ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research
Vol. 08, Issue, 07, pp.21574-21578, July, 2018

ORIGINAL RESEARCH ARTICLE OPEN ACCESS

IDENTIFICATION AND CLASSIFICATION OF CRIMINALS IN THE LIGHT OF ARTIFICIAL INTELLIGENCE

[1,*]Priscilla Labanca and [2]Rodrigo Garcia Topan Moreira

[1]Faculdade de Direito Damásio de Jesus – Departamento de Perícias Criminais, Rua da Glória, 195, Liberdade, São Paulo, Brazil
[2]Faculdade de Tecnologia do Estado de São Paulo - Unidade Praia Grande CEETEPS – Centro Estadual de Educação Tecnológica Paula Souza, Praça 19 de Janeiro, Boqueirão, Praia Grande, São Paulo, Brazil

ARTICLE INFO

Article History:
Received 22nd April, 2018
Received in revised form
19th May, 2018
Accepted 21st June, 2018
Published online 30th July, 2018

Key Words:

Criminology,
Criminal Mind, Identification,
Artificial Intelligence,
Machine Learning.

ABSTRACT

This study aimed to present a prototype system to assist the scientific police to identify criminals. This is a descriptive applied research, with a qualitative approach using the hypothetical-deductive method; developed during the specialization course in Criminal Forensics, in Brazil. The subjects of the study to form the database were extracted from documents published by the Department of Public Safety. Although recognizing that it is not always possible to accurately identify the causal relationship of a particular crime and suspect, the work of the scientific police is of great importance for the resolution of crimes in our society. In order for the work of this important police force to succeed, it is necessary, in addition to the diagnosis concluded through analysis and collection of traces, a computer program that is capable of integrating all the information and traces that were collected during the scientific investigation and offers an investigation and, consequently, giving greater support to the causal link between crimes and suspects. Crime is a social problem, for this it is necessary to invest in information and discuss / negotiate ways to deal with it, sustaining the possibility of reducing the burden of society and the maintenance of justice.

Citation: Priscilla Labanca and Rodrigo Garcia Topan Moreira, 2018. "Identification and classification of criminals in the light of Artificial Intelligence", *International Journal of Development Research*, 8, (07), 21574-21578.

INTRODUCTION

Establish the identity of a person has become a major challenge since the early times. The identification is the most frequent and elementary act of social life. We use all our senses, vision, smell, hearing, touch and taste, constantly in the process of identification, be it with people or things(Araujo, 2003). However, when faced with a specific need to impute responsibility to a person, and this is the goal of the Scientific Police, the term "identification" must be differentiated from "rating". For this identification procedure it is fundamental that there is a method capable of establishing a univocal relation between the elements in question, creating a set of own characters that can differentiate people or things from each other.

Corresponding author: **Priscilla Labanca,**
Faculdade de Direito Damásio de Jesus – Departamento de Perícias Criminais, Rua da Glória, 195, Liberdade, São Paulo, Brazil.

After all, more than just recognizing a person, it is necessary to individualize it, establishing an identity(Araujo, 2003). The classification procedure is the one capable of establishing a relationship between the elements found and analyzed in a particular place where the crime was committed. Identification, as a process designed to establish the identity of criminals has been used since the earliest times and has been modified to obtain identification. Together with the classification, it is possible to verify the profile of the person who committed a crime and its respective criminal type established by the legislation of a particular country.At present, it is possible to verify the existence of some information systems developed to assist in the identification of people whose profiles are criminal. This segment has been mapping crime not only in the State of São Paulo but also in other countries. In the state of São Paulo, for example, cities such as Diadema can be mentioned, where the highest rates of intentional homicides (with intent to kill) are recorded;

