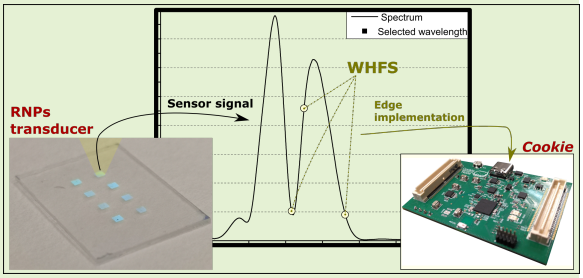
A Machine Learning-based Methodology for in-Process Fluid Characterisation with Photonic Sensors

Rodrigo Marino, *Student Member, IEEE,* Sergio Quintero, Andres Otero, *Member, IEEE,* Jose M. Lanza-Gutierrez, and Miguel Holgado

***Abstract*— This paper proposes a novel methodology for run-time fluid characterization through the application of machine learning techniques. It aims to integrate sophisticated multi-dimensional photonic sensors inside the chemical processes, following the Industry 4.0 paradigm. Currently, this analysis is done offline in laboratory environments, which increases the decision-making times. As an alternative, the proposed method tunes the spectral-based machine learning solutions to the requirements of each case enabling the integration of compound detection systems at the computing edge. It includes a novel feature selection strategy that combines filters and wrappers, namely Wavelength-based Hybrid Feature Selection, to select the relevant information of the spectrum (i.e., the relevant wavelengths). This technique allows providing different trade-offs involving the spectrum dimensionality, complexity, and detection quality. In terms of execution time, the provided solutions outperform the state-of-the-art up to 61.78 times using less than 99% of the wavelengths while maintaining the same detection accuracy. Also, these solutions were tested in a real-world edge platform, decreasing up to 68.57 times the energy consumption for an ethanol detection use case.**



***Index Terms*— Chemical Monitoring, Edge Computing, Feature Selection, Machine Learning, Optical Sensors.**

I. Introduction

PROCESS monitoring plays a significant role in Industry 4.0 since it allows fault detection, prevents harmful accidents, and minimizes quality losses in the final products [1]. In these smart factories, the production stage adapts continuously according to the information provided by the processing monitoring. For some physicochemical parameters, such as temperature or pressure, conventional gauges usually offer a direct mathematical model between the traced process variables and the physical phenomena measured by the transducer. However, the development of sophisticated acquisition systems has increased data dimensionality and complexity, making it difficult to provide a mathematical model relating it to the manufacturing process variables. Artificial Intelligence (AI) techniques based on Machine Learning (ML) have gained importance to process multi-dimensional measurements in this scenario [2]. However, ML-based processing solutions have high computational demands, which forces moving data processing to centralized in-factory or even remote computing clouds. This fact goes against the decentralization objective in Industry 4.0 paradigm, which stands for keeping data processing techniques near the sensors since the response time and costs are reduced when the decision-making tasks of the operations are close to the production lines [3].

In the particular case of the chemical industry, product quality is evaluated typically in specialized laboratories, usually placed far from the operation field. Thus, manufacturers usually require to send the laboratory multiple samples of the products and byproducts generated during the fabrication process to analyze their chemical composition. Performing the analytics in the laboratory includes some shortcomings, such as the process is error-prone due to high human intervention, time-consuming, and expensive due to the usage of specialized instrumentation and highly trained personnel [4], [5]. This analysis process feeds directly into the decision-making tasks affecting its performance. Therefore, moving the instrumentation from the laboratory to the operation field is a fundamental challenge in the chemical industry.

In this regard, optical sensors based on micro and nanostructures could be the technology required to move the instrumentation to the operation field thanks to their reduced size, excellent sensitivity, adequate resolution, short response time, and excellent usability [6]. In contraposition to standard optical sensors based on classical spectroscopy techniques,

R. Marino and A. Otero are with the Centro de Electrónica Industrial, Universidad Politécnica de Madrid, 28006 Madrid, Spain (email: {rodrigo.marino, joseandres.otero}@upm.es).

J.M. Lanza is with the Department of Computer Science, Universidad de Alcalá, 28871, Alcalá de Henares, Spain (e-mail: jm.lanza@uah.es).

S. Quintero and M. Holgado are with the Center for Biomedical Technology ; Optics, Photonics and Biophotonics Group, Universidad Politécnica de Madrid, Campus Montegancedo, 28223 Pozuelo de Alarcón, Madrid, Spain; (email: {sa.quintero, m.holgado}@upm.es).

those based on micro and nanostructures decreases the cost of the equipment in exchange for poorer sensing capabilities, making sensor factory distribution a viable option. Note that whereas conventional optical transducers pursue a general chemical analysis, transducers based on nanostructures can be specialized to detect specific compounds for each particular stage of the process [7]. However, optical transducer improvements alone are not enough to enable distributed run-time process monitoring of fluid media, as will be discussed below.

In the literature, there is a wide variety of optical sensors based on micro and nanostructures. For instance, ring resonators [8], surface plasmon resonators [9], and Mach-Zehnder, Young, and Fabry-Perot interferometers [6]. In the particular case of Fabry-Perot interferometers, there is one especially appropriate for in-field implementation thanks to its easy alignment and high endurance. These transducers, dubbed as Resonant NanoPillars (RNPs) arrays, produce a sensing signal composed of the spectral response at a certain wavelength range, usually in the visible field [10]. Therefore, output data is a high dimensional vector, which hards to extract relevant information to the autonomous decision taken in the application. Thus, it is necessary to count on specific near-sensor processing algorithms to translate the multi-dimensional measurements into useful information. In this regard, this paper proposes using near-sensor ML techniques to extract patterns from the gathered RNPs spectral responses and correlate them with the desired property, particularly the amount of a specific compound in a fluid media. Embedded ML techniques for optical transducers within an edge device reduce the system latency and enable early fault detection strategies [10]. Moreover, the distribution of these advanced sensing systems alongside the factory would increase the productivity and the robustness of the whole process.

Thus, a framework termed the Wavelength-based Hybrid Feature Selection (WHFS) method is proposed in this paper as a feature selection technique for optical transducer wavelengths. It aims at identifying the more relevant wavelengths from the spectral response produced by the RNPs. By relevance, the authors mean how each wavelength, also known as a feature using ML terminology, correlates with the property to be predicted. The method ensures that the wavelengths selected are distributed alongside the spectral response and that they are not concentrated in a particular bandwidth region, guaranteeing the diversity of the obtained information. Also, ordered features are removed increasingly to ensure that the minimum level of performance quality is maintained [11]. The proposal is fully scalable. Different sets of wavelengths and trained ML models are provided at design time, so the designer might choose the solution that fits application needs better, depending on the expected quality and the available processing resources.

As the use case in this work, the percentage identification of ethanol existing in a fluid matrix was selected. The spectral response during the experimentation was registered using the RNPs optical transducers. This specific optical technology usage does not compromise the proposed method generality because no particular assumption was taken about the industrial process and the optical transducer. Compared to the state-of-the-art [10], the solutions provided in this paper achieve equal performance metrics, even improving them in particular cases. However, the solutions proposed in this paper outperform the state-of-the-art in terms of the number of wavelengths used, needing less than 99 % of the available wavelengths to obtain the same performance metrics values, resulting in less energy cost for the capturing process. Furthermore, the feasibility of ethanol detection during run-time according to the proposal is proved by implementing the selected solutions in a low-power edge computer.

Therefore, this paper presents the following contributions. i) A novel proposal for compound detection based on ML techniques using nano-structure optical transducers is presented. ii) A wavelength-based ML solution framework for wavelengths selection according to the application requirements is proposed. iii) The achieved solutions obtained by the framework are implemented in a low-power edge device, that is, the trained ML models for the selected particular wavelengths.

The rest of this paper is organized as follows. In Section II, state-of-the-art of photonic transducers and ML photonic data are reviewed. In Section III, the proposed method is detailed. In Section IV, the dataset and the evaluation procedure are depicted, whereas in Section V the results of the evaluation are presented. In Section VI concluding remarks and future perspectives are highlighted.

II. State-of-the-art

This section first reviews significant examples showing the use of photonic transducers to analyze fluid media. Second, a survey of the recent studies in edge-compatible ML techniques for compound detection using photonic sensors is provided.

A. Photonic Transducers for Chemical Analysis

Photonic sensors offer the resolution, response-time, and usability required for in-field chemical analysis of fluid samples [4], [12]. Besides, they are immune to electromagnetic noise, simplifying its installation and usability in certain environments. In this sensing technology, the properties of the chemical sample modify the sensor response when stimulated with specific types of light. Thereby, these transducers can be interrogated in different wavebands to extract specific properties, including Mid-InfraRed (MIR) [4], Near-InfraRed (NIR) [13], visible [14], or UltraViolet (UV) [12].

Commercial optical sensing equipment, as reported in [4], [12] and [15], is oriented to be installed in specialized laboratories, where they are employed with different types of samples and analyses. This flexibility implies a high cost and complexity of the system, meaning that this equipment can not be integrated into real-world process monitoring systems.

These practical limitations preventing the in-field use of photonic sensors can be overcome by miniaturizing the transducers using micro-and nano-technology, with benefits in cost, response time, size, and high usability. However, they still have limitations in resolution and sensing capacity when compared to conventional spectroscopy techniques. A novel photonic transducer technology based on nanostructures, the RNPs transducers, is emerging in this context. RNPs are composed

of several, usually hundreds, nanometric pillars, in which each of them acts as a single optical resonator in where the light is coupled. The reduced size of each nanopillar and their proximity produce an evanescence electric field outside the RNPs. Thus, the RNPs make a unique optical interference pattern, that is, the sensor footprint, at a particular wavelength range, which depends on the refractive index of the media the RPNs are immersed in. Note that, since the sensing principle of this sensor is based on the evanescence field, the sensing signal is less sensitive to temperature changes than sensors based on optical gratings. However, temperature changes can affect the chemical sample properties. Thus it is highly recommended to perform all the experiments at stable temperature conditions. The combination of a high light confinement effect and a highly-defined shape of the optical response offered by RNPs are behind their high sensing capacity, being reported their use in different applications in the chemical and biochemical fields [7], [16]–[21].

