IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received February 4, 2020, accepted February 13, 2020, date of publication February 17, 2020, date of current version March 25, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2974623

A Machine Learning Approach for Lamb Meat Quality Assessment Using FTIR Spectra

ROCÍO ALAIZ-RODRÍGUEZ[1] AND ANDREW C. PARNELL[2]
[1]Department of Electrical, Systems and Automation Engineering, Universidad de León, 24071 León, Spain
[2]Hamilton Institute, Maynooth University, Maynooth, W23 F2H6 Ireland

Corresponding author: Rocío Alaiz-Rodríguez (rocio.alaiz@unileon.es)

The work of Andrew C. Parnell was supported in part by the Science Foundation Ireland Career Development Award under Grant 17/CDA/4695 and in part by the SFI Centre under Grant 12/RC/2289_P2.

ABSTRACT The food industry requires automatic methods to establish authenticity of food products. In this work, we address the problem of the certification of suckling lamb meat with respect to the rearing system. We evaluate the performance of neural network classifiers as well as different dimensionality reduction techniques, with the aim of categorizing lamb fat by means of spectroscopy and analyzing the features with more discrimination power. Assessing the stability of feature ranking algorithms also becomes particularly important. We assess six feature selection techniques (χ^2, Information Gain, Gain Ratio, Relief and two embedded techniques based on the decision rule 1R and SVM (Support Vector Machine). Additionally, we compare them with common approaches in the chemometrics field like the Partial Least Square (PLS) model and Principal Component Analysis (PCA) regression. Experimental results with a fat sample dataset collected from carcasses of suckling lambs show that performing feature selection contributes to classification performance increasing accuracy from 89.70% with the full feature set to 91.80% and 93.89% with the SVM approach and PCA, respectively. Moreover, the neural classifiers yield a significant increase in the accuracy with respect to the PLS model (85.60% accuracy). It is noteworthy that unlike PCA or PLS, the feature selection techniques that select relevant wavelengths allow the user to identify the regions in the spectrum with the most discriminant power, which makes the understanding of this process easier for veterinary experts. The robustness of the feature selection methods is assessed via a visual approach.

INDEX TERMS Meat quality assessment, feature selection, machine learning, neural networks, feature selection robustness.

I. INTRODUCTION

Sucking lamb meat, a characteristic product of the European Mediterranean area, is highly appreciated for its gastronomic quality on the basis of its tenderness and juiciness [1], [2]. For that reason, this meat is very common in the markets and has been protected by several Protected Geographical Indication (PGI) European Union's quality labels.

It has been found that the type of rearing has influence on meat quality [3]–[6]. In Mediterranean Europe, due to common incorporation of vegetable fat in milk replacers, considerable differences have been found between fat composition of lambs fed with ewe milk and commercial milk replacers [3], [7]. Therefore, meat quality can be adversely modified with the use of milk replacers, mainly by means of decreasing fatty acid dietetic value.

The development of new and sophisticated techniques to establish authenticity of food products is becoming more and more necessary because of the growing consumer concern about quality and food safety. Likewise, some types of food are subject to strict regulations such as quality labels. In the case of quality labels, the production has to be controlled in a precise and systematic manner.

There is a need to develop control methods to categorize suckling lamb meat according to the rearing system. The discrimination between animals reared with maternal milk and those fed with artificial milk replacers is crucial for the fulfilment of some quality label regulations. One of these Spanish quality labels is ''Lechazo de Castilla y León'' Protected Geographical Indication (PGI) (Council Regulation 2081/92/EC) [7], [8]. This can be done by analytical methods

The associate editor coordinating the review of this manuscript and approving it for publication was Shagufta Henna.

in a laboratory, but they need to be carried out manually by experts, thus being very expensive and slow.

The development of fast, automatic and efficient tools to be implemented in traceability programs is very important for the food industry. It has attracted great interest in the last decade (a comprehensive revision in [9]–[12] and references therein).

Fourier-transform infrared (FTIR) spectroscopy has become a powerful tool and attractive alternative to the existing analytical techniques for food quality assessment [13]–[17]. FTIR spectroscopy is considered as an ideal technique for fast screening of meat due to its unique spectral fingerprint [18], since, in theory, there are no two meats having the same FTIR spectra.

Authenticity control by spectroscopic techniques has several advantages over other analytical laboratory methods, namely increased sensitivity, speed, resolution, versatility and ease of use. In the case of suckling lambs, if the information contained in the FTIR spectra could be used to discriminate suckling lamb meat originating from different nutritional regimes, this tool could be implemented in the food industry.

The FTIR spectra are characterized by thousands of features which presents particular challenges since the analysis has to be conducted in a high-dimensional feature space. Many of these features are likely to be irrelevant or redundant and many others inevitably affected by noise. This becomes a common problem to all classification models [19]–[23]. Moreover, the number of instances tend to be small due to the cost of data collection, which worsens this scenario. Therefore, a dimensionality reduction step seems crucial prior to any attempt to build any classification model. Reducing the data dimensionality can avoid over-fitting and improve model performance [20]. Additionally, we can gain a deeper insight into the problem when the unwanted noisy and irrelevant features are withdrawn.

The aim of this work is: (i) to develop a classification model of fat samples according to the rearing system based on the FTIR spectra and (ii) to provide a deeper insight to veterinarian experts about the wavelengths that provide more information for the discrimination of fatty tissues in suckling lamb meat.

We assess a neural classifier with MLP (Multi-Layer Perceptron) architecture able to model non linear input-ouput relationships. Additionally, a comparative study of different feature selection techniques with regard to both their stability and their effect on classification performance is carried out. Our neural classifier is compared with the PLS model as well as neural classification model with PCA features extracted from the original set of features.

The rest of the paper is organized as follows: In Section II we formulate the problem. In Section III we present the proposed methodology and Section IV includes the empirical study of our proposal. We perform a comparison with other common approaches in Section V. Finally, Section VI summarizes our main conclusions and future work.

II. BACKGROUND

Machine learning techniques have been used for the assessment of lamb meat. Some previous works aim at the prediction of lamb meat tenderness using either hyper-spectral imaging [24] or instrumental and sensory measurements [25]. Thus, an SVM is combined in [25] with a feature selection procedure based on sensitivity analysis for the assessment of lamb meat tenderness. There are also some proposals to tackle the problem of categorizing lamb muscle types [26]–[28] applying machine learning techniques like artificial neural networks, SVM, logistic regression or linear discriminant analysis. The problem of estimating intramuscular fat content in lamb muscle based on ultrasound images and using artificial neural networks has also been tackled [29]. Other studies have been carried out for predicting lamb meat adulteration using multiple linear regression [30].

The application of FTIR spectroscopy techniques offers both qualitative and quantitative analysis of a wide range of materials. The spectrum information content is very specific permitting fine discrimination between similar materials [31]. Typical applications of FTIR are: environment analysis [32], food industry [33], polymer science [34], pharmaceutical industry [35] as well as forensics [36], among others.