according to one of the articles published in the Diário do Grande ABC (Brasil, 2003; Mauro, 2011). It is possible to cite as one of the systems of the public security segment Infocrim (www.saopaulo.sp.gov.br), which carries out thematic mapping for the development of preventive actions of great social involvement, used in Diadema. In addition to Infocrim (www.saopaulo.sp.gov.br), other information systems help the public security segment fight crimes. What is perceived is that these systems do not have full integration, standardization regarding the formatting of stored information, some of them are only used as registers, and others only play the role of storing photos of these criminals, burdening expert procedures when they are requested. In this context, a prototype of an "intelligent" computer program was proposed to identify and classify individuals with a criminal mind profile, using methodology and concepts from the Machine Learning area, which can contribute to improve the analysis of (and their probable remnants left at the crime scene), providing greater assertiveness in forensic procedures, posting their reports in police inquiries.

The forensic police in Brazil: Brazilian forensic police date from the end of the nineteenth century, a period of transition between the Monarchist and Republic regimes. In this period there were many changes, especially in economic areas (with the advent of industry) and social ("social confusion") (Brasil, 1841; Brasil, 1954). To reorganize and educate society the State decided to value and invest in scientific work, hiring academic professionals from the most diverse areas of knowledge. The purpose of these scholars was to identify and recognize among those citizens who participated in the society, citizens whose profiles were violators of the laws and morals and consequently thus making it more efficient, consistent and reliable police investigative activity at the time, giving greater support to the judicial system. Over the decades, the work of these academics has become increasingly relevant, with the need for modernization of scientific research techniques and the hiring of more human resources that are specialized (Brasil, 1994). In 1998 with a reformulated structure, its professionals were named technical-scientific police (forensic police) (Brasil, 2003; Brasil, 2017). The administration of the Public Security Department, assigning it the mission of "being one of the pillars of maintaining the Democratic State of Law, ensuring every citizen within the due process of law, the exercise of its constitutional guarantees backed by evidence in science and obtained with the application of high technology (Conduta, 2010)" Justified the various forms of organization and constant maintenance of scientific knowledge of public safety because the crime was also organized and specialized.

The forensic computing: The A discipline that dates back to the 1980s, it originated in the United States when it was noted that, with the popularization of computers, criminals also emerged who specialize in committing cyber-crimes(Brasil, 1871). Forensic Computing is nothing but the interpretation of information obtained through observation and collection elements (evidence) of computer systems under the criminal point of view. It assists in the composition of consistent and reliable evidence, enhancing support for judgment determination in the courts. It can be considered that the basis of this discipline is closely related to the information security discipline due to strong evidence of crimes committed through the Internet network. Not only of systems for mapping the Internet Network consists of the discipline of forensic computing. Observing the meanings of the word forensics, one finds that one of them is "the use of science and technology for reconstituting and obtaining evidence of crimes." In this sense, it is clear that systems that are not necessarily made to map virtual crimes, can also be classified as a forensic provided, of course, they have the objective of contributing to the identification, classification and identification of criminals.

The Artificial Intelligence: Artificial Intelligence is a subject that is dedicated to searching for computational methods or devices that possess or multiply the rational capacity of the human being to solve problems, to think or, in a broad way, to be intelligent. Its origin dates back to the 1950s at Carnegie Mellon University, where it sought to build machines that are capable of reproducing the human capacity to act and think. In order for this goal to be achieved, researchers at Carnegie Mellon University had to look in other areas for concepts such as mind, consciousness, thinking, acting, and their forms. The main areas studied were: Philosophy (more specifically Philosophy of Mind with special attention to: Identity Theory, Functionalism, Simulated Consciousness), Psychology (in particular Behavioral Psychology, Cognitive Psychology and Social Psychology), Sociology, Anthropology, Neuroscience.
The Artificial Intelligence has areas adjacent such as neural networks, genetic algorithms, robotics, machine learning, Distributed Artificial Intelligence and others. Nowadays one can count on several applications in its most diverse adjacencies, each one with its set of specialized algorithms. This subarea describes the adjacency called Machine Learning, a concept used to construct the prototype of this study. It is a branch dedicated to the construction of algorithms that allow the computer to absorb and reproduce actions carried out solely by humans. In order to reproduce these actions, there are currently many algorithms and techniques, such as Linear Regressions and Logistics, Bayesian Reasoning, Markov Chains (HMM), Transformation Based Learning (TBL), Vector Support (SVM) etc. To make the computer "learn" something, a procedure is performed that consists of three phases:

Training phase: in a very simplified way, in this phase the data is stored in a structure (database) that is inside the memory of the computer.