The novel methodology proposed in this paper is applied to the experimental results obtained in the previous work [10], where all the experimental data was gathered at the laboratory, in which temperature was controlled with an accuracy of ± 1 degree. In that paper, the RNP sensing technology was used to analyze the ethanol concentration in different fluid matrices, i.e., deionized water and white wine. The employed RNPs transducer presented nanopillars with a diameter and height of 250 and 2000 nm, respectively. The nanopillars array had a squared distribution with a pitch of 500 nm. When analyzed with a white light source and spectrometer, the RNPs produced a high-resolution interference pattern in the visible band, generating a high data dimensionality footprint, which was considered to design and train an ML system to detect the ethanol concentration. Differently, in this work, the authors propose a novel methodology. It consists in analyzing the changes in the RNPs footprint by measuring the intensity of the RNPs spectra in specific wavelengths. This approach is based on an ML solution, maintaining the same performance in the detection as using the whole spectra.

B. Machine Learning for Photonic Sensors Data Analytics

The output data produced by optical transducers interrogated by broadband light source and spectrometers, including the RNPs-based sensors selected in this work, is directly their spectral response in the waveband of interest, e.g., visible or NIR bands [22]. This continuous response produced by the transducers is then sampled in several wavelengths that determine the resolution of the acquisition system. The sampled optical response is finally available, from which a model relating it with the desired property can be created. Traditionally, this model is generated by trained personnel based on previous expert knowledge, identifying the implicit physical phenomena between the measurement and the property. An example of this process was reported in [19] to model the percentage of sodium chloride in a water matrix. Nonetheless, the higher is the acquisition resolution, the higher is data dimensionality, making it more difficult to understand the relationship between the input variables and the produced response.

This increased complexity in dimensionality makes then convenient to apply data-driven ML techniques to create multivariate models [23]. In this regard, ML permits extracting a pattern from data, then correlating the intensity values to the desired property. Nevertheless, processing ML techniques might require a high computational and energy cost, which are not desirable in an edge layer. Therefore, strategies, such as dimensionality reduction, might be explored. To this end, there are two main dimensionality reduction strategies in ML: Feature Extraction (FE) and Feature Selection (FS) [24], [25]. FE compresses the input data, transforming a high-dimensional data space into a low-dimensional space [26]. However, this strategy requires the acquisition of the complete signal (the waveband) to later apply the transformation. FS chooses a relevant subset of features (wavelengths) from the problem input space (the waveband), which can result in reducing the number of wavelengths acquired [27]. This latter dimensionality reduction approach fits with the focus of this paper.

There are three main FS strategies: filter, wrapper, and hybrid methods [25]. Filter strategies select features by statistical procedures, such as measuring the correlation between variables and the property, then removing redundant or irrelevant features. The purpose is to reduce the noise produced by these redundant or irrelevant features, which negatively affect the prediction quality. The authors may cite some works considering filter approaches [28]–[30]. In [28], they applied the *t-test* statistical method to identify liver fibrosis. In [29], they estimated the redundancy of the variables (i.e., wavelengths) and removed those which surpassed a previously fixed threshold for protein detection. Filter techniques include a relevant limitation related to the selection obtained, that is, this strategy is independent of the ML model in the system. This fact might entail that some irrelevant features are maintained, which could decrease the prediction performance. On the contrary, both wrapper and hybrid strategies tailor FS to application needs by considering the ML algorithms to be included in the final system and use case data.

Wrapper strategies evaluate combinations of features iteratively, searching for the combination whose performance is better according to the use case. The search finishes when a combination meets the stop condition, which is a threshold related to a specific ML performance metric, such as accuracy [31]–[34]. Opposite to filter strategies, wrapper ones evaluate a given combination considering the ML model in the system, usually relying on classification/regression supervised learning. Note that classification consists of predicting a discrete variable and regression involves predicting a continuous variable. The authors may cite some works considering wrapper approaches, determining if they removed or added features during the selection iterative process. In [31] the authors applied a backward elimination algorithm that started with the initial set of features and in each iteration removed a subset. In [33] the method began with no feature, adding a new feature in each iteration, namely forward selection. The advantage of forwarding selection is that the resultant set of solutions has lower dimensionality than backward elimination.

Wrapper strategies also include shortcomings, such as that

the iterative strategy hinders the convergence of the method in high-dimensional problems, such as when processing optical responses in a high dimensionality spectrum. As a solution, some authors proposed splitting the spectrum into regions, which can be discarded regarding their influence on the use case. For instance, in [32], the authors proposed that a domain expert decided the dimension and the number of the regions. Therefore, the spectrum division depended on the previous use case knowledge of the designer. Similarly, in [31], the authors proposed an algorithm that automatically divided and selected the regions. This algorithm alleviated the computational cost. However, the spectrum division based its procedure on combinatory as it is a pure wrapper instead of attending to the relationship between the region and the application use case. Hence, there was no guarantee that discarded regions did not have relevance, which would decrease the ML performance.

Finally, hybrid approaches combine filter and wrapper methods to select features. The authors in [30] created a hybrid approach that starts with an iterative filter process for redundancy removal, based on the Pearson correlation, deleting the redundant features. The filter stage also required establishing a threshold defined by a domain expert. Afterward, it applied a wrapper strategy based on particle swarm optimization for choosing the best combination of features for its use case.

Focusing on works considering FS techniques in the specific optical domain, classification-based ML models were used to aid FS techniques for choosing features. The authors in [32] built an FS approach based on convolutional neural network classifiers for fluid characterization, such as beer identification. In [31], the authors developed an FS based on Support Vector Machine (SVM) classifiers to identify steel aging. The authors in [29] applied the Partial Least-Square Regression (PLSR) for protein determination in milk powder. In [33], they considered artificial neural networks to identify antioxidants in certain oils. In the regression domain, the authors in [34] applied a transformation of classifiers to regression techniques (in particular, SVM regression) to estimate the rice root density. This type of approach usually has a classifier structure, but its output stage aims to convert the prediction into a regression.

As far as the authors know, the state-of-the-art regarding ethanol detection by photonic transducers based on nanostructures consists of using Principal Component Analysis (PCA) for FE [35]. Before the ML inference, PCA compresses the input space with 3468 wavelengths into a few variables, also known as Principal Components (PCs), using the Gram-Schmidt orthogonalization. This approach has a lack related to the high dimensionality of the input space. This fact might prevent its implementation in low-power resource-aware edge platforms. Differently, this paper aims to provide a framework based on FS techniques for ethanol detection with photonic sensors based on nanostructures, trying to reduce the wavelengths required for the input space, whereas state-of-the-art results are achieved. The proposed method follows a hybrid FS for regression models to meet this requirement. Although the focus is on the RNPs transducer, this method can be generalizable to any photonic transducer as it works directly with the transducer output, i.e., the intensity of consecutive wavelengths in a particular waveband. In a previous work, the authors of this paper proposed a hybrid FS approach for a fault detection use case in the Industry 4.0 domain [35]. Beyond the completely different use case, this previous work differs from the one provided in this paper as follows. The proposal incorporates an automated region split method for identifying the relevant spectral regions for the application, addressing the wrapper shortcoming discussed before. It also integrates a filter strategy to reduce the redundancy of the features selected by the wrapper. Moreover, a relevant aspect of this proposal is that instead of returning a unique solution to the problem, the framework returns a set of trade-off solutions for ML performance and the number of wavelengths. This focus will help system designers to develop photonics-based edge-oriented ML systems according to the application needs, such as maximizing the ML performance or reducing the acquisition system cost by minimizing the number of wavelengths.

III. Proposed Method

This section starts by describing the methodology overview for the application of the WHFS method. Next, a detailed explanation of the WHFS method and the wavelength removal filter procedure is presented. Finally, a discussion about WHFS parameterization is included.

A. Methodology Overview

The proposed ML-based methodology for processing the response of photonic sensors encompasses both training and inference stages, as shown in the flowchart in Fig. 1. Training generates a set of ML models, relating the raw data of sensors with the wanted physical property. This stage requires high computational resources and is performed offline, in a workstation, before the field deployment. The multidimensional nature of the photonic sensor response makes it interesting to select a subset of optical wavelengths during the training instead of using all the sensor wavelengths. Thus, only those which are relevant will feed the ML model. This fact could result in better performance and lesser sensor costs. In the ML context, each wavelength is treated as a feature, and then both terms are used indistinctly along with the discussion. One of the trained models is used to estimate the expected properties from future raw sensor measurements in the inference. This latter stage is executed online in the edge node near the sensors, implying that computational complexity should be reduced as much as possible.

The training stage is performed considering the WHFS algorithm. WHFS is a hybrid method, including both filter and wrapper stages. The filter stage evaluates the relevance of the wavelengths. Moreover, as the WHFS method works in the spectral domain, an analysis of spectral dependencies is also integrated into the filter stage to remove redundant wavelengths. During the wrapper stage, it is needed to train the ML model several times during the selection process. The model is trained using supervised learning techniques, meaning that a labeled dataset is required. Each instance of the training dataset must include the wavelengths (i.e., the intensity values produced by the sensor in each waveband) and the label (i.e., the theoretical quantity value of the compound

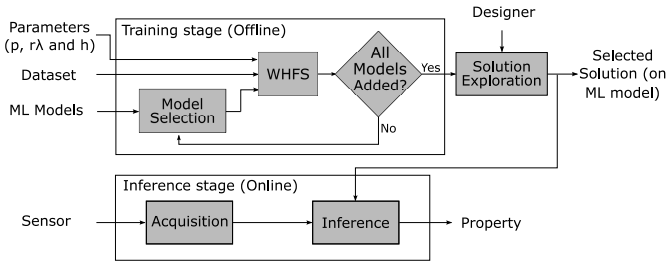


Fig. 1: Methodology flowchart.

to be analyzed). The execution of the WHFS provides a space of solutions, being a solution defined in this context as the trained ML model and its subset of selected wavelengths. In this regard, this space shows a trade-off between the compound detection quality and the number of chosen wavelengths in each solution (i.e., the complexity of the model). The designer might consider this trade-off to select the solution that suits better to the application requirements.

The inference stage focuses on online liquid property prediction. To this end, an ML model is selected among the previously generated, then implementing in an embedded system to be deployed in the field.

