UNIVERSITÀ POLITECNICA DELLE MARCHE

Faculty of Engineering
Department of Information Engineering
Master of Science in Biomedical Engineering

Cyber risk assessment of complex infrastructures through machine learning-based techniques

Supervisor
Prof. Marco Baldi

Candidate
Simone Compagnoni

Co-Supervisor
Massimo Battaglioni

Academic Year 2021-2022

Abstract

Modern society is trending really fast towards a digital future, in which data are produced, collected and exchanged digitally. Organizations are the primary players who gather and process data, carrying with them all the issues related with data protection. To solve these issues, cyber risk assessment methodologies are essential for determining the current level of data protection as well as the eventual requirement for a cyber security update. In the healthcare field, hospital and healthcare infrastructures in general have become a target of cyber attacks, due to the sensitivity of the data that are stored and processed, making it necessary to implement cyber security procedures. The goal of this work is to create a machine learning model that can perform cyber risk assessment in the most objective and simple way possible, considering as inputs maturity, complexity, and attractiveness of an organization and 2 classes of risks as output. The development of such a model has been carried out through MATLAB environment, exploiting the "classification learner" functionality. A database containing 420 companies has been used to train and test the machine learning model. As a result, exploiting different algorithms, several machine learning models have been obtained, the best of which is characterized by an accuracy of the 78.6%. These results show the possibility to perform cyber risk assessment with a machine learning approach, pointing out possible improvements exploiting a database containing more entries.

List of contents

List of figures

List of tables

1. Introduction

The current society, to work properly, is based on the processing and exchange of data. The term data is referred to information, that can be collected and analyzed, used to drive algorithms involved in decision-making or stored into computers to be used when necessary. In the past decade, the world has developed towards a digital direction, with the advent of computers and mobile phones and their increasingly large-scale use. This has led to an increase in the production rate of information on a daily basis: an estimation has stated that roughly 2.5 quintillion bytes of data are generated every day [1]. After being produced, data can be exchanged, processed, and utilized for the most disparate purposes. Data is collected by governments, businesses, and individuals. It can be used to track trends, understand human behavior, and make predictions.

On the one hand, it is possible to think about data that people exchange with each other, for instance the information contained in messages that are exchanged using the messaging services between users. On the other hand, a great quantity of data is processed, treated, and exchanged between companies. Companies collect an extensive amount of data from people: data about patient's health, financial data, technological data, and so forth. Companies can use data to track trends, understand human behavior, and make predictions. Data can be bought, sold, and traded. They are used to make decisions about everything, starting from what products to sell up to where to build new roads, just to make example.

Data processing is an essential component of many companies, since it is the act of taking data and manipulating it in order to extract meaning from them. In most cases, this process involves changing the data to provide them with a better suit with respect to the purpose for which they have been collected. Every time a user is connected to the services provided by a company or by an app, the company collects information about the user. To process the collected data, the individual must give their consent. The consent can be asked by the company in an explicit way, such as when someone is ticking a box to opt-in to a service, or in an implicit way.

The permission for the processing of the data is not enough to ensure its security. Consequently, information collected by companies is regulated by state legislation. Regardless of the differences that can arise between laws of different States concerning data

treatment, the basic principles are the same. Companies that process data must guarantee the users pseudonymization, availability, confidentiality, integrity, and protection [2].

Pseudonymization is the data management and de-identification procedure by which personally identifiable information fields within a data record, one or more fictitious identifiers, or pseudonyms, are substituted. This identifier can be used to link data across different datasets, but it cannot be used to identify an individual. Pseudonymization can be achieved through different techniques such as encryption, scrambling, masking, and so forth.

Availability is an important property that is closely linked to the fact that the only owner of the data is the one who produces them. Hence, companies who collect data are forced to have this data available when requested by the user who produced them and, therefore, at any moment.

Confidentiality can be defined as the practice of keeping sensitive information private unless the owner or custodian of the data gives explicit consent for it to be shared with another party. Consequently, companies aiming to share data with third parties must always ask the owner for consent.

Integrity refers to the consistency and security of the data in a physical and logical way. With physical integrity it is intended the protection of hardware that allows the storage of a large quantity of information. On the other hand, logical integrity aims to keep data unchanged, protecting them from human error and hackers.

Data protection is the process of safeguarding data from corruption, compromission or loss. Data protection assures that data is not corrupted, is accessible for authorized purposes only, and is in compliance with applicable legal or regulatory requirements. Companies should not lose data and more importantly they should prevent any unauthorized person from getting hold of them, in particular a malicious attacker [2].

The massive exploitation of data and information systems in companies and organizations is motivating an increasing attention to cyber security and its management.

Cyber security is an informatic field that concerns all the tools and technology whose function is to protect data from attacks from an entity located outside the company that is processing them [2].

Strictly correlated with the concept of cyber security, cyber risk can be defined as risks correlated to information and technology assets with implications for the confidentiality, availability, integrity of information or information systems [3]. Accordingly, since nowadays all the information is stored and treated by informatic systems, cyber security has

become a tool of primary importance for the defense of companies against possible cyber attacks aimed at the subtraction of data. As a result, in order to be safe from cyber attacks, each organization should implement a cyber security system commensurate to the sensitivity and value of the processed data. It is important to mention that cyber security does not reduce the risk of being targeted by a successful cyber attack to zero, but it lowers the probability that this can happen. A good cyber security system provides a company with a shield to protect against hackers and acts as a deterrent: a hacker is more likely to target a company which does not a have a good security system.

Cyber security assessment can be considered a tool of primary importance to determine any weaknesses in the IT system (Information Technology) of a company, which can be used by hackers to carry out successful attacks and, consequently, to data loss. The cyber risk assessment procedure is aimed at analyzing the whole company IT system, allowing to have awareness about the security status of the IT system and to improve some of its aspects [4], if necessary. The reasons that can lead an attacker to attempt a cyber attack against a particular company are correlated with the importance of the data processed by the IT system and its security posture. It is, in fact, more likely that companies that process high level data (bank, hospital, finance company) and that do not possess an appropriate cyber security system may become target of a hacker.

This work attempts to develop a machine learning model that can classify companies according to their level of cyber risk and provide information about it. The first step in the development of the model has been to understand which parameters might have a relationship with the security posture of a specific organization. The factors that have been selected are the company's attractiveness, employee count, and security rating.

In the second stage, the relationship of these parameters with the security posture of a company has been outlined. Security score represents the level of security of a company; the higher it is the safer a company is against cyber attacks. The number of employees for the purposes of this work has been correlated to the security posture of the company with an inverse relationship: a great value of this parameter has a negative impact on the security posture of a company. Attractiveness refers to the nature of data that a company is processing. Since there are companies that process sensitive data, it is more likely that they will be more targeted by cyber attacks. Finally, a machine learning model that can classify organization into classes of risk has been developed and trained using these factors in conjunction with proper parameterization.

1.1 Thesis organization

This work is organized as follows. In section 2, a definition and description of cyber attack is given, providing some examples of possible attacks against healthcare organizations. In section 3, the main available cyber risk assessment methods are mentioned, pointing out their functionalities and the possible interaction between each other. In section 4, the materials that have been employed in the development of the machine learning models are listed, while section 5 lists the approaches used to create the machine learning model. The findings of this research are provided in section 6. Section 7 presents two case examples to demonstrate the significance of the complexity index and the output of the machine learning model to changes in complexity and attractivity. The limits of this work are addressed in section 8. Finally, section 9 shows the conclusion of this work.

2. Cyber attacks

The current society is rapidly moving towards a complete shift to digital technologies. A digital society has already arrived in many ways. Digital platforms are already being used to communicate with friends, access information, work and study, shop, and bank. The Internet connects the entire world, resulting in the creation and maintenance of a massive network that generates billions of dollars annually [5]. Vital and essential parts of modern society are either created or implemented in cyber space [6], a virtual space in which users are connected through telecommunication networks. In the cyber space, data is collected, processed, and used to perform the most disparate tasks. Operations within cyber space are controlled by a small number of individuals, and users do not have the possibility to deeply control the software they are using [7]. The absence of controls that citizens have when utilizing the internet poses a hazard to them since a hacker may seize control of the software they are using and gather their sensitive data. As a result, there is now a connection between the life of citizens and the security of the data processed in the cyber space [8], due to the sizeable volume of government or citizen information that is processed there today. This new trend has subjected world governments to new challenges concerning the tracking of information and actors dealing in cyber space. The anonymity, which is the key on which the cyber space and the cyber security in general are based, has as drawback the difficulty that governments have when it comes to deal with hacker, organized terrorist groups and treats as cyber crime and cyber warfare. For what concerns cyber attacks, there is a widespread scenario regarding the kind of attacks and the damage, both in a physical or economical point of view, they can cause. Cyber attacks can be defined as a particular category of digital attacks that target computers, networks, or data. A cyber attack can be carried out in many forms: phishing, malware, virus, etc. Phishing is the practice of an attacker tricking a user into disclosing personal data, passwords or credit card details. Malware seeks to interfere with or take down a system. A virus is a unique piece of computer code that can replicate and propagate within a computer and to other computers connected to the same network. Cyber attacks are becoming more sophisticated and destructive, and the possibility of a successful attack cannot and should not be underestimated. They may cause financial harm by stealing a particular quantity of private information, or they may cause physical harm by sending false messages that prevent one or more organizational components from functioning properly

and create structural damage [9]. It is also worth mentioning that cyber attacks do not only represent a problem for world's governments, but also a tool: they can be employed to lay the groundwork for any unrest and popular uprising, or directly to damage or disabling equipment and facilitating physical aggression of a country by another country or a minority of people [10]. Foreign intelligence services also exploit the possibility related to cyber space to gather information and perform espionage activities [11], by collecting information about other governments. In addition, groups or individuals with malicious intents, hackers, can exploit the knowledge and the weakness of a particular space within the cyber space to steal data, from which they can make money [12], or express themselves, stating their purposes and the way in which they operate to achieve it. As mentioned in [13], another source of cyber attacks can be internal dissatisfied agents. In this case, the attacker does not even need to be particularly skilled when it comes to cyber attacks because, by virtue of being a part of the system they wish to target, they most likely already know how it operates, its flaws, and the credentials needed to access a specific sort of data.

The way a hacker or any group of people can perpetrate a cyber attack towards a particular organization depends on the purpose of the attack and on the security posture of the organization. In particular, modalities with which a cyber attack can be executed are [13]:

- Denial of service
- Logical bomb
- Abuse tool
- Sniffer
- Trojan horse
- Virus
- Worm
- Botnet

In the denial of services, the authorized access to a system is lost, due to a message spam that an attacker may send blocking the normal data flow [14]. The logical bomb is basically a code, which is inserted inside a system by an attacker. The code is then implemented inside the system and, when the system executes that part of the code, the logical bomb performs a destructive activity [15], which can damage the code of the system and its functionality with it. Abuse tools are correlated to internal dissatisfied agent, in which the attacker is part of the system that he is attacking. Therefore, he has knowledge about the tool that he can use to cause damage to the system or subtract data from it, without the requirement to have any

relevant knowledge about cyber attacks. Sniffer attacks are based on the eavesdrop of routed information with the aim of finding passwords and credentials [16], that can be used to get the unauthorized access to a particular component of a system. A trojan horse hides a dangerous code that is hidden under the shape of a useful program that the user may be intended to run [17]. When the user runs the program, he unconsciously activates also the code of the trojan, allowing its execution. Worms and viruses are both based on the multiplication and the spreading inside the system in which they are implemented. The difference between worms and viruses is that worms do not require human intervention to multiply and spread, while viruses do [18]. Viruses spread by attacking a particular code and becoming part of it. When the code is run by the user for any reason, the virus is capable of infecting other components of the system or other systems. Lastly, botnet is referred to an infected remote-control system that is installed into a user's computer. Therefore, the attacker can control and perform activities on the remote computer [19].

2.1 Cyber attacks against healthcare organizations

Since society is trending more and more towards a digital future, in which data are processed and stored digitally, also healthcare organizations are adopting new technologies to deal with medical and personal data and to allow better screening and operating techniques to be performed. The Internet of Medical Things (IoMT), which refers to devices connected to healthcare IT systems via network connections, is something that the pharmaceutical and healthcare industries are already part of or are transforming into [20]. The definition of IoMT is strictly related to the one of Internet of Things (IoT) which is: "A dynamic global network infrastructure with protocol-based self-configuration capabilities, standardized and interoperable communication systems, in which physical and virtual objects have an identity, physical attributes, virtual personalities, which use intelligent interfaces, and which are perfectly integrated with computer networks" [21].

IoMT refers to the application of this definition to the healthcare industry and it is spreading rapidly among healthcare organizations, since doctors and patients increasingly utilize connected devices for routine medical functions [20]. The integration of new technologies has upgraded the quality of services of hospitals in terms of better diagnoses and better treatments. In the last decades, computers have been introduced into healthcare environment, allowing medical imaging, the early discovery and better treatment of many diseases [22].

In early 2000, imaging techniques have been ranked as one of the most important discoveries in the medical field in the past 1000 years [23]. At the same time, implantable medical devices, such as cardiac and neurological implants with wireless connections that can be programmed and controlled, began to be adopted alongside medical imaging techniques. [24]. Medical imaging techniques, together with implantable devices and IoMT, have been an important improvement in the medical field, allowing better and faster services delivery to the patient. Together with the benefits mentioned so far, these new developments also brought drawbacks concerning cyber security.

2.1.1 IoMT-based attacks

For what concerns IoMT, the interconnections of devices through the global network can be exploited by groups or by a single individual to perform cyber attacks. The vulnerability of the infrastructures is related to the lack of proper cyber security algorithms to manage or prevent cyber attacks, which polarizes the attention of possible attacker [25]. In most cases, the global network-connected devices are released in a dangerous state, putting patients and medical professionals at considerable risk in case of a cyber attack.

