Behavioral Analysis for Child Protection in Social Network through Data Mining and Multiagent Systems

Mário Sérgio Rodrigues Falcão Jr., Enyo José Tavares Gonçalves, Tciciana Linhares Coelho da Silva and Marcos de Oliveira
Universidade Federal do Ceará (UFC), Quixadá, Brazil

Keywords: Child Sex Grooming, Multiagent Systems, Data Mining, Cognitive Analysis.

Abstract: The Internet connects millions of people worldwide, enabling diverse ways of interaction and social organization. Online Social Networks such as Facebook, MySpace, and Twitter, have created a new form of socialization that can provide good experiences for users. However, such systems, as well as connecting people, expose their users lives to others, making them subject of exploitation in many ways. This work explores specifically children and teenagers degree of exposition on Facebook. Due to the risk offered in distinct layers of the Internet, the aim of this work is to develop a smart tool that helps to avoid the action of individuals that are possibly a risky for children and teenagers, users of the social network Facebook, applying Data Mining techniques in a Multiagent System.

1 INTRODUCTION

The Internet is a global network that connects millions of computers around the world, serving as a major factor of communication and social integration (Belloni, 2001). A group of applications for internet are built based on the ideological and technological foundations of Web 2.0, which allow the creation and exchange of User Generated Content (UGC) (Kietzmann et al., 2011), i.e. blogs, social networking pages, chats, professional networks (LinkedIn), community networks (social networks in neighbourhoods or cities), political networks and especially electronic social networks such as Facebook, Twitter, Google+, MySpace, etc. They compose the large set of social media (Lemieux et al., 2008).

It is relevant that in our modern society people are closer to technology than before, especially children. A child from the 80's was more used to handle toys, and nowadays children have better skills on handling information technology, due to the routine contact with it (Bombonatto, 2012).

Online Social Networks are virtual organizations composed by people, and give the opportunity for them to interact with distinct types of individuals (Mazman and Usluel, 2009). That allows for the possibility of development of certain relationships that match certain characteristics indicating that the children and teenagers are being sexually groomed. Children with emotional distress, emotionally dysfunctional, and with low self-esteem are more likely to be tricked and open to feelings of defencelessness, a factor that some child sex offenders use as advantage in its investees. Dialogs involving subjects such as family problems, secrets and participation in controversial issues, are some of the varied resources adopted by these offenders to gradually seduce their victims. After that follows events such as sending photos with sexual connotations, improper conversations on pornography are dominant in other dialogues. Thus, the responsible by the child can be assured that a seduction process was started, and protection measures should be taken.

There is a noted lack of inspection by most of the parents related to the use of the Internet by their children. Many children are exposed to a world before unknown, and that put together different types of personalities (Pereira, 2009). A survey, conducted by Minor Monitor Company, says that about 38% of users in the social network Facebook do not have the allowed age for their use, and 30% of parents allow unsupervised use (Silvio, 2012). One possible reason for this lack of inspection may be associated with the shortage of computational tools to help them.

Jr., M., Gonçalves, E., Silva, T. and Oliveira, M.
Behavioral Analysis for Child Protection in Social Network through Data Mining and Multiagent Systems.
In *Proceedings of the 18th International Conference on Enterprise Information Systems (ICEIS 2016) - Volume 2*, pages 306-313
ISBN: 978-989-758-187-8

To reduce this deficit of inspection, it is interesting to have available automated tools that can identify irregular behaviour patterns or potential suspects in the social network. Therefore, the group of users, composed by children and teenagers using these networks, would be less vulnerable.

Given this gap, it is necessary to propose mechanisms that can be easily used by parents, to ensure the dignity of children and teenagers that use social networks, as they are still in psychological growth, and are subject to individuals with bad intentions on the Internet (Buckingham, 2000).