In particular, FTIR spectroscopy has also been applied for lamb meat with the aim to detect lard mixed with lamb body fat [37]. In closely related work, near-infrared hyperspectral imaging has been applied to detect adulteration in minced lamb meat [30], [38], the assessment of tenderness [24] or lamb muscle discrimination [26], [28].

Feature subset selection is of great importance in the field of machine learning. Feature selection techniques evaluate the importance of a feature or a subset of features according to a given measure. Three main benefits can be drawn from feature selection [19], [20]. Firstly, it reduces the risk of overfitting, in particular if our data is defined in a high dimensional feature space and too few training instances are available. Second, it allows to gain a deeper insight into the underlying processes by determining which features are the most correlated with the class labels. Third, it makes possible an increase in efficiency providing faster and more cost-effective prediction models.

Two main approaches have been followed to address the high dimensionality problem of FTIR spectra. Dimensionality reduction has been conducted either extracting factors - Principal Component Analysis (PCA), Principal Components of the Partial Least Square (PLS) model - [39]–[42] or applying feature selection methods for identifying the most relevant features (wavelengths in this context) prior to the model calibration [21], [43], [44]. The variables chosen with the first alternative do not have a direct physical meaning with respect to the sample being analyzed, while the approaches based on selected wavelengths are often preferred because they use original variables and therefore with physical interpretation.

Following a feature selection approach, a minimum Redundancy-Maximum Relevance (mRMR) filter has been proposed [43] to select relevant wavelengths in FTIR spectra of multi-component chemical mixtures. The wrapper approach based on sequential forward selection has outperformed the feature selection method based on a t-statistic for the detection of salmonella contamination in beef FTIR spectra [21]. In [44], the SELECT algorithm, based on stepwise decorrelation of variables, was used for wavelength selection in olive oil FTIR spectra.

Feature selection methods can be categorized into three groups: *filter*, *wrapper* and *embedded* approaches [20], [45], [46]. The *filter* techniques rely on general characteristics of the training data to rank the features according to a metric being independent of the classifier. The *wrapper* approaches select candidate subsets of features and assess their fitness based on the classification model performance. Finally, in the *embedded* techniques, the feature search mechanism is incorporated into the classifier objective function and are, therefore, specific to a given inductive learning algorithm.

Consider each sample $\mathbf{x}_i$, defined in a t-dimensional vector $x_i = (x_{i1}, x_{i2}, \ldots x_{it})$ where each component x_{ij} represents the value of a given feature f_j for that example i, that is, $f_j(\mathbf{x}_i) = x_{ij}$.

From a functional point of view, the output of a feature selection algorithm may be a ranking (weighting-score) on the features or a feature set. Obviously, representation changes are possible and thus, a feature subset can be extracted from a full ranked list by selecting the most important features.

Consider now a feature ranking algorithm that provides a ranking vector r with components defined in (1)

$$r = (r_1, r_2, r_3, \ldots, r_t) \tag{1}$$

where $1 \leq r_i \leq t$. Note that 1 is considered the highest rank. Consider also a feature subset (as denoted in (2)) with k elements as the outcome of a feature selection technique

$$s = (s_1, s_2, s_3, \ldots, s_t), \quad s_i \in \{0, 1\} \tag{2}$$

where 1 indicates the presence of a feature and 0 the absence and $\sum_{i=1}^{t} s_i = k$ for a top-k list.

Lists with a full ranking of features can be converted into top-k lists that contain the most important k features. Converting a ranking output into a feature subset is easily conducted according to

$$s_i = \begin{cases} 1, & \text{if } r_i \leq k \\ 0, & \text{otherwise} \end{cases}$$

A. STABILITY OF FEATURE SELECTION METHODS

A problem that arises in many real world problems is that small variations in the available data set lead to different outcomes of the feature selection method. In particular, it is common when dealing with high dimensional data and few samples such as the problem we are working on. This fact makes the conclusions derived from the study unreliable, in particular when knowledge is to be extracted from the ranking of features or the analysis of the top-k features. For this reason the topic of stability has attracted great interest over the last few years [47]–[54].

Suppose we ran a feature ranking algorithm K times. Results are gathered in a matrix $\mathcal{A}$ with elements r_{ij} with $i = 1, \ldots, t$ and $j = 1, \ldots, K$ that indicate the rank assigned in the run-j for feature-i. The same applies to a feature selector. In general, stability is quantified as follows: Given a set of rankings (subsets) of the same feature selector on slightly different datasets, pairwise similarities are computed and then, reduced to a single metric by averaging [49], [55], [56]. Examples of these metrics are the Jaccard distance [55] or Kuncheva stability index [56].

A visual approach can also be followed. In order to study the stability with a visual-based approach, different alternatives could be used, depending on the amount of information available. Note that even simple visualization approaches like histograms or scatter graphs allow the depiction of the results in a convenient way to ease result interpretation. However, they have some limitations as the number of dimensions increases.

A dimensionality reduction technique like MultiDimensional Scaling (MDS) [57], that preserves as much of the original data structure as possible, may be more convenient. It allows the projection of data from a high dimensional space to a 2D or 3D space while preserving the distance in the original high dimensional space. This technique was first used in the context of machine learning for classifier comparison with respect to multiple metrics and multiple domains [58], [59].

Thus the outcome of a feature ranking algorithm can be interpreted as a point in a high dimensional space (with t dimensions). As mentioned above, stability is assessed computing pairwise similarities between points in that high dimensional space and averaging the results. In this case, the ranking data is turned into a single number (projected to one dimension) and the algorithms are compared on the basis of this metric.

These scalar metrics can be seen as projections to one dimensional space and its use only shows where the feature selector stands in relation to the stable and the random ranking algorithm. If we change from a projection to a one-dimensional space, into a space with two or more dimensions, we have a visual representation that allows to establish comparisons with respect to the random selector as well as comparisons of each feature selector to the others [60]. Conducting a visual-based stability analysis allows the evaluation of the similarity between feature ranking algorithms as well as their stability.

III. METHODS

Our aim is to calibrate a classification model able to discriminate fat samples according to the rearing system based on their FTIR spectra. We consider a training dataset $\mathcal{D} = \{(\mathbf{x}_i, d_i), i = 1, \ldots, M\}$ with M examples and a two-class label d associated with each sample ($d = 0$ and $d = 1$).

The feature selection methods studied in this work are briefly described below.

A. FEATURE SELECTION WITH FILTERS

Within this category, we consider the well-known Relief algorithm together with the metrics from the information Theory field.