The solutions generated by this methodology are based on data whose origin is a specific source, i.e., a transducer in this use case. That means that the models developed are specialized to work with this specific photonic transducer, which is the origin of the data. This fact is due to the differences observed in the transducer performance regarding the manufacturing process. Although the differences in performance are minimal, they could still affect the parameterization of the models obtained. Thus, the recommendation is to apply the methodology for each transducer to be used in the system.

B. Wavelength-based Hybrid Feature Selection

As shown in Algorithm 1, the WHFS computation comprises two stages: the filter stage (lines 1-2) and the wrapper stage (lines 3-22), both described below. The algorithm requires a matrix $X \in \mathbb{R}^{m \times n}$ with m instances for n wavelengths and a vector $V \in \mathbb{R}^{m \times 1}$ with m regression values (labels), one per instance. Let $models$ be the vector with the t regression ML models to evaluate during the algorithm execution, $models \in \mathbb{R}^{t \times 1}$. Let $p \in \mathbb{R}^1$ be the maximum number of wavelengths that a model might have. As will be discussed below, p can also be defined as the number of solutions that will be explored for each ML model in $models$. Let $h \in \mathbb{R}^1$ be the number of wavelengths evaluated during the selection. Note that $models$, p, and h parameters configure the WHFS execution. The tuning of these parameters is described in Section III-D. The $r\lambda \in \mathbb{R}^1$ and $sRes \in \mathbb{R}^1$ parameters are associated with the distance between wavelengths to establish regions within the waveband. A further description of these two parameters is found in the splitting process discussion in Section III-C.

1) Filter stage: It performs an initial removal of the wavelengths according to their intensity value. The filter stage first applies the Fisher-score algorithm, which is a filter feature selection technique based on the relevance of each wavelength in the use case data (X and V). The relevance is obtained by analyzing the features and labels of the dataset, studying similarities and divergences [36]. In this regard, this technique permits the removal of irrelevant wavelengths according to the use case perspective, alleviating the search in the wrapper stage. Fisher-score was applied in a wide variety of applications. For instance, in [37], the authors applied this technique to a biomedical application, minimizing the number of features required for the detection of sleep apnea using ML algorithms. In [38], the technique was applied in the industry domain, obtaining the features that monitored the condition of the bearings in wind turbines. However, this initial filtering action might not ensure the decrease of redundancy. This fact is because wavelengths with close positions in the spectrum share information related to the physical process, and then they might have the same relevance. As relevance highly influences the wrapper stage, close wavelengths can mask other wavelengths with relevant use case information. To address this problem, the filter stage splits the waveband into various regions based on the $r\lambda$ parameter. This splitting process results in removing those wavelengths which do not give relevant information to the system based on the region information. This removal method is explained in Section III-C. The filter stage ends gathering a subset of wavelengths, $Wv \in \mathbb{R}^{n' \times 1}$ (line 1). This Wv vector is the ranked set of n' features represented by their IDs in X, with $n' \in 1, \ldots, n$. Next, a copy of Wv is saved in Wv_{old} to restore in the wrapper stage the information obtained during the filter (line 2), as will be discussed below.

2) Wrapper stage: It performs an exploration of the solution space for each model using a modified Sequential Forward Selection (SFS) method. In the literature, SFS was applied in a wide variety of applications. For instance, it was used in the biochemical industry to choose the minimum set of features to identify antioxidants proteins [39]. It was also used in hyperspectral cameras to select the most suitable wavebands to monitor rice seeds [40]. Regarding the combination of Fisher-score and SFS in WHFS, it was shown as advantageous in previous work for a problem related to the Industry 4.0 field [35]. Focusing on the proposal in this paper, the purpose of the modified SFS considered is to find an optimal minimum subset of wavelengths without distorting data [40], where each feature selected maximizes the prediction performance of the ML algorithm. As a result, the solution provided by this technique utilizes the minimum number of wavelengths, which results in reducing the energy and execution time of the models generated. This wrapper stage requires four parameters:

- The ranked ID vector (Wv) generated in the filter stage.
- The set of regression models ($models$) to evaluate.
- The maximum number of wavelengths (p) that a model might have.
- The maximum number of wavelengths evaluated (h) from the Wv vector in each z-th iteration, with $z \in 1 \ldots p$.

In this stage, the algorithm will seek p solutions for each ML model (lines 3-22). The solution space for the y-th model, with $y \in 1, \ldots, length(models)$, follows an SFS approach

(lines 5 -18). Thus, the first solution of the model will have one wavelength, the second one, two wavelengths, and the p-th one, p wavelengths. This fact is the reason for the double meaning of the p parameter. Besides, during the construction of a solution, only one wavelength is searched to speed up the process, reusing the wavelengths found in the previous solution. That is, the z-th solution uses the wavelengths of the $(z-1)$-th, with $z \in 2, \ldots, p$). As expected, when z equals 1, the algorithm does not reuse any wavelength. To implement this functionality, it is required to save the wavelengths found in each z iteration, namely D_z, to apply then a union operation (line 16). Note that it is also required to define an empty vector D_0 to maintain the union operation coherence when z equals 1 (line 4).

Once a model is chosen, the algorithm starts searching the p solutions (line 5). Before beginning the wavelength search, the algorithm initializes the performance metric threshold (q_{th}, line 6), which saves the minimum ML performance that a solution should provide. The ML performance metric should be selected by the system designer among those available in the literature. Next, it looks for the wavelength that adapts better to the model and the use case (lines 7 - 14). To this end, in each iteration, it selects the j-th wavelength from Wv to evaluate its performance from the ML metrics perspective. As stated before, this value, Wv_j, points out a position inside the original dataset, X. Thus, the algorithm creates an iteration dataset, X', which combines the dataset of the previous solution, D_{z-1}, and the instances of the particular Wv_j wavelength, $X\{Wv_j\}$ (line 8). Then, the model is trained and evaluated while adding the Wv_j wavelength, obtaining its performance metric values (line 9). The model performance metric is compared to the threshold q_{th} (lines 10 -13). If the Wv_j wavelength outperforms the previously selected one, q_{th} is updated (line 11), and this wavelength is marked (line 12). Once the h wavelengths for this solution are evaluated, the marked wavelength is added to the Ws vector (line 15) that stores the selected wavelengths. Besides, the instances of the marked wavelength are added to the model dataset (D_z, line 16). This wavelength is removed from the Wv vector to avoid redundancy so that it is not repeated in the $(z+1)$ iteration (line 17). Before running the exploration of the $(y+1)$ solution, the Wv vector is restored with its original value W_{old} (line 21), which was already saved in line 2. When the execution ends, WHFS returns two arrays of arrays, W_y and S_y (lines 20 - 21). The first includes the individual wavelengths selected for the z solutions generated for the y models. The second includes all the wavelengths that compose the z solutions generated for the y models.

Algorithm 1 Wavelength-based Hybrid Feature Selection.

Require: $X, V, models, p, h, r\lambda, sRes$
Ensure: S_y, W_y
1: $Wv \leftarrow RegWaveRemoval(X, V, r\lambda, sRes)$
2: $Wv_{old} \leftarrow Wv$
3: **for** $y = 1$; $y \leqslant length(models)$; $y++$ **do**
4: $D_0 \leftarrow \emptyset$
5: **for** $z = 1$; $z \leqslant p$; $z++$ **do**
6: $q_{th} \leftarrow 0$
7: **for** $j = 1$; $j \leqslant min(h, size(Wv))$; $j++$ **do**
8: $X' \leftarrow D_{z-1} \cup X(Wv_j)$
9: $q_{metric} \leftarrow evaluateModel(X', V)$
10: **if** $q_{metric} > q_{th}$ **then**
11: $q_{th} \leftarrow q_{metric}$
12: $mark \leftarrow Wv_j$
13: **end if**
14: **end for**
15: $Ws_z \leftarrow mark$
16: $D_z \leftarrow D_{z-1} \cup X(mark)$
17: $Wv \leftarrow Wv - mark$
18: **end for**
19: $S_y \leftarrow \{D_1, \ldots, D_z\}$
20: $W_y \leftarrow \{Ws_1, \ldots, Ws_z\}$
21: $Wv \leftarrow Wv_{old}$
22: **end for**

C. Wavelength Removal Filter Stage

In a high-resolution spectrum, there are regions (subsets of consecutive wavelengths of the spectrum) that concentrate more relevance from a filter point of view. A wrapper stage that treats each wavelength independently will prioritize those wavelengths with higher relevance, which might drive to gather redundant wavelengths. To tackle this problem, it is required to incorporate in the analysis the region relevance. In this regard, the proposed wavelength-based filter method identifies potential regions to search wavelengths based on relevance metrics, such as Fisher-score and the spectral wavelength dependencies. This method also attends to the redundancy in each region. Since the spectral domain is continuous, closer wavelengths share more information, which implies more redundancy between them. Therefore, when the wavelength-based filter selects a particular wavelength from a region, it also removes an interval of wavelengths adjacent to the chosen one from that region. The $r\lambda$ parameter establishes the set of adjacent wavelengths to be deleted. Let $r\lambda \in \mathbb{R}^1$ be defined as the range of the spectral region used during the wavelength removal method, whose value is the size in nm of that region. The value assignment for this parameter is described in Section III-D. As shown in Algorithm 2, the wavelength-based filter works as follows:

- Step 1 (line 1): the fisher-score technique calculates the relevance of each wavelength (i.e., the fisher coefficient) from the dataset, X. Thus, the fisher-score vector of the waveband, $FScore \in \mathbb{R}^{nx1}$, contains the relevance value of each wavelength in X, as given by

$$FScore = [FScore_1, FScore_2, \ldots, FScore_n]^T, \quad (1)$$

  where $FScore_j \in FScore$ is the fisher-score relevance value for the j-th wavelength in the waveband, with $j \in 1, \ldots, n$. Note that the position that each wavelength occupies in this vector is the same they occupy within the waveband.
- Step 2 (lines 5-15): it comprises a technique to divide the spectrum into regions according to the relevance of the wavelengths. Fig. 2 shows an example dividing the

spectrum into five regions, where each one is characterized by its wavelength number and relevance, both factors influencing the selection. It is required to provide a quantitative definition for the regions to perform this division. A region is defined whenever the relevance of the spectral response (Fisher-score values of consecutive wavelengths) surpasses a given threshold. Thereby, first, the algorithm has to establish a threshold (line 5), which is chosen as the mean value of the whole set of Fisher-score values, $mFScore \in \mathbb{R}^1$, also represented in Fig. 2. Then, an iterative process for region split is performed in lines 6-15. Thus, each wavelength of the spectrum is analyzed consecutively, checking if its relevance value surpasses the spectrum mean fisher-score value upwards (line 7) or downwards (line 9). The starting and stopping points of the region are defined in the first and second conditions, respectively. Thus, the number of regions, $nReg \in \mathbb{R}^1$, is established dynamically as the algorithm discovers a new region, $Reg \in \mathbb{R}^{nRegx2}$. The highest relevant value of each region is stored in $maxFSReg \in \mathbb{R}^{nRegx1}$ (line 11).