This article proposes an alternative to parents before the scenario presented. We implemented a Multiagent System that applies Data Mining to automatically classify the exposition of children in the Facebook. That is made based on the interactions developed in it. This paper is organized as follows: Section 2 presents the theoretical foundation, the Section 3 shows some related works, Section 4 explores the architecture of the systems developed, observing the Data Mining modules that were adopted in this work, section 5 describes the results, and Section 6 presents the conclusion and lists possible future work.

2 RELATED WORK

2.1 Automatic Text Analysis in Instant Messages to Detect Child Sexual Grooming

In a study by (Santin, Freitas and Paraiso, 2011), a software service was developed to classify conversations stages in chat rooms, through a set of pre-selected words and rules. The algorithm SVM (support vector machine) was used to classify the stages of interaction between entities, children and possible suspects. They used the database found in the website "www.perverted-justice.com", which has real talks between pedophiles and children. This database was not used directly in this work, but served as examples of how sex offenders interact with their victims and persuade them.

Inconveniently, the words of the chat should be exactly the same to those found in the set of words in the database, which greatly limits the accuracy to identify the stages, since words can vary from region to region or from person to person.

The main difference between this work and (Santin, Freitas and Paraiso, 2011) is the analysis of events that happen in the context of the social network Facebook, and do not depend as much from accuracy as analysing a given string, not just focusing on a literal analysis as the automatic text analysis that they do.

2.2 Fake Profile Identification in Online Social Networks

The Social Privacy Protector Software tool for Facebook (SPP), aims to identify "fake" profiles on the social network Facebook, and improve the privacy and security settings of users (Fire, Kagan, Elyashar and Elovici, 2014).

The SPP has three layers of protection that enhance user privacy through the implementation of different methods. The system first displays a possible profile that can pose as threat, and then immediately provides the means to restrict personal information to the suspect profile. Then the second layer allows the owner of the profile on the social network to adjust your privacy settings according to your personality type.

The third layer of the system alerts the user about the number of third-party applications installed in his profile that have access to his private data

Similarly to the work presented in this paper, the social network used by the SPP is the Facebook, and it makes use of data mining techniques for profile classification. However, it should be noted that the focus of the SPP is to identify "fake" profiles, that is, profiles that are note true in relation to the identity of the their owners regardless of their goals. Similarly, the approach described here seeks to identify people who threaten children. This children group refers to people under 12 years of age incomplete, according to Brazilian law.

2.3 A Learning Model for Intelligent Agents Applied to Poultry Farming

The work in question relates to a system responsible for analysing data on poultry farming and to provide results that meet the basic needs for birds like amount of food being consumed, temperature, amount of water and energy to be consumed among other factors (Ribeiro et al, 2015).

The system run practical examples, and achieved good results, consistent with reality, and equivalent to data that poultry professionals would consider correct and appropriate for a given situation. The experiments were conducted in two parts: the first in which the agents worked according to the average weight of chickens and the second according with the weekly food. The agents with the best accuracy

classified the chickens that obtained the best weight and disease resistance them took the indicated supplements and ate correctly taking into consideration temperature, humidity, time, light, etc.

Similarly, the work presented in this paper also used MAS and Data Mining, the logic of both works is somewhat similar in the sense that agents form part of the system responsible for making decisions, with predefined heuristics, and through data mining, data classification becomes possible.

2.4 A Problem-solving Agent to Test Rational Agents a Case Study with Reactive Agents

In this work described by (Silveira, Campos and Cortés, 2014), a set of tests was conducted with a rational agent, owner of knowledge and judgment skills in certain subjects. Several tests were conducted with the objective of obtaining the results and thereby try to improve agent performance on their weaknesses.

The agent was composed of behaviours previously defined by the project designer and sensors for detecting aspects in accordance with the environment. The tests are designed to identify weaknesses that the agent would probably have when certain events happen.

The results were satisfactory for research, but not for the agent tested, some vulnerabilities have been found and immediately reported to the project designer.

