Islamic University of Gaza
Deanery of Higher Studies
Faculty of Information Technology
Information Technology  Program

الجامعة الاسلامية-غزة
عمادة الدراسات العليا
كلية تكنولوجيا المعلومات
برنامج تكنولوجيا المعلومات

# *Detection Model for Pharming Attack Based on IP-Address Check and Website Predictability*

*Prepared by*

***Areej N. El-Buhaisi***
***220093710***

*Supervised by*

***Dr. Tawfiq S.Barhoom***

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science In Information Technology

*2013- 1434H*

بسم الله الرحمن الرحيم

الجامعة الإسلامية – غزة
The Islamic University - Gaza

مكتب نائب الرئيس للبحث العلمي والدراسات العليا    هاتف داخلي: 1150

ج س غ/35/

Ref ..........2013/10/07م الرقم

Date...................... التاريخ

## نتيجة الحكم على أطروحة ماجستير

بناءً على موافقة الدراسات العليا بالجامعة الإسلامية بغزة على تشكيل لجنة الحكم على أطروحة الباحثة/ **أريج نصر حمدان البحيصي** لنيل درجة الماجستير في **كلية *تكنولوجيا المعلومات*** **برنامج تكنولوجيا المعلومات** وموضوعها:

**Detection Model for Pharming Attack Based on IP Address Check and Website Predictability**

وبعد المناقشة التي تمت اليوم الاثنين 02 ذو الحجة 1434هـ، الموافق 2013/10/07م الساعة العاشرة صباحاً، اجتمعت لجنة الحكم على الأطروحة والمكونة من:

| | | |
|---|---|---|
| **د. توفيق سليمان برهوم** | **مشرفاً ورئيساً** | ............................. |
| **أ.د. نبيل محمود حويحي** | **مناقشاً داخلياً** | ............................. |
| **د. تامر سعيد فطاير** | **مناقشاً خارجياً** | ............................. |

وبعد المداولة أوصت اللجنة بمنح الباحثة درجة الماجستير في **كلية *تكنولوجيا المعلومات*/ برنامج تكنولوجيا المعلومات**.

***واللجنة إذ تمنحها هذه الدرجة فإنها توصيها بتقوى الله ولزوم طاعته وأن تسخر علمها في خدمة دينها ووطنها.***

والله ولي التوفيق ،،،

**مساعد نائب الرئيس للبحث العلمي و للدراسات العليا**

**أ.د. فؤاد علي العاجز**

بسم الله الرحمن الرحيم

وما توفيقي إلا بالله

## Dedication

*To the soul of my brother:*  
*Eng. Basem El-Buhaisi*  
*To my beloved mother and father*  
*To my lovely husband*  
*To my children Shayma'a and Salah*  
*To my brother*  
*To my sisters*  
*To all friends*

# Acknowledgements

All praise is for Allah, the Almighty for providing me with the strength to complete this work and for guiding me at every stage of my life.

This thesis is the result of years of work whereby I have been accompanied and supported by many people. It is wonderful that I now have the opportunity to express my gratitude to all of them.

I'd like to thank everyone who has helped me in completing this work. I submit my highest appreciation to my thesis advisor **Dr. Tawfiq S. Barhoom**, who helped in every step of the way. I would like to express my deep and sincere gratitude to him. His understanding and personal guidance have provided a good basis for the present thesis.

I would also like to thank **Eng. Mahmoud El-Hoby**, for his valuable scientific and technical notes; I also extend my thanks to **Ms. Rana Shubair** for proof reading my thesis.

Also, I would like to take this opportunity to express my profound gratitude to my beloved family, especially my husband, and to my mother, father, brother and sisters - without whom I would never have been able to achieve so much.

I offer my thanks and appreciation to all of those who supported me in any respect during the completion of the research.

Last, but certainly not least, I want to thank my colleagues, for their support during my years of study.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AIWL | Automated Individual White-List |
| AOL | America Online |
| APWG | Anti Phishing Working Group |
| DNS | Domain Name System |
| DT | Decision Tree |
| FTP | File Transfer Protocol |
| HTTP | Hyper Text Transfer Protocol |
| IM | Instant Messaging |
| IP | Internet Protocol |
| IPALT | Internet Protocol Alternate |
| IPLOC | Internet Protocol Locally |
| KNN | K- Nearest Neighbor |
| LAN | Local Area Network |
| LUIs | Login User Interfaces |
| ML | Machine Learning |
| NB | Naïve Bayes |
| P2P | Peer to Peer |
| SSL | Secure Sockets Layer |
| TCP | Transport Control Protocol |
| URL | Uniform Resource Locator |
| W3C | World Wide Web Consortium |

# نموذج لمكافحة هجمات التصيد بالاعتماد على فحص عنوان بروتوكول الانترنت والتنبؤ بصفحة الانترنت المصابة

مع الانتشار الهائل والرغبة الدائمة للاتصال بشبكة الإنترنت، ومع أهمية وخصوصية حفظ المعلومات الشخصية والمالية الحساسة والتي أصبحت أكثر إثارة للقلق لاصحابها، لذلك فقد أصبح من الضروري البحث عن حلول مناسبة وفعالة للحفاظ على هذه المعلومات من السرقة والعبث من قبل المهاجمين والعابثين على الشبكة العالمية الانترنت. هجمات التصيد هي واحدة من أكبر المخاوف بالنسبة للمستخدمين الذين يتعاملون مع معلومات حساسة وشخصية مهمة، وذلك لأن المهاجم في هذا النوع من الهجمات "التصيد" يقوم بعمل نسخة مثالية للموقع، ويقوم بالتلاعب في الخادم المزود لأسماء النطاقات وذلك لتوجيه المستخدم إلى موقع احتيالي مزور ولكنه مطابق تماما للموقع الاصلي بحيث لا يشعر المستخدم بانه مزور وينخدع به ويقوم بادخال معلوماته الحساسة من حسابات خاصة او بنكية ويقوم المهاجم حينها بسرقتها. في هذا العمل الذي نقدمه قمنا بتعريف نموذج جديد لإحباط هجمات التصيد، وهذا النموذج يعتمد على طريقة للحل مكونة من خطوتين، الأولى التحقق من عنوان بروتوكول الانترنت للتأكد من عدم وجود أي تلاعب على الخادم المزود لأسماء النطاقات، ثم الخطوة الثانية هي عمل تصنيف لمواقع الانترنت للتنبؤ بوجود أي مواقع مزورة ومصابة بهجمات التصيد إن وجدت. الجزء الثاني من الحل هو تعزيز وتقوية للحل، وذلك باستخدام واحدة من تقنيات تنقيب البيانات. مع أكثر من 96.7٪ نسبة نجاح قمنا بتقييم وفحص نموذجنا.

**كلمات مفتاحية:** هجمات التصيد، تنقيب البيانات، التصنيف ،خادم أسماء النطاقات.

# Detection Model for Pharming Attack Based on IP-Address Check and Website Predictability

## Abstract

With the deployment of broadband Internet access, the importance of saving sensitive personal and financial information becomes more troubling, so it has become necessary to search for appropriate and effective solutions to keep this information from theft and tampering by attackers. Pharming attack -the sophisticated version of phishing attack - is one of the greatest concerns for users who deal with sensitive information, because the attacker makes an ideal copy of the site, and manipulates the DNS (Domain Name System) to route the user to that fraudulent site, then steels the information. In this work, we have defined a new model to defeat pharming attacks on client side. This model depends on a two-step solution; first IP address check to be sure there is not any manipulation on DNS server, then the second step is to make classification on the websites to predict the pharming sites if any exists. The second part of the solution is to strengthen the solution using Naïve Bayes classification approach which is one of the Data Mining techniques. With over 96.7% success we evaluated our model.

**Key words:** Phishing attack, Pharming attack, DNS server, Data Mining, Classification.

# Chapter (1): Introduction

Phishing attacks are a major concern for saving Internet users privacy. By combining social engineering and website forgery techniques, phishing attacks spoof the identity of a company typically a bank or an e-commerce site, to trick internet users to give sensitive information like login, password, and credit card number. The ideal phishing attack creates a website very similar to the legitimate one by using the same structure, images and so on. However, if the user carefully examines the URL displayed in the address bar of the web browser, he should notice that the URL (especially the domain name) is not the same as the original one. On the other hand, pharming attacks are much more complex to detect because the user cannot distinguish between the visited URL and the original website, because the forgery 'URL' website is similar to the legitimate site [5].

Pharming is an attack aiming to redirect a website's traffic to another forged site. Pharming can happen either by changing the hosts file on a victim’s computer or by exploitation of a vulnerability in DNS (Domain Name System) server software. DNS servers are computers responsible for resolving internet names into their real addresses (IP addresses). Modification of DNS servers are sometimes called "poisoned"[24].

The term pharming is new based on farming and phishing. In recent years both pharming and phishing have been used for online identity theft information. Pharming has become a major concern to businesses hosting ecommerce and online banking websites.

## 1.1 Pharming History:

The term Phishing appeared in early 1996, but it was not until the end of 2003 that email based phishing attacks began to become a popular attack vector for cyber criminal as a means to conduct financial fraud and identity theft [8].

As for Pharming attack, in 2005 was the first warning about new attacks by that corrupt some DNS server, in this attack the number of (.com) sites were directed to forgery servers maintained by attackers [8].

## 1.2 DNS server:

The Domain Name System (DNS) is a standard technology for managing the names of Web sites and other Internet domains. DNS technology allows the user to type names into the Web browser like google.com and the browser automatically finds that address (IP address) on the Internet. A key element of the DNS is a worldwide collection of DNS servers.

DNS server is any computer registered to join the Domain Name System. A DNS server runs special-purpose networking software, features a public IP address, and contains a database of network names and addresses for other internet hosts [9].

The DNS plays a critical role in supporting the internet infrastructure by providing a distributed and fairly robust mechanism that resolves Internet host names into IP addresses and IP addresses back into host names [9].

## 1.3 Data Mining and Classification:

Data mining involves the use of advanced data analysis tools to discover previously unknown, valid patterns and relationships in a large data set. These tools may include statistical models, mathematical algorithm and machine learning methods. This means that, data mining consists of more than collecting and managing data, it also includes analysis and prediction [10].

Classification is a data mining (machine learning) technique used to predict group membership for data instances. Classification technique is capable of processing a wide variety of data than regression and is growing in popularity [10].

In this research, the work of defeating pharming attack on the client side is performed done in two steps; first IP address check to be sure that the DNS server is not infected with any pharming attack, then the classification of the website to predict the phishing sites if any should exist.

## 1.4 Statement of the problem

Pharming attacks are dangerous to users since they are used to steal sensitive information like usernames, passwords, and credit card numbers, and it's much more dangerous to users, since the user can't notice the difference between normal site and forged site in URL like phishing attack. The previous detection methods use classification to classify the phishing emails or sites. These methods focused on phishing attacks. But pharming attacks are much more complicated than phishing attacks. Other methods use IP address check and page contents analysis to detect pharming website. The problem caused by employing these methods is that they lack combination of phishing and pharming attack detection, and they need more rules on their classifier to improve it.

## 1.5 Research Objectives

With the deployment of broadband Internet access, the importance of saving sensitive personal and financial information becomes more troubling, so our objectives of this research is focused on these essentials. We can split the objectives to main objective and specific objectives.

### 1.5.1 Main Objective:

The main objective of this proposal can be stated as *"To design and develop a new Model for protecting web users from phishing and Pharming attack with more efficient and power using dual approach: IP address check and website predictability"*.

### 1.5.2 Specific Objectives:

These are specific objectives which could be extracted from the main objective:

- Literature reviews of the recent researches to defeat phishing and pharming attack, to analyze and compare between them in order to take advantage from them in our research.
- Finding the appropriate classifiers to predict phishing and pharming websites.
- Analyzing our model for defeating pharming attacks.
- Designing our model; which combines IP address check and web page predictability.
- Implementation of the proposed model.
- Testing of the proposed model.
- Evaluating the accuracy of our implemented model.

## 1.6 Importance of the research

Pharming attack is very dangerous to users who are dealing with sensitive accounts like user name, password and credit card number, since it is important information relating to personal and financial data. That's why keeping users aware attention of forgery websites is an important issue.

The research will help users to avoid fraud and theft of personal and financial information and accounts by using dual step IP address check and web page predictability. These two steps will make the solution more powerful.

## 1.7 Scope and limitations

In this research the scope and limitations can be summarized as follow:

- In the classification method of this research we used a specific dataset downloaded from the site phishtank[28] and we collected phishing characteristics and indicators to do this research.

- The age of the dataset is the most important problem, which is relevant to the phishing websites. The average phishing site stays live about 2.25 days [4].

- Also this defeating model is done at client side to save users from forgery of pharming sites.

- In our experiments we used PHP files and scripts in collecting pharming characteristics.

- I don’t consider changes in IP address between host and domain.

## 1.8 Outline of the Thesis

This thesis has been divided into six major chapters, which are structured around the objectives of the research. The thesis is organized as follows:

**Chapter 2,** Presents literature review of phishing and pharming detection approaches. Also, this chapter presents details about machine learning and data mining techniques, classification methods, and classification algorithms used on our model. DNS server is the important part of our research.

**Chapter 3,** Presents some related work of phishing and pharming detection, and highlights its main setbacks which are to be avoided and solved in our work.

**Chapter 4,** Includes the methodology steps and the architecture of our model. An explanation about the data sets used in the experiments, preprocessing of these data set, and the experiment cases are included as well. Also, this chapter presents the baseline experiments to choose the optimal classifiers algorithms.

**Chapter 5,** Gives the details about the sets of experiments, and analyzes the experimental results. Also discussion for each set of experiments, and reviews some experiment scenarios for comparison goals.

**Chapter 6**, Draws the conclusion and summarizes the research achievement of experiments and suggests future work.

# Chapter (2): Literature Review

In this chapter, we will identify phishing and pharming attacks, the history of these attacks, how the attacker can trick the user and steal their identity, and we will talk about DNS server, finally Data mining techniques.

## 2.1 Phishing attack

Phishing is a social engineering technique used to take advantage of human ignorance. It allows people to exploit the weaknesses in web security technologies.
The word “phishing” originally came from the analogy of early Internet criminals using email lures to “fish” for passwords, credit cards, and financial data from unsuspecting Internet users [19].
Phishing is defined as "attempting to steal sensitive information, by disappearing in a trusted entity in an electronic communication"[19]. As a social engineering technique it's a logical continuation of dressing up in a fake uniform to gain access to a certain area. But because of the multicast nature of email and other electronic communications, it is far easier to find a target. In practice there is little chance of the people operating phishing scams getting caught, and hence fooling some of the people they target which bring in a perfectly adequate revenue stream.