1) χ^2 STATISTIC

The χ^2 statistic evaluates the worth of each feature individually by computing a value with respect to the class (two, in our work) [61]–[63]. Numeric features are discretized in several intervals. Then, the χ^2 for each one is computed following (3) as

$$\chi^2 = \sum_{i=1}^{p} \sum_{j=0}^{1} \frac{\left(A_{ij} - \frac{R_i M_j}{M}\right)^2}{\frac{R_i M_j}{M}} \tag{3}$$

where p is the number of intervals, M total number of instances, R_i the number of examples in the i-th interval, M_j the number of examples in the j-th class and A_{ij} the number of examples in the i-th interval, j-th class. Finally, the features are either ranked according to the χ^2 statistic or the features with highest values are selected as the most relevant.

2) INFORMATION GAIN (IG)

The worth of an attribute x_r is evaluated by measuring the Information Gain (IG) with respect to the class C_i [62], [64], [65]. This metric is given by (4)

$$IG(C_i, x_r) = H(C_i) - H(C_i|x_r). \tag{4}$$

where H represents the entropy.

3) GAIN RATIO (GR)

This filter evaluates the worth of an attribute by measuring the Gain Ratio (GR) with respect to the class. It is a modification of the IG filter that reduces certain bias issues found with IG attribute selection filter [62], [65], [66] as stated in (5):

$$GR(C_i, x_r) = \frac{H(C_i) - H(C_i|x_r)}{H(x_r)}. \tag{5}$$

4) RELIEF

The basic idea of the Relief algorithm is to re-weight features according to their ability to distinguish examples of the same and different classes that are near to each other [62], [67].

B. FEATURE SELECTION WITH WRAPPER APPROACHES

Wrapper methods use the performance of a learning algorithm to assess the usefulness of a feature set. Either they iteratively discard features with the least discriminant power or they add the best features according to model performance [20]. Wrapper approaches often achieve better results than filters due to the fact that they are calibrated with the interaction between training data and a classification model. However, they are much slower and more computationally intensive than filter methods since it requires the calibration of a classification model. Blocks of features can be removed or added simultaneously to make this process faster. In this work, we evaluate a Neural Network wrapper approach (NN-Wrapper).

C. FEATURE SELECTION WITH EMBEDDED APPROACHES

In this work we evaluate two embedded feature ranking strategies based on SVM and decision rules 1R, respectively.

1) SVM WITH RECURSIVE FEATURE ELIMINATION

This feature selection technique [20] determines which features provide the best contribution to the precision of the model while it is being created. This increases the performance in terms of time compared with the wrapper techniques.

2) DECISION RULE 1R

1R [68] builds rules based on single features (predictive 1-rules). They can be seen as single level decision trees. We can use 1R as a feature selector [69], on the assumption that the classification accuracy is a good indicator of the feature relevance.

D. NEURAL CLASSIFICATION MODEL

A Multilayer Perceptron (MLP) has been used to classify the spectra. We evaluate a three layer network with a logistic sigmoid activation function for the hidden and the output layers.

The cost function used in training was the Mean Square Error (MSE), minimized by the backpropagation algorithm with adaptive learning rate and momentum.

Several combinations of neurons in the hidden layer and different number of training cycles have been assessed with all the descriptors, in order to find the optimal network configuration.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we build a lamb meat categorization model and assess several feature ranking algorithms. The evaluation is conducted in terms of the classification performance (predictive power) and the robustness (or stability) of the feature selection algorithms.

A. FAT SAMPLES

Fourier-transform infrared spectroscopy (FTIR) has been used as a powerful tool for food quality assessment. We conducted our experiments on a real world spectral dataset intended to link lambs' diets to meat quality. Lambs came from the flocks of three farms affiliated to the 'Asociación Nacional de Criadores de Ganado Ovino de Raza Churra', which is a Churra breeders association from the region of 'Castilla-Leon' (Spain). Lambs were reared either exclusively on Ewe Milk (EM) or on a Milk Replacer (MR) (from up to three days after birth to slaughter).

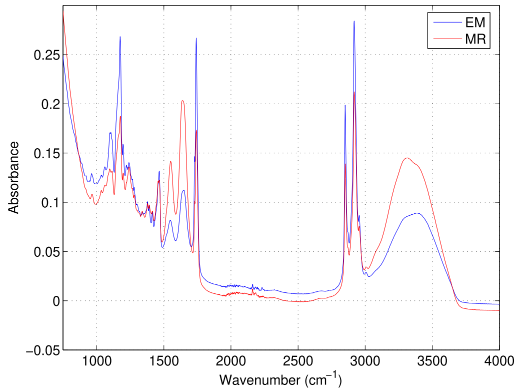

FIGURE 1. Average FT-IR spectrum of omental fat samples for lambs reared with a Milk Replacer (MR) and Ewe Milk (EM).

TABLE 1. Fat lamb spectral dataset.

Number of attributes	Class EM # instances	Class MR # instances
1687	68	66

Approximately four hours after slaughter, an omental fat sample was obtained from each of the carcasses. We collected omental fat samples from carcasses of suckling lambs [70]. The whole dataset has 134 instances: 66 from lambs being fed with a MR, while the other 68 are reared on ewe milk EM. Fat samples were packed individually in Ziploc freezing plastic bags (SC Johnson, Racine, WI, USA) and frozen and stored at −40°C for up to three months prior to analysis.

After thawing (overnight at 4°C), fat samples were homogenized using an IKALabortechnik A10 blender (IKA, Staufen, Germany) and analyzed using FTIR. The MB100 FT-IR spectrometer (Arid-ZoneTM, Quebec, Canada) was used to record the spectra.

All FTIR spectra were recorded from 4000 to 750cm^{-1} with a resolution of 4cm^{-1}, which leads to a total of 1687 features. Spectra of fat samples were taken using an Attenuated Total Reflectance (ATR) device with a Durascope diamond crystal (SensIR Technologies, Norwalk, CT, USA). The fat samples were squeezed against the ATR diamond crystal. A total of 32 scans were collected for each spectrum and the average calculated and subtracted from the background spectrum using an empty ATR diamond crystal. The data acquisition and processing software used was that of Win-Bomem Easy (Galactic Industries Corp., Salem, NH, USA).

The average spectra for each class (Milk replacer and Ewe milk) is shown in Figure 1 and the dataset description is displayed in Table 1.

B. CATEGORIZATION WITH THE WHOLE SPECTRA

We used MATLAB to train the neural classifier. We assessed network architectures with 1687 nodes in the input layer, sigmoidal activation functions and different number of nodes in the hidden layer as well as different number of training cycles.

Performance was estimated by 10-fold cross-validation whereby the data is randomly split into 10 folds and for each iteration one fold is used for testing and the other nine for training. The results shown next are the average over the 10 runs. In order to select the model complexity and training conditions, we used a 5-fold cross-validation method applied on the training data (that is, with the data belonging in each iteration to the 9-folds mentioned before).