Furthermore, connection between the different components of an IoT system are achieved through session keys, which are generated by a cryptographic algorithm. Session keys are used to identify a user, confirm that the user has authorization to use the system, encrypt data, and ensure the security of data transfer across systems. Each time an authorized user connects into the system, a session key is generated and used to encrypt the data being processed by the user so that only the user and the server can decrypt it. The key is lost after the user logs out. The hijacking of session keys [25] is a cyber attack that can be carried out in this situation since it allows a hacker to obtain the session keys needed to access various network components. If an unauthorized person obtains a session key, this knowledge can be exploited to access information from a healthcare database. Another possible attack against IoMT infrastructures listed in [25] is the ransomware, which can block the access to a system component of a hospital by encrypting it. If the decryption cannot be performed in a small amount of time or at all, this can lead to a delay of scheduled appointments of patients. Another possible attack is the denial of service [25], in which an attacker aims to paralyze the hospitals services. This can be achieved by stopping the functioning of

computers inside the network or by sending a big load of data traffic to a device to clog it and therefore stop its correct functioning.
As explained in [25], the main problems that cyber attacks can generate are:

- System Failures
- Network Failures.

For what regards system failures, a cyber attack can cause failure of a medical device, which, depending on the device, may lead to an interruption of real time data collections, as blood pressure or glucose concentration monitors. Therefore, due to the interconnection of all the devices belonging to the IoMT, a failure of a medical device can also have an important impact on the functionality of the entire system.
Network failures are serious issues, since all the devices connected to IoMT rely on the connection to a global network for the proper functioning. A network failure can stop, for example, the correct functioning of cardiac or radiologic devices.
The damages that a cyber attack can cause to a healthcare facility connected to the global network are, as reported in [25], the paralysis of the instrumentation, paralysis of the information system and, in the worst-case scenario, the loss of human life.

2.1.2 Attacks against medical devices

Medical Imaging Devices (MID) have been a fundamental implementation of the early 2000', upgrading the quality of the examinations and consequently the quality of the possible follow up surgery. Two of the most crucial medical imaging technologies, computer tomography (CT) and magnetic resonance imaging (MRI), are more linked, computerized, and vulnerable to cyber attacks now than ever before [19]. Since the current society is trending to become a digital society, the number of possible cyber attacks is exponentially increasing: each year more than 120 million of new malwares are discovered [22]. Medical imaging equipment is at risk from the development of new malwares and, subsequently, new ways to conduct cyber attacks, necessitating continuous updating. Unfortunately, neither the CT nor the MRI are cyber security oriented and performing a complete upgrade can take years [22], making these devices vulnerable. An example of an attack towards MID is the WannaCry attack [26], which has been carried out in May 2017 and it infected more than 200,000 devices in 150 nations [27], targeting also MID devices such as MRI and CT [28].

The attack was based on the encryption of devices, with the aim of making them non-operational. This caused a lot of logistic problems such as rejection of patient [29] and diversion of ambulance routes, leading to chaos and delays [30]. Another example of attacks against MID was carried out by the hacker group “the shadows Brokers” [31]. This attack managed to infect Microsoft systems in such a way that a code could be run into computer by remote. This attack had a strong impact on all the MID devices. Attacks that target CT scanners in particular may be more harmful than attacks that target MRI since ionizing radiations is used during a CT scan. Because the host computer controls the CT scan process, infecting it might allow a hacker to manage the process remotely, jeopardizing the patients. [22]. Another possible attack that may be performed involves the control from hackers of the movement of the component of a CT scan, causing the focus of the x-ray beam to move away from the target area [22]. This effect might be hazardous since the amount of x-rays that should be applied to a certain place relies on the characteristics of that area and applying the same amount to another site could be harmful. CT scans use a reconstruction algorithm to provide imaging; an attack can be carried out to alter the proper outcome of the reconstruction procedures[22]. Small changes in the outcome of the reconstruction procedures are impossible to tell, leading to a wrong evaluation of the clinical situation of a patient. Once in possession of the system that permits the reconstruction, a hacker can also link a specific reconstruction to a different patient, which results in a misestimate of the patient's clinical status [22].

2.1.3 Attacks against implantable devices

Implantable devices are nowadays widely used in different clinical applications. The newer developments in this field have led to the control and monitoring of implantable medical devices (IMD) through smartphone, employing wireless technologies [24]. As for what concerns MID devices, also implantable devices and their controls are not cyber security oriented, which make them vulnerable to possible cyber attacks. A cyber attack directed towards an IMD may allow a hacker to change the settings on which the IMD is working, which makes the IMD dangerous for the patient [24]. Together with the possibility to harm and kill patients, another issue that may be caused by a cyber attack is the loss of patient confidence to undergo a procedure for installing an IMD. Although there are currently no documented attacks against IMD [24], given how widely used they are in clinical practice,

it is conceivable that they will become a target in the future. The communication between the base-station and the IMD device is usually not protected by encryption and even if it is, the fact that these connections follow a static pattern can be exploited by expert hackers to decrypt and collect information [24]. The information that IMD devices contains not only concerns the setting on which the device is working, but also some personal information of the patient such as the name and birthdate [24].

2.1.4 Data Mining

Healthcare facilities, in particular hospitals, maintain a large amount of data online in databases. Although they are designed with some degree of cyber security due to the high sensitivity of the contained data, these databases are a target for nefarious attackers. In the Community Health Systems (CHS) data breaches, more than 4.5 billion of records have been stolen with an economical repercussion of about 150 million dollars [32]. As reported in [33], from 2012 to 2015 U.S. hospitals have been subjected to data breaches 196 times. The data that are usually stolen are the patients' name, birthday and address; in particular, the patient's name is stolen in almost all data breaches [33]. The authors of [33] have investigated the possibility of hackers making money by selling the data they have stolen. The platform in which it is easier to sell stolen data is the dark web, which contains all the information or sites that are not indexed on the common search engines. The dark web hosts a variety of information, including both illicit operations as drug trafficking and money laundering and less dangerous ones as file sharing and online forums. To get the access to dark web, particular browsers are required, such as TOR [33]. It is difficult to monitor and regulate the operation on the dark web due to its ominous qualities. Because of this, it is challenging to find criminals who engage in illicit activity on it. In [33] it is explained that after the access to dark web has been gained, it has been possible to search the price for personal information such as name, birthday, address, and other personal information. The cost for a single record is around 1.25 dollars [33]. Stolen data can be used for the most disparate activities, for example getting illegal access to medical devices. Taking as example the data breach of CHS, the value of the stolen data on the dark web is around 2.7 million dollars [33].

3. Cyber risk assessment

Organizations that operate digitally, processing data, have the necessity to be provided with proper cyber security algorithm to carry out their tasks as safely as possible. The degree of cyber security procedures that an organization should adopt depends on different factors, such as the sensitivity of the processed data. Consequently, organizations are interested on assessing their cyber security posture, to check whether the processed data are protected against cyber attacks or not.

The procedure to be performed to evaluate the level of security of a particular organization is named cyber risk assessment, which aims at identifying the possible vulnerabilities in confidentiality, integrity and availability of an organization's data and systems. Cyber risk assessment seeks to identify potential threats and evaluate their impact on the organization. This process combines technical and business analysis to assist organizations in understanding the threat landscape and determining the most effective way to protect their data. Therefore, the demand for cyber risk assessment tool is increasing, especially for quantitative probabilistic model [34].

A unique definition for "risk" does not exist, but according to the National Institute of Standard Technology (NIST) it can be defined as how much an organization is threatened of receiving a cyber attack, taking into account the possibility of success related to the cyber attacks and the amount of damage (economical or structural) that it can cause [34].

A unique tool or methodology for cyber risk assessment does not exist; consequently, different probabilistic models based on quantitative or qualitative approaches have been developed during the past years. Despite the variety of methodology to perform cyber risk assessment, usually all models are based on the following steps [34]:

- Risk identification
- Risk analysis
- Magnitude estimation

The risk identification phase is a procedure of primary importance in the development of tools for cyber risk assessment and it represent a critical step in any organization's cybersecurity strategy. It entails analysing the current environment and identifying any potential security risk or vulnerability. Potential dangers that might result in a cyber attack should be identified and prioritized after an evaluation of the present security posture of an

organization. There are both internal and external risks, such as lax security procedures or careless employees, which include harmful software, hackers, and malware.
Risk analysis involves assessing the likelihood of occurrence and potential impact based on an organization's security posture. Each threat should be evaluated based on its likelihood of occurrence and potential for harm.
Finally, magnitude estimation, in terms of economical or structural damage, is estimated.
As previously mentioned, a cyber risk assessment tool can be based on a quantitative or a qualitative approach, in function of the way in which the model is constructed. As stated in [34], in a qualitative approach, a nonnumerical method is exploited for the development of the model, making the development procedure easy and fast. The drawback of these approaches is that the results are not reproducible since they are based on a certain level of subjectivity. On the other hand, quantitative approaches are based on numerical methods that follows the probability theory. Despite the development of these models is complex, the main advantages are the reproducibility of the results and the low degree of subjectivity [36].
As reported in [34], one example of a cyber risk assessment model based on a quantitative approach is MAGIC (Method for AssessinG cyber Incidents oCcurrence), which allows the assessment of the likelihood of occurrence of cyber attacks by exploiting numerical solutions. For instance, other two probabilistic models based on a quantitative approach are FAIR (Factor Analysis of Information Risk) [35] and HTMA (How To Measure Anything in cyber security risk) [36], despite they rely on a certain level of subjectivity since for the evaluation of some parameters the judgement of experts is taken into account.

3.1 Cyber risk assessment models

Cyber risk assessment models are methods for identifying, assessing and evaluating potential risks and vulnerabilities of an organization's information technology systems and assets. To determine the overall level of risk, these models typically analyse various factors such as the organization's assets, threats, vulnerabilities, and impacts [34]. Threat modelling, vulnerability assessment, and risk assessment are examples of common cyber risk assessment models. These models can be used to take risk management and security control decisions, as well as to assess the effectiveness of existing security measures.

3.1.1 MAGIC

As mentioned, MAGIC is a probabilistic model based on a quantitative approach to determine the likelihood of occurrence of a cyber incident, based on the cyber posture of the target organization [34].
In this paragraph, the MAGIC's components are briefly defined and described [34]:

- Cyber Threat

A cyber threat can be defined as an event which can negatively impact the ability of an organization to perform its normal tasks.

- Cyber attack

A cyber attack is an intentional action carried out by a hacker or a group of hackers aimed to steal data, create disorder, or paralyze a company's information system.

- Awareness of the employees

It can be defined as the knowledge that employees have about the cyber security risk, which can be increased through training programs.

- Maturity of the organization

It can be defined as the level at which cyber security practices are implemented in order to reduce the probability of receiving a cyber attack.

- Complexity of the organization

It can be defined as the intricacy of the technological infrastructure of an organization and how the processes are managed.

- Attractiveness of the organization

It can be defined as the level of interest that a possible attacker could have in performing an attack to a certain organization, to obtain profit. It is related to the sensitivity of the data that the organization processes.

- Maturity of the attacker

For the cyber risk assessment topic, it is important to consider that cyber attacks are carried out by capable hackers, in order to implement optimal algorithms.

- Probability of success of an attack

It is related to the probability that an attack breaches the organization, since not all attacks manage to do it.

- Number of attacks

It is the number of attacks that a hacker may attempt to perform in a specific time period.

- Likelihood of occurrence of a cyber incident

It is the probability that an organization is targeted by a certain number of attacks.

The relations among the components of MAGIC are depicted in Figure 1 [34].

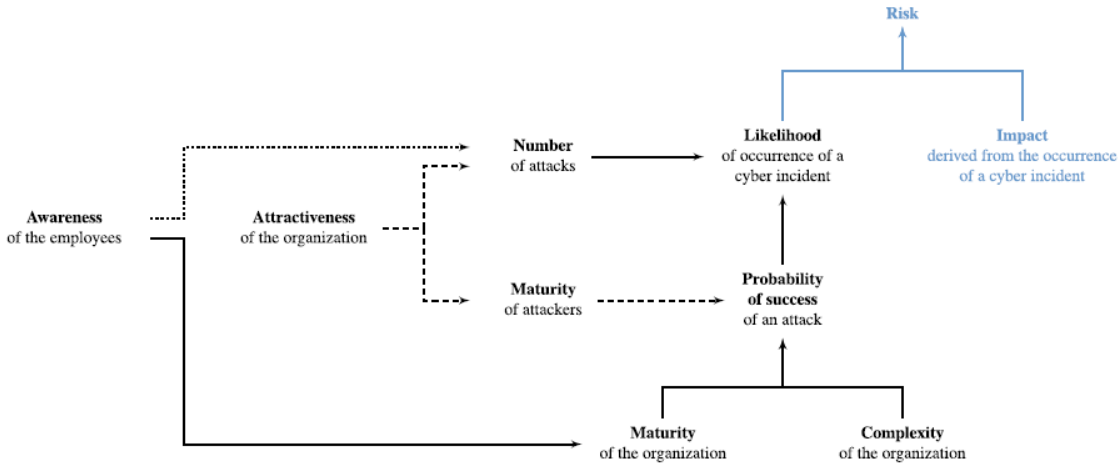

Figure 1: Relation among MAGIC component [34]

In order to compute the likelihood of occurrence of a cyber incident, two different approaches have been taken into consideration. The first approach is based on the assumption that after receiving one or more successful cyber attacks the company does not change its posture. Going into further details, it means that all the cyber security procedures remain the same without any upgrade or change.

The second approach is based on the expectation that the company will change posture after being successfully targeted by a cyber attack. Going deeper, this means that the organization upgrades or changes the cyber security procedures.

3.1.2 FAIR and HTMA

FAIR and HTMA are two probabilistic models for cyber risk assessment.

As explained in [35], the FAIR methodology can be described through the following steps:

- Definition of the situation under examination and decomposition in sub-scenarios.
- Evaluation of the parameters for each sub-scenario.
- Monte Carlo approach
- Results interpretation.

The FAIR approach explains how risk assessment can be obtained by defining the ontology of the risk. Risk elements are measurable and evaluable. As explained in [34], in this approach the risk factor is computed as a combination of Loss Event Frequency (LEF) and Loss Magnitude. The Loss Event Frequency can be defined as how often an event related to an economical loss happens while loss magnitude is related to the entity of economic losses. As it can be seen in Figure 2, both Loss Event Frequency and Loss Magnitude can be decomposed into other factors.

