A machine learning proposal method to detect milk tainted with cheese whey

Juliana S. Lima,[1] Daniela C. S. Z. Ribeiro,[1] Habib Asseiss Neto,[2] Sérgio V. A. Campos,[3] Mônica O. Leite,[1] Márcia E. de R. Fortini,[1] Beatriz Pinho Martins de Carvalho,[1] Marcos Vinícius Oliveira Almeida,[1] and Leorges M. Fonseca[1]*

[1]Department of Food Technology and Inspection, School of Veterinary Medicine, Universidade Federal de Minas Gerais (UFMG),
Belo Horizonte, MG, Brazil 31270-901
[2]Federal Institute of Mato Grosso do Sul, Três Lagoas, MS, Brazil 79641-162
[3]Department of Computer Science, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil 31270-901

ABSTRACT

Cheese whey addition to milk is a type of fraud with high prevalence and severe economic effects, resulting in low yield for dairy products, nutritional reduction of milk and milk-derived products, and even some safety concerns. Nevertheless, methods to detect fraudulent addition of cheese whey to milk are expensive and time consuming, and are thus ineffective as screening methods. The Fourier-transform infrared (FTIR) spectroscopy technique is a promising alternative to identify this type of fraud because a large number of data are generated, and useful information might be extracted to be used by machine learning models. The objective of this work was to evaluate the use of FTIR with machine learning methods, such as classification tree and multilayer perceptron neural networks to detect the addition of cheese whey to milk. A total of 520 samples of raw milk were added with cheese whey in concentrations of 1, 2, 5, 10, 15, 20, 25, and 30%; and 65 samples were used as control. The samples were stored at 7, 20, and 30°C for 0, 24, 48, 72, and 168 h, and analyzed using FTIR equipment. Complementary results of 520 samples of authentic raw milk were used. Selected components (fat, protein, casein, lactose, total solids, and solids nonfat) and freezing point (°C) were predicted using FTIR and then used as input features for the machine learning algorithms. Performance metrics included accuracy as high as 96.2% for CART (classification and regression trees) and 97.8% for multilayer perceptron neural networks, with precision, sensitivity, and specificity above 95% for both methods. The use of milk composition and freezing point predicted using FTIR, associated with machine learning techniques, was highly efficient to differentiate authentic milk from samples added with cheese whey. The results indicate that this is a potential method to be used as a high-performance screening process to detected milk adulterated with cheese whey in milk quality laboratories.
Key words: fraud, cheese whey, infrared spectroscopy, machine learning, artificial neural networks

INTRODUCTION

Raw milk tampering with cheese whey is a serious problem for the dairy industry, especially in developing countries. This fraudulent practice is of concern for food inspection agencies and consumers because of the nutritional value reduction of milk and some derivatives and even safety issues (Brandao et al., 2010; Robim et al., 2012; Tibola et al., 2018). For example, milk protein dilution after cheese whey tampering may motivate the addition of cheaper materials, such as urea or even hazardous chemicals such as melamine to disguise lower protein composition (Handford et al., 2016; Poonia et al., 2017).

However, suitable analytical methods to investigate this fraud usually are expensive, time consuming, and limit precision and accuracy (De La Fuente and Juárez, 2005; de Carvalho et al., 2015; Tibola et al., 2018). Notably, one of the most known methods for cheese whey detection, based on the quantification of the caseinomacropeptide (**CMP**) by HPLC, is a subject of controversy due to false-positive results and accuracy issues (Lenardon et al., 2017; de Pádua Alves et al., 2018; Raymundo et al., 2018; Lobato et al., 2020).

Current advances in the use of machine learning in analytical methods may be an answer to this problem because new and innovative techniques could be created and incorporated in a laboratory routine (De La Fuente and Juárez, 2005). Predictive methods such as machine learning algorithms are powerful modeling tools which can detect complex, nonlinear relationships between inputs and outputs (Alves da Rocha et al., 2015; Morota et al., 2018; Skansi, 2018; Neto et al., 2019). Their use has expanded to many fields such as

information technology, linguistics, medicine, finance, marketing, and so on. In analytics, several applications have been developed such as the use of neural networks associated with infrared analysis to investigate the addition of extraneous substances to milk, such as sugar and starch, among others (Liakos et al., 2018; Conceição et al., 2019; Neto et al., 2019).

Decision tree is a popular machine learning method for classification or regression problems. A decision tree is a predictive technique that involves splitting the values of the predictors based on a set of splitting rules, which divide the data into homogeneous subsets. A prediction, or decision, can be made by navigating from the root to a terminal node. Several algorithms can be used, depending on the type of tree. Decision tree learning is powerful, although simple and efficient, and can be easily understood, interpreted, and controlled (Wu et al., 2008; Ertel, 2017).

Classification and regression trees (**CART**) cover the use of trees as a data analysis method, and was developed by Leo Breiman (Breiman et al., 1984). Despite its simplicity and analysis power, CART use as a tool in food analytical methods has been limited (Hansen and Ferrao, 2020). In a CART tree, the dependent variable can be either categorical (classification trees) or continuous (regression trees), whereas the predictors can be both continuous and categorical (Bramer, 2016). For binary classification trees, the value of each terminal node is the mode of the observations in the corresponding subset, and the prediction accuracy is given by the percentage of correctly classified cases.

Artificial neural networks are a powerful learning method based on the components of the biological brain. An artificial neural network is composed by connected nodes called neurons, and each connection transmits a signal from one neuron to the other neurons in a deeper layer. An artificial neuron receives a signal, which is a data input, and processes the signal by evaluating a computational function with a specified weight value. Finally, the neuron transmits the result as another signal to other connected neurons. Neural networks can detect complex, nonlinear relationships between inputs and outputs, and their uses are found in different fields, such as finance (Xu and Zhang, 2021; Ghaffarian et al., 2022), marketing (Guiné et al., 2020), physics (Schiassi et al., 2021; Ma et al., 2022), linguistics (Lakretz et al., 2021), medicine (Dunnmon et al., 2019; Koo et al., 2021; Zhu et al., 2022), and so on. (Alpaydin, 2014; Witten et al., 2016).

Multilayer perceptron (**MLP**) is a feedforward network with 3 or more layers (one input, one output, and one or more hidden layers), which usually employs a sigmoid or a hyperbolic tangent function as an activation function. Indexes, such as accuracy, precision, sensitivity, specificity, may be valuable tools to estimate the performance of a prediction method, together with the receiver operating characteristics curve (Neto et al., 2019).

The most used infrared equipment for raw milk analysis today is based on Fourier-transform infrared (**FTIR**) spectroscopy in the mid-infrared range, 649 to 3,999 cm^{-1} (Ho et al., 2021). This type of equipment is used worldwide for daily compositional analyses for millions of raw milk samples, aiming for quality control and inspection in the dairy industry, and dairy herd improvement programs. Association of this technology with machine learning algorithms might be an optimization tool to screen milk compositional data for authenticity (Oliveira et al., 2012; Gondim et al., 2017; Neto et al., 2019; Brito et al., 2020).

Although multivariate analysis has been used to detect milk authentication from cheese whey adulteration (Valente et al., 2014; Vinciguerra et al., 2019), the novelty of our study is underpinned by the use of supervised machine learning methods applied to a large set of real samples of bulk tank raw milk with potential for use as a quality control tool in a milk quality laboratory routine. Artificial neural networks may be classified as a robust nonlinear multivariate analysis technique (Witten et al., 2016; Kubat, 2017).

The objective of this work was to discriminate between raw milk and milk adulterated with cheese whey using machine learning methods applied to FTIR results. This is an innovative screening method with the possibility to optimize analytical speed of raw milk samples with practical implications.

MATERIALS AND METHODS

Ethical approval was waived by the local Ethics Committee of the University (CEP-UFMG) in view of the nature of the study and all the procedures.

