



Detection of the Primary User's Behavior for the Intervention of the Secondary User Using Machine Learning

Deisy Dayana Zambrano Soto¹, Octavio José Salcedo Parra^{1,2(✉)},
and Danilo Alfonso López Sarmiento¹

¹ Faculty of Engineering, Intelligent Internet Research Group,
Universidad Distrital "Francisco José de Caldas", Bogotá D.C., Colombia
ddzambranos@correo.udistrital.edu.co,
{osalcedo,dalopez}@udistrital.edu.co

² Department of Systems and Industrial Engineering, Faculty of Engineering,
Universidad Nacional de Colombia, Bogotá D.C., Colombia
ojsalcedop@unal.edu.co

Abstract. The predictive analysis for the spectral decision with automatic Learning is a task that is currently challenging. Some automatic Learning techniques are shown in order to predict the presence or absence of a primary user (PU) in Cognitive Radio. Four machine learning methods are examined including the K-nearest neighbors (KNN), the support vector machines (SVM), logistic regression (LR) and decision tree (DT) classifiers. These predictive models are built based on data and their performance is compared with the purpose of selecting the best classifier that can predict spectral occupancy.

Keywords: Cognitive Radio · KNN · LR · Machine learning · Primary users (PU) SVM

1 Introduction

Cognitive radio authorizes users with a license (PUs) to Exchange the exceeding spectrum and temporarily transfer the use of spectrum to non-licensed users (SUs). In consequence, the PU need a resource management scheme that allows them to optimally assign a fixed amount of spectrum offered between different types of services and adapt to the changes in the network conditions.

To detect its environment in a Cognitive Radio node, it examines the parameters in the air and makes decisions for the allocation and dynamic management of resources in time-frequency-space to improve the use of the radio spectrum. For an efficient real time process, cognitive radio is combined with artificial intelligence and automatic learning techniques to allocate adaptive and intelligently [1]. By using the spectrum detection techniques, the inactive channels are located and allocated to the secondary users.

Spectrum detection techniques are classified into three types such as: cooperative system, non-cooperative system and interference-based system [3] as seen in Fig. 1.



Fig. 1. Spectrum classification techniques (Author).

The advantages of cooperative techniques include: reduction of the sensibility requirements, minimization of costs and power, etc. The main disadvantage of this technique lies in the periodical demand in terms of the detection of the spectrum.

The centralized detection has a cognitive radio controller (Fig. 2). When the main user is found, the cognitive radio informs the controller (Fig. 2). Then, the controller transmits all the cognitive radio users over the main user. The centralized coordinated techniques have two types: partially cooperative techniques and fully cooperative techniques. In the partially cooperative techniques, all nodes cooperate between them on the channel detection while the fully cooperative technique the nodes cooperate between them both in the channel detection and in the transference of information between nodes [3].

The purpose of the cognitive controller is to coordinate the actions of the CR using Machine Learning algorithms. However, only in recent years, has there been a growing interest in the application of learning algorithms in Cognitive Radio [4]. The main challenge is the complexity and convergence of these techniques under a limited amount of time. With the purpose of reducing the complexity and achieving an efficient real time resource allocation, the CRs use automatic learning and artificial intelligence to make decisions based on models built using the cognitive radio's learning capacity. However, this is not complete or precise due to limited training data [2].

For machine learning in CR, the main steps involve the observation of the environment and analyzing its responses in terms of feedback, learning, keeping decisions and observations which update the model. This leads to highly accurate decision-making over management of resources and adjusting the transmission errors as can be seen in Fig. 3.

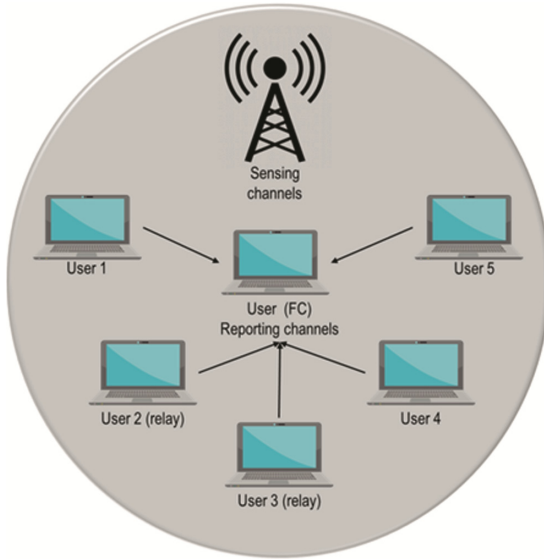


Fig. 2. Representation of the centralized spectrum detection coordinated with Cognitive Radio users (Author).

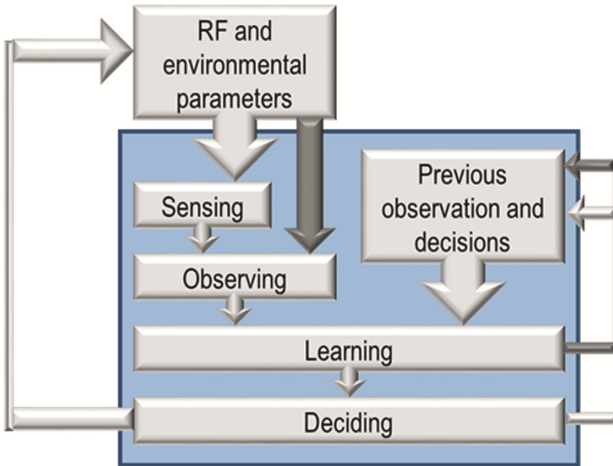


Fig. 3. Principales pasos de Machine Learning en RC [1].

The contributions and results of the document are hereby summarized: A new learning strategy is introduced to analyze spectrum behavior by applying ANFIS and time series to obtain prediction results based on comparisons of the obtained data.

The rest of the document is organized as follows: Sect. 2 offers the architecture of the CR model; the recollection process of data is described in Sect. 3, Sect. 4 presents

the ANFIS model, Sect. 5 shows the results and Sect. 6 discusses them. Conclusions are shown in Sect. 7.

2 Background

In the “Least-square support vector machine-based learning and decision making in cognitive radios” [5] a model for decision-making is exposed in a Cognitive Radio scenario based on learning and Support Vector Machines (SVM). For the methodology, the authors focus on pre-processing the data which implies their normalization, reduction, grouping and selection of the learning method which determines the setting parameters. A peer-to-peer wireless communication is built with orthogonal frequency division multiplexing (OFDM) technology. C is established as $C' = d(K_I, K_R, E', R')$ where $d(*)$ indicates the decision-making function, E' is the current condition of the environment and R' is the user requirement. The decision model or function uses the learned knowledge (K_I) and the previous knowledge (K_R) to obtain adequate setting parameters. By recollecting communication instances from MATLAB simulations and taking into account the calculated parameters, it is concluded that their selection and configuration is highly important in smart decision-making in CR systems so that there is no insufficient or excessive adjustment of data. This leads to a comparison-based multiclass classification method and a parameter search through Genetic Algorithms (GA).

In the article exposed in 2013 by Bkassiny et al., a variety of learning problems in CR is shown as well as the importance of the application in artificial intelligence in solving them; The two main categories of learning in CR are established such as decision-making (where rules and decision policies are made for CR) and classification (where different observation models are identified). The three basic conditions for intelligence in a Cognitive Radio system are: perception (obtained through the sensorial observations to classify and organize them into adequate categories) and finally reasoning from what has been learned. The main classification methods in the CRs are exposed such as Support Vector Machines used in pattern recognition and the classification of objects characterized by the absence of local minimum and the mapping of input vectors in a linear-based separated space of high dimension. The authors establish a protocol where every secondary user can independently bet on each main channel and the offer that saves the most energy is chosen. There is also a distributed algorithm where each secondary user updates his strategy based on local information in order to converge to the equilibrium point; this is seen within a decentralized Cognitive Radio architecture [4].

In the article called “Machine Learning Techniques with Probability Vector for Cooperative Spectrum Sensing in Cognitive Radio Networks” [6] a low dimension probability vector (n-dimensional energy vector in a CR network with one primary user and n secondary users) is proposed as a characteristics vector for classification based on automatic learning, improving the spectrum detection and establishing the duration of the training and the classification delay. By testing on some scenarios (varying the number of SU and simulating with some learning algorithms), the spectrum detection's accuracy is highlighted with short training duration and a lower classification time in comparison to the energy vector and the two-dimensional probability vector.

Sharma and Bohara in the article “Exploiting Machine Learning Algorithms for Cognitive Radio” discuss the application of Machine Learning algorithms (such as the Genetic Algorithm, the Artificial Neural Networks (ANN) and the Hidden Markov Models (HMM) in the resolution of specific problems in diverse CR scenarios; so the spectrum management can focus on different algorithms where prediction depends on the size of the data as well as the complexity of the space [7].

In “MAC Protocol Selection Based on Machine Learning in Cognitive Radio Networks”, Qiao et al. gather external and internal network parameters by building a dataset that will be trained with a classification algorithm. It will be used to decide which MAC protocol will adapt to the network’s current situation in dynamic networks; The candidates for classification protocols are the DFC (Distributed Function Coordination) as the competitive one and the TDMA (Time Division Multiple Access) as the non-competitive one. This assures that the established classifier can swiftly and robustly select an adequate MAC protocol. For each protocol, when the performance curve descends, the DFC classifies the samples as modifiable. When the curve has not reach the saturation region, the TDMA is modifiable.

The simulation results show that the average probability of correct classification (PoCC) of the proposed method exceeds the baseline. The classification model manages to choose the appropriate MAC protocol that fits the current circumstance of the network [8].

In an article published in 2017 [9], the authors present an architecture for the cognitive management of 5G networks with two approaches: ‘SLA enforcement’ (level of service) and ‘Mobile Quality Predictor’. The first approach applies LSTM (Long Short Term Memory) which basically works as an automatic learning technique that contains cells with three doors to manage the state of the memory (forget, remember and update). The purpose is to anticipate and make sure that recovery actions are followed which can be seen as an early alarm system. The second approach uses automatic learning to precisely predict the bandwidth in real time and improve the quality of the service; this establishes that although the strategy is likely, it requires some modifications in the mobile nodes since the servers must support the transport protocol.

For “Profit optimization in multi-service cognitive mesh network using machine learning”, the authors analyze the capacity of cognitive wireless mesh networking (CWMN) to maximize the PU’s income, maintain the QoS and attend to the majority of SU. The CWMN uses the spectrum rented from the PU to overlap its traffic. The inferred CWMN was modeled, analyzed and simulated. The authors propose a scheme so that primary users control the spectrum’s traffic for the spectrum’s emergent state. The objective was to adjust the size and the price of the spectrum in order to maximize the net income of the PU while still keeping the QoS of the PU. Simulations revealed the capacity of the algorithm to establish the requirements of the SU and achieve the potential profit through the application of cognitive radio. This confirms that the cognitive scenarios can have additional users without affecting the QoS. The capacity of this architecture was also revealed when maintaining the QoS for users by adapting the size and price of the offered spectrum under different conditions. The Primary Users share a spectrum based on demand so they need to borrow spectrum from their neighbors while complying with interference rules; the proposed scheme leads to higher earnings [10].

3 Methodology

The GSM 850 MGZ band is used in this study. The measures correspond to one day in the week of March 2016. Since the channel distribution goes from 128 to 151 for GSM, there are 124 channels. In this research, two days are analyzed which includes patterns of use in one day of the week and the weekend.

To keep track of the previously required data, specialized equipment was used. Records of the power of the absolute frequency of a radio-electric channel in dB mW obtained in 290 ms intervals throughout 15 h per day. For the aspects regarding this document, 35725 data will be averaged every 5 min, i.e. 1034 records will be averaged to determine the following probabilities [14].

3.1 Activity Percentage

Activity Probability

$$P(\text{Occupied}) = \frac{\text{Number of times occupied}}{\text{Total PU activities in the 0,300 s interval}}$$

Inactivity Probability

$$P(\text{Availability}) = \frac{\text{Number of times available}}{\text{Total PU activities in the 0,300 s interval}}$$

3.2 Permanence Probability

It is the probability of permanence within a time interval. It determines the percentage of occupied tendency and comes from averaging the activity and inactivity data where 40 represents the total number of users found.

$$P(\text{Activity Tendency}) = \frac{\# \text{ Duration tendencies}}{40}$$

$$P(\text{Inactivity Tendency}) = \frac{\# \text{ Duration tendencies}}{40}$$

3.3 Probability of the “Amplitude Node” (Presence or Absence of the PUs)

It is the probability of the “Amplitude” node to represent the presence or absence of the PUs which is represented as the probability of having a high value of P(High) and the probability of having a low value P(Low) based on a threshold established as 89. The percentage of high and low amplitude are defined as:

$$P(\text{High}) = \frac{\text{Average of high frequencies}}{\text{Total activities of the PU in the 0, 300 s interval}}$$

$$P(\text{Low}) = \frac{\text{Average of low frequencies}}{\text{Total activities of the PU in the 0, 300 s interval}}$$

The probabilities of the nodes “State”, “Amplitude” and “Occupation tendency” are generated in this way which complement the Machine Learning techniques established as follows.

K-nearest Neighbors

The K-nearest neighbors classifies an object based on the vote of the majority of its closest neighbors. In other words, the class of a new instance is predicted based on some distance metrics. The distance metric used in the nearest neighbors for numerical attributes can be a simple Euclidian distance [15]:

$$d(x, y) = \sqrt{\sum_{i=1}^k (x_i - y_i)^2}$$

The Euclidian distance $d(x, y)$ is used to measure the distance to find the closest examples in the pattern’s space [16].

Support Vector Machine

The vector machine models are defined as vectorial spaces of finite dimension where each dimension represents a “feature” of a particular object. It has been proven to be an efficient approach in highly-dimension spatial problems [15]:

$$\max Q(a) = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_i a_j d_i d_j x_i x_j$$

where $0 \leq a_i \leq C$ for $i = 1, 2, \dots, n$.

SVM uses the decision function $f(x)$ defined as a kernel function to calculate the output as:

$$f(x) = \text{sign} \left[\sum_{i=1}^i a_i d_i K(x, x_i) + b \right]$$

where $K(x, x_i)$ is the kernel function.

Decision Tree

A decision tree is a diagram flow structure includes a root node, branches and leaf nodes. The attributes of the dataset are defined with the internal nodes. The branches are the

result of every test against every node. The data is divided in classes based on the value of the attribute found in the training sample described in Algorithm 1.

Algorithm 1: Decision Tree

1. create a node N ;
 2. **if** tuples in D are all of the same class, C **then**
 return N as a leaf node labeled with the class C ;
 3. **if** $attribute_list$ is empty **then**
 return N as a leaf node labeled with the majority class in D ;
 4. apply **Attribute_selection_method**($D, attribute_list$) to find the "best" $splitting_criterion$;
 5. label node N with $splitting_criterion$;
 6. **if** $splitting_attribute$ is discrete-valued and multiway splits allowed **then**
 $attribute_list \leftarrow attribute_list - splitting_attribute$;
 7. **for each** outcome j of $splitting_criterion$
 let D_j be the set of data tuples in D satisfying outcome j ;
 if D_j is empty **then**
 Attach a leaf labeled with the majority class in D to node N ;
 8. **else** attach the node returned by **Generate_decision_tree** ($D_j, attribute_list$) to node N ;
 9. return N ;
-

The purpose of the decision tree is to from a set of rules to analyze a dataset sample with the purpose of making decisions on the classification of unknown data. The real operation flow of the decision tree algorithm is shown in Fig. 4.

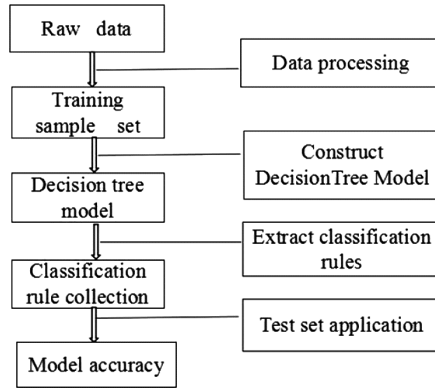


Fig. 4. Flow of the decision tree algorithm for Machine Learning [Taken from [11].

Logistic Regression

Logistic regression (LR) is a model where the result variable in the logic regression is binary and this is reflected both in the selection of a parametric model and the suppositions [12]. LR calculates the distribution between the example X and the boolean class label Y by $P(X)$. Logistic regression classifies the Boolean class label as:

$$P(Y = 1|X) = \frac{1}{1 + \exp(w_0 + \sum_{i=1}^n w_i X_i)}$$

$$P(Y = 0|X) = \frac{\exp(w_0 + \sum_{i=1}^n w_i X_i)}{1 + \exp(w_0 + \sum_{i=1}^n w_i X_i)}$$

4 Proposed Method

This article compares four types of mechanical learning techniques to predict the absence or presence of primary users (PUs). The proposed construction processes of the predictive models are shown in Fig. 5.



Fig. 5. Data training process [17].

This training flow includes two steps:

1. Data training which contains the data insertion where pre-processing is carried out through feature extraction. Then, the classification methods will be chosen to learn from our data. This must be iterative and regression-based since it requires to go back to the data pre-processing phase. Different automatic learning algorithms are tested with different parameters.
2. Using the data model where the pre-processing steps will be reused and the model will allow the prediction over the new data.

5 Results

The performance of the tests based on verisimilitude can be hard to estimate since large amount of data have to be simulated or recorded [13]. This research generates a confusion matrix which is used in Machine Learning to summarize the performance of classification algorithm. The calculation of a confusion matrix might be helpful during feature selection since it offers a better idea of what the classification method is doing

right and what type of errors are being committed regarding the classifiers [18]. This matrix can be seen for every technique used (Fig. 6):

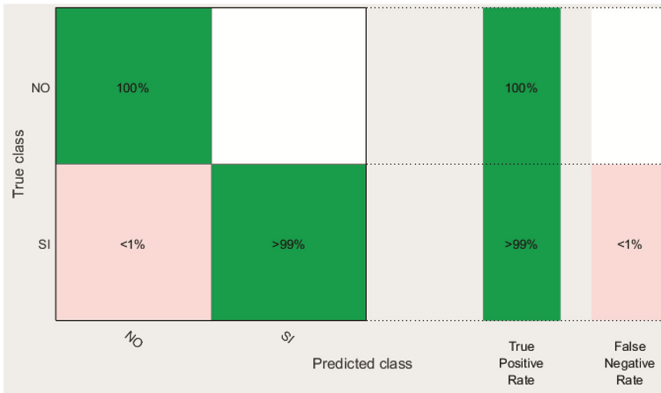


Fig. 6. Confusion matrix with the decision tree [Matlab by author].

The previous figure shows how the decision tree correctly classifies the channels that are available which represent close to 18593 observations while it makes 1 mistake over the 1873 observations in the occupied channels. However, the algorithm classifies does have a 100% classification rate for the GSM occupied channels. There is a clear difference with the SVM algorithm which wrongfully classifies 132 observations out of 18593 for the available channels. For the occupied channels, 1655 observations out of 1874 are correctly classified with a classification percentage of 98.3% as shown in Fig. 7.



Fig. 7. Confusion matrix using SVM [Made in MatLab by author].

In terms of the KNN and logistic regression techniques, percentages of 99.9% and 97.8% are established respectively which implies that the best classification technique for spectral occupancy are the decision trees followed by KNN, SVM and logistic regression where KNN classifies 11 channels as occupied within the 18593 observations.

This means that at some time instance where the channel is unoccupied, the algorithm establishes that it is not case. Out of 1873 occupied channels, it shows that 13 are unoccupied which is not true and can be seen in Fig. 8.

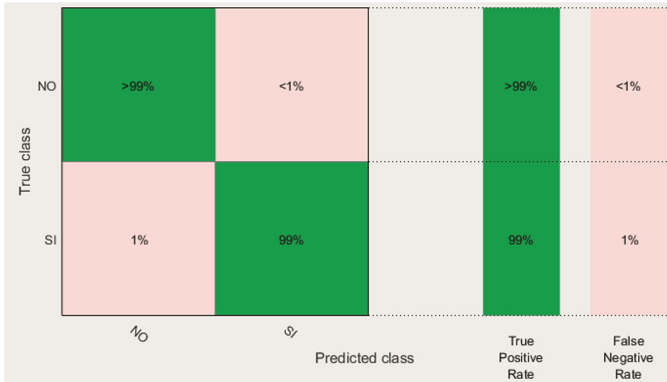


Fig. 8. Confusion matrix using KNN [Made in Matlab by author].

The technique that is least advised with a 97.8% accuracy is the Logistic Regression that detects 147 observations as occupied out of 18593 and 305 as available out of 1873 which implies a high risk in an embedded application. If the algorithm makes these types of mistakes, it will lead to interventions of the SU when the channel is really occupied by the PU. The confusion matrix for this technique is shown (Fig. 9).

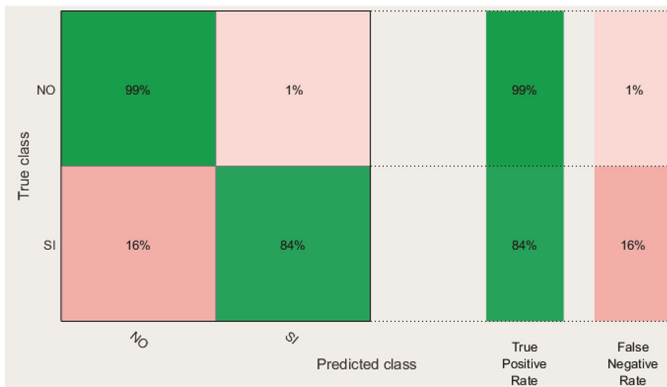


Fig. 9. Confusion matrix using Logistic Regression [Made in Matlab by author].

6 Discussion of Results

In consequence and according to [5] and [4], the Support Vector Machine algorithm is a good option to classify the occupation in Cognitive Radio channels with a 98.3% accuracy. As stated by [7], it is important to establish that spectrum management can be

approached with different algorithms by establishing the architectural needs in CR. The results show that the best classification is developed by the decision tree algorithm. However, in terms of the execution time of the algorithm, SVM, KNN and LR spend an average of 42.18 s which is a lot less than the decision trees which take about 184 s. This implies three times more machine time than the other methods. In agreement with [19] which show 95.96% for KNN and 87.34% for SVM, our methods show 99.9% and 98.3% respectively marking a formidable performance for both methods. If there is a rigorous analysis in terms of accuracy and execution time (resources on a machine level), the best option would be KNN with 99.9% accuracy, an execution time of 44.073 s and a prediction speed of 5900 observations/second.

7 Conclusions

In this work, four classification models are proposed with automatic learning techniques which are K-nearest neighbors (KNN), support vector machines (SVM), logistic regression (LR) and decision tree classifiers to predict the occupation of 124 GSM channels by establishing 7 predictors and 2 types of responses: whether the channel is occupied or not so that it can be used by the secondary user. Based on experimental results, the decision trees offer 100% accuracy on channel occupancy on behalf of the primary user which strictly accomplishes the initial goal of this study which is the classification of the GSM channels. The prediction speeds reach 46000 observations/second with a training time of 185.59 s. This research contributes to the development of embedded applications that can maximize the efficient use of the Cognitive Radio spectrum.

References

1. Abbas, N., Nasser, Y., El Ahmad, K.: Recent advances on artificial intelligence and learning techniques in cognitive radio networks. *EURASIP J. Wirel. Commun. Netw.* **2015**(1), 174 (2015)
2. Elhachmi, J., Guennoun, Z.: Cognitive radio spectrum allocation using genetic algorithm. *EURASIP J. Wirel. Commun. Netw.* **2016**(1), 133 (2016)
3. Senthilkumar, S., GeethaPriya, C.: A review of channel estimation and security techniques for CRNS. *Autom. Control Comput. Sci.* **50**(3), 187–210 (2016)
4. Bkassiny, M., Li, Y., Jayaweera, S.K.: A survey on machine-learning techniques in cognitive radios. *IEEE Commun. Surv. Tutor.* **15**(3), 1136–1159 (2013)
5. Wu, C., Yu, Q., Yi, K.: Least-squares support vector machine-based learning and decision making in cognitive radios, pp. 2855–2863, April 2012
6. Lu, Y., Zhu, P., Wang, D., Fattouche, M.: Machine learning techniques with probability vector for cooperative spectrum sensing in cognitive radio networks. In: *WCNC*, pp. 1–6 (2016)
7. Sharma, V.: Exploiting machine learning algorithms for cognitive radio, pp. 1554–1558 (2014)
8. Qiao, M., Zhao, H., Wang, S., Wei, J.: MAC Protocol selection based on machine learning in cognitive radio networks, pp. 453–458 (2016)
9. Grida, I., Yahia, B., Bendriss, J.: Cognitive 5G net works: comprehensive operator use cases with machine learning for management operations, pp. 252–259 (2017)

10. Alsarhan, A., Agarwal, A.: Profit optimization in multi-service cognitive mesh network using machine learning, pp. 1–14 (2011)
11. Liu, X., Li, B., Shen, D., Cao, J., Mao, B.: Analysis of grain storage loss based on decision tree algorithm. *Procedia Comput. Sci.* **122**, 130–137 (2017)
12. Hosmer Jr., D.W., Lemeshow, S., Sturdivant, R.X.: *Applied Logistic Regression*. Wiley Series in Probability and Statistics, pp. 15–19 (2013)
13. Lingenfelter, D.J., Fessler, J.A., Scott, C.D., He, Z.: Predicting ROC curves for source detection under model mismatch, pp. 1092–1095. *IEEE Explorer* (2010)
14. Ordoñez, J.: Caracterización de usuarios primarias para la implementación de un modelo predictor para la toma de decisiones en redes inalámbricas de radio cognitiva. In: 2016 Trabajo de Grado, Universidad Distrital Francisco José de Caldas (2016)
15. Pouriyyeh, S.: A comprehensive investigation and comparison of machine learning techniques in the domain of Herat disease. In: 2017 IEEE International Conference Signal Processing, Informatics, Communication and Energy Systems (SPICES), February 2015
16. Charleonnán, A.: Predictive analytics for chronic kidney disease using machine learning techniques. In: 2014 International Conference on Advanced Computing and Communication Technologies (ICACACT 2014), pp. 1–5 (2014)
17. Mathworks: MatLab (2017), vol. 21, no. 8, pp. 680–693 (2010). <https://es.mathworks.com/products/matlab.html>
18. Teixeira, H.: Contextual game design: from interface development to human activity recognition. Facultad de Ingeniería Universidad de Porto, Tesis maestría en Ingeniería Electrónica y de computadores (2017)
19. Bernal, C., Distrital, U., José, F., Hernández, C., Distrital, U., José, F.: Intelligent decision-making model for spectrum in cognitive wireless networks, vol. 10, no. 15, pp. 721–738 (2017)