After assessing by cross-validation several architectures (up to 30 nodes in the hidden layer) and different numbers of training cycles (up to 10000), the optimal configuration found was a neural classifier with 10 nodes in the hidden layer and and 700 training cycles. The performance in terms of accuracy for this classifier using the whole spectrum (1678 wavelengths) is $A_{1687} = 89.70\%$.

C. RELEVANT WAVELENGTH IDENTIFICATION

We now assess the six feature selectors. These are based on filters (χ^2, Information Gain (IG), Gain Ratio (GR), Relief) and another two embedded techniques based on the parameter values of an independent classifier (Decision Rule 1R and SVM).

The dataset was randomly split in ten folds, launching the feature ranking algorithm with nine out the ten folds, in a consecutive way. Five runs of this process resulted in a total of $K = 50$ rankings. Feature ranking was carried out with WEKA [62] and the assessment of the stability with MATLAB.

All the results gathered in the experiment can be organized as a set of 300 points (6 algorithms x 50 runs each one) defined a 1687-dimensional space.

MultiDimensional Scaling (MDS) [57] is used in this section to visualize both the feature selectors as well as a completely random selector in a graph so that comparisons between all of them can be established.

These points are projected to a 2D space using MDS. The distance between points is calculated with the Spearman's rank correlation coefficient and the stress criterion is normalized with the sum of squares of the dissimilarities.

After the projection, each outcome of the algorithm is represented by two coordinates (x,y) and the similarities among feature selectors can be seen in Figure 2. In terms of stability we can see that the points that correspond to Relief are much less scattered than the rest. In other words, this is the most stable algorithm. The outcomes of the other feature selectors appear to be more scattered, which indicates their lower stability. This graph allows us to see that IG generates similar rankings to GR and χ^2 and that they can be considered equivalent in this context. Figure 2 clearly shows which feature selection algorithms tend to be unstable. Thus, a single run of one of these feature ranking algorithms is unlikely to be representative or reliable.

The results gathered in the experiment can be also organized as a set of 6 points (6 feature selection methods) defined in 84350-dimensional space, where 84350 is the result of multiplying 50 runs by the 1687 ranking vector

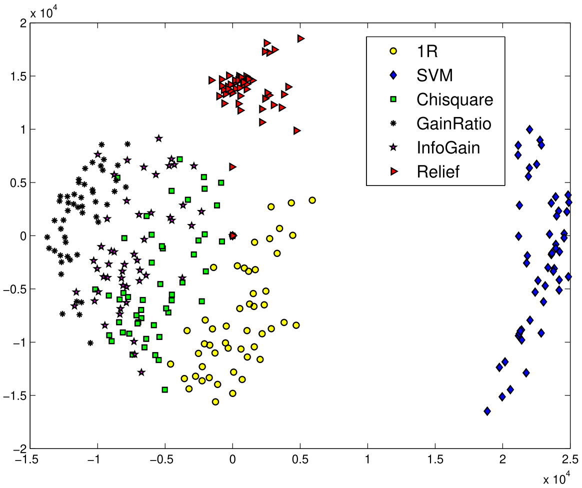

FIGURE 2. MDS plot of the feature selection algorithms for fat spectra data.

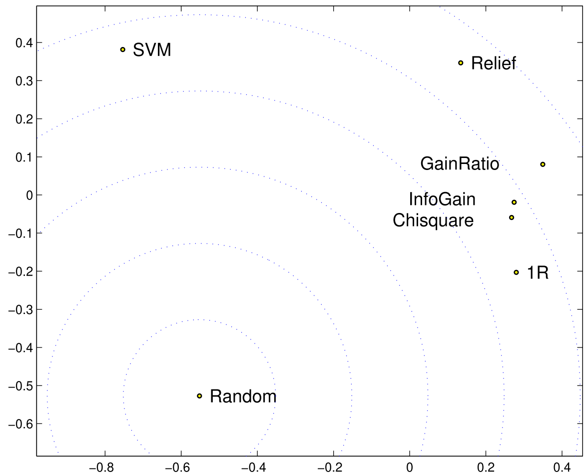

FIGURE 3. MDS plot of the feature selection algorithms together with a random feature selector.

obtained in each run. An extra point corresponding to a random feature ranking algorithm can also be generated through simulation.

After the projection with MDS to a two dimensional space, we plot the results in Figure 3. It depicts the the distance to the Random selector as well the distances among the different feature ranking algorithms. Relief is the most distant to a trivial random feature ranking algorithm. The figure also indicates that the ranking yielded by Relief is very different from the one generated by SVM and the other equivalent groups (IG, GR, χ^2).

Since Figure 3 shows dissimilarities among the rankings, they should be evaluated in order to determine the quality of the selected features to predict the target class (EM,MR). This is crucial in order to provide the veterinarian experts with reliable information about the most important regions of the spectrum, and not only with the most stable top-k list.

TABLE 2. Accuracy (in %) for a neural network classifier with feature subsets of different cardinality selected by several feature selection algorithms.

Number of features	χ^2	IG	GR	Relief	1R	SVM
5	84.54	84.54	84.54	83.39	81.06	87.40
10	82.79	81.85	83.55	82.48	80.52	90.39
15	80.76	80.68	80.61	82.05	82.38	88.58
20	83.36	80.52	82.31	84.08	82.70	88.89
30	85.65	81.79	83.08	85.37	82.94	88.40
40	85.54	86.46	83.19	84.74	83.92	89.38
50	84.62	85.46	81.75	85.48	84.49	88.61
60	83.63	84.76	81.52	84.92	84.73	88.15
80	85.92	86.25	80.58	83.88	85.39	89.45
100	85.38	86.26	82.73	83.45	85.85	88.48
120	86.96	86.93	82.13	83.43	86.95	89.05
140	86.55	85.52	82.62	82.40	86.80	89.85
160	87.44	86.89	84.33	82.40	87.60	90.57
180	86.30	86.77	83.55	82.96	87.05	89.69
200	87.10	86.33	85.07	82.64	86.32	90.80
300	87.36	87.19	82.82	82.79	86.58	**91.80**
600	86.89	86.82	86.13	87.58	89.14	90.93
900	86.54	86.42	87.29	89.55	**90.50**	91.29
1200	88.24	88.39	87.83	**90.46**	89.76	91.08
1500	**89.71**	**90.30**	**89.73**	90.00	90.46	90.54

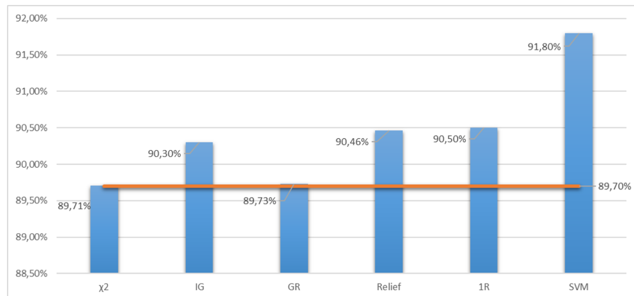

FIGURE 4. Best performance (accuracy) achieved by the MLP classifier with different feature selection techniques. Reference accuracy with the full feature set.