According to a Symantec Intelligence Report issued in February 2012, the global phishing rate increased by 0.01 percentage points since January 2012, taking the global average rate to one in 358.1 emails (0.28%), (See Figure 2.1).

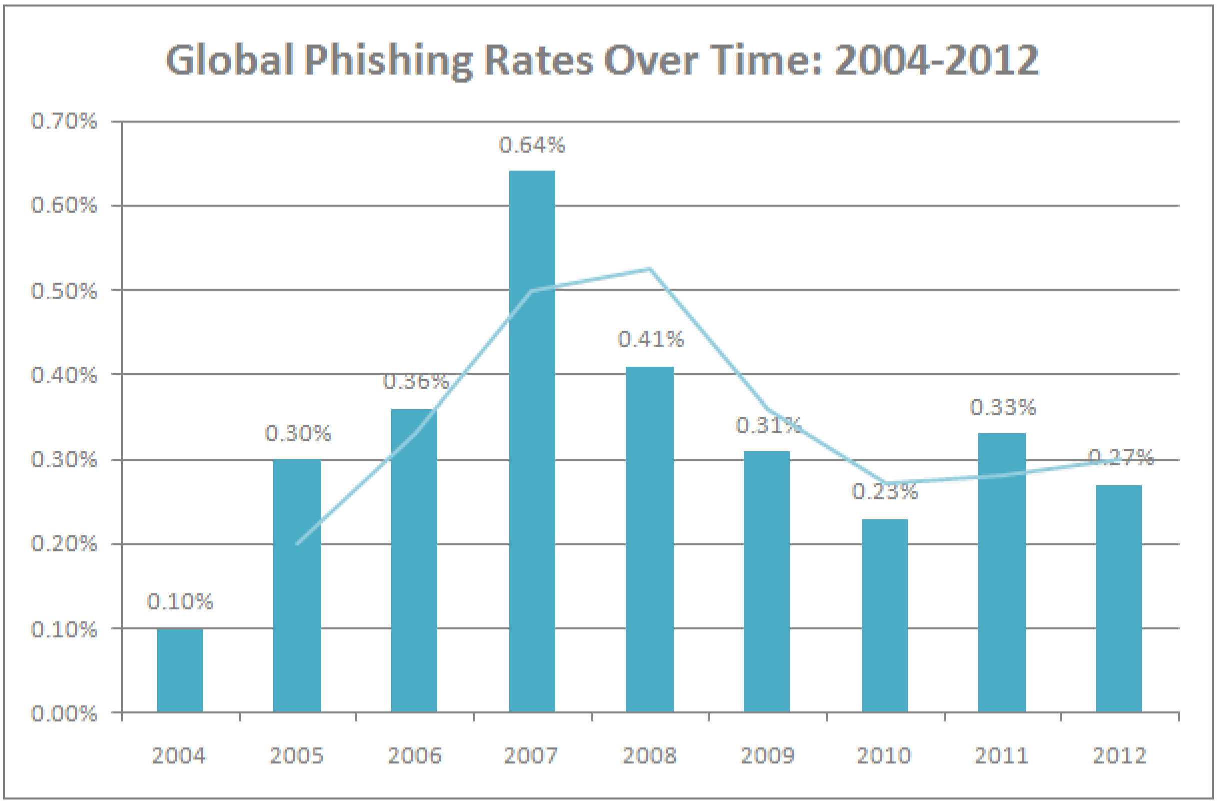


*Figure 2.1: Global Phishing Rates over time,[31]*

## 2.2 Pharming attack

Pharming is a form of domain spoofing. In simple terms, rather than spamming you with email requests to confirm your financial or personal information, pharmers work invisibly. They change your local DNS server to redirect your Web request to a fake site. This means that when you enter a web address, such as www.iugaza.edu; you will be taken to a fake website rather than the legitimate website.

As far as you know, you're connected to the correct site. No email is involved, and if they copied the appearance of the real site well, you would have no way to know that anything was wrong [19].

## 2.3 History of Phishing and Pharming attacks

The term phishing was stated when the America Online (AOL) accounts were stolen by attackers using email in year 1996. The term phishing was derived from the concept of fishing hook in which the attackers use email to lure the user's AOL password. The character "f" of fishing is then being replaced by "ph" to keep it compatible with the computer hackers' tradition. Phishing works by using social engineering to lure consumers to detect their sensitive personal information at fake websites or known as spoofed site, sending email, through instant messaging (IM), Peer to Peer (P2P) network, search engines and etc [8].

Pharming was the evolution of phishing that is also used to steel consumer's sensitive information by using technical tricks like sending email containing viruses or Trojan horse that will install a small application program to the targeted victims' computer. The term Phishing appeared in early 1996, but it was not until the end of 2003 that email based phishing attacks began to become a popular attack vector for cyber criminals as a means to conduct financial fraud and identity theft [8].

In 2005 there was first warning about new attacks by some corrupt DNS servers. In this attack the number of (.com) sites was directed to forgery servers maintained by attackers [8].

The application program will redirect users to a fraudulent website when they visit an authentic official website. Beside this, the attacker will also use those well known traditional techniques like DNS cache poisoning, domain spoofing and other techniques to redirect users to the fraudulent website when users want to visit an authentic website.

The term pharming is new based on farming and phishing. In recent years both pharming and phishing have been used for online identity theft information. Pharming has become a major concern to businesses hosting ecommerce and online banking websites [24].

## 2.4 Techniques of Pharming

Pharming, relies on changing the DNS entries of the organization’s website. There are multiple ways to do this. These are:

### 2.4.1 Sending Email

Pharming is carried out by attackers in several ways. The attacker will send email to the targeted victim that contains viruses or Trojan horse that will download and run on the user's computer. The recipient of the email can be duped by the attackers even if they did not open or download the attachment in the email. The viruses or Trojan horse contained in the email will install a small application in the recipient's computer that will try to redirect the recipient to the fraudulent website when the recipient tries to visit an authentic website [19].

### 2.4.2 DNS Cache Poisoning

DNS cache poisoning can be carried out by using malicious responses or taking DNS software vulnerability to "poison" the cache that stores queries made by users in a certain amount of time in order to speed up the user response time for frequently used domains in order to enhance the user experience. After the cache being "poisoned", when the user makes queries at the DNS, the user will be redirected to the fake website where they are asked to update their personal information [19][29].

### 2.4.3 Domain Hijacking

Domain hijacking is performed by skipping the confirmation of the old domain registrar and the domain owner where the change of domain registrar can only be made with the confirmation of three parties, the domain owner, old registrar and new registrar [19].

### 2.4.4 DNS Server Hijacking

Pharming may be performing through DNS server hijacking. To hijack a DNS server, the attacker will first target the DNS server on the LAN or DNS server hosted by the ISP to change the IP address of an authentic website's domain name to the IP address of the fake website. When the user tries to visit the authentic website, queries will be made on the DNS server for the IP address of the domain name. Since the IP address of the domain name has been changed, it will redirect the user to the fraudulent website. When the user is being redirected to the fraudulent website, they will perform the activities that they wish to perform at the website because the address displayed in the address bar remains the same as the authentic website's address and they think that they are accessing the authentic website. Through the activities that are performed by the user, the attacker will be able to obtain the information that they addresses that start with HTTP but not HTTPS because the website is without SSL protection and this is the summary of the steps:

1. Attackers target the DNS server on the LAN or DNS server hosted by the ISP to change the IP address of an authentic website's domain name to the IP address of fraudulent website
2. User tries to visit the authentic website
3. Queries will be made on the DNS server for the IP address of the domain name
4. The IP address gathered from the DNS server is the IP address of fraudulent website
5. User is being redirected to the fraudulent website

### 2.4.5 Static domain name spoofing

The pharmer may attempt to take advantage of slight misspellings in domain names to trick users into visiting the malicious websites [19].

## 2.5 Domain Name System (DNS):

Humans can't think like computers. Humans simply can't remember dozens of IP addresses. They need easy-to-remember names to locate their mail server or their favorite web sites. To make things easy on the internet, DNS was therefore invented. And with it came a new place for hackers of all sorts to have fun.

The purpose of DNS makes it a very sensitive area; for this is the place the client connection is orientated. The possibilities a black-hat can have by succeeding in hacking DNS are tremendous (a user can be directed to a host controlled by a hacker, whatever service he might be using: http, ftp, telnet) [24].

DNS server is any computer registered to join the Domain Name System. A DNS server runs special-purpose networking software, features a public IP address, and contains a database of network names and addresses for other Internet hosts [9].

The DNS plays a critical role in supporting the internet infrastructure by providing a distributed and fairly robust mechanism that resolves internet host names into IP addresses and IP addresses back into host names [9].

### 2.5.1 How DNS work:

DNS stands for Domain Name Service. All in all, what it does is to translate a host's name into its IP address. Internet is an IP network. Every host is affects an IP address that must be known to any other host willing to communicate. But it would be impossible for a human being to remember all the IP addresses it will use on the internet. It would be possible to create the mappings between IP addresses and names locally to each computer. But the update of those tables would be very complex and slow given the number of computers on the internet and how fast a modification in their address or name cans occur.DNS provides a way to know the IP address of any host on the Internet. It is no different than any other directory service [24]. Figure (2.2) showed how DNS work.

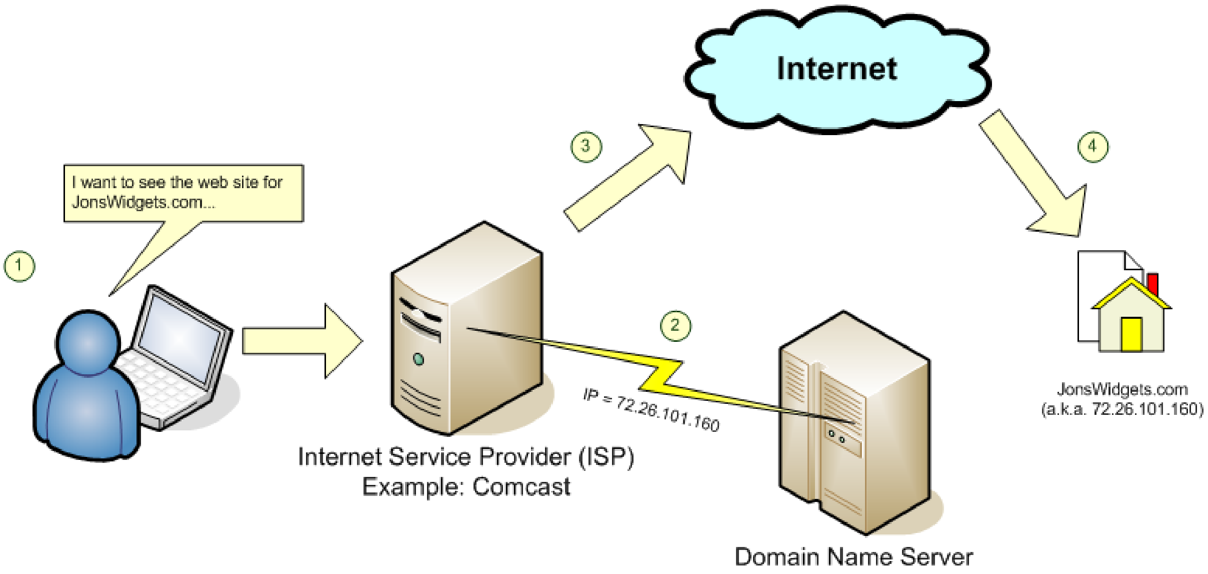


*Figure 2.2 The work of DNS [30].*

## 2.6 Data mining:

It is considered as one of the applications of supervised machine learning, and it plays an important role in the process of retrieving the lost information [18] [20]. Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data set. These tools can include statistical models, mathematical algorithm and machine learning methods. Consequently, data mining consists of more than collection and managing data, it also includes analysis and prediction. Classification technique is capable of processing a wider variety of data than regression and is growing in popularity.

There are several applications for Machine Learning (ML), the most significant of which is data mining. People are often prone to making mistakes during analyses or, possibly, when trying to establish relationships between multiple features.

This makes it difficult for them to find solutions to certain problems. Machine learning can often be successfully applied to these problems, improving the efficiency of systems and the designs of machines.

Many terms carry a similar or slightly different meaning to data mining, such as knowledge mining from data, knowledge extraction, data/pattern analysis, data archaeology, and data dredging [18].

Data mining functionalities are used to specify the type of patterns to be found in the data mining tasks. In general data mining tasks can be classified into two main categories: descriptive and predictive. Descriptive mining tasks characterize the general properties of the data. Predictive mining tasks perform inferences on the current data in order to make predictions [18]. Most of data mining tasks can be one or combination of the following:

1. Classification: used for predictive mining tasks. This method is intended for learning different functions that map each item of the selected data into one of a predefined set of classes. Given the set of predefined classes, a number of attributes, and a "learning (or training) set," the classification methods can

automatically predict the class of other unclassified data of the learning set [18].

2. Prediction: used for predictive mining tasks. Analysis is related to regression techniques. The key idea of prediction analysis is to discover the relationship between the dependent and independent variables. For example, by using historical data from both sales and profit, either linear or nonlinear regression techniques can produce a fitted regression curve that can be used for profit prediction in the future [20].
3. Association Rules: used for descriptive mining tasks. It aims to find out the relationship among valuables in database, and produce a set of rules describing the set of features that are strongly related to each other’s, so that the relationship of a particular item in a data transaction on other items in the same transaction is used to predict patterns [18].
4. Clustering: used for descriptive mining tasks. It is unsupervised, and does not require a learning set. It shares a common methodological ground with Classification. It ungroupes data and uses automatic techniques to put this data into groups [20]. In other words, finds groups of data points (clusters) so that data points that belong to one cluster are more similar to each other than to data points belonging to different cluster.
5. Outlier Analysis: used for predictive mining tasks. Discovers all data points that are different from the rest of data. Such points are known as exceptions or surprises. While outliers can be considered noise and discarded in some applications, they can reveal important knowledge in other domains, and thus can be very significant and their analysis valuable. So it is very important to identify the outliers [19].