Milk and Cheese Whey

The experiment was done in the Laboratory for Milk Quality Analysis (ISO/IEC 17025 accredited), School of Veterinary Medicine, Universidade Federal de Minas Gerais (**UFMG**), Brazil. Five batches of refrigerated raw milk were collected from a refrigerated farm bulk tank, from May to December 2019, in a research farm, with a herd containing about 100 lactating cows with different genetic ratio of Holstein and Gyr cattle. The milk was processed to obtain Minas cheese (a typical Brazilian cheese) by rennet addition (chymosin) and coagulation (Andreatta et al., 2009). Briefly, 10-L
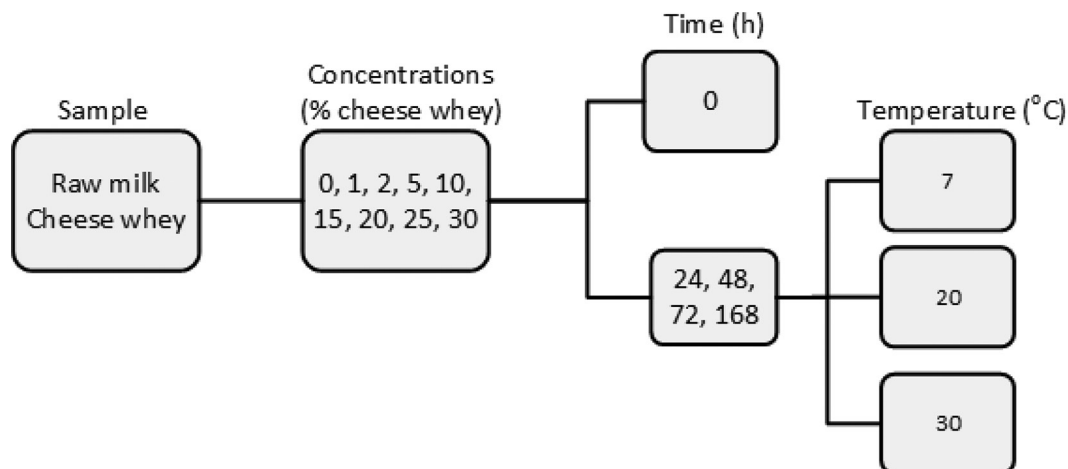
Figure 1. Sample preparation scheme with cheese whey addition to milk (1, 2, 5, 10, 15, 20, 25, and 30%, and control without cheese whey addition), stored at times 0, 24, 48, 72, and 168 h, under 7°C, 20°C, and 30°C.

batches of raw milk were pasteurized (low temperature, long time) at 64°C for 30 min, and after cooling to approximately 35°C, liquid rennet (Ha-la, Christian-Hansen) was added (0.8 mL/L). After coagulation (about 40 min), the gel was cut into cubes with sides of approximately 1.5 cm and stirred for 30 min. At the end, whey was collected and filtered in qualitative filtration paper (11 μm). The resulting cheese whey was heated to 72–75°C for 10 min to denature chymosin, and then refrigerated to 20°C for immediate experimental use. For each repetition, whey was added to raw milk into 50-mL vials at different concentrations (0, 1, 2, 5, 10, 15, 20, 25, and 30%), and added with bronopol as a preservative (Broad Spectrum MicroTabs, 8 mg of bronopol and 0.30 mg of natamycin; Advanced Instruments). After randomization, samples were stored at 7, 20, and 30°C for a period of 0, 24, 48, 72, and 168 h, generating a total of 585 samples (Figure 1). Samples from each treatment were randomly positioned in racks specific for the FTIR equipment. Complementary results of 520 samples of authentic bulk tank raw milk, analyzed in the years 2019 and 2020, were collected from the laboratory server, generating a total of 1,105 samples.

Fourier-Transform Infrared Analyses

Fourier-transform infrared analyses were done in the Laboratory for Milk Quality Analysis, Veterinary School, UFMG, Brazil. This is an ISO 17025 accredited laboratory, which can analyze about 80,000 samples of raw milk per month.

Raw milk and raw milk added with cheese whey were analyzed for composition and freezing point using an FTIR equipment (CombiScope FTIR 400 Delta Instru-

ments) containing a validated multivariate calibration model (partial least squares; Delta Instruments, 2009). Instrument verification was based on standard milk samples (Valacta). Sample results included composition (fat, protein, lactose, TS, SNF, casein, MUN, and freezing point (°C)). The mid-infrared region was used for FTIR measurement (900–3,000 cm^{-1}). Statistical analysis for the analytical FTIR measurement was done according to ISO/IDF (2013) official method. Briefly, the following instrumental and analytical factors were verified for compliance to ISO 9622:2013 (ISO/IDF, 2013) repeatability, reproducibility, zero stability, homogenization, linearity, and carryover. The quality of the analytical procedure was done with control charts (ISO/IDF, 2013).

Machine Learning

Machine learning workflows usually divide data in specific sets: training, validation, and test. The training set consists of samples used to fit the model. During training, the model can split the training set and define a validation set, which is used to provide an unbiased evaluation and guide the algorithm into tuning the model hyperparameters. Finally, the test set consists of samples to which the model is applied (James et al., 2017).

The CART Classification Tree (Minitab 19.2020) was used as the classification method, and the resulting predictive algorithms were applied to the test data set with the objective of classifying the authentic milk from the adulterated one. For binary classification, CART algorithm, the value of each terminal node was the mode of the observations in the corresponding subset, and the prediction accuracy was given by the percent-

age of correctly classified cases. The tree structure was composed by a root node, internal nodes, and terminal nodes. Each internal node divided the instance space into 2 or more spaces according to a discrete function of the input attributes. Each terminal node represented a decision on the target attribute. The included parameters: probabilities matching sample frequencies; a ratio of training, validation, and test sets of approximately 55:25:20, respectively, were randomly split from 1,105 samples; Gini splitting method, with one standard error of minimum misclassification cost.

The procedure for a MLP neural network (MLP;IBM SPSS Modeler 18.2) included the following parameters: Training, validation, and test samples were randomly split at a rate of approximately 55:25:20, randomly chosen by the software algorithm; input layer with the selected features as covariates; algorithm optimization based on scaled conjugate gradient for training; maximum training time of 15 min; training epochs computed automatically. With one hidden layer containing 3 units, excluding the bias unit, the activation function for this layer was the hyperbolic tangent function and the activation function for the output layer was Softmax. Cross-entropy was used as a loss function for optimization of the neural network.

Because the reference CMP index method, using HPLC, detects levels with certainty above 1% of added cheese whey, treatments with low levels of whey addition (1% and below) were additionally tested as nondetectable cheese whey in both methods, CART and MLP (noted as CART1 and MLP1). All features as input included protein, casein, lactose, SNF, TS, fat, freezing point (°C), and MUN (CART all features and MLP all features, respectively). A simpler model with exclusion of fat, TS, freezing point, and MUN was tested due to lower relative importance.

Statistics

Statistical analyses of the compositional data included descriptive and multivariate (SPSS 22.0, IBM; JMP 16.0.0, SAS Institute Inc.). Tukey's test was used for post hoc comparison in the treatments at the significance level of 5% (Dean et al., 2017).

Performance metrics were evaluated based on accuracy, precision, sensitivity, and specificity, where

$$Accuracy = \frac{samples\ correctly\ predicted}{total\ samples},$$

$$Precision = \frac{true\ positive}{(true\ positive + false\ positive)},$$

$$Sensitivity = \frac{true\ positive}{(true\ positive + false\ negative)},$$

$$Specificity = \frac{true\ negative}{(true\ negative + false\ positive)}.$$

The receiver operating characteristics curve will plot the true positive rate, also known as power, on the y-axis, and the false-positive rate, also known as type 1 error, on the x-axis. Hence, in a hypothetical situation when a classification tree can perfectly separate the classes, the area under the curve would be 1. On the other hand, if the tree does not classify better than a random process, the area under the curve would be 0.5 (Neto et al., 2019).