- Step 3 (lines 16-18): all the spectral values outside the selected regions ($!(Reg)$) are marked (line 16) and then removed from the $FScore$ vector (line 17). Then, the regions are sorted in descending order attending to their $maxFSReg$ value (line 18).
- Step 4 (line 19): the $r\lambda$ parameter cannot be applied directly to the wavelength removal method because a spectral region is also defined by the spectral resolution ($sRes$), i.e., the distance in nm between two consecutive wavelengths. This resolution is typically determined by the acquisition system. Hence, $r\lambda u$ stores the number of wavelengths of the interval defined by $r\lambda$ (line 19).
- Step 5 (lines 20-34): the wavelength removal is an iterative process in which the regions are processed according to their relevance, which is defined in Reg. In each iteration, the wavelength with the highest relevance in the corresponding region is selected. The position of the selected wavelength in X is saved as $pos\lambda$ (line 23). Once the relevant wavelength from the region is obtained, an interval ($int\lambda$) inside the region is defined by $r\lambda u$ centered in $pos\lambda$ (line 24). This interval gathers the most relevant wavelength within the adjacent ones, which have similar information. Hence, to avoid redundancy, the entire interval is marked (line 25) and then removed from the $FScore$ vector (line 26). Next, to prevent wavelength masking from regions with high fisher-score values, the process is forced to search for a wavelength into another region (lines 27 - 30). Thus, as the regions are sorted by their relevance in descending order, the process chooses a region with a relevance value lower than the previous one (line 27). However, when it reaches the least relevant region, it resets the count, going to the most relevant region (line 29). The process is completed when there is no wavelength in the $FScore$ to be processed, i.e., $sumFScore$ equals zero. As a result, the wavelength vector, $Wv \in \mathbb{R}^{n'x1}$, with $n' \in 1, \ldots, n$, is generated and contains the dataset positions of the subset of selected wavelengths (line 31) as given by

$$Wv = [Wv_1, Wv_2, \ldots, Wv_i]^T, \tag{2}$$

where $Wv_j \in Wv$ is the j-th most relevant selected wavelength, with $j \in 1, \ldots, i$.

Algorithm 2 Wavelength-based filter.

Require: $X, V, r\lambda, sRes$
Ensure: $Wv_i, i \in 1, \ldots, n'$
1: $FScore \leftarrow fisherScore(X, V)$
2: $n_{att} \leftarrow length(X(1,:))$
3: $nReg \leftarrow 0$
4: $val_{prev} \leftarrow -1$
5: $mFScore \leftarrow mean(FScore)$
6: **for** $z = 0$; $z \leqslant n$; $z++$ **do**
7: **if** $FScore(z) \geqslant meanFScore$ && $val_{prev} < mFScore$ **then**
8: $Reg(nReg, 1) \leftarrow z$
9: **else if** $FScore(z) < meanFScore$ && $val_{prev} \geqslant mFScore$ **then**
10: $Reg(nReg, 2) \leftarrow z$
11: $maxFSReg(nReg) \leftarrow max(FScore, intReg)$
12: $nReg \leftarrow nReg + 1$
13: **end if**
14: $val_{prev} \leftarrow FScore(z)$
15: **end for**
16: $FScore_{mark} \leftarrow FScore(!(Reg))$
17: $FScore \leftarrow FScore - FScore_{mark}$
18: $Reg \leftarrow Sort(Reg, maxFSReg)$
19: $r\lambda u \leftarrow r\lambda / sRes$
20: $nReg \leftarrow 0$
21: $i \leftarrow 0$
22: **while** $sumFScore! = 0$ **do**
23: $pos\lambda \leftarrow max(FScore, Reg(nReg))$
24: $int\lambda \leftarrow pos\lambda - (r\lambda u/2) : pos\lambda + (r\lambda u/2)$
25: $FScore_{mark} \leftarrow FScore(int\lambda)$
26: $FScore \leftarrow FScore - FScore_{mark}$
27: $nReg \leftarrow nReg + 1$
28: **if** $nReg > length(Reg(1,:))$ **then**
29: $nReg \leftarrow 0$
30: **end if**
31: $sumFScore \leftarrow sum(FScore)$
32: $Wv_i \leftarrow pos\lambda$
33: $i \leftarrow i + 1$
34: **end while**

D. Wavelength-based Hybrid Fisher Wrapper Parameterization

As discussed before, the $models$ parameter comprises the set of ML models evaluated during the WHFS execution. The system designer should make a previous selection of the ML models that might suit the application use case.

The $r\lambda$ parameter defines the size (nm) of the spectral region that will be removed in each iteration. Furthermore, this region determines the minimum distance between wavelengths

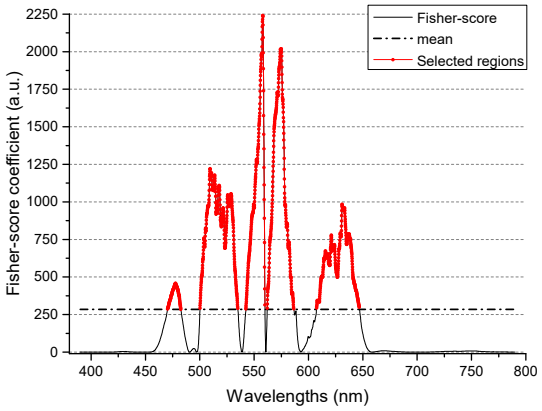


Fig. 2: Fisher-score relevance coefficient for each wavelength and its mean value.

provided in a solution. Therefore, using both criteria, the designer might establish its value according to the application's needs.

The p parameter establishes the maximum number of solutions generated for a particular ML model. As the wrapper stage in WHFS follows an accumulative strategy during the search, the p parameter impacts the number of wavelengths obtained. Hence, the first solution will have one wavelength; the second, two wavelengths, and so on, until reaching the p-th solution that will have p wavelengths. This fact influences the ML inference quality, the WHFS execution time, and the system acquisition cost. The designer must evaluate the application requirements to set a proper p value according to this trade-off.

The search process done in the wrapper stage in WHFS is shaped by the h parameter. This parameter has a more substantial impact on the inference quality and the WHFS execution time than p because it directly affects the search process. On the one hand, the WHFS algorithm with $h = 1$ would only choose the most relevant wavelength for each iteration. The same performance would be achieved by running a pure wavelength removal method. On the other hand, the WHFS algorithm with $h = n'$ (see n' definition in step 5 of Section III-C) would run a pure SFS algorithm that might improve the model performance metrics but increasing the WHFS execution time. This time is mainly affected by the total number of trained combinations, $C \in \mathbb{R}^1$, during the execution, which is given by

$$C = \sum_{j=1}^{p} \binom{n' - (j-1)}{1}, \tag{3}$$

where the j-th element of the sum contains the number of combinations for the j-th iteration. It is worth noting that C depends on the maximum number of iterations, p. This parameter directly affects the execution of the wrapper stage. Therefore, in high-dimensional applications, such as the use case selected in this paper, the wrapper stage might not converge when a model requires high training time. As a solution, the authors propose a general methodology, which might be used to set the h value to establish a maximum number of $\binom{h}{1}$ combinations per iteration.

The general methodology proposed uses a pure wavelength-based filter algorithm for all the available wavelengths, where WHFS is configured with $h = 1$ and $p = n'$. Then, this mode chooses $\binom{1}{1}$ combinations for n' iterations. As a result, the algorithm will provide n' solutions, which are analyzed by a threshold defined from an objective function value and the number of the solution wavelengths. From those solutions that surpass the threshold, the one with the lowest number of wavelengths will be selected, being h defined as this number of wavelengths. The objective value might be established as a value defined by the user that the most restrictive performance metric might achieve. The user can redefine this value, based on the experience, changing the objective function or its value. In Section V, the process of selecting the objective function and the value to obtain a proper h value will be detailed for the selected use case.

IV. Experimental Setup

This section first describes the ethanol detection database used for the evaluation of the proposed method. Furthermore, the regression metrics used to evaluate the proposed methodology are stated.

A. Dataset

The proposed framework is evaluated using the database generated in [10]. This database includes two datasets where the ethanol concentration is measured: the water-ethanol and wine-ethanol experiments. A single dataset consists of samples gathered with a spectrometer that registers the optical response of the RNPs in the visible waveband. Each experiment includes approximately 250-minutes spectrometer data, acquired at a sampling frequency of 1 Hz. The spectrometer has a resolution of $0.1nm$ ($sRes = 0.1$), getting 3648 wavelengths per sample.

In both experiments, pure ethanol (Sigma-Aldrich >99%) was added every 10 minutes by $1\%vol.$ to the base fluid. The goal of this 10-min measurement is to stabilize the mixture, reducing the outliers produced by the inertia of the process. Also, during the dataset creation, we identify and remove the transients from ethanol addition. The ethanol percentage is the dataset label annotated during the experimentation. For each experiment, a range exists for the property to be predicted. The water-ethanol experiment started at $1\%vol.$ of ethanol and ended at $25\%vol.$; thus, it had 25 regression labels. In turn, the wine-ethanol experiment started at $11\%vol.$ of ethanol and ended at $27\%vol.$; thus, it had 16 regression labels.

