EUROPEAN ACADEMIC RESEARCH
Vol. X, Issue 11/ February 2023

Impact Factor: 3.4546 (UIF)
DRJI Value: 5.9 (B+)

ISSN 2286-4822
www.euacademic.org

Data Mining and Construction of a Reference Database of Physicochemical Water Parameters for Anomaly Detection in the Measurements in the Amazonas River

NETO, MANOEL FERREIRA
CARDOSO, JÉSSICA FARIAS
Software Engineer of Embedded Systems Laboratory at Amazonas State University, Brazil
RENDEIRO, MANOEL FERNANDES BRAZ
Master in Computer Networks of Embedded Systems Laboratory at Amazonas State University, Brazil
CRUZ, LUCAS FARIAS DE
Graduating in Information Systems at the Federal University of Amazonas, Brazil
SOUZA, PEDRO FARIAS GÓES DE
Software Engineer of Embedded Systems Laboratory at Amazonas State University, Brazil
SOUZA, JOSE CAMILO DE
ALBUQUERQUE, CARLOSSANDRO CARVALHO DE
PhD in Geography of Embedded Systems Laboratory at Amazonas State University, Brazil
BATISTA, IEDA HORTÊNCIO
PhD in Biology of Embedded Systems Laboratory at Amazonas State University, Brazil
TEIXEIRA, THIAGO ALMEIDA
Electronics Engineer of Embedded Systems Laboratory at Amazonas State University, Brazil
MONTEIRO, GABRIELLA RABELO
Undergraduate student in Computer Engineering at Amazonas State University, Brazil
CISNEROS, EDRY ANTONIO GARCIA[1]
Doctor, Professor of Mechanical Engineering of Embedded System Laboratory at Amazonas State University, Brazil
SILVA, LUANA PAULA DA SILVA
Bachelor of Applied Mathematics of Embedded Systems Laboratory at Amazonas State University, Brazil

Abstract

The use of context in remote monitoring is a necessity in any area of research. Currently, the identification of anomalies of physicochemical parameters in the Amazon River is performed manually, which generates delays in the comparison and analysis of data for reference. The need to create a database from defined boundaries in conjunction with the use of data mining techniques can help for a more effective and faster analysis. This paper aims to specify, construct, and validate a database acquired from manual collections in the Amazon River. As a result we obtained a database that was worked and tested using the concepts of Percent Split and Cross Validation with an accuracy of 93.60% of the J48 algorithm and that obtained the best performance from the tests performed, being then proposed for use in the detection of anomalies of physical-chemical parameters of water in measurements made in the Amazon River.

Keywords: Remote Monitoring, Data Mining, Physicochemical Parameters of the Water, Anomaly Detection, Split Percentage, Cross Validation.

1. INTRODUCTION

Water resources are essential to the entire world population, so we must find ways to improve their management and, if possible, innovate. To this end, discovering new ways

[1] Corresponding author: edry1961cu@gmail.com

to monitor the quality of these waters and prevent problems with this essential resource are priority actions. We must take care of the water due to its great significance for human existence (GEROLIN, 2018).

According to Sales et al. (2014), numerous factors lead to the deterioration of water quality, and it is possible to classify its sources as point and diffuse. Point sources correspond mainly to wastewater (domestic and industrial), while diffuse sources include agricultural waste (fertilizers, herbicides, pesticides, fungicides, and others).

However, there are different ways to analyze water quality, such as by checking water parameters, physical and chemical. According to Santos et al. (2014), identifying anomalies in the physical-chemical parameters of river water and identifying their peaks is performed manually and often requires a certain amount of time for data comparison, depending on the region.

Rodrigues (2006) states that even though the tools for data analysis have evolved, the diagnosis of anomalies in physical-chemical parameters has not followed this process.

For this reality, it is possible to use data mining to analyze these collected physical-chemical parameters. Data Mining is a tool of great relevance to analyze large volumes of data with different algorithms, each of which implements its techniques.

In this research, we proposed the mining of a database of physical-chemical parameters of the water of the Amazon River, built from a historical series of analyses performed by two pieces of equipment: a commercial multiparameter probe composed of sensors (pH, electrical conductivity, dissolved oxygen, total dissolved solids, salinity, water temperature, and air temperature), analyzed directly in the water (except for the last parameter), and a turbidimeter where the analyses take place outside the water, with sample collection, which provides the turbidity level of the river.

The data were added to the WEKA tool for analysis and choice of algorithms with better performance and efficiency to create a classification model to identify anomalies and detect, in advance, unacceptable patterns in the water quality of the Amazon River.