### 2.6.1 Classification

It is one of the data mining techniques that fall under supervised machine learning techniques classification. The classifier needs to be trained with labeled input examples, so that it could understand the characteristics of different classes, and then, it could map new data items to different classes [20]. There are many classification algorithms in data mining. We will describe some of those algorithms in order to be used in our research such as Naïve Bayes (NB), Decision Tree (DT), and K- Nearest Neighbor (KNN) algorithms. Following is a brief overview about the classification algorithms mentioned above [18].

#### 2.6.1.1 Naïve Bayes (NB)

Naïve Bayes is a technique for estimating probabilities of individual variable values, given a class, from training data and to then allow the use of these probabilities to classify new entities, which is a term in Bayesian statistics dealing with a simple probabilistic classifier based on applying Bayes' theorem (from Bayesian statistics) with strong (naive) independence assumptions. In simple terms, a naïve Bayes classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature. The naïve Bayesian classifier, works as following derivation [18]:

1. Let D be a training set of tuples and their associated class labels. Each tuple is represented by an n-dimensional attribute vector, X = (x1, x2, ….. , xn), n

measurements made on the tuple from n attributes, respectively, A1, A2, … , An.

2. Suppose that there are m classes C1, C2, …. , and Cm. Given a tuple, X, the classifier will predict that X belongs to the class having the highest posterior probability, conditioned on X. That is, the naïve Bayesian classifier predicts that tuple X belongs to the class Ci if and only if

$$P(C_i|X) > P(C_j|X) \quad \text{for } 1 \le j \le m, j \ne i. \quad \longrightarrow \quad (1)$$

The maximize P(Ci|X). The class Ci for which P(Cj|X) is maximized is called the maximum posteriori hypothesis. By Bayes' theorem (Next equation).

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)}. \quad \longrightarrow \quad (2)$$

3. Since P(X) is constant for all classes, only (P(Ci|X) = P(X |Ci)P(Ci)) need to be maximized.
4. Based on the assumption is that attributes are conditionally independent (i.e., no dependence relation between attributes), the computing of P(X|Ci) using the following equation:

$$P(X|C_i) \quad = \quad \prod_{k=1}^{n} P(x_k|C_i) \quad \longrightarrow \quad (3)$$

Reduces the computation cost by Equation (P(Ci|X) = P(X |Ci)P(Ci), only counts the class distribution. If Ak is categorical, P(xk|Ci) is the no. of tuples in Ci having value Xk for Ak divided by |Ci, D| no. of tuples of Ci in D. And if Ak is continuous-valued, P(xk|Ci) is usually computed based on Gaussian distribution with a mean μ and standard deviation σ and P(xk|Ci) is:

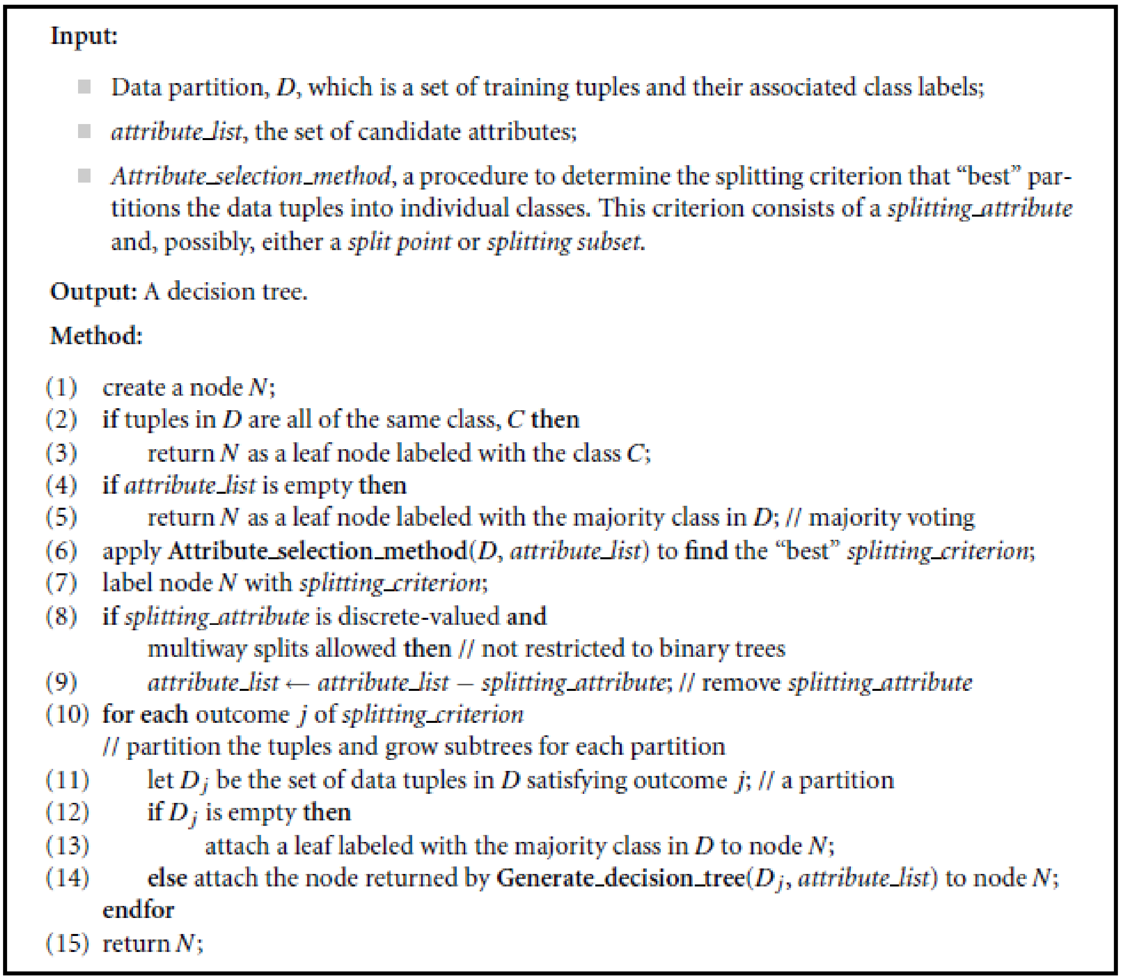
$$g(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \longrightarrow \quad (4)$$

$$P(x_k|C_i) = g(x_k, \mu_{C_i}, \sigma_{C_i}). \quad \longrightarrow \quad (5)$$

Where μ is the mean and σ is the variance. If an attribute value doesn't occur with every class value, the probability will be zero, and a posteriori probability will also be zero.

### 2.6.1.2 Decision Tree (DT)

Decision Tree is a common method used in statistics, data mining and machine learning, where it is an efficient method for producing classifiers from data. It is considered as a tree-structured plan of a set of attributes to be tested in order to predict the output. In these tree structures, leaves represent class labels and branches represent conjunctions of features that lead to those class labels. Moreover, it is a type of tree-diagram used in determining the optimum course of action, in situations having several possible alternatives with uncertain outcomes. A decision tree classifier is modeled in two phases: tree building and tree pruning. In tree building, the decision tree model is built by recursively splitting the training data set and assigning a class label to leaf by the most frequent class. Pruning a sub tree with branches if lower training error is obtained. Figure 2.1 presents decision tree algorithm [18].



**Input:**

- Data partition, $D$, which is a set of training tuples and their associated class labels;
- *attribute_list*, the set of candidate attributes;
- *Attribute_selection_method*, a procedure to determine the splitting criterion that "best" partitions the data tuples into individual classes. This criterion consists of a *splitting_attribute* and, possibly, either a *split point* or *splitting subset*.

**Output:** A decision tree.

**Method:**

```
create a node $N$;
if tuples in $D$ are all of the same class, $C$ then
      return $N$ as a leaf node labeled with the class $C$;
if attribute_list is empty then
      return $N$ as a leaf node labeled with the majority class in $D$; // majority voting
apply Attribute_selection_method($D$, attribute_list) to find the "best" splitting_criterion;
label node $N$ with splitting_criterion;
if splitting_attribute is discrete-valued and
      multiway splits allowed then // not restricted to binary trees
      attribute_list $\leftarrow$ attribute_list $-$ splitting_attribute; // remove splitting_attribute
for each outcome $j$ of splitting_criterion
// partition the tuples and grow subtrees for each partition
      let $D_j$ be the set of data tuples in $D$ satisfying outcome $j$; // a partition
      if $D_j$ is empty then
              attach a leaf labeled with the majority class in $D$ to node $N$;
      else attach the node returned by Generate_decision_tree($D_j$, attribute_list) to node $N$;
endfor
return $N$;
```

***Figure 2.3: Basic structure of Decision Tree algorithm (ID3 algorithm)[18]***

### 2.6.1.3 K- Nearest Neighbor (KNN):

The k-nearest neighbor algorithm (k-NN) is a non- parametric method for classifying objects based on the closest training examples in the feature space. K-nearest neighbor is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification. The k-nearest neighbor algorithm is amongst the simplest of all machine learning algorithms: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors (where k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of its nearest or closeness neighbor [18].

"Closeness" is defined in terms of Euclidean distance, where the Euclidean distance, where the Euclidean distance between two points, X=(x1,x2,……,xn) and Y=(y1,y2,….,yn) is:

$$d(X, Y)=\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad \longrightarrow \quad (6)$$

Nearest neighbor classifiers can also be used for prediction, that is, to return a real-valued prediction for a given unknown sample. In this case, the classifier returns the average value of the real-valued associated with the k nearest neighbors of the unknown sample. The k-nearest neighbors' algorithm is amongst the simplest of all machine learning algorithms. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k-nearest neighbors. K is a positive integer, typically small. If k = 1, then the object is simply assigned to the class of its nearest neighbor. In binary (two class) classification problems, it is helpful to choose k to be an odd number as this avoids tied votes [22].

The same method can be used for regression, by simply assigning the property value for the object to be the average of the values of its k nearest neighbors. It can be useful to weigh the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones [22].

## 2.7 Conclusion:

Phishing and pharming attacks are social engineering techniques used to take advantage on human ignorance. The term Phishing appeared in early 1996, but it was not until the latest of 2003 that email based phishing attacks began to become a popular attack vector for cyber criminals as a means to conduct financial fraud and identity theft. Pharmer works by changing DNS server information to redirect the web request to a fake site. In this work we design a defeating model to detect pharming attacks on client side.

Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. These tools can include statistical models, mathematical algorithm and machine learning methods. Data mining tasks can be: classification, prediction, association, clustering, and outlier. In this work we used classification task from these data mining tasks in our defeating model.

# Chapter (3): Related works

## 3.1 Phishing attack:

In 2010, P. sengar and V. kumar [2], presented a solution for phishing attacks by classification of web pages. They proposed "PageSafe" which is an anti-phishing tool that prevents access to phishing web pages and also detects DNS poisoning attacks. PageSafe tool uses a machine learning approach (Neural Network) for automatic classification. Authors also build pharming detection module to detect pharming attacks. Pharming detection module compares the IP addresses from Local DNS and Network DNS with the IP address from remote DNS for the requested URL. If they do not match, then pharming is detected and alert is made to the user. Authors introduced the PageSafe -a novel tool that does not completely rely on automation to detect phishing. Instead, PageSafe relies on user input to decide on the legitimacy of a URL. So users may not be comfortable with that, and they may not rely well on those tools to protect themselves well.

In 2011, A. Martin et al. [4], presented a framework for predicting phishing web sites using Neural Networks. A neural network is a multilayer system which reduces the error and increases the performance. This paper describes a framework to better classify and predict the phishing sites using neural networks. Authors take E-Banking phishing sites as a case study, and they take 27 features to make their classification. Authors use neural network to predict phishing websites using some criteria and indicators like URL and domain identity, but some of that isn't useful in the case of pharming attacks.

In 2010, M. Aburrous et al. [7], presented a new approach to passing the difficulty and complexity in detecting phishing websites especially e- banking website. This approach includes an effective model based on using association and classification data mining algorithms. These algorithms were used to characterize and identify all the factors and rules in order to classify the phishing website. Authors implemented six different classification algorithms and techniques to extract the phishing training data sets criteria to classify their legitimacy. A Phishing Case study was applied to illustrate the website phishing process. The rules generated from the associative classification model and showed the relationship between some important characteristics like URL and Domain Identity, and Security and Encryption criteria in the final phishing detection rate.

In 2010, M. Aburrous et al. [25], presented a new approach to overcome the difficulty and complexity in detecting and predicting e-banking phishing websites. They proposed an intelligent and effective model that is based on using association and classification Data Mining algorithms. These algorithms were used to characterize and identify all the factors and rules in order to classify the phishing websites. They implemented six different classification algorithm and

techniques to extract the phishing training data sets criteria to classify their legitimacy, and then they compared their performances, accuracy, number of rules generated and speed. The authors experimented their model and they found that the results had better performance as compared to other traditional classifications algorithms. Authors used types of classifiers for the prediction of phishing websites, the error rate of was 12.622%.
In 2008, Y. Cao et al. [13], presented an anti-phishing approach named Automated Individual White-List (AIWL). This approach automatically maintained a white list of the user’s all familiar Login User Interfaces (LUIs) of web sites. When a user tries to submit the confidential information to an LUI that is not in the white-list, AIWL will alert the user to the possible attack. Then, AIWL can efficiently defend against pharming attacks, since AIWL will alert the user when the legitimate IP is changed. For the legitimate IP addresses, authors used Naïve Bayesian classifier to automatically maintain the white-list in AIWL.
In this paper, authors depended on individual white list to store familiar login user interface, but that white list may become corrupt and infected with any type of attacks, or may get lost for any reason.

In 2012, R.Sumathi, R.Vidhya Prakash [14], presented a new approach to overcome the difficulty and complexity in predicting and detecting phishing websites. In this proposed system the authors implemented the PSO (Particle Swarm Optimization) algorithm for predicting Phishing Websites. In this work, they presented a novel approach to overcome the ‘fuzziness’ in the phishing website estimate and proposed an intelligent flexible and effective model for phishing websites. The experimental results demonstrated the feasibility of using Association and Classification techniques and Particle Swarm Optimization (PSO) real applications and its better performance. Authors used two types of classifiers to compare between them to prove that their classifier is the best. However they should use much more classifiers to prove that their theory is the best for detection of phishing websites.

