*International Journal of Engineering & Technology*, 7 (3.20) (2018) 932-934

International Journal of Engineering & Technology

*Website: www.sciencepubco.com/index.php/IJET*

***Research paper***

# Identify Attacking Attempts through Logging on Websites and Database Systems by Machine Learning

**[1] Hasan Jameel Alwan, [2] Barakat Saad Ibrahim, [3] Ahmed Hamid Shaker**

*[1]Centre computer /Al-Muthanna University*
*[2]Al-Muthanna University*
*[3]College of Education for Pure Science / Al-Muthanna University*
**Corresponding Author Email: Ba2011sa@yahoo.com*

**Abstract**

we will focus on the possibility of using and applying multiple technologies of information technology including, bank accounts, pattern recognition, follow_up data, and data mining. Improve the performance of these systems and techniques through new methods of learning and training, and thus improve the performance by reducing the costs of educational sources and the preservation, study and evaluation of the previous results through conversion to machines and machine learning technology.

***Keyword:*** *machine learning; cognitive computing; data mining; artificial intelligence; pattern recognition*.

## 1. Introduction

Information technology plays an important role in modern society, nowadays, it is used in any field of human life. Among the digital and informational technologies in our world today, a special and growing position is being taken by machine learning technology. The beginning of this technology where in the article "an essay towards solving a problem in the doctrine of chances"(1). After that, Markuv series opened and the smaller grid method. The real beginning of the development of machine learning is in 1950 when Turing published an essay "computing machinery and intelligence"(2), which suggests that it can be learnt and the idea formed, this was the prediction of the genetic algorithm. The learning technology of machine learning will also be analysed and applied in the field of optimal problem solving. The basic purpose of intelligent data collection is to increase the amount of information obtained(according to a selected evaluation method) and reduce the costs of preparing of interesting situations. In this research, we are talking about a large part of the decision making power needs for a set of measurement, computing, and data analysis based on the concept( intelligent autonomous agent)(3) and (multi_agent systems)(4). With the provide of machine learning and data security along with a summary of machine learning structures and to apply its methods in the classification and the security of data and to discuss and segmentation of the problem to find the best solution by applying these methods on a vast amount of data and how to reach the electronic objective through the use of electronic models biometric using the algorithms of machine learning to reduce the applied error in a given data sample or training. In addition, talking about the most famous methods of machine learning algorithms in cybersecurity by (learning under a supervision) through pairs(object, response) classified on the basis of objective structures and functions of learning or training algorithms.

## 2. Research methods:

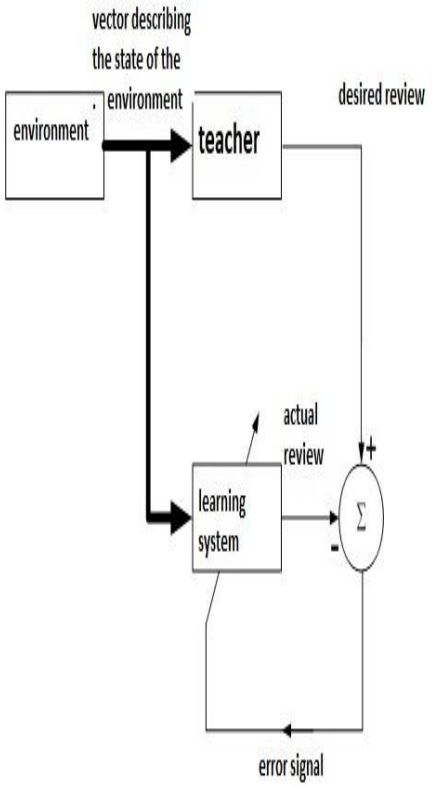
It includes a set of methods of factor_ based modelling of mathematical economics, aggregate optimization, and machine learning. In addition to the advantages of building the information society, the risks associated with data security are increasing. As the security of information resources from unauthorized access and the removal of information by technical conclusions, information technology and communication systems are also one of the main concern in the field of information technology. In this regard, there is a need to develop modern methods and systems to secure information from different types of threats in all these systems. A large number of tools and information security systems are created on the basis of mathematical models using digital calculations on this purpose.

## 3. Main topic

Machine learning is a broad section of artificial intelligence that studies methods of building algorithms that can be learned. Machine learning is at the intersection of mathematical statistics, methods of optimization, and classical mathematical disciplines, but has its own specifies related to the problems of computing efficiency, relearning and training. Many methods of inductive learning have been improved as an alternative to classical statistical approaches. Many methods are closely related to the clouds of information and the intellectual analysis of data. Machine learning is not only mathematical, but also practical, it is an engineering system(4). Textual theory as a base does not immediately lead to methods and algorithms that can be used in practice to make it work well. We have to propose additional mechanisms to the discrepancy between the theoretically presented assumption and the circumstances of the real problem(5). Practically, there is no research in machine learning

that can be done without a test on a real data model, which emphasizes the practical ability to work. Learning programs are conducted through relevant and appropriate tasks that divides into four major categories, depending on learning characteristics:

1_Learning under supervision(learning with a teacher): The idea is that there is some correlation between the responses and the objects, but it is not known here only as a limited set of precedentspairs(object, response)which is called the sample of learning. On the basis of this data, it needs to recover accreditation. This will be to build an algorithm capable of being able to each object be given a fairly accurate response. To measure the accuracy of the responses in a particular way, the quality of the function is presented. The teacher understands either the learning sample itself those who have an object that knows exact responses.




**Figure(1):** Education of the teacher.

2_ Learning without a teacher (spontaneously):

Learning without a supervisor is not the solve of the problem in what we do not know but describes a group of objects. Examples of these functions are data collection and visualization.

3_ Learning with enhancements, enhanced learning(learning enhanced): study how a worker must work in an environment to maximize long gains(6). The algorithms of micro_learning are trying to find a strategy to make sure of selecting the best agent in each case of the environment in economics, and the theory of learning games seen as learning to enhance it as an explanation of how to create balance. Formally, the simplest model of learning with enhancements should consist of the following:

A_ sets of the environment situation S

B_ sets of the actions

C_ sets of responses

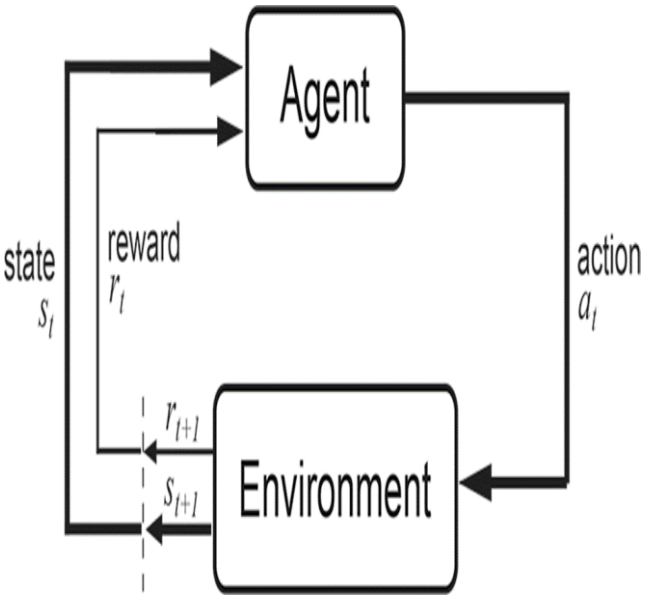
In general, enhanced learning works as follows: (figure)

1_ The agent comes to the situation the environment.

2_ The agent selects the actions.

3_ The mediator sends the reponse to the agent and the next situation.

4_ The agent processes the response and corrects its actions.




**Figure(2):** Interactions between client and environment.

4_ Differentiation and other categories of machine learning: this includes learning on training, evolutionary learning, and robotics learning, etc. A basic example is one of the main ways to data entry and login in websites is the model of this type of learning. When you click on the button( log in, for example), the data is entered to be sent for further processing by the server.

## 4. Methods of data security using machine learning:

Forms are an attractive and highly desirable target for promoters of spam and unwanted messages. In order to limit this exploitation, we will examine the messages, and what was the entry of data in any part of these forms in this type of message to be harmful? Which leads to the profits to be brought to the promotional messages or does not correspond on the purpose of the website and its subject matter.

## 5. System architecture:

The system can be included with the system of logging in sending messages ordering orders and other systems for most of the website resources. It consists of two sub_system: the agent and the server.

The client part is executed in script java language. Its main purpose is to collect data on user hand movement and send it for further processing and classifying on the server(7). The server part receives the input data collected the decision on assigning the user to one of the two " person or robot". The registration sub_ system will activated when the user presses the send button on the server. At this moment, a special task appears on the screen before recording all its movements for further classification. This task must have a particular element to avoid the situation in which spammers have passed even once by their own hands, recording all their movements with the help of a special program to track and monitor the change.

## 6. How to solve the problem:

The point pi will be the triple ($x_i$, $y_i$, $t_i$), where xi ∈ Z is the coordinate of the i-point along the abscissa, $t_i$ ∈ Z is the coordinate of the i-point along the coordinate axis, $t_i$ ∈ R* is the registration time i-points.

Definition 1. π a length N we mean a complete sequence of points (p) N, for which the following property is satisfied: $\forall_i = 2$, N: $t_{i-1} < t_i$

The length of the path π is denoted | π |. The set of all paths of length n € N is denoted ∏n. Then $\prod_* = \bigcup_{n=1}^{\infty} \prod n$ is the set of all paths of arbitrary length. We denote $\prod* = \bigcup_{n=1}^{\infty} \prod_n*$ the set of all tuples of arbitrary length, whose elements are paths of arbitrary length.

For convenience, we denote the classes of users as follows: +1 _ " HUMAN ", -1 _"ROBOT". Suppose that there is an unknown y (π *) dependence between the set of paths received from the user and the class to which it belongs: y: Π* → {-1, + 1}.

Then, in order to construct the recognition system, we need to construct an algorithm α (π *) that can calculate the value γ (π*) ∀π* ∈ Π * from a finite number of steps. Now we need to decide how we can construct the algorithm α (π*).

It is obvious that there exist an infinite number of tuples π* ∈ Π*, which are characteristic both for a robot and for a human. In addition, the structure of such an object is difficult to construct an explicit and precise algorithm that could unambiguously classify it as one of the classes. Therefore, we will choose another strategy: we will form the set $\prod_{train}$ ב $\prod*$ which will consist of both data characteristic for people and according to the characteristics of bots and try with its help to construct an algorithm α (π*) that would approximate the true the dependence γ (π*) over the whole set Π *, in order to estimate the quality of the algorithm obtained at a certain stage of construction, we will use the set Π train for which the correct answers are not known [9]. Complex mathematical tools capable of solving this problem is called machine learning, and our task most accurately corresponds to a class of problems, called supervised learning.

In order to reduce the data security, in this situation to determine the paths of the machine learning task with the teacher, it is necessary to follow these steps:

–Collecting a training sample consisting of persons (human) and robotic(march) data.

–Include description tasks of object properties.

–Select a classification algorithm that would give the best.

–Select a scale to evaluate the quality and grade if the model obtained. A variety of machine learning techniques have the ability of an impressive space of a large database, and this vast space is rapidly evolving. According to the new concept, most(ml) is a subset of artificial intelligence(AI). It can be learnt and applied effectively in anti_spam and using it to detect fraud, security risks, secure personal information, and for many other purposes(8). Here are some examples that can be referred to as result of the use of machine learning practically in database security:

_machine learning is used in banks and other financial institutions for two main purposes: defining data relationships and preventing fraud. Acquired knowledge can help identify investment risks which guide users to the most profitable ones. Intelligent data analysis facilitates the identification of in solve customers, it can also be used as an electronic monitoring tool to prevent and detect fraud.

–Security government agencies work with many heterogeneous data sources, so the applying of machine learning is appropriate and solves important tasks in this field such as, detecting fraud and removing its consequences with minimal harm.

–The advanced applying of machine medical learning methods is due to the large number of devices and sensors that help to track the patients situations in real time, which generates using machine learning techniques that help doctors to find important factors in finding the right diagnosis.

## 7. The experience risk:

The function of the algorithm on the learning sample distinguishes the quality of the approximation of a given function on the sample and equals the average value of the loss(total lass function values of the learning sample divided by the sample length)(11). In anther words, the experimental risk is the average error value of an algorithm in the learning sample. Thus, the educational objectives are to process the problem of optimization, and to find the value of the minimum and loss function and the approximation function, that is to reduce the problem of the digital optimization techniques.

## 8. Conclusions:

In this paper, a system was proposed to identify a follow_up model based on the biometric characteristics of human hand movement when the input button is pressed during the log in process in websites using machine learning to secure the input data. All phases of solving the problem of machine learning were implemented, and a model capable of detecting plagiarism was obtained with very high accuracy. As indicators of organisms, statistical indicators were used of the main characteristics o the movement. The general structure of the follow_up model identification system was proposed based on the model developed in this paper. The results obtained indicates such a system will be able to provide a high level of security against spam messages and will be simple and convenient to use.

## References:


[1] Bishop C.M. Pattern Recognition and Machine Learning / C.M. Bishop. — NY: Springer. — 2016.

[2] Bishop Christopher. M. Pattern Recognition and Machine Learning / Christopher M. Bishop.– New York:. Springer-Verlag, 2016. - 738 p

[3] Fukunaga K. Introduction to statistical pattern recognition / K. Fukunaga // Academic press, 2013. — 591 p.

[4] Ghamisi P. Advanced Spectral Classifiers for Hyperspectral Images: A review / P. Ghamisi, J. Plaza, Y. Chen et al // IEEE Geoscience and Remote Sensing Magazine. — Vol. 5, N 1. — 2017. — P. 8–32.

[5] Dua, S., & Du, X. (2016). Data mining and machine learning in cybersecurity. Auerbach Publications.

[6] LeCun Y. Deep learning / Y. LeCun, B. Yoshua, H. Geoffrey // Nature. — Vol. 521, N 7553. — 2015. — P. 436–444

[7] Lorenc C. Sensitivity to observations in the Met Office Global numerical weather prediction system / A. Lorenc, R. T. Marriott - Meteor. Soc., 2014.

[8] Maulik U. Remote Sensing Image Classification: A survey of support-vectormachine-based advanced techniques / U. Maulik, D. Chakraborty // IEEE Geoscience and Remote Sensing Magazine. — Vol. 5, N 1. — 2017. — P. 33–52.

[9] Paul Nolan. Impacts of climate change on mid-twenty-first-century rainfall in Ireland: a high-resolution regional climate model ensemble approach / Paul Nolan, John O'Sullivan, Ray McGrath-International Journal of Climatology 26., 2017

[10] Лутц М., Программирование на Python, том II, 4-е издание. – Пер. с англ. – СПб.: Символ-Плюс, 2011. – 992 с.

[11] Профессиональный информационно-аналитический ресурс, посвященный машинному обучению..2018

[12] Саттон Р.С Обучение с подкреплением / Саттон Р.С, Э. Г. Барто // БИНОМ, Лаборатория знаний, 2014 – С.42-96.