Computers and Electronics in Agriculture 127 (2016) 368–375

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Storage time prediction of pork by Computational Intelligence

Ana Paula A.C. Barbon [a,*], Sylvio Barbon Jr. [b,1], Rafael Gomes Mantovani [c], Estefânia Mayumi Fuzyi [b], Louise Manha Peres [a], Ana Maria Bridi [a]

[a] Department of Zootechnology, Londrina State University (UEL), Londrina 86057-970, Brazil
[b] Department of Computer Science, Londrina State University (UEL), Londrina 86057-970, Brazil
[c] Sciences Institute of Mathematics and Computers (ICMC), University of São Paulo (USP), São Carlos, Brazil

ARTICLE INFO

Article history:
Received 3 February 2016
Received in revised form 6 June 2016
Accepted 23 June 2016
Available online 4 July 2016

Keywords:
Fuzzy rule based system
Machine learning
Meat quality
Classification
Post mortem

ABSTRACT

In this paper, a storage time prediction of pork using Computational Intelligence (CI) model was reported. We investigated a solution based on traditional pork assessment towards a low time-cost parameters acquisition and high accurate CI models by selection of appropriate parameters. The models investigated were built by J48, Naïve Bayes (NB), k-NN, Random Forest (RF), SVM, MLP and Fuzzy approaches. CI input were traditional quality parameters, including pH, water holding capacity (WHC), color and lipid oxidation extracted from 250 samples of 0, 7 and 14 days of post mortem. Five parameters (pH, WHC, L^*, a^* and b^*) were found superior results to determine the storage time and corroborate with identification in minutes. Results showed RF (94.41%), 3-NN (93.57%), Fuzzy Chi (93.23%), Fuzzy W (92.35%), MLP (88.35%), J48 (83.64%), SVM (82.03%) and NB (78.26%) were modeled by the five parameters. One important observation is about the ease of 0-day identification, followed by 14-day and 7-day independently of CI approach. Result of this paper offers the potential of CI for implementation in real scenarios, inclusive for fraud detection and pork quality assessment based on a non-destructive, fast, accurate analysis of the storage time.

1. Introduction

The parameters involving meat quality are of most importance for the meat processing industry. Research projects are often developed to assess improvements in measurements and quality assessments, as well as factors that influence pork quality such as environmental conditions, pre-slaughter management and purchasing decisions of consumers (Rosenvold and Andersen, 2003). A perspective of pork quality evaluation is based on the muscle to meat conversion. This process includes several enzymatic and protein denaturation processes that directly influence pH and other quality attributes (Salmi et al., 2012). These parameters are play a major role on quality and are related to post mortem period, mainly because the rate of glycolysis, affecting the technological quality of meat (Hammelman et al., 2003).

Determination of post mortem period is relevant for the industry because it allows identification of aging period and freshness evaluation, as well as indicating the consumer preferences. Another advantage is to identify potential fraud during the food storage period. Besides, it can indicate storage problems as temperature deviation in cold rooms, freezers, and refrigerators that can lead to meat deterioration and shelf-life reduction.

Nowadays, consumers are more demanding for food quality, as they are looking for clear and reliable information about product origin, production method and food preservation (Sentandreu and Sentandreu, 2014).

Fraud in the meat sector is constantly described and (Ballin, 2010) describes that fraud can be categorized according to the possibility of occurrence: origin of meat, meat replacement, and meat processing. Moreover, within each of these frauds there are subcategories: post mortem period, meat cuts, animal breed, meat freshness, among others.

However, during post mortem period, some meat quality parameters may be modified, e.g. pH, Water Holding Capacity (WHC), color and lipid oxidation (Tarsitano et al., 2013). The Meat freshness determines the choice of the product by the consumer (Xiong et al., 2015). Moreover, this assessment is also measured by quality parameters mentioned before, and depends directly on the storage time. Nonetheless, laboratory evaluation parameters are costly, time consuming, dependent on trained persons and subjective evaluation. In this context arise alternative methods and non-destructive analysis of food using computational tools (Chen et al., 2011).

[1] http://www.uel.br/grupo-pesquisa/remid/.

* Corresponding author.
E-mail addresses: apbarbon@gmail.com (A.P.A.C. Barbon), barbon@uel.br (S. Barbon Jr.), rgmantov@icmc.usp.br (R.G. Mantovani), emfuzyi@uel.br (E.M. Fuzyi), louise_mp@zootecnista.com.br (L.M. Peres), ambridi@uel.br (A.M. Bridi).

http://dx.doi.org/10.1016/j.compag.2016.06.028
0168-1699/

The use of Computational Intelligence (CI) for food quality classification has been widely discussed (Kodogiannis and Alshejari, 2014; Shan et al., 2015; Przybylak et al., 2015; Ravikanth et al., 2015; Zapotoczny et al., 2016). The main advantage of CI is the capacity of handling multiple parameters, facilitating the evaluation in an industrial environment; being faster; more accurate; not requiring reagents that could damage the environment and having low costs (Qiao et al., 2007).

Among the various techniques of CI for assessing the food quality, some standout, e.g. Clustering Algorithms (CA), Support Vector Machine (SVM), k-Nearest Neighbors (k-NN), Fuzzy Rule-Based Systems, Multilayer Perceptron Neural Network (MLP) and others (Liu et al., 2013).

Some of these techniques have been applied to several kinds of food quality research, mainly on prediction and classification tasks. Often combined or not, CI techniques have been widespread for evaluating spoilage, defects, fraud or predict the most important quality parameters to assess the meat (Kamruzzaman et al., 2013; Argyri et al., 2013; Liu et al., 2014; Ropodi et al., 2015).

The aim of this paper was to classify the pork storage time through the use of CI techniques (J48, Naïve Bayes, k-NN, Random Forest, SVM, MLP and Fuzzy approaches) and to investigate the most appropriated method to perform the meat quality classification by traditional parameters (pH, color, WHC and TBARS). Secondly, it was observed which quality parameters were relevant and capable to provide the identification of storage time.

This paper is organized as follow: Section 2 presents the Materials and Methods of this work, followed by Section 3 which presents experiments results and discussions they promote. Conclusions and closing remarks can be found in Section 4.

2. Materials and methods

2.1. Data samples and analytical measurements

The samples used in our experiments were obtained from pigs slaughtered in Federal Inspection and transported to the Food Analysis Laboratory (LANA) at UEL, Londrina, Brazil for further analysis. Data samples were about two hundred and fifty (250) of *longissimus dorsi et thoracis* muscle from different animals. The samples were vacuum packed and stored at 1 °C for periods of 0, 7 and 14 days.Considering the quality parameters, the pH, color (CIELab), WHC and lipid oxidation (TBARS) value were chosen to describe the samples.

The pH was measured 24 h after cooling (ultimate pH) with insertion electrodes into the meat sample using TESTO 205 pH meter (TESTO, Hampshire, UK).

After a 30 min blooming period, the color was obtained as the average of 3 consecutive measurements at random locations of samples using the Colorimeter (Konica Minolta Color reader CR10) calibrated against a standard white tile. The color was expressed in terms of values for lightness (L^*), redness (a^*), and yellowness (b^*) using the Commission Internationale de lEclairage (CIE) color system (de lEclairage, 1978; Honikel, 1998).

The evaluation of WHC was measured by water pressure loss technique according to described by Barbut (1996). In this analysis, 2 g of sample was weighed in a semi-analytical balance. This sample was placed between two paper filters and two acrylic plates and then applied a weight of about 10 kg for 5 min under the sample. After the pressing time, the weight of the sample was checked again to calculate losses.

The lipid oxidation were analyzed by the methodology indicative of thiobarbituric acid reactive substances (TBARS) described by Pikul et al. (1989).

Statistical information of traditional parameters for the Dataset is exhibited in Table 1. The Dataset was composed by 208 pale, firm, and non-exudative (PFN) samples and 42 red, firm, and non-exudative (RFN), classified based on Faucitano et al. (2010).

Table 1
Statistical summary of entire Dataset (storage times 0, 7 and 14) composed by 250 samples.

Parameter	Mean	St. Dev.	Min	Max
L	52.292	2.252	47.000	56.600
a^*	5.238	1.186	2.700	8.800
b^*	13.191	1.511	10.000	17.900
WHC	26.740	3.126	20.020	34.520
pH	5.716	0.137	5.350	6.020
TBARS	0.446	0.122	0.109	1.136

In this work, we employed statistical tests and information theory-based aiming the same. Pearson's statistical correlation was applied to depicts linear relationships, and Spearman's correlation was computed to discovering monotonic relationships, both between storage time and each parameter. Considering information's theory approaches we computed: Information Gain (*IG*), Gain Ratio (*GR*) and Symmetrical Uncertainty (*SU*), all based on Shannon Entropy (*H*) are showed in Table 2, where *P* is the parameter, *A* is the storage time and *p* the probability of *P* showing up in a specific storage time. Chi^2 was applied to test the independence between the storage time and quality parameters. Correlation Feature Selection (CFS) was used to obtain the best subset based on individual attributes using the symmetrical uncertainty based on subset merit (Yu and Liu, 2004). Subset merit, $Merit_s$, is calculated as in Table 2, where *s* is the subset composed by *k* features, r_{cf} is correlation between feature and class based on entropy, and r_{ff} is the inter-correlation between features.

2.2. Pork quality parameters

2.2.1. pH

The evaluation of pH value in fresh meat is one of the most important parameters to measure the meat quality and in many situations influence on other parameters such as WHC, color and shelf life. The rate of pH decline will dictate the final quality features during the post mortem period, and may modify the storage time meat (Liao et al., 2012).

Holmer et al. (2009) reported regression equations to predict the shelf life of pork in a certain pH range for 28 days. In this study, researchers assert that the pH might interfere with the shelf life after post mortem time and that regression equations could predict that higher pH with longer days of storage of meat had shorter shelf life.

2.2.2. CIELab color system

Color is an important attribute that relates to the first consumer perceptions about the meat quality at the time of purchase of the product (Chmiel et al., 2011).The color of the meat comes from main factors, myoglobin, hemoglobin and cytochrome C (Mancini and Hunt, 2005). In general, the color is measured objectively using a colorimeter device and the most common evaluation system is

Table 2
Entropy-based metrics and Merit for parameters evaluation.

Metric	Equation
Entropy	$H(P) = -\sum_{x=P} p(x) \log p(x)$
Information Gain	$IG(P,A) = H(A) + H(P) - H(A\|P)$
Gain Ratio	$GR(P,A) = \frac{H(A)+H(P)-H(A\|P)}{H(P)}$
Symmetrical Uncertainty	$SU(P,A) = \frac{H(A)+H(P)-H(A\|P)}{H(P)+H(A)}$
Merit	$Merit_s = \frac{k(r_{cf})}{\sqrt{k+k(k-1)(r_{ff})}}$

the CIELab system recommended by International Commission on Illumination (CIE) (CIE 1976). It is widely used because it is reliable and evaluates the brightness (L) in the range of gray tones, ranging from dark to light. The coordinate a^* ranging from red to green and the coordinate b^* varies from yellow to blue (Rodriguez-Pulido et al., 2013).

2.2.3. Water holding capacity

Water is the main constituent of the food and it is of extreme importance to the muscle fibers of the meat, assisting in lubrication and transport of metabolites within the fibers (Puolanne and Halonen, 2010).

In general, the muscle tissue contains approximately 75% water and the other components are proteins, lipids, minerals and vitamins. Most of this water present in muscle (85%) is present in the intracellular and the rest extracellular medium. The intracellular water can be linked to proteins muscle, immobilized or free (Pearce et al., 2011).

The mobility of such water within the muscle can influence some quality parameters such as tenderness, juiciness, and appearance of meat. It is also indicative of cell membrane integrity and early protein denaturation.Thus, knowledge about the muscle WHC is necessary, mainly aspects related to the conservation, llleting, cooking, and processing of meat (Prevolnik et al., 2010; Tejerina et al., 2012).

2.2.4. Thiobarbituric acid reactive substances

The lipid oxidation is one of the leading causes of loss of freshness and meat deterioration. The assessment of 2-thiobarbituric acid reactive substances (TBARS) is an important oxidative parameter that may reflect the oxidation degree of lipids (Xiong et al., 2015). It measures the amount of malondialdehyde; a product of degradation of peroxides formed during the oxidation of polyunsaturated fatty acids.

2.3. Classification approaches

A classification problem consists of taking an input vector with data and deciding which of *N* classes they belong to. It follows a supervised learning process based on training from exemplars of each class (Marsland, 2009). The most important learning feature is the *generalization*: the algorithm should produce sensible output for inputs that were not encountered during learning.

A classification task is discrete, each example belongs to one class, and the set of classes covers the whole possible output space. These constraints are not necessary realistic, sometimes examples might belong partially to two different classes (and fuzzy classifiers may be used to solve this problem). The methods of performing classification vary in the ways they learn about the solution, but in essence they aim to do the same thing: find *decision boundaries* that can be used to separate out the different classes. The machine learning classification algorithms used in experiments are described below. All of them were implemented in the `R` environment[2]. The corresponding packages used to implement each classification algorithm are summarized in Table 3.

Table 3
ML classification algorithms used in experiments and corresponding R packages.

Algorithm	Id	R package
Decision Tree	J48	`RWeka`
Naïve Bayes	NB	`e1071`
k-Nearest Neighbors	k-NN	`kknn`
Random Forest	RF	`randomForest`
Support Vector Machine	SVM	`e1071`
Multilayer Perceptron	MLP	`RWeka`
Fuzzy-RBS (Chi)	Fuzzy-Chi	`frbs`
Fuzzy-RBS (Ishibuchi)	Fuzzy-W	`frbs`

2.3.1. J48 decision tree

Decision Trees (DT) algorithms are one of the most common and powerful methods in ML field (Marsland, 2009). They are every popular and easy to understand. Following a tree to get an answer is transparent, in contrast with many 'black-box' solutions, such as Neural Networks. The J48 algorithm is the C4.5 Quinlan's DT algorithm (Quinlan, 1993) optimized implementation in JAVA. This algorithm explores the concept of information gain of each attribute, separating data with orthogonal planes in the output space. The main idea is to decompose a complex problem in simple sub-problems, applying the same criteria recursively in sub-problems until a stopping criteria is satisfied.

2.3.2. Naïve Bayes

In ML field, Naïve Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features (Mitchell, 1997). NB has been studied extensively since the 1950s, introduced into the text retrieval community, and remains a popular baseline method for text categorization and data classification.

2.3.3. k-Nearest Neighbors

The *k*-Nearest Neighbor algorithm (*k*-NN) is the most basic instance-based method (Mitchell, 1997). This algorithm assumes all instances corresponds to points in the *n*-dimensional space. The *nearest neighbors* of an instance are defined in terms of the standard Euclidean distance, but some adaptations can be made to deal with other distance functions. The distance-weighted *k*-NN is a highly effective inductive method for many practical problems, because it can smooth out the impact of isolated noisy training examples. One practical issue in applying *k*-NN is that the distance between instances is calculated based on all attributes of the instance. This lies in contrast to methods such as rule and DT learning algorithms that select only a subset of features when forming the hypothesis (Maimon and Rokach, 2010).

2.3.4. Random forest

The Random Forest (RF) method was proposed by Breiman (2001) combining DT classifiers in an ensemble. This classifiers is computationally efficient only when each tree is built independently from each other. Further, the trees must be built with maximum deep, without pruning. The number of attributes in each node and the number of tree are hyper-parameter defined by user. This classifier is based on Bagging approach (Breiman, 1996), thus, each new tree uses a totally different training set composed by random sampling instances.

2.3.5. Support Vector Machines

The Support Vector Machine (SVM) is a state-of-the-art classification method introduced by Vapnik (1995), belonging to the general category of kernel based methods. SVMs are widely used in ML field due to its high accuracy, ability to deal with high-dimensional data, and flexibility in modeling diverse sources of data (Ben-Hur and Weston, 2010).

2.3.6. Multilayer Perceptron

The multilayer feed-forward networks are an important class of neural networks (Haykin, 1999). Typically, the network consists of: a set of sensory units that constitute the input layer; one of more hidden layers of computational nodes; and an output layer of

[2] https://www.r-project.org/.

computation nodes. The input signal propagates through the network in a forward direction, on a layer-by-layer basis. These neural networks are also commonly referred as Multilayer Perceptrons (MLPs). MLPs have been applied successfully to solve some difficult and diverse problems using the back-propagation learning algorithm, due to the general nature of the network learning process. The generality offered by this common process allows arrangement and connectivity of individual units within the network that can vary dramatically (Freeman and Skapura, 1991).

2.3.7. Fuzzy rule-based systems

Fuzzy rule-based systems (FRBSs) are a well-known methods within soft computing, based on fuzzy concepts (Riza et al., 2015). They have become a powerful method to address various problems such as uncertainty, imprecision, and non-linearity. FRBSs are an extension of classical rule-based systems. Basically, they are expressed in the form "IF A THEN B" where A and B are fuzzy sets. In experiments, we used two FBRSs algorithms: Fuzzy-Chi, a system based on Chi's technique; and Fuzzy-W, a weighted-factor fuzzy algorithm that applies Ishibuchi's strategy (Ishibuchi et al., 1999).

3. Results and discussion

3.1. Classification based on all parameters

We ran in our dataset all of the classification algorithms previously described in Section 2.3 and Table 3. Each technique have performed 30 times using a 10-fold Cross Validation resampling strategy on the same data partitions. We extracted some performance measures such as: predictive accuracy, misclassification error, precision, recall and f-measure; all of them balanced according to the number of classes. The results, mean and standard deviation values, are summarized in Table 4.

Looking at the Table 4 we may note that the RF algorithm was the best one. It presented an averaged balanced accuracy value of 0.9440 and a f-score value of 0.9434, over-performing the other algorithms. It also presented the best averaged values for all the performance measures. The Fuzzy ruled-based-systems algorithms, k-NN, MLP, J48 and SVM presented good prediction values, with f-score values between 0.9385 and 0.8274, but they were not so good as RF was. Among them, Fuzzy approaches were the best, probably due to their capacity to deal with non-linearity and uncertainties, that select a parameter when it express a desirable truth degree. At last, the NB was worst one. NB algorithm assumes independence between features, and this fact may compromise the accuracy. Nevertheless, NB obtained the lesser standard deviation in all results, showing stability along every repetitions.

In our experiments we used the default hyper-parameters values suggested by R packages for all the classification algorithms. However, we conducted an additional small experiment to see if a hyper-parameter tuning step would increase or not the performance of the MLP and k-NN algorithms. Fig. 1 left-side depicts MLP performances changing the number of units in the hidden layer. The right-side shows k-NN performances changing the number of the nearest neighbors. In the k-NN scenario, growing the number of neighbors decreases all of the performance measures, and the best results were obtained with $k = 3$ (the default value). For MLPs, using a number of hidden units between 15 and 50 leads to small variations in the performance measures with no significant difference, so the default value ($hidden = 15$) is a good approach.

3.2. Hits and misses

The Fig. 2 depicts all the misclassification predictions accumulated over the 30 algorithms' executions. The x-axis represents all the 250 examples of the dataset, each one representing a pork

Table 4
Algorithms results by cross-validation strategy sorted by accuracy.

Algorithm	Accuracy		Precision		Recall		F-Score	
	mean	std	mean	std	mean	std	mean	std
RF	0.9440	0.0105	0.9440	0.0101	0.9505	0.0092	0.9434	0.0105
Fuzzy Chi	0.9393	0.0143	0.9394	0.0139	0.9461	0.0128	0.9385	0.0143
Fuzzy W	0.9376	0.0160	0.9371	0.0165	0.9450	0.0142	0.9365	0.0165
3-NN	0.9256	0.0131	0.9257	0.0125	0.9340	0.0112	0.9250	0.0132
MLP	0.9104	0.0131	0.9123	0.0132	0.9184	0.0115	0.9100	0.0133
J48	0.8432	0.0193	0.8438	0.0187	0.8565	0.0180	0.8418	0.0192
SVM	0.8283	0.0106	0.8289	0.0107	0.8411	0.0113	0.8274	0.0105
NB	0.7938	0.0057	0.7937	0.0055	0.8116	0.0070	0.7941	0.0055

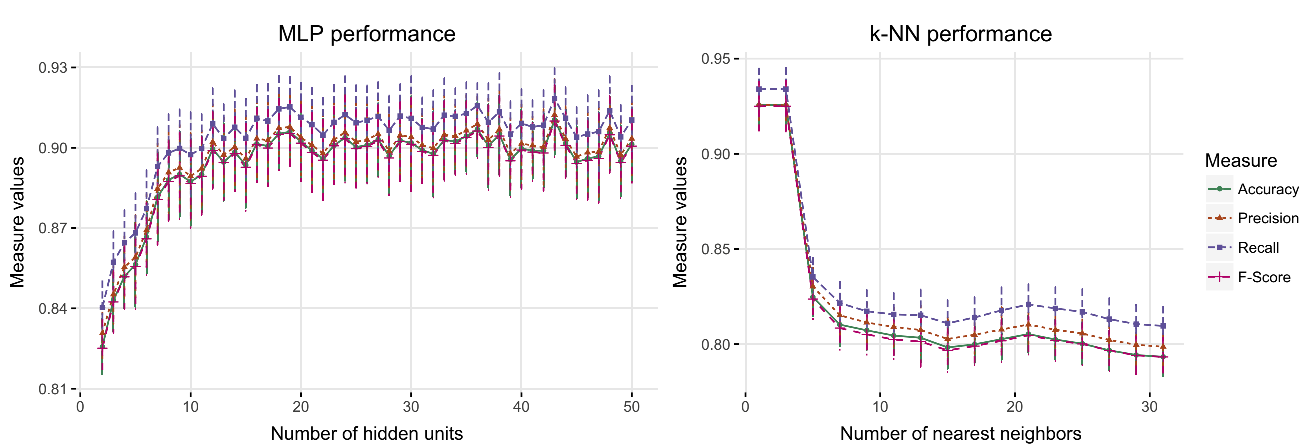

Fig. 1. MLP and K-NN performances analysis according to their hyper-parameters values.

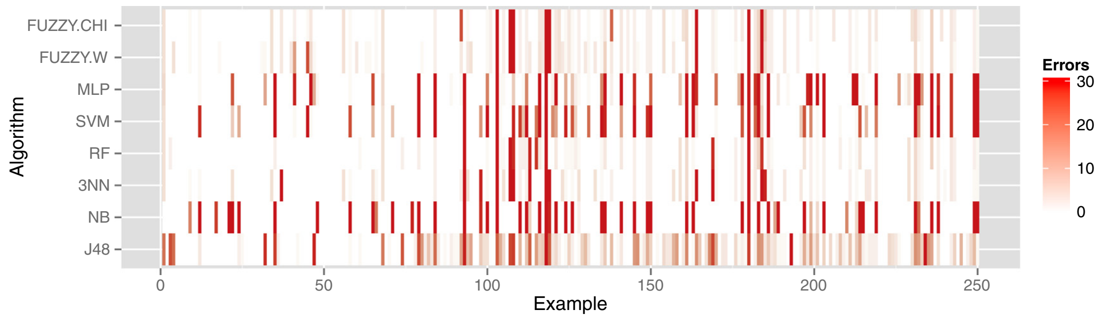

Fig. 2. Accumulated misclassification predictions over 30 executions.

muscle sample. The *y*-axis represents the classification algorithms: all of them that performed on the dataset. In the figure, the whiter a cell, the less is the number of incorrect predictions that the classifier did for that particular sample over all the executions. Higher error values are represented by red.

Reflecting the results on Table 4, the RF's predictions line showed few misclassifications. There are just around 4 problematic samples (96, 103, 118 and 180) that presented a high accumulated error rate. These problematic samples constitute samples of animals that does not match with quality assessment of a specific storage time, in other words, the meat quality attributes were different from other samples, on a expected day. In order to analyze the problematic samples errors, we calculate the accumulated normalized difference of attributes between the problematic ones and each storage time, as Table 5 shows. We observed that samples 96 and 103 presented the minor difference with 14-day, as is possible to see in Fig. 3, when all classification algorithms misclassified these samples as 7-day, so these samples are similar to 14-day. Sample 118 was predicted as 7-day, but it presents minor difference with 0-day, corroborating with classification algorithms. Finally, Sample 180 that had the minor difference with 7-day, but is a sample of 14-day. These problems occur in all predictors and are illustrated in Fig. 3.

Regarding the days' distance, highlighted in Table 5, it was observed that had natural behavior of sample degradation.For example, 0-day presented 3.763 of accumulated mean difference for 7-day. This difference increases when calculated related to 14-day, obtaining 7.077 of difference. Thus, on 7-day the accumulated mean difference of normalized attributes were more balanced, 0-day and 14-day with 3.763 and 4.986, respectively. This variability in the age of meat can be explained by several factors related to individual characteristics of each animal or the biological and operational processes that directly affect the transformation of muscle in meat. The post mortem period involved various enzymatic reactions and intrinsic characteristics of the meat and may be varied according to the animal. The consequences of apoptosis and changes during the post mortem affecting a different set of

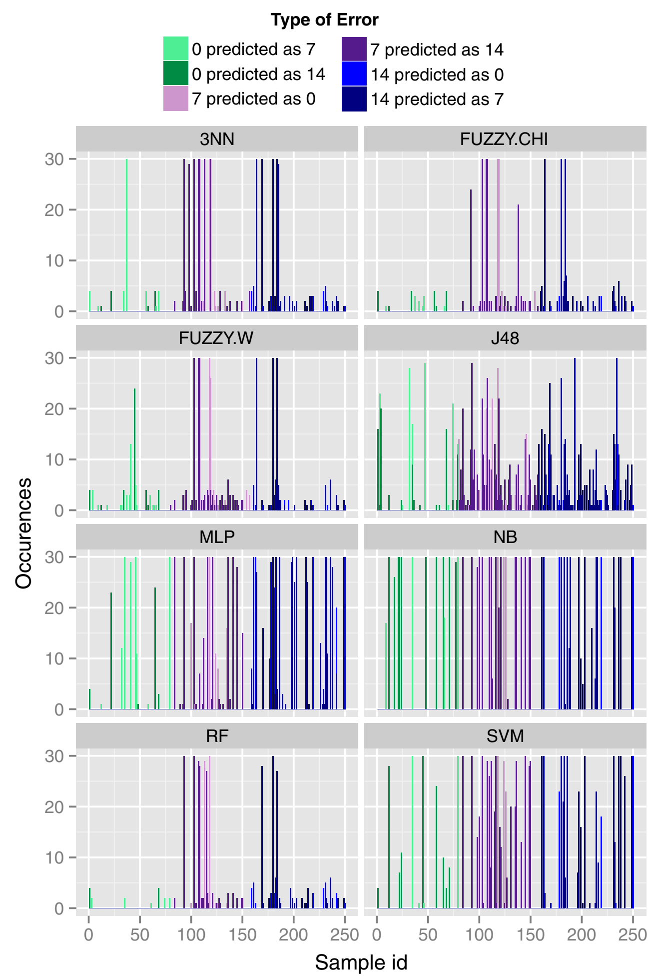

Fig. 3. Type of errors observed during all the executions separated by classifier.

Table 5
Problematic samples comparison with different days.

Predicted as	Mean			Samples			
	0-day	7-day	14-day	96 7-day	103 7-day	118 7-day	180 14-day
L^*	51.80	51.50	54.10	52.40	53.70	51.00	51.90
a^*	4.70	5.15	5.70	7.20	6.60	4.40	5.50
b^*	13.00	12.50	13.50	15.60	14.50	12.1	12.20
WHC	25.14	27.42	28.22	27.79	26.03	29.56	27.17
pH	5.59	5.79	5.77	5.75	5.80	5.60	5.77
TBARS	0.44	0.41	0.42	0.36	0.41	0.54	0.41
Diff 0-Day	–	3.763	7.077	1.0644	0.6923	**0.5558**	0.4029
Diff 7-Day	3.763	–	4.986	0.8055	0.5488	0.6266	**0.1317**
Diff 14-Day	7.077	4.986	–	**0.6234**	**0.3479**	0.7447	0.7447

muscle characteristics and parameters of quality (Herrera-Mendez et al., 2006). These changes can still be influenced by procedures before the transformation of muscle in meat as stress during the pre-slaughter, microbial load and sanitary conditions of employees and slaughterhouses.

RF predicted accurately the class 0-day (first third of samples), but did some errors when predicting the class 7-day. This classification perspective "day-by-day" provided a comprehension about the storage time misclassified. We investigated specific prediction errors following the hypothesis of some samples had particular characteristics that imply in a different time degradation. Thus, our hypothesis supposes that predictions mistake would be similar to 0-day as 7-day, 7-day as 0-day or 14-day and finally 14-day as 7-day. To detail these predictions, in the Fig. 3 we exposed all the classification algorithms to evaluate "day-by-day" results. Given that RF's algorithm achieved the highest accuracy, we observe in Fig. 3 that just two predictions errors (samples 1 and 71) predicted 0-day as 14-day (light green), the majority of errors with more occurrences, in a total of five, were 0-day as 7-day (samples 34, 61, 79, 83 and 84) represented by dark green bar. Endorsing our hypothesis, the dark blue bars are high and numerous in comparison to light blue means that 14-day predicted as 0-day were minor errors, eight against twenty-three of 14-day predicted as 0-day. Analyzing the 7-day prediction, we observe in Fig. 3 on RF's results a seventeen dark purple against six light purple, in other words, the great part of 7-day was predicted as 14 day. One final observation is about the accumulated error rate (length of bars), 7-day shows the highest accumulated error, thus a more complex classification task.

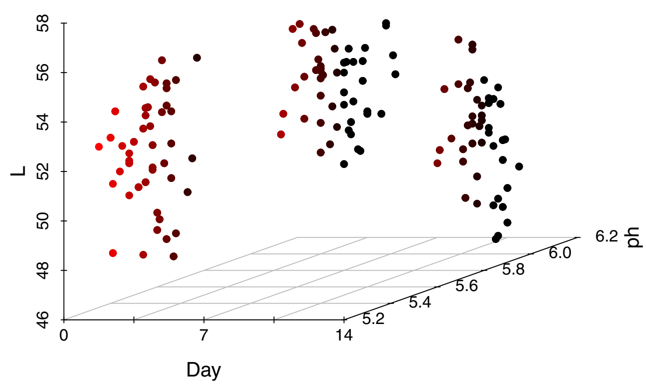

Fig. 4. Dataset distribution of pH, L^* and storage time.

Table 6
Ranking of parameter significance by information and statistical metrics.

Parameter	Entropy			Chi^2	Correlation	
	IG	GR	SU		Pearson	Spearman
pH	0.30	0.45	0.34	0.74	0.55	0.53
L^*	0.16	0.16	0.15	0.36	0.38	0.38
WHC	0.15	0.15	0.14	0.37	0.37	0.37
a^*	0.13	0.18	0.14	0.35	0.40	0.36
b^*	0.09	0.28	0.13	0.41	0.28	0.25
TBARS	0.00	0.00	0.00	0.00	0.03	0.00

3.3. Classification without TBARS

Some researches, e.g. Holmer et al. (2009), identified a significant relation between pH and storage time. In order to clarify this relation in our Dataset, we plot the Fig. 4 to show based on storage time the distribution of pH and L^* color value. It is possible to observe that the low values of pH ($5.35 \leqslant \text{pH} \leqslant 5.77$, mean of 5.58) were characteristic of 0-day, illustrated on Fig. 4 by reddish points, independent of L^* value. 7-day and 14-day showed a reduction of samples with this pH characteristic with values about $5.60 \leqslant \text{pH} \leqslant 5.95$ and $5.53 \leqslant \text{pH} \leqslant 6.02$ respectively. Thus, observing the behavior of pH across different storage times, illustrated by decreased reddish points, we had evidence of relation between pH and storage time, without an apparent relation to L^*.

As mentioned in Section 2.1, to identify the parameters significance for classification purposes, we applied some statistical tests and information theory-based metrics. We obtained similar results to that observed by Holmer et al. (2009) where pH is the most significant parameter, followed by L^*, WHC, a^*, b^* and finally TBARS, as in Table 6.

It is important to observe that pH was the best results in all tests (Entropy, Chi^2 and Correlation). Except TBARS, all others achieved intermediate results, changing positions in each test.

CFS obtained the best subset composed by L^*, a^*, b^*, WHC and pH with $Merit_s = 0.369$. TBARS was not included in the optimal subset. Tests based on exhaustive search which every combination of features across all possible dimensions were performed with 30 repetitions for each combination. The result corroborates with CFS indicating the removal of TBARS from the original set of attributes.

The methodology of TBARS has the disadvantage of being extremely labor intensive and requires more time for its implementation. Besides, it is a biochemical analysis where the malondialdehyde (which quantify this when we conducted analysis), be a molecule which has a high instability.The assessment without TBARS can lead to a fast storage time identification considering the acquisition time of others parameters and respective measures as input in an accurate classification algorithm. Table 7 shows the equipment used for the evaluation of the quality parameters (pH, color, WHC and TBARS) and their respective analysis times.

We notice that some assessments are extremely fast, e.g. seconds for pH. While the evaluation of TBARS can take hours to obtain the result. Thus, the purpose of this work in classify storage time fast way can be confirmed; once, the TBARS analysis showed non-significance in obtaining the storage time.

In our experiments, the lipid oxidation was irrelevant to explain the storage time. To evaluate this fact, we conduct alike experiments as reported in Section 3.1, however based on a dataset without TBARS. Table 8 shows the classification without TBARS and is possible to see a similar rank presented in Table 4.

RF algorithm was the best one (accuracy value of 0.9441), again. The remaining of algorithms were k-NN, Fuzzy approaches, MLP, J48, SVM and NB, sorted by accuracy. Except MLP, that decrease accuracy in 0.0269, the others algorithms obtained an average difference inferior to 5×10^{-3} when compared with TBARS featured dataset.

Table 7
Acquisition time for parameters and respective equipments.

Parameter	Acquisition time	Equipments	References
Color (CIELab)	Minutes	Colorimeter	de lEclairage (1978)
pH	Seconds	pHMeter	Tarsitano et al. (2013)
WHC	Minutes	Semi-analytical balance, paper filters, two acrylic plates and weight of about 10 kg	Barbut (1996)
TBARS	Hours	Spectrophotometer, chemical reagents and laboratory equipments	Pikul et al. (1989)

Table 8
Algorithms results sorted by accuracy executed without the TBARS feature.

Algorithm	Accuracy		Precision		Recall		F-Score	
	mean	std	mean	std	mean	std	mean	sd
RF	0.9441	0.0118	0.9444	0.0117	0.9512	0.0112	0.9437	0.0120
3-NN	0.9357	0.0136	0.9374	0.0132	0.9434	0.0113	0.9356	0.0140
Fuzzy Chi	0.9323	0.0155	0.9331	0.0153	0.9392	0.0134	0.9317	0.0160
Fuzzy W	0.9235	0.0147	0.9226	0.0144	0.9331	0.0126	0.9222	0.0151
MLP	0.8835	0.0127	0.8867	0.0123	0.8952	0.0115	0.8826	0.0127
J48	0.8364	0.0207	0.8384	0.0200	0.8505	0.0220	0.8355	0.0204
SVM	0.8203	0.0101	0.8240	0.0099	0.8302	0.0105	0.8169	0.0104
NB	0.7826	0.0061	0.7847	0.0062	0.7984	0.0072	0.7828	0.0059

4. Conclusion

In the current research eight different CI algorithms were used to identify three post mortem storage times from pork samples. The parameters of samples included pH, CIELab colors, WHC, and TBARS. Based on significance evaluation, TBARS was non-significant in statistical ($Chi^2 = 0.00$, Pearson's coefficient 0.03, Spearman's coefficient 0) and Entropy (IG 0.00, GR 0.00, and SU 0.00). The most significative parameter was pH based on statistical ($Chi^2 = 0.74$, Pearson's coefficient 0.55, Spearman's coefficient 0.53) and entropy (IG 0.30, GR 0.45 and SU 0.34) measurements.

Comparison of algorithms results showed RF as the best classifier, and its accuracy and precision were 94.40%. After RF, Fuzzy Chi (93.93%), Fuzzy W(93.76%), k-NN with k equals 3 (92.56%), MLP with 8-3-1 topology (91.04%), J48 (84.32%), SVM (82.83%) and NB (79.38%). If we consider ±5% as an accuracy tolerance to compare the classification results, this can be concluded RF, Fuzzy approaches, k-NN and MLP presented a good performance to be applied in real scenarios.

It is important to highlight that samples of day 0 showed the best classification results, with few occurrences of day 0 identified as day 7. Days 7 and 14 presented similar complexity for storage time identification.

Therefore, the storage time of pork could be predicted with satisfactory accuracy without TBARS parameter in appropriate condition of storage as in the vacuum package. As exposed in our experiments without TBARS, the classifiers achieve similar prediction results with the same classification ranking. We introduce in this paper the use of CI for fast storage time identification only by pH, CIELab and WHC, all of them feasible to local assessment and rapidly measured by portable devices.

Acknowledgments

The authors would like to thank CNPq, CAPES and FAPESP (Brazilian Agencies) for their financial support, specially the grant #2012/23114-9, São Paulo Research Foundation (FAPESP).

References

Argyri, A.A., Jarvis, R.M., Wedge, D., Xu, Y., Panagou, E.Z., Goodacre, R., Nychas, G.-J. E., 2013. A comparison of raman and ft-ir spectroscopy for the prediction of meat spoilage. Food Control 29 (2), 461–470.
Ballin, N.Z., 2010. Authentication of meat and meat products. Meat Sci. 86 (3), 577–587.
Barbut, S., 1996. Estimates and detection of the PSE problem in young turkey breast meat. Can. J. Anim. Sci. 76 (3), 455–457.
Ben-Hur, A., Weston, J., 2010. A user's guide to support vector machines. In: Data Mining Techniques for the Life Sciences. Methods Mol. Biol., Vol. 609. Humana Press, pp. 223–239.
Breiman, L., 1996. Bagging predictors. Mach. Learn. 24 (2), 123–140.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Chen, Q., Cai, J., Wan, X., Zhao, J., 2011. Application of linear/non-linear classification algorithms in discrimination of pork storage time using fourier transform near infrared (FT-NIR) spectroscopy. LWT-Food Sci. Technol. 44 (10), 2053–2058.
Chmiel, M., Słowiński, M., Dasiewicz, K., 2011. Lightness of the color measured by computer image analysis as a factor for assessing the quality of pork meat. Meat Sci. 88 (3), 566–570.
de lEclairage, C.C.I., 1978. International commission on illumination, recommendations on uniform color spaces.
Faucitano, L., Ielo, M., Ster, C., Fiego, D.L., Methot, S., Saucier, L., 2010. Shelf life of pork from five different quality classes. Meat Sci. 84 (3), 466–469.
Freeman, J.A., Skapura, D.M., 1991. Neural Networks: Algorithms, Applications, and Programming Techniques. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA.
Hammelman, J., Bowker, B., Grant, A., Forrest, J., Schinckel, A., Gerrard, D., 2003. Early postmortem electrical stimulation simulates PSE pork development. Meat Sci. 63 (1), 69–77.
Haykin, S., 1999. Neural Networks: A Comprehensive Foundation. Prentice Hall.
Herrera-Mendez, C.H., Becila, S., Boudjellal, A., Ouali, A., 2006. Meat ageing: reconsideration of the current concept. Trends Food Sci. Technol. 17 (8), 394–405.
Holmer, S., McKeith, R., Boler, D., Dilger, A., Eggert, J., Petry, D., McKeith, F., Jones, K., Killefer, J., 2009. The effect of ph on shelf-life of pork during aging and simulated retail display. Meat Sci. 82 (1), 86–93.
Honikel, K.O., 1998. Reference methods for the assessment of physical characteristics of meat. Meat Sci. 49 (4), 447–457.
Ishibuchi, H., Nakashima, T., Murata, T., 1999. Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems. IEEE Trans. Syst. Man Cybern. Part B 29 (5), 601–618.
Kamruzzaman, M., Sun, D.-W., ElMasry, G., Allen, P., 2013. Fast detection and visualization of minced lamb meat adulteration using NIR hyperspectral imaging and multivariate image analysis. Talanta 103, 130–136.
Kodogiannis, V.S., Alshejari, A., 2014. An adaptive neuro-fuzzy identification model for the detection of meat spoilage. Appl. Soft Comput. 23, 483–497.
Liao, Y., Fan, Y., Cheng, F., 2012. On-line prediction of pH values in fresh pork using visible/near-infrared spectroscopy with wavelet de-noising and variable selection methods. J. Food Eng. 109 (4), 668–675.
Liu, M., Wang, M., Wang, J., Li, D., 2013. Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: application to the recognition of orange beverage and chinese vinegar. Sensor. Actuator. B: Chem. 177, 970–980.
Liu, Y., Peng, Y., Zhang, L., Dhakal, S., Wang, C., 2014. Nondestructive detection of pork comprehensive quality based on spectroscopy and support vector machine. In: SPIE Sensing Technology+ Applications. International Society for Optics and Photonics, p. 91080R.
Maimon, O., Rokach, L., 2010. Data Mining and Knowledge Discovery Handbook, second ed. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
Mancini, R., Hunt, M., 2005. Current research in meat color. Meat Sci. 71 (1), 100–121.
Marsland, S., 2009. Machine Learning: An Algorithmic Perspective, first ed. Chapman & HallCRC.
Mitchell, T.M., 1997. Machine Learning, first ed. McGraw-Hill, Inc., New York, NY, USA.
Pearce, K.L., Rosenvold, K., Andersen, H.J., Hopkins, D.L., 2011. Water distribution and mobility in meat during the conversion of muscle to meat and ageing and the impacts on fresh meat quality attributes—a review. Meat Sci. 89 (2), 111–124.
Pikul, J., Leszczynski, D.E., Kummerow, F.A., 1989. Evaluation of three modified TBA methods for measuring lipid oxidation in chicken meat. J. Agric. Food. Chem. 37 (5), 1309–1313.
Prevolnik, M., Čandek-Potokar, M., Škorjanc, D., 2010. Predicting pork water-holding capacity with NIR spectroscopy in relation to different reference methods. J. Food Eng. 98 (3), 347–352.
Przybylak, A., Ślósarz, P., Boniecki, P., Koszela, K., Przybył, K., Wojcieszak, D., Szulc, R., Ludwiczak, A., Górna, K., 2015. Marbling classification of lambs carcasses with the artificial neural image analysis. In: Seventh International Conference on Digital Image Processing (ICDIP15). International Society for Optics and Photonics, p. 963113.
Puolanne, E., Halonen, M., 2010. Theoretical aspects of water-holding in meat. Meat Sci. 86 (1), 151–165.
Qiao, J., Wang, N., Ngadi, M., Gunenc, A., Monroy, M., Gariepy, C., Prasher, S., 2007. Prediction of drip-loss, pH, and color for pork using a hyperspectral imaging technique. Meat Sci. 76 (1), 1–8.

Quinlan, J.R., 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
Ravikanth, L., Singh, C.B., Jayas, D.S., White, N.D., 2015. Classification of contaminants from wheat using near-infrared hyperspectral imaging. Biosyst. Eng. 135, 73–86.
Riza, L.S., Bergmeir, C., Herrera, F., Benítez, J.M., 2015. frbs: Fuzzy rule-based systems for classification and regression in R. J. Stat. Softw. 65 (6), 1–30 <http://www.jstatsoft.org/v65/i06/>.
Rodríguez-Pulido, F.J., Gordillo, B., González-Miret, M.L., Heredia, F.J., 2013. Analysis of food appearance properties by computer vision applying ellipsoids to colour data. Comput. Electron. Agric. 99, 108–115.
Ropodi, A., Pavlidis, D., Mohareb, F., Panagou, E., Nychas, G.-J., 2015. Multispectral image analysis approach to detect adulteration of beef and pork in raw meats. Food Res. Int. 67, 12–18.
Rosenvold, K., Andersen, H.J., 2003. Factors of significance for pork quality—a review. Meat Sci. 64 (3), 219–237.
Salmi, B., Trefan, L., Bünger, L., Doeschl-Wilson, A., Bidanel, J.P., Terlouw, C., Larzul, C., 2012. Bayesian meta-analysis of the effect of fasting, transport and lairage times on four attributes of pork meat quality. Meat Sci. 90 (3), 584–598.
Sentandreu, M.Á., Sentandreu, E., 2014. Authenticity of meat products: tools against fraud. Food Res. Int. 60, 19–29.
Shan, Z., HaiBin, W., FengBing, Z., Jun, B., et al., 2015. Study on hyperspectral image technology based on manifold fuzzy clustering for pork quality classification. J. Food Safe. Quality 6 (4), 1421–1428.
Tarsitano, M.A., Bridi, A.M., da Silva, C.A., Constantino, C., Andreo, N., Peres, L.M., Youssef, E.Y., Cardoso, T. d. A.B., 2013. Pork meat matured for different periods of time in vacuum-packaging system. Semina: Ciências Agrárias 34 (6Supl2), 4015–4024.
Tejerina, D., García-Torres, S., Cava, R., 2012. Water-holding capacity and instrumental texture properties of m. Longissimus dorsi and m. Serratus ventralis from iberian pigs as affected by the production system. Livest. Sci. 148 (1), 46–51.
Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. Springer-Verlag New York, Inc., New York, NY, USA.
Xiong, Z., Sun, D.-W., Pu, H., Xie, A., Han, Z., Luo, M., 2015. Non-destructive prediction of thiobarbituricacid reactive substances (TBARS) value for freshness evaluation of chicken meat using hyperspectral imaging. Food Chem. 179, 175–181.
Yu, L., Liu, H., 2004. Efficient feature selection via analysis of relevance and redundancy. J. Mach. Learn. Res. 5, 1205–1224.
Zapotoczny, P., Szczypiński, P.M., Daszkiewicz, T., 2016. Evaluation of the quality of cold meats by computer-assisted image analysis. LWT-Food Sci. Technol. 67, 37–49.