In 2011, M. Alkhozae, O. Batarfi [15], proposed a phishing detection approach based on checking the webpage source code. They extracted some phishing characteristics out of the W3C standards to evaluate the security of the websites, and check each character in the webpage source code if they find a phishing character, and they decreased weight from the initial secure weight. Finally they calculated the security percentage based on the final weight, the high percentage indicates secure website and others indicate the website is most likely to be a phishing website. They checked two webpage source codes for legitimate and phishing websites and compared the security percentages between them. They found that the phishing website has a less security percentage than the legitimate website. Their approach can detect the phishing website based on checking phishing characteristics in the webpage source code.

In 2008, M. Aburrous, M.A. Hossain [16], presented an approach, which is an intelligent Phishing Website Detection System using Fuzzy Techniques. It is based on fuzzy logic and produces six criteria of website phishing attack. There are many characteristics and factors that can distinguish the original legitimate website from the forged faked phishing website like spelling errors, long URL address and abnormal DNS record. Website phishing detection rate is performed based on six criteria and there are different numbers of components for each criterion. The criteria are: URL & Domain Identity, Security & Encryption, Source Code & Java script, Page Style & Contents, Web Address Bar, and Social Human Factor. Authors do not state any analysis about their results.

In 2009, S. Afroz and R. Greenstadt [19], proposed a phishing detection approach, called "PhishZoo", which used profiles of trusted web sites' appearances built with fuzzy hashing techniques to detect phishing. Authors evaluated their approach on 636 phishing web sites simulating 20 real web sites and showed that it provides similar accuracy to blacklist approaches. Their approach had advantage that it can classify new attacks and targeted attacks against smaller web sites. Authors made the assumption that SSL is supported by these web sites of interest and secure in both the underlying protocol and the trust model used by the browser. But SSL sometimes is not supported by the website.

In 2007, T. Moore and R. Clayton [26], estimated the number of phishing victims by examining web server logs. They estimated that 311,449 people fall for phishing scams annually, costing around 350 million dollars. Authors did not present a formal model of the costs and benefits of phishing attacks at this stage of their work.

## 3.2 Pharming attack:

In 2007, S. Stamm et al. [1], described the concept of the attack: "Drive by Pharming". In this attack an attacker sets up a web page so that when the victim user views, in the case of javascript enabled browser, the attacker changes the DNS server settings on victim's home broadband router .The authors explain and describe scenarios for that type of attack, and they talk about new attacks like, pharming, growing zombies, and viral spread. So, researchers recommended for any user to change the default password of his own router, and to disable javascript in the browser to avoid these types of attacks. Authors do not introduce an effective solution to protect users from phishing and pharming attacks, they only explain some new attacks, and introduce some advice for users to protect themselves from those attacks.

In 2010, B. Aslam et al.[3], presented a solution to protect users from phishing and pharming attacks. This solution is based on a hashed password which is the hash value of the user-typed password and the authentication server's IP address. The solution rests on the fact that the server connected by a client using TCP connection can't fraud about its IP address. If a user goes to a malicious server (by a Phishing or a Pharming attack), the password obtained by the malicious server will be the hashed password (tied to the malicious server's IP address) and will not be usable by the attacker at the real server, thus defeating Phishing or Pharming attack.  Authors used PwdIP-Hash for specific browser, but they do not useful for other famous web browsers such as Firefox, Chrome, etc.

In 2011, S. Prevost et al. [5], proposed a dual approach to detect pharming attack at the client side. This approach combines the IP address check and the webpage content analysis, using information provided by multiple DNS servers (Local DNS server and Alternate DNS server like GoogleDNS or OpenDNS). The approach is integrated within web browser of the user. Authors validated their proposed approach by conducting a first set of experimentations from continents (North America, South America, Europe, Africa, Asia and Australia). The same third party DNS server was asked to resolve many homepages of legitimate domain names in order to check the IP addresses changes. Authors used webpage content analysis as a second approach of their solution, but the main drawbacks of this solution are to maintain an up-to-date database as well as to protect it against any compromising attacks.

In 2011, S. Prevost et al. [6], defined an advanced approach to strengthen their first approach in [5] to alert the end-user in case of pharming attacks at the client-side. They had a success rate over 95%, and they validated their solution that helps users to differentiate legitimate from fraudulent login websites, based on a dual-step analysis (IP address check and webpage content comparison) performed using multiple DNS servers information. Authors used the same webpage content analysis in their first paper in [5], as a second approach of their solution, but still the main drawbacks of this solution are to maintain an up-to-date database as well as to protect it against any compromising attacks

In 2006, O. Mahmood [12], proposed a method to identify, warn and protect users from the attacks: phishing, pharming, and man-in-the-middle attacks, by authenticating the site before the user actually shares personal information. The presented method is based on the use of a browser plug in which enables the user to validate the website and provides visual feedback. The plugin also automatically visually notifies the user of possible attack in case of pharming, where Domain Name System (DNS) is infected. The presented method is divided into three processes: IP address check, SSL certificate validation and verification, and Friend of a Friend evaluation process. Author depend on validating and verifing the SSL certificate by using locally stored information, but that local information may be infected or corrupted on any type of attack.

### 3.3 Conclusion:

In this chapter we reviewed a related work about our research on phishing and pharming attacks. Most of the authors [2][4][7][13][14][16][19][25][26], focused on phishing attack and the techniques and mechanisms employed to detect and avoid fraud for this type of attack (phishing attack).Some of them[1][3][5][6][12], talk about pharming attack; the Sophisticated version of phishing attack. Some authors [4][16],  used a classification techniques to detect phishing attack. Some [4], used neural network to predict phishing website, while others [3] used hashed password and others [15] used checking the webpage source code to detect phishing attack.

Most of these papers do not talk about a success rate for their approaches or models, and they may not have applied their work, except on papers [5][6], which estimated their accuracy to about 95%, which is a low percentage for the sensitivity of this attack, and we could achieve better results from that .

# Chapter (4) Methodology and Implementation

In this chapter, we present and explain the proposed model (Detection Model for Pharming Attack Based on IP-Address Check and Website Predictability) and the methodology which we followed in this research.

## 4.1 Overall Methodology

The solution to defeat pharming attack is in two steps, first IP address check, and the second is classification of the websites. So the methodology to achieve the objectives consists of the following which is shown in Figure 4.1:

1. IP address check step, by designing a java program using NetBeans program that can check the IP addresses. A DNS request is sent to two DNS servers. The first is to compare default IP address with IP addresses returned from third party DNS (public DNS like, OpenDNS, GoogleDNS, HadaraDNS). If the default IP address is included in the IP addresses for third party DNS, the site is considered legitimate, otherwise the site is regarded as suspicious.

2. Classification of the web page in which the following is done :

    a. Searching for an appropriate dataset which must contain phishing websites addresses and other features of that site so we can classify those pages as safe or as phishing, However we could not find those features in the dataset we downloaded from phishtank [28] website, and from APWG-Anti-Phishing working group [27].

    b. Collecting the features of the phishing sites and on the other hand we collected the same features from saved sites which proved to be safe.

    c. After we had a complete data set we tried to find the best classifier, by using Rapid Miner program.

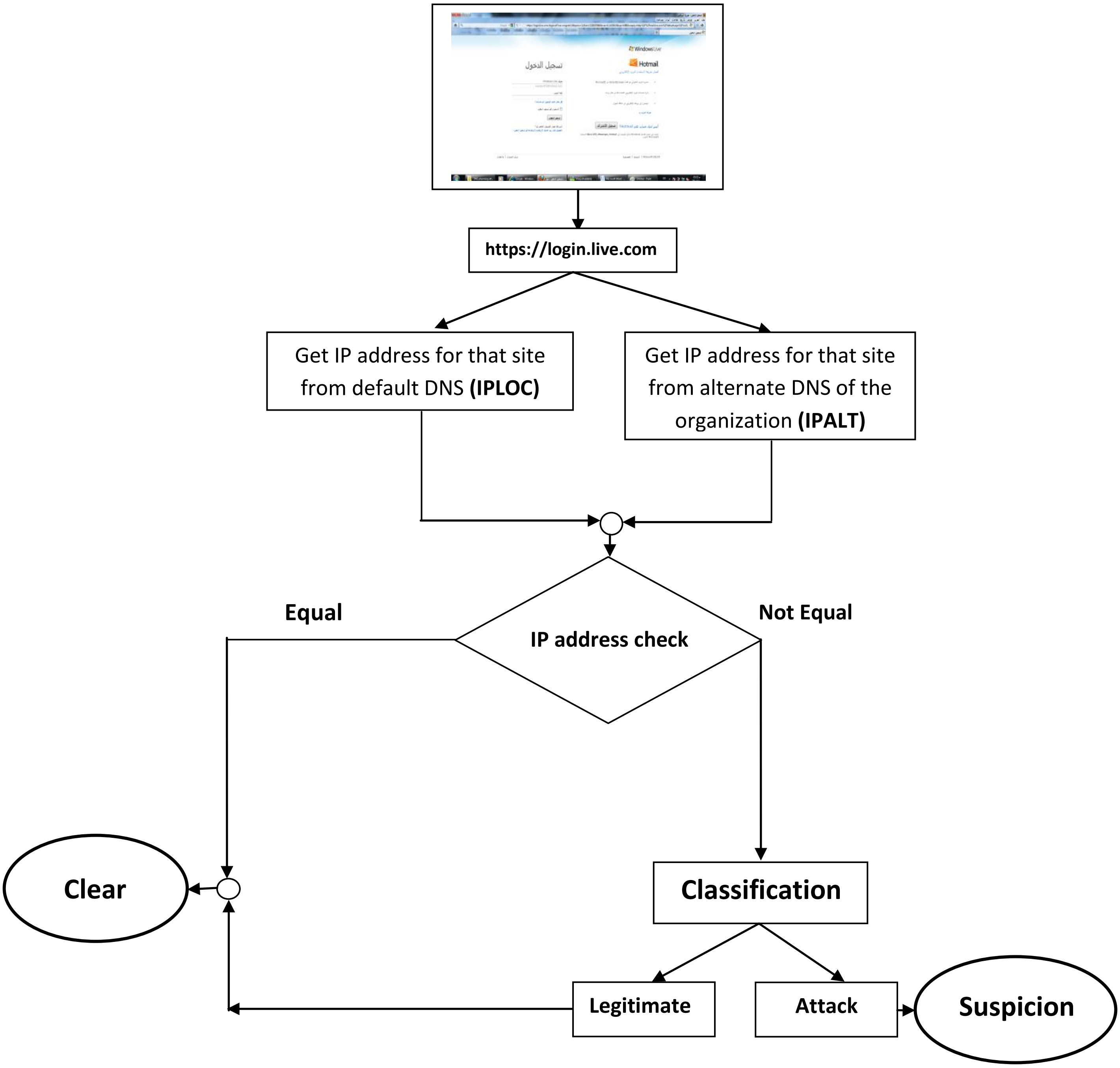

https://login.live.com
Get IP address for that site from default DNS (IPLOC)
Get IP address for that site from alternate DNS of the organization (IPALT)
Equal
IP address check
Not Equal
Clear
Classification
Legitimate
Attack
Suspicion

*Figure 4.1: The proposed model*

## 4.2 First step: IP address check:

Each time the web browser accesses a URL, the domain name of the visited website is checked out. Then, a DNS request is sent to two DNS servers. The comparison of the default IP address with IP addresses returned from third party DNS (public DNS like, OpenDNS, and GoogleDNS). If the default IP address is included in the IP addresses for third party DNS, the site is considered as legitimate, otherwise the site is suspicious.

To perform the previous step, we designed the program using NetBeans program. This program checked the IP address from local DNS, and from GoogleDNS. As a third party DNS, we used GoogleDNS because it is the most famous and effective. The flowchart of this program is shown in Figure 4.2.

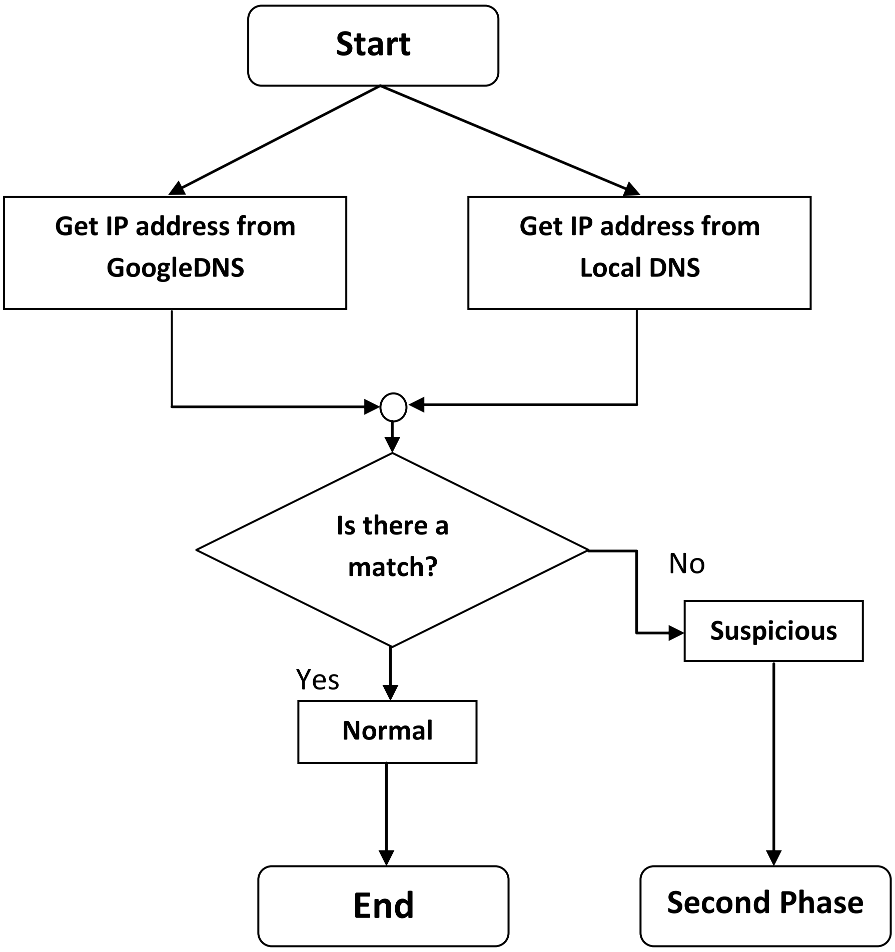
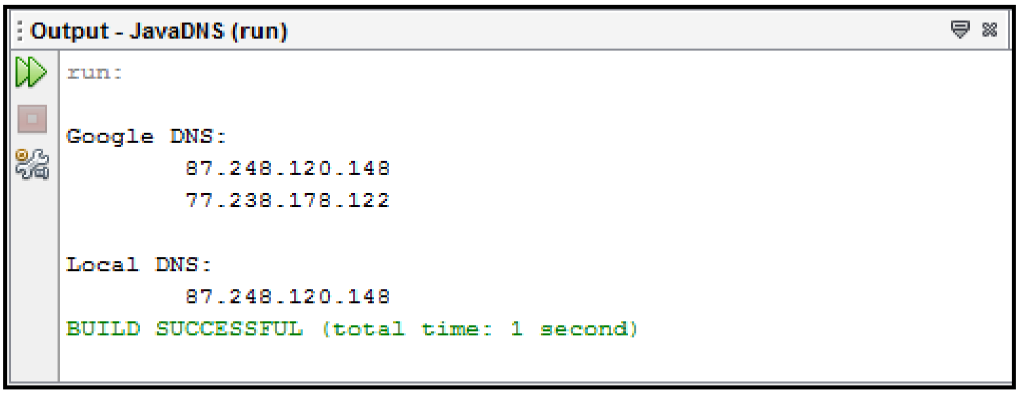


*Figure 4.2: Program Flowchart*

The action result of that program is showed in Figure 4.3. If the website is clean from any pharming attack, there will be a match between two IP addresses.



