Data Mining Generating Decision Trees to Alert System Against Death and Losses in Egg Production

Maria Elena Silva Montanhani [1]; Leda Gobbo de Freitas Bueno[1]; Érik dos Santos Harada[1]; Danilo Florentino Pereira[2]; Mario Mollo Neto[2]

[1]São Paulo State University (UNESP), Faculty of Agricultural and Technological Sciences, Dracena/SP Brazil - Rod. Cmte João Ribeiro de Barros, km 651 - Bairro das Antas - Dracena/SP ZIP code 17900-000Phone: (55 18) 3821-8200

[2]São Paulo State University (UNESP), Faculty of Science and Engineering, Tupã/SP Brazil Domingos da Costa Lopes Street, 780 - Jd. Itaipu - Tupã / SP - ZIP Code 17602-496 Phone: (55 14) 3404-4200

Corresponding author: Mario Mollo Neto[2], **mario.mollo@unesp.br**

Abstract

Climatic changes and high temperatures have been affecting animal production and the well-being of laying birds, with heat stress and high mortality rates, generating economic losses. Legacy databases can contain information to help model thermal comfort at climatic extremes. They can enable decision trees to be created through the use of data mining to prevent mortality and production losses. Thus, the objective of this study is to seek to develop decision trees, for application as an alert system, for the incidence of caloric stress in the production of layers. We used a database of three aviaries located in the city of Bastos-SP, collected in 2013. The data were organized in Excel® spreadsheets, and processed with the Weka® software with the J48 (C4.5) algorithm for mining of the data. The technique allowed the construction of decision trees that in the chosen sheds were classified with respectively 99.73%, 99.61%, and 98.71% of correct answers and with Kappa indexes equal to 0.9958, 0.9907 and 0.9663, which indicate that the three classifiers built are excellent. Thus, the proposed system, with the decision trees built, can serve as a basis for the construction of an alert system to be applied to the three warehouses simultaneously.

Keywords: Sustainability. Climatic Extremes. Data Mining. Layer Poultry.

1. INTRODUCTION

Climate change has been affecting agricultural systems. INMET- Brazilian National Institute of Meteorology [1] issued a heatwave warning with a degree of severity of danger with temperatures of 5 ° C above average for a period of 3 to 5 days that began on September 12, 2019, the one where it reaches several regions of the states of São Paulo and Mato Grosso do Sul, with Bastos-SP being one of the affected municipalities.

At the same time, the demand for this animal production increases in the world due to the growth of the population and the increase in per capita income in developing countries, according to [2], the world population by 2050 should surpass 9 billion, consequently the demand for products of animal origin (meat, milk, eggs) also increases. Therefore, it is necessary to understand how animal production systems are affected by extreme events, which are becoming more frequent and intense with climate change.

There is a discussion within the global context, which presents itself as a major interdisciplinary challenge of studies applied to propose strategies to mitigate the impacts of climate change on production systems [3], [4] and [5].

These advances must occur in a combination of expanding current knowledge, innovative research, and effective dissemination of research [6].

One of the main causes of losses in modern poultry is the bioclimatic and aerial environment, being one of the main causes. According to [7] the climatic extremes that have been happening can generate severe losses in production, due to heat stress, severely impacting the use of new approaches in the genetic selection of animals. In addition to thermal stress being harmful to laying birds as it leads to increased mortality, it also causes loss of body condition that directly affects the number of eggs that would be produced, and it is important that geneticists can predict the impact of environmental factors on climate adaptation [8].

Thermal stress causes a drop in feed intake, lower growth rate, higher water consumption, an acceleration of the heart rate, changes in feed conversion. And it leads to a drop in egg production and a higher incidence of eggs with soft shells, a decrease in egg size, a noticeable decrease in quality standards and a reduction in fertility and hatchability of eggs [9], [10] and [11].

Searching for solutions to mitigate these problems that directly affect productivity and mortality, consequently, animal welfare, producers, and researchers began to collect a large volume of data regarding their production. For this, much data is currently collected through research, but often they are not used or still do not offer the expected result.

Data mining is a very promising technology in this context, making use of this collected data generating useful information for the development of decision trees, giving the producer the decision making in a given problem through the use of the data mining tool in the database that the producer has collected. Is also, among other tools aimed at decision making and more precise actions, has contributed to the advancement and speed of research in animal production.

Data mining is a database analysis technique that stands out, being very promising in several areas of knowledge, according to [12] involves tasks of classification, association, or grouping. The classification encompasses several interesting application possibilities, derived from the ability to develop patterns that can be represented graphically through a decision or classification tree.

Given the presented scenario, it is assumed that, based on data mining, decision trees can be developed, generating the bases for the creation of an alert system for the incidence of caloric stress, minimizing the losses caused by heat that affect the well-being and productivity of laying hens in production sheds.

2. LITERATURE REVIEW

According to [13], the increase in productivity is related to several factors, such as improvement of lines and inputs, investments in the automation of the production process, improvement of employees concerning management, control of sanitary conditions, among others.

External environmental factors can interfere with the microclimate generated inside the facilities, thereby causing impacts on production, with consequent economic losses to exploration. Besides, heat stress can affect the metabolism of laying hens, generating reflexes, and impacts on the production amounts and egg quality.

As indicated by [14], maximum productivity in laying hens is achieved when they are in an environment that provides minimal energy changes and meets welfare needs, where the bird has the complete state of its physical and mental health, comfort state for the animal.

For this, in the last ten years, a lot of data has been collected through research, however, these legacy databases are often not fully used or still do not offer the expected result.

Data mining is a very promising technology and in this context it has been making use of this collected and unused data in its entirety, thus allowing it to be generating useful information for the development of decision trees, giving the producer the support for decision making about some of the problems that arise in the day-to-day processes through the use of the data mining tool in the database that the producer has previously collected.

This large volume of data requires computational means and specific techniques to interpret them [15].

Several Data Mining Techniques have been used to detect relationships between the different attributes found in large databases [16].

Decision trees classify cases using their rules based on the values of each determinant data. Each node in a decision tree represents a rule in a classification instance, and each branch represents a value that the node can assume. A decision tree starts at the root node which is the most relevant attribute and is ordered based on their values up to the leaves [17].

According to the authors [18] and [19], data mining is a database analysis technique that has been applied in several areas of knowledge. This fact, according to [12], involves classification, association, or grouping tasks.

In their work, [20] related, the occurrence of heatwaves with the incidence of high mortality of layers created in different types of facilities, using a database. The classification was made on normal and high mortality through data mining using the J48 algorithm. The authors find a classification tree with 71% accuracy for high mortality. Thus, the use of mining with the generation of the classification, made it possible to associate the occurrences of a heatwave with the increase in the mortality of laying hens.

3. MATERIAL AND METHODS

This study was divided into two stages, wherein the first stage the survey of the database resulting from a collection of data from Dry Bulb Temperature (ºC) and Relative Air Humidity (%), carried out in 2013, was carried out. in three commercial laying hens, two of the vertical system type and one of the pyramid type.

The aviaries belonged to the same commercial farm, located in the municipality of Bastos-SP (latitude 21º55'19 "south and longitude 50º44'02" west, the altitude of 445 meters).

The region's climate is subtropical Aw (rainy tropical with dry winter and coldest month with an average temperature above 18ºC. The driest month has precipitation below 60 mm and the rainy season is late for autumn).

The layers used in the study were from the Dekalb White line, all of the same age and with the same dietary composition, which entered the production system in March 2013, at 17 weeks of age.

The aviaries differed concerning the typology, construction material, and air conditioning systems (without air conditioning, the tunnel with negative pressure, and sprinkling on the roof), named A1, A2, and A3 respectively.

This experiment was approved by the Commission for the Use of Animals in Research, Teaching, and Extension (CEUA) of FCAT / Unesp under protocol number 11/2012.

The monitoring of the external environment was carried out through the collection of meteorological data from the station installed on the Tupã Campus, which is about 20 km from the municipality of Bastos.

During the study, the year of production of 2013 was evaluated.

Data loggers were distributed inside each aviary to measure the variables.

To characterize the internal environment of the warehouses, dry bulb temperature (Tbs) and relative humidity (UR) were monitored, using thirty-one HOBO U12-012 data loggers (Onset®).

Five HOBO data loggers were separated for possible substitutions.

Seven HOBO data loggers were distributed in aviary A1 and twelve HOBO data loggers in shed A2 and A3.

These data loggers recorded bioclimatic variables every thirty minutes, installed in each corridor of the warehouses, and at three different heights.

Operation of shed A1 (Pyramidal): The production system is of the pyramidal type, with three floors of cages and two batteries.

Feeding was carried out automatically, water was supplied ad libitum and carried out by nipple drinkers.

This shed did not have any type of air conditioning system.

Operation of shed A2 (Negative tunnel): It was a production system of the vertical type, containing three batteries with six floors of cages.

The feed was supplied by automatic feeder and water was available in ad libitum drinkers, three for every two cages. The feed was the same provided in shed A1 due to the birds being of the same age and lineage.

In this warehouse, twelve HOBO data loggers were installed, distributed in three units per corridor.

In this shed, the air conditioning system was carried out by exhaust fans and pad cooling (made of cellulose). With fourteen exhaust fans at one end of the building and the other, the pad cooling also presents on the right and left sides of the end.

As for the handling of the curtains of this shed, they remained closed throughout the day, due to the air conditioning system that requires total sealing of the environment for the adiabatic tunnel to function and so that the internal conditions were not affected by the external ones.

Operation of the A3 shed (without air conditioning): The A3 shed was of the vertical type and had no air conditioning system, its natural ventilation, and sprinkling on the roof.

Feeding was performed by the automatic feeder, nipple drinkers, and water was offered ad libitum. The diet offered was the same as in sheds A1 and A2.

There were two batteries with six levels of cages, one battery consisting of wire cages and the other made of polyethylene. Twelve Hobos were installed in this shed.

Calculation of the THI: For each assessment, both those carried out by the HOBO data loggers from the data from the experiment, the Temperature and Humidity Index (THI) was calculated using the equation proposed by [21].

Comfort bands were used according to the THI, these bands were obtained through the 4 stress zones. The definition in 4 zones, were based on the thermoneutrality center using air temperature (minimum and maximum) and relative humidity (RH), which are based on [22] and [23], and the laying lineage manual [24].

Table 1 describes each of them:

Table 1 - Comfort zones according to THI.

Zone	THI	Stress Level
1	< 67,16	Mild Stress
2	67,16 to 76,96	Thermoneutrality
3	76,97 to 89,85	Heat Stress
4	> 89,96	Emergency

Source: Adapted from [22] and [23], and the laying lineage manual [24].

Data Mining: The second stage of the study consisted of carrying out the Data Mining process and this was divided into distinct phases, which were cyclically followed, to allow for new considerations and reassessments concerning the domain and the inclusion or exclusion of attributes, according to the methodology proposed by the CRISP-DM consortium [25].

The data were computationally processed using the Weka® program version 3-8-3 [17] and using the J48 algorithm as applied by [20]. The task that should be used in the analyzes will be the classification, with the construction of decision trees.

The models induced with a variation in the number of instances (or observations), per sheet, were evaluated using the cross-validation method. Pruning of attributes that did not contribute to the performance results of decision trees was carried out. The selection of the best model was made based on the measures: precision; the number of leaves (rules) and the Kappa coefficient.

The Kappa coefficient is used to describe the measure of agreement between the predicted and observed classes. This coefficient ranges from 0 to 1, representing very poor to excellent classification results, respectively [26].

As a result of the decision tree model induction, it was possible to obtain the confusion matrix, which according to [27] is widely used in the statistical analysis of concordance.

From the confusion matrix, according to [26], it was possible to obtain performance evaluation measures.

4. RESULTS AND DISCUSSION

Through the data that was organized for the period, in an Excel® spreadsheet for each aviary and classified in THI ranges, a file compatible with the Weka® software was built, which is a text file, formatted to organize the collected data, with its attributes corresponding to the values and the comfort range understudy that was determined, this file receives the extension ".arff".

With the ".arff" files completed, they were executed with the Weka® software, applying the classification procedure with the choice of the J48 algorithm as applied by [20], which resulted in the construction of the answer outputs for the construction of the decision trees of this research.

Annual analysis for THI indicator

Pyramidal aviary (A1):

The file with extension ".arff" referring to the data of the pyramidal aviary (A1) of the summer for THI (temperature and Humidity Index), used the attributes of the database referring to the sensor readings for dry bulb temperature, relative humidity, THI and stress status of the birds and the result of the processing in the software generated the corresponding decision tree that is presented in algorithm format:

(The numbers indicated in parentheses indicate the number of processing instances that were selected with this classification indicated in the line of the algorithm).

```
THI <= 76.94
|   THI <= 67.13: Mild Stress (171.0)
|   THI > 67.13: Thermoneutrality (330.0)
THI > 76.94: Heat Stress (242.0)
```

The algorithm obtained as a mining model leads to classification with 99.73% correct answers.

The Kappa index has a value of 0.9958, which according to the methodology proposed by [26] allows us to infer that the classifier obtained from mining is excellent, as the value is very close to 1, which indicates the best condition.

Air-conditioned vertical aviary (A2):

The file with extension ".arff" referring to the data of the air-conditioned vertical aviary (A2) for THI (temperature and Humidity Index), used the attributes of the database referring to the sensor readings for dry bulb temperature, relative humidity, THI and stress status of the birds and the result of the processing in the software generated the corresponding decision tree that is presented in algorithm format:

(The numbers indicated in parentheses indicate the number of processing instances that were selected with this classification indicated in the line of the algorithm).

```
THI <= 76.96
|       THI <= 67,08: Mild Stress (49.0)
|       THI > 67.08: Thermoneutrality (580.0)
THI > 76.96: Heat Stress (157.0)
```

It is observed that the algorithm obtained as a mining model leads to classification with 99.61% correct answers. The Kappa index, in this case, also reached the value of 0.9907, which according to the methodology proposed by [26] allows us to infer that the classifier obtained from mining is also excellent.

Vertical unheated poultry house (A3):

The file with extension ".arff" referring to the data of the non-conditioned vertical aviary (A3) for THI (temperature and Humidity Index), used the attributes of the database referring to the sensor readings for dry bulb temperature, relative humidity, THI and stress status of the birds and the result of the processing in the software generated the corresponding decision tree that is presented in algorithm format:

(The numbers indicated in parentheses indicate the number of processing instances that were selected with this classification indicated in the line of the algorithm).

```
THI <= 67.14: Mild Stress (78.0)
THI > 67.14
|     THI <= 76.22: Thermoneutrality (227.0)
|     THI > 76.22
|     |     dbt <= 26.755: Thermoneutrality (4.0)
|     |     dbt    >26.755: Heat Stress (2.0)
```

It is observed that the algorithm obtained as a mining model leads to classification with correct answers of 98.71%. The Kappa index obtained, in this case, had a value of 0.9663 which according to the methodology proposed by [26] allows us to infer that the classifier obtained from mining, in this case, is also excellent.

In this research, the three aviaries, after pruning the attributes that did not contribute to the results, it is observed that only the THI indicator was used in the decision trees, which shows that the Temperature and Humidity Index using based on an adaptation of the equation proposed by [21], it is a tool that helps a lot in the process of controlling thermal comfort conditions for production warehouses in countries with a tropical climate like Brazil.

Thus, the proposed alert system, in the year 2013, can be applied to warehouses (A1), (A2) and (A3) based on the generation of trees for the three breeding models as a basis for algorithms for the generation of an alert system with high precision that was indicated by the results of the Kappa indices.

As presented in theory by [17], the trees corresponding to warehouses (A1) and (A2) have a size equal to 5 with three decision branches.

The tree generated for the shed (A3), as indicated by [17], the size obtained was equal to 7 with 4 decision branches, thus being a little more complex in terms of processing.

In summary, the technique finally allowed the construction of decision trees which, in the chosen sheds, were classified with respectively 99.73%, 99.61% and 98.71% of correct answers and with Kappa indexes equal to 0.9958, 0, 9907 and 0.9663, which indicate that the three classifiers constructed were excellent as indicated by the authors [26].

The warehouses (A1) and (A2) are very close in their results to the classifiers with accuracy in their hits always greater than 99%, although the construction characteristics are very different, which can be a point to be considered by producers in the region when the definition of the type of house to be built aiming at the thermal comfort of birds based on THI, meeting the comfort indications made by [14].

The shed (A3), on the other hand, brought slightly fewer precision results, but even more than 96%.

The Kappa indicators in the three cases were excellent, which indicates that the trees allowed the design of a support system for the decision of high-quality discrimination regarding the monitoring of THI.

Thus, we observed that the proposed system, having in its motor algorithm the construction of the coding with the decision trees constructed as indicated by [15] due to the computational application, can serve as a basis for the construction of an alert system with great accuracy to be applied to the three warehouses simultaneously.

From this study, producers can monitor and act to mitigate heat stress problems for their three breeding models simultaneously with a single computer program that will generate THI-based alerts.

5. CONCLUSION

Through the legacy database used, it was possible to identify the thermal comfort bands that were used for the development of the scenarios created for the construction of decision trees with the return of classification.

The generated models allowed us to identify and predict environmental conditions for laying hens in the production phase that presented excellent Kappa coefficient performance in the study cases, using the THI indicator as a priority.

The study allowed prospecting by an indicator that allowed trees to be generated for the three shed models simultaneously, with great accuracy in classification, taking advantage of a legacy database and which provided the basis for building a future information system for supervision and control of the sheds understudy in the city of Bastos in São Paulo/Brazil.

6. ACKNOWLEDGMENT

This work was supported by the Brazilian National Council for Scientific and Technological Development – CNPq (No. 313339 / 2019-8) and São Paulo Research Foundation – FAPESP (2018/17447-1).

7. REFERENCES

[1] INMET- INSTITUTO NACIONAL DE METEOROLOGIA. **Avisos meteorológicos Brasil**, 2019. Disponível em: <http://alert-as.inmet.gov.br/cv/emergencia/cap/10648?lang=pt_BR>. Acesso em 19 de setembro de 2019.

[2] GARG, M.R.; SHERASIA, P.L.; SHELKE, S.K.; PHONDBA, B.T. Productivity enhancement and methane emission reduction through ration balancing. **Indian Dairyman**. 64(8):54-58. Indain Dairy Association, New Delhi, 2012.

[3] VIGODERIS, R. B., SILVA, J.M., GUISELINE, C., PANDORFI[3], H., & VIEIRA, D. V. Broilers thermal comfort and performance utilizing two different wood-burning heating systems. Acta Scientiarum. **Animal Sciences**, 40, e39194, 2018. Epub 16 de Agosto de 2018.https://doi.org/10.4025/actascianimsci.v40i1.39194

[4] COELHO, D.J. DE R., TINÔCO, I. F. F., SOUZA, C.F., BAPTISTA, F. J. F., BARBARI, M., & OLIVEIRA, K.P... Thermal environment of masonry-walled poultry house in the initial life stage of broilers. **Rev. bras. eng. agríc. ambient**., Campina Grande , v. 23, n. 3, p. 203-208, mar. 2019. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-43662019000300203&lng=pt&nrm=iso>. Acessos em 10 jun. 2020. https://doi.org/10.1590/1807-1929/agriambi.v23n3p203-208.

[5] BARRETT, N. W.; ROWLAND, K.L.; SCHMIDT, C. J.; LAMONT, S. J.; ROTHSCHILD, M. F.; ASHWELL, C. M.; PERSIA, M. E. Effects of acute and chronic heat stress on the performance, egg quality, body temperature, and blood gas parameters of laying hens. **Poultry Science**, v. 98, n. 12, p. 6684-6692, 2019.

[6] EDENHORFER, O. et al. (Ed.). **Climate change 2014: mitigation of climate change: Working Group III contribution to the fifth assessment report of the intergovernmental panel on climate change**. New York: Cambridge University Press, 2014. Disponível em: <http://www.ipcc.ch/pdf/assessment-report/ar5/wg3/ipcc_wg3_ar5_frontmatter.pdf>. Acesso em: 21 ago.2018.

[7] NARDONE, A. et al. Climatic effects on productive traits in livestock. **Veterinary Research Communications**, v.30, n.1, p.75-81, 2006.

[8] PEREIRA, D. F.; VALE, M. M.; ZEVOLLI, B. R.; SALGADO, D. D. Estimating mortality in laying hens as the environmental temperature increases. **Brazilian Journal of Poultry Science**, v. 12, n. 4, p. 265-271, out. /Dez. 2010.

[9] JÁCOME, I. M.; FURTADO, D. A.; LEAL, A. F.; SILVA, J. H.; MOURA, J. F. Avaliação de índices de conforto térmico de instalações para poedeiras no nordeste do Brasil. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 11, n. 5, p. 527-531, set./out. 2007.

[10] TINÔCO, I. F. F. Avicultura industrial: Novos conceitos de materiais, concepções e técnicas construtivas disponíveis para galpões avícolas brasileiros: **Revista Brasileira de Ciência Avícola**, v.3, p.1-26, 2001.

[11] MORENG, R. E; AVENS,. **Ciência e produção de aves**. Livraria Roca, 1990.

[12] LAVRAC, N. Machine learning for data mining in medicine. **Lecture Notes in Computer Science**, v.1620, p.47-62, 1999.

[13] OLIVEIRA, D.R.M.S.; NÄÄS, I.A. **Issues of sustainability on the Brazilian broiler meat production chain**. In: INTERNATIONAL CONFERENCE ADVANCES IN PRODUCTION MANAGEMENT SYSTEMS, 2012, Rhodes. Anais. Competitive Manufacturing for Innovative Products and Services: proceedings, Greece: Internacional Federation for Information Processing, 2012.

[14] ANDRADE, R.R.; TINOCO, I.F.F.; SOUZA, C.F.; OLIVEIRA, K.P.; BARBARI, M.; CRUZ, V.M.F.; BAPTISTA, F.J.F.; VILELA, M.O.; CONTI, L.; ROSSI, G. Effect of thermal environment on body temperature of early-stage laying hens. **Agronomy Research** 16(2), 320 327, 2018.

[15] ANID, **Associação Nacional Para Inclusão**, 2015. A anid atua pela inclusão social e digital no brasil Disponível em: < https://www.anid.org.br/site/sobre.html >.

[16] FAYYAD, U; PIATETSKY-SHAPIRO, G; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. **American Association for Artificial Intelligence**, 1996.

[17] WITTEN, I.H.; FRANK, E.; HALL, M.A. **Data mining: practical machine learning tools and techniques**. São Francisco, CA: The Morgan Kaufmann series in data management systems, 2011. 665p.

[18] ZHAO, Y.; ZHANG, C.; ZHANG, Y.; WANG, Z.; LI, J. A review of data mining technologies in building energy systems: Load prediction, pattern identification, fault detection and diagnosis. **Energy and Built Environment**, v.1, n.2, p.149-164, 2020.

[19] ISSAD, H. A.; AOUDJIT, R.; RODRIGUES, J.J.P.C. A comprehensive review of Data Mining techniques in smart agriculture. Engineering in Agriculture, **Environment and Food**, v.12, n.4, p.511-525, 2019.

[20] RIQUENA, R. S.; PEREIRA, D. F.; VALE, M. M.; SALGADO, D. D'A. Mortality prediction of laying

hens due to heat waves. **Revista Ciência Agronômica**, v. 50, n. 1, p. 18-26, 2019.

[21] THOM, E. The discomfort index. **Weatherwise**, 1959, vol. 12, N° 1, p. 57-60.

[22] UBA - UNIÃO BRASILEIRA DE AVICULTURA. **Protocolo de bem-estar para aves poedeiras**. São Paulo: UBA, 2008.

[23] VALE, M.M.; MOURA, D.J.; NÄÄS, I.A.; PEREIRA, D.F. Characterization of Heat Waves Affecting Mortality Rates of Broilers Between 29 Days and Market Age. **Brazillian Journal of Poultry Science**. v.12, n.4, p.279-285, Oct-Dec 2010.

[24] **MANUAL DE MANEJO DAS POEDEIRAS DEKALB WHITE**. GRANJA PLANALTO, Minas Gerais. Disponível em: < https://www.fcav.unesp.br/Home/departamentos/zootecnia/NILVAKAZUESAKOMURA/manual_dekalb_white.pdf > Acesso em 20 de setembro de 2019.

[25] VALE, M.M.; MOURA, D.J.; NÄÄS, I.A.; OLIVEIRA, R.M.; RODRIGUES, L.H.A. Data mining to estimate broiler mortality when exposed to heat wave. **Sci. Agric.** (Piracicaba, Braz.), v.65, n.3, p.223-229, May/June 2008.

[26] LIMA, E. S.; SOUZA, Z. M.; MONTANARI, R.; OLIVEIRA, S. R. M.; LOVERA, L. H.; FARHATE, C. V. V. Classification of the initial development of eucaliptus using data mining techniques. **CERNE**, Lavras, v. 23, n. 2, p. 201-208, Jun 2017.

[27] HAN, J.; KAMBER, M.; PEI, J. **Data Mining: concepts and techniques**. 3rd edition. San Francisco: Morgan Kaufmann Publishers. 2011.