$\boldsymbol{n}_t = \frac{\boldsymbol{f}_t}{\boldsymbol{s}_t} \tag{3}$$

Where $\boldsymbol{n}_t$ is the feature vector after the normalization at time t, $\boldsymbol{f}_t$ is the feature vector at time t and the $\boldsymbol{s}_t$ represent the value of EMA at t.

B. Classification

Our final goal is to build a multiclass classifier system with the capability of distinguish among a known set of substances flowing through the wastewater. Given the complexity of the real scenario where every day there are plenty of spills of different nature (e.g. organic materials, soaps, industrial products etc.), we have to face a problem known as "open-set recognition" [22], [23]. In a real context, we must avoid to input samples of unknown classes to our final multiclass classifier system, otherwise we will end up by giving the wrong results, which would make our system useless. Accordingly, here we propose the use of One-Class classifiers for the Anomaly Detection [24], [25], used as a kind of false positive reduction system.

In order to find the most suitable classifier we compared three solutions using (according to the used algorithm) two different approaches:

- *Novelty Detection:* the training set contains only positive samples and we are interested to find out if a new sample is an outlier (aka novelty) [26];
- *Outlier Detection:* the training set is contaminated with some outlier samples and we are interested to characterize the region where most of the training data are concentrated [27].

In particular, for the novelty detection we used the One-Class Classifier SVM [28] whereas for the outlier detection we used Isolation Forest and Elliptic Envelope [29].

[1]NB: a smaller α decreases the weights of older data slowly.

C. Data Set

The dataset consists of 10 data acquisitions for each of the 9 substances of interest (8 pollutants plus water), totaling 90 data acquisitions carried out with the measurement protocol described in Section III. Each acquisition contains 1,600 samples, from which we have removed the first 600 samples, necessary to achieve a "sensor's stabilization" phase (first 300 samples) and to build a robust baseline (the remaining 300 samples), and thus the last 1,000 samples of each acquisition have been used to build up our dataset. Consequently, the entire dataset contains 90,000 samples.

To evaluate the ability of our classifiers to correctly predict the class of a new unseen sample we used 10–fold cross-validation. In each of the 10 iterations, 9 acquisitions (one for each substance) were used as validation set, and the remaining 81 acquisitions (9 for each substance) as training set.

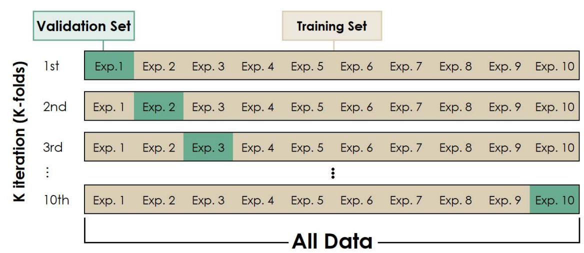

Fig. 5. K-Fold Cross-Validation

To use the outlier detection classifiers we enriched the training set of each fold with the addition of 9,000 samples of interfering substances (aka outliers) consisting of washing machine detergent, dishwasher detergent and sodium chloride, so that they represent 10% of the training set.

In summary, as concerns the novelty detection, each fold was divided in two parts:

- *training set:* 81 acquisitions each formed by 1,000 samples containing only positive data for a total of 81,000 samples;
- *validation set:* 9 acquisitions for a total of 9,000 samples containing only positive data;

whereas for the outlier detection classifiers each fold consisted of:

- *training set:* 81 acquisitions each formed by 1,000 samples containing only positive data, and 9,000 outlier samples taken from three different interfering substances, for a total of 90,000 samples;
- *validation set:* 9 acquisitions for a total of 9,000 samples containing only positive data;

The performance of each classifier was evaluated on a separate test set composed of 10,000 samples of 3 interfering substances, for a total of 30,000 test samples.

Authorized licensed use limited to: Zhejiang University of Technology. Downloaded on September 10,2024 at 07:46:40 UTC from IEEE Xplore. Restrictions apply.

TABLE I
TPR RESULTS USING BEST HYPERPARAMETERS.