Loss Event Frequency can primarily be related to the Threat Event Frequency (TEF) which is how often a malicious attacker tries to breach the organization in a specific time period and to the vulnerability which is the probability that an attack breaches the organization. For what concerns Loss Magnitude, it depends on the sum of the impacts caused by the primary risk plus those generated by the secondary risk.

This procedure allows the estimation of risk starting by the known factor, considering a particular level of the FAIR ontology.

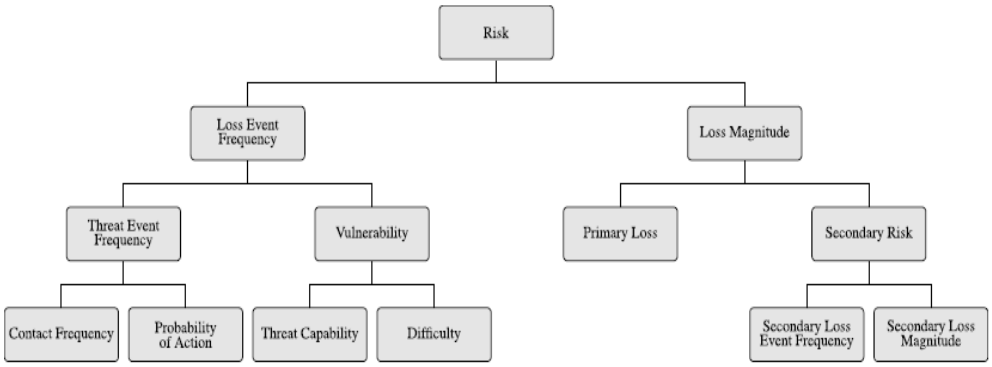

Figure 2: Ontology of FAIR model [37]

The steps of HTMA are as follows, as also described in [36]:

- Definition of potential cyber risks.
- Evaluation of the likelihood of occurrence of each cyber risk individuated in the previous step.
- Monte Carlo simulation to develop a scenario.
- Results interpretation.

The results obtained by the Monte Carlo evaluation are used to obtain the Loss Exceedance Curve (LEC) which is the plot of the complementary cumulative distribution function of the annualized loss expectancy [36].

3.1.3 MAGIC implementation with HTMA and FAIR

The drawback of HTMA and FAIR method is that they rely on a certain degree of subjectivity. For what concerns HTMA, the likelihood of occurrence of a cyber attack is determined by an expert. For what regards FAIR, the expert judgement is required to determine the parameters of the LEF.

In this framework, MAGIC can be employed to derive the inputs to drive HTMA and FAIR reducing the subjectivity.

The MAGIC approach can be utilized, together with HTMA, to quantitatively derive the likelihood of occurrence of a cyber-attack. For each cyber-attack, a certain value of maturity is calculated, and it is used together with the complexity and the attractiveness of the organization to calculate the likelihood, as shown in Figure 3.

ID	Threat	Maturity index	p^*	p_m	p_M	L^C
1	Malware	4.3	0.50	0.28	0.72	0.86
2	Web-based attacks	5.6	0.23	0.11	0.43	0.61
3	Denial of services	3.6	0.66	0.43	0.83	0.92
4	Malicious insiders	1.9	0.90	0.79	0.95	0.97
5	Phishing and social engineering	3.6	0.66	0.43	0.83	0.92
6	Malicious code	6.0	0.17	0.08	0.34	0.52
7	Stolen devices	4.8	0.38	0.19	0.62	0.78
8	Ransomware	5.1	0.32	0.16	0.55	0.72
9	Botnets	4.3	0.50	0.28	0.72	0.86

Figure 3: Table listing the parameters to estimate likelihood [36]

MAGIC can also be employed to quantitatively calculate the parameters of the LEF. Using MAGIC, it is possible to quantitatively compute the parameters of the LEF taking into account also TEF and Vulnerability, reducing the subjectivity of the process.

4. Materials

In this section the materials that have been used to perform this work have been listed.

4.1 Dataset

The dataset that has been used in this work is taken from [37]. It has been imported in Excel format and it contains a list of attacks that companies have experienced in a period of time ranging from 2005 to 2018. It also contains information about the sector in which the companies operate and specifics about the cyber attacks they suffered.

4.2 UpGuard

UpGuard [38] is a platform specialized in third-parties risk assessment and management. UpGuard monitors thousands of companies and billions of data every day, using them to prevent third parties' data breach. For the purposes of this work, UpGuard has been utilized to collect security scores and the number of employees of each company. The cyber security assessment is performed through third parties security ratings, questionnaires about security and threat intelligence capability.

4.3 Probabilistic Models

This work has been carried out by using a probabilistic model, (called MAGIC, indeed) that has been used to calculate the likelihood of occurrence of a cyber incident, based on the evaluation of the cyber posture of an organization [34]. MAGIC allows to derive the inputs that can drive two different probabilistic assessment methods, FAIR [35] and HTMA [36].

4.4 MATLAB

The machine learning model has been developed on the MATLAB environment (*MATLAB 2022b, The MathWorks, Inc., Natick, Massachusetts, United States.)*. In particular, the Classification Learner app has been used to develop the model. Using this app, it has been possible to explore data of the dataset, select features, specify validation schemes, train models, and assess results.

5. Methods

In this work, machine learning models have been developed to assess the estimation of the cyber risk related to a particular organization. The features that have been chosen to power up and train the ML model are maturity, complexity, and attractiveness. The maturity, complexity, and attractiveness scores have been derived from UpGuard since, for practical use, they could not be obtained directly from the companies. In addition, the MAGIC method has been driven by the previously described parameters, as well as HTMA or FAIR, to create a fourth feature, “Likelihood” or “Expected Attacks”.

5.1 Parameters

Before delving into the steps taken to develop the models, it is necessary to discuss the reasoning that led to the decision on which parameter to consider.

The choice has been based on the information provided by [34]:

- Security score (Maturity)
- Number of employees (Complexity)
- Attractiveness

The security score parameter has been chosen to represent the maturity of an organization because it is a data-driven, objective, and dynamic measurement of an organization's security posture [38].

Complexity is related to the organization’s dimension, since, considering the same condition of maturity and attractiveness, smaller organizations are less likely to be targeted by attacks [34]. In this work, the number of employees has been considered as a mirror for the complexity of an organization.

The attractiveness of an organization is mainly related to the sensitivity and the importance of the data that it is processing. The more sensitive data are, the more profit an attacker can accomplish stealing them or depriving the organization of being able to access them. Attractiveness is not correlated to the complexity, since also small organizations with a small number of employees can process sensitive data.

These parameters have been considered as strictly related to the number of cyber attacks that can be perpetuated against a certain company.
The time period in which a company has been targeted by a certain number of attacks has not been taken into consideration as a parameter as its insertion would have had the effect of polarizing the machine learning model since, given the conformation of the dataset, it has a too strict correlation with the number of attacks.

5.2 Construction of the dataset

The first step in the development of a machine learning model capable of classifying organizations into classes of risk has been the research for a suitable dataset [37].

The dataset that has been considered in this work contains the name, the sector in which a specific organization operates, and a list of attacks that have targeted a specific company during the considered time period. Information about the specific category of cyber attack perpetrated was not considered for the purposes of this work. UpGuard has provided new information for the development of a slightly modified dataset. The sector in which each company operates, the number of cyber attacks received, the security score, and the number of employees have all been considered and added to the new dataset for each company listed in UpGuard and simultaneously listed in the original database. Because both UpGuard and the original database contain sector information, it has been decided to utilize only the information about the sectors listed in UpGuard. The reason behind choice is the greater specificity that the UpGuard site reports in terms of sectors. As it can be seen in Table 1, the used parameters have been called respectively security score, number of employees and sector. In order to be called maturity, complexity and attractiveness a parameterization of these parameters needs to be carried out, to make them suitable to be processed by MAGIC.

Table 1: Dataset excerpt

ID	Security Score	Number of employees	Sector	Attacks
Facebook	874	25000	IT and Telecommunication	12
Starbucks	855	238000	Hotels and Hospitality	8
CVS Health	817	158000	Healthcare	10

5.3 Parameterization

The next step has been the parameterization of sector, security score, and number of employees to derive attractiveness, maturity, and complexity.

Inspired by the information reported in [34], attractiveness has been categorized in five levels.

The classification of the sectors in which companies operates, is a function of the percentage of the attacks (π) received by a specific sector or group of similar sectors in terms of the total amount of attacks. Consequently, sectors that have received a higher percentage of attacks inside the dataset have been classified as "very highly attractive (5)", while sectors who had received a lower percentage of attacks as "highly attractive (4)", "averagely attractive (3)", "lowly attractive (2)" and "very lowly attractive (1)"; furthermore, each of these levels has been associated with a numerical value, as shown in Table 2. The classification of sectors into attractiveness levels has been shown in Table 3. For instance, companies that operate in the financial field have been classified as level 5 for what concerns attractiveness, since they have been targeted by the 27% of attacks present in the dataset, while healthcare services have been classified as 3, since healthcare organizations were targeted by the 10% of attacks.

Table 2: Classification of the organization based on their attractiveness level

Percentage of attacks	**Attractiveness Levels**	**Numerical Value**
$\boldsymbol{\pi}$ <2.50%	Very Lowly attractive (1)	0.6
2.50≤ $\boldsymbol{\pi}$<5%	Lowly Attractive (2)	0.7
5%≤$\boldsymbol{\pi}$<10%	Averagely attractive (3)	0.8
10%≤$\boldsymbol{\pi}$<20%	Highly attractive (4)	0.9
$\boldsymbol{\pi}$>20%	Very highly attractive (5)	1

Table 3: Classification of the companies based on the number of attacks received

Sector	#	Number of attacks		Attractiveness Level
IT and Telecommunications	1	360	20%	4
Retail	2	197	11%	4
Hotels and Hospitality	3	131	7%	3
Manufacturing	4	119	7%	3
Finance and insurance	5	500	27%	5
Professional Services	6	114	6%	3
Tech	7	4	0%	1
Utilities	8	21	1%	1
Healthcare	9	174	10%	3
Education	10	32	2%	1
Government	11	15	1%	1
Media	12	34	2%	1
Transportation and Warehousing	13	54	3%	2
Dating	14	3	0%	1
Mining and Oil & Gas	15	53	3%	2
Real Estate	16	5	0%	1
Agriculture, Forestry, Fishing and Hunting	17	4	0%	1
Construction	18	3	0%	1

Subsequently, the security score and the number of employees has been parameterized in order to derive maturity and complexity and make them suitable for driving the MAGIC algorithm. For what concerns the security score, different parameterizations have been tested, all based on the categorization of this parameter into 11 classes, from 0 to 10. The

more relevant parameterizations tested have been based on two different approaches: the first on a linear classification and the second on the creation of classes with roughly the same number of elements (companies). The first parameterization has revealed a significant disparity in the distribution of maturity value across all companies in the dataset. The reason for this is that only a few companies had security score values lower than 500, and almost no companies had values lower than 400. As a result, the classification based on a linear approach have been ruled out.

In Table 4 and 5 the thresholds that have been used to discriminate between classes and the minimum, maximum, mean, and median value of the security score are shown.

As it can be seen in Table 5 the mean value for maturity is 777, leading to the impossibility to use a linear classification to get an even distribution of the maturity the dataset.

Table 4: Security score classes

Thresholds	**Maturity Classes**
0	0
650	1
680	2
710	3
740	4
770	5
800	6
830	7
860	8
890	9
920	10

Table 5: Security score distribution

Security score distribution			
Min	Max	Mean	Median
152	950	777	789

In regard of the last parameter, the number of employees, three parameterizations have been employed to create classes among the companies of the dataset: a linear one, another linear but with the upper limit of the class $i + 1$ being the double of the one of the class i, and last one based on having the same amount of elements for each class, similarly to what has been done for security score.

In this case the second approach has been used because it allowed a better distribution of the complexity parameter among the datasets.

In Table 6 the thresholds used in the classification are shown, while in Table 7 the minimum, maximum, mean and median value of the number of employees have been represented.

Table 6: Complexity classes

Thresholds	**Complexity Classes**
1	0
1000	1
2000	2
4000	3
8000	4
16000	5
32000	6
64000	7
128000	8
256000	9
512000	10

Table 7: Complexity distribution

Number of employees distribution			
Min	Max	Mean	Median
5	2300000	56669	18960

After the parameterizations of sector, security score, and complexity into attractiveness, maturity and complexity have been carried out, the next step concerned the definition of the relation that links these three parameters to the number of attacks that a company can receive.

A high maturity score implies that all components of a company are applying cyber security policies effectively, lowering the likelihood of a successful cyber attack. A high value of the complexity reflects an intricate connection between the various technological infrastructures of an organization, leading to a complication of how the services and activities are managed. Therefore, a high complexity increases the possibility to receive successful attacks, since a company with a high complexity is most likely to have some trouble with the application of the cyber security procedure in all its components.

For what concern attractiveness, as it has been mentioned in the previous pages, it reflects the sensitivity of the data processed by a certain company: the more sensitive the data are, the more profit a hacker will make from getting them. Accordingly, a high value of this parameter implies that a company is more likely to be targeted by cyber attacks.

5.4 Dataset analysis

In this section the analysis of the dataset is shown. Figures 4, 5, 6 report the distribution of maturity, attractiveness, and complexity in relation to the number of attacks respectively, in order to show the correlation that links these parameters to the cyber risk.