D. CATEGORIZATION WITH RELEVANT WAVELENGTHS

In this section, we evaluate the performance of a neural classifier trained with feature subsets of different cardinality selected by the feature ranking algorithms: χ^2, IG, GR, Relief, 1R and SVM.

In order to make the feature ranking process more robust a final ranking is generated as the result of aggregating the 50 rankings by computing their median value.

The classifier accuracy is estimated using 10-fold cross validation and the results shown in Table 2 are the average of 10 runs. Table 2 records the accuracy of a classifier trained with the top-k features (with k from 10 to 1500) selected for the different ranking algorithms.

Figure 4 shows the best performance (accuracy) achieved by the MLP classifier with different feature selection techniques together with the reference accuracy with the full feature set. As the previous analysis based on the MDS projections revealed, GI, GIR and χ^2 lead to similar performance since their rankings are very similar. Selecting the features according to Relief allows to get a classifier with

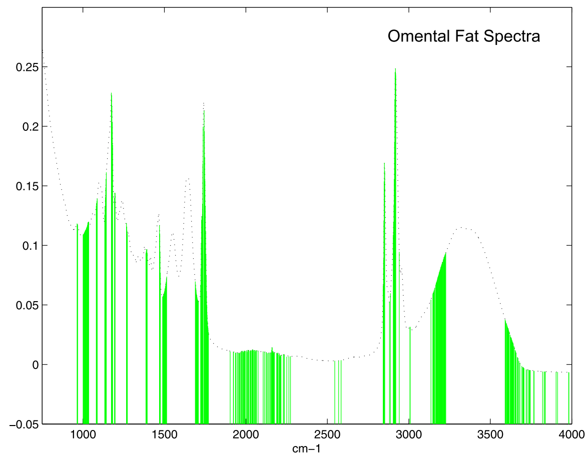

FIGURE 5. Fat spectra together with the highest ranked wavelengths selected by the feature selection algorithm based on SVM.

higher accuracy than the metrics based on Information theory (90.46% with 1200 features). The ranking generated with 1R leads to similar performance (90.50% accuracy). However, the ranking carried out with SVM, achieves the highest accuracy. Thus, the accuracy is 91, 80% with the top-300 features selected by the SVM feature selector algorithm.

Figure 5 plots the average spectrum of the fat dataset. It also shows the top-300 wavelengths selected by the feature selection algorithm based on SVM (the most effective in terms of classifier performance).

Information about the most discriminant regions of the spectra is useful to interpret which fatty acids establish differences between animals reared by maternal milk and by milk replacers. The initial guess of the veterinarian experts was that the region between 1700cm^{-1} and 2700cm^{-1} was a noisy band with little information. However, our analysis reveals that the region of the spectrum between 2000cm^{-1} and 2200cm^{-1} is crucial for the fat categorization task. Likewise, the region between 1100cm^{-1} and 1200cm^{-1}. This information becomes useful to interpret which fatty acids are responsible for the differences in the fat samples that come from feeding with maternal milk with respect to a rearing system with milk replacers. Some of the relevant features selected by SVM were related to the major wavelength bands attributed to functional groups of fatty acids and glycerol, which comes from differences in the rearing system.

V. COMPARISON WITH OTHER APPROACHES

In this section we compare the performance of the neural classifier with other approaches not based on feature selection, but on feature extraction and therefore generating new features with no possibility of physical interpretation. We evaluated the feature extraction method based on Principal Component Analysis (PCA) and the Partial Least Square (PLS) model.

Additionally, the feature selection approach based on wrappers is also assessed. This alternative has not been recommended for problems defined in high dimensional spaces with few samples such as the one we deal with in this work. We implemented, though, a light version of the wrappper feature selector for this problem.

TABLE 3. Accuracy (in %) for an MLP classifier with different number of PCA components.

Number of components	PCA
4	80.14
8	81.80
12	90.20
16	92.23
20	90.99
24	90.37
28	93.52
32	**93.89**
36	92.92
40	91.78

A. PRINCIPAL COMPONENT ANALYSIS (PCA)

Principal component analysis (PCA) is a well-known statistical method for dimensionality reduction that extract new variables (principal components) from linear combinations of the original variables in an unsupervised way. A few principal components may retain the same information as many original variables and explain most of the data variability.

The classification accuracy using features extracted by PCA is shown in Table 3. The best result is achieved with 32 PCA components, which is an accuracy of 93.89%. Classification performance with this set of features outperformed the model that employs the feature set selected by SVM (91.80% against 93.89%).

B. PARTIAL LEAST SQUARE (PLS)

The PLS (Partial Least Square) model, which is widely used in the field of chemometrics, works by extracting a few orthogonal latent components that are linear combinations of the original ones. The classification with PLS was carried out with the MATLAB Toolbox [39]. Spectral data were pre-processed using Multiplicative Signal Correction (MSC) followed by mean centering [71]. The number of PLS components was automatically determined by cross-validation. The model performance was estimated by cross validation with 10 folds.This PLS model with 4 components achieves an accuracy of 85.60%.

C. THE WRAPPER APPROACH

In this work we also evaluate NN-Wrapper, a wrapper approach that measures the relevance of a feature set based on the performance of a Neural Network with a MultiLayer Perceptron architecture. Model performance is estimated by the accuracy in classification.

Feature selection is carried out by backward elimination. We start with all the features and then remove those that affect the classification the least. Given the high number of features, these are grouped (according to wavelength) in sixteen windows of feature sets. Assessing by cross-validation the impact in classification, we remove iteratively those windows that reduced or do not significantly improve (lower than 0.5%) the classification accuracy. Reducing the dimensionality with this approach, the neural classifier yields an accuracy of 89.88%, and so does not show an improvement with respect the whole set of features.

VI. CONCLUSION AND FUTURE WORK

In this work, we address the problem of authentication of suckling lamb meat with respect to the type of feeding. The rearing system determines the difference in prices and quality.

It has been shown that the diet has effect on the composition of fat tissues. For this reason, FTIR spectroscopy has been assessed to categorized suckling lamb fat according to the rearing system. It provides several advantages to the existing analytical techniques mainly its speed, cost and versatility. The FTIR spectra comprise, however, a large number of irrelevant and redundant information. Appropriate feature selection becomes mandatory when building a categorization model for these fat samples. It helps to avoid overfitting and is an aid to identify the wavelengths with more prediction power.

We have conducted some experiments on a real world spectral dataset of fat samples from suckling lambs. In this work we evaluated different feature ranking algorithms: χ^2, IG, GR, Relief, and another two embedded methods based on 1R and SVM. We have studied their effect on classification performance at the same time their stability is also evaluated.

We discovered that the feature selectors tend to be unstable for this real world application, likely due to the combination of high dimensionality and relatively few samples. Relief turned out to be the most stable algorithm, while the other techniques provided more scattered outcomes. We faced this problem by aggregating (computing the median value) 50 rankings generated from slightly different training data sets. The experimental results show that the selection of an appropriate subset of wavelengths can considerably improve the classification accuracy. Thus, the neural classifier with the whole spectrum yields an accuracy of $A_{1687} = 89.70\%$ while it increases to $A_{900} = 90.50\%$ with the 900 most relevant wavelengths selected by 1R or to $A_{300} = 91.80\%$ with the 300 most relevant ones selected by SVM.

The feature selection methods together with the neural classifier have been compared with the PLS model based on feature extraction and widely used in chemometrics. We observed that the MLP classifier significantly increases the accuracy: from an accuracy of 85.60% with he PLS model to accuracy $A = 93.89\%$ (MLP and 32 PCA components) or $A = 91.80\%$ (MLP and 300 features selected by SVM).

Additionally, the neural network classifier together with the feature selection by SVM allows us to identify which regions of the spectrum have more discriminant power (and absorbance is directly related to the chemical information). This has a particular interest for veterinarian experts. The PLS model, however, can not offer this analysis capabilities for the veterinarian professionals since it creates a reduced new set of features from the original ones.

Considering the decision costs linked to the classification errors, important for a quality label system, is part of our future research work. Additionally, the proposal and evaluation of different techniques to increase stability becomes also important in this real world application.

REFERENCES

[1] C. Sañudo, M. Alfonso, R. San Julián, G. Thorkelsson, T. Valdimarsdottir, D. Zygoyiannis, C. Stamataris, E. Piasentier, C. Mills, P. Berge, E. Dransfield, G. R. Nute, M. Enser, and A. V. Fisher, "Regional variation in the hedonic evaluation of lamb meat from diverse production systems by consumers in six European countries," *Meat Sci.*, vol. 75, no. 4, pp. 610–621, Apr. 2007.

[2] V. Gkarane, N. P. Brunton, P. Allen, R. S. Gravador, N. A. Claffey, M. G. Diskin, A. G. Fahey, L. J. Farmer, A. P. Moloney, M. J. Alcalde, P. Murphy, and F. J. Monahan, "Effect of finishing diet and duration on the sensory quality and volatile profile of lamb meat," *Food Res. Int.*, vol. 115, pp. 54–64, Jan. 2019.

[3] M. Lanza, M. Bella, A. Priolo, D. Barbagallo, V. Galofaro, C. Landi, and P. Pennisi, "Lamb meat quality as affected by a natural or artificial milk feeding regime," *Meat Sci.*, vol. 73, no. 2, pp. 313–318, Jun. 2006.

[4] G. Ripoll, M. J. Alcalde, A. Argüello, M. D. G. Córdoba, and B. Panea, "Effect of the rearing system on the color of four muscles of suckling kids," *Food Sci. Nutrition*, vol. 7, no. 4, pp. 1502–1511, Apr. 2019.

[5] S. J. Ward, M. Campo, and G. Liste, "The effects of artificial rearing and fostering on the growth, carcass and meat quality of lambs," *Small Ruminant Res.*, vol. 149, pp. 16–22, Apr. 2017.

[6] S. W. Erasmus, M. Muller, and L. C. Hoffman, "Authentic sheep meat in the European union: Factors influencing and validating its unique meat quality," *J. Sci. Food Agricult.*, vol. 97, no. 7, pp. 1979–1996, May 2017.

[7] M. T. Osorio, J. M. Zumalacárregui, A. Figueira, and J. Mateo, "Physicochemical properties of perirenal and omental fat from suckling lamb carcasses evaluated according to the type of milk source," *Small Ruminant Res.*, vol. 72, nos. 2–3, pp. 111–118, Oct. 2007.

[8] M. T. Osorio, J. M. Zumalacárregui, A. Figueira, and J. Mateo, "Fatty acid composition in subcutaneous, intermuscular and intramuscular fat deposits of suckling lamb meat: Effect of milk source," *Small Ruminant Res.*, vol. 73, nos. 1–3, pp. 127–134, Nov. 2007.

[9] A. M. Jiménez-Carvelo, A. González-Casado, M. G. Bagur-González, and L. Cuadros-Rodríguez, "Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity—A review," *Food Res. Int.*, vol. 122, pp. 25–39, Aug. 2019.

[10] M. P. Callao and I. Ruisánchez, "An overview of multivariate qualitative methods for food fraud detection," *Food Control*, vol. 86, pp. 283–293, Apr. 2018.

[11] V. Babushkin, A. Spiridonov, and A. Kozhukhar, "Application of NIR and FTIR in food analysis," *J. Phys. Sci. Appl.*, vol. 6, no. 2, pp. 47–50, Feb. 2016.

[12] E. Hong, S. Y. Lee, J. Y. Jeong, J. M. Park, B. H. Kim, K. Kwon, and H. S. Chun, "Modern analytical methods for the detection of food fraud and adulteration by food category," *J. Sci. Food Agricult.*, vol. 97, no. 12, pp. 3877–3896, Sep. 2017.

[13] A. Bendini, L. Cerretani, F. Di Virgilio, P. Belloni, M. Bonoli-Carbognin, and G. Lercker, "Preliminary evaluation of the application of the FTIR spectroscopy to control the geographic origin and quality of virgin olive oils," *J. Food Qual.*, vol. 30, no. 4, pp. 424–437, Aug. 2007.

[14] L. E. Rodriguez-Saona and M. E. Allendorf, "Use of FTIR for rapid authentication and detection of adulteration of food," *Annu. Rev. Food Sci. Technol.*, vol. 2, no. 1, pp. 467–483, Apr. 2011.

[15] S. M. Obeidat, A. Y. Hammoudeh, and A. A. Alomary, "Application of FTIR spectroscopy for assessment of green coffee beans according to their origin," *J. Appl. Spectrosc.*, vol. 84, no. 6, pp. 1051–1055, Jan. 2018.

[16] M. Lucarini, A. Durazzo, J. Sánchez del Pulgar, P. Gabrielli, and G. Lombardi-Boccia, "Determination of fatty acid content in meat and meat products: The FTIR-ATR approach," *Food Chem.*, vol. 267, pp. 223–230, Nov. 2018.

[17] U. Khan, M. Afzaal, M. Arshad, and M. Imran, "Non-destructive analysis of food adulteration and legitimacy by FTIR technology," *J. Food Ind. Microbiol.*, vol. 1, no. 1, pp. 1–7, 2015.

[18] A. Rohman, "The employment of Fourier transform infrared spectroscopy coupled with chemometrics techniques for traceability and authentication of meat and meat products," *J. Adv. Vet. Animal Res.*, vol. 6, no. 1, pp. 9–17, 2019.
[19] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Jan. 2003.
[20] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, *Feature Extraction: Foundations and Applications* (Studies in Fuzziness and Soft Computing). Secaucus, NJ, USA: Springer-Verlag, 2006.
[21] J. K. Amamcharla, S. Panigrahi, C. M. Logue, M. Marchello, and J. S. Sherwood, "Application of vapour-phase Fourier transform infrared spectroscopy (FTIR) and statistical feature selection methods for identifying *salmonella enterica typhimurium* contamination in beef," *Biosyst. Eng.*, vol. 107, no. 1, pp. 1–9, Sep. 2010.
[22] R. M. Balabin and S. V. Smirnov, "Variable selection in near-infrared spectroscopy: Benchmarking of feature selection methods on biodiesel data," *Analytica Chim. Acta*, vol. 692, nos. 1–2, pp. 63–72, Apr. 2011.
[23] C. S. W. Miaw, C. Assis, A. R. C. S. Silva, M. L. Cunha, M. M. Sena, and S. V. C. de Souza, "Determination of main fruits in adulterated nectars by ATR-FTIR spectroscopy combined with multivariate calibration and variable selection methods," *Food Chem.*, vol. 254, pp. 272–280, Jul. 2018.
[24] M. Kamruzzaman, G. ElMasry, D.-W. Sun, and P. Allen, "Non-destructive assessment of instrumental and sensory tenderness of lamb meat using NIR hyperspectral imaging," *Food Chem.*, vol. 141, no. 1, pp. 389–396, Nov. 2013.
[25] P. Cortez, M. Portelinha, S. Rodrigues, V. Cadavez, and A. Teixeira, "Lamb meat quality assessment by support vector machines," *Neural Process. Lett.*, vol. 24, no. 1, pp. 41–51, Aug. 2006.
[26] J. A. Sanz, A. M. Fernandes, E. Barrenechea, S. Silva, V. Santos, N. Gonçalves, D. Paternain, A. Jurio, and P. Melo-Pinto, "Lamb muscle discrimination using hyperspectral imaging: Comparison of various machine learning algorithms," *J. Food Eng.*, vol. 174, pp. 92–100, Apr. 2016.
[27] M. Kamruzzaman, G. ElMasry, D.-W. Sun, and P. Allen, "Application of NIR hyperspectral imaging for discrimination of lamb muscles," *J. Food Eng.*, vol. 104, no. 3, pp. 332–340, Jun. 2011.
[28] H. Pu, A. Xie, D.-W. Sun, M. Kamruzzaman, and J. Ma, "Application of wavelet analysis to spectral data for categorization of lamb muscles," *Food Bioprocess Technol.*, vol. 8, no. 1, pp. 1–16, Jan. 2015.
[29] P. Slósarz, M. Stanisz, P. Boniecki, A. Przybylak, D. Lisiak, and A. Ludwiczak, "Artificial neural network analysis of ultrasound image for the estimation of intramuscular fat content in lamb muscle," *Afr. J. Biotechnol.*, vol. 10, no. 55, p. 11792, 2011.
[30] M. Kamruzzaman, D.-W. Sun, G. ElMasry, and P. Allen, "Fast detection and visualization of minced lamb meat adulteration using NIR hyperspectral imaging and multivariate image analysis," *Talanta*, vol. 103, pp. 130–136, Jan. 2013.
[31] B. C. Smith, *Fundamentals of Fourier Transform Infrared Spectroscopy*. Boca Raton, FL, USA: CRC Press, 2011.
[32] C. M. Simonescu, "Application of FTIR spectroscopy in environmental studies," in *Advanced Aspects of Spectroscopy*, M. A. Farrukh, Eds. Rijeka, Croatia: IntechOpen, 2012, ch. 2, doi: 10.5772/48331.
[33] D.-W. Sun, *Infrared Spectroscopy for Food Quality Analysis and Control*. New York, NY, USA: Academic, 2009.
[34] S. Kalia, B. S. Kaith, and I. Kaur, "Pretreatments of natural fibers and their application as reinforcing material in polymer composites—A review," *Polym. Eng. Sci.*, vol. 49, no. 7, pp. 1253–1272, 2009.
[35] A. Rahman, "Application of Fourier transform infrared spectroscopy for quality control of pharmaceutical products: A review," *Indonesian J. Pharmacy*, pp. 1–8, Jan. 2012.
[36] C.-M. Orphanou, "The detection and discrimination of human body fluids using ATR FT-IR spectroscopy," *Forensic Sci. Int.*, vol. 252, pp. e10–e16, Jul. 2015.
[37] Y. B. Che Man and M. E. S. Mirghani, "Detection of lard mixed with body fats of chicken, lamb, and cow by Fourier transform infrared spectroscopy," *J. Amer. Oil Chemists' Soc.*, vol. 78, no. 7, pp. 753–761, Jul. 2001.
[38] X. Zheng, Y. Li, W. Wei, and Y. Peng, "Detection of adulteration with duck meat in minced lamb meat by using visible near-infrared hyperspectral imaging," *Meat Sci.*, vol. 149, pp. 55–62, Mar. 2019.
[39] L. Nørgaard, A. Saudland, J. Wagner, J. P. Nielsen, L. Munck, and S. B. Engelsen, "Interval partial least-squares regression (*i*PLS): A comparative chemometric study with an example from near-infrared spectroscopy," *Appl. Spectrosc.*, vol. 54, no. 3, pp. 413–419, 2000.
[40] P. Geladi and B. R. Kowalski, "Partial least-squares regression: A tutorial," *Analytica Chim. Acta*, vol. 185, pp. 1–17, Jan. 1986.
[41] S. Maitra and J. Yan, "Principle component analysis and partial least squares: Two dimension reduction techniques for regression," *Applying Multivariate Stat. Models*, vol. 79, pp. 79–90, Jun. 2008.
[42] A. Pomerantz, Y. Cohen, E. Shufan, Y. Ben-Naim, S. Mordechai, A. Salman, and M. Huleihel, "Characterization of phytophthora infestans resistance to mefenoxam using FTIR spectroscopy," *J. Photochemistry Photobiol. B, Biol.*, vol. 141, pp. 308–314, Dec. 2014.
[43] A. Vergara and E. Llobet, "Feature selection versus feature compression in the building of calibration models from FTIR-spectrophotometry datasets," *Talanta*, vol. 88, pp. 95–103, Jan. 2012.
[44] M. Casale, N. Sinelli, P. Oliveri, V. Di Egidio, and S. Lanteri, "Chemometrical strategies for feature selection and data compression applied to NIR and MIR spectra of extra virgin olive oils for cultivar identification," *Talanta*, vol. 80, no. 5, pp. 1832–1837, Mar. 2010.
[45] V. Bolón-Canedo, N. Sánchez-Maroño, and A. Alonso-Betanzos, "A review of feature selection methods on synthetic data," *Knowl. Inf. Syst.*, vol. 34, no. 3, pp. 483–519, Mar. 2013.
[46] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Comput. Electr. Eng.*, vol. 40, no. 1, pp. 16–28, Jan. 2014.
[47] S. Nogueira and G. Brown, "Measuring the stability of feature selection," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases* Springer, 2016, pp. 442–457.
[48] T. Abeel, T. Helleputte, Y. Van de Peer, P. Dupont, and Y. Saeys, "Robust biomarker identification for cancer diagnosis with ensemble feature selection methods," *Bioinformatics*, vol. 26, no. 3, pp. 392–398, Feb. 2010.
[49] S. Nogueira, K. Sechidis, and G. Brown, "On the stability of feature selection algorithms," *J. Mach. Learn. Res.*, vol. 18, no. 174, pp. 1–54, 2018.
[50] W. W. B. Goh and L. Wong, "Evaluating feature-selection stability in next-generation proteomics," *J. Bioinf. Comput. Biol.*, vol. 14, no. 05, Oct. 2016, Art. no. 1650029.
[51] É. Perthame, C. Friguet, and D. Causeur, "Stability of feature selection in classification issues for high-dimensional correlated data," *Statist. Comput.*, vol. 26, no. 4, pp. 783–796, Jul. 2016.
[52] J. González, J. Ortega, M. Damas, P. Martín-Smith, and J. Q. Gan, "A new multi-objective wrapper method for feature selection—Accuracy and stability analysis for BCI," *Neurocomputing*, vol. 333, pp. 407–418, Mar. 2019.
[53] B. Pes, "Ensemble feature selection for high-dimensional data: A stability analysis across multiple domains," *Neural Comput. Appl.*, pp. 1–23, Feb. 2019.
[54] V. Bolón-Canedo and A. Alonso-Betanzos, "Ensembles for feature selection: A review and future trends," *Inf. Fusion*, vol. 52, pp. 1–12, Dec. 2019.
[55] A. Kalousis, J. Prados, and M. Hilario, "Stability of feature selection algorithms: A study on high-dimensional spaces," *Knowl. Inf. Syst.*, vol. 12, no. 1, pp. 95–116, May 2007.
[56] L. Kuncheva, "A stability index for feature selection," in *Proc. 25th Conf. Proc. 25th IASTED Int. Multi-Conf., Artif. Intell. Appl.* Calgary, AB, Canada: ACTA Press, 2007, pp. 390–395.
[57] T. Cox and M. Cox, *Multidimensional Scaling*. London, U.K.: Chapman & Hall, Oct. 1994.
[58] R. Alaiz-Rodríguez, N. Japkowicz, and P. Tischer, "A visualization-based exploratory technique for classifier comparison with respect to multiple metrics and multiple domains," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Springer, 2008, pp. 660–665.
[59] R. Alaiz-Rodríguez, N. Japkowicz, and P. Tischer, "Visualizing classifier performance on different domains," in *Proc. 20th IEEE Int. Conf. Tools Artif. Intell.*, vol. 2, Nov. 2008, pp. 3–10.
[60] N. Cueto-López, M. T. García-Ordás, V. Dávila-Batista, V. Moreno, N. Aragonés, and R. Alaiz-Rodríguez, "A comparative study on feature selection for a risk prediction model for colorectal cancer," *Comput. Methods Programs Biomed.*, vol. 177, pp. 219–229, Aug. 2019.
[61] H. Liu and R. Setiono, "Chi2: Feature selection and discretization of numeric attributes," in *Proc. 7th IEEE Int. Conf. Tools Artif. Intell.*, Nov. 1995, pp. 388–391.
[62] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques With Java Implementations*. San Mateo, CA, USA: Morgan Kaufmann, Oct. 1999.
[63] C. S. Kumar and R. Sree, "Application of ranking based attribute selection filters to perform automated evaluation of descriptive answers through sequential minimal optimization models," *ICTACT J. Soft Comput.*, vol. 5, no. 1, pp. 860–868, Oct. 2014.

[64] J. Novakovic, "Using information gain attribute evaluation to classify sonar targets," in *Proc. 17th Telecommun. Forum TELFOR*, 2009, pp. 1351–1354.
[65] M. A. Hall and L. A. Smith, "Practical feature subset selection for machine learning," in *Proc. 21st Australas. Comput. Sci. Conf. (ACSC)*, Perth, WA, Australia, C. McDonald, Ed. Berlin, Germany: Springer, Feb. 1998, pp. 181–191.
[66] R. P. Priyadarsini, M. Valarmathi, and S. Sivakumari, "Gain ratio based feature selection method for privacy preservation," *ICTACT J. Soft Comput.*, vol. 1, no. 4, pp. 201–205, 2011.
[67] K. Kira and L. A. Rendell, "A practical approach to feature selection," in *Proc. Mach. Learn.*, 1992, pp. 249–256.
[68] R. Holte, "Very simple classification rules perform well on most commonly used datasets," *Mach. Learn.*, vol. 11, no. 1, pp. 63–90, Apr. 1993.
[69] G. Holmes and C. Nevill-Manning, "Feature selection via the discovery of simple classification rules," in *Proc Symp. Intell. Data Anal. (IDA)*, Baden-Baden, Germany, 1995, pp. 75–79.
[70] M. Osorio, J. M. Zumalacárregui, R. Alaiz-Rodríguez, R. Guzman-Martínez, S. Engelsen, and J. Mateo, "Differentiation of perirenal and omental fat quality of suckling lambs according to the rearing system from Fourier transforms mid-infrared spectra using partial least squares and artificial neural networks analysis," *Meat Sci.*, vol. 83, no. 1, pp. 140–147, 2009.
[71] H. Martens, J. P. Nielsen, and S. B. Engelsen, "Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures," *Anal. Chem.*, vol. 75, no. 3, pp. 394–404, Feb. 2003.

ROCÍO ALAIZ-RODRÍGUEZ received the B.S. degree in electrical engineering from the University of Valladolid, Spain, in 1999, and the Ph.D. degree from the Carlos III University of Madrid, Spain, in 2005. She is currently an Associate Professor with the Universidad de León, Spain. Her research interests include machine learning, statistical pattern recognition, neural networks, and the dataset shift problem.

ANDREW C. PARNELL is currently a Hamilton Professor with the Hamilton Institute, Maynooth University. His research interest includes statistics and machine learning for large structured data sets in a variety of application areas. He has coauthored more than 70 peer-reviewed articles in relevant journals. He is also a funded Investigator with the SFI Insight Centre for Data Analytics and a Leader of the Advanced Analytics and Engineer Feedback Pillar of the SFI I-Form Advanced Manufacturing Centre.