Classification phase: From the previous phase, the computer will compare and, according to the stored information, will classify the stored data.

Test phase: from external data, i.e. data reported by the operator of the computer program, these will be analyzed, identified and classified according to the operations performed between the data already stored and the external data. In this phase, it is verified if the computer "learned" correctly. The technique chosen to construct the prototype was the Bayesian Reasoning (more precisely, the Naive Bayes Theory that is a derivation of this reasoning), because it is simpler to implement and easy to understand to the reader. It was emphasized that the main objective of this study is to show how Artificial Intelligence - more precisely Machine Learning - may be able to assist in the area of Forensic Computing.

The Naïve Bayes Theorem: Based on the Bayes Theory (also known as Naive Bayes Classifier), it is one that does not take into account the dependencies that may exist between the data, therefore being called naive. In theory, a Naive Bayes prediction will only be considered correct if all independent

variables are statistically independent of each other (which is often false). This statement is justified by the fact that this theorem is (Conduta, Bruno, 2010)

- Based on the assumption that attribute values are conditionally independent given the target value, it
- Assumes that the presence (or absence) of a particular characteristic of a class to which it is not related to the presence (or absence) of any other characteristic, given the class variable, and
- Trained very efficiently in a supervised learning environment.

The estimation of parameters for Naive Bayes models uses the maximum likelihood method, in other words, one can work with the "naive" Bayes model, without believing Bayesian probability or using any Bayesian methods. In this method, independent variables are considered, meaning that only the variations of the variables for each class need to be determined and not the covariance matrix (Bolstad, 2007; Korb, 2011).The structure of the probability model of a classifier (whose origin is a conditional model), in mathematical terms, can be represented as follows (Bolstad, 2007; Korb, 2011):

$(C\backslash F1 \ldots Fn)$

Where:

p is the probability
C is the class

$F_1,\ldots,F_n$ are the dependent variables that to know the result. The problem is that if the number of evidence n is very large, then this application of this formula becomes infeasible. For this to be avoided, it is possible to reformulate it, using the Bayes Theorem (Bolstad, 2007; Korb, 2011):

$$P(C \mid F_1, \ldots, F_n) = \frac{p(C)p(F1, \ldots, Fn \mid C)}{p(F1,\ldots,Fn)}$$

In practice, there is interest only in the numerator of the fraction, because the denominator does not depend on C and the values of the characteristics F_i (where *i* is a number belonging to the set of Real numbers) are given, so that the denominator is effectively constant. The final result of developing this formula will be used for programming the prototype (Bolstad, 2007; Korb, 2011):

$$p(C|F_1, \ldots, F_n) = \frac{1}{Z} p(C) \prod_{i=1}^{n} p(F_i|C)$$

Where:

Z is (the evidence) a scale factor dependent only on F1, ..., Fn, that is, it will be a constant value if the values of the characteristic variables are known.p(C) is the priori probability of the a priori class.

$(F_i\backslash C)$ are the independent probability distributions.

This theorem when applied, has a small imperfection. This imperfection addresses the following question: depending on the amount of evidence contained in the training set, the result of interactions with these calculations results in zero.

To circumvent this problem, an adjustment is made using the Gaussian distribution hypothesis.

MATERIALS AND METHODS

This is a descriptive applied research, with a qualitative approach using the hypothetic-deductive method. It was developed during the specialization course in Criminal Expertise, in São Paulo, Brazil. According to (Minayo, 2016) studies of a qualitative nature seek to understand human phenomena as part of social reality. Brazil had 59,103 people murdered in 2017 (an increase of 2.7% over the previous year). The death toll of every 100,000 people has risen to 28.5 and the number of homicides and personal injuries followed by death has increased, but robbery (theft followed by death) has declined. For example, Ceará is the state with the highest number of deaths in absolute numbers (1,677 more deaths in one year) and percentage (48.5%). Rio Grande do Norte has the highest death rate: 64 per 100,000. Studies show that violence in Brazil kills more than the war in Syria. Many of these crimes do not have a conclusive outcome, due to the great difficulty in relating the causal link with the suspected person. This sometimes causes an innocent to be putting in jail and frees the aggressor. The method for constructing the prototype has the following phases:

Training phase: For this phase the activities were performed - storing the data (creating the database): The subjects of the study to form the database were extracted from documents published by the public security body to form the training base. The data collected were: gender, foot size, height, weight and age. Calculate the probabilities of the classes and make adjustments (apply Gaussian distribution).

Classification Phase: For this phase the activity was performed - calculate the probabilities of the unknown sample (each attribute has a probability).

Test Phase: For this phase the result verification activity was performed. Is the answer to the question: *What is the class most likely to have for the unknown sample?*

RESULTS AND DISCUSSION

The execution of the prototype: To evaluate the effectiveness of the prototype, a crime scene described. The description of the crime scene is a deserted beach with a body identified as male, and the surroundings of this body are traces of barefoot footprints. A couple of these trails were identified as being from the victim, but there are other pairs of traces that were not identified. Below are the photographs that document the location: The first question is whether there really was a homicide crime or whether there was suicide. Starting from the hypothesis that it is homicide then the second question is whose are those traces (footprints)? Suppose that, using the appropriate methodology for the collection of this trace (photographs numbers 2 and 3), the following characteristics observed: Analyzing the victim's state of the body and found traces (footprints), it is clearly ruled out that the victim had not been attacked by an animal. Removing this hypothesis, soon this crime may have been experienced by both a man and a woman, that is, any person. By inserting in the prototype the data collected in the place, the prototype compared the data collected with those that are already stored (training base).

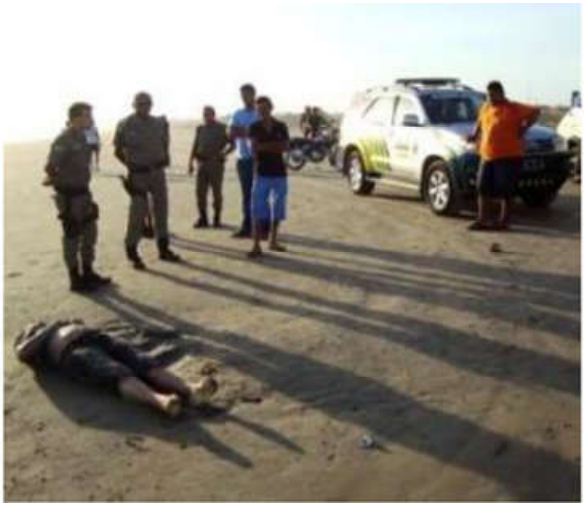

Figure 1. Victim and the scene of the crime scene (GLOBO)

Figure 2. Traces of footprints [www.grupocorreiodosul.com.br/jornal/corpo-de-mulher-e-encontrado-na-beira-da-praia/].

Figure 3.Trail of enlarged footprints (vestige) [https://br.freepik.com/fotos-gratis/pegada-israel-praia_627721.htm]

Table 1.Data collected at the crime scene

Genre	Height (cm)	Size of feet (cm)	Weight (kg)	Age (years)
????	166	37	100	38

The result of this comparison the computer obtained the following answer: *It can be verified that it is a woman*. The response given by the prototype took into account the diversity of physical measures between men and women. Age was estimated according to the vestigial footprint analysis performed by forensic police. Note that with few data it was possible to minimize the universe of possibilities, that is, instead of starting from the presupposition that can be any person, the solution offered a direction, reducing this universe to a group (a class) of people (in the case some offender is female).The idea is that the more data that is stored in the database, the classes of this solution will be more accurate, making it easier to make decisions about investigations and reducing the time taken to capture those individuals whose profiles are criminal.

Final Considerations

The various forms of organization and the constant maintenance of scientific knowledge in the area of public security is justified by the fact that crime has also organized and specialized, consequently making the forms of transgressing the most complex laws and requiring that the techniques of police investigations were clearer. To this end, law enforcement agencies are increasingly counting on computational solutions to assist them, not limited to, and investigations of virtual crimes. Using the concepts of Machine Learning (adjacency of the discipline of Artificial Intelligence), it was observed that it is possible to contribute with a support solution. Proof of this was that with few data informed in the prototype developed, it was possible to minimize the universe of possibilities before a criminal investigation process, that is, instead of starting from the presupposition that can be any person, the solution offered a direction, universe for a group (a class) of people. The more data that is stored in the database, the more the prototype classes will be accurate, making it easier for decision making and targeting of investigation processes, reducing the time taken to capture those individuals whose profiles are criminal.

REFERENCES

Araujo, 2003. Marcos Elias Claudio. Histórico dos Processos de Identificação. Instituto Nacional de Identificação. Departamento da Polícia Federal, 2003.

Bolstad, 2007. William M. Introduction to Bayesian Statistics. 2nd ed. New Jersey: John Wiley & Sons.

Brasil, São Paulo. Decreto nº 47.694 de 7 de março de 2003. Autoriza a Secretaria da Segurança Pública a, representando o Estado, celebrar convênios com Municípios do Estado de São Paulo, objetivando o desenvolvimento de programas municipais para prevenção do crime e da violência. Administração Pública. Convênios. Diário Oficial v.113, n.46, 08/03/2003.

Brasil, São Paulo. Lei Complementar nº 756 de 27 de junho de 1994. Organiza a Superintendência da Polícia Técnico-Científica e dá outras providências correlatas.

Brasil, São Paulo. Lei nº 2.034 de 30 de dezembro de 1924. Reorganiza a Polícia do Estado.

Brasil. Lei nº 2.033 de 20 de setembro de 1871. Altera diferentes disposições da Legislação Judiciária

Brasil. Lei nº 261 de 3 de dezembro de 1841. Reforma do Código de Processo Criminal.

Brasil. São Paulo, Superintendência da Polícia Técnico-Científica do Estado de São Paulo. SPTC – Superintendência. www.polcientifica.sp.gov.br

Conduta, Bruno C. *et al*. Aprendizagem de Máquina. UNICAMP. Faculdade de Tecnologia – FT. Programa de Pós graduação modalidade stricto sensu em Computação. Disciplina de Inteligência Artificial, 2010.

Correio Do Sul. Corpo de mulher é encontrado na beira da praia. www.grupocorreiodosul.com.br/jornal/corpo-de-mulher-e-encontrado-na-beira-da-praia/

Costa, Marcos Ferreira Guedes. Evolução dosDireitos Humanos nas ações das forças de segurança latino-americanas: uma análise comparada das polícias das Repúblicas da Argentina e do Brasil. USP. Programa de Pós graduação stricto sensu em Integração da América Latina, 2001.

Freepik. Pegada Israel. https://br.freepik.com/fotos-gratis/pegada-israel-praia_627721.htm

Globo. Mulher encontrada morta na praia em Luís Correia. radioglobоparnaiba.blogspot.com

Korb, Kevin B., Nicholson, Ann E.. Bayesian Artificial Intelligence. 2nd ed. EUA: CRC Press, 2011.

Mauro, Mário César de. Convênio traz Infocrim a Diadema. Diário do Grande ABC.www.saopaulo.sp.gov.br

****