Similar to this work some tests were performed with our agents, the black-box type, different input types and amount of data to see the reaction of the agents. During the process, some inconsistencies were found, such as misclassification of information, and then treated so that the agents could work cohesively and efficiently.

3 MULTI-AGENT SYSTEM AND DATA MINING FOR CHILDREN PROTECTION ON FACEBOOK

The system consists of Autonomous Agents developed in the Jade framework (Java Agent Development Framework) (Bellifemine et al., 2007), and a data-mining module that uses the WEKA's classification algorithms (Hall et al, 2009) over the data collected from Facebook using the Facebook4j framework (Yamashita, 2009).

3.1 System Architecture

Multi-agent systems were used to make decisions with precision, working together on data from Facebook. The system has more than one agent in order to modularize the functions of analyzers and distribute responsibilities, so that each agent has a specific function, for example, the agent analyzer of posts will check the times, the kind of privacy, the content present in the text and among other heuristics defined in its scope. The agents have a structure with predefined heuristics by professionals of psychology aimed at detecting irregularities on data found on Facebook of children. Formulas that calculate the level of privacy in accordance with what the child posts, likes or shares are defined, moreover, these agents have access to dictionaries full of indecent language to compare with the words that child uses.

The autonomous agents interact among them collecting data from the Facebook server via the Facebook4j framework (Yamashita, 2009). Files in JSON format are requested and read by autonomous agents. These filter the information and work together to create an object that is capable of being classified by the mining module, such an object generated has references to relevant information for the data mining module, then the latter is in charge of receiving this object by means of a method call, and use of the classification algorithm to classify this instance according with the model defined in the system. After that, the vulnerability level at which the child is found is returned to the central module of the server, and this elaborates a report to the end user by explaining the reasons why the child is or not classified as at risk.

After all processing of information the central server prepares the data considered suspect and returns it in an HTML page.

The information provided in the final report are: friends considered suspicious, posts made by the child, or for someone where she was marked, with descriptions or comments considered not advisable for his/her age, amount of pictures in high exposure, books , videos, TV channels, groups and music not recommended, registered family in the social network or not, amount of pokes received after midnight, shared albums with third parties and that have a subject not suitable for children and a privacy note stating a degree for the child overall exposure within the social network, i.e., indicating how exposed is his/hers data to friends, friends of friends,

strangers and games or applications that obtain data in a way not so perceptive to human eyes.

The architecture used did not cause problems unless on the synchronization from the central server with the MAS. In that case, implementation of thread control was necessary so that the server can wait for all the autonomous agents response and then immediately proceed to code execution.

3.2 Facebook Social Network Data Collection

So, to care on with the investigation it is necessary the existence of content. Given that the social network Facebook has a large volume of data and different types of user profiles, some attributes play an important role in the interpretation of sentiment, habits, tastes, etc. As a result some features are relevant for analysis and decision of the potential risks posed to the child's account. These attributes can include: news feed, interests, music, videos, pages, liked or shared links, date of birth, inspirations, type of education, conversations, posts, improper dates and times for registered activities in social networking, games, events, TV channels, family, shared albums and photos in excess, bad comments on photos and posts, private messages from friends or others containing sexual connotation, privacy notice that tells you how your personal information is exposed and friends who might pose a threat. All of this information can be obtained through the Facebook Graph API, the primary way for apps to read and write to the Facebook social graph. The target audience are children under 12 years. All available information for Facebook will be obtained in a matter of a minute and following will be analyzed by the System.

3.3 Technologies Used

The developed system consists of autonomous agents developed in the JADE framework, and uses the standard protocols from FIPA (Foundation for Intelligent Physical Agents) for communication among agents.

The Weka data mining tool was chosen because it is known as one of the "top 10 free" tools of the current business intelligence market, according to the website "predicts analytics today" (Nyce and CPCU, 2007). The classification algorithm used in this work was decision tree (Quinlan, 1986) implemented in Weka as j48. The basic idea of the decision tree algorithm is recursively choose the best attribute to divide the nodes of the tree. After selecting an attribute, the data is divided into multiple partitions according to the value of the chosen attribute. For each partition recursively computed the best attribute is chosen to split the data in the current node of the tree. The decision rules are stored and new rules are generated.

3.4 Data Used and Validated Classification Model

The data to generate the classification model are obtained from Facebook and organized into tuples that have the following columns: posts, (if presents only private albums), private albums, preferences on books, music, videos, if updates the news feed, family registered in the account or not, games, videos with no children connotation in your comments or descriptions, pokes after midnight and conversations with adult content, and the last column responsible for classifying the tuple as "risk" or "no risk". The data that will fill the columns are analyzed by intelligent agents, which assign predefined values according to the discovered content, which will be used for the classification model. The values chosen by the agents will be explained further at the end of this section. The Decision Tree model is shown in Figure 1, so new instances can be classified as "no"(is not at risk), or "yes"(is at risk) warning the child to a possible danger based on the paths of the Tree.

Initially, the pattern obtained had a high rate of accuracy and was named template best case scenario, i.e., had weak presence of events uncommon as false positives and false negatives. At first it had a high correction rate to classify instances not to be correct but for being an addict model, that is, only one column with a risk offerer value the generated model already rate this instance as "risk" and do not bother to go further in the other columns aimed at generating a reasoned result in the maximum possible columns, as you can see at the Figure 1.

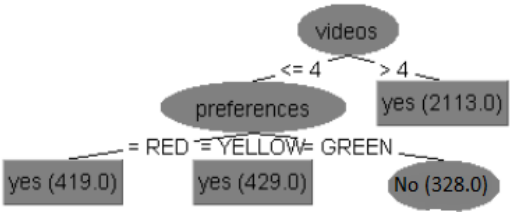

Figure 1: Vicious Model generated by the J48 algorithm.

However after a test sequence file composed of data responsible for generating the model yielded a

pattern suitable to receive different types of bodies, i.e., no longer grounded in only one column but in much as possible according to the child data be received, as you can see at the Figure 2.

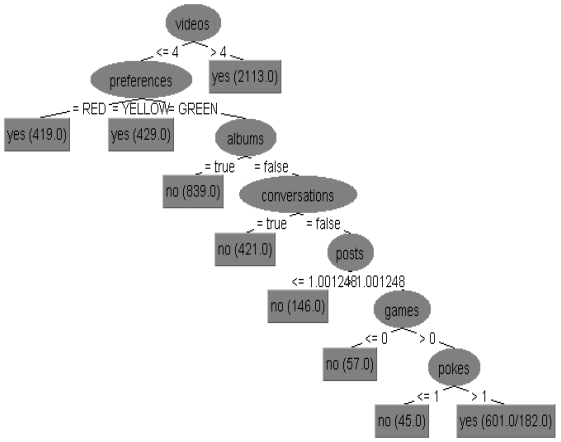

Figure 2: Model generated by the J48 algorithm.

A random algorithm created five thousand lines, with certain defined heuristics, to generate the classification model. The training of the model was as follows: instances are generated according to the values assigned to the tree fields in Figure 2, and according to these values at the end, it is classified as "no"(is not at risk) or "yes"(is at risk).

In order to refine the model and increase its accuracy some tests were run to have a satisfactory model. Among the tests used and provided by Weka, we can mention the Use Training Set, Supplied Test Set, Cross-Validation and the Percentage Split. The tests run served to refine the model and make it more able to categorize each instance of less incorrectly possible. As a result the inevitability of adding outliers, identification data that should follow an expected pattern but do not (Campos, 2014), became important to improve the model to less trivial cases.

The database used to form the classification model has 5270.00 lines (30% were used for model testing and 70% for training) composed of false positives, false negatives, true positives and true negatives. The tests conducted with J48 algorithm presented precision of 94.53% with 4,982 instances as correct and 288 classified as incorrect. The SMO algorithm achieved an accuracy of 90.13% with 4750 instances as correct and 520 classified as incorrect. The IBK algorithm had an accuracy of 94.03% with 4,954 instances as correct and 316 classified as incorrect. Finally, the NaiveBayes algorithm had an accuracy of 82.86% with 4,367 instances as correct and 903 classified as incorrect.

Even the result with lowest accuracy (generated by Naive Bayes) presents a quite satisfactory approach. We believe it is possible to use any of the tested approaches, however we kept in this experiments the one with greater accuracy, i.e. decision tree (J48).

The tree fields are filled with values defined by autonomous agents, as mentioned previously, according to their analysis; the field posts informs the safety of posts made by the child and the ones where he/she was marked. Any value above 3.0 is already considered as unsafe, the value is obtained by using the following formula F:

F = ((quantity of dangerous dates) * 3+ (privacy note * 1) + (quantity of dangerous messages * 2) + (amount of suspicious friends * 2) + (number of suspicious descriptions and comments in the post * 2)) / 10.

This formula came after a series of analyzes and tests with different Facebook accounts. For this reason, a weighted average was made in accordance with the information provided by the Facebook. In all cases in which this value was greater than 3.0 actually exist irregularities in the data.

The games field is a numeric value that tells you how much inadvisable games or apps are often used by children, for example, Tinder, Cupid, Interesting, Skout, Let's date, Meetmoi and Catrachos. Any amount in excess of five is already considered as unsafe. The videos field is a numeric value that tells you how much the child saw not advisable videos. The videos deemed not advisable are identified by name or description, if any keyword contains pejorative terms or sexual context video is classified as dangerous.

The field conversations are a Boolean that receives "true" or "false" indicating whether the private conversations of the child have any disagreeable talk or inappropriate content. These talks are extracted from Facebook and analysed by agents that are referred to in a database with pre-defined words that might pose a risk to the child.

4 VALIDATION

Two children's profiles were selected on Facebook and, with the permission of their parents, their data were used for analysis. The tool carried out the classification in the two instances of the case study fully automatically.

The two children have 12 years old and are female, the first child was born in Floriano City - PI and the second in Quixadá City - CE. Between these

Table 1: Model columns for classification.

Album	Preferences	Feed News	Pokes	Family	Games	Videos	Conversations
True|False	Green|Yellow|Red	True| False	Numeric	True| False	Numeric	Numeric	True|False

Table 2: Results.

Child	Times at dawn	Not advisable TV show	Not advisable Singers	Family Presence	Inbox Conversa-tions	Pokes after 12 AM	Privacy Rate	Posts
Child from Floriano	Logged 224 times after 12 AM	Malha-ção	6 different singers.	No family member is registered in your account.	Fuck, motherfucke, dirty and the expression: I want to see you naked told by a friend.	22	12,85	"And what if it is to be alone and happy without him, my life continued; D #Happy #Funny #Girl #Top (08 / April / 2014 to 12:31:57)."
Child from Quixadá	Nothing	Nothing	3 different singers.	No family member is registered in your account.	Nothing	Nothing	1,86	Nothing

two selected children, the child from Floriano did not know she was being monitored; on the other hand, the child from Quixadá knew about the experiment and was present during the work's validation next to his mother.

The Parents of the children in question provided the authorization for the analysis already knowing that there would not be disclosure of personal information from children, as the name, Facebook profile or any data that could expose their personal information would not be revealed. However, the results are listed observing the content found in children profiles, as shown in Table 2.

4.1 Considerations about Child from Floriano City

All the words in the column "Private Conversations" were recorded in mostly during the night, between 07:00 pm until 02:00 am the next day. Of this same child in 2014 Carnival period some friends asked if itself would have gone to two bands carnival celebrations. In addition a friend in particular called for a party at his home at 08:24 pm on 23 August 2013.