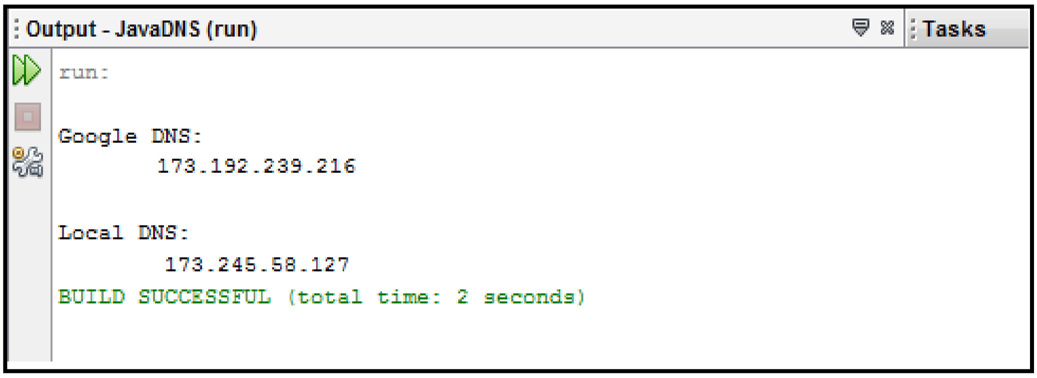
```
run:

Google DNS:
        87.248.120.148
        77.238.178.122

Local DNS:
        87.248.120.148
BUILD SUCCESSFUL (total time: 1 second)
```

*Figure 4.3: Result of IP address check of clean site*

Otherwise, if there is no match between two returned IP addresses, the site will be considered suspicious like Figure 4.4.



```
run:

Google DNS:
       173.192.239.216

Local DNS:
       173.245.58.127
BUILD SUCCESSFUL (total time: 2 seconds)
```

*Figure 4.4: Result of IP address check of suspicious site*

## 4.3 Second step: classification of websites:

In the second step the classifiers on the dataset (K-NNs, decision tree, Naïve Bayesian, etc), will be used to classify the pharming websites. So for the implementation we shall use two publicly available datasets to train and test the classifiers: the "phishtank" from the phishtank.com [28]. The Phishtank database records the URL for the suspected website that has been reported, the time of that report, and sometimes further detail such as the screenshots of the website, which is publicly available.

The second source of the dataset is the Anti Phishing Working Group (APWG) [27] which makes "Phishing Archive" describing phishing attacks. We setup an account with the site administrator; which was between the site and the Islamic University Gaza. Note that if necessary these datasets or some other attributes from these datasets will be collected.

Note that the pharming websites are temporary sites, which means that the pharming site may stay online from few hours to few days, so we periodically download the dataset when necessary.

Collecting the features of the pharming sites by designing a program using Visual Studio program which collects the features from any site by putting the URL of that site in the bar panel in execution of that program shown in Figure 4.4.

Our program for collecting our features is called "Pharming Detector". That program can find eight features from any website from which we can determine if that site is an attack or not.

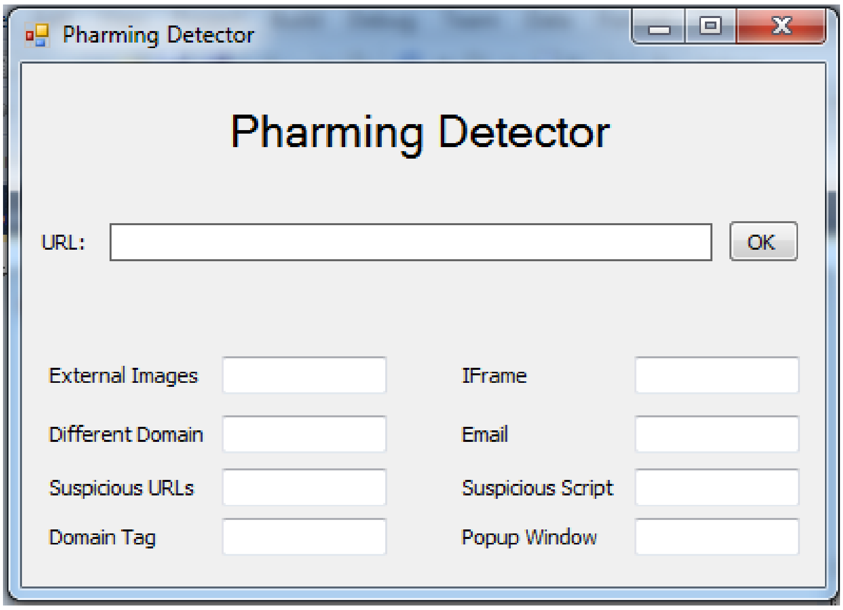


*Figure 4.4: Pharming Detector Program Interface*

Those features or characteristics [15] can check the webpage source code file, and the extraction of these characteristics out of W3C standards to evaluate the website security.

### 4.3.1 Pharming Attacks Characteristics:

Pharmers use some tricks and traps to fool and tempt users, so our method is to check for these tricks and factors in the webpage source code and calculate the security levels based on these factors to classify the webpage if it is secure or not [15]:

We can summarize these factors and their risk level, in table 4.1.

***Table 4.1: pharming characteristics classification [15]***

| **Pharming characteristics** | **Brief explanation** | **Pharming characteristics risk** |
|---|---|---|
| Https | Secured protocol which is used to tell us if this website is secured | Medium |
| Images | Images in the website should load from the same URL in the website | Low |
| Suspicious URLs | Phishers may use symbol @ to hide their host names | High |
| Domain | Phisher use external domains | Medium |
| Email | It's a function on PHP called mail or email which takes the information we enter in the forms like Password, etc. | High |
| Iframe | HTML tag code which is used to embed another webpage into the current webpage | Low |
| Script | PHP files, some phisher use scripts to send personal information to them | High |
| Popup window | Phishers use popup windows to get personal information | Low |

**1) Https**: It is the secured protocol which is used to tell us that this website is secure but it should be in "URL" of the website not in the body source of the webpage because pharmers use Https inside their source code file to tell us that these images or links are secured which is not true. The normal page should be like this <img src="example.png" />. But there are some phishers who use the SSL certificate in the source code like this <img src=https:// www.xxx.com/example.png/>. Phishers use https to make us think they are secured website which is in fact false. Many similar phishing attacks in which phishing websites use a certificate that can be expected to trigger a browser warning [15].

**2) Images:** All images in the website including website logo should load from the same URL of the website not from another website, so all links should be internal links not external. Therefore, we must check the links to detect any external links inside the source code like this : <img src=https: //www.xxx.com/logo.jpg>" it is a pharming character [15].

**3) Suspicious URLs:** Most of the pharmers use an IP address instead of using the actual domain name. Other pharmers may use symbol @ to make their host names unclear [15].

**4) Domain:** It is the external domains mean: if a user logs onto a website under the name: www.example1.com and we find some URLS of links in the source code like this "www.example-1.com" which is not the source URL. This means that this website is trying to steal our information. Pharmer use forward domain also called domain redirection; it is a technique on the World Wide Web for making a webpage available under many URLs [15].

**5) Email:** There is a function on PHP called mail or email and it takes the information we enter in the forms like "Password, etc. "and sends it when we press the pay button through e-mail to the pharmers e-mail. Pharmer can insert PHP code inside Html code and use this function to send our information [15].

**6) iframe:** It is HTML tag code and is used to embed another webpage into a current webpage. It creates a frame or window on a webpage so that another page can load inside this frame. Pharmers use the iframe and make it invisible i.e. without frame borders. When the user goes to the website, he/she is unaware that there is another page also loading in the iframe window. It is a big problem which is unknown to most people. It is like a small website that opens in the current webpage for example: we can open www.google.com in the page www.example.com by using iframe so when people enter our website they will see the secured website opened, but it is not the same page that opens in the iframe.

Example: http://www.Pharmer.com/index.php?search=""> <iframe src= http://google.com ></iframe>, Replace http://google.com by the phishing page [15].

**7) Script:** It is PHP files; some pharmers use scripts to send personal information to them, and some scripts send viruses or load from external websites. Scripts tag use to put any external file in the page like jquery or CSS and if it is with start and end tag, it is legal because this is the correct and standard script tag. Example: <script type="text/javascript" src="includes/jscripts/jquery.min.js"></script>, it is now a load file which makes the page's appearance good. When there are tags like this <script> and codes between them (not links) they are suspicious tags because this script code is javascript or any other languages which may be used to send personal information or PC information to pharmers. So if we find <script> tags and their end tags </script> they are legal tags, otherwise it is a pharming character [15].

**8) Popup window:** Pharmers use popup windows to steal personal information. Often, these popups may ask to update, validate or confirm account information and it is like official organizations websites. If the user enters his information in the popup windows, the pharmer then steals this private information. There are two kinds of popup windows. One of them is used to confirm or to tell us something and this window has a special way to fill it in html or JavaScript like: < onClick="window.open('example.html')"> and it is by html and legal window . The other is an illegal popup window because it is a javaScript file used in like: "Open Popup" onClick="javascript:popUp('example.html')"> . It is illegal because it opens a new page from another website such as registration page or asks the user information. It is a new full page and it opens automatically when the user opens a page or clicks on any link so it is an illegal popup window [15].

After collecting the pharming characteristics, we record these data on the dataset that we get from phishtank [28], and from APWG-Anti-Phishing working group [27] websites, so now we have a full dataset with all records of pharming attacks, to these records we must add clear records without attack, so we either record websites we are sure are not pharming websites (we can check the safety of these websites from phishtank websites [28]). After that we check these sites by using our program (Pharming Detector), and record the results in our dataset, in order to have a full dataset with pharming and clean dataset. Our dataset contains 1503 records, 400 records (26.6%) pharming, and the remaining (1103, 73.4%) are clean.

## 4.4 Implementation and Evaluation:

This Model is dedicated to people who deal with sensitive data and accounts such as personal and financial accounts Therefore, the accuracy is the main goal to be achieved by this solution, not efficiency or performance. That's why the evaluation of this work will depend on measuring the accuracy.

### 4.4.1 Data Preprocessing and Feature Selection:

We download a dataset from phishtank website [28] and from APWG-Anti-Phishing working group [27]; these datasets contain the following attributes:

**Phish_id:** The ID number by which Phishtank [28] refers to a phish submission. All data in PhishTank is tied to this ID. This will always be a positive integer.

**URL:** The URL address for that website.

**Phish_detail_url**: This is the phishtank URL which contains the details of that phishing websites.

**Submission time:** The date and time at which this phish was reported to Phishtank. This is an ISO 8601 formatted date.

**Verified:** Whether or not this phish has been verified by our community. In these data files, this will always be the string 'yes' since we only supply verified phishes in these files.

**Verification time**: The date and time at which the phish was verified as valid by our community. This is an ISO 8601 formatted date.

**Online**: Whether or not the phish is online and operational. In these data files, this will always be the string 'yes' since we only supply online phishes in these files.

**Target:** The name of the company or brand the phish is impersonating, if it's known.

In addition to these attributes we added the previous detailed attributes (Https, Images, Suspicious URLs, Domain, Email, Iframe, Script, and Popup window).

The full attributes and their importance to our classification methods are listed in table 4.2.

*Table 4.2: Full attributes of our dataset*

| Attributes | Important? |
|---|---|
| id | No |
| URL | No |
| detail_url | No |
| Submission time | No |
| Verified | No |
| Verification time | No |
| Online | No |
| Target | Yes |
| Https | Yes |
| Images | Yes |
| Suspicious URLs | Yes |
| Domain | Yes |
| Email | Yes |
| Iframe | Yes |
| Script | Yes |
| Popup window | Yes |

As preprocessing and feature selection steps, we did an exception of ID, detail_url, Submission time, Verified, Verification time, and Online, attributes because it's not useful in our classification methods, we take target attribute, so we can make classifier by domain type, and we put "1" in pharming rows, and "0" in clear rows, on the pharming attributes.

### 4.4.2 Apply the classifier algorithms:

This section describes the types of classifiers algorithms which we tested in our model: Naïve Bayes (NB), Decision Tree (DT), and KNN (K Nearest Neighbor), which are provided by RapidMiner [29] program.

### 4.4.2.1 Decision Tree algorithm:

We apply decision tree algorithm on our dataset, then we use x-validation, X-Validation encapsulates a cross-validation in order to estimate the performance of a learning operator. We used it as shown in Figure 4.5.