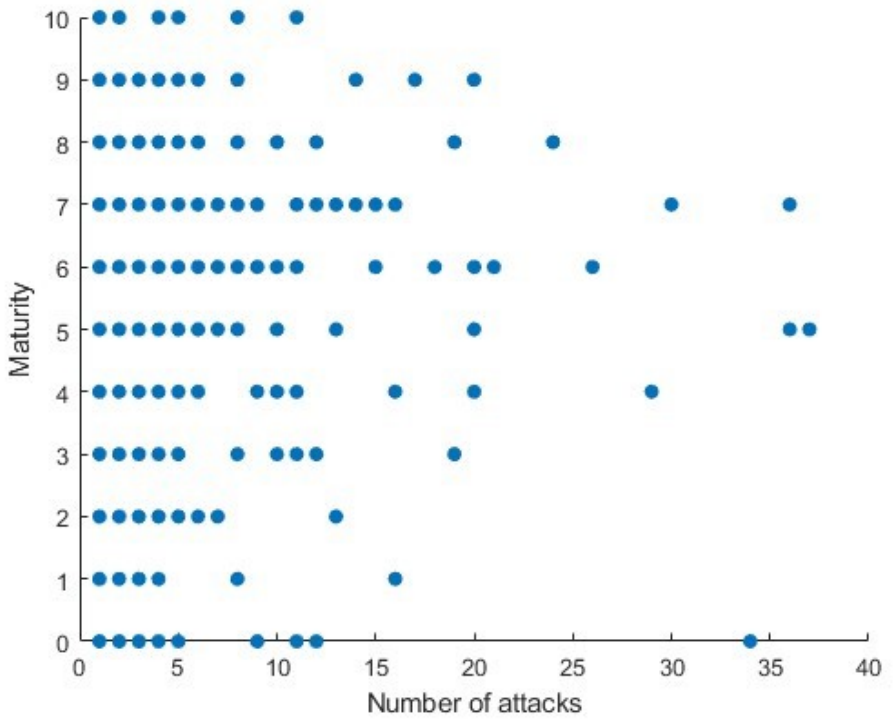

Figure 4: Maturity distribution

Figure 4 shows on one hand that for maturity score ≥ 5, an increase in maturity leads to a drop in the number of successfully perpetuated attacks, but on the other hand show one weakness of the database, since for low value of maturity the number of successfully perpetuated attacks is very low.

Figures 5 and 6 instead show a similar trend. In particular, for an increase of attractiveness or complexity level, we have an increase in the number of attacks that have been received. These plots confirm the hypothesis on which this work has been based on, i.e., the relation between maturity, complexity, and attractiveness with cyber risk.

An analysis on the distribution of the attacks among all the companies of the dataset is shown in Figure 7.

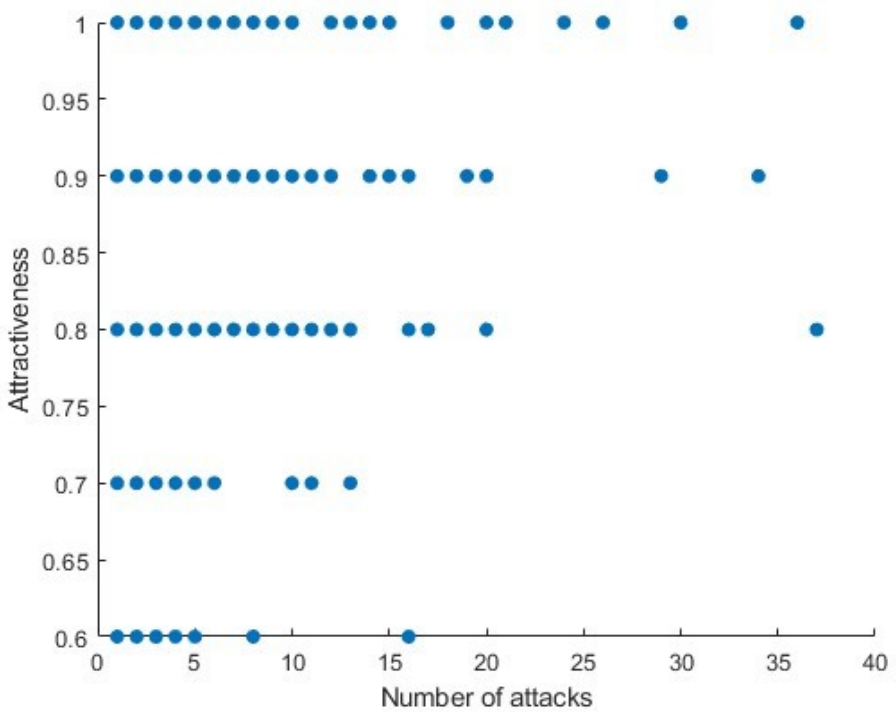

Figure 5: Attractiveness distribution

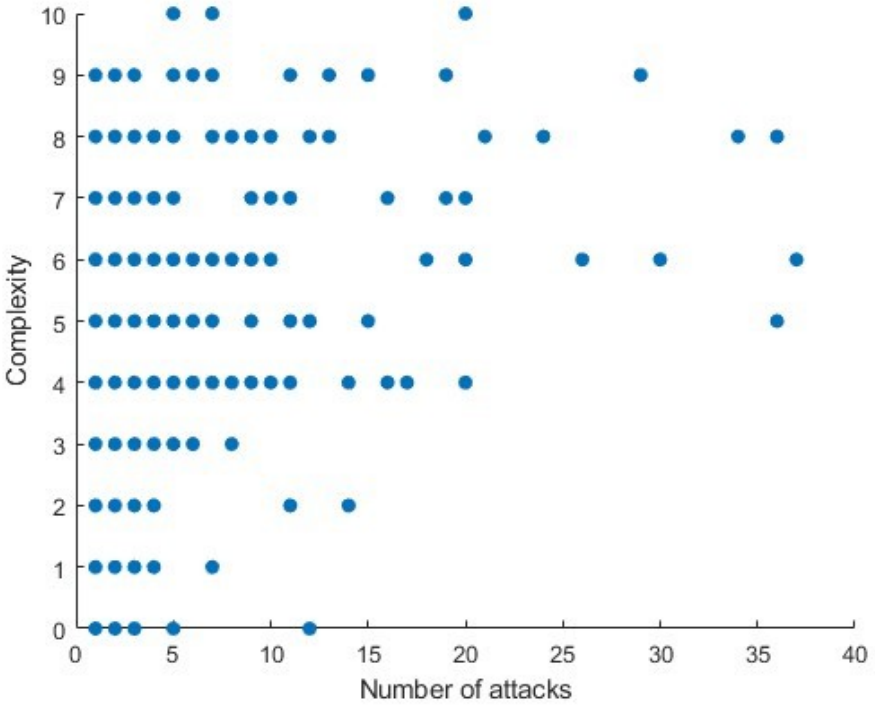

Figure 6: Complexity distribution

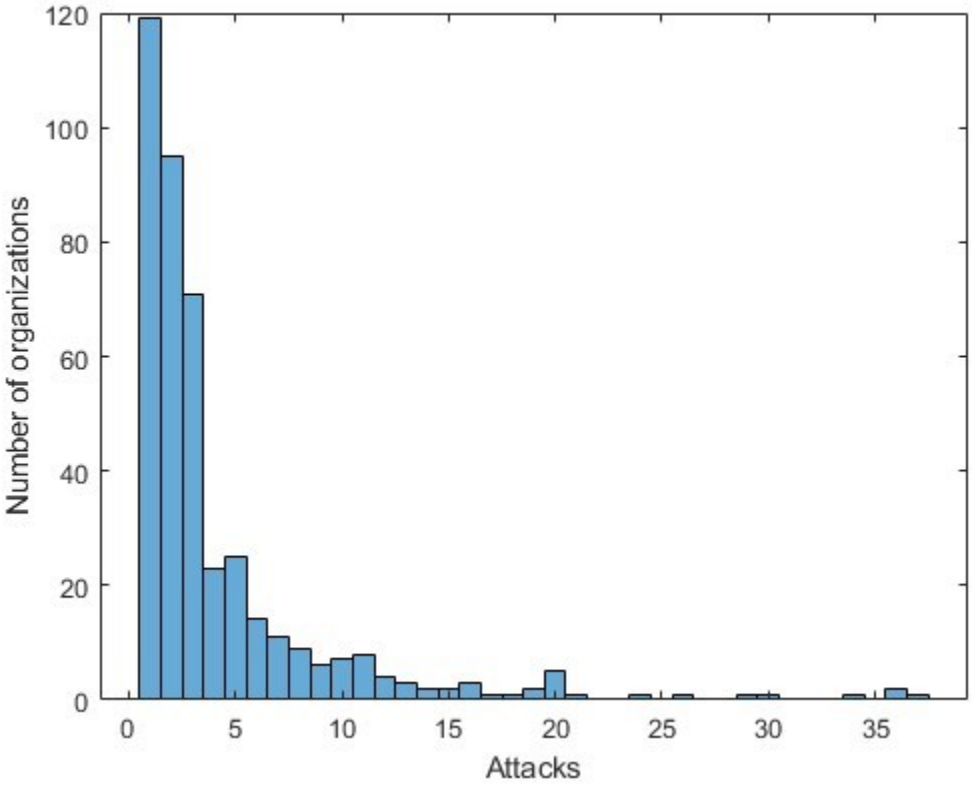

Figure 7: Histogram representing on the y-axis the number of companies and on the x-axis the number of attacks.

As shown in the histogram in Figure 7, the distribution of attacks among companies is completely biased towards low number of attacks, for instance from 1 up to 3. This shape of the histogram shows an important limitation of the database, which represents an obstacle to the correct training of a machine learning model.

As a result, a classification of the number of attacks has also been carried out. The reason is the difficulty that a machine learning model may face when dealing with a classification problem in which the dimension of the predictor variable ranges between 1 and 37 attacks and attempting to create classes with roughly the same number of elements in order to reduce database unbalancing.

The classification has been based on the histogram reported in Figure 7. The similarities that organizations with a similar amounts of received attacks may have in terms of maturity, complexity, and attractiveness have also been taken into account when performing the classification into classes. It is indeed likely , that a company that in 13 years has received 1 attack shares similar parameters to those companies that in the same span of time have received 2 or 3 attacks. Thus, categorizing in two different classes organizations that have received 1, 2 or 3 attacks will lead to confusion in the training of the machine learning model. Consequently, the number of attacks has been classified as it follows:

- Class 1: $x \leq 3$
- Class 2: $x > 3$,

with x representing the number of attacks.

These two classes have been named “Classes of risk”, low risk and high risk respectively, as they represent the risk to receive a certain amount of cyber attack, and they have been added for each company listed in the database. As it is possible to witness looking at Figure 8, due to the impossibility to insert companies that have received 1, 2 or 3 attacks into different classes, this procedure was not been able to solve the unbalance of the dataset, while it solved the problematic related to the range of the response variable, reducing it from 37 to 2.

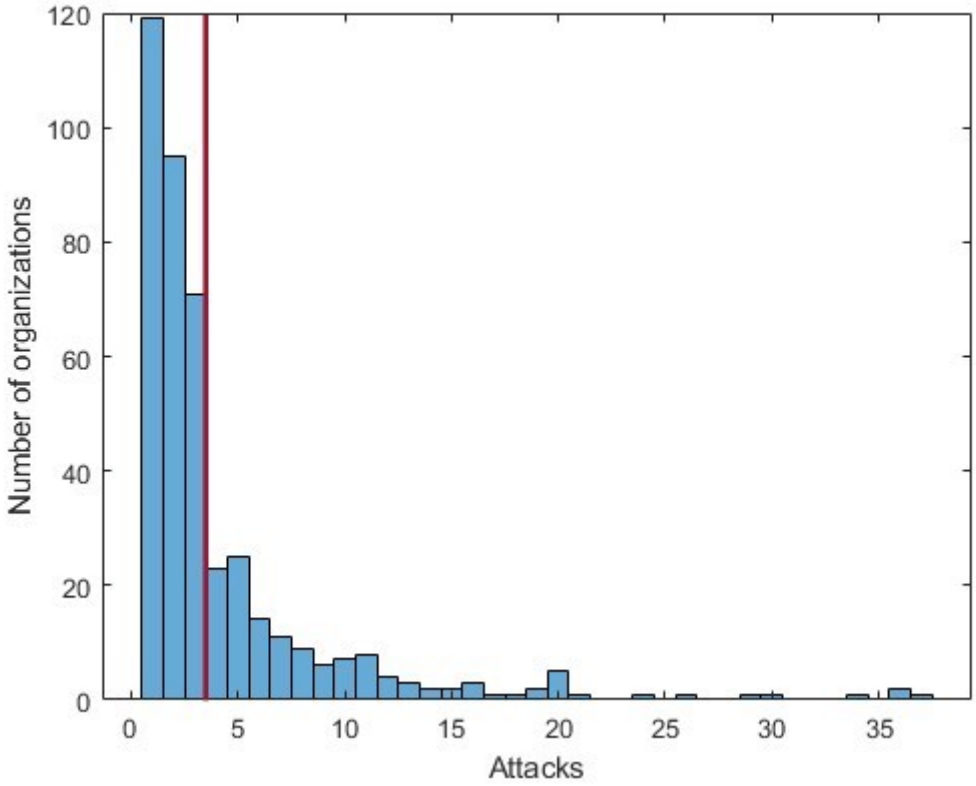

Figure 8: Division of the database into two classes

5.5 Oversampling

To deal with the unbalance of the dataset, that would result in a poor training of the machine learning models, an oversampling procedure has been employed. The oversampling procedure solves the unbalance of the dataset by generating elements of the minority class until they become equal in number to the elements of the majority class. There are several methods available to oversample a dataset used in typical classification problem, in this case SMOTE [39] (Synthetic Minority Over-sampling Technique) has been utilized. This function synthesizes new observations based on existing (input) data, and a k-nearest neighbour approach. The SMOTE algorithm can only be applied to data that is intended to be used as a training set by the machine learning models. As a result, the standard dataset of 420 companies has been divided into a training set (80%) and a test set (20%). The training set, the number of elements to synthesize, and the number of nearest elements to consider are the inputs to this MATLAB function. The output is the oversampled training set, which

includes the starting one as well as the newly synthesized elements, as well as a vector containing the newly synthesized elements.

SMOTE has been used in our framework to generate fictitious minority-class companies. These new elements are made up of the same variables parameterized in the same way as previously described, as shown in Table 8.

Therefore, with the risk classes being the two mentioned in the previous pages, SMOTE algorithm synthesized 147 companies belonging to class 2, that is high risk class. The new dataset, composed of 484 companies for the training set and 84 companies for the test set, has been utilized for the development of machine learning models.

Table 8: Example of new synthetized organizations

ID	**Maturity**	**Complexity**	**Attractiveness**	**Attacks**	**Class of Risk**
373	4	5	0.7	5	2
384	3	3	0.8	5	2
435	8	6	1	8	2
570	7	7	0.8	14	2

As can be noted in Table 8, the new synthetized companies have been denoted with an ID and all of them belongs to the same class of risk. Furthermore, all the new synthetized organizations share a high level of attractiveness together with a number of attacks greater than 3. A histogram of the new dataset is represented in Figure 9.

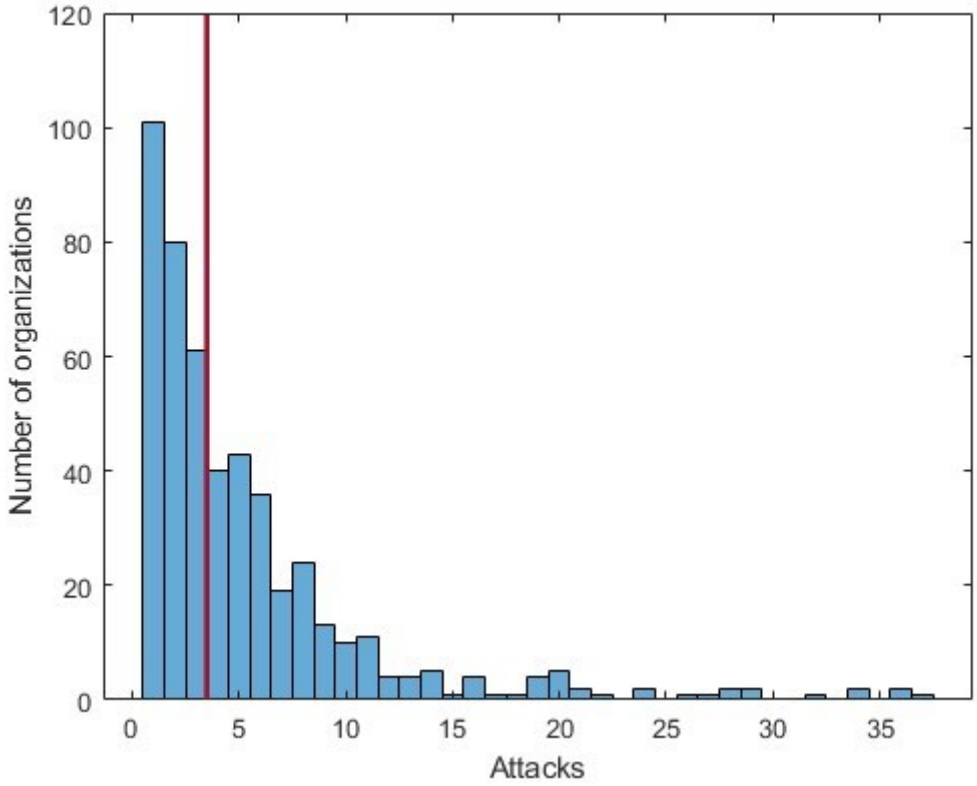

Figure 9: Histogram representing on the y-axis the number of companies and on the x-axis the number of attacks of the new dataset

As shown in Figure 9, the new dataset has a greater number of organizations that have been targeted by 4 and 5 attacks in particular, allowing the organization to be placed into two different risk categories with the same amounts of elements.

5.6 Quantitative cyber risk assessment

In the contest of this work, the MAGIC procedure has been utilized to derive inputs to drive HTMA or FAIR. In order to do so, MAGIC has been fed with maturity, complexity, attractiveness, and another parameter, “n_avg”, i.e., the number of attacks attempts that an organization is expected to receive in the considered period of time.

This parameter has been implemented in three different configurations:

- “n_avg mean” that represents the average amount of attacks among all the companies.

- "n_avg attractiveness" that has been defined based on the attractiveness of a certain company.
- "n_avg vec" which has been also based on the attractiveness, but with the addition of a random component.

In the second configuration, different possibilities for the associations have been explored. Once the parameters have been tuned and made suitable to drive the MAGIC algorithm, both the approaches of MAGIC with HTMA and MAGIC with FAIR have been employed.

For the purpose of this work, the HTMA approach has been implemented, together with MAGIC, to derive the probability of receiving one attack (Likelihood). Therefore, to avoid having too many possible values of "Likelihood" (%), which take values in an infinite set, this parameter has been categorized in 10 classes, as shown in Table 9.

For what concerns the implementation of MAGIC with FAIR, it has been utilized to derive the most likely number of attacks that could target a certain organization in a specified time period. This new parameter, which has been called "Expected Attacks," has been used without any categorization.

Table 9: Classes of likelihood

Likelihood	**Class**
Ω≤10%	1
10%< **Ω** ≤20%	2
20%< **Ω** ≤30%	3
30%< **Ω** ≤40%	4
40%< **Ω** ≤50%	5
50%< **Ω** ≤60%	6
60%< **Ω** ≤70%	7
70%< **Ω** ≤80%	8
80%< **Ω** ≤90%	9
Ω >90%	10

5.7 Development of the models

The models development has been carried out on MATLAB environment, using the app "*classification learner*". MATLAB's classification learner allows to import variables from an Excel sheet and to interpret the columns as predictors or response. For the purposes of this work, maturity, complexity, attractiveness, and one of the parameters given as output by FAIR or HTMA have been used as predictors, and the class of risk as responses. The training set has been used to train the different machine learning models and the test part to check the accuracy of the trained models. The classification learner also allows to train all the possible machine learning models simultaneously. This characteristic has been utilized to train each time all the available models utilizing the predictors given as input. For what concerns the training phase, different approaches have been attempted to train the machine learning algorithm, as it has been listed in Table 10. MAGIC has been combined with HTMA to derive three different values of "Likelihood", one for each "n_avg" implementation. The same can be said for the implementation of MAGIC with FAIR, for which three different "Expected Attacks" have been calculated.

Therefore, six different Excel files have been developed and used to drive the machine learning algorithms, however, in all the cases, each Excel file has been characterized by the same maturity, complexity, attractiveness, and class of risk.

Table 10: Representation of the 6 different approaches: 3 with FAIR and 3 with HTMA

HTMA approach	**FAIR approach**
N_avg mean	N_avg mean
N_avg attractiveness	N_avg attractiveness
N_avg attractiveness + random component	N_avg attractiveness + random component

Cross fold validation has been used in the training of the models. Cross-validation is a resampling procedure used in statistics to determine how well the results of a statistical investigation would generalize to an independent data set. The purpose of cross-validation is to provide a dataset for evaluating the model during the training phase, to limit problems

such as overfitting, and to provide insight into how the model will generalize to an independent dataset.

6. Results

The best results obtained with both the standard and the SMOTE dataset are reported in this section. All the models reported in the following pages have been obtained with FAIR approaches, since in this framework it has proven to work slightly better than HTMA.

6.1 Standard Dataset

In this section the result obtained with the old database containing 420 organization are shown. The best model, in terms of pure accuracy, in the classification of organizations into risk classes obtained with the standard databased is the Fine Gaussian SVM, whose characteristics are reported in the Table 11 and shown in Figures 10 and 11. To obtain "Expected Attacks", a value of Navg equal to 10 has been utilized.

Table 11: Model characteristics

Training Results	
Accuracy	72.3%
Total Cost	93
Prediction Speed	3400 obs/sec
Training Time	10.848 sec
Test Results	
Accuracy	71.4%
Total Cost	24

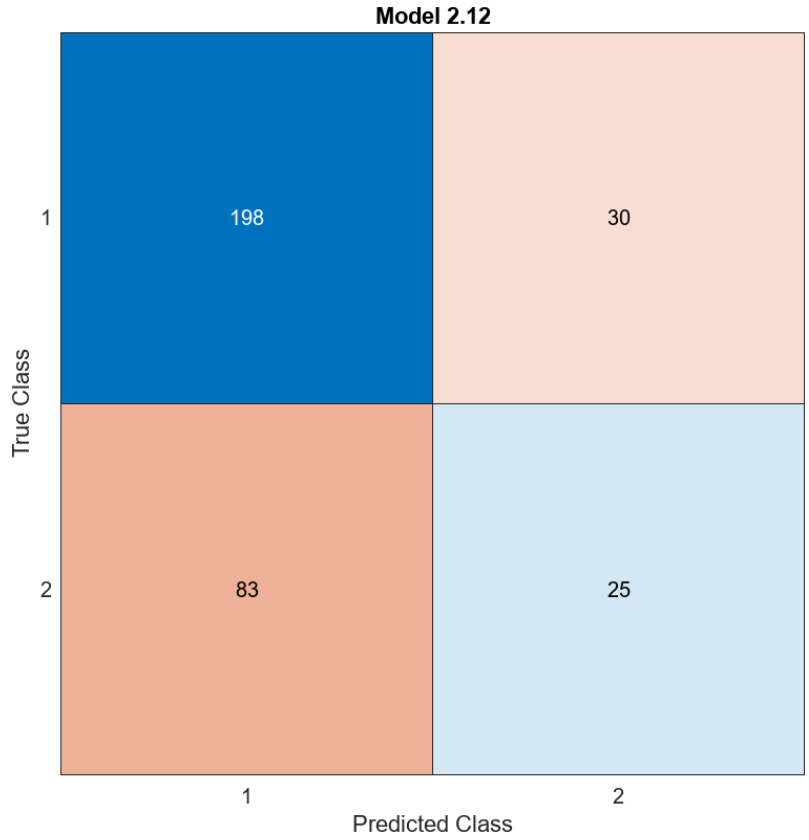

Figure 10: Train confusion matrix

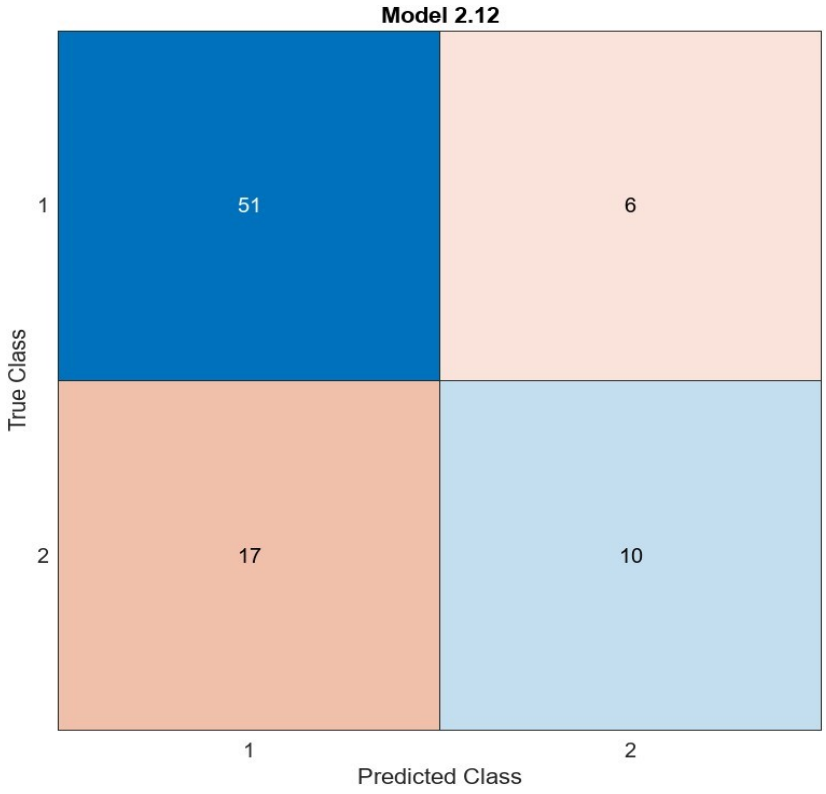

Figure 11: Test confusion matrix

As it can be seen in the previous figures, the result obtained with the standard dataset are polarized towards class 1 in both training and test. This is the result of the unbalance of the dataset, which have led into a poor training and consequently a poor test. The accuracy value (71.4%) is not reliable, since the model classifies almost every entry as class 1.

In order to try to get a better training and test phase, another attempt has been made by increasing the value of Navg to 40. The best model resulted to be the KNN, whose characteristics have been listed in Table 12 and shown in Figures 12 and 13.

Table 12: Model characteristics

Training Results	
Accuracy	66.5%
Total Cost	118
Prediction Speed	5500 obs/sec
Training Time	1.0409 sec
Test Results	
Accuracy	77.4%
Total Cost	19

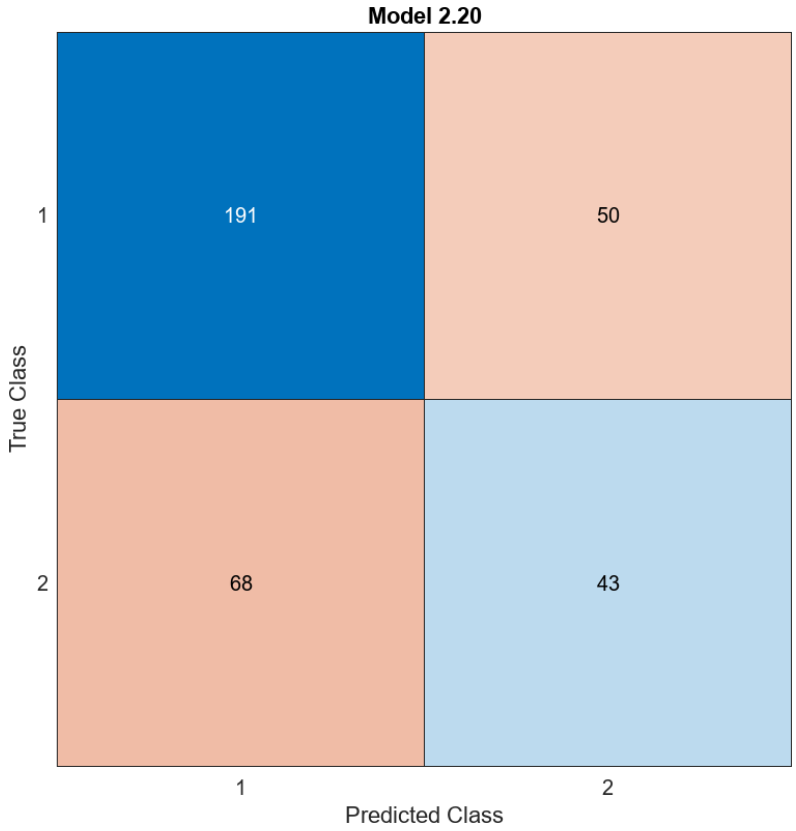

Figure 12: Train confusion matrix

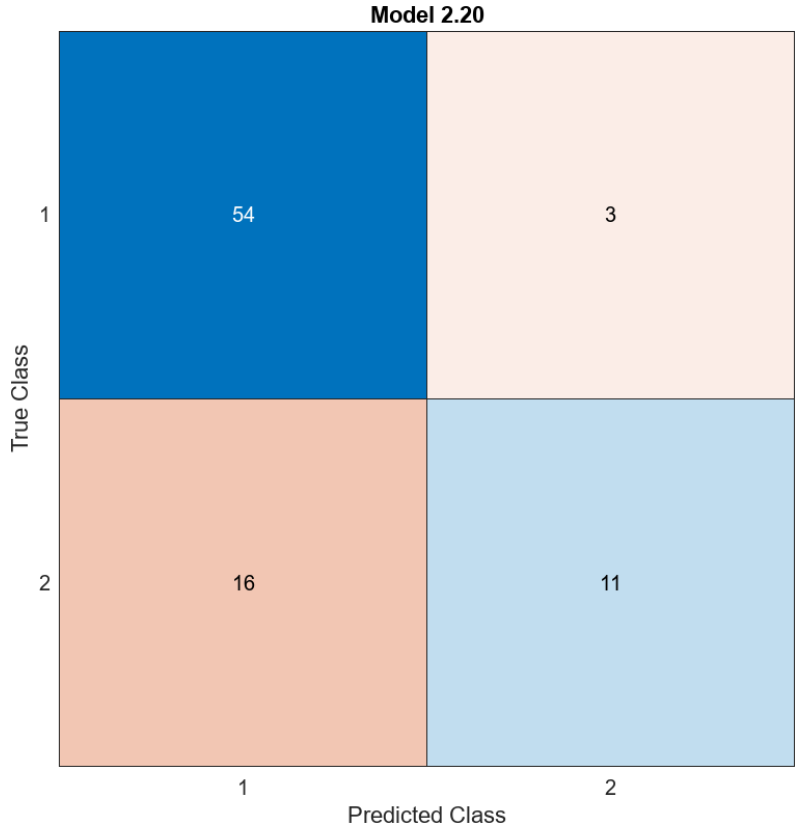

Figure 13: Test confusion Matrix

As it can be seen looking at Figure 12 and 13, also in this case both training and test are completely polarized towards class 1, underlying the impossibility to have a reliable result employing the standard dataset.

6.2 SMOTE dataset

This section reports the findings obtained using the new dataset acquired using the SMOTE procedure. The first model has been obtained employing a Medium Tree approach. The "Expected Attacks" parameter has been derived from the FAIR approach using a value of Navg for each level of attractiveness. In particular, starting from 5, an increase in the attractivity level yields a $x2$ in the value of the Navg. In Table 13 the value of Navg associated to each level of attractiveness has been represented. The characteristics of this model have been reported in Table 14 and in Figures 14 and 15.

Table 13: Navg correlation to attractiveness

Attractiveness	**Navg**
0.6	5
0.7	10
0.8	20
0.9	40
1	80

Table 14: Medium Tree characteristics

Training Results	
Accuracy	64.7%
Total cost	171
Prediction Speed	3900 obs/sec
Training speed	2.0981 sec
Test Results	
Accuracy	71.4%
Total cost	24

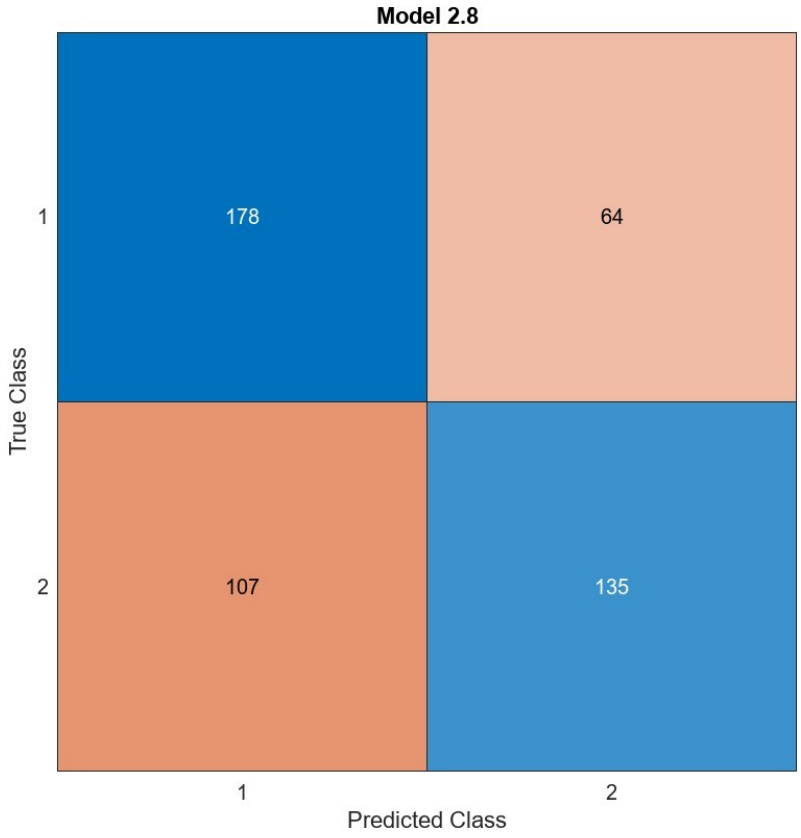

Figure 14: Train confusion matrix

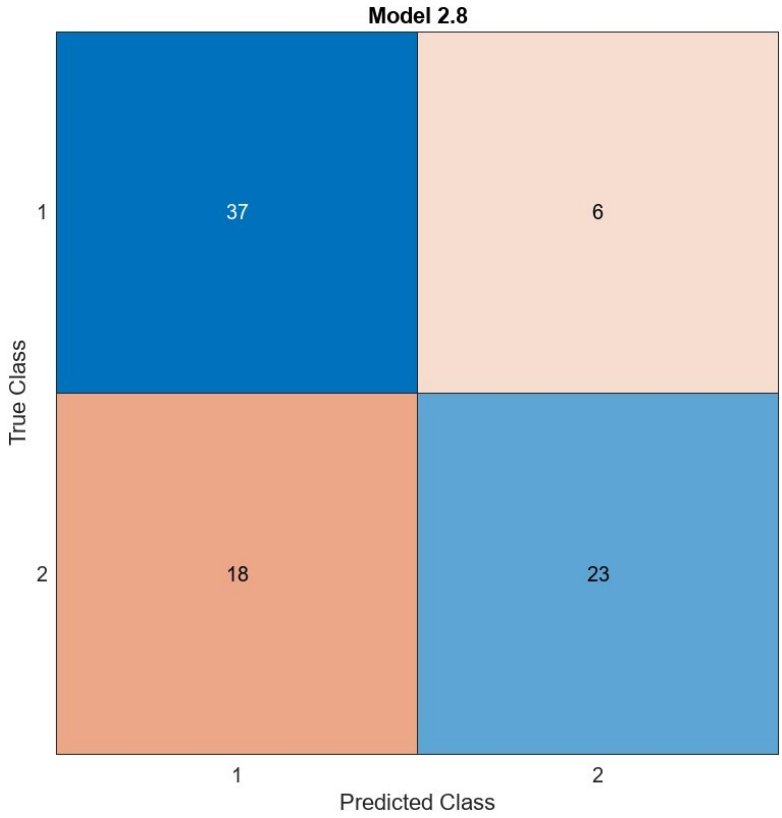

Figure 15: Test confusion matrix

As it is possible to see looking at Figures 14 and 15, with the SMOTE dataset the classifier has been able to classify companies in the two classes, without any polarization. Even if in the training phase almost 180 companies have been misclassified, this allowed the algorithm to learn and resulted into a 24 misclassification out of the 84 companies belonging to the test set. The accuracy (71.4%) this time is a reliable value to evaluate the performance of the model since the classifier has provided to perform well in the classification task.

The second model has been developed using a Coarse Tree algorithm. In this case "Expected Attacks" has been derived utilizing a constant Navg, equal to 50. The reason behind this choice is related to the dataset composition. In particular, among the training set, 39 companies have an attractiveness of 0.6, 37 of 0.7, 148 of 0.8, 138 of 0.9 and 122 of 1. Since organization with a high value of attractiveness are more likely to be target of cyber attacks, the value of 50 attempted attack within a year is considered a good prediction. In Table 15 and Figures 16 and 17 the characteristics of the model have been listed.

Table 15: Characteristics of the model

Training Results	
Accuracy (Validation)	65.9%
Total Cost	165
Prediction Speed	23000 obs/sec
Training Speed	2.3142 sec
Test Results	
Accuracy (Test)	77.4%
Total cost	19

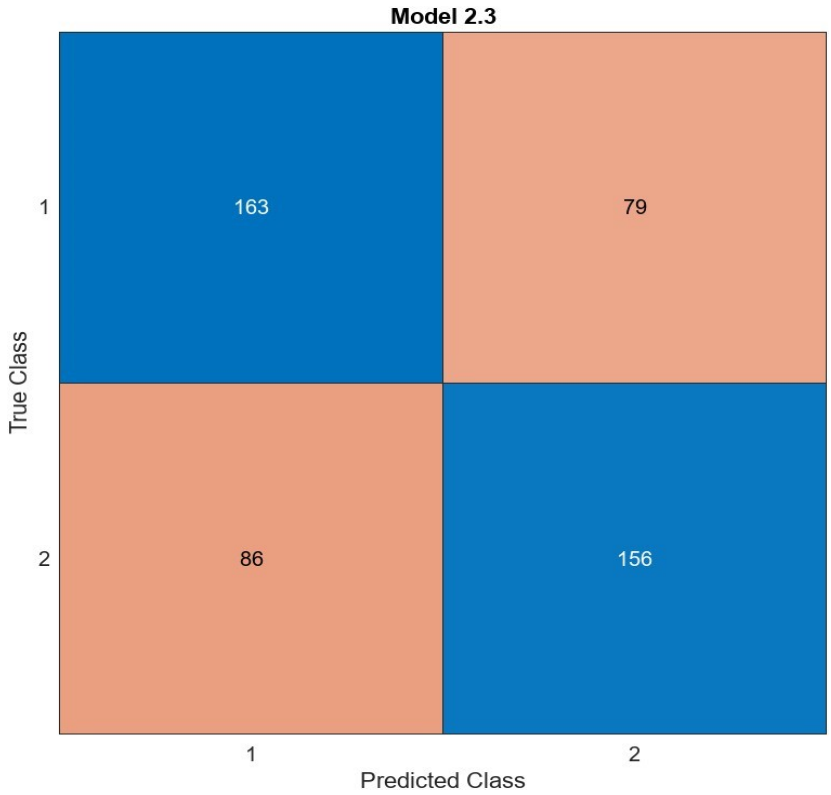

Figure 16: Train Confusion Matrix

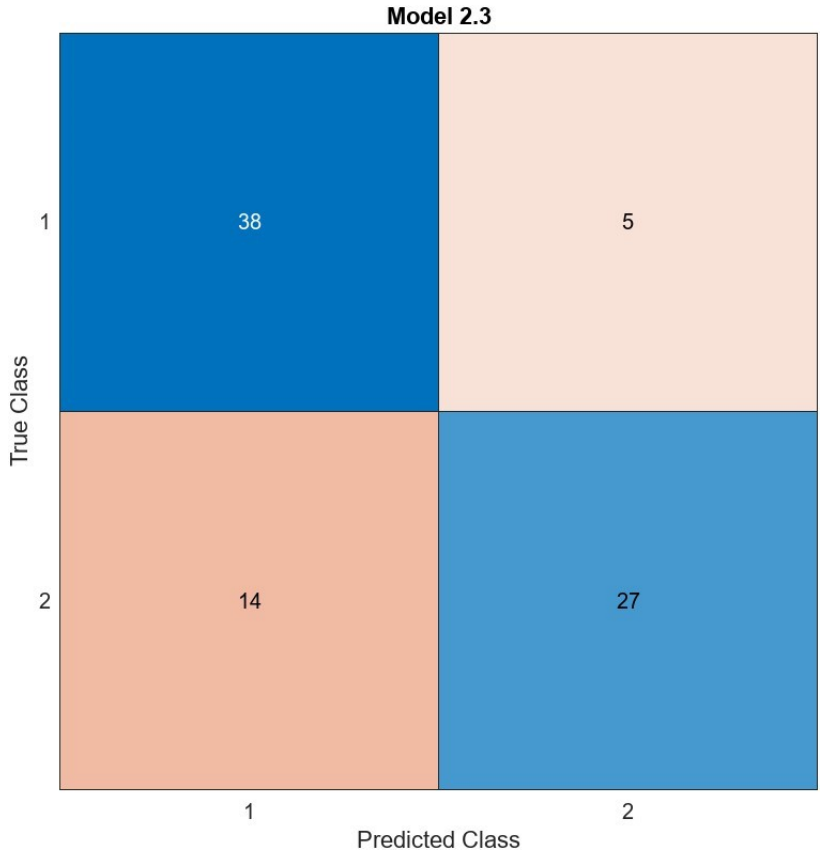

Figure 17. Test Confusion Matrix

As it is possible to witness looking at Figure 17, this model has resulted to be able to discriminate between companies with a low risk to be successfully targeted by a cyber attack and companies with a high risk. The accuracy value of 77.4% reflects Figure 17, in which it is possible to see that only 19 out of 84 companies have been misclassified.