RESULTS AND DISCUSSION

Fourier-Transform Infrared Compositional Results

Compositional FTIR results are shown in Table 1. Cheese whey addition to raw milk resulted in the reduction of components concentration, except for lactose, with increasing concentration correlated with increasing amounts of added cheese whey ($R^2 = 0.60$; $P < 0.001$). No difference was found for freezing point in the different treatments. However, milk components concentration, fat, protein, casein, TS, SNF, lactose, and MUN were affected by cheese whey addition ($P < 0.05$), noticeably dilution effect for fat and protein (Table 1). It is important to observe that the raw milk samples without cheese whey addition were obtained from the bulk tank raw milk samples used for the treatments and from authentic bulk tank raw milk from routine analysis. These findings are expected because a significant amount of the milk solids components will be retained in the curd during renneting. Consequently, whey addition to the milk will result in lower solids concentration due to a dilution effect (Lou and Ng-Kwai-Hang, 1992; Cortez et al., 2010; Condé et al., 2020).

The raw and adulterated milk with cheese whey did not present distinct bands, because their positions overlapped due to the same absorption wavelength number. However, differences were observed regarding the absorption intensity in some of the bands because the intensity of the vibrational modes is proportional to the concentration of the constituents (Figure 2).

It is noteworthy that, for some high levels of cheese whey addition, average concentration for components remained in the legally accepted range for raw milk (Brazil, 2018; Table 2). For example, protein concentration was within the legal Brazilian requirements for all

Table 1. Composition and freezing point (FP; mean and SD) of raw milk and cheese whey added to raw milk analyzed with Fourier-transform infrared spectroscopy

Component (g/100 g)		Cheese whey added to raw milk[1] (% vol/vol)								
		0	1	2	5	10	15	20	25	30
Fat	Mean	3.67[a]	3.59[ab]	3.57[ab]	3.49[ab]	3.33[c]	3.19[d]	3.06[e]	2.90[f]	2.78[g]
	SD	0.26	0.22	0.21	0.23	0.23	0.23	0.22	0.23	0.22
Protein	Mean	3.29[b]	3.42[a]	3.40[a]	3.33[b]	3.22[c]	3.11[d]	3.01[e]	2.90[f]	2.80[g]
	SD	0.11	0.08	0.08	0.08	0.08	0.07	0.08	0.08	0.08
Casein	Mean	2.56[b]	2.67[a]	2.66[a]	2.60[b]	2.50[c]	2.40[d]	2.31[e]	2.22[f]	2.13[g]
	SD	0.11	0.08	0.08	0.08	0.08	0.07	0.08	0.08	0.08
Solids	Mean	12.46[b]	12.65[a]	12.63[a]	12.50[ab]	12.25[c]	12.03[d]	11.85[e]	11.61[f]	11.43[g]
	SD	0.34	0.25	0.24	0.25	0.27	0.28	0.25	0.28	0.27
SNF	Mean	8.92[ef]	9.33[a]	9.33[a]	9.27[ab]	9.17[bc]	9.08[cd]	9.03[de]	8.93[ef]	8.87[f]
	SD	0.32	0.09	0.09	0.09	0.09	0.09	0.10	0.10	0.11
Lactose	Mean	4.52[f]	4.67[e]	4.69[de]	4.71[de]	4.74[cd]	4.77[c]	4.83[b]	4.85[ab]	4.90[a]
	SD	0.14	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.03
MUN	Mean	13.42[a]	11.58[b]	11.84[b]	11.62[b]	11.14[b]	11.38[b]	11.28[b]	10.98[b]	10.70[b]
	SD	3.32	2.14	2.06	2.05	2.17	2.03	1.75	1.89	1.85
FP (°C)	Mean	0.519[a]	0.526[a]	0.527[a]	0.527[a]	0.525[a]	0.524[a]	0.526[a]	0.525[a]	0.525[a]
	SD	0.024	0.003	0.002	0.003	0.003	0.004	0.003	0.005	0.005

[a–g]Means within a row with different superscripts differ using Tukey's test ($P < 0.05$).

[1]65 samples for each treatment.

samples (at least 2.9 g/100 g) even after 15% of cheese whey addition to milk. For treatments with the addition of 20 and 25% of cheese whey, about 80% of the samples remained within legal parameters. Only at 30% of whey addition, the majority of the samples were noncompliant (98.5%) with the legal requirements. Similar trends were found for fat. This is a concerning finding, because even at high levels of adulteration with cheese whey, gross composition might be in the acceptable range for components, impairing routine surveillance to detect these samples.

Similar results were found in another study, with protein average values in the legal range even after the addition of 30% cheese whey to milk, despite showing a tendency to reduce concentration. Lactose concentration, however, showed slight reduction. Fat values were 2.7% after the addition of 15% cheese whey. This will decrease the chances to identify fraudulent addition of cheese whey to milk (Cortez et al., 2010).

In several countries, HPLC based on CMP index is the standard method to detect cheese whey addition to dairy products (Olieman and Bedem, 1983; Olieman and Riel, 1989; Brazil, 2019). However, some reports have demonstrated accuracy and performance problems of this method due to several factors, such as whey acidity and storage conditions. For more reliable results using the CMP method, milk samples should be immediately analyzed or frozen until analysis. If not, proteases from microbial origin may hydrolyze the κ-CN close to the same cleavage point of chymosin, which results in the pseudo-CMP formation and, consequently, overestimation of the cheese whey addition (de Pádua Alves et al., 2018; Raymundo et al., 2018; Lobato et al., 2020).

Hence, an alternative method, not susceptible to such factors, is a major need for the dairy industry.

Because of this potential pseudo-CMP production due to quality problems, some countries establish acceptable limit levels of CMP. For example, Brazilian levels of CMP are up to 30 mg/L for an equivalent of liquid milk (Brazil, 2019). Nevertheless, raw milk CMP levels remain within this legal limit even after 1% of cheese whey addition to raw milk. Based on this fact, additional predictive and classification methods were processed, assigning samples with 1% of cheese whey added to milk as "no whey detected." In fact, after processing the samples for the evaluated predictive methods, the best prediction results were reached with the treatment using 1% of whey added to milk being treated as nondetectable.

Classification and Regression Trees Classification

The CART method was processed with all input features, except milk urea nitrogen. The nodes that are mostly blue indicate a strong proportion of the event level (chance of whey addition to milk), contrasting with the mostly red nodes, which indicate a strong proportion of the nonevent level (chance of authentic milk; Figure 3).

Although CART results from several variables with positive importance, the relative rankings provide information about how many of these variables are needed for a certain application, as the relative importance values from one variable to the next variable can be useful for decision making about which variables to control or monitor. This metric helps us to explain

Table 2. Relative number for samples of raw milk and cheese whey added to raw milk within minimum concentration for selected parameters, analyzed by Fourier-transform infrared spectroscopy

| Component %[1] (g/100 g)	Cheese whey added to raw milk[2] (% vol/vol)								
	0	1	2	5	10	15	20	25	30
Fat ≥3.0%	100	100	100	100	98.5	80.0	56.9	20.0	20.0
Protein ≥2.9%	100	100	100	100	100	100	80.0	78.5	1.5
Lactose ≥4.3%	100	100	100	100	100	100	100	100	100
SNF ≥8.4%	100	100	100	100	100	100	100	100	100
TS ≥11.4%	100	100	100	100	100	100	100	69.2	33.8

[1]Minimum composition according to Brazilian legislation (Brazil, 2018).

[2]65 samples for each treatment.

the predictive power of each feature in the data set. Relative importance values range from 0 to 100%. The more important variable is assigned with a relative importance value of 100%. Low relative variables are not important and automatically eliminated from the tree.