B. Machine Learning Performance Metrics

This section depicts the metrics used for the analysis of the ML system. These metrics are calculated during the training and testing stages of each model (line 9, Algorithm 1). The evaluation focus on establishing the ML performance in terms of predictability from a regression perspective:

- The coefficient of determination (R^2) analyses how a group of m observations is correlated to the model, i.e., the correlation between the label values and the model predictions. This metric is calculated as

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(V_i - \hat{V}_i)^2}{\sum_{i=1}^{m}(V_i - \overline{V})^2}, \tag{4}$$

  where, for the i-th observation, V_i is the label value, $\hat{V}_i$ is the prediction of the ML system, and $\overline{V}$ is the mean of the prediction values. The range of R^2 is $[0, 1]$. An R^2 value near 0 indicates no correlation between the data and the model, and an R^2 value near 1 indicates a strong correlation.
- Mean Absolute Error (MAE) is an estimator that focuses on the differences between the prediction and the label values. It is defined as the average of the absolute value of the errors (differences) and is calculated as

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|V_i - \hat{V}_i|. \tag{5}$$

  The range of the MAE is $[0, \infty)$, where the closer to 0, the better will be the model.
- Root-Square Mean Error (RMSE) is an estimator that also focuses on the differences between the prediction and the label values. Still, it highlights the influence of large errors. It is calculated as the square root of the quadratic mean of the errors, as given by:

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(V_i - \hat{V}_i)^2}. \tag{6}$$

  Similarly to MAE, the range of the RMSE is $[0, \infty)$. As the outliers and large errors have a higher impact on the RMSE metric than the one provided by MAE, the near to 0 the RMSE is, the lesser outliers and large errors have the model.

The evaluation of the models follows the k-fold cross-validation strategy to avoid overfitting. The k parameter determines the number of portions the dataset is split and the number of iterations the model is tested. The dataset is split randomly into k subsets of the same number of instances, approximately. Then, $(k-1)$ subsets are used for training, and the remaining subset is used for testing. This process is repeated k times, testing with a different subset in each iteration. The performance metrics of the model are the average of the performance metrics obtained in each cross-validation iteration. In this use case, k was set equal to 10 to obtain robust performance metrics [41], [42].

C. Target Platform

The edge node selected for the implementation is the *cookie* platform, a modular platform created at the Universidad Politécnica de Madrid [43]. It has a 4-layer hardware structure: power supply, sensing/actuation, processing, and communications. Each layer is a hardware board with a standard vertical connector that acts as the bridge between layers. Remarkably, the processing layer is based on an ultra-low-power, 32-bit medium performance ARM Cortex M4, featured with 256 MB external RAM.

Simplicity studio tool is considered to embed the solution into the platform. This tool integrates an energy profiler, which permits measuring the energy consumed by the solution. This fact allows studying the energy consumption of the system while running. Furthermore, this target uses C/C++ language for the applications. In particular, only the chosen solutions to be tested in the platform are programmed in this language to reduce development time.

V. Results and Discussion

In this section, the proposed methodology is applied to water-ethanol and wine-ethanol experiments. First, the WHFS parameters are tuned to the use case. Then the WHFS method is applied to generate the space of solutions. One solution for each experiment is then selected following a set of use case requirements. Finally, the chosen solutions are implemented in an edge computing node, showing the feasibility of the proposed methodology for real-time fluid characterization with photonic sensors.

A. WHFS Parameterisation for Ethanol Detection

In this subsection, the WHFS is parameterized for the ethanol detection use case following the methodology proposed in Section III-D. This parameterization is applied separately for water and wine experiments.

First, the $models$ parameter is set. The authors choose the following regression ML models: linear, interactions linear, robust linear, and stepwise linear. This selection is based on the proposal in [10]. Note that the Linear SVM regression model was discarded because the results shown in [10] were outperformed with simple models.

Second, the p parameter indicates the maximum number of solutions the WHFS method will provide and establishes the maximum number of wavelengths explored, as they are intrinsically related. In the ethanol detection use case, the objective is to provide a system with the minimum number of wavelengths to reduce the amount of information the acquisition system must gather in each sample because it affects the computational requirements imposed on the edge nodes. In this regard, a low p value is chosen, $p = 10$, which is an experimental threshold established below the 1 % of the initial dataset (3648 wavelengths). This p value is empirical, so it could be increased if the solutions obtained would not achieve an expected performance for the use case.

Third, the $r\lambda$ parameter reduces the information redundancy in each solution obtained. When one wavelength is picked, it removes the consecutive wavelengths. However, as seen in Fig. 2, it is possible to find areas with nearby local maxima by applying a Fisher-score relevance analysis, and then, the removal should be carefully performed. Note that a local maximum indicates that this wavelength might be a possible selected wavelength by the WHFS method due to its relevance. Experimentally, it was observed that the minimum distance between local maxima was $3nm$. Thus to avoid removing relevant wavelengths but reducing redundancy, $r\lambda/2$, which

TABLE I: h parameterization for each experiment and ML model by considering $p = 10$ (uds), $r\lambda = 5$ nm, and RMSE ($\%_{eth}$) as objective function (its objective values are defined as thresholds).

		h value of each ML regression model			
Experiment	**Objective value**	**linear**	**Interactions**	**Robust**	**Stepwise**
Water-ethanol	0.33	10	4	10	4
Wine-ethanol	0.25	5	4	5	4

represents the distance between the chosen wavelength and the last wavelength of one side, is fixed to $2.5nm$. Consequently, $r\lambda$ is fixed to $5nm$, relating to the distance to both sides of the wavelength selected.

Fourth, the combination of a particular experiment and ML model has a specific value of h. This value is automatically obtained by applying the methodology exposed in Section III-D, requiring an objective function and a threshold. The RMSE was selected as the objective function because it is the most restrictive metric for our use case. Note that the goal is to achieve a state-of-the-art prediction quality (i.e., state-of-the-art RMSE) using few wavelengths. The threshold is defined by getting at least an RMSE value equal to 75% of the state-of-the-art RMSE. Note that improving RMSE means bringing this metric close to 0. In this regard, the goal value is calculated as given by

$$GV = SV + 75\% * SV, \tag{7}$$

where GV is the RMSE goal value, and SV is the state-of-the-art RMSE value. In the water-ethanol experiment, state-of-the-art RMSE is equal to 0.19 $\%_{eth}$ [10]. Note that the notation $\%_{eth}$ means the percentage of ethanol in the fluid matrix. Thus the goal value is 0.33 $\%_{eth}$ in this first experiment. In the wine-ethanol use case, state-of-the-art RMSE is equal to 0.14 $\%_{eth}$ [10]. Therefore the goal value is 0.25 $\%_{eth}$ in the second experiment. According to these goal values, the h parameterization resulted in the h values shown for each ML model and experiment in Table I.

B. WHFS Execution for Ethanol Detection

Table II presents the space of solutions generated for the water-ethanol experiment by executing the WHFS strategy with the parameterization discussed in Section V-A. As established in Section III-B, it is worth noting that WHFS follows an accumulative approach, where the wavelengths of a particular solution are added to the next solution. Therefore, the i-th solution includes the current chosen wavelength (*Freq(nm)* field in the i-th row) and the wavelengths selected in the previous solutions in the same column, i.e., from 1 to $i - 1$. For instance, in the linear model, the second solution is composed of its selected wavelength (552.32 nm) and the wavelength of the first solution (569.43 nm). This table also shows the performance metrics (RMSE, R^2, and MAE), the number of wavelengths considered, which corresponds to the *Solution* field, and the training time required to create that solution (*Training time(s)* field). Analyzing this table, all the R^2 values are close to 1.00, meaning that all the models are suited to the water-ethanol experiment [44]. Also, in all

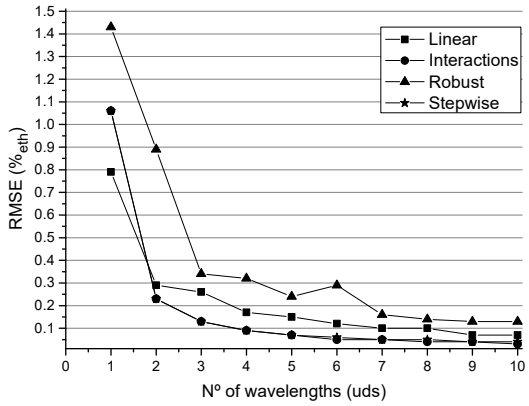


Fig. 3: Solution space exploration for the water-ethanol experiment, with $p = 10$.

solutions, MAE results are better than RMSE, which is less than 0.4 $\%_{eth}$. That means that, in this use case, the model is slightly affected by the outliers. In this regard, the exploration focuses on the RMSE metric as is the most restrictive one. Regarding the training time, it shows the complexity in terms of computational cost. Thus, the higher the training time, the higher the computational cost. This timing metric was obtained using an Intel i7-7700 processor with 16 GB of RAM, running in Matlab 2020b, with the Statistics and Machine Learning toolbox, over Windows 10 operating system. Note that the stepwise linear model increases two times the training time compared to interactions linear.

The evolution of the RMSE metric of each model for the water-ethanol experiment is shown in Fig. 3. It tries to be an aid to determine which solution is better. As designers, the authors opted for establishing two conditions in this regard. First, the solution will have a lower or equal value than the state-of-the-art, $0.19\%_{eth}$ in the water-ethanol experiment. Second, the model and solution with the least number of wavelengths will be selected to be implemented in an edge node. This latter criterion is due to the focus on minimizing the edge computing cost. Therefore, the third solution based on the interactions linear model was selected. This solution appears shaded in Table II.

In turn, the space of solutions for the wine-ethanol experiment is shown in Table III. Similar to the water-ethanol experiment, all the models and their solutions fit the data with $R^2 \geqslant 0.95$. The solutions are also slightly affected by the outliers, being RMSE the most restrictive metric. Besides, Fig. 4 exposes the RMSE value of each model solution for the experiment. The same conditions were followed to choose the solutions to be implemented in the edge, as in the water-ethanol experiment. In the wine experiment, the state-of-the-art RMSE value is $0.14\%_{eth}$. In this regard, the second solution of the linear model is chosen. This solution has the same RMSE value as the interactions linear and stepwise linear models. However, the interactions linear complexity is lesser than the stepwise linear model. The selected solution appears shaded in Table III.

TABLE II: Solution space for the water-ethanol experiment.