6.3 Reviewed SMOTE dataset

This attempt has been made by eliminating from the training set of the SMOTE dataset 10 companies characterized by a combination of maturity, complexity, attractiveness and number of received attacks that would most certainly lead the machine learning model to a poor training procedure. The reason behind these combinations has been explained in the limitation's paragraph. The best model is based on a Medium Tree algorithm whose characteristics have been listed in Table 16 and shown in Figures 18 and 19.

Table 16: Characteristics of the model

Training Results	
Accuracy (Validation)	61.3%
Total Cost	184
Prediction Speed	20000 obs/sec
Training Speed	2.6046 sec
Test Results	
Accuracy (Test)	78.6%
Total cost	18

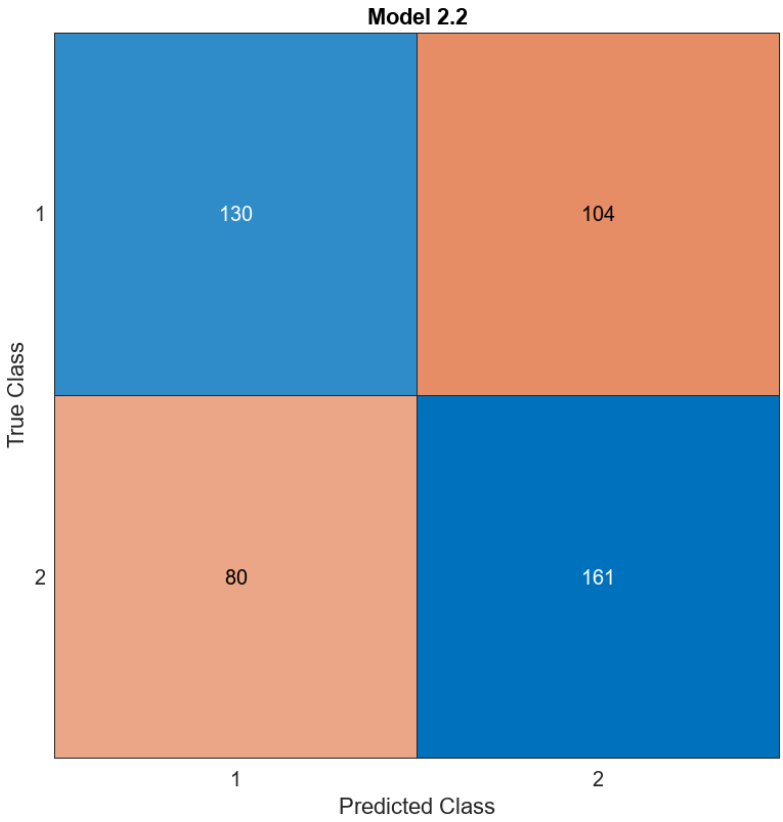

Figure 18: Train confusion matrix

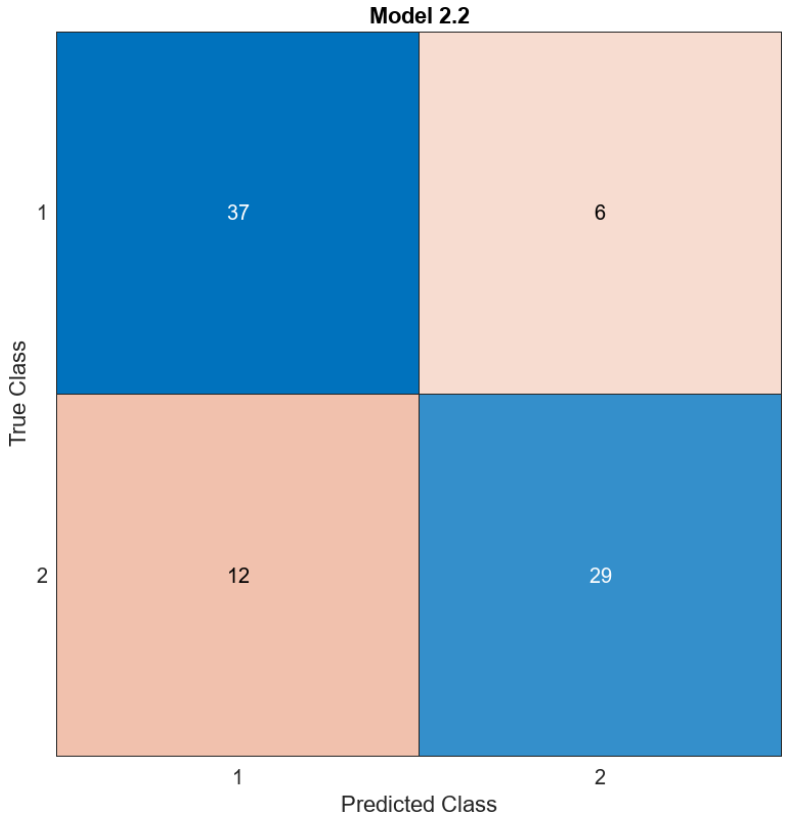

Figure 19: Test confusion matrix

As is can be seen looking at Figure 19, the model is able to discriminate companies between low risk and high risk to be subjected by a cyber attack, even though there is still confusion on high risk companies being classified as low risk companies. The accuracy level of 78.6% together with the confusion matrix shown in Figure 18 and 19, underly the improvement due to the utilization of the reviewed SMOTE dataset with respect of the others, in particular to the standard one.

7. Case Study

In this section two case studies are presented.

7.1 Complexity impact on cyber risk

This section is focused on healthcare organizations specifically, both indirectly and directly. As stated in the section on cyber risk for healthcare organizations, conducting a cyber risk assessment on hospitals is essential in order to implement an adequate degree of cyber security procedure. A successful cyber attack can result in anything from data theft to medical instrument paralysis, resulting in patient delays or even patient death. The adoption of IoMT functionality has increased the complexity of hospital infrastructure, making it more vulnerable. As a result, the impact of the complexity index on cyber risk has been investigated in this section while fixing the other parameters.

Table 17: Fixed parameters

Fixed parameter	
Maturity	7
Attractiveness	0.8
Navg	30

In table 17 the fixed parameter has been listed.

To avoid taking extreme cases into account, the maturity index has been set to 7, the attractiveness index to 0.8, to simulate healthcare infrastructure, and Navg to 30. The results were obtained using both the MAGIC/FAIR and MAGIC/HTMA approaches.

7.1.1 Magic integration with HTMA

The result of the integration of MAGIC/HTMA are depicted in Figure 20. Figure 20 shows that for complexity indexes ≤ 3, the probability of receiving a successful attack remains constant at 51%. The probability of receiving a successful attack increases as complexity increases, particularly between 4 and 7. For complexity 8, it remains constant at 100%. This analysis found that low complexity companies are less likely to be successfully targeted by a cyber attack, whereas complex organizations have a nearly 100% chance, highlighting the importance of implementing the best cyber security algorithms available.

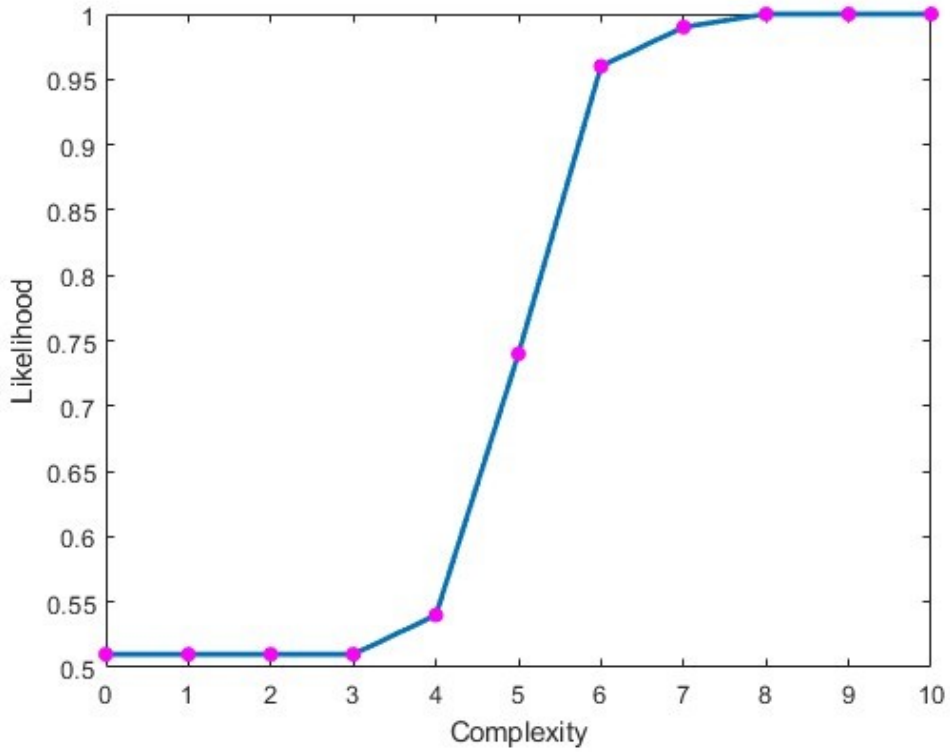

Figure 20: Likelihood in function of the complexity level

7.1.2 MAGIC integration with FAIR

In Figure 21, the results obtained with MAGIC/FAIR are shown. The figure shows that for organizations with a complexity index up to 4, the most likely number of successful attacks received in a year is 0. This value increases for complexity values of 5, reaching a maximum value for a complexity value of 10. This trend shows that low complexity organizations are far less likely to be targeted by successful attacks, whereas highly complex organizations

have a maximum number of expected successful attacks of 23. Furthermore, this demonstrates the importance of adopting the best cyber security algorithms available for high-complexity businesses.

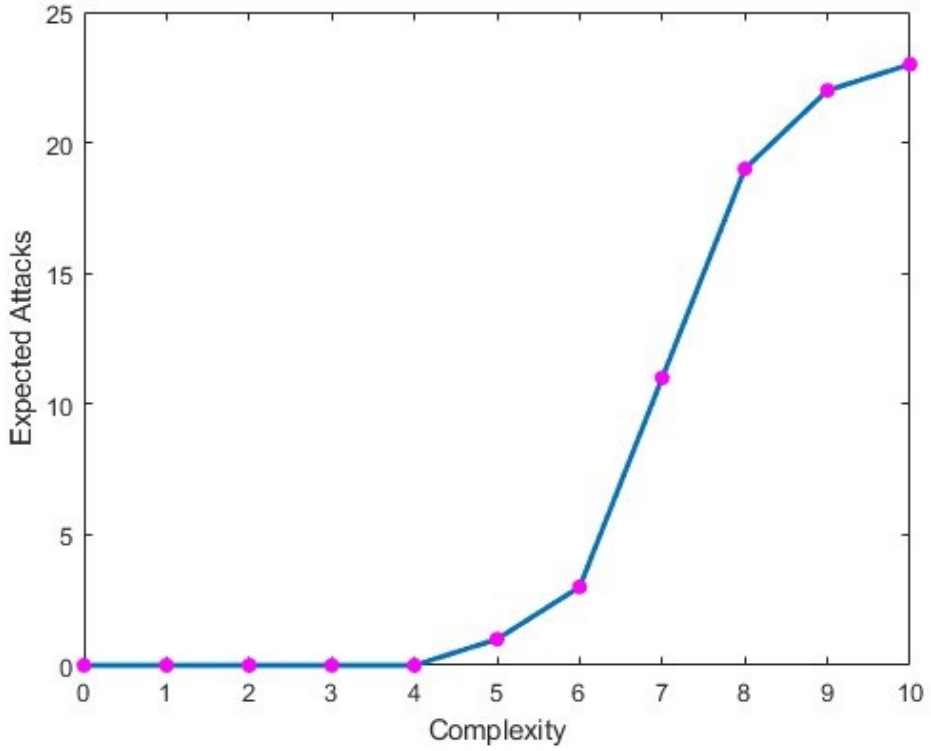

Figure 21: Expected Attacks trend as a function of complexity

7.2 Output of the model as complexity and maturity change

The output of the model based on a Medium Tree algorithm generated using the reviewed SMOTE dataset has been analysed in this section, beginning with the parameters related to the Akorn International Hospital and then by changing maturity and complexity to assess how the output changes. Table 18 lists the parameters associated with the Akorn International Hospital.

Table 18: Parameters related to Akorn International Hospital

ID	Maturity	Complexity	Attractiveness	Attacks
Akorn International Hospital	5	1	0.8	2

In this case, the Akorn International Hospital is a private infrastructure, which explains the low value of the complexity index, as it employs only 1.100 people.
As a first attempt, this hospital was included as the sole entry for the machine learning model test. The model correctly predicted the risk class associated with the Akorn International Hospital, classifying it as low risk (class 1).

For the next attempt, a complexity value of 5 have been considered, maintaining maturity to 5. In this case the implementation of FAIR together with MAGIC has given as output an Expected Attacks of 32, still leading to a classification to the class 1, hence low risk.
The complexity index was then raised to 9, simulating a very complex infrastructure with a low level of maturity. In this case, the FAIR algorithm combined with MAGIC yielded a "Expected Attacks" of 38, and as a result, the model classified the hospital as a high risk infrastructure (class 2).
The model will continue to classify the Akorn International Hospital as low risk based on the results of this test until the complexity index reaches a low or medium value. When the complexity exceeds 7, the ML model begins to classify it as class 2, indicating a high risk.
As final attempt, complexity was set to a starting value of 1 and maturity to 1 and 3 respectively. It is worth noting that a maturity index of 1 represents a security score that can range between 650 and 680, indicating that an organization has poor cyber security coverage but is not completely unprotected.
The previously mentioned combinations are listed in Table 19.