If you notice that the posts column there are certain reserved words that pose no danger, but when together the context of the sentence has subtended goal, in this case, it comes to an end a dating. At first few sentences were analysed manually to form the classification model.

It is noted that although Facebook account analysed in question had been inactive from the beginning of 2014 to June 2015, still some data could be retrieved and identified as unsafe for a child 12 years. The data-mining model classified the child as subject to risks.

4.2 Considerations about Child from Quixadá City

There were no suspicious events after midnight (00:00 to 06:00) in the second child account, the mother just made sure to point out that she does not

allow the use of social networking during late periods of the night.

No family member is registered in her account, which may represent a kind of lack of protection thus unknown people may feel more free to approach the child.

However, the privacy score for the child in question is low, 1.86, meaning that the child's data is "protected" and less accessible to applications or people.

In relation to music, three singers were identified as with song lyrics not advisable for children, which led the tool to classify them as irregular to the child. The following sentences were posted in the three singer's pages:

- **First Singer:** Come sing with me my people Ouro Branco-MG ... today has a lot of music and joy in the heating Hot Party Country! I want to see everyone there huh! From 22h. #ourobranco #escarpasfolia.
- **Second Singer:** Left Ponta Grossa-PR! Today the party is right there! Let's go! #goWithGod # Tour2014.
- **Second Singer:** And the party starts… I was homesick! Today the party will be in Porangatu-GO city. #goWithGod#carnival2014#hotParty#causingEffects#HotSituation#caseUndefined.
- **Third Singer:** It is coming! With full house, let's go with Thaeme and Thiago Victor Hugo and all rich guys!! # # ticketsAreOver #hotParty.
- **Third Singer:** And the party continues, today is the day to sing and cheer a lot with the guys in Atlanta! Who goes there? # TourUSA2015.

The data-mining model classified the child as not subject to risks.

After the results, those responsible for the child considered no problem in relation with her musical taste. This is one among the advantages that the tool offers, identify the suspicious events and providing to the responsible for the child the opportunity to consider stopping the access to information identified in the mining process, such as music, book, video, and other possible improper things through the Facebook.

The results were shown to the children's parents and served to improve the approach to people. Through interaction between parents, children and our system, we can see that curiosity by both parties was notorious, especially the children's parents about specific words, pictures posted and friends of their children. So we conclude that the approach should be taken with great care not to hurt third of human rights or harm the image of someone because of a spoken word, picture posted or specific action taken within the social network.

5 CONCLUSIONS AND FUTURE WORKS

This work proposes an intelligent tool capable of analyse data from a child's Facebook page, and inform his/hers possible level of vulnerability.

Child sex grooming besides being present on the Internet also takes up space within social networks, this is because of large amount of information and resources offered by these social interaction services, and parents often fail to monitor or track the risks present in this virtual world.

The software developed to combat these risks uses the Jade framework and Weka data-mining tool. In addition to a previous version, three new agents were added (private conversations scanner, external conversations checker and news feed tester) to provide a broader classification model to accept different data from the child's account.

The agents are the brain of the system and make most of the work, by modelling the child's instance profile data so that he/she can be classified as vulnerable or not. All agents follow a communication protocol and the heuristics previously defined in order to find irregularities in the children's data.

The algorithm used for classification and applied in the tests, the J48, was elected to represent the data in a binary tree, thus the visualization and interpretation of how the model works became more noticeable to the project's authors.

Initially, the model used for the instances classification caused differences in the results due to conflicts of information, i.e., instances with similar data were classified as different classes. After a filtering sequence and correct choice of the types of data to be used in the model, the algorithm classification began to return consistent results.

A plausible option for future work would be the test in new children's profiles, with preference for children who have frequent activity on Facebook or contain relevant amount of data for analysis. Moreover, always check for new data and aim to improve the classification process to dirty language within the texts analysed, thus the accuracy of the classification model can be more accurate. The way the data is presented to the responsible for the child could be improved as well to something more

automated, for example, generating a daily report in the form of a pdf file that could be sent daily to them.

For the classification model a possible future work would be adding a sentiment analysis, a model that will be trained with data from Facebook of children (especially messages from chats and posts), for texts made by the child, thus, the verifier would not be so manual and an expansion of the sample space for detection of irregularities would be completed. Based on the training set, the model is able to identify words in new texts related connotations classes for which he was trained, for example, if classes for sexual harassment classification are positive, negative, ambiguous and neutral, the task to classify a word in particular would be unnecessary and thus the context would bring us a more precise idea about the risk.

The use of heat maps or graphics could possibly improve the data presentation in the final report, easing understanding and providing more utility to the analysed data. In conjunction to this, psychology advice on how parents should approach and advise their children would be a way to improve the work.

REFERENCES

Belloni, M.L., 2001. *O que é mídia-educação (Vol. 78)*. Autores Associados.

Bombonatto, Q. 2012. Associação Brasileira de Psicopedagogia. XVI Encontro de Psicopedagogia do Ceará, na *UNICHRISTUS*. Fortaleza, Brasil.

Buckingham, D., 2000. After the Death of Childhood: Growing Up in the Age of Electronic Media.

Bellifemine, F.L., Caire, G. and Greenwood, D., 2007. *Developing multi-agent systems with JADE (Vol. 7)*. John Wiley & Sons.

Fire, M., Kagan, D., Elyashar, A. and Elovici, Y., 2014. Friend or foe? Fake profile identification in online social networks. Social Network Analysis and Mining, 4(1), pp.1-23.

Kietzmann, J.H., Hermkens, K., McCarthy, I.P. and Silvestre, B.S., 2011. Social media? Get serious! Understanding the functional building blocks of social media. *Business horizons, 54(3),* pp.241-251.

Campos, L.M.L., 2014. Mineração de Dados com Detecção de Outliers em Tarefas de Predição de Séries Temporais, *XI Simpósio de excelência em gestão e tecnologia. SEGET.*

Mazman, S.G. and Usluel, Y.K., 2009. The usage of social networks in educational context. *World Academy of Science, Engineering and Technology, 49(1)*, pp.404-408.

Nyce, C. and CPCU, A., 2007. Predictive analytics white paper. *American Institute for CPCU*. Insurance Institute of America, pp.9-10.

Pereira, S.E.F.N., 2009. Redes sociais de adolescentes em contexto de vulnerabilidade social e sua relação com os riscos de envolvimento com o tráfico de drogas.

Quinlan, J.R., 1986. *Induction of decision trees. Machine learning, 1(1)*, pp.81-106.

Yamashita, R., 2009. Facebook4J A most easily usable Facebook API wrapper in Java. Retrieved from http://facebook4j.org/en/index.html.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I.H., 2009. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter, 11(1),* pp.10-18.

Ribeiro, R., Teixeira, M., Wirth, A., Borges, A. and Enembreck, F. A., 2015. A learning model for intelligent agents applied to poultry farming. In: International Conference on Enterprise Information Systems (ICEIS). *Proceedings of the 17th International Conference on Enterprise Information Systems (ICEIS)*, Barcelona, Spain.

Silveira V., Campos, G., Cortés, M., 2014. A Problem-solving Agent to Test Rational Agents - A Case Study with Reactive Agents. In: *International Conference on Enterprise Information Systems (ICEIS)*. Lisbon, Portugal.

Santin, P. L. L., Freitas, C. O. A., Paraiso, E. C., 2011. Análise automática de textos de mensagens instantâneas para detecção de aliciamento sexual de crianças e adolescentes. V. 2, n. 2, p. 43-59, PUC - Paraná,.

Silvio, C., 2012. 38% das crianças no Facebook têm idade abaixo do permitido, *LeiaJá, v. 21, n.7, edição 344*, p.18-22, São Paulo.

Lemieux, V., Ouimet, M. and Pereira, S., 2008. *Análise estrutural das redes sociais.*