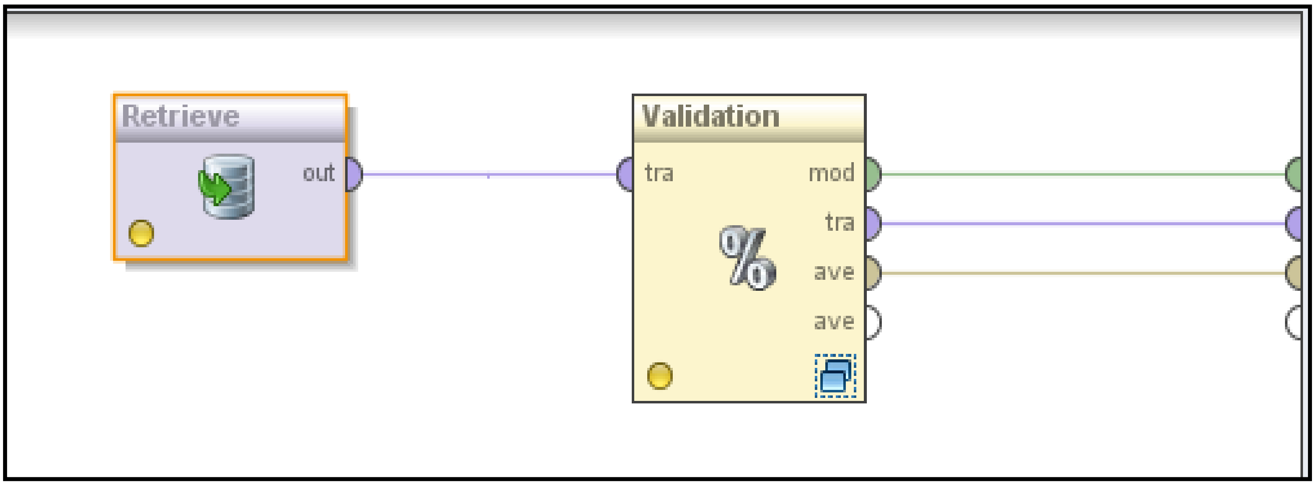


*Figure 4.5: The X-validation method*

Inside the cross validation method, we can put our algorithm of classification to classify if the record is an attack or not. In this case we used decision tree method as shown in Figure 4.6.

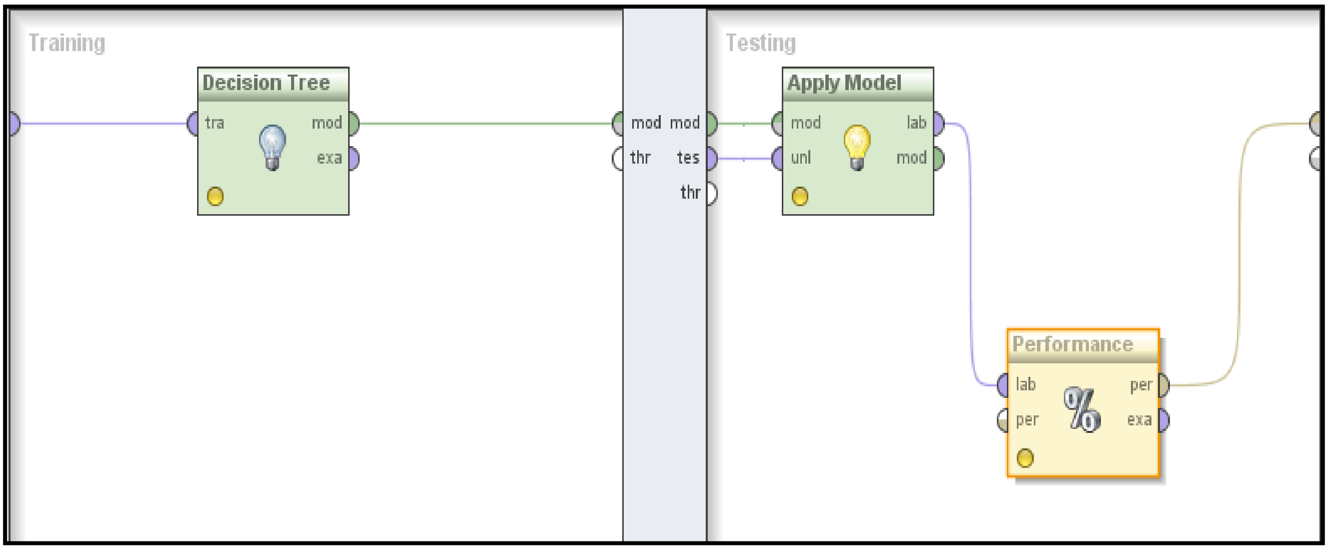
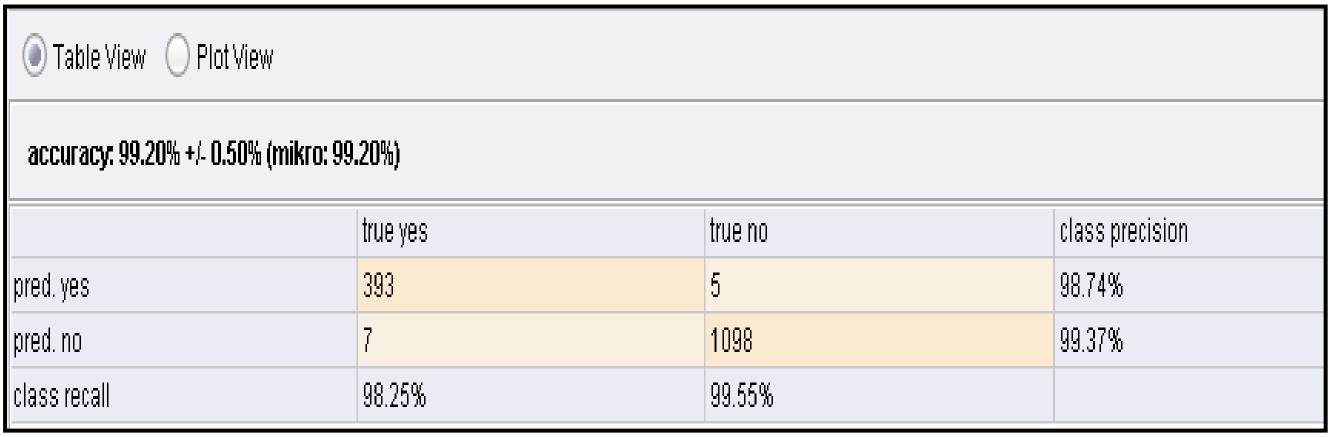


*Figure 4.6: Applying decision tree algorithm*

The accuracy result of applying decision tree algorithm was 99.2%, so the true yes was 98.25%, true no was 99.5%, these results are shown in Figure 4.7.



● Table View ○ Plot View

accuracy: 99.20% +/- 0.50% (mikro: 99.20%)

| | true yes | true no | class precision |
|---|---|---|---|
| pred. yes | 393 | 5 | 98.74% |
| pred. no | 7 | 1098 | 99.37% |
| class recall | 98.25% | 99.55% | |

*Figure 4.7: Result of decision tree algorithm*

### 4.4.2.2 KNN algorithm:

We applied KNN algorithm on our dataset, then we used x-validation. X-Validation encapsulates a cross-validation in order to estimate the performance of a learning operator, as shown in Figure 4.5.

Inside the cross validation method, we can put our algorithm of classification, to classify if the record is an attack or not. In this case we use KNN method with k=1, as shown in Figure 4.8.

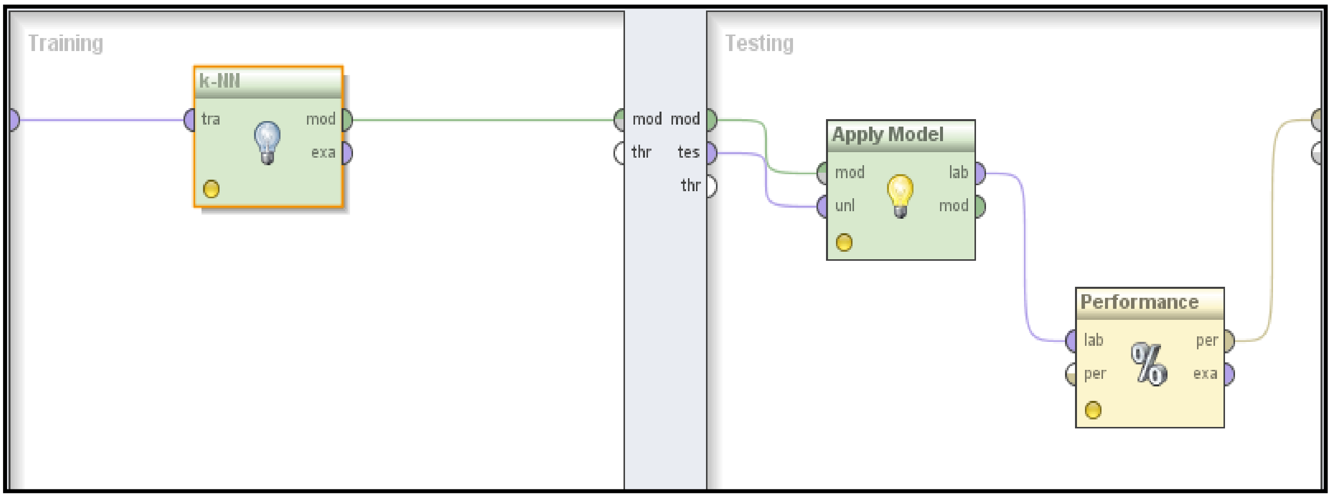
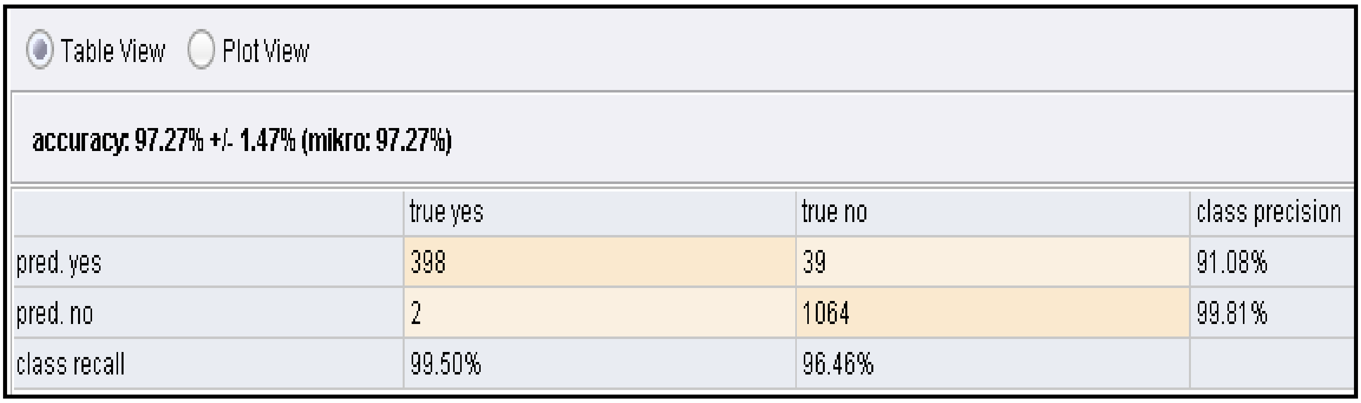


*Figure 4.8: Applying KNN algorithm*

The accuracy result of applying KNN algorithm was 97.27%, so the true yes was 99.5%, true no was 96.4%. Results are shown in Figure 4.9.



● Table View ○ Plot View

accuracy: 97.27% +/- 1.47% (mikro: 97.27%)

| | true yes | true no | class precision |
|---|---|---|---|
| pred. yes | 398 | 39 | 91.08% |
| pred. no | 2 | 1064 | 99.81% |
| class recall | 99.50% | 96.46% | |

*Figure 4.9: Result of KNN algorithm*

### 4.4.2.3 Naïve Bayes algorithm:

We applied Naïve Bayes algorithm on our dataset, then we used x-validation. X-Validation encapsulates a cross-validation in order to estimate the performance of a learning operator, as shown in Figure 4.5.

Inside the cross validation method, we can put our algorithm of classification. To classify if the record is an attack or not, in this case we use Naïve Bayes method as in Figure 4.10.

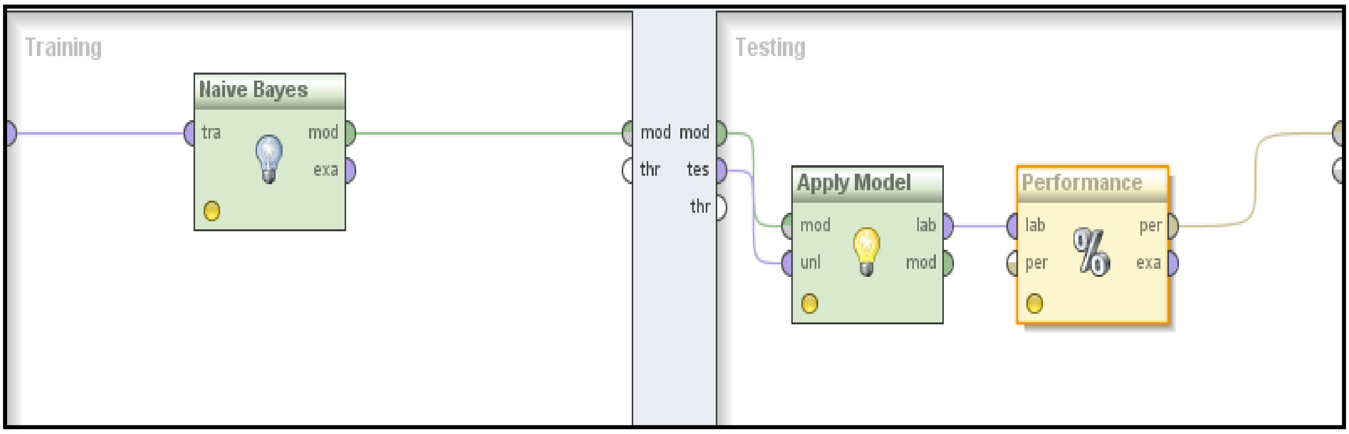
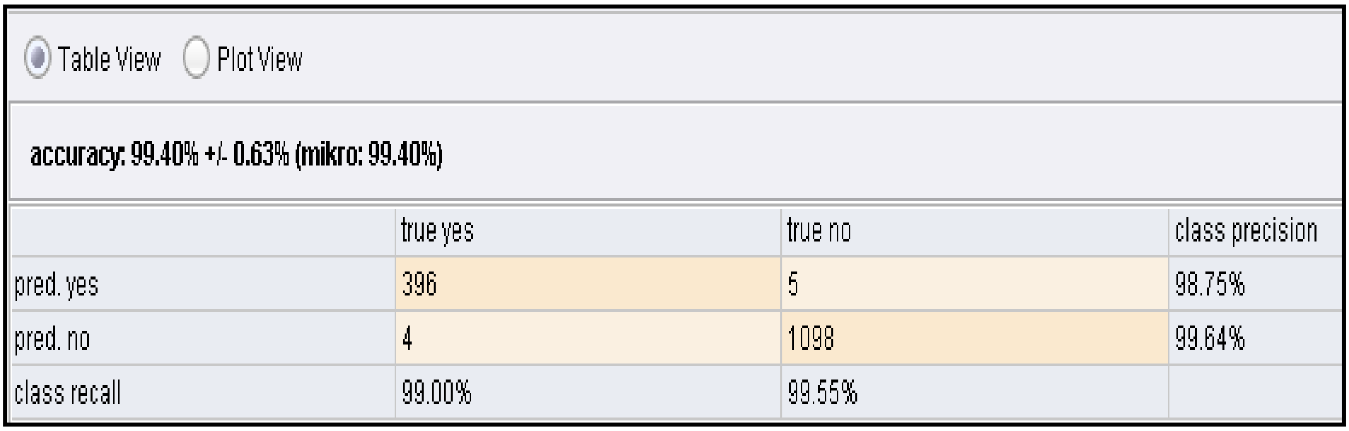


*Figure 4.10: Applying Naïve Bayes algorithm*

The accuracy result of applying Naïve Bayes algorithm was 99.4%. So the true yes was 99%, true no was 99.5%. Results are shown in Figure 4.11.



● Table View ○ Plot View

accuracy: 99.40% +/- 0.63% (mikro: 99.40%)

| | true yes | true no | class precision |
|---|---|---|---|
| pred. yes | 396 | 5 | 98.75% |
| pred. no | 4 | 1098 | 99.64% |
| class recall | 99.00% | 99.55% | |

*Figure 4.11: Result of Naïve Bayes algorithm*

So, from these three data mining classifiers, we can conclude that the Naïve Bayes algorithm is the best one for our case of classification of phishing websites, because the accuracy is the best (99.4%).

### 4.4.3 Evaluating the classification methods:

Performance evaluation of the classification methods is one of the most important tasks in the research. For the purpose of evaluating the results, we used confusion matrices in which the commonly evaluation measures were visualization tools used in supervised learning, and created for each classifier. Each column of the confusion matrix represents the instances in a predicted class, while each row represents the instances in an actual class. The following four which define the members of the matrix are: the True Positive rate (TP), False Positive rate (FP), True Negative rate (TN), False Negative rate (FN). Also, accuracy is considered the most important to evaluate classification performance. In our research, there are other measures used to evaluate classifiers performance, which are: recall, precision, and overall accuracy, which can be defined as follows:

**Confusion matrices**: In the classification problem, the primary source of performance measurement is confusion matrix. The confusion matrix is a useful tool for analyzing how well your classifier can recognize data of different classes [19].

In our work it's created for each classifier using the actual and predicted responses [18] [19]. The following four estimates define the members of the matrix (as showed in Table 4.3:

*Table 4.3: Confusion matrix table [18][19]*

| | | True Class | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Predicted Class** | **Positive** | True Positive<br>(TP) | False Positive<br>(FP) |
| | **Negative** | False Negative<br>(FN) | True Negative<br>(TN) |

**True positive (TP):** refers to positive instances that correctly label the classifier.

$$\text{True Positive rate} = TP/(TP+FN)$$

**True negatives (TN):** refers to negative instances that correctly label the classifier.

$$\text{True Negative rate} = TN/(TN+FP)$$

**False Positive (FP):** are the negative instances that were incorrect.

$$\text{False Positive rate} = FP/(TN+FP)$$

**False Negative (FN):** are the positive instances that were incorrect.

$$\text{False Negative rate} = FN/(TP+FN)$$

**Accuracy:** we can calculate the accuracy which refers to the percentage of test set rows that are correctly classified by the classifier [19].

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

We can calculate either:

**Recall:** refers to number of positive instances that correctly label the classifier

$$\text{Recall} = TP/(TP + FN)$$

**Precision:** refers to the percentage of retrieved instances that are relevant

$$\text{Precision} = TP/(TP + FP)$$

**F-measure:** defines as the harmonic mean of precision and recall [26].

$$F = 2*(\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

We can determine the best classifier from those three classification methods (Decision tree, Naïve bayes, and KNN), by using the previous definitions and calculations.

## 4.5 Conclusion:

Our solution to defeat pharming attack is in two steps; first IP address check, and the second is classification of the websites. First step: IP address check step, performed by a java program using NetBeans program that can check the IP addresses. A DNS request is sent to two DNS servers. The first is to compare default IP address with IP addresses returned from third party DNS (public DNS, GoogleDNS). If the default IP address is included in the IP addresses for third party DNS, the site is considered legitimate, otherwise the site is regarded suspicious. The second step classification of the web page in which we found an appropriate dataset which must contain a pharming website address and other features of that site so we could classify that page as safe or as pharming page. However we could not find those features in the dataset we downloaded from phishtank [28] website, and from APWG-Anti-Phishing working group [27], so we collected the features of the phishing sites and on the other hand we collected the same features from other sites which proved to be safe. After we had a complete data set we tried to find the best classifier. We could conclude that the Naïve Bayes algorithm is the best one for our case of classification of phishing websites because it has the best accuracy (99.4%).

# Chapter (5): Experimental Results and Evaluation:

In this chapter we presented and analyzed the experiments results. We explained the machine environment and tools used in our research. Also we presented the evaluation measurements for classifications model during sets of experiments by using the equation of accuracy, recall, precision, and f-measure which are illustrated in section 4.5, and finally we extracted the overall accuracy of our model.

## 5.1 Experiments Setup

In this section, we describe the experimental environment and tools used in experiments, measures of performance evaluation of our model.

### 5.1.1 Experimental Environment and Tools

We had ran the experiments on a machine with properties that are Intel Pentium Core i3 M330 @ 2.13 GHz processor and 2.00 GB of RAM. To carry out our work (including the experimentation), special tools and programs were used:

- **RapidMiner 5**: It is an application program, used to choose the best classifier from the three choosing classifiers. We experimented our dataset on the RapiMiner program and extracted the required results.

- **Microsoft Visual Studio 2010:** Used to design a program that can collect the characteristics or features of the web pages which we can use to classify the page on phishing or clean page.

- **NetBeans IDE 6.8:** Used to design a program that can compare two IP addresses from local DNS server and from alternate DNS server (GoogleDNS, OpenDNS). We used this program to verify that the web page is clear from pharming attack as a first phase of our model.

- **Microsoft Excel 2007:** Used Excel to partition, organize and store datasets in tables, and to do some simple preprocessing and analyze the results.

## 5.2 Measurements for Experiments:

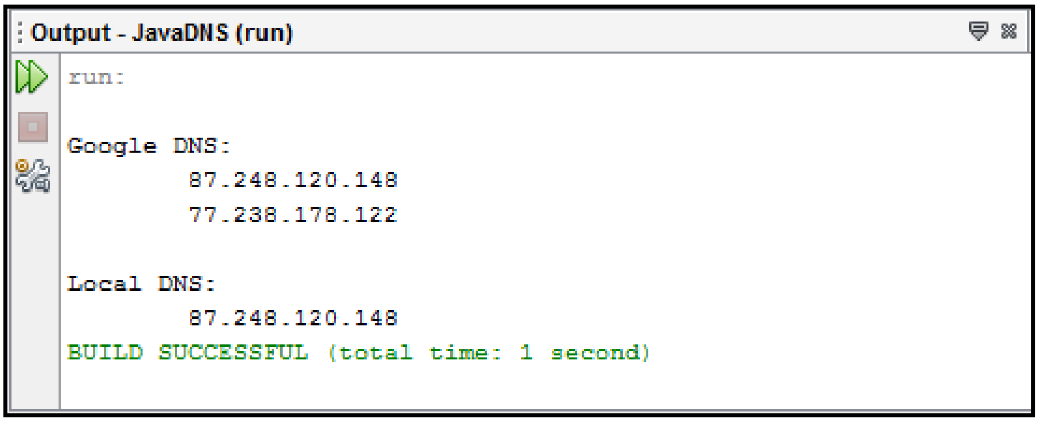
We explained and measured the results of our two steps or phases of the model, IP address check and classification in the following sections.

### 5.2.1 First Phase: IP address check:

The IP address check step is done by sending a DNS request to two DNS servers and comparing the default IP address with IP addresses returned from third party DNS (GoogleDNS or OpenDNS). If the default IP address is included in the IP addresses for third party DNS, the site is considered as legitimate, otherwise the site suspicious.

The implementation of the IP address check is done by a program written in NetBeans in java language. This program checks the IP address from local DNS and from GoogleDNS.

The output snap of the implementation is shown in Figure 5.1. If the website is clean from any pharming attack, there will be a match between two IP addresses. We can note in this figure that the IP address returned from local DNS server is included in the IP addresses returned from GoogleDNS (87.248.120.148 in that example).



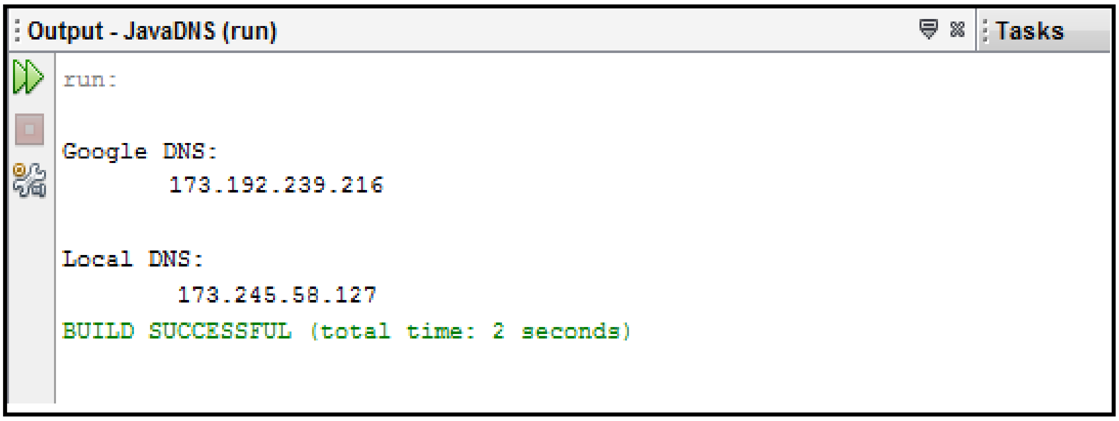
```
run:

Google DNS:
        87.248.120.148
        77.238.178.122

Local DNS:
        87.248.120.148
BUILD SUCCESSFUL (total time: 1 second)
```

*Figure 5.1: Result of IP address check of clean site*

The output of the implementation shown in Figure 5.2, in case site is suspicious. So the two IP addresses returned from local DNS and from GoogleDNS will be unmatched.



```
run:

Google DNS:
       173.192.239.216

Local DNS:
        173.245.58.127
BUILD SUCCESSFUL (total time: 2 seconds)
```

*Figure 5.2: Results of IP address check of suspicion site*

Otherwise, if there isn't a match between two returned IP addresses, the site will be considered suspicious.

We tested the program with 1503 sites. Among those sites there are 400 (26.6%) sites identified as pharming sites, and the other 1103 (73.4%) are clean sites. The results from our testing on those sites were: for these pharming sites, the program can get 391 sites of which 97.8% percentage are success. And for clean sites the program can get 1087 sites of which 98.55% percentage are success.

### 5.2.2 Second Phase: Classification of websites:

The measures of evaluating the second phase of our model are confusion matrices. Also to perform the comparisons of the tested algorithms, the result of each classifier was evaluated using the detection rate, classification error (misclassification) rate, accuracy, and F-measure. Based on the equations in section 4.5, we extracted the results of our experiments as follows:

We extracted our result on two scenarios, first: when our dataset contained 400 (26.6%) records as attacks, and 1103 (73.4%) records as clean sites (balanced dataset which means a dataset with about 30% attack and 70% clean).

Second: when our dataset contained 10 (0.9%) records as attacks, and 1103 (99.1%) records as clean sites (unbalanced dataset which means a dataset with less or over than 30% attack ).

#### 5.2.2.1 First scenario (balanced dataset):

We ran our experiment on the dataset containing 400 records as attacks (pharming sites), and the other (1103 records) clean sites and the results were as follows:

- **Decision tree classifier:**
  We applied decision tree algorithm on our dataset. Then we used x-validation, X-Validation encapsulates a cross-validation in order to estimate the performance of a learning operator. We used as shown in Figure 4.6. The confusion matrix for the result was similar to Table 5.1:

***Table 5.1: Confusion matrix table for decision tree classifier***

| | | **True Class** | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Predicted Class** | **Positive** | 393<br>(TP) | 5<br>(FP) |
| | **Negative** | 7<br>(FN) | 1098<br>(TN) |

From Table 5.1 we can calculate the accuracy and recall for that classifier as follows:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

$$= (393 + 1098) / ((393 + 1098 + 5 +7)$$

$$= 1491 / 1503 * 100$$

= 99.2%

Recall = TP/(TP + FN)

= 393 / ( 393 + 7 ) *100

= 98.25%

Precision = TP/(TP +FP)

= 393 / (393 + 5) *100

= 98.74%

True Positive rate = TP/(TP+FN)

= 393/(393+7)*100

= 98.25%

True Negative rate = TN/(TN+FP)

=1098/(1098+5)*100

=99.5%

F-measure = 2*(Precision * Recall) / (Precision + Recall)

= 2*(98.74*98.25)/( 98.74+98.25)

= 98.49%

- **KNN classifier:**
  We applied K nearest neighbor (KNN) algorithm on our dataset, then we used x-validation. X-Validation encapsulates a cross-validation in order to estimate the performance of a learning operator. We used it as shown in Figure 4.8. The confusion matrix for the result was similar to Table 5.2:

***Table 5.2: Confusion matrix table for KNN classifier***

| | | **True Class** | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Predicted Class** | **Positive** | 398<br>(TP) | 39<br>(FP) |
| | **Negative** | 2<br>(FN) | 1064<br>(TN) |