The methodological path began with a literature review to build the initial topics of the article, followed by an exploratory study of the database collected (historical series), seeking to show that it is possible to perform data mining and develop the reference model for anomaly detection to complement the information on the analysis of water quality.

This paper has the following division. In section 2, the literature review takes place. In section 3, we present the materials and methods containing: the description of the equipment used, the process flow, and the anomaly identification method. In section 4, the results and discussions that make up this study. Finally, in section 5, we present our conclusions based on the success percentages of the selected algorithm and the reference base created to verify the anomalies detected in the readings of physical-chemical water quality parameters collected in the Amazon River, followed by proposals for future work.

2. LITERATURE REVIEW

2.1 Data mining

Data Mining is a procedure of revealing new correlations, models, and trends from the selection of large amounts of data stored in databases with specific content, which uses current and recognized tools of informational analysis for the identification of patterns, as well as statistics, and mathematics (KRIVDA, 1996). According to MacLennan et al. (2011), Data Mining has several techniques, such as classification, regression, segmentation, association, projection, and anomaly detection.

Data mining also enables the use of the Cross Validation approach on the database, as well as applying the Split Percentage approach to create training and testing bases with classification techniques (HALL et al., 2009).

The Split Percentage consists in separating the database into training subsets with the percentage of the base defined during the pre-processing phase, to train pattern recognition in search of improved statistical results, making a comparison of different pieces of training of the base (HALL et al., 2009).

Cross Validation consists in performing n repetitions of i iterations of subsets of the database provided in the input called folds, where the number of pairs in the training and testing set is the folds. The number of folds influences the outcome of the tests and should be alternated to higher or lower in search of the best result (HALL et al.,2009).

The approaches can be performed in systems such as WEKA (Waikato Environment for Knowledge Analysis), which is recognized by the scientific community as a reference software in machine learning and data mining (WITTEN; FRANK, 2005). It consists of a set of algorithm implementations of various machine learning techniques and is implemented in the Java programming language, making it accessible on all major computing platforms.

WEKA includes algorithms for regression, classification, clustering, association rules, and parameter selection (attributes). Currently, it is in version 3.8.6 and to use it, the data must be converted to one of the file formats supported by the system. In this work, the format adopted is that of WEKA itself, called ARFF (Attribute Relationship File Format).

3. MATERIALS AND METHODS

This section presents the equipment used, the system architecture, and the anomaly analysis method.

3.1. Study area and physicochemical parameters

The study area of the research was the Amazon River, more specifically the part of the river that passes through the port of Parintins, in the state of Amazonas - Brazil. The collections were made between the months of February and September 2022.

Resolution 357/2005 of the National Council of the Environment (CONAMA), which regulates physical-chemical analysis in Brazil, was used as a reference to measure the water quality in the port area of Parintins. The physical-chemical water

quality parameters analyzed in the Amazon River, with their measurement units, were: Hydrogen Potential (pH), Dissolved Oxygen (ppm), Electrical Conductivity (µs/cm), Air Temperature (ºC), Water Temperature (ºC), Turbidity (NTU), Total Dissolved Solids (ppm) and Salinity (PSU) (BRASIL, 2005).

3.2 Equipment used for the measurement of physical-chemical parameters

1) Portable Turbidimeter: we used the Hanna® turbidimeter, model HI98703, which according to the equipment instruction manual, it was developed for reliable and accurate water quality measurements for low turbidity values. It has an algorithm that calculates and converts the readings into NTU (Nephelometric Turbidity Units), in a range from 0.00 to 1000 NTU, faithfully following the reading criteria according to the EPA (United States Environmental Protection Agency) (HANNA 2020).

2) Portable probe: the Hanna® multiparameter probe, model HI98194, was used. This is a water-resistant meter, ideal for field work (lakes, rivers and oceans). It originally has four sensors: hydrogen potential (pH)/redox potential (ORP), electrical conductivity (EC), dissolved oxygen (DO), and temperature, in addition to an integrated barometer for DO concentration compensation. These sensors monitor up to 12 water quality parameters, 6 measured and 6 calculated (HANNA, 2020).

The sensors are color-coded (pH - red; EC - blue; DO - white), and the inputs are identified with colored triangles, making them easy to identify and install. The main operating modes of the HI98194 probe are measurement, recording, and configuration. The measurement screen can display one parameter or up to 12 at once, depending on the configuration. The parameter measurement units are user selectable, for example, for the standard temperature unit of the equipment, you can select an option other than °C (default), such as: °F and K. This possibility of changing the default unit can be applied to the other existing parameters.