For example, in these data, the most important predictor for this model was lactose concentration with a relative importance of 100% compared with protein which had a relative importance of 60.8%. This means that protein has a relative importance close to half of the lactose in this classification tree. The feature with the lowest relative importance was MUN (18.8%). The misclassification cost for this simulation was 0.027 for the training samples and 0.054 for the test samples.
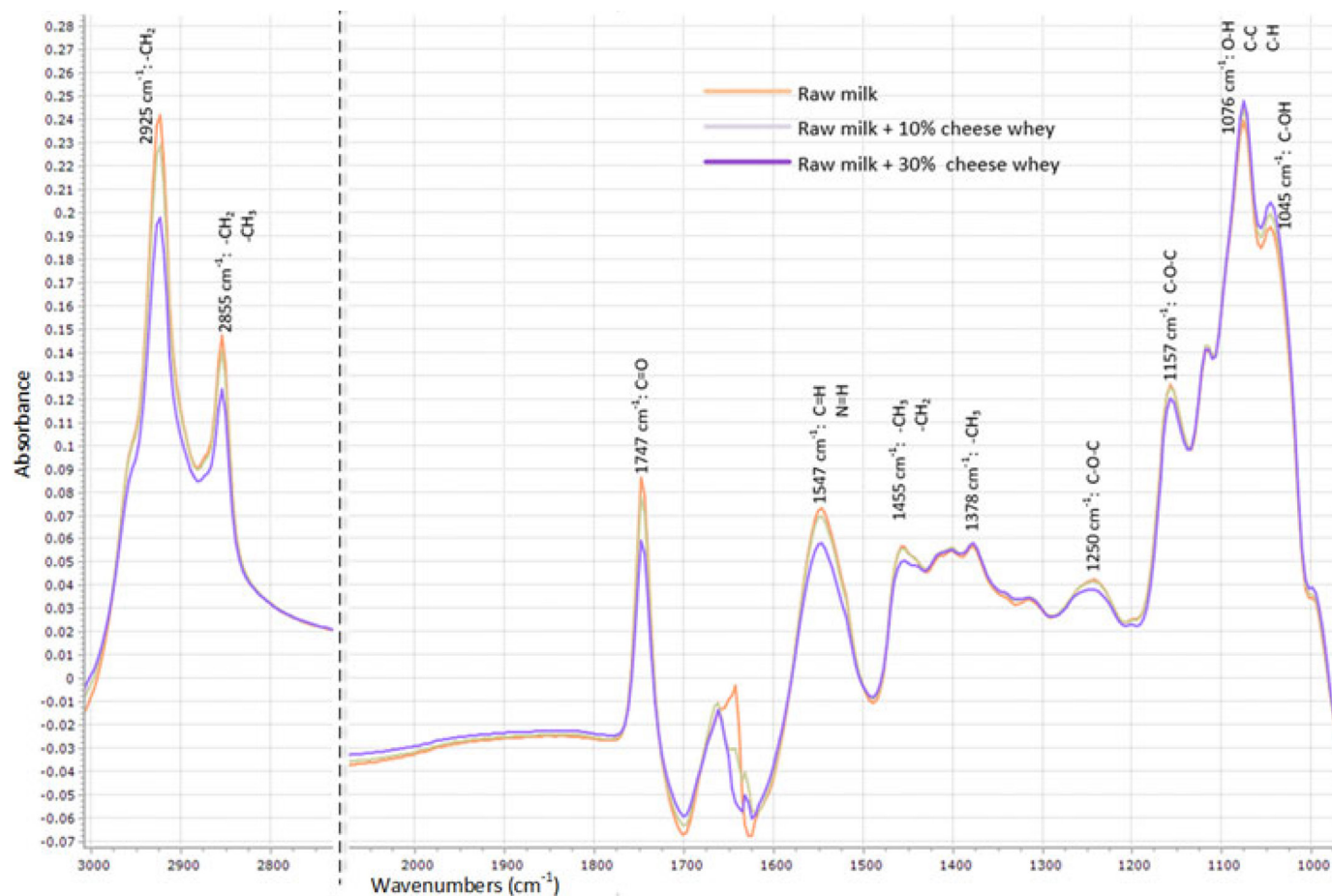
Figure 2. Fourier-transform infrared spectra of authentic milk and milk adulterated with cheese whey (10 and 30%). Main functional groups are pointed out.
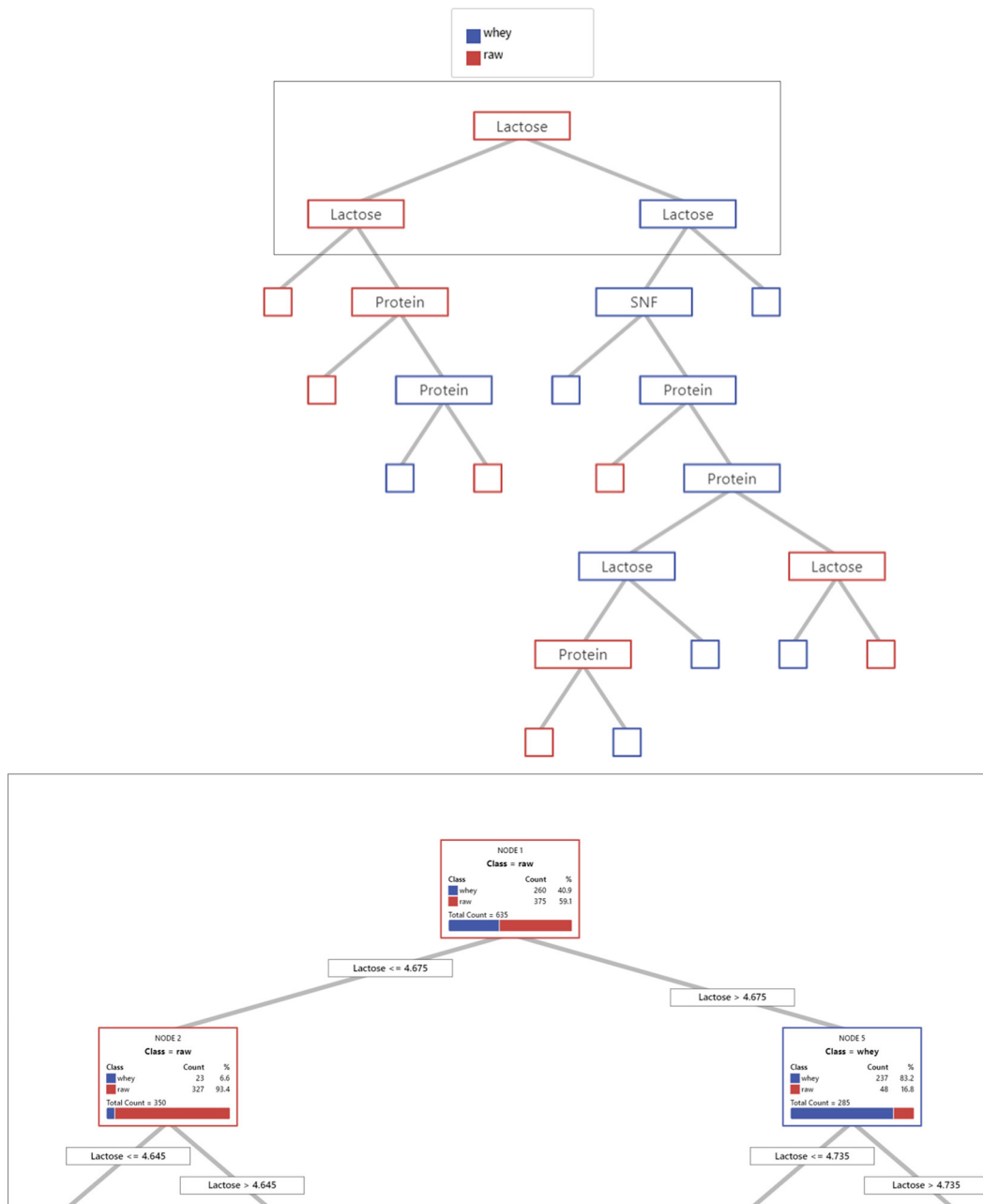
Figure 3. Example of CART (classification and regression trees) classification method for milk and milk with cheese whey added based on composition obtained using Fourier-transform infrared spectroscopy. Node view is complete and part is amplified below. Gray nodes in the node view presented a stronger influence (rectangular part below is amplified from tree above).

The most accurate tree is the one with the lowest misclassification cost. Misclassification may occur due to selection of property which is not suitable for classification (IBM SPSS Modeler 18.2).

The receiver operating characteristics curve is an important visualization tool for the method performance (Figure 4). With an area under the curve of 0.994 for the training and 0.980 for the test data, this receiver operating characteristics curve indicates an optimal classification performance of the model which may be applied for prediction purposes, because the model presents high levels of correct predictions for each class (Dunnmon et al., 2019; Neto et al., 2019).

This can be confirmed using the classification matrix (Table 3) which indicates high rates of sample correct prediction and low rates of misclassification, both in

Table 3. Classification matrix for a CART (classification and regression tree) with fat, protein, casein, lactose, TS, SNF, and freezing point as input features, and binomial output as raw milk and cheese whey added to raw milk (2, 5, 10, 15, 20, 25, and 30%)

| Sample	Observed	Predicted		
		Raw milk	Whey added	Percent correct
Training	Raw milk	365	10	97.3
	Whey added	10	250	96.2
	Overall percent			96.9
Validation	Raw milk	141	12	92.2
	Whey added	3	104	97.2
	Overall percent			94.2
Testing	Raw milk	121	1	99.2
	Whey added	83	5	94.3
	Overall percent			97.1

the training and the test set. Correct predictions to detect tainted samples were as high as 96.2% in the training set and 97.2% in the test set of samples.

The best model was used in the test set of samples (Figure 5). The performance of the best model rendered a test with high-performance, with an accuracy of 0.962, and precision, sensitivity, and specificity as high as 0.965, 0.943, and 0.975, respectively.

Decision tree algorithms similar to CART can provide a better understanding of the whole classification process and also provide meaningful information about each feature, such as feature importance. Other ensemble methods, similar to random forest, would make explaining the algorithm's decisions much more complex. This makes even more sense in our application proposal, which is to be used for milk screening in laboratories, where a simple and explainable solution is desired. The main advantages of decision trees are that they are easy to visualize and interpret, can handle all type of predictors, work well in the case of nonlinear relationship between variables, and make no assumption about the variables distribution (decision tree learning is a nonparametric method). However, some of the disadvantages may include the possibility of overfitting in the training set and a smaller predictive accuracy in the holdout set (test set; Miller and Miller, 2010).

Multilayer Perceptron Networks

The same trends were found using MLP networks, with the best results for the MLP with protein, casein, lactose, SNF, TS, and freezing point as input features and treatment of 1% of cheese whey assigned as "no detectable whey" in the training data set. As in the CART, MUN was eliminated as a feature because it worsened the performance index in the models (Figure

Area Under Curve: Training = 0.9937, Test = 0.9799
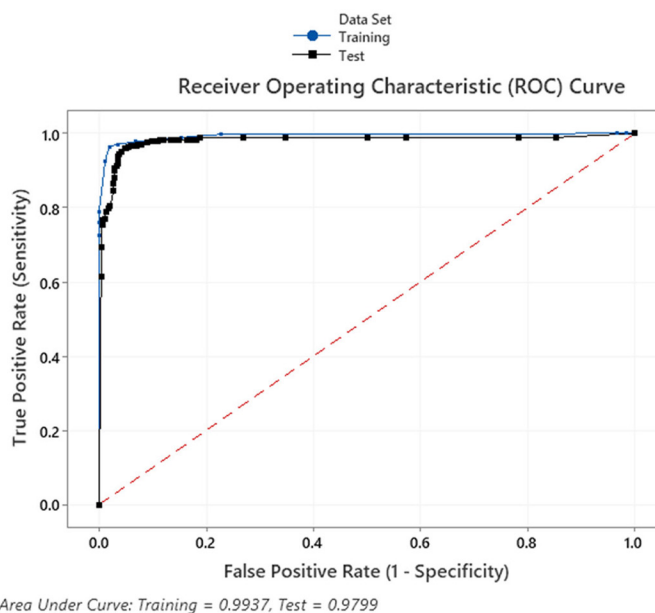
Figure 4. Receiver operating characteristic (ROC) curve for CART (classification and regression trees) of milk and milk with cheese whey added, and features based on composition (g/100 g) of fat, protein, lactose, TS, SNF, and casein, and freezing point (°C), measured with Fourier-transform infrared spectroscopy.

6). The neural network architecture is exemplified with 8 input neurons related to the milk, and milk and cheese whey components, and an additional input neuron for bias. Each neuron from the input layer is connected to each neuron in the second layer (hidden), but they are not interconnected in the same layer. This second layer with a bias node is processed for a final classification as raw milk or raw milk added with cheese whey (Skansi, 2018).

The MLP model created with raw milk added with different levels of cheese whey was validated with real samples analyzed in the laboratory routine.

The training set resulted in 2.6% of incorrect predictions, whereas for the testing set, incorrect predictions were 1.6% as shown in Table 4. These results indicate that this MLP network has an excellent prediction power, with huge potential for testing real samples, as further indicated. To our knowledge, reports of the use of machine learning methods to detect fraudulent addition of cheese whey to milk are scarce. The use of artificial neural networks was reported elsewhere, using the compositional results of routine analyzes in milk samples as input variables. Cheese whey was added to milk at levels of 0, 5, 10, and 20%, and samples were analyzed for fat, SNF, density, protein, lactose, minerals, and freezing point, totaling 164 samples, of which 60% were used for network training, 20% for network validation, and 20% for neural
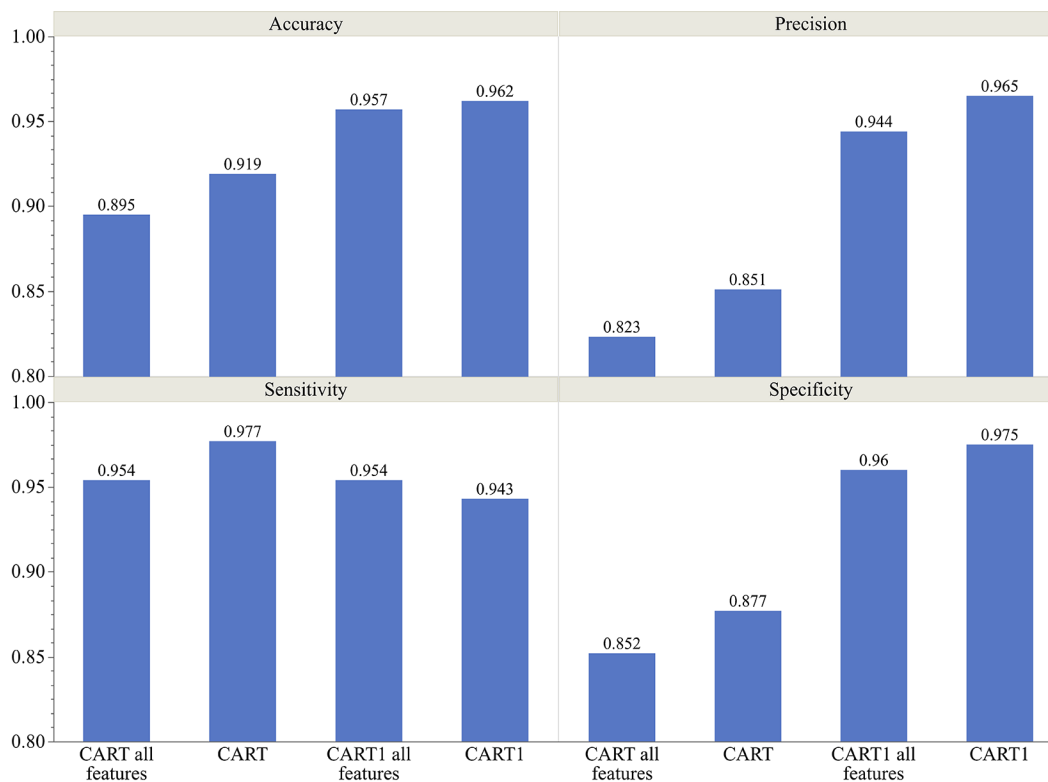
Figure 5. Test performance of CART (classification and regression trees) method to detect cheese whey added to raw milk based on compositional data obtained with Fourier-transform infrared spectroscopy method.

network testing. Although the authors stated that the use of neural networks proved to be efficient, they suggested the use of a more accurate method to confirm the fraud. Despite being a model for quantification, model performance metrics were not presented (Condé et al., 2020).

The implemented models were applied to the same ratios of randomly split data set with similar perfor-

Table 4. Classification matrix for a multilayer perceptron network with protein, casein, lactose, TS, fat composition, and freezing point as input features, and binomial outputs as raw milk and cheese whey added to raw milk (2, 5, 10, 15, 20, 25, and 30%)

| Sample | Observed | Predicted | | |
		Raw milk	Whey added	Percent correct
Training	Raw milk	358	8	97.8
	Whey added	8	249	96.9
	Overall percent			97.4
Validation	Raw milk	153	1	99.4
	Whey added	3	94	96.9
	Overall percent			98.4
Testing	Raw milk	126	4	96.9
	Whey added	1	100	99.0
	Overall percent			97.8

mance outcomes from MLP and CART. The algorithm for the MLP model, considering 1% of added whey to milk as nondetectable was applied to the test set of samples, and total accuracy, precision, sensitivity, and specificity were, respectively, 0.978, 0.96, 0.99, and 0.973, which indicate very good performance for a screening method (Figure 7). It is important to note that current FTIR analytical equipment for raw milk has reached production of up to 600 samples/h. So, the use of a screening method with such performance combined with database processing with this type of MLP algorithm would be an important tool for more strict surveillance of suspected farms and dairy plants, and to reduce the use of more expensive and time-consuming precision methods, such as HPLC.

Although we did not find reports using mid-infrared FTIR associated with neural networks to detect cheese whey addition to milk, other techniques have been studied. For example, radial function and MLP were applied to analytical results of milk and cheese whey added milk obtained using ultrasound analyzer. Classification error was reported as less than 5%; however, sample number was limited (101, 33, and 33 samples for training, validation, and testing, respectively; Valente et al., 2014).
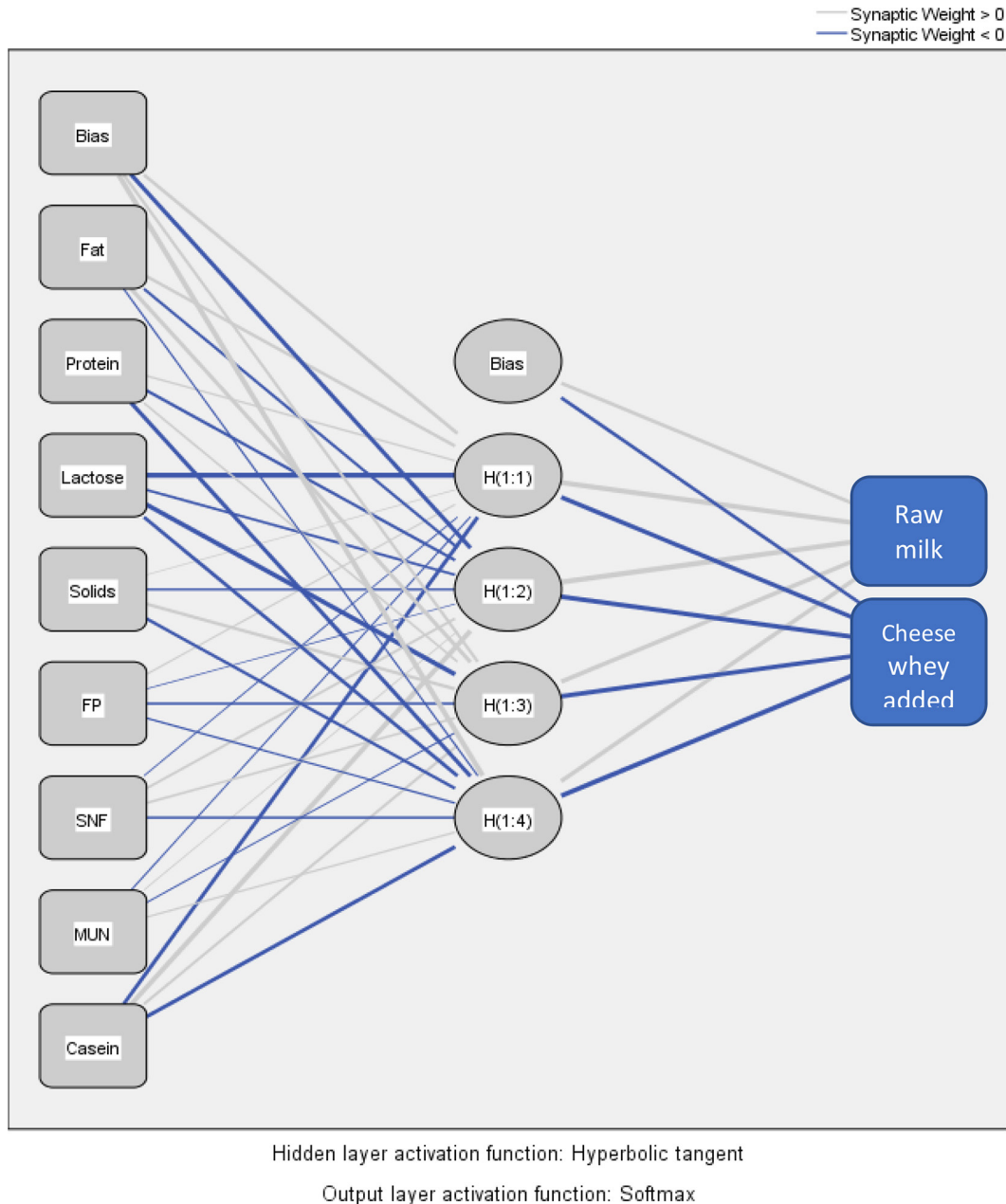
Figure 6. Network diagram for a multilayer perceptron network with fat, protein, casein, lactose, TS, SNF, and freezing point (FP) as input features, and binomial output as raw milk and raw milk with cheese whey added (2, 5, 10, 15, 20, 25, and 30%). H = hidden layer.

None of the previous reported studies, which evaluated the identification of fraud by cheese whey in raw milk by FTIR spectroscopy, used a similar machine learning methodology or obtained results superior to those of this work.

Overall strength of both methods is the use of compositional data, easily obtained through FTIR or other analytical methods, and the optimal performance without additional data preprocessing. However, the raw milk samples represented a specific population of cows with a different genetic ratio of Holstein and Gyr cattle. Hence, different milk origin profiles might require a different structural approach for the evaluated machine learning methods. It is important to note that this work was aimed at bulk raw milk, not individual milk, whose composition is more variable.
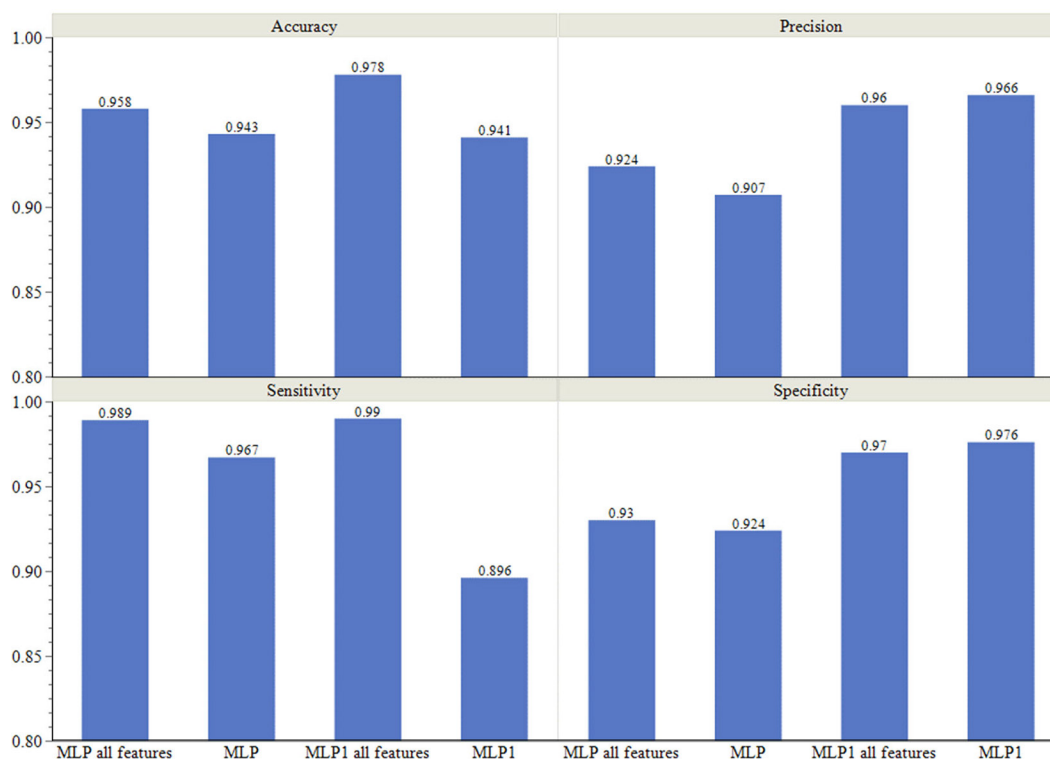
Figure 7. Test performance of multilayer perceptron networks (MLP) to detect cheese whey added to raw milk based on compositional data.

CONCLUSIONS

The CART and MLP network, associated with milk features predicted with FTIR spectroscopy, presented high-performance metrics to detect cheese whey added to raw milk, with high levels of correctly predicted samples and reduced misclassifications. Such performance is practically relevant because it might allow future implementation of both techniques in a laboratory routine for milk quality analysis to screen suspected milk samples, which can be directed for complementary analyses to confirm fraud.

ACKNOWLEDGMENTS

REFERENCES

Alpaydin, E. 2014. Introduction to Machine Learning. 3rd ed. The MIT Press.

Alves da Rocha, R., I. M. Paiva, V. Anjos, M. A. M. Furtado, and M. J. V. Bell. 2015. Quantification of whey in fluid milk using confocal Raman microscopy and artificial neural network. J. Dairy Sci. 98:3559–3567. https://doi.org/10.3168/jds.2014-8548.

Andreatta, E., A. M. Fernandes, M. V. Santos, C. Mussarelli, M. C. Marques, M. L. Gigante, and C. A. F. Oliveira. 2009. Qualidade de queijo minas frescal preparado com leite com diferentes quantidades de células somáticas. Pesqui. Agropecu. Bras. 44:320–326. https://doi.org/10.1590/S0100-204X2009000300014.

Bramer, M. 2016. Principles of Data Mining. 3rd ed. Undergraduate Topics in Computer Science. Springer. https://doi.org/10.1007/978-1-4471-7307-6.

Brandao, M. C. M. P., A. P. Carmo, M. J. V. Bell, and V. C. Anjos. 2010. Characterization of milk by infrared spectroscopy. Rev. Inst. Laticínios Cândido Tostes 65:30–33.

Brazil. 2018. Diário Oficial da União. Vol. 76/2018. Brasília.

Brazil. 2019. Manual de métodos oficiais para análise de alimentos de origem animal. 2nd ed. Rev. ed. Brasília.

Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen. 1984. Classification and Regression Trees. Taylor & Francis.

Brito, R. F., R. Rodrigues, S. A. Diniz, L. M. Fonseca, M. O. Leite, M. R. Souza, R. S. Conrrado, S. A. O. Veríssimo, G. L. S. Valente, and M. M. O. P. Cerqueira. 2020. Analysis of the freezing point of milk by precision method and by Fourier transform infrared (FTIR) spectroscopy. Arq. Bras. Med. Vet. Zootec. 72:1713–1718. https://doi.org/10.1590/1678-4162-11961.

Conceição, D., B.-H. Gonçalves, F. da Hora, A. Faleiro, L. Santos, and S. Ferrão. 2019. Use of FTIR-ATR spectroscopy combined with multivariate analysis as a screening tool to identify adulterants in raw milk. J. Braz. Chem. Soc. 30:780–785. https://doi.org/10.21577/0103-5053.20180208.

Condé, V. A., G. F. S. Valente, and E. C. Minighin. 2020. Milk fraud by the addition of whey using an artificial neural network. Cienc. Rural 50:e20190312. https://doi.org/10.1590/0103-8478cr20190312.

Cortez, M. A. S., V. G. Dias, R. G. Maia, and C. C. A. Costa. 2010. Physicochemical characteristics and sensorial evaluation of pasteurized milk added with water, cheese whey, 0.9% sodium chloride solution and 5.0% dextrose solution. Rev. Inst. Laticínios Cândido Tostes 65:18–25.

de Carvalho, B. M. A., L. M. de Carvalho, J. S. dos Reis Coimbra, L. A. Minim, E. de Souza Barcellos, W. F. da Silva Júnior, E. Detmann, and G. G. P. de Carvalho. 2015. Rapid detection of whey in milk powder samples by spectrophotometric and multivariate calibration. Food Chem. 174:1–7. https://doi.org/10.1016/j.foodchem.2014.11.003.

de La Fuente, M. A., and M. Juárez. 2005. Authenticity assessment of dairy products. Crit. Rev. Food Sci. Nutr. 45:563–585. https://doi.org/10.1080/10408690490478127.

de Pádua Alves, É., A. L. D. A. de Alcântara, A. J. K. Guimarães, E. H. W. de Santana, B. G. Botaro, and R. Fagnani. 2018. Milk adulteration with acidified rennet whey: A limitation for caseinomacropeptide detection by high-performance liquid chromatography. J. Sci. Food Agric. 98:3994–3996. https://doi.org/10.1002/jsfa.8846.

Dean, A., D. Voss, and D. Draguljić. 2017. Design and Analysis of Experiments. 2nd ed. Springer Texts in Statistics. Springer. https://doi.org/10.1007/978-3-319-52250-0.

Delta Instruments. 2009. Datascope with 20 FT components. Delta Instruments.

Dunnmon, J. A., D. Yi, C. P. Langlotz, C. Ré, D. L. Rubin, and M. P. Lungren. 2019. Assessment of convolutional neural networks for automated classification of chest radiographs. Radiology 290:537–544. https://doi.org/10.1148/radiol.2018181422.

Ertel, W. 2017. Introduction to Artificial Intelligence. 2nd ed. Undergraduate Topics in Computer Science. Springer. https://doi.org/10.1007/978-3-319-58487-4.

Ghaffarian, S., M. van der Voort, J. Valente, B. Tekinerdogan, and Y. de Mey. 2022. Machine learning-based farm risk management: A systematic mapping review. Comput. Electron. Agric. 192:106631. https://doi.org/10.1016/j.compag.2021.106631.

Gondim, C. S., R. G. Junqueira, S. V. C. Souza, I. Ruisánchez, and M. P. Callao. 2017. Detection of several common adulterants in raw milk by MID-infrared spectroscopy and one-class and multi-class multivariate strategies. Food Chem. 230:68–75. https://doi.org/10.1016/j.foodchem.2017.03.022.

Guiné, R. P. F., A. C. Ferrão, M. Ferreira, P. Correia, M. Mendes, E. Bartkiene, V. Szűcs, M. Tarcea, M. M. Sarić, M. Černelič-Bizjak, K. Isoldi, A. El-Kenawy, V. Ferreira, D. Klava, M. Korzeniowska, E. Vittadini, M. Leal, L. Frez-Muñoz, M. Papageorgiou, and I. Djekić. 2020. Influence of sociodemographic factors on eating motivations - modelling through artificial neural networks (ANN). Int. J. Food Sci. Nutr. 71:614–627. https://doi.org/10.1080/09637486.2019.1695758.

Handford, C. E., K. Campbell, and C. T. Elliott. 2016. Impacts of milk fraud on food safety and nutrition with special emphasis on developing countries. Compr. Rev. Food Sci. Food Saf. 15:130–142. https://doi.org/10.1111/1541-4337.12181.

Hansen, L., and M. F. Ferrão. 2020. Classification of milk samples using CART. Food Anal. Methods 13:13–20. https://doi.org/10.1007/s12161-019-01493-9.

Ho, P. N., T. D. W. Luke, and J. E. Pryce. 2021. Validation of milk mid-infrared spectroscopy for predicting the metabolic status of lactating dairy cows in Australia. J. Dairy Sci. 104:4467–4477. https://doi.org/10.3168/jds.2020-19603.

ISO/IDF. 2013. ISO 9622/IDF 141, Milk and liquid milk products — Guidelines for the application of mid-infrared spectrometry. 2nd ed. ISO.

James, G., D. Witten, T. Hastie, and R. Tibshirani. 2017. An introduction to statistical learning: With applications in R. Springer. https://doi.org/10.1007/978-1-4614-7138-7.

Koo, T., M. Kim, and M. Jue. 2021. Automated detection of superficial fungal infections from microscopic images through a regional convolutional neural network. PLoS One 16:e0256290. https://doi.org/10.1371/journal.pone.0256290.

Kubat, M. 2017. An Introduction to Machine Learning. 2nd ed. Springer. https://doi.org/10.1007/978-3-319-63913-0.

Lakretz, Y., D. Hupkes, A. Vergallito, M. Marelli, M. Baroni, and S. Dehaene. 2021. Mechanisms for handling nested dependencies in neural-network language models and humans. Cognition 213:104699. https://doi.org/10.1016/j.cognition.2021.104699.

Lenardon, L., L. Z. Meneghini, R. B. Hoff, T. M. C. Motta, T. M. Pizzolato, M. F. Ferrão, and A. M. Bergold. 2017. Determination of caseinomacropeptide in Brazilian bovine milk by high-performance liquid chromatography-mass spectrometry. Anal. Lett. 50:2068–2077. https://doi.org/10.1080/00032719.2016.1263644.

Liakos, K. G., P. Busato, D. Moshou, S. Pearson, and D. Bochtis. 2018. Machine Learning in Agriculture: A Review. Sensors (Basel) 18:2674. https://doi.org/10.3390/s18082674.

Lobato, P. R., J. P. M. Heringer, M. E. R. Fortini, L. F. Ferreira, F. A. C. Feijó, M. O. Leite, M. M. O. P. Cerqueira, C. A. M. Penna, M. R. Souza, and L. M. Fonseca. 2020. Índice de CMP em leite pasteurizado comercializado em Minas Gerais, Brasil, durante os anos de 2011 a 2017. Arq. Bras. Med. Vet. Zootec. 72:641–646. https://doi.org/10.1590/1678-4162-10735.

Lou, Y., and K. F. Ng-Kwai-Hang. 1992. Effects of protein and fat levels in milk on cheese and whey compositions. Food Res. Int. 25:445–451. https://doi.org/10.1016/0963-9969(92)90169-6.

Ma, L., S. Kashanj, S. Xu, J. Zhou, D. S. Nobes, and M. Ye. 2022. Flow reconstruction and prediction based on small particle image velocimetry experimental datasets with convolutional neural networks. Ind. Eng. Chem. Res. 61:8504–8519. https://doi.org/10.1021/acs.iecr.1c04704.

Miller, J. N., and J. C. Miller. 2010. Statistics and Chemometrics for Analytical Chemistry. Pearson Education Limited.

Morota, G., R. V. Ventura, F. F. Silva, M. Koyama, and S. C. Fernando. 2018. Big data analytics and precision agriculture symposium: Machine learning and data mining advance predictive big data analysis in precision animal agriculture. J. Anim. Sci. 96:1540–1550. https://doi.org/10.1093/jas/sky014.

Neto, H. A., W. L. F. Tavares, D. C. S. Z. Ribeiro, R. C. O. Alves, L. M. Fonseca, and S. V. A. Campos. 2019. On the utilization of deep and ensemble learning to detect milk adulteration. BioData Min. 12:13. https://doi.org/10.1186/s13040-019-0200-5.

Olieman, K., and J. Bedem. 1983. A sensitive HPLC method of detecting and estimating rennet whey total solids in skim milk powder. Int. Dairy J. 37:27–36.

Olieman, K., and J. Riel. 1989. Detection of rennet whey solids in skim milk and buttermilk powder with reversed-phase HPLC. Neth. Milk Dairy J. 43:171–184.

Oliveira, M. C. P. P., N. M. A. Silva, L. P. F. Bastos, L. M. Fonseca, M. M. O. P. Cerqueira, M. O. Leite, and R. S. Conrrado. 2012. Fourier transform infrared spectroscopy (FTIR) for MUN analysis in normal and adulterated milk. Arq. Bras. Med. Vet. Zootec. 64:1360–1366. https://doi.org/10.1590/S0102-09352012000500037.

Poonia, A., A. Jha, R. Sharma, H. B. Singh, A. K. Rai, and N. Sharma. 2017. Detection of adulteration in milk: A review. Int. J. Dairy Technol. 70:23–42. https://doi.org/10.1111/1471-0307.12274.

Raymundo, N. K. L., H. Daguer, S. C. Osaki, and L. S. Bersot. 2018. Correlating mesophilic counts to the pseudo-CMP content of raw milk. Arq. Bras. Med. Vet. Zootec. 70:1660–1664. https://doi.org/10.1590/1678-4162-10014.

Robim, M. S., M. A. S. Cortez, A. C. O. Silva, R. A. T. Filho, N. H. Gemal, and E. B. Nogueira. 2012. Research fraud in UHT whole milk marketed in the state of Rio de Janeiro and comparison between the methods of physicochemical officers and the method of ultrasound. Rev. Inst. Laticínios Cândido Tostes 67:43–50. https://doi.org/10.5935/2238-6416.20120077.

Schiassi, E., R. Furfaro, C. Leake, M. De Florio, H. Johnston, and D. Mortari. 2021. Extreme theory of functional connections: A fast physics-informed neural network method for solving ordinary and partial differential equations. Neurocomputing 457:334–356. https://doi.org/10.1016/j.neucom.2021.06.015.

Skansi, S. 2018. Introduction to Deep Learning: From Logical Calculus to Artificial Intelligence. Springer. https://doi.org/10.1007/978-3-319-73004-2.

Tibola, C. S., S. A. da Silva, A. A. Dossa, and D. I. Patrício. 2018. Economically motivated food fraud and adulteration in Brazil:

Incidents and alternatives to minimize occurrence. J. Food Sci. 83:2028–2038. https://doi.org/10.1111/1750-3841.14279.

Valente, G. F. S., D. C. Guimarães, A. L. A. Gaspardi, and L. A. Oliveira. 2014. Applying artificial neural networks as a test to detect milk fraud by whey addition. Rev. Inst. Laticínios Cândido Tostes 69:425–432.

Vinciguerra, L. L., M. C. A. Marcelo, T. M. C. Motta, L. Z. Meneghini, A. M. Bergold, and M. F. Ferrao. 2019. Chemometric tools and FTIR-ATR spectroscopy applied in milk adulterated with cheese whey. Quim. Nova 42:249–254. https://doi.org/10.21577/0100-4042.20170337.

Witten, I., E. Frank, M. Hall, and C. Pal. 2016. Data Mining: Practical Machine Learning Tools and Techniques. 4th ed. Morgan Kaufmann..

Wu, X., V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z.-H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg. 2008. Top 10 algorithms in data mining. Knowl. Inf. Syst. 14:1–37. https://doi.org/10.1007/s10115-007-0114-2.

Xu, X., and Y. Zhang. 2021. Corn cash price forecasting with neural networks. Comput. Electron. Agric. 184:106120. https://doi.org/10.1016/j.compag.2021.106120.

Zhu, X., B. Zheng, W. Cai, J. Zhang, S. Lu, X. Li, L. Xi, and Y. Kong. 2022. Deep learning-based diagnosis models for onychomycosis in dermoscopy. Mycoses 65:466–472. https://doi.org/10.1111/myc.13427.