Table 19: Machine learning model output to changes in maturity and complexity

Maturity	Complexity	Navg	Expected Attacks	Class of risk
Fixed maturity				
5	1	50	1	1
5	5	50	32	1
5	9	50	38	2
Fixed complexity				
3	1	50	2	1
1	1	50	21	1

As it is possible to witness looking at Table 19, the maturity index has been always lower than 5, since in this case, having a maturity index greater than 5 would most certainly lead to a classification as low risk. For greater values of the complexity, a poor value of the maturity index led to a classification of high risk, as is it possible to see when the maturity is 5 and the complexity index is 9.

Also in this case, it is worth nothing that a maturity index of 5 reflects a security score between 770 and 800, underlying the necessity to employ optimal cyber security procedure when it comes to cover a really complex infrastructure.

Another interesting case would be to check, fixing the complexity index to 9, the maturity value that would result in a classification to low risk by the ML learning model. Maturity values ≥ 7 have been taken into consideration, and the results are listed in Table 20.

Table 20: Machine learning model output to variation of maturity

Maturity	Complexity	Navg	Expected Attacks	Class of risk
7	9	50	37	2
8	9	50	31	2
9	9	50	16	2
10	9	50	1	2

As can be seen in Table 20, for extremely complex healthcare infrastructure, even a maturity value of 10 yielded a high risk classification. The complexity of the interconnection between components of an infrastructure represents an insurmountable barrier to the correct application of the optimal degree of cyber security procedures, which explains these

findings. Nonetheless, a complexity value of 9 represents a range of employees from 256.000 to 512.000, which cannot be correlated to a single hospital, but to a healthcare infrastructure that includes multiple healthcare facilities.

Finally, due to the limitations listed in the following paragraph, this behaviour may be a reflection of some issues encountered during the training phase.

8. Limitations

In this section, the limitations of this work have been listed.

8.1 Polarization of the Dataset

The standard dataset is completely polarized towards organizations that have been successfully targeted less than three times. This would result is a poor training of the machine learning model, since its ability to recognize between low/high risk would also be completely polarized towards low risk. For this reason, an over fitting procedure has been employed. Despite the fact that the SMOTE algorithm has been demonstrated to perform successfully, the 147 synthesized organizations may have been a challenge for the accurate creation of the model, given their fictitious nature.

8.2 Dataset

Another problem related to the dataset is the number of elements. With the SMOTE procedures, 147 companies have been added reaching 484 companies for the training set. It is not a low value of elements, but for the development of a full-fledged model a greater amount of element would result in an increase of the accuracy, due to a greater training set.

8.3 UpGuard

In terms of UpGuard, the security score parameter assigned to each organization solely considers the previous year. Consequently, it is possible that some businesses with a history of many attacks have improved their cyber security procedures and now have a high security score. Similar to this, businesses with a poor security score and a low number of successful

attacks may have changed their cyber security protocol in the past but not in 2022. This restriction may also reduce accuracy since it interferes with the training process.

9. Conclusions

In this work, the development of a machine learning model to perform cyber risk assessment has been carried out. Various characteristics have been chosen to reflect the likelihood of being the target of a cyber attacks, such as maturity, complexity and attractiveness .Different approaches have been taken, one that take into consideration that after receiving a cyber attack an organization does not change posture, FAIR, and the other taking into consideration that after receiving a cyber attack the organization changes posture. The dataset that has been employed counted 420 companies and the number of successful cyber attacks that targeted each company. Maturity, complexity, and attractiveness parameters have been added for each company and subsequently parametrized. More than one model have been developed, exploiting the possibility to perform over-fitting procedures to obtain better results. The overfitted dataset, with the addition of 147 fictitious companies, counted 484 companies as training set and 84 as test set. The best result has been obtained exploiting the over-fitted dataset and the FAIR methodology integrated with the MAGIC approach, yielding an accuracy of 78.6%. This result underlies the possibility to perform cyber risk assessment with a machine learning approach, while on the other hand, points out some limitation of this approach. The dataset, in particular, is critical, and a better-suited dataset must be used to achieve a better outcome. The results achieved in this paper, using an unbalanced dataset, demonstrate the full potential of this method and its potential future applications.

10. Bibliography

[1] J. Bulaq. (Jan 12, 2023), How much data is created everyday in 2023? [Online] Available:https://techjury.net/blog/how-much-data-is-created-every-day/#.~.text=In%202021%2C%20people%20created%202.5,cloud%20storage% 20around%20the%20globe

[2]M. Baldi, Privacy and data protection, Marche Polytechnic University Department of Information Engineering, 2020-2021

[3] J. cedula, L.R. Young. (2010, Dec.) "A Taxonomy of Operational Cyber Security Risks," *Software Engineering Institute, Carnegie Mellon University*, Pittsburgh, Pennsylvania, Technical Note CMU/SEI-2010-TN-028.
Available: http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=9395

[4] L. Xiaroung, D. Yulong, Y. Shuang-Hua (2019, Apr.), "Safety and security risk assessment in cyber-physical systems", " *IET Cyber-Physical Systems: Theory & Applications*, Vol. 4, pp. 221-232.
Available: https://doi.org/10.1049/iet-cps.2018.5068

[5] M. A. Judge, A. Manzoor, C. Maple, J.P.C Rodrigues, S. u. Islam. (2021, Nov.), "Price-based demand response for household load management with interval uncertainty", " *Energy Reports*, vol.7, pp.8493-8504.
Available: http://www.journals.elsevier.com/energy-reports/

[6] H. Akhavan-Hejazi, H. Mohsenian-Rad. (2018, Nov.), "Power systems big data analytics: An assessment of paradigm shift barriers and prospects," *Energy Reports*, vol. 4, pp. 91-100.
Available: http://www.journals.elsevier.com/energy-reports/

[7] X. Zhang, M. Xu, G. Da, P. Zhao. (2021, Oct.), "Ensuring confidentiality and availability of sensitive data over a network system under cyber threats," *Reliability Engineering & System Safety*, vol. 214, p. 107697.
Available: https://www.sciencedirect.com/science/article/pii/S0951832021002337

[8] M. Amir, T. Givargis. (2020, Dec.), "Pareto optimal design space exploration of cyber-physical systems," *Internet of Things*, vol. 12, p. 100308.
Available: https://www.sciencedirect.com/science/article/pii/S2542660520301402

[9] M. Snehi, A. Bhandar. (2021, May.), "Vulnerability retrospection of security solutions for software-defined Cyber–Physical System against DDoS and IoT-DDoS attacks," *Computer Science Review*, vol. 40, no. C, p. 23.

Available: https://www.sciencedirect.com/science/article/pii/S1574013721000113

[10] C. Ma. (2021, Nov.), “Smart city and cyber-security; technologies used, leading challenges and future recommendations,” *Energy Reports*, vol. 7, pp. 7999-8012.
Available: https://www.sciencedirect.com/science/article/pii/S2352484721007265

[11] M. I. Alghamdi. (2021, Apr.) “Effects of knowledge of cyber security on prevention of attacks,” Materials Today: Proceedings.
Available: https://www.sciencedirect.com/science/article/pii/S2214785321029941

[12] M. Beechey, K. G. Kyriakopoulos, S. Lambotharan. (2021, Nov.), “Evidential classification and feature selection for cyber-threat hunting,” Knowledge-Based Systems, vol. 226, p. 107120.
Available: https://www.sciencedirect.com/science/article/pii/S095070512100383X

[13] L. Yuchong, L. Qinghui. (2021, Nov.), “A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments,” *Energy Reports,* vol. 7, pp. 8176-8186
Available. https://www.sciencedirect.com/science/article/pii/S2352484721007289.

[14] C. Topping, A. Dwyer, O. Michalec, B. Craggs, A. Rashid. (2021, Sept.) “Beware Suppliers Bearing Gifts!: Analysing coverage of supply chain cyber security in critical national infrastructure sectorial and cross-sectorial frameworks,” *Computers & Security,* vol. 108, p. 102324.
Available: https://www.sciencedirect.com/science/article/pii/S0167404821001486

[15] L. Jian, S. Chaowei, S. Qingyu. (2021, Apr.) “Analysis of cascading failures of power cyber-physical systems considering false data injection attacks,” *Global Energy Interconnection,* vol. 4, pp. 204-213. Available: https://www.sciencedirect.com/science/article/pii/S2096511721000402

[16] D. C. Patel, M. F. Berry, P. Bhandari, L. M. Backhus, S. Raees, W. Trope, A. Nash, N. S. Lui, D. Z. Liou, J. B. Shrager. (2021, June) “Paradoxical Motion on Sniff Test Predicts Greater Improvement Following Diaphragm Plication,” *The Annals of Thoracic Surgery,* vol. 111, no. 6, pp. 1820-1826.
Available: https://www.sciencedirect.com/science/article/pii/S0003497520316180

[17] D. Al Shaer, O. Al Musaimi, B. G. de la Torre, F. Albericio. (2020, Dec.) “Hydroxamate siderophores: Natural occurrence, chemical synthesis, iron binding affinity and use as Trojan horses against pathogens”. ” *Eur J Med Chem,* vol. 208, p.112791.
Available: https://www.sciencedirect.com/science/article/pii/S0223523420307637

[18] A.A. Aziz, Z. Amtul. (2021, Nov.) “Developing Trojan horses to induce, diagnose and suppress Alzheimer's pathology,” *Pharmacol Res,* vol. 149, p. 104471.
Available: https://www.sciencedirect.com/science/article/pii/S1043661819316500

[19] N. Kharlamova, S. Hashemi, C. Træholt, (2021, Sept.), "Data-driven approaches for cyber defense of battery energy storage systems," *Energy and AI,* vol. 5, p. 100095.
Available: https://www.sciencedirect.com/science/article/pii/S2666546821000495

[20] R. Bernard, G. Bowsher, R. Sullivan. (2020, Dec.) **"**Cyber security and the unexplored threat to global health: a call for global norms", "*Global Security: Health, Science and Policy,* vol.5, pp.164-141.
Available: https://doi.org/10.1080/23779497.2020.1865182

[21] M. A. Spirito, M. T. Delgado*, "*IoT Innovation: from application scenarios to pilots and deployments*", "in Building the Hyperconnected Society- Internet of Things Research and Innovation Value Chains, Ecosystems and Markets, O.* Vermesan*, P. Friess, Eds. New York, 2015*

[22] T. Mahler, N. Nissim, E. Shalom, I. Goldenberg, G. Hassman, A. Makori, I. Kochav, Y. Elovici, Y. Shahar, "Know Your Enemy: Characteristics of Cyber-Attacks on Medical Imaging", " *in Conf. RSNA Conference*, Chicago, IL, 2017.

[23] G. R. Saade. (2000, Jan.), *"*Looking Back on the Millennium in Medicine*," N. Engl. J. Med.,* vol. 342, no. 1, pp. 42–49.
Available: https://www.nejm.org/doi/full/10.1056/NEJM200001063420106

[24] L. Pycroft, T. Z. Aziz. (2018, June) "Security of implantable medical devices with wireless connections: The dangers of cyber-attacks", "*Expert Review of Medical Devices*, vol.15, pp. 403-406.
Available: https://www.tandfonline.com/journals/ierd20

[25] A. Djenna, D. E. Saïdouni, "Cyber Attacks Classification in IoT-Based-Healthcare Infrastructure," in Conf. *2018 2nd Cyber Security in Networking Conference (CSNet)*, Paris, France, 2018, pp. 1-4.

[26] *S. A. Millar, T. P. Marshall, and N. A. Cardon. (2017, May.22), WannaCry: Are Your Security Tools Up to Date?,* [Online].
Available: https://www.natlawreview.com/article/wannacry-are-your-security-tools-to-date.

[27] B. Brenner. (2017, May.17), *WannaCry: The Ransomware Worm That Didn't Arrive on a Phishing Hook* [Online]*,*
Available: https://nakedsecurity.sophos.com/2017/05/17/wannacry-the-ransomware-worm-that-didnt-arrive-on-a-phishing-hook/

[28] J. Ungoed-Thomas, R. Henry, D. Gadher. (2017, May.14), *Cyber-attack guides promoted on YouTube* [Online].

Available: https://www.thetimes.co.uk/article/cyber-attack-guides-promoted-on-youtube-972s0hh2c.

[29] A. Hern, S. Gibbs. (2017, May.12), *What is 'WanaCrypt0r 2.0' ransomware and why is it attacking the NHS?* [Online].
Available: https://www.theguardian.com/technology/2017/may/12/nhs-ransomware-cyber-attack-what-is-wanacrypt0r-20

[30] BBC news. (2017, May.13), *NHS cyber-attack: GPs and hospitals hit by ransomware* [Online],
Available: https://www.bbc.com/news/health-39899646

[31] Microsoft Corporation. (2017), Vulnerabilities, CVE-2017-0144, NIST.
Available: https://nvd.nist.gov/vuln/detail/cve-2017-0144

[32] D. Munro. (31, Mar. 2016), *Assessing the Financial Impact of 4.5 Million Stolen Health Records* [Online].
Available: https://www.forbes.com/sites/danmunro/2014/08/24/assessing-the-financial-impact-of-4-5-million-stolen-health-records/?sh=f801f087da00

[33] T. Floyd, M. Grieco, E. F. Reid, "Mining hospital data breach records: Cyber threats to U.S. hospitals," in Conf. *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*, Tucson, AZ, 2016, pp. 43-48.

[34] M. Battaglioni, G. Rafaiani, F. Chiaraluce, M. Baldi, "MAGIC: A Method for Assessing Cyber Incidents Occurrence," *IEEE Access*, vol. 10, pp. 73458-73473, 2022.

[35] J. Freund, J. Jones, *Measuring and Managing Information Risk: A FAIR Approach*, Oxford, U.K: Butterworth-Heinemann, 2014

[36] D. W. Hubbard, R. Severson, *How to Measure Anything Cybersecurity Risk*, Hoboken, NJ, USA: Wiley, 2016

[37] Front. Big Data, 26 January 2021 Sec. Cybersecurity and Privacy
Available:https://doi.org/10.3389/fdata.2020.521132

[38] ©2022 UpGuard, Inc

[39] B. S. Larsen (2023). Synthetic Minority Over-sampling Technique (SMOTE)
Available:https://github.com/dkbsl/matlab_smote/releases/tag/1.0)