The measurement ranges of the parameters read by the probe, according to standard units, are: temperature from -5 to 55 °C; pH/mV from 0 to 14 pH ± 600 mV; ORP 2000 mV; DO from 0 to 500 % and 0 to 50 ppm; conductivity 0 to 200 µs/cm; resistivity from 0 to 999999 Q-cm; TDS (total dissolved solids) from 0 to 400000 ppm; salinity from 0 to 70 PSU; seawater sigma from 0 to 50 σ; atmospheric pressure from 8. 702 to 16,436 psi.

3.3 Database

For the analysis of the collected data, it was necessary to use a level that determines the alteration in the physical-chemical parameters in the waters of the Amazon River. For this it was necessary to perform a learning process in Data Mining, called algorithm learning, where the system is shown the rule that it will use in its decision tree algorithms or NaiveBayes algorithm to show us the expected result.

With this the database was generated from a historical series of collections made manually in the Amazon River, presenting two states according to each parameter collected: Normal (within the expected range for the river) and Anomaly (outside the expected range for the river). The collections were performed mostly in the morning, on some days of the week, according to a schedule, but there was even uninterrupted collection for up to 12 hours, with 1-hour intervals, as part of this follow-

up. Table 1 below shows the number of times each state was identified according to the behavioral pattern matrix generated in the WEKA tool.

Table 1 Values of behavioral patterns

Label	Behavioral Pattern	Values Obtained
A	Normal	224
B	Anomalia	1781
C	Indefinido	137

Image Credit: Autor's

The values in the table above are labels applied by the WEKA tool from the database we used, where A refers to the data within the acceptance range "Normal", B are the data outside the defined ranges, the "Anomalies". And finally, the label C refers to data of "undefined" values, where the tool could not initially identify the data as "Normal" or "Anomaly".

The experiments were set up in two ways, first using the Split percentage approach with the database split in two, later the same experiments were done using the cross-validation approach.

3.4 Process Flow

The Process Flow is divided into three modules, where for each one the development of an accomplished process step will occur. The modules are described in figure 1 below:

Figure 1: Processes Flow

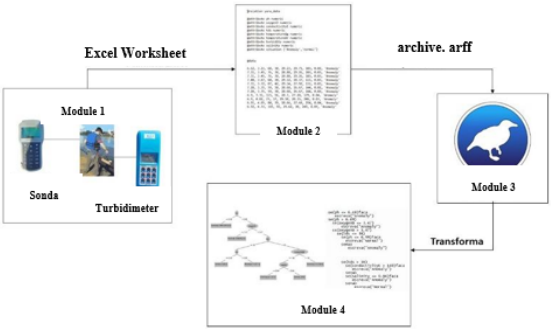

Image Credit: Autor's

Module 1 consists of manual collections performed by the chemistry team using a probe and a turbidimeter. The collection of physical-chemical parameters occurred on alternate days at the port, at a specific point of access to the Amazon River. The data were recorded in an Excel spreadsheet and sent to the next module.

Module 2 consists of receiving the data from module 1 and converting it into files that store the described data in a list of instances that have a shared set of attributes.

Module 3 uses the algorithms of the WEKA tool for performance evaluation regarding the detection of anomalies in the monitored parameters, in checking the accuracy and precision of matching by these algorithms.

Module 4 receives the data provided in the format of: decision tree, pseudocode and confusion matrix. These results will be used to make a decision regarding the algorithm to be chosen and will serve to develop a baseline for anomaly detection after this mining.

4. RESULTS AND DISCUSSION

In the first experiments, initial tests were performed with four classifiers, being three of the tree type (J48, RandomTree and RepTree) and one of the bayes type (NaiveBayes) using the same database with 2,142 data without filters and without adjustment. The trials were set up according to Hall et al. (2009) in two ways, first using the cross-validation approach and then the Split percentage approach with splitting the database into two: training and testing.

The WEKA data mining tool was used to assist in generating the decision tree model to be used in the application. Figure 2 shows the structure of the data in a text file with extension ".arff", used by the mining tool to extract the features from the database.

The file used by the mining tool has an annotated structure. The character "@" indicates the beginning of the attribute declaration. In line 11, the classification attribute determines, for example, the pre-classified data Anomaly or Normal mode. Here we recall that these two states are linked to the references used to analyze the values, coming from the CONAMA resolution No. 357/2005, where whenever they exceed the established limits they will be classified as Anomaly and otherwise they will be classified as Normal.

Figure 2: Data Structure

```
@relation yara_data

@attribute ph numeric
@attribute oxygenD numeric
@attribute conductivityE numeric
@attribute tds numeric
@attribute temperatureAg numeric
@attribute temperatureAr numeric
@attribute turbidity numeric
@attribute salinity numeric
@attribute situation {'Anomaly','normal'}


@data

6.62, 3.21, 68, 30, 29.23, 29.71, 103, 0.02, 'Anomaly'
7.31, 3.45, 76, 38, 28.88, 29.26, 103, 0.03, 'Anomaly'
7.31, 3.45, 76, 38, 28.88, 29.26, 103, 0.03, 'Anomaly'
7.40, 3.87, 60, 30, 29.12, 28.37, 115, 0.03, 'Anomaly'
7.31, 3.72, 87, 42, 29.14, 27.92, 131, 0.03, 'Anomaly'
7.28, 3.35, 59, 30, 28.68, 26.67, 144, 0.02, 'Anomaly'
7.28, 3.35, 59, 30, 28.68, 26.67, 144, 0.02, 'Anomaly'
6.9, 3.31, 115, 56, 28.3, 27.28, 179, 0.04, 'Anomaly'
6.9, 4.82, 73, 37, 29.58, 28.35, 246, 0.03, 'Anomaly'
6.91, 4.85, 80, 39, 28.66, 27.68, 254, 0.04, 'Anomaly'
6.92, 4.33, 183, 92, 29.62, 30, 249, 0.07, 'Anomaly'
```

Image Credit: Autor's

Each line marked from the @data annotation, line 16, corresponds to an instance passed to the classifier to train and generate the decision tree models and rules. The example in Figure 3 teaches the algorithm to identify the Anomaly classification.

Figure 3: Data Instance Example

```
@data

6.62, 3.21, 68, 30, 29.23, 29.71, 103, 0.02, 'Anomaly'
```

Image Credit: Autor's

The parameters to evaluate the quality of the application were the attributes that can be extracted from the Confusion Matrix automatically generated by the mining tool. The evaluated characteristics were: True Positive (VP), True Negative (VN), False Positive (FP) and False Negative (FN). These attributes were applied to the metrics used to evaluate the quality of the proposed method which is described in the formulas in Figure 4 below:

Figure 4: Formulas

$$Accuracy = \frac{TP + VN}{Total}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

Image Credit: Autor's

For the validation experiments we used cross-validation with K fold, which is a computational technique that uses all available samples as training and test samples.

The cross-validation database has a total of 10,020 records and by setting k=5 the database was partitioned into 5 subsets, where each subset will have approximately 2,004 records each. After the partitioning, one subset was used to validate the model and the remaining sets were used as training. The cross-validation process was then repeated K (5) times, so that each of the K subsets was used exactly once as a test to validate the proposed model.

The goal was to choose the most suitable algorithm to solve the problem by analyzing the issues of accuracy, runtime, and computational cost. Table 2 shows the data and features evaluated on the four algorithms using the cross-validation method.
With the feature analysis, it was observed that the NaiveBayes algorithm had the absolute and quadratic processing error higher than the other algorithms used: J48, RandomTree and RepTree.

After cross-validation it was found that the tree algorithms showed a much more significant response, because this model uses a graph-based theory, thus allowing us to develop a structure for decision making from the labels inserted in our database.

Table 2: Result with cross-validation

Algoritm	Time(s)	Absolute Erro	Quadratic erro	Accuracy	Precision	Recall	F1 Score
J48	0	0.1191	0.2442	0,927	0,929	1,0	0,96
RandomTree	0.02	0.1247	0.2567	0,927	0,929	1,0	0,96
RepTree	0.03	0.1191	0.2441	0,927	0,929	1,0	0,96
NaiveBayes	0.01	0.2397	0.3881	0,877	0,888	0,877	0,88

Image Credit: Autor's

The Naivebayes algorithm is based only on the repetition of data patterns, totally ignoring the correlation of variables which is the main feature used to create decision trees. But because it performs a probabilistic classification of observations, its hit rate stood out from the other Bayes-type algorithms, and because of this it was used in the experiments.

For the Splits Percentage experiments, the algorithms were fragmented into 50%, 70% and 80% splits for training the base. Table 3 shows the results using 50% of the database.

Table 3: Test results using 50% of the database for training

Algoritm	Time(s)	Absolute Erro	Quadratic erro	Accuracy	Precision	Recall	F1 Score
J48	0	0.1184	0.2472	0,925	0,927	1,0	0,96
RandomTree	0	0.1286	0.2651	0,925	0,927	1,0	0,96
RepTree	0.01	0.1183	0.2473	0,925	0,927	1,0	0,96
NaiveBayes	0.02	0.2401	0.3518	0,915	0,895	0,915	0,90

Image Credit: Autor's

Observing Table 3, it can be seen that using 50% for training and 50% for testing, the J48, RandomTree and RepTree algorithms are compatible. However, the square error of the NaiveBayes algorithm is higher than the others, which for this amount of data is already possible to identify differences.

We also split the data with 70% for training and 30% for testing indicated by the literature, where the results show a significant change. The algorithms J48, RandomTree and RepTree have the same number of hits of 92.9% as can be seen in Table 4 below:

Table 4: Test results using 70% of the training database

Algoritm	Time(s)	Absolute Erro	Quadratic erro	Accuracy	Precision	Recall	F1 Score
J48	0	0.1166	0.2411	0,927	0,929	1,0	0,96
RandomTree	0	0.1254	0.2553	0,927	0,929	1,0	0,96
RepTree	0.02	0.1165	0.2414	0,927	0,929	1,0	0,96
NaiveBayes	0	0.2274	0.3564	0,895	0,877	0,895	0,88

Image Credit: Autor's

You can also see that the execution time of the NaiveBayes algorithm was equal to the J48 and RepTree algorithms. The RepTree algorithm had a higher value in response time, but the accuracy remained the same as the tree algorithms.

To finish the analysis of the Splits Percentage approach, the data from the tests with 80% of the database for training and 20% for testing are shown in Table 5 below.

Table 5: Test results using 80% of the database for training

Algoritm	Time(s)	Absolute Erro	Quadratic erro	Accuracy	Precision	Recal 1	F1 Score
J48	0	0.1228	0.2586	0,915	0,917	1,0	0,95
RandomTree	0	0.1306	0.2727	0,915	0,917	1,0	0,95
RepTree	0.01	0.1225	0.2586	0,915	0,917	1,0	0,95
NaiveBayes	0	0.2276	0.3626	0,895	0,869	0,895	0,88

Image Credit: Autor's

In addition to evaluating the number of correct classifiers and the execution time, the mean square error was also evaluated as a first metric for a possible choice of the algorithm that best fits the decision problem addressed in the experiments.

By analyzing the Percent Splits tests performed with different parameters, it is possible to infer that the tests performed with 50%, 70% and 80% of training base obtained equal accuracy results and that the decision tree-based algorithms are the most suitable for the proposed problem.

At the end of the analysis, it is possible to say that both in the Cross-Validation and Splits Percentage approaches, the NaiveBayes algorithm did not obtain good results in data classification and this is justified due to the concept of the algorithm itself that does not work with data correlation which is essential in the correct classification of the spreadsheet base. In this case the J48 algorithm, RandomTree and RepTree appear with 93.60% the highest value achieved among the algorithms checked.

In Table 6 below, we show a general comparison of the results obtained.

Table 6: Comparison of Analyzed Data

Algoritm	Percentage (%)
J48	93.60
RandomTree	93.60
RepTree	93.60
NaiveBayes	80.57

Image Credit: Autor's

The algorithms J48, RandomTree and RepTree, obtained the same hit percentage in their classifications, where it can be observed that the tree algorithms are ideal for this type of application, because besides maintaining a good hit percentage they still offer a decision tree presented in Figure 8.

Figure 8: Decision Tree

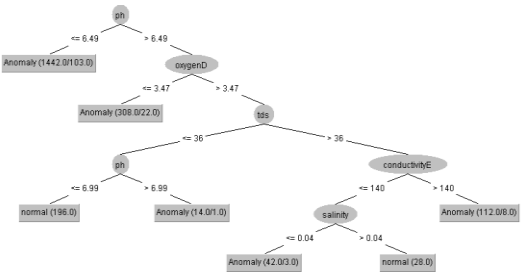

Image Credit: Autor's

From the creation of the decision tree an algorithm was implemented, the pseudocode in Figure 9.

Figure 9: Pseudocode

```
se(ph <= 6.49)faca
   escreva("Anomaly")
se(ph > 6.49)
  se(oxygenD <= 3.47)
     escreva("Anomaly")
  se(oxygenD > 3.47)
    se(tds <= 36)
      se(ph <= 6.99)faca
        escreva("Normal")
      senao
         escreva("Anomaly")


      se(tds > 36)
        se(conductivityE > 140)faca
           escreva("Anomaly")
        senao
        se(salinity <= 0.04)faca
           escreva("Anomaly")
        senao
            escreva("Normal")
```

Image Credit: Autor's

With this decision tree made by algorithm J48 in WEKA tool, its pseudocode was the one that had the best performance in all test splits of 50%, 70% and 80%, with better response time and less quadratic errors, which became a differential to the other tree algorithms, since the others obtained the same percentage of accuracy. With this, we can implement an algorithm that as the data is collected it already identifies if there are anomalies in the collection of parameters according to the base training performed and the CONAMA specifications used.

5. CONCLUSIONS

Brazil is a country that has great water resources, in this context, the analysis of the quality of these waters, such as the Amazon River, is a vital activity in order to ensure that this resource will be available for future generations. In this sense, seeking to contribute to this process, the study points to the use of data mining techniques for a better analysis of the physical and chemical parameters collected, in the detection of anomalies, bringing more reliable data for decision making.

The data of physical-chemical parameters collected from the waters of the Amazon River, in Parintins, by a turbidimeter and a commercial multi-parameter probe, generated a historical series that served as a basis for this research. From this base and the limits established by CONAMA, the concepts of Data Mining were applied, using the WEKA tool, in the Split-Percent and Cross-Validation classification approaches, where it was possible, at the end of a set of tests, to choose the most efficient algorithm for the analysis and identification of anomalies.

The J48 decision tree algorithm, from the WEKA tool, implemented a solution that best fits the measurements of the collected data, identifying changes in the parameter data according to the training of the base run and the specifications of the limits used. Thus, with the use of this algorithm it is possible to analyze databases of this format faster, saving time in the analysis of physicochemical parameters in our region and detecting anomalies in the collections more efficiently.

REFERENCES

BRASIL. Ministério do Meio Ambiente. Conselho Nacional de Meio Ambiente 'Resolução CONAMA nº 357, de 15 de junho de 2005'. Disponível em: http://pnqa.ana.gov.br/Publicacao/Resolução_CONAMA_n_357.pdf. Acesso em: 15 out. 2022.

GEROLIN, P. H., & RECCO, C. 'Mineração de Dados na Gestão de Recursos Hídricos Subterrâneos: Estudo de caso'. 2018.

HALL, M., FRANK, E., HOLMES, G., PFAHRINGER, B., REUTEMANN, P., & WITTEN, I. H. 'The weka data mining software: an update'. *ACM SIGKDD explorations newsletter*, v. 11, n. 1, p. 10-18, 2009.

HANNA, 'Instruments. Manual de instruções: hi 98194, hi98195, hi98196. HI 98194, HI98195, HI98196'. 2020. Disponível em: https://hannainst.com.br/wp-content/uploads/2020/06/Cat%C3%A1logo-Multipar%C3%A2metros-HI98194-HI98195-e-HI98196-Hanna-Instruments-Brasil.pdf. Acesso em: 10 set. 2022.

KRIVDA, C. D. 'Unearthing Underground Data'. *LAN – The Network Solutions Magazine*. Cidade, v. 11, n. 5, p. 42-48, mai.1996.

MACLENNAN, J., TANG, Z., & CRIVAT, B. (2011). 'Data Mining with Microsoft SQL Server', 2008. Wiley.

RODRIGUES, M. I. M. de C. 'Agricultura peri-urbana e ecossistemas mediterrânicos', 2006. Tese de Doutorado. FCT-UNL.

SALES, L. M., PRADO, R., & GONÇALVES, A. (18-20 de novembro de 2014). 'Análise comparativa entre sondas multiparamétricas para'. *Simpósio Nacional de Instrumentização Agropecuária*.

SANTOS, A. D. F. dos et al. 'Proposta de gerenciamento de dados para monitoramento de saúde estrutural utilizando redes de sensores ópticos FBG'. 2014.
WITTEN, I. H., & FRANK, E. (2005). 'Data mining: practical machine learning tools and techniques. 2nd Edition, Morgan Kaufmann Publisher, Burlington.

INFORMATION ABOUT AUTHORS

MANOEL FERREIRA NETO
Affiliation: Amazonas State University
Graduated in Software Engineering - UFAM; Technician in Informatics - IFAM, conclusion in 2014; currently attending post-graduation course in Control and Automation Engineering - Unopar.

JÉSSICA FARIAS CARDOSO
Graduanda do Curso de Engenharia de Software na Universidade Federal do Amazonas (UFAM). Participação em eventos científicos como: Semana Nacional de Ciências e Tecnologia (SNCT) e Semana de Informática CESIT/UEA.
Affiliation: Amazonas State University

MANOEL FERNANDES BRAZ RENDEIRO
Affiliation: Amazonas State University
Graduated in Data Processing from the University of Amazonia - UNAMA (1997), Specialist in Computer Networks from the University of Amazonia - UNAMA (1999) and Master in Science Education in the Amazon from the Amazonas State University - UEA (2014). He has worked in several Federal, State and Municipal agencies in technical areas related to Computer Science, with emphasis on Computer Networks, as a specialist and in the area of Education he worked as a temporary teacher of Technical and Higher Education until 2008. Since 2009, he works as a teacher at the Amazonas State University - UEA, at the Centro de Estudos Superiores de Parintins - CESP, performing his function as an effective public servant of this State Higher Education Institution. He is a researcher member of the Group of Study and Research in Science Education in Non-Formal Spaces - GEPECENF and the Group of Studies and Research in Mathematics and Technologies - COMPLEXUS. He develops research in the areas of Information and Communication Technologies - ICT, Mathematics Teaching, Scientific Dissemination, Non-Formal Spaces and Science Education.

LUCAS FARIAS DA CRUZ
Affiliation: Amazonas State University
Graduating in Information Systems (UFAM). Experience in extension projects, scientific initiation and research and development. Participated in research laboratory and the opportunity to work with industrial automation using Ladder language and SCADA (Supervisory Control And Data Acquisition System), creating system for controlling motors and valves. Has scientific initiation projects with Bio-inspired Algorithms and application development for analytical chemistry.

PEDRO FARIAS GÓES DE SOUZA
Affiliation: Amazonas State University
Graduated in Software Engineering at the Federal University of Amazonas - UFAM. Post-graduating in Information Systems Architecture at Centro de Estudos de Especialização e Extensão - CENES.

JOSÉ CAMILO RAMOS DE SOUZA
Affiliation: Amazonas State University
Graduated in Geography at the Federal University of Amazonas (1995), Bachelor's Degree in Geography at the Federal University of Amazonas (1998), Specialization in Management in Ethno-Development at the Federal University of Amazonas (2002 - 2003), Master's Degree in Education at the Federal University of Amazonas - FACED (2004 - 2006) and doctorate in Sciences obtained in the Geography Program (Physical Geography) - Area of Concentration: Physical Geography, at the University of São Paulo-USP, on 04/07/2013. Professor at the Amazonas State University. Has experience in the area of Geography and Education. Geography: Economic Geography, Agrarian Geography, Teaching Methodology in Geography, Teaching Practice in Geography, Geography of Tourism, Cartography applied to the teaching of Geography etc. Education: Supervised Internship, Curriculum, Research Methodology and Studies. Study on Japanese immigrants in the Amazon, Amazonian riverine people, and Amazonian geographic thought. Study on Water Management and Governance and regulation of water resources.

CARLOSSANDRO CARVALHO DE ALBUQUERQUE
Affiliation: Amazonas State University
PhD in Geography from the Federal University of Ceará (2012). Master in Environmental Sciences and Sustainability in the Amazon UFAM (2003). Specialization in Environmental Engineering - UFAM (1998). Degree in Geography from the Federal University of Amazonas - UFAM (1997). Adjunct Professor at the Amazonas State University - UEA. Professional experience in Geography, Environment, Water Resources and Tourism. Coordinator of the Professional Master in Management and Regulation of Hydric Resources. Works in higher education and research in the master's course and graduation in geography, water resources, environment and tourism.

IEDA HORTÊNCIO BATISTA
Affiliation: Amazonas State University
Post-doctorate in Environmental Education from the Federal University of Ceará (UFC - 2010), PhD in Biotechnology from the Federal University of Amazonas (UFAM - 2009), Master in Environmental Sciences and Sustainability in the Amazon (UFAM - 2000), specialization in Biotechnology for Sustainable Development (UFAM - 2000) and graduation in Biological Sciences (UFAM - 1996). Currently she is an Adjunct Professor at the Amazonas State University (Universidade do Estado do Amazonas), teaching Biological Sciences. She is a professor in the Graduate Program in Management and Regulation of Water Resources. She has experience in General Biology, with emphasis on Cellular and Molecular Biology, Environment and Environmental Microbiology. She works mainly on the following topics: Environment, Water Resources, Cellular and Molecular Biology, microbiology and Biotechnology.

THIAGO ALMEIDA
Affiliation: Amazonas State University
Contact Information: tat.tai@uea.edu.br
Thiago Almeida Teixeira: Graduating in Electronic Engineering from the University of the State of Amazonas and in Quality Management from Laureat International, 8 years of experience in the area of process quality, 2 years at Moto da Honda da Amazônia as administrative apprentice in monitoring processes performed by third parties, 2 years at Requinte Eireli as administrative assistant, Currently working at HUB Technology and Innovation as a test developer and Scrum Master in technological projects in the R&D area.

GABRIELLA RABELO MONTEIRO
Affiliation: Amazonas State University
Undergraduate student in Computer Engineering - UEA, Mechatronics Technician - Nokia Foundation Conclusion in 2017. Experience in Web development, knowledge in Python, Javascript, C, Java. Currently student-researcher at the Amazonas State University.

EDRY ANTONIO GARCIA CISNEROS
Affiliation: Amazonas State University
Contact Information: ecisneros@uea.edu.br. http://lattes.cnpq.br/1105882720421386
He has a degree in Mechanical Engineering from the Volgograd State Polytechnic University, Russia (1985), a master's degree in Higher Education from the University of Camagüey, Cuba, (1999), a master's degree in Technical Sciences from the Volgograd State Polytechnic University, Russia (1985), a doctorate in Technical Sciences from the University of Holguín Oscar Lucero Moya, Cuba (2004) and a doctorate in Mechanical Engineering from the Federal University of Pará, Brazil (2016). He is currently a Professor at Marta Falcão Wyden College, Adrianópolis, Manaus-Amazonas and Voluntary Professor at the Amazonas State University. Has experience in the area of industrial maintenance and automotive transport; CAD projects of mechanical engineering applied to industry, energy use in the mechanical industry and automotive transport. Research Group Modeling and Intelligent Identification of Technological and Biotechnological Systems. Research lines: Industrial maintenance and automotive transport; CAD developments of machinery and equipment, efficient use of energy in transport and industry. Advised more than 70 students in undergraduate, 20 in masters and 1 in doctorate.
- Author of the paper:

1. CISNEROS, Edry Antonio Garcia. Rational Structure of the System of Machines in the Transport Process Rice Harvest. International Journal of Research in Engineering &Technology., Vol 6, page 9 - 19, 2019.
2. GARCIA, Cisneros Edry Antonio; PRINTES André Luis; SOUZA, Gomes Raimundo Claudio; CARDOSO, Fabio de Souza; FERREIRA, Sobrinho Angilberto Muniz; BARBOSA, Martins Karolayne; PEDRAÇA, Júnior Neirival Rodrigues; ABREU Furtado Diogo; DA COSTA, Barbosa Isaias; MARTINS, da Costa João Carlos; SICCO, de Oliveira João Victor Reis. Evaluation of Helical Springs Behavior Submitted to Compression. ISSN 2286-4822, ISSN-L 2286-4822. European Academic Research, International Multidisciplinary Research Journal. v. IX, ISSUE 12 MARCH 2022. p.6878 - 6891.

3. GARCIA Cisneros Edry Antonio; TORNÉ Israel Gondres, RAMÍREZ Neeldes Matos, PRINTES André Luiz, SOUZA Gomes Raimundo Cláudio, CARDOSO Fábio de Souza.Analysis of maintenance in mass transportation vehicles through world class maintenance indicators. Conjecturas, ISSN: 1657-5830, Vol. 22, No. 2.2022. DOI: 10.53660/CONJ-836-F19

LUANA PAULA DA SILVA E SILVA
Affiliation: Amazonas State University
Bachelor's degree in Applied Mathematics from the Federal University of Amazonas - UFAM. Graduated in Pedagogical Education in Mathematics at the Leonardo da Vinci University Center - Uniasselvi. Specialist in Mathematics Teaching Methodology - Uniasselvi. MBA in progress in Project Management and specialization in Industry 4.0 Management at Centro Universitário Leonardo da Vinci - Uniasselvi. Participated in the Project of Scientific Initiation, with the theme: Analysis of the premiums of the options market of cattle at BM&FBOVESPA. Has complementary training in the areas of People Management and Corporate Accounting. Experience in LaTeX, production and elaboration of technical reports. Currently she is a scholarship student at the development center HUB - Technology and Innovation, working on the preparation of technical reports of the PD&I projects.
Contact Information: luanapaula590@gmail.com