	Linear					Interactions Linear				
Solution	**Freq(nm)**	**RMSE ($\%_{eth}$)**	**R^2**	**MAE ($\%_{eth}$)**	**Training time (s)**	**Freq (nm)**	**RMSE ($\%_{eth}$)**	**R^2**	**MAE ($\%_{eth}$)**	**Training time (s)**
1	569.43	0.79	0.99	0.60	1.45	575.02	1.06	0.98	0.71	0.84
2	552.32	0.29	1.00	0.22	3.18	630.96	0.23	1.00	0.18	1.76
3	558.02	0.26	1.00	0.19	4.91	558.02	0.13	1.00	0.10	2.78
4	637.54	0.17	1.00	0.14	6.70	509.55	0.09	1.00	0.07	3.95
5	472.14	0.15	1.00	0.12	8.77	552.32	0.07	1.00	0.05	5.30
6	580.62	0.12	1.00	0.10	10.82	637.54	0.05	1.00	0.04	6.85
7	546.73	0.10	1.00	0.08	12.63	472.14	0.05	1.00	0.03	8.64
8	477.74	0.10	1.00	0.08	14.48	517.77	0.04	1.00	0.03	10.7
9	563.84	0.07	1.00	0.05	16.39	546.73	0.04	1.00	0.03	13.1
10	575.02	0.07	1.00	0.05	18.39	580.62	0.03	1.00	0.02	15.8
	Robust Linear					**Stepwise Linear**				
Solution	**Freq(nm)**	**RMSE ($\%_{eth}$)**	**R^2**	**MAE ($\%_{eth}$)**	**Training time (s)**	**Freq (nm)**	**RMSE ($\%_{eth}$)**	**R^2**	**MAE ($\%_{eth}$)**	**Training time (s)**
1	630.96	1.43	1.00	1.15	2.76	575.02	1.06	0.98	0.71	1.09
2	558.02	0.89	1.00	0.73	5.57	630.96	0.23	1.00	0.18	2.42
3	575.02	0.34	1.00	0.25	8.75	558.02	0.13	1.00	0.10	6.78
4	472.14	0.32	1.00	0.23	12.5	509.55	0.09	1.00	0.08	19.70
5	552.32	0.24	1.00	0.16	16.1	552.32	0.07	1.00	0.06	50.31
6	509.55	0.29	1.00	0.17	20.0	637.54	0.06	1.00	0.05	114.36
7	528.85	0.16	1.00	0.14	23.5	472.14	0.05	1.00	0.04	231.43
8	477.74	0.14	1.00	0.12	26.5	477.74	0.05	1.00	0.04	419.27
9	621.09	0.13	1.00	0.11	29.8	580.62	0.04	1.00	0.04	739.57
10	569.43	0.13	1.00	0.11	33.7	569.43	0.04	1.00	0.03	1133.87

TABLE III: Solution space for the wine-ethanol experiment.

	Linear					Interactions Linear				
Solution	**Freq(nm)**	**RMSE ($\%_{eth}$)**	**R^2**	**MAE ($\%_{eth}$)**	**Training time (s)**	**Freq (nm)**	**RMSE ($\%_{eth}$)**	**R^2**	**MAE ($\%_{eth}$)**	**Training time (s)**
1	582.70	0.48	0.99	0.37	0.70	582.70	0.48	0.99	0.37	0.80
2	632.39	0.12	1.00	0.09	1.48	632.39	0.12	1.00	0.09	1.69
3	477.08	0.11	1.00	0.09	2.35	561.86	0.10	1.00	0.08	2.77
4	520.40	0.10	1.00	0.08	3.12	550.13	0.06	1.00	0.05	3.89
5	550.13	0.08	1.00	0.07	4.04	637.98	0.05	1.00	0.04	5.10
6	572.39	0.07	1.00	0.05	5.01	572.39	0.03	1.00	0.03	6.56
7	637.98	0.05	1.00	0.04	5.98	520.40	0.03	1.00	0.03	8.31
8	555.72	0.04	1.00	0.03	6.88	555.72	0.03	1.00	0.02	10.2
9	532.36	0.04	1.00	0.03	7.75	512.29	0.02	1.00	0.02	12.3
10	561.86	0.04	1.00	0.03	8.67	532.36	0.02	1.00	0.02	14.6
	Robust Linear					**Stepwise Linear**				
Solution	**Freq(nm)**	**RMSE ($\%_{eth}$)**	**R^2**	**MAE ($\%_{eth}$)**	**Training time (s)**	**Freq (nm)**	**RMSE ($\%_{eth}$)**	**R^2**	**MAE ($\%_{eth}$)**	**Training time (s)**
1	582.70	0.48	1.00	0.37	1.40	582.70	0.48	0.99	0.37	1.91
2	561.86	0.17	1.00	0.14	3.02	632.39	0.12	1.00	0.09	4.19
3	632.39	0.12	1.00	0.08	4.99	561.86	0.10	1.00	0.08	10.94
4	512.29	0.11	1.00	0.08	7.08	550.13	0.07	1.00	0.05	28.40
5	572.39	0.08	1.00	0.06	9.55	520.40	0.06	1.00	0.05	63.72
6	637.98	0.06	1.00	0.04	11.7	572.39	0.05	1.00	0.03	124.91
7	477.08	0.06	1.00	0.04	14.1	637.98	0.04	1.00	0.03	217.58
8	550.13	0.05	1.00	0.04	16.5	555.72	0.04	1.00	0.03	328.99
9	520.40	0.05	1.00	0.04	19.1	477.08	0.04	1.00	0.03	532.86
10	482.67	0.05	1.00	0.04	21.9	482.67	0.03	1.00	0.03	790.99

The suitability of the WHFS method compared to pure filter and wrapper methods is also analyzed in this section to justify the hybridization. Thus, the methodology is applied using a pure filter based on Fisher-score and a pure wrapper based on SFS for both experiments. As one of the drawbacks of pure wrapper methods is their convergence, a maximum convergence time of a week for each solution (tcmax =10080 min) is established, running the algorithm on the same desktop machine considered for the rest of the experiments. Under these circumstances, no solution has been obtained using this method as it surpasses in each case the timeout. On the contrary, applying a pure Fisher-score method, all the p solutions were obtained. In the water-ethanol, only the interactions linear model reaches the performance of the state-the-art with a minimum number of wavelengths equal to 7. Nonetheless, the 7-wavelengths interaction linear model has an RMSE value of 0.18 %eth; therefore, it does not outperform our proposal performance in RMSE value nor the number of wavelengths. In the wine-ethanol experiment, the best RMSE value was 0.14 %eth using 10-wavelengths interactions linear model, achieving the state-of-the-art performance, but not improving the proposal provided by the WHFS algorithm. Besides, in both experiments, the use of a pure Fisher-score method resulted in solutions with high redundancy, where the maximum distance between wavelengths was approximately 1 nm. Therefore, the WHFS generates a more suitable solution in terms of ML performance and redundancy for the application.

The methodology followed in this section permits a scalability analysis showing the impact of the number of wavelengths and the ML model for identifying the compound percentage in a fluid media. For both experiments, the goal was selecting the solutions providing at least a given RMSE, whereas the

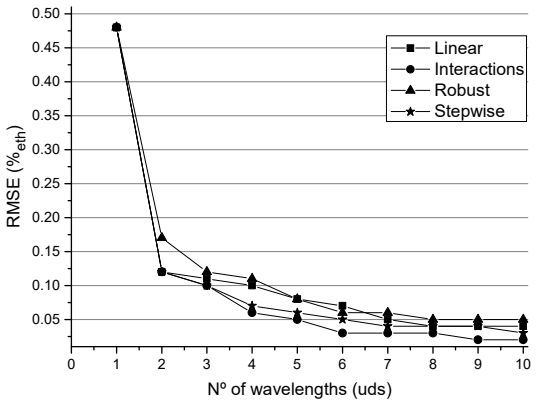


Fig. 4: Solution space exploration for the wine-ethanol experiment, with $p = 10$.

number of wavelengths is minimized. Nevertheless, these conditions can be modified to achieve the requirements of other use cases. Hence, this framework is generalizable to different applications.

Finally, the two selected solutions are represented in Figs. 5a and 5b for the water-ethanol and wine-ethanol experiments, respectively. In both figures, a spectrum acquired during the experimentation is shown. Note that a spectrum consists of the RNPs reflectivity measurements for the 3648 wavelengths. Moreover, the wavelengths selected by the framework proposed are marked using black squares. These figures permit to have a quality perspective of the dimensionality reduction and the position of the wavelengths selected. Thus, comparing the wavelengths selected for both experiments, two wavelengths come from the same regions, one close to $580nm$ and the other close to $630nm$, showing a correlation between both media.

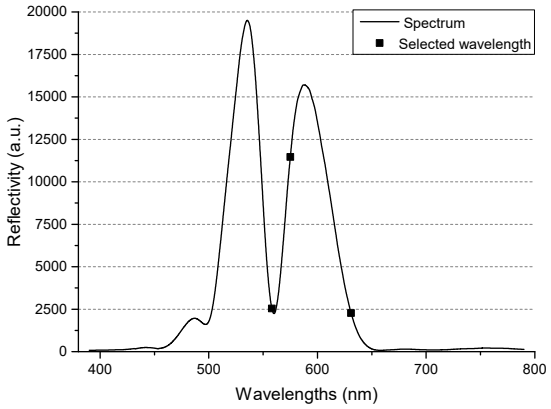
C. Evaluating WHFS solutions at the Edge

After applying the WHFS design exploration method, the selected solutions surpassed the state-of-the-art results in terms of performance metrics, as reported in the previous section. Besides, the methodology enables choosing those solutions that focus on alleviating the ML system computational load, e.g., solutions with fewer wavelengths. This subsection analyzes the performance of the state-of-the-art and the proposed solutions during inference for each experiment in a resource-constraint IoT node. The solutions selected in the previous section appear summarized in Table IV. The state-of-the-art implementations selected for comparative purposes are based on the solutions described in [10] for the same dataset. These state-of-the-art implementations consider the whole spectrum (the 3648 wavelengths) reduced by PCA to several PCs, which are used to infer the property using a regression model. Among the different configurations provided in this previous work, the authors selected two solutions for comparative purposes, one for each experiment, according to the following criteria: i) the solution might surpass the state-of-the-art metrics in terms of ML performance and ii) only the solution with the lowest number of wavelengths and the lightest ML regression model must be chosen. Note that any of the ML models evaluated could be used because there are configurations in which they outperform the state-of-the-art paper. However, it will incur in increasing the number of wavelengths required (see Tables II and III).

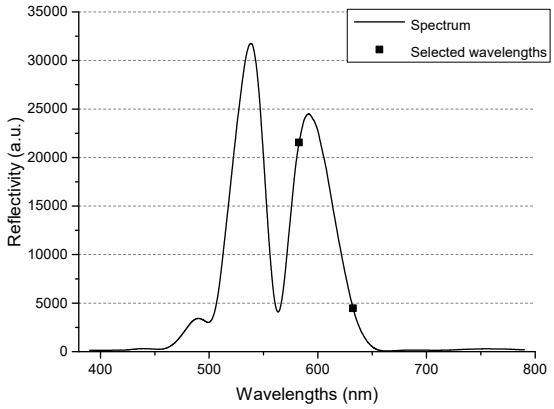
Table IV also presents the inference execution time and energy consumption measured in the *cookie* platform for each solution tested. Comparing the proposed solutions for both experiments, the differences between execution time and energy consumption are 6 ms and 0.3 mJ, respectively. Analyzing the water-ethanol experiment while using the same regression model for the proposed and the state-of-the-art solutions, the execution time and energy consumption are decreased up to 43.59 times the processing time and 43.90 times the energy consumption. As both solutions consider the same ML model fed by the same number of variables (three variables in this experiment), the differences observed lie on that the state-of-the-art solution uses the PCA technique, which compresses the whole set of wavelengths into three variables. The PCA algorithm used in the compression consumes 97.77% of energy from the entire set of calculations necessary to make the prediction. The number of wavelengths is reduced by 1216 times, also lowering the energy consumption in the acquisition system. Focusing on the wine-ethanol experiment while using the same regression model for the proposed and the state-of-the-art solutions, the execution time and energy consumption are decreased to 67.18 and 68.57 times, respectively. As in the previous experiment, the PCA technique consumes 98.56% of energy to make the prediction. The number of wavelengths is decreased to 1824. As a result, it can be stated that the proposed solutions outperform the state-of-the-art ones in terms of resources, execution time, and energy consumption while maintaining the same prediction performance. Comparing both proposed solutions (2-wavelengths linear model and 3-wavelengths interactions linear model), it can be observed that the energy consumption raises in 0.04 mJ. In the water-ethanol experiment, the 2-wavelengths linear model proposal decreases the ethanol detection performance, up to 0.29 %eth of RMSE, not surpassing the state-of-the-art. On the contrary, in the wine-ethanol experiment, the performance is improved, obtaining an RMSE of 0.10 %eth, however, the improvement is not higher enough (0.02 %eth) to assume an increment of energy consumption. Therefore, the system to be implemented remains invariant.

VI. Conclusion

This work proposes a novel spectral methodology for miniaturizing ML systems for fluid characterization based on photonic sensors, particularly RNPs transducers. It presents an alternative to traditional techniques, which demand high computational costs and resources. This methodology comprises the use of a hybrid FS procedure-oriented to the wavelength domain, providing a method for wavelength selection in photonic transducers. The WHFS method aims at reducing the number of wavelengths that an ML model needs while maintaining a high prediction quality. Its configuration permits



(a) Wavelengths selected for water-ethanol experiment.



(b) Wavelengths selected for wine-ethanol experiment.

Fig. 5: Resulting wavelengths selected for both experiments.

TABLE IV: Comparison between the proposed and state-of-the-art solution for each experiment.

Experiment	ML Model	Wavelengths (uds)	Wavelengths (nm)	PCs (uds)	RMSE ($\%_{eth}$)	R^2	MAE ($\%_{eth}$)	Time (ms)	Energy (mJ)	Reference
Water-ethanol	Interactions	3	575.02, 630.96, 558.02	-	0.13	1.00	0.10	17	0.11	This paper
Water-ethanol	Interactions	3648 (All)	-	2	0.19	1.00	0.14	741	4.83	[10]
Wine-ethanol	Linear	2	582.70, 632.39	-	0.12	1.00	0.09	11	0.07	This paper
Wine-ethanol	Linear	3648 (All)	-	2	0.14	1.00	0.11	739	4.80	[10]

the adaptation of the technique to the use case requirements by modifying the restrictions for the wavelength selection. As a result, the authors verified that i) the solutions proposed by the methodology can achieve state-of-the-art prediction quality results, and ii) they fit better in a resource-aware platform for the edge layer due to its energy consumption, inference execution time, and resource utilization. The proposed solutions outperform the state-of-the-art up to 67.18 times the inference execution time and 68.57 times the energy consumption. From the sensor design point of view, the proposal can also help to simplify the sensor readout equipment for specific applications. For instance, in the considered case of study, the analysis performed suggests that the light source and spectrometer can be replaced by several lasers and photodetectors, which can be cheaper and more precise in determining the intensity of the sensor signal in the selected ranges.

Future work is oriented to apply these solutions in different edge nodes from the hardware perspective, such as FPGA-based nodes, whose parallelization capabilities might allow integration prognosis inside this IoT layer. From a chemical industry standpoint, the next objective is to analyze the performance of this methodology in different media and applications, such as the determination of other chemical compounds of interest in the industry in liquid or gas states. Also, in the smart agriculture domain, this WHFS method might be applied to choose the relevant wavebands among those produced by a hyperspectral camera to reduce the computational needs.

REFERENCES

[1] D. Wu, S. Liu, L. Zhang, J. Terpenny, R. X. Gao, T. Kurfess, and J. A. Guzzo, "A fog computing-based framework for process monitoring and prognosis in cyber-manufacturing," Journal of Manufacturing Systems, vol. 43, pp. 25–34, 2017.

[2] S. Ahmad, A. Badwelan, A. M. Ghaleb, A. Qamhan, M. Sharaf, M. Alatefi, and A. Moohialdin, "Analyzing critical failures in a production process: Is industrial iot the solution?," Wireless Communications and Mobile Computing, vol. 2018, 2018.

[3] M. Wollschlaeger, T. Sauter, and J. Jasperneite, "The future of industrial communication: Automation networks in the era of the internet of things and industry 4.0," IEEE industrial electronics magazine, vol. 11, no. 1, pp. 17–27, 2017.

[4] H. Rahmania, A. Rohman, et al., "The employment of ftir spectroscopy in combination with chemometrics for analysis of rat meat in meatball formulation," Meat science, vol. 100, pp. 301–305, 2015.

[5] M. Khorasani, J. M. Amigo, P. Bertelsen, C. C. Sun, and J. Rantanen, "Process optimization of dry granulation based tableting line: Extracting physical material characteristics from granules, ribbons and tablets using near-ir (nir) spectroscopic measurement," Powder Technology, vol. 300, pp. 120–125, 2016.

[6] L. Tu, L. Huang, and W. Wang, "A novel micromachined fabry-perot interferometer integrating nano-holes and dielectrophoresis for enhanced biochemical sensing," Biosensors and Bioelectronics, vol. 127, pp. 19–24, 2019.

[7] Á. Lavín, R. Casquel, F. J. Sanza, M. F. Laguna, and M. Holgado, "Efficient design and optimization of bio-photonic sensing cells (bicells) for label free biosensing," Sensors and Actuators B: chemical, vol. 176, pp. 753–760, 2013.

[8] A. Mohebzadeh Bahabady, S. Olyaee, and H. Arman, "Optical biochemical sensor using photonic crystal nano-ring resonators for the detection of protein concentration," Current Nanoscience, vol. 13, no. 4, pp. 421–425, 2017.

[9] C. Chen, X. Hou, and J. Si, "Protein analysis by mach-zehnder interferometers with a hybrid plasmonic waveguide with nano-slots," Optics Express, vol. 25, no. 25, pp. 31294–31308, 2017.

[10] S. Quintero, R. Marino, J. M. Lanza-Gutierrez, F. J. Sanza, T. Riesgo, and M. Holgado, "A novel data processing technique for expert resonant nano-pillars transducers: A case study measuring ethanol in water and wine liquid matrices," IEEE Access, vol. 7, pp. 129778–129788, 2019.

[11] J. Miao and L. Niu, "A survey on feature selection," Procedia Computer Science, vol. 91, pp. 919–926, 2016.

[12] Y. Lu, A. Memon, P. Fuerst, A. Kizonas, C. Morris, and D. Luthria,

"Changes in the phenolic acids composition during pancake preparation: Whole and refined grain flour and processed food classification by uv and nir spectral fingerprinting method—proof of concept," Journal of Food Composition and Analysis, vol. 60, pp. 10–16, 2017.

[13] L. Wimonsiri, P. Ritthiruangdej, S. Kasemsumran, N. Therdthai, W. Chanput, and Y. Ozaki, "Rapid analysis of chemical composition in intact and milled rice cookies using near infrared spectroscopy," Journal of Near Infrared Spectroscopy, vol. 25, no. 5, pp. 330–337, 2017.

[14] D. Lorente, P. Escandell-Montero, S. Cubero, J. Gómez-Sanchis, and J. Blasco, "Visible–nir reflectance spectroscopy and manifold learning methods applied to the detection of fungal infections on citrus fruit," Journal of Food Engineering, vol. 163, pp. 17–24, 2015.

[15] S. H. F. Scafi and C. Pasquini, "Identification of counterfeit drugs using near-infrared spectroscopy," Analyst, vol. 126, no. 12, pp. 2218–2224, 2001.

[16] R. Espinosa, M. Garrido-Arandia, A. Romero-Sahagun, P. Herreros, L. Tramarin, M. Laguna, A. Díaz-Perales, and M. Holgado, "A new optical interferometric-based in vitro detection system for the specific ige detection in serum of the main peach allergen," Biosensors and Bioelectronics, vol. 169, 2020.

[17] R. Casquel, M. Holgado, M. F. Laguna, A. L. Hernández, B. Santamaría, Á. Lavín, L. Tramarin, and P. Herreros, "Engineering vertically interrogated interferometric sensors for optical label-free biosensing," Analytical and bioanalytical chemistry, vol. 412, no. 14, pp. 3285–3297, 2020.

[18] Z. Díaz-Betancor, M.-J. Bañuls, F. J. Sanza, R. Casquel, M. F. Laguna, M. Holgado, R. Puchades, and Á. Maquieira, "Phosphorylcholine-based hydrogel for immobilization of biomolecules. application to fluorometric microarrays for use in hybridization assays and immunoassays, and nanophotonic biosensing," Microchimica Acta, vol. 186, no. 8, pp. 1–11, 2019.

[19] A. L. Hernandez, F. Dortu, T. Veenstra, P. Ciaurriz, R. Casquel, I. Cornago, H. V. Horsten, E. Tellechea, M. V. Maigler, F. Fernández, et al., "Automated chemical sensing unit integration for parallel optical interrogation," Sensors, vol. 19, no. 4, p. 878, 2019.

[20] S. Quintero, R. Casquel, M. F. Laguna, and M. Holgado, "Optical vapor sensors based on periodic resonant nanopillar structures," ACS omega, vol. 5, no. 40, pp. 25913–25918, 2020.

[21] J. Gil-Rostra, S. Quintero-Moreno, V. J. Rico, F. Yubero, F. J. Sanza, R. Casquel, E. Gallo-Valverde, M. E. Jara-Galán, P. Sanz-Sanz, M. Holgado, and A. R. González-Elipe, "Photonic sensor systems for the identification of hydrocarbons and crude oils in static and flow conditions," Sensors and Actuators B: Chemical, vol. 344, p. 130265, 2021.

[22] Y.-H. Yun, H.-D. Li, B.-C. Deng, and D.-S. Cao, "An overview of variable selection methods in multivariate analysis of near-infrared spectra," TrAC Trends in Analytical Chemistry, 2019.

[23] D. Ballabio, V. Consonni, A. Mauri, M. Claeys-Bruno, M. Sergent, and R. Todeschini, "A novel variable reduction method adapted from space-filling designs," Chemometrics and Intelligent Laboratory Systems, vol. 136, pp. 147–154, 2014.

[24] X. Zeng, S.-B. Yin, Y. Guo, J.-R. Lin, and J.-G. Zhu, "A novel semi-supervised feature extraction method and its application in automotive assembly fault diagnosis based on vision sensor data," Sensors, vol. 18, no. 8, p. 2545, 2018.

[25] R. Zhang, F. Nie, X. Li, and X. Wei, "Feature selection with multi-view data: A survey," Information Fusion, vol. 50, pp. 158–167, 2019.

[26] P. Peets, I. Leito, J. Pelt, and S. Vahur, "Identification and classification of textile fibres using atr-ft-ir spectroscopy with chemometric methods," Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, vol. 173, pp. 175–181, 2017.

[27] J. Li, K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang, and H. Liu, "Feature selection: A data perspective," ACM Computing Surveys (CSUR), vol. 50, no. 6, pp. 1–45, 2017.

[28] C. Huffman, H. Sobral, and E. Terán-Hinojosa, "Laser-induced breakdown spectroscopy spectral feature selection to enhance classification capabilities: A t-test filter approach," Spectrochimica Acta Part B: Atomic Spectroscopy, vol. 162, p. 105721, 2019.

[29] H. Chen, C. Tan, Z. Lin, and T. Wu, "Classification and quantitation of milk powder by near-infrared spectroscopy and mutual information-based variable selection and partial least squares," Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, vol. 189, pp. 183–189, 2018.

[30] C. Yan, J. Liang, M. Zhao, X. Zhang, T. Zhang, and H. Li, "A novel hybrid feature selection strategy in quantitative analysis of laser-induced breakdown spectroscopy," Analytica chimica acta, vol. 1080, pp. 35–42, 2019.

[31] S. Lu, S. Shen, J. Huang, M. Dong, J. Lu, and W. Li, "Feature selection of laser-induced breakdown spectroscopy data for steel aging estimation," Spectrochimica Acta Part B: Atomic Spectroscopy, vol. 150, pp. 49–58, 2018.

[32] J. Acquarelli, T. van Laarhoven, J. Gerretzen, T. N. Tran, L. M. Buydens, and E. Marchiori, "Convolutional neural networks for vibrational spectroscopic data analysis," Analytica chimica acta, vol. 954, pp. 22–31, 2017.

[33] S. Abbasi, S. Gharaghani, A. Benvidi, and A. Latif, "Identifying the novel natural antioxidants by coupling different feature selection methods with nonlinear regressions and gas chromatography-mass spectroscopy," Microchemical Journal, vol. 139, pp. 372–379, 2018.

[34] S. Xu, Y. Zhao, M. Wang, and X. Shi, "Determination of rice root density from vis–nir spectroscopy by support vector machine regression and spectral variable selection techniques," Catena, vol. 157, pp. 12–23, 2017.

[35] R. Marino, C. Wisultschew, A. Otero, J. M. Lanza-Gutierrez, J. Portilla, and E. de la Torre, "A machine learning-based distributed system for fault diagnosis with scalable detection quality in industrial iot," IEEE Internet of Things Journal, 2020.

[36] L. Sun, T. Wang, W. Ding, J. Xu, and Y. Lin, "Feature selection using fisher score and multilabel neighborhood rough sets for multilabel classification," Information Sciences, 2021.

[37] S. Taran and V. Bajaj, "Sleep apnea detection using artificial bee colony optimize hermite basis functions for eeg signals," IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 2, pp. 608–616, 2019.

[38] L. Fu, T. Zhu, K. Zhu, and Y. Yang, "Condition monitoring for the roller bearings of wind turbines under variable working conditions based on the fisher score and permutation entropy," Energies, vol. 12, no. 16, p. 3085, 2019.

[39] A. Ahmad, S. Akbar, M. Hayat, F. Ali, S. khan, and M. Sohail, "Identification of antioxidant proteins using a discriminative intelligent model of k-space amino acid pairs based descriptors incorporating with ensemble feature selection," Biocybernetics and Biomedical Engineering, 2020.

[40] I. Baek, M. S. Kim, B.-K. Cho, C. Mo, J. Y. Barnaby, A. M. McClung, and M. Oh, "Selection of optimal hyperspectral wavebands for detection of discolored, diseased rice seeds," Applied Sciences, vol. 9, no. 5, p. 1027, 2019.

[41] J. Han, J. Pei, and M. Kamber, Data mining: concepts and techniques. Elsevier, 2011.

[42] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2016.

[43] P. Merino, G. Mujica, J. Señor, and J. Portilla, "A modular iot hardware platform for distributed and secured extreme edge computing," Electronics, vol. 9, no. 3, p. 538, 2020.

[44] G. Vishwakarma, A. Sonpal, and J. Hachmann, "Metrics for benchmarking and uncertainty quantification: Quality, applicability, and a path to best practices for machine learning in chemistry," arXiv preprint arXiv:2010.00110, 2020.

Rodrigo Marino received the BSc degree in Industrial Electronics Engineering and Automation from Universidade de Vigo, Spain, in 2015, and the MSc in Industrial Engineering in 2017 from the same University. He is currently working toward a Ph.D. degree in Industrial Electronics at Universidad Politécnica de Madrid. His current research area is in the field of machine learning applied to embedded systems, also known as expert embedded computing.

He is participating in a national research project, PLATINO, related to enhancing acquisition systems, combining machine learning techniques with embedded systems, to develop expert sensors for the agro-food industry. He has also participated in an industrial project, REMO, with Indra and Repsol companies, to create a framework for chemical detection.

Sergio Quintero received the bachelor's degree in electronics and automation engineering from Miguel Hernández University, in 2015, and the master's degree in biomedical engineering from the Universidad Politécnica de Madrid, in 2016. He is currently pursuing a Ph.D. degree with the Group of Optics, Photonics, and Biophotonics, Center for Biomedical Technology (CTB-UPM), Universidad Politécnica de Madrid. His research interests include the development of an optical sensor for chemical, agro-food, and medical applications. He has also participated in an industrial project, REMO, with Indra and Repsol companies, so as to create a framework for chemical detection.

Andrés Otero received his M.Sc. degree in Telecommunication Engineering from the University of Vigo, where he graduated with honors in 2007. He received his Master of Research and Ph.D. degrees in Industrial Electronics from Universidad Politécnica de Madrid (UPM), in 2009 and 2014, respectively. He is currently an Assistant Professor of electronics with the UPM, as well as a researcher in the Centro de Electrónica Industrial (CEI). His current research interests are focused on Embedded System Design, Reconfigurable Systems on FPGAs, Evolvable Hardware, and Embedded Machine Learning. During the last years, he has been involved in different research projects in these areas, and he is the author of more than 30 papers published in international conferences and journals. He has served as the Program Committee member of different international conferences in the field of reconfigurable systems, such as SPL, ERSA, ReConFig, DASIP, and ReCoSoC.

Jose M. Lanza-Gutierrez received the B.S and M.S degrees in Computer Science from the University of Extremadura, Caceres, Spain, in 2008 and 2009 respectively. In 2010, he obtained a master's degree in grid computing and parallelism at the same University. In 2015, he received a Ph.D. degree in Computer Science from the University of Extremadura under the guidance of Prof. Dr. Juan A. Gomez-Pulido. Currently, he is an assistant professor at the University of Alcalá, Spain. He has authored or co-authored more than 50 publications, including Journal Citation Report (JCR) papers in journals, such as IEEE Access, IEEE Internet of Things, Applied Soft Computing, Expert Systems with Application, BMC Bioinformatics, Soft Computing, and Reliability Engineering and System Safety. His main research interests include metaheuristics, digital embedded systems, cognitive computing, machine learning, and the Internet of Things.

Miguel Holgado received the degree in industrial engineering from the Technical University of Madrid (UPM), Madrid, Spain, in 1996, and the Ph.D. degree from the Technical University of Madrid and Institute of Material Science, Spanish National Research Council, Madrid, in 2000. He is currently the Deputy Vice-Rector of innovation and the Group Leader of the Optics, Photonics, and Biophotonics Group, Center for Biomedical Technology, UPM, and an Associate Professor with the Applied Physics and Material Engineering Department, Industrial Engineering School, UPM. He has led and participated in 34 research projects: 9 European, 19 National and regional, and other industrial and RD initiatives. He has authored or coauthored more than 150 scientific contributions, which have been cited more than 2200 times, and is the inventor of six patents applications. In addition, he is also the Founder of Bio Optical Detection; a spin-off company (BIOD S.L.) that develops optical Point-of-Care devices and offers IVD screening services