Classifier	Fold1	Fold2	Fold3	Fold4	Fold5	Fold6	Fold7	Fold8	Fold9	Fold10	Mean	SD
OCCSVM	85.32	100.0	99.82	98.06	100.0	87.99	100.0	64.61	82.54	70.39	88.87	12.49
Isolation Forest	89.63	100.0	100.0	97.23	100.0	98.2	99.52	87.22	79.99	82.81	93.46	7.42
Elliptic Enveloper	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	91.98	93.69	98.57	2.89

TABLE II
TNR RESULTS ON TEST SET.

	Fold1	Fold2	Fold3	Fold4	Fold5	Fold6	Fold7	Fold8	Fold9	Fold10	Mean	SD
OCCSVM	99.67	99.67	99.67	99.67	99.67	99.69	99.67	99.69	99.68	99.68	99.68	0.008
Isolation Forest	66.84	60.73	52.57	52.22	61.00	60.07	62.77	64.68	63.82	68.26	61.3	5.09
Elliptic Enveloper	82.24	83.73	82.73	81.18	82.55	80.45	83.27	85.35	92.41	90.97	84.49	3.83

D. Performance Evaluation

As performance evaluation metrics, we computed the True Positive Rate (TPR, where a True Positive is a sample classified as positive and labeled as positive in the ground truth too) for the validation set and the True Negative Rate (TNR, where a True Negative is a sample classified as negative and labeled as outlier in the ground truth too) for the test set (as shown in eq. 4) of each classifier over all the training/validation sets and test sets.

$$\begin{aligned} \mathit{TPR} &= \frac{\mathit{True\ Positive}}{Positive\ Samples} \\ \mathit{TNR} &= \frac{\mathit{True\ Negative}}{Negative\ Samples} \end{aligned} \tag{4}$$

E. Hyperparameters

As regards the Isolation Forest and Elliptic Envelope, there are no real hyperparameters, since the only parameter to be set is the *contamination*, that represents the proportion of outliers in the data set. In our case, since we have contaminated the training set with 10% (see Section IV-C) of outlier data, we have used this value as contamination parameter for both classifiers. Regarding the One-Class Classifier SVM (OCCSVM) we have different hyperparameters to choose:

- ν: the probability of finding a new, but regular, observation outside the frontier and thus it represents an upper bound on the fraction of training errors and a lower bound of the fraction of support vectors;
- γ: the inverse of the radius of influence of samples selected by the model as support vectors (i.e. high value causes overfitting);
- *Kernel*: the kernel function used by the algorithm.

Since the nature of our problem is not linearly separable, we have chosen the Radial Basis Function (RBF) as kernel function. As to ν and γ, at each cross-validation step the pair of values that maximised the TPR was selected. At the end of the iterative process we compute the mean and the standard deviation of the TPR over all the 10 folds.

V. RESULTS

In Table I we report the 10–fold cross-validation results in terms of TPR (mean and standard deviation) related to the three used classifiers, obtained by applying the best hyperparameters found as described in Section IV-E. In Table II, we report the results of the trained models on the test set, in terms of TNR. As can be seen, despite the results related to Table I, the best performance on test set has been obtained by One-Class Classifier SVM while the worst one by the Isolation Forest.

VI. CONCLUSIONS

In this paper an anomaly detection in a water monitoring system has been presented. Two approaches have been proposed: a Novelty Detection system and an Outlier Detection system respectively. The problem at hand is inherently an Open Set Classification problem where many substances could be encountered and identified as an outlier. For the experimental phase three substances have been considered as an outlier: washing machine detergent, dishwasher detergent and sodium chloride.

The TNR obtained on test set (99.68 % with OCCSVM) gives the evidence that the proposed approach is promising. In future works the number and typology of substances utilized as an outlier will be expanded and the proposed system will be combined with the final classifier in order to evaluate the performances from an overall point of view.

VII. ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement SYSTEM No. 787128. The authors are solely responsible for it and that it does not represent the opinion of the Community and that the Community is not responsible for any use that might be made of information contained therein.

This work was also supported by MIUR (Minister for Education, University and Research, Law 232/216, Department of Excellence).

REFERENCES

[1] S. Zhuiykov, "Solid-state sensors monitoring parameters of water quality for the next generation of wireless sensor networks," *Sensors and Actuators B: Chemical*, vol. 161, no. 1, pp. 1 – 20, 2012.

Authorized licensed use limited to: Zhejiang University of Technology. Downloaded on September 10,2024 at 07:46:40 UTC from IEEE Xplore. Restrictions apply.

[2] M. H. Gholizadeh, A. M. Melesse, and L. Reddi, “A comprehensive review on water quality parameters estimation using remote sensing techniques,” *Sensors*, vol. 16, no. 8, p. 1298, 2016.
[3] S. N. Zulkifli, H. A. Rahim, and W.-J. Lau, “Detection of contaminants in water supply: A review on state-of-the-art monitoring technologies and their applications,” *Sensors and Actuators B: Chemical*, vol. 255, pp. 2657 – 2689, 2018.
[4] C. Desmet, A. Degiuli, C. Ferrari, F. S. Romolo, L. Blum, and C. Marquette, “Electrochemical sensor for explosives precursors’ detection in water,” *Challenges*, vol. 8, no. 1, 2017.
[5] J. K. Atkinson, M. Glanc, M. Prakorbjanya, M. Sophocleous, R. P. Sion, and E. Garcia-Breijo, “Thick film screen printed environmental and chemical sensor array reference electrodes suitable for subterranean and subaqueous deployments,” *Microelectronics International*, Apr. 2013.
[6] T. P. Lambrou, C. C. Anastasiou, C. G. Panayiotou, and M. M. Polycarpou, “A low-cost sensor network for real-time monitoring and contamination detection in drinking water distribution systems,” *IEEE Sensors Journal*, vol. 14, no. 8, pp. 2765–2772, Aug 2014.
[7] A. M. Syaifudin, K. Jayasundera, and S. Mukhopadhyay, “A low cost novel sensing system for detection of dangerous marine biotoxins in seafood,” *Sensors and Actuators B: Chemical*, vol. 137, no. 1, pp. 67–75, 2009.
[8] X. Li, K. Toyoda, and I. Ihara, “Coagulation process of soymilk characterized by electrical impedance spectroscopy,” *Journal of Food Engineering*, vol. 105, no. 3, pp. 563–568, 2011.
[9] P. Geng, X. Zhang, W. Meng, Q. Wang, W. Zhang, L. Jin, Z. Feng, and Z. Wu, “Self-assembled monolayers-based immunosensor for detection of escherichia coli using electrochemical impedance spectroscopy,” *Electrochimica Acta*, vol. 53, no. 14, pp. 4663 – 4668, 2008.
[10] G. Charulatha, S. Srinivasalu, O. Uma Maheswari, T. Venugopal, and L. Giridharan, “Evaluation of ground water quality contaminants using linear regression and artificial neural network models,” *Arabian Journal of Geosciences*, vol. 10, no. 6, p. 128, Mar 2017.
[11] S. N. Dean, L. C. Shriver-Lake, D. A. Stenger, J. S. Erickson, J. P. Golden, and S. A. Trammell, “Machine learning techniques for chemical identification using cyclic square wave voltammetry,” *Sensors*, vol. 19, no. 10, 2019.
[12] M. Ferdinandi, M. Molinara, G. Cerro, L. Ferrigno, C. Marrocco, A. Bria, P. Di Meo, C. Bourelly, and R. Simmarano, “A novel smart system for contaminants detection and recognition in water,” in *2019 IEEE International Conference on Smart Computing (SMARTCOMP)*, June 2019, pp. 186–191.
[13] M. Molinara, M. Ferdinandi, G. Cerro, L. Ferrigno, and E. Massera, “An end to end indoor air monitoring system based on machine learning and sensiplus platform,” *IEEE Access*, vol. 8, pp. 72 204–72 215, 2020.
[14] A. Bria, G. Cerro, M. Ferdinandi, C. Marrocco, and M. Molinara, “An iot-ready solution for automated recognition of water contaminants,” *Pattern Recognition Letters*, vol. 135, pp. 188–195, 2020.
[15] G. Betta, G. Cerro, M. Ferdinandi, L. Ferrigno, and M. Molinara, “Contaminants detection and classification through a customized iot-based platform: A case study,” *IEEE Instrumentation Measurement Magazine*, vol. 22, no. 6, pp. 35–44, Dec 2019.
[16] G. Cerro, M. Ferdinandi, L. Ferrigno, and M. Molinara, “Preliminary realization of a monitoring system of activated carbon filter rli based on the sensiplus® microsensor platform,” in *2017 IEEE International Workshop on Measurement and Networking (M N)*, Sep. 2017, pp. 1–5.
[17] P. Bruschi, G. Cerro, L. Colace, A. De Iacovo, S. Del Cesta, M. Ferdinandi, L. Ferrigno, M. Molinara, A. Ria, R. Simmarano, F. Tortorella, and C. Venettacci, “A novel integrated smart system for indoor air monitoring and gas recognition,” in *2018 IEEE International Conference on Smart Computing (SMARTCOMP)*, June 2018, pp. 470–475.
[18] G. Cerro, M. Ferdinandi, L. Ferrigno, M. Laracca, and M. Molinara, “Metrological characterization of a novel microsensor platform for activated carbon filters monitoring,” *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 10, pp. 2504–2515, Oct 2018.
[19] D. Capriglione, G. Cerro, L. Ferrigno, and G. Miele, “Analysis and implementation of a wavelet based spectrum sensing method for low snr scenarios,” in *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2016, pp. 1–6.
[20] I. Nopens, C. Capalozza, and P. A. Vanrolleghem, “Stability analysis of a synthetic municipal wastewater,” *Department of Applied Mathematics Biometrics and Process Control, University of Gent, Belgium*, 2001.
[21] H. Janna, “Characterisation of raw sewage and performance evaluation of al-diwaniyah sewage treatment work, iraq,” *World Journal of Engineering and Technology*, vol. 4, no. 2, pp. 296–304, 2016.
[22] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T. E. Boult, “Toward open set recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1757–1772, 2013.
[23] W. J. Scheirer, L. P. Jain, and T. E. Boult, “Probability models for open set recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 11, pp. 2317–2324, 2014.
[24] B. Schölkopf, R. Williamson, A. Smola, J. Shawe-Taylor, and J. Platt, “Support vector method for novelty detection,” in *Proceedings of the 12th International Conference on Neural Information Processing Systems*, ser. NIPS’99. Cambridge, MA, USA: MIT Press, 1999, p. 582–588.
[25] P. Bodesheim, A. Freytag, E. Rodner, M. Kemmler, and J. Denzler, “Kernel null space methods for novelty detection,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2013, pp. 3374–3381.
[26] B. Lamrini, A. Gjini, S. Daudin, P. Pratmarty, F. Armando, and L. Travé-Massuyès, “Anomaly detection using similarity-based one-class svm for network traffic characterization,” in *DX@Safeprocess*, 2018.
[27] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” *ACM Comput. Surv.*, vol. 41, no. 3, Jul. 2009. [Online]. Available: https://doi.org/10.1145/1541880.1541882
[28] M. Amer, M. Goldstein, and S. Abdennadher, “Enhancing one-class support vector machines for unsupervised anomaly detection,” in *Proceedings of the ACM SIGKDD Workshop on Outlier Detection and Description*, ser. ODD ’13. New York, NY, USA: Association for Computing Machinery, 2013, p. 8–15. [Online]. Available: https://doi.org/10.1145/2500853.2500857
[29] Xulei Yang, Qing Song, and A. Cao, “Weighted support vector machine for data classification,” in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, vol. 2, July 2005, pp. 859–864 vol. 2.

Authorized licensed use limited to: Zhejiang University of Technology. Downloaded on September 10,2024 at 07:46:40 UTC from IEEE Xplore. Restrictions apply.