ABSTRACT
Classification of poultry meat quality relies on inconsistent, time-consuming, and laborious techniques often conducted subjectively. Measurement of attributes such as water holding capacity (WHC) and interpretation of other parameters such as pH and colour features depends on trained personnel and specific equipments for each test. The current work presents a combined attribute selection step and classification step for poultry meat samples. Samples from different quality attributes were analysed to comprise a large set of values for each parameter. REPTree decision trees predictor exploit the optimal parameters for classification tasks of different quality grades of poultry meat. The proposed methodology was conducted with a Support Vector Machine algorithm (SVM) with standard parameters to compare model precision without a pipeline of processes. Experiments were performed on colour (CIEL*a*b*, chroma and hue), water holding capacity (WHC), and pH of each sample analysed. Results show that the best method was a REPTree based on 2 parameters, allowing for classification of poultry samples on quality grades with 98.1% precision.

KEY WORDS
Decision Table, Decision Tree, Quality, M5P, SVM

1 Introduction

Chicken meat processing industry has an important economic role worldwide, as production and consumption of chicken meat has increased significantly in recent years. The search for high quality chicken meat is essential to keep the supply chain and a major demand from consumers.

Traditional analytical methods are still widely used to evaluate meat quality. Parameters such as colour, water holding capacity (WHC) and pH are often used to detect defects in chicken such as pale, soft and exudative (PSE) and dark, firm and dry (DFD), normal (N) or pale (P) samples [1]. However, these techniques are laborious and time-consuming for the accelerated routine of the meat industry [2]. The disadvantages of traditional methods of evaluation performed by human experts, is that these methods are often conducted subjectively and inconsistently [3]. Thus, novel methods for objective and reliable meat quality classification are needed.

Selecting a few essential parameters related to the response information can reduce significantly the amount of data to be analysed and is a major step for development of specialist systems. Implementation of specialist systems as a process analytical technology (PAT) to the food industry involves a multidisciplinary approach where Computational Intelligence (CI), specially Machine Learning (ML) [4, 5, 6, 7, 8] have been investigated. The main advantage of CI is its capacity of handling multiple parameters, facilitating the evaluation in an industrial environment with a faster, more accurate approach with low costs [9]. In addition, it is a sustainable alternative since requires no chemicals that can harm the environment and are hazardous to human beings.

Nowadays, ML techniques have been employed in different food scenarios, mainly for prediction and assessment of food quality [10, 11, 12, 13, 14, 15, 16]. [11] showed a study to predict the total viable counts (TVC) in pork using Support Vector Machines (SVMs), showing the advantage of a rapid and readily performed analysis obtaining a coefficient of correlation of \( r = 0.88 \).

[10] applied a Multilayer Perceptron (MLP) neural network to correlate Fourier transform infrared (FTIR) spectral data with beef spoilage, with good performance of the classifier with 10 neurons in the hidden layer providing a 90.5% overall correct classification.

Also based on SVM, [13] explored beef samples under different packaging conditions by spectroscopy and sensory analysis in order to predict fresh, semi-fresh and spoiled samples. It was reported that the ML techniques
(including ANN) provided better prediction models with 87.50% of accuracy, for the various groups when compared to multivariate statistical methods. [16] presented a study of predicting of beef eating quality parameters, including color, ultimate pH and SSF. This study was based on spectroscopy, and prediction results using SVM on three datasets was over 96%.

However, a limitation of SVM is the selection of the kernel function and its hyper-parameters’ values. SVM are not optimal to identify more than two label (binary classification) data efficiently [17] [18] without problem-related parameters. Another disadvantage is complex training for a very large training set, because each class must be taught against all data from all other classes [19]. This fact exposes an unfeasible use of SVM in the meat industry [11]. The disadvantage of approach parametrization also occurs in ANN architecture, which uses empirical experimentation. Besides, a choice of a kernel demands experience on the ANN types, such as Back-propagation Neural Network (BPNN), Multilayer Perceptron (MLP) or Radial Basis Function (RBF). Hence, both approaches (ANN and SVM) can only handle numerical attributes, which in practical terms demands data pre-processing and normalization.

ANN limitations were reduced by hybrid models that provide means of encoding prior knowledge about a process [20]. The introduction of of Wavenet-Based Neural Network (WNN) as a special type RBFN can overcome some MLP and RBFN disadvantages [21].

Decision Tree (DT) induction algorithms allows to handle the limitation of a high number of classes of SVM, empirical selection and setup of ANN type [18]. The most DT induction algorithms also can handle different types of attributes (numerical/categorical), and even missing values. Additionally, DTs yields accurate multiclass classifiers and improve predictive performance with multiclass problems directly [22].

REPTree is a fast learner DT induction algorithm that builds a decision tree based on information gain as the splitting criterion, and prunes it by reduced-error pruning [23]. REPTree is a useful decision tree method due to be better fit to deal with different scenarios [23]. In [24] was performed a comparison among different DT and other high accurate algorithms (e.g., Random Forest), and REPTree obtained the best performance of DT algorithms. M5P is another algorithm to build induction trees, where leaves are linear models [22]. It is useful for wavelength selection and has advantages compared to traditional prediction algorithms, including the ability to deal with both categorical and continuous variables and to handle variables with missing values [25]. However, M5P is a tree for regression task, due to this is not feasible with our problem.

Thus, the purpose of the current study is to use machine learning approaches to classify chicken meat samples according to quality grades on different parameters obtained such as colour (CIEL*a*b*, chroma and hue), WHC and pH. Considering the limitations of some ML approaches and avoiding the use of a pipeline of process, in this paper we perform the quality assessment based on Decision Tree as in REPTree algorithm. In order to compare the selected approaches with current literature, we use an SVM based on training by Sequential Minimal Optimization (SMO) with a Normalized PolyKernel, enhanced for multi-class scenarios based on feature subsets using best-first search. To evaluate the performance of the supervised ML approaches, the classification of selected wavelengths was based on class labels (PSE, DFD, N and P).

2 Materials and Methods

2.1 Preparation of poultry samples

Slaughtered chicken breast fillets (*pectoralis major muscle*) were selected by an experienced analyst in order to comprise as large variation in quality features as possible. All samples were supplied by local retailer in two batches (*n*total = 158 samples) within 5 h after slaughtering, and transported under refrigerated conditions to the Laboratory of Food Science at State University of Londrina, Londrina-PR, Brazil, for further analysis [26].

2.2 Quality of poultry meat

Chicken quality attributes were measured at 48h post-mortem, since it is within this time that most of the biochemical changes in meat take place [27, 28]. After a 30 min blooming period, ultimate pH values were measured using a Testo 205 (Testo AG, Lenzkirch, Germany); the average of 2 measurements were taken for each sample. Colour features were calculated as the average of 4 consecutive measurements at random locations of samples using a Minolta colorimeter (CR 400, D65 illuminant and 10 observer, Konica-Minolta Sensing Inc., Osaka, Japan) after calibration with a standard ceramic tile. Colour was expressed in terms of values for lightness (*L*+), redness (*a*+), and yellowness (*b*+) using the Comission Internationale de Léclairage (CIE) colour system [29, 30]. Water holding capacity (WHC) was calculated based on meat water loss when pressure is applied on the muscle [31]. Cubic pieces of samples weighing 2g were laid between two filter paper circles placed on acrylic plates, on which a 10 kg weight was put for 5min, with the samples then removed from the filter papers, and weighted. Water loss was calculated as the difference between initial and final weight. Results were expressed as the percentage of drip loss relative to initial sample weight.

All samples were pre-classified into four different quality grades based on colour reflectance, ultimate pH and water holding capacity, namely P (*L*+ > 53, *pH* > 5.8) PSE (*L*+ > 53, *pH* < 5.8), DFD (*L*+ < 46, *pH* > 6.1), and normal (46 < *L*+ < 53, *pH* > 5.8). Values are based on information adapted from [32], with samples within these ranges representing the four classes. Thus, we labeled each sample to perform a supervised classification.
2.3 Classifier evaluation

This work involved two major tasks for comparison of several techniques: wavelength selection and quality classification. Wavelength selection was compared by Correlation Coefficient (CC), Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The first illustrates a quantitative value of the relationship between the subset of wavelengths, $W$, and the most significant traditional attributes found, $C_a$. MAE measures the average magnitude of the relationship between $C_a$ and $W$. Finally, RMSE is a quadratic scoring of the average magnitude of the error. In other words, the greater difference between MAE and RMSE illustrates a greater variance in individual errors of samples. A similar value of RMSE and MAE means that all the errors present the same magnitude.

The accuracy of the model for quality classification was assessed by confusion matrices as final output from 10-fold Cross-Validation (CV) over 30 repetitions. This was adopted because it was used a four-class dataset (namely PSE, DFD, P and N) without overlapping. Confusion matrices represent an appropriate evaluation criteria to select the most suitable classifiers (REPTree and M5P).

True Positive (TP) and True Negative (TN) means that a sample was correctly recognized, it is the desired result. False Positive (FP) and False Negative (FN) occurs when the ML approach, incorrectly, classify a sample assigning a wrong class.

The average per-class effectiveness of a classifier was measured by Average Accuracy. The average per-class agreement of the data class label with those of classifiers by Precision and the average per-class effectiveness of a classifier to identify class labels was estimated by Recall. $F – Measure$ is the harmonic mean of Precision and Recall and allow us to determine if one algorithm is superior to another as a particular goal.

In order to compare statistically our results, we applied the Friedman’s statistical test [33] with significance level at $\alpha = 0.05$. The null hypothesis here states that the performances of the induced classifiers is equivalent in terms of the averaged accuracy per class. Any time the null hypothesis was rejected, the Nemenyi post-hoc test was applied, stating that the performance of two different algorithms is significantly different whether the corresponding average ranks differ by at least a Critical Difference (CD) value. When multiple classifiers are compared in this way, the results can be represented graphically with a simple Critical Difference (CD) diagram.

3 Results and discussion

3.1 Parameter selection

The poultry dataset consisted of 158 samples, where there were 24, 86, 41 and 7 samples of the respective classes PSE, P, N or DFD. It was applied CFS in order to find the best subset of attributes based on subset merit. The selected attributes were $A_s = \{pH, L^*\}$ with $Merit_s = 0.816$.

The same approach was performed on the balanced dataset to corroborate the subset found. The balanced dataset was created using a technique of resampling that removes the sample of maximal classes, in our case PSE (24), P (86) and N (41). After resampling, the balanced dataset was composed by 12, 7, 6 and 6 samples of PSE, P, N and DFD, respectively. The selected attributes were $pH$ and $L^*$ that found $B_m = \{pH, L^*\}$ with $Merit_s = 0.870$. The $pH$ values in the database were between 5.62 and 6.43, a mean of 5.95 having a standard deviation of 0.15. The luminosity value $L^*$ were in the range between 43.53 and 67.53, a mean of 57.32 and a standard deviation about 6.43.

However, even M5P composed by a minor number of nodes, achieves superior correlation results, justified by the use of a Linear Model to identify a class in the leaf node.

CFS evaluation was based on Merit and was not compared to the others, which were evaluated by correlation coefficient and errors rates. Independent of measurement, the higher correlation (or merit) was obtained by selected wavelengths of $L^*$.

3.2 Poultry quality classification

It was observed that DFD and PSE for SVM, without correct classification, were the minority class. There were much less DFD samples compared to P samples with 86 samples. To investigate the balancing effect over SVM and REPTree we performed the same experiments on the balanced dataset and the results were the same. SVM algorithm is limited to handle a multi-class scenario, which is better for binary problems.

To support the hypothesis of high dimensionality influence on classification accuracy, it was compared SVM and REPTree. The results of SVM applied to whole wavelengths is available on Table 1. It was observed the same behavior of limitation on multi-class scenarios (not recognition of PSE and DFD) and lower results in SVM precision (0.621).

REPTree was applied on traditional parameters, which obtained excellent results on classification task. It was the higher value of TP and the lower value of FP (desirable) with an F-Measure of 0.981. One important observation that emerges from results exposed in Table 1 is the results of REPTree based on traditional attributes and only $pH$ and $L^*$. The results are the same, once REPTree in both cases selects only $pH$ and $L^*$ to construct the regression three. In other others, the REPTree algorithm identity $pH$ and $L^*$ as sufficient to solve the classification problem. It is possible to see in Table 1 that SVM algorithm over traditional parameters shows the same behavior of previous experiments, which could not handle the multi-class problem with robustness obtaining 0.621 of average precision.

The step of attribute selection contributed to SVM prediction, diminishing dimensionality of the problem and mainly to build a classification model with reduced overfit.
Table 1. Traditional attributes and poultry classification results (pale, soft and exudative (PSE), dark, firm and dry (DFD), normal (N) or pale (P)) obtained by Support Vector Machine (SVM) and REPTree in mean of True Positive (TP), False Positive (FP), Precision, Recall and F-Measure.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Attributes</th>
<th>Class</th>
<th>TP</th>
<th>FP</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>[pH, L* , a*, b*, a*/b*, WHC, texture, chroma, hue]</td>
<td>PSE</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P</td>
<td>0.965</td>
<td>0.361</td>
<td>0.761</td>
<td>0.965</td>
<td>0.851</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>0.951</td>
<td>0.085</td>
<td>0.796</td>
<td>0.951</td>
<td>0.867</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DFD</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Weighted AVG</td>
<td></td>
<td>0.772</td>
<td>0.219</td>
<td>0.621</td>
<td>0.772</td>
<td>0.688</td>
</tr>
<tr>
<td>SVM</td>
<td>[pH, L*]</td>
<td>PSE</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P</td>
<td>0.965</td>
<td>0.347</td>
<td>0.769</td>
<td>0.965</td>
<td>0.856</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>0.976</td>
<td>0.085</td>
<td>0.8</td>
<td>0.976</td>
<td>0.879</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DFD</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Weighted AVG</td>
<td></td>
<td>0.778</td>
<td>0.211</td>
<td>0.626</td>
<td>0.778</td>
<td>0.694</td>
</tr>
<tr>
<td>REPTree</td>
<td>[pH, L* , a*, b*, a*/b*, WHC, texture, chroma, hue]</td>
<td>PSE</td>
<td>0.958</td>
<td>0.007</td>
<td>0.958</td>
<td>0.958</td>
<td>0.960</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P</td>
<td>0.988</td>
<td>0.028</td>
<td>0.977</td>
<td>0.988</td>
<td>0.968</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>0.976</td>
<td>0.000</td>
<td>1.000</td>
<td>0.976</td>
<td>0.986</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DFD</td>
<td>1.000</td>
<td>0.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Weighted AVG</td>
<td></td>
<td>0.981</td>
<td>0.016</td>
<td>0.981</td>
<td>0.981</td>
<td>0.981</td>
</tr>
<tr>
<td>REPTree</td>
<td>[pH, L*]</td>
<td>PSE</td>
<td>0.958</td>
<td>0.007</td>
<td>0.958</td>
<td>0.958</td>
<td>0.960</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P</td>
<td>0.988</td>
<td>0.028</td>
<td>0.977</td>
<td>0.988</td>
<td>0.968</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>0.976</td>
<td>0.000</td>
<td>1.000</td>
<td>0.976</td>
<td>0.986</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DFD</td>
<td>1.000</td>
<td>0.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Weighted AVG</td>
<td></td>
<td>0.981</td>
<td>0.016</td>
<td>0.981</td>
<td>0.981</td>
<td>0.981</td>
</tr>
</tbody>
</table>

Figure 1. Decision Tree for poultry classification created by REPTree algorithm.

An advantage of parameter selection is the complexity reduction of decision tree generated. In Figure 1 it is possible to see a 2 – levels DT with two nodes, a simpler solution obtained by REPTree algorithm after the training step. The complexity can be exemplified by the number of paths to reach the leaves of the tree. Based on DT presented in Figure 1 we can classify PSE, DFD and N samples follow just one path for each class, just checking two parameters.

Figure 2 shows the results of the statistical tests over the experiments. Two different methods are connected when there is no statistically significance difference between them (at $\alpha = 0.05$ and CD = 2.69). Lower rank values indicate the most accurate method, while the higher the less one.

Thus, according to the Figure 2, statistically differences were not found when comparing REPTree method against SVM. The SVM with all traditional parameters has the worst rank value over all the methods, while the REPTree (with $pH$ and $L^*$) was the best one. Another important observation is about PSE and DFD samples. SVM presented problems to handle PSE and DFD class. Thus, a plausible choice is the REPTree model that achieved the same performance (Table 1) in experiments with different parameters, once the algorithm performs attribute selection. Therefore, the results based on: all parameters or $pH$ and $L^*$ is the same, the model is shown in Figure 1.

This suggests that the DTable method is the best approach, but this algorithm cannot handle PSE samples.

Based on the methodology applied in this work, we can affirm that ML approaches could have substantial impact on quality data analysis. Differently of traditional work based on SVM and ANN, which require several parametrization and produce a black box output model, our results showed the potential of Decision Table and Decision Tree application. These ML algorithms are alternative...
Figure 2. Comparison of the Averaged Accuracies per class values of the classifiers according to the Nemenyi test. Groups of classifiers that are not significantly different (at $\alpha = 0.05$) are connected.

approaches to performing a multivariate spectral subset selection and offering an opportunity to construct an accurate and visual prediction models.

4 Conclusion

Poultry meat quality assessment is possible based on a subset of traditional parameters. Based only on two parameters it is possible to classify with 98.1% precision considering just pH and $L^*$. This fact allow the use of equipments, just observing $L^*$ colour component by colourimeter and pH by pHmeter.

Another important consideration touches the traditional SVM and ANN approaches. It was proposed the use of Decision Trees induction algorithms to avoid complex configurations or necessity of expertise on a particular technique. In the reported study, it was handled a multi-class scenario composed by DFD, P, N and PSE poultry meat assessments. Based on the literature, it was applied SVM with dimension reduced by CFS to create a reduced set of parameters and perform the classification. It was observed some limitations of SVM on multi-class scenarios. DT models (REPTree) obtained superior results of performance and achieved a comprehensive and intuitive model, that describes the optimal parameters, respective to values and classes by a Tree.

References