From Table 5.2 we can calculate the accuracy and recall for that classifier as follows:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

$$= (398 + 1064) / ((398 + 1064 + 2 +39)$$

$$= 1462 / 1503 * 100$$

$$= 97.27\%$$

$$\text{Recall} = TP/(TP + FN)$$

$$= 398 / ( 398 + 2 ) *100$$

$$= 99.5\%$$

$$\text{Precision} = TP/(TP +FP)$$

$$= 398 / (398 + 39) *100$$

$$= 91.07\%$$

$$\text{True Positive rate} = TP/(TP+FN)$$

$$= 398/(398+2)*100$$

$$= 99.\ 5\%$$

$$\text{True Negative rate} = TN/(TN+FP)$$

$$=1064/(1064+39)*100$$

$$=96.46\%$$

$$\text{False Negative rate} = FN/(TP+FN)$$

$$=2/(398+4)*100$$

$$=0.5\%$$

$$\text{False Positive rate} = FP/(TN+FP)$$

$$=39/(1064+39)*100$$

$$=0.45\%$$

$$\text{F-measure} = 2*(\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

$$= 2*(91.07*99.5)/( 91.07+99.5)$$

$$= 95.1\%$$

- **Naïve Bayes classifier:**
  We applied Naïve Bayes algorithm on our dataset, then we used x-validation. X-Validation encapsulates a cross-validation in order to estimate the performance of a learning operator. We used it as shown in Figure 4.10. The confusion matrix for the result was similar to Table 5.3:

***Table 5.3: Confusion matrix table for Naïve Bayes classifier***

| | | **True Class** | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Predicted Class** | **Positive** | 396<br>(TP) | 5<br>(FP) |
| | **Negative** | 4<br>(FN) | 1098<br>(TN) |

From Table 5.3 we can calculate the accuracy and recall for that classifier as follows:

Accuracy = (TP + TN) / (TP + TN + FP + FN)

= (396 + 1098) / ((396 + 1098 + 4 +5)

= 1494 / 1503 * 100

= 99.4%

Recall = TP/(TP + FN)

= 396/ ( 396 + 4 ) *100

= 99.0%

Precision = TP/(TP +FP)

= 396 / (396 + 5) *100

= 98.75%

True Positive rate = TP/(TP+FN)

= 396/(396+4)*100

= 99.0%

True Negative rate = TN/(TN+FP)

=1098/(1098+5)*100

=99.54%

False Negative rate = FN/(TP+FN)

=4/(396+4)*100

=1%

False Positive rate = FP/(TN+FP)

=5/(1098+5)*100

=0.45%

F-measure = 2*(Precision * Recall) / (Precision + Recall)

= 2*(98.75*99.0)/( 98.75+99.0)

= 98.87%

From the previous calculations we can construct Table 5.4, which contains the experimentations results from three classifiers for the first dataset.

***Table 5.4: Compares the results of three classifiers for balanced dataset***

| | Recall | Precision | Accuracy |
|---|---|---|---|
| Decision Tree | 98.25 | 98.74 | 99.2 |
| KNN | 99.5 | 91.07 | 97.27 |
| Naïve Bayes | 99.0 | 98.75 | 99.4 |
| Best Result (Best Classifier) | 99.5<br>(KNN) | 98.75<br>(Naïve Bayes) | 99.4<br>(Naïve Bayes) |

Figure 5.3 shows these results in a detailed manner.

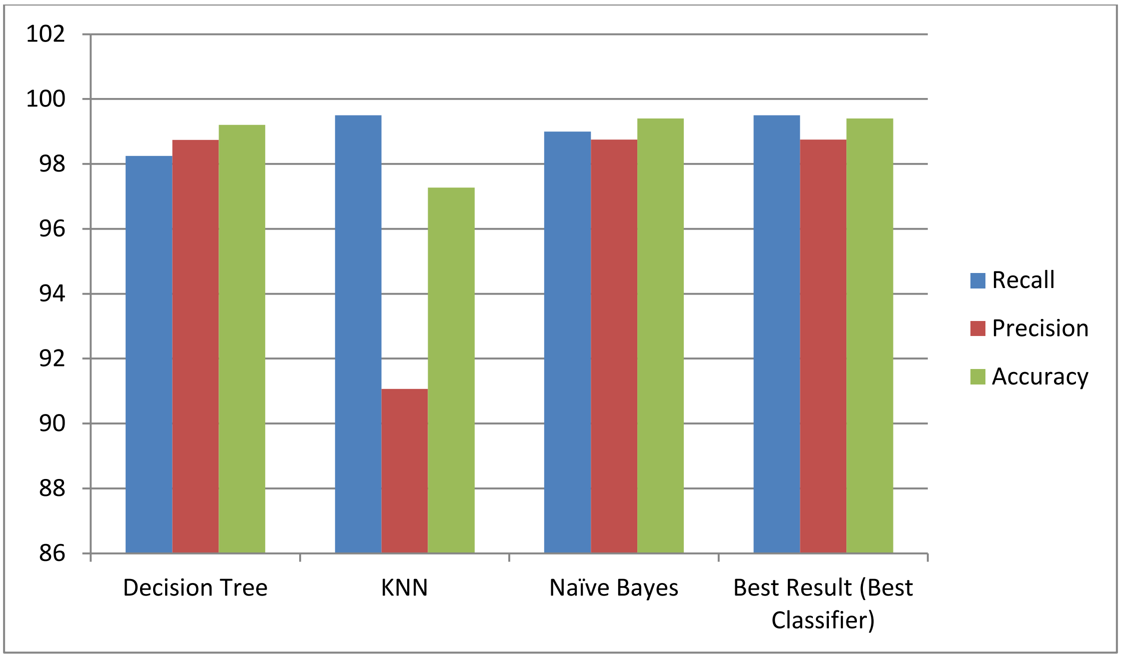


*Figure 5.3: Compares the results of three classifiers for balanced dataset*

From this comparison we can conclude that Naïve Bayes classifier is the best one from these three classifiers, because it has the best accuracy (99.4%), and the best precision (98.75%). So we will use it in our model to detect pharming attack.

### 5.2.2.2 Second scenario (unbalanced dataset):

We ran our experiment on the dataset containing just 10 records as attacks (pharming sites), and the other (1103 records) clean sites and the results were as follows:

- **Decision tree classifier:**
  We applied decision tree algorithm on our dataset, then we used x-validation. X-Validation encapsulates a cross-validation in order to estimate the performance of a learning operator. We used it as shown in Figure 4.6. The confusion matrix for the result was similar to Table 5.5:

*Table 5.5: Confusion matrix table for decision tree classifier*

| | | **True Class** | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Predicted Class** | **Positive** | 4<br>(TP) | 2<br>(FP) |
| | **Negative** | 6<br>(FN) | 1101<br>(TN) |

From Table 5.5 we can calculate the accuracy and recall for that classifier as follows:

$$\begin{aligned}\text{Accuracy} &= (TP + TN) / (TP + TN + FP + FN)\\ &= (4 + 1101) / (4 + 1101 + 2 + 6)\\ &= 1105 / 1113 * 100\\ &= 99.28\%\end{aligned}$$

$$\begin{aligned}\text{Recall} &= TP/(TP + FN)\\ &= 4 / ( 4 + 6 ) *100\\ &= 40\%\end{aligned}$$

$$\begin{aligned}\text{Precision} &= TP/(TP +FP)\\ &= 4 / (4 + 2) *100\\ &= 66.67\%\end{aligned}$$

$$\begin{aligned}\text{True Positive rate} &= TP/(TP+FN)\\ &= 4/(4+6)*100\\ &= 40\%\end{aligned}$$

$$\begin{aligned}\text{True Negative rate} &= TN/(TN+FP)\\ &=1101/(1101+2)*100\\ &=99.8\%\end{aligned}$$

$$\begin{aligned}\text{False Negative rate} &= FN/(TP+FN)\\ &=6/(4+6)*100\\ &=60\%\end{aligned}$$

$$\begin{aligned}\text{False Positive rate} &= FP/(TN+FP)\\ &=2/(1101+2)*100\\ &=0.18\%\end{aligned}$$

$$\begin{aligned}\text{F-measure} &= 2*(\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})\\ &= 2*(66.67*40)/( 66.67+40)\\ &= 50\%\end{aligned}$$

- **KNN classifier:**
  We applied K nearest neighbor (KNN) algorithm on our dataset, then we use x-validation. X-Validation encapsulates a cross-validation in order to estimate the performance of a learning operator. We used it as shown in Figure 4.8. The confusion matrix for the result was similar to Table 5.6:

***Table 5.6: Confusion matrix table for KNN classifier***

| | | **True Class** | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Predicted Class** | **Positive** | 9<br>(TP) | 5<br>(FP) |
| | **Negative** | 1<br>(FN) | 1098<br>(TN) |

From Table 5.6 we can calculate the accuracy and recall for that classifier as follows:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

$$= (9 + 1098) / ((9 + 1098 + 5 +1)$$

$$= 1107 / 1113 * 100$$

$$= 99.46\%$$

$$\text{Recall} = TP/(TP + FN)$$

$$= 9 / ( 9 + 1 ) *100$$

$$= 90.0\%$$

$$\text{Precision} = TP/(TP +FP)$$

$$= 9 / (9 + 5) *100$$

$$= 64.29\%$$

$$\text{True Positive rate} = TP/(TP+FN)$$

$$= 9/(9+1)*100$$

$$= 90.0\%$$

$$\text{True Negative rate} = TN/(TN+FP)$$

=1098/(1098+5)*100

=99.5%

False Negative rate = FN/(TP+FN)

=1/(9+1)*100

=10%

False Positive rate = FP/(TN+FP)

=5/(1098+5)*100

=0.45%

F-measure = 2*(Precision * Recall) / (Precision + Recall)

= 2*(64.29*90)/( 64.29+90)

= 75%

- **Naïve Bayes classifier:**
  We applied Naïve Bayes algorithm on our dataset, then we used x-validation. X-Validation encapsulates a cross-validation in order to estimate the performance of a learning operator. We used it as shown in Figure 4.10. The confusion matrix for the result was similar to Table 5.7:

***Table 5.7: Confusion matrix table for Naïve Bayes classifier***

| | | **True Class** | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Predicted Class** | **Positive** | 9<br>(TP) | 0<br>(FP) |
| | **Negative** | 1<br>(FN) | 1103<br>(TN) |

From Table 5.7 we can calculate the accuracy and recall for that classifier as follows:

Accuracy = (TP + TN) / (TP + TN + FP + FN)

= (9 + 1103) / ((9 + 1103 + 0 +1)

= 1112 / 1113 * 100

= 99.91%

Recall = TP/(TP + FN)

= 9/ ( 9 + 1 ) *100

= 90.0%

Precision = TP/(TP +FP)

= 9 / (9 + 0) *100

= 100%

True Positive rate = TP/(TP+FN)

= 9/(9+1)*100

= 90.0%

True Negative rate = TN/(TN+FP)

=1103/(1103+0)*100

=100%

False Negative rate = FN/(TP+FN)

=1/(9+1)*100

=10%

False Positive rate = FP/(TN+FP)

=0/(1103+0)*100

=0%

F-measure = 2*(Precision * Recall) / (Precision + Recall)

= 2*(100*90)/( 100+90)

= 94.74%

From the previous calculations we can construct Table 5.8, which contains the experimentations results from three classifiers for the first dataset.

***Table 5.8: Compares the results of three classifiers for unbalanced dataset***

| | Recall | Precision | Accuracy |
|---|---|---|---|
| Decision Tree | 40 | 66.67 | 99.28 |
| KNN | 90.0 | 64.29 | 99.46 |
| Naïve Bayes | 90.0 | 100 | 99.91 |
| Best Result (Best Classifier) | 90.0<br>(Naïve Bayes) | 100<br>(Naïve Bayes) | 99.91<br>(Naïve Bayes) |

Figure 5.4 shows these results in a more detailed manner.

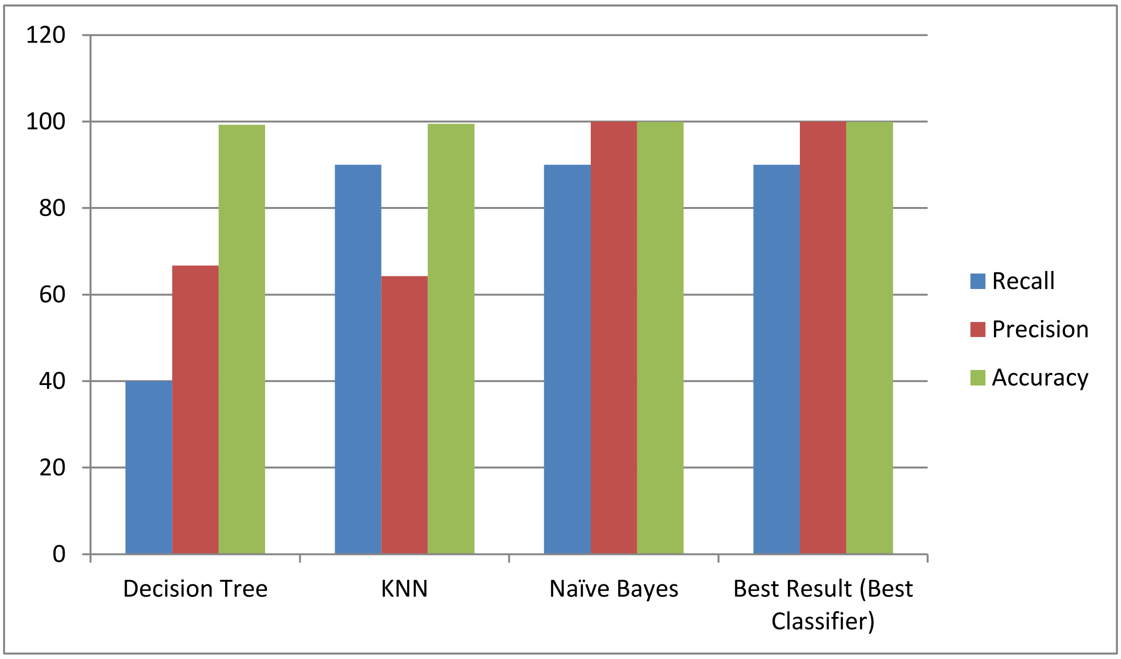


***Figure 5.4: Compares the results of three classifiers for unbalanced dataset***

From this comparison we can conclude that Naïve Bayes classifier is the best one from these three classifiers, because it has the best accuracy (99.91%), the best precision (100%) and the best recall (90.0%).Sso we will use it in our model to detect pharming attack.

## 5.3 The overall evaluation of our model:

The overall evaluation of our model depended on a two-steