A Mobile Health Solution for Diseases Control Transmitted by *Aedes Aegypti* Mosquito using Predictive Classifiers

Oton C. Braga[1], Olimária C. Fonsêca[1], Mário W. L. Moreira[1,2], Joel J. P. C. Rodrigues,[2,3,4] Francisca R. V. Silveira[1], Antônio M. B. Oliveira[1], Augusto J. V. Neto[5]

[1]Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE), Aracati, CE - Brazil

[2]Instituto de Telecomunicações, Universidade da Beira Interior, Covilhã - Portugal

[3]Instituto Nacional de Telecomunicações (Inatel), Santa Rita do Sapucaí, MG - Brazil

[4]Universidade de Fortaleza (UNIFOR), Fortaleza, CE - Brazil

[5]DIMAP, Universidade Federal do Rio Grande do Norte (UFRN), Natal, RN - Brazil

{otoncbraga, olimariacastro, amauroboliveira}@gmail.com

raquel_silveira@ifce.edu.br, augusto@dimap.ufrn.br

{joeljr, mario.moreira}@ieee.org

Abstract. *In healthcare, uncertainty moments are frequent, especially when they come from diseases with similar signals and symptoms. This work proposes a mobile health application based on predictive classifiers as inference mechanism capable to support health professionals in the identification of diseases transmitted by the Aedes Aegypti mosquito. The proposed system identifies the most probable disease in the case of dengue and chikungunya, given a set of symptoms presented by a patient. This work evaluates the experiments by cross-validation using real data, and the results show that decision tree perform well for the proposed solution.*

1. Introduction

Latest technologies have transformed the way people live and communicate. In healthcare [Manirabona et al. 2017], for example, developments of prostheses and mechanical arms used in surgeries are already a reality. In information and communication technologies, smart decision support systems (DSSs) try to find solutions capable to improve the outcome of extremely complex processes.

For several reasons, moments of uncertainty often occur in medicine and healthcare. Some diseases have similar or identical symptoms, requiring specific evaluations for a more accurate diagnosis. In some cases, tests are not able to provide a correct diagnosis. DSSs propose interesting contributions for problem-solving involving lack of information. DSSs embedded in mobile technologies can infer a more reliable result in an automated manner at any time and any where [Laguna and Finat 2013]. The application of these systems on health has brought several benefits to both physicians and patients, improving medical care and supporting professionals in the decision-making process [Gardini et al. 2013, Raffaeli et al. 2016, Joseph and Brown 2017]. Although it is recommended that patients have a medical appointment when present symptoms of dengue or *chikungunya*, this process may be time-consuming in several countries, such as Brazil, given the pressure of population that demands healthcare services. The *Aedes Aegypti* mosquito mitigation, which is the transmitter of these diseases, has become the primary target of public health campaigns, in Brazil, last times. According to the Ministry of Health, more than R$ 20 million (Brazilian real) were released, in 2016, to combat the mosquito [Brazil 2016b]. Some initiatives have been taken to stem its progress. However, mosquitoes develop rapidly and replicate quickly in favorable environments.

DSSs can assist health specialists and general population in the process of identifying mosquito-related diseases and their severity, which are commonly confused, due to the similarity of their symptoms. In the same way, using mobile technology, together with intelligent systems, can contribute to the reduction of queues in hospitals when related to these diseases. In this context, this study proposes a mobile system to identify diseases transmitted by the *Aedes Aegypti* mosquito, using classification methods. From the user-informed symptoms, via the user interface provided by the application, the system calculates the probability of the patient being infected with dengue or *chikungunya* through a supervised machine learning (ML) algorithm. The system can be used by both doctors and users in uncertainty moments on triage. The objective of the proposed system is to contribute for fighting against this mosquito, as well as the diseases that it transmits.

The rest of the paper is organized as follows. Section 2 deals with works related to smart DSSs, both in several knowledge areas and applied to healthcare. Section 3 deals with the use of classifiers, highlighting those based on probabilistic models and decision trees. Data pre-processing and sorting are considered in Section 4. Section 5 presents and describes the created mobile solution and, finally, Section 6 concludes the paper and suggests further works.

2. Intelligent Approaches to Support Decision-making Process

Decision Support Systems (DSSs) use several strategies to solve a given problem related to the topic under study. Some approaches apply inductive inference to adapt to new situations while other methods use probability-based mathematical models for information and knowledge discovery in large data sets. Another well-known method is ML, which is a well-known artificial intelligence tool under use. This approach uses pattern recognition to perform deductions from a set of examples. A learning model classifies a particular class through its attributes, which helps prediction rules to treat new situations, for instance, to generate a precise diagnosis [Stange and Neto 2010]. Both approaches have evolved rapidly and many methods already produce reliable results in different situations [Faceli et al. 2015]. Thus, this work uses supervised ML algorithms based on probabilistic methods and search. This approach involves data mining (DM) statistical models that identify the data arrangement in a sample. The classifiers based on Bayes' theorem are examples of this strategy. The search-based methods represent other leading techniques group. The tree-based models determine a hierarchical data description. Decision trees and adaptive systems are approaches that fit this strategy. Therefore, it is possible to find several applications for each approach and the above-mentioned data classification strategies. In Morais and Fechine [Morais and Fechine 2013], the authors use a system to support the teacher in the decision-making process. The paper analyzes the most relevant factors in a distance education course. This system uses DM techniques in a virtual learning environment (VLE) to classify educational data, to know, decision trees and Bayesian networks (BNs). Both techniques presented similar results. Thanathornwong *et al.* use a system for predicting dental whitening procedure results. This work applies multiple regression equations in a set data of CIELAB color coordination, before and after the procedure. Results show this method can predict other cases precisely [Thanathornwong et al. 2016]. Teles *et al.* present a DSS focused on the dengue diagnosis and its severity. This system uses BNs to support dengue diagnosis in uncertainty cases. The system analyzes user data (symptoms) and infers about the disease severity [Teles et al. 2014]. In [Moreira et al. 2016a] the authors use BNs to support decision making in uncertainty moments. This research develops an inference mechanism based on ML techniques using the Naïve Bayes (NB) classifier in a health database to classify hypertensive disorders in pregnancy focusing on the preeclampsia care. This system analyzes the data disposition and sort them in the network. From the symptoms presented by the pregnant woman, the system infers the severity of the case using statistical data. Results show this approach can assist the specialist in the preeclampsia diagnosis. This method proved to be accurate even with a small amount of data. Ayyaz *et al.* use mathematical models to simulate epidemics, creating preventive measures to combat diseases with epidemic characteristics. The presented models can identify the spread disease particularities and predict with reasonable efficiency where can occur epidemic problems. The work shows an application for *Aedes Aegypt* mosquito propagation [Ayyaz et al. 2015]. Moreira *et al.* compare the Naive Bayes method with the J48 decision tree-based classifier. This paper analyzes a data set related to hypertensive disorders in pregnancy to evaluate complications in gestation using a confusion matrix and its predictive parameters. Although the two classifiers have presented close values, the results show that J48 decision tree algorithm is the classifier with greater precision for this situation [Moreira et al. 2016c].

3. Using Classifiers in Smart DSSs as strategy to control the *Aedes Aegypti* Mosquito

A physician needs to follow a particular procedure to give a diagnostic. Firstly, the physician analyzes the information presented by the patient to formulate an hypothesis; then, requests exams to validate it, providing a more accurate diagnosis. However, many factors can influence the physician decision-making process, affecting the final diagnosis. The environment, fatigue, stress, excess of patients, emotional factors, among others, can contribute negatively to consultation result. Classifiers work in a similar way, though in a less complex procedure. These approaches analyze a data set and separate it into individual classes that rely on a set of specific or shared characteristics. The classifier seeks data in the electronic medical records to analyze the symptoms presented in a new case. Then, it calculates the probability for each class.

3.1. Epidemiological Context

Dengue is a problem of greater significance in Brazil due to its high mortality rate. The vector *Aedes Aegypti* transmits this disease. This mosquito spreads quickly in favorable environments, and it has been reproducing mainly because of inadequate sanitary conditions and standing water exposure. The Brazilian government has already taken several steps to mitigate the problem. Despite high expenses, the Brazilian health ministry had obtained no significant results, and the population infected with dengue grows each year. This issue has become an object of the leading public campaign in the country. Notwithstanding, this mosquito also carries virus from other diseases, such as Zika virus and chikungunya. The incidence of these diseases has increased in recent years and has caused significant harms to public health. The cases evolved mainly in the rainy seasons usually by the first weeks of the year. The health ministry publicized more than 800,000 dengue cases only in 2016 [Brazil 2016a].

3.2. The Naïve Bayes Classifier

Statistical assumptions characterize the Bayesian classifiers. This approach calculates the frequency of event occurrence to define a mathematical model suitable for predicting the result of a new event still unknown. This proposal use the Bayes' theorem, which calculates the probability of an event c_i given an event x, $P(c_i|x)$. For example, a patient likelihood having dengue fever if it has a fever, back pain, among others symptoms. Equation 1 shows this theorem.

$$P(c_i|x) = \frac{P(x|c_i)P(c_i)}{P(x)} \tag{1}$$

Where, $x = (x_1, x_2, ..., x_n)$ represents the set of attributes (symptoms) and $c = (c_1, c_2, ..., c_m)$ the classes (diseases). $P(x_j), P(c_i)$ are the *a priori* probabilities. Like this, $P(c_i|x)$ are the conditional probabilities of the attributes for each class and $P(x|c_i)$ the likelihood for new events.

The NB classifier is one of the most used Bayesian classifiers in DM. Although using a simple premise and considering the independent attributes of each other, it presents accurate results for cases appropriate to its context. From the Bayes theorem showed in (1), this work disregards the term $P(x)$, since it will be the same for all the classes, thus simplifying this theorem by obtaining (2).

$$P(c_i|x) \propto P(c_i) \prod_{j=1}^{n} P(x^j|c_i) \tag{2}$$

Equation 3 shows the theorem application to the diseases context, considering the independence of symptoms among themselves.

$$P(Disease_i|Symp) = P(Symp_1|Disease_i) \cdot ... \cdot P(Symp_n|Disease_i) \cdot P(Disease_i) \tag{3}$$

Figure 1 presents the nodes relationship constructed by the NB classifier. It gives the graphical structure of the relationship between input nodes (symptoms) and output nodes (diseases). In this case, each disease depends on the *a priori* probability of its symptoms group.

3.3. The Decision Tree-based Classifier

Decision trees represent directed graph structures. Its nodes can be root nodes, where the tree begins, nodes that divide a given attribute and generate branches, and leaf nodes that

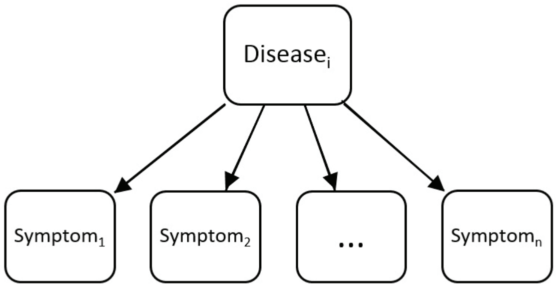

Figure 1. The NB classifier Representation.

contain the classification information of the algorithm. This method uses search-based algorithms to achieve the best possible representative graphical model of the observed experience. Figure 2 shows a decision tree construction example.

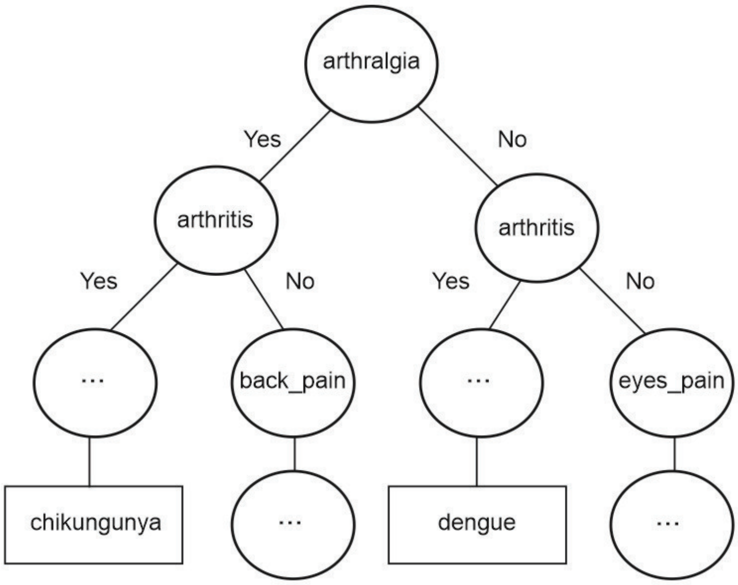

Figure 2. Example of a Decision Tree Graph.

To obtain this graphic representation it is necessary to perform several search functions. Equation 4 calculates the purity degree of an attribute. The value found on attribute divisions represents the amount of class information, *e.g.*, the relationship between fever and dengue fever. Thus, performing a weighted sum of this sample set, it is possible to discover the purity degree of that attribute.

$$IG(X, A) = Entropy(X) - \sum_{i=1}^{n} \frac{|n_i|}{|N|} Entropy(X) \tag{4}$$

Where $X = (n_1, n_2, ..., n_n)$ represents the attributes, n_i expresses the size of X, and N is the training set. The term $Entropy(X)$ measures the variable variation, *i.e.*, how difficult is to predict this variable. It is necessary to separate the attributes into distinct classes to find a maximum entropy. Otherwise, the entropy is zero. Thus, as much as fever cases are distributed between dengue and *chikungunya*, more difficult will be their prediction. However, as much as fever cases are distributed for one of the diseases, it improves its prediction. Equation 5 shows the entropy equation.

$$Entropy(X) = -\sum_{j=1}^{Nclass} p(c_j|n_i)\log_2 p(c_j|n_i) \tag{5}$$

The difference between the entropy of the first sets case and the partitions entropy defines the information gain of an X attribute. Equation 5 presents this relationship. The node that has the highest information gain determines the best attribute to be used, simplifying the tree structure. Thus, the symptoms that have the best entropy will be at the top of the tree.

4. Data Classification of Dengue Fever and *Chikungunya*

The identification difficulty of these diseases is due to its symptoms similarity. It is an interesting research topic for artificial intelligence context. Thus, this study uses data on diseases caused by the dengue fever mosquito for training the classifiers.

4.1. Data Analysis and Preprocessing

The data analyzed were extracted from the open data portal of the Recife prefecture, Brazil [Recife 2016]. These data present clinical and laboratory attributes. It requires technical or invasive procedures on its acquisition. Therefore, this work disregarded such data, paying attention to the characteristics related to the diseases symptoms under study. Table 1 shows the main symptoms presented by patients.

Table 1. Main Symptoms Presented by Patients.

Symptoms
Fever
Nausea
Vomiting
Arthritis
Conjunctivitis
Headache
Back pain
Muscle pain
Arthralgia intense
Pain around the eyes
Red spots on the skin
Red dots on the skin

Data about the patient's health history can also be extremely relevant for a disease diagnosis. Therefore, some attributes express preexisting conditions. Table 2 shows these conditions. This study concerned to verify only prevalent diseases, with the intention to reach ordinary people. This work truncated cases that contained missing data.

Table 2. Preexisting Diseases.

Diseases
Diabetes
Hepatitis/cirrhosis
Chronic kidney disease
Hypertension

The ”Month of the Year” field was separated into four periods to highlight rainy seasons. Thus, the first weeks of the year, when more cases occur, the diseases under study are clearly distinguished.

The database includes $1,274$ *chikungunya* cases and $4,687$ dengue fever cases. Hence, performance evaluation uses the SMOTE technique to perform a data balancing to obtain a better classification [Chawla et al. 2002]. This evaluation also performs a balancing percentage ranging between 200% to 350% depending on the algorithm behavior. Thus, *chikungunya* cases numbers are close to the dengue fever cases. Some algorithms presented better results after an attribute selection, which truncates the less relevant attributes to highlight the most significant ones [Hall and Holmes 2003]. Thus, this study also applies this strategy to some algorithms.

4.2. Data Classification

An intense search from related works to diseases' diagnosis in doubt moments allowed choosing the most appropriate classifier for the problem. For this, the search used parameters such as uncertainty, disease, classifiers, and DSS. It highlighted some classifiers such as J48, NB, Random Forest (RF), and Bayes Net [Webb 2011] [Bradley 1997] [Moreira et al. 2016b].

These classifiers present satisfactory results for specific situations but lose their effectiveness when applied to other circumstances [Chakraborty et al. 2016, Parida and Dehuri 2014]. Thus, this research concludes that there is a classifier that is the most suitable for each situation. Therefore, a 10-fold cross-validation test indicated the most appropriate classification for the dataset used in this work. It consists of dividing the database into ten subsets and selecting one for testing and the remainder for learning. This validation method uses each set only one time for the test, *i.e.*, this procedure occurs ten times. This method developed by [Browne 2000] is widely used in validation tests.

4.3. Performance Evaluation and Results Analysis

Performance evaluation used two preprocessing strategies: balancing and attribute selection. Table 3 shows the used mechanisms in each algorithm as well as its configurations. For this, the use of different balancing and attribute selection settings reached to 20 experiments. The harmonic mean used in this study is a performance measure widely used in forecasting tasks. Combining precision and recall, avoiding disadvantages of simple metrics such as error rate, especially in cases of unbalanced class distributions [Busa-Fekete et al. 2015]. Precision and recall are measures obtained by the ratio of the positive examples and the success rate of the real class respectively. Table 3 shows the best results achieved by each algorithm.

Table 3. F-Measure Classification Results.

F-Measure				
Algorithm	**Dengue Fever**	***Chicungunya***	**Attibute Selection**	**Balancing**
Random Tree	74.2	80.2	No	350%
Random Forest	75.4	79.7	No	320%
Naive Bayes	70.2	70.4	Yes	300%
Bayes Net	73.1	78.5	No	350%

According to table 3, the algorithms based on decision tree presented better results for the data set considered in this work. Table 4 shows the confusion matrix for the two considered algorithms. The confusion matrix shows the relationship between the real cases and the cases classified by the algorithm. This model allows to discover the sensitivity and specificity.

Table 4. Confusion Matrix.

Bayes Net			**Random Forest**		
Real \ Classified	D	C	Real \ Classified	D	C
Dengue Fever	3387	1300	Dengue Fever	3426	1261
Chicungunya	1192	4741	*Chicungunya*	972	4378

The symptoms presented by the diseases are very similar and are easily confused. The matrices show the relationship between the classes predicted by the algorithms and their real values. Both algorithms make mistakes in the classification of some cases. However, the results are relevant, indicating excellent predictors for identifying significant number of cases. The more specific used attributes would certainly improve this prediction. However, since the objective proposed by this work is to perform a first patient screening, it would not be feasible to use such attributes.

Table 5. Random Forest Classifier Results.

Precision Measurements				
	Precision	Recall	F-Measure	ROC Area
Dengue Fever	77.6	81.8	75.4	85.6
Chikungunya	77.9	73.1	79.7	85.6

The RF classifier performed well, in average. The ROC area presents appropriate values. However, the precision, recall, and F-Measure show the classifier's deficiency in handling the furnished data. The relationship between sensitivity and specificity provides these measures, calculated from false positive, false negative, true positive, and true negative values. These indicators estimate the system reliability [Powers 2011].

5. Mobile Solution to Support Experts in the Decision-making Process

From the analysis of the algorithms, it was possible the development of a remote system dedicated to first care systems for patients with suspicion of the diseases under study. The system receives HTTP requests from any application (containing the patient disease's symptoms) and returns the probability for each illness and explanation for the suggested diagnosis results. The RF classifier processes the data. This approach had the best adaptation for each case under study, showing better results in the classification process. This research also developed a mobile application that captures a set of symptoms reported by the patient from an objective questionnaire. After filling the survey, the users' interface provides a button for classifying the user' symptoms. After, it sends the obtained classification from the patient to a remote server, as shown in Figures 3 and 4.

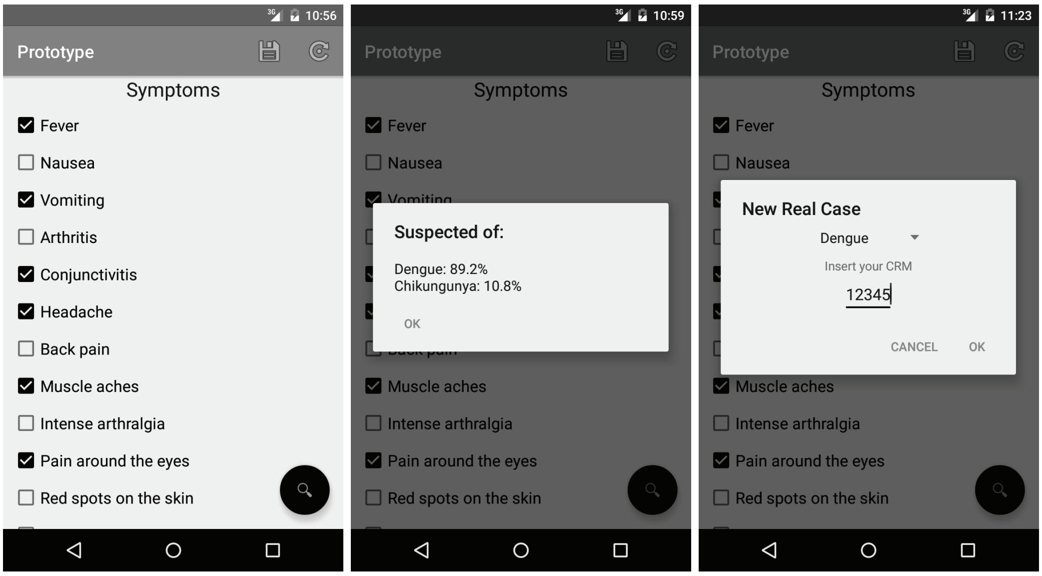

Figure 3. Some Mobile Application User's Interface.

The system offers a collaboration option of experts on healthcare. After validation of the CRM (Brazilian Regional Medical Record), experts can increase the system with new cases. The remote system database receives the increased data. Figure 3 (third screen) shows this functionality. Figure 4 shows the proposed architecture for the model. With more cases registered in the database, the algorithm can train on different situations and improve its prediction accuracy. The dataset formed by $5,961$ cases of diseases under study was used to prepare the application. The Apache Tomcat framework, available at `http://tomcat.apache.org` was used to develop the Web system using JAVA EE, and the WEKA API performed the classifications. This API can be found on "Data Mining Software in Java" [Frank et al. 2016]. The API has several classification algorithms widely used in DM. The Data Server uses MySQL, available in `https://www.mysql.com`. The system runs in the cloud, supporting multiples requests simultaneously. The mobile application development in JAVA uses the IDE platform Android, named Android's Official Studio. It uses the SDK Android 22. It is available for devices with versions from 4.0 (Jelly Bean). Other applications can also send HTTP requests to the system.

6. Conclusion and Future Work

The analysis of different strategies is essential in prediction processes and DM. The algorithms comparison applied to particular problems are determinant in any ML practice. The uniqueness of data precludes a deduction efficient because the algorithms can behave

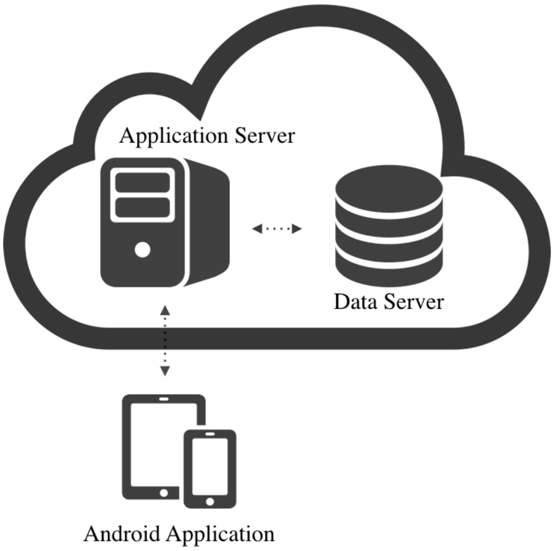

Figure 4. Simplified Illustration of the System Architecture.

differently in individual cases. Finding a classifier that presents better results for a given data set is substantial for a system success. The decision tree-based classifier, named RF, although showed high values to predict all the diseases proved to be very useful for the system requirement. Each symptom can be presented in various illness transmitted by the mosquito and it has an absolute randomness, which complicates the classification. However, the accuracy given by the system has significantly relevance to control the population uncertainty and help health professionals about these diseases. This work concludes that the proposed system offers a significant contribution to the society, in general, supporting also health professionals to combine the patient symptoms to the right disease precisely, functioning as a query tool for users in general.

The created application only considers attributes and symptoms identified in a case analysis available on a single database. In this way, further works propose the creation of an adaptive system for new attributes or classes in the network, improving the data prediction. Other diseases transmitted by the *Aedes Aegypt* mosquito is an important research topic. Zika fever represents a dangerous disease transmitted by the same vector and has caused serious harm in the population. There is an increased occurrence of microcephaly in the country associated with this disease (Brazil), considered as a national urgency [Luz et al. 2015]. The public records lack for this illness made it impossible for their inclusion in the classification process of this research.

Acknowledgments

This work was supported by the by National Funding from the FCT - Fundação para a Ciência e a Tecnologia through the $UID/EEA/50008/2013$ Project, by Finep, with resources from Funttel, Grant No. $01.14.0231.00$, under the Centro de Referência em Radiocomunicações - CRR project of the Instituto Nacional de Telecomunicações (Inatel), Brazil, by Ciência sem Fronteiras of CNPq, Brazil, through the process number $207706/2014 - 0$, and by FUNCAP (Fundação Cearense de Apoio ao Desenvolvimento

Científico e Tecnológico) under the Program of Research Productivity Grants, Incentive for Interiorization and Technological Innovation - BPI, FUNCAP Edital No. 09/2015.

References

[Ayyaz et al. 2015] Ayyaz, A., Muaz, U., Awan, S., and Ayyaz, M. N. (2015). Simulation model for counter-measures against *Aedes Aegypti*. In *2015 13th International Conference on Frontiers of Information Technology (FIT), December 14-16, Islamabad, Pakistan*, pages 98–103.

[Bradley 1997] Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145–1159.

[Brazil 2016a] Brazil (2016a). Ministério da saúde. monitoramento dos casos de dengue, febre de chikungunya e febre pelo vírus zika até a semana epidemiológica 13. http://portalsaude.saude.gov.br/images/pdf/2016/abril/26/2016-014---Dengue-SE13-prelo.pdf. Accessed: 2017-04-29.

[Brazil 2016b] Brazil (2016b). Ministério da saúde. prevenção e combate: Dengue, chicungunya e zica. http://portalsaude.saude.gov.br/. Accessed: 2017-04-29.

[Browne 2000] Browne, M. W. (2000). Cross-Validation Methods. *Journal of Mathematical Psychology*, 44:108–132.

[Busa-Fekete et al. 2015] Busa-Fekete, R., Szörényi, B., Dembczynski, K., and Hüllermeier, E. (2015). Online f-measure optimization. In *Advances in Neural Information Processing Systems (NIPS 2015), December 7-12, Montreal, Canada*, pages 595–603.

[Chakraborty et al. 2016] Chakraborty, C., Gupta, B., and Ghosh, S. K. (2016). Chronic wound characterization using bayesian classifier under telemedicine framework. *International Journal of E-Health and Medical Communications (IJEHMC)*, 7(1):76–93.

[Chawla et al. 2002] Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357.

[Faceli et al. 2015] Faceli, K., Lorena, A. C., Gama, J., and de Carvalho, A. C. (2015). *Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina*. LTC, Rio de Janeiro, Brazil.

[Frank et al. 2016] Frank, E., Hall, M. A., and Witten, I. H. (2016). The weka workbench. online appendix for "data mining: Practical machine learning tools and techniques". http://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf. Accessed: 2017-04-29.

[Gardini et al. 2013] Gardini, L. M., Braga, R., Bringel, J., Oliveira, C., Andrade, R., Martin, H., Andrade, L. O. M., and Oliveira, M. (2013). Clariisa, a context-aware framework based on geolocation for a health care governance system. In *2013 IEEE 15th International Conference on e-Health Networking, Applications Services (Healthcom), October 9-12, Lisbon, Portugal*, pages 334–339.

[Hall and Holmes 2003] Hall, M. and Holmes, G. (2003). Benchmarking attribute selection techniques for discrete class data mining. *IEEE Transactions on Knowledge and Data Engineering*, 15(6):1437–1447.

[Joseph and Brown 2017] Joseph, R. and Brown, P. (2017). The cloud gets personal: Perspectives on cloud computing for personalized medicine. *International Journal of E-Health and Medical Communications (IJEHMC)*, 8(2):1–17.

[Laguna and Finat 2013] Laguna, M. A. and Finat, J. (2013). Personalized mobile applications for remote monitoring. *International Journal of E-Health and Medical Communications (IJEHMC)*, 4(1):1–11.

[Luz et al. 2015] Luz, K. G., Santos, G. I. V. d., and Vieira, R. d. M. (2015). Febre pelo vírus zika. *Epidemiologia e Serviços de Saúde*, 24(4):785–788.

[Manirabona et al. 2017] Manirabona, A., Fourati, L. C., and Boudjit, S. (2017). Investigation on healthcare monitoring systems: Innovative services and applications. *International Journal of E-Health and Medical Communications (IJEHMC)*, 8(1):1–18.

[Morais and Fechine 2013] Morais, A. M. and Fechine, J. (2013). Mineração de dados educacionais no apoio ao processo de tomada de decisão do docente. In *Congresso da Sociedade Brasileira de Computação, July 23-26, Maceió, AL, Brazil*, pages 478–483.

[Moreira et al. 2016a] Moreira, M. W., Rodrigues, J. J., Oliveira, A. M., Ramos, R. F., and Saleem, K. (2016a). A preeclampsia diagnosis approach using bayesian networks. In *2016 IEEE International Conference on Communications (ICC), May 23-27, Kuala Lumpur, Malaysia*, pages 1–5.

[Moreira et al. 2016b] Moreira, M. W. L., Rodrigues, J. J. P. C., Oliveira, A. M. B., Saleem, K., and Neto, A. V. (2016b). An inference mechanism using bayes-based classifiers in pregnancy care. In *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom), Semptember 14-17, Munich, Germany*, pages 1–5.

[Moreira et al. 2016c] Moreira, M. W. L., Rodrigues, J. J. P. C., Oliveira, A. M. B., Saleem, K., and Venancio Neto, A. J. (2016c). Performance evaluation of predictive classifiers for pregnancy care. In *2016 IEEE Global Communications Conference, December 4-8, Washington, DC, USA*, pages 1–5.

[Parida and Dehuri 2014] Parida, S. and Dehuri, S. (2014). Review of fmri data analysis: A special focus on classification. *International Journal of E-Health and Medical Communications (IJEHMC)*, 5(2):1–26.

[Powers 2011] Powers, D. M. (2011). Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1):37–63.

[Raffaeli et al. 2016] Raffaeli, L., Spinsante, S., and Gambi, E. (2016). Integrated smart tv-based personal e-health system. *International Journal of E-Health and Medical Communications (IJEHMC)*, 7(1):48–64.

[Recife 2016] Recife (2016). Prefeitura de recife. casos de dengue, zika e chikungunya. `http://dados.recife.pe.gov.br/dataset/casos-de-dengue-zika-e-chikungunya`. Accessed: 2017-04-29.

[Stange and Neto 2010] Stange, R. L. and Neto, J. J. (2010). Reconhecimento de padrões em classificadores – comparação de técnicas e aplicações. In *IV Workshop de Tecnologia Adaptativa (WTA 2010), January 21-22, São Paulo, SP, Brazil*, pages 63–67.

[Teles et al. 2014] Teles, G., Oliveira, C., Braga, R., Andrade, L., Ramos, R., Cunha, P., and Oliveira, M. (2014). Using bayesian networks to improve the decision-making process in public health systems. In *2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom), October 11-15, Natal, RN, Brazil*, pages 565–570.

[Thanathornwong et al. 2016] Thanathornwong, B., Suebnukarn, S., and Ouivirach, K. (2016). Decision support system for predicting color change after tooth whitening. *Computer methods and programs in biomedicine*, 125:88–93.

[Webb 2011] Webb, G. I. (2011). Naïve bayes. In *Encyclopedia of Machine Learning*, pages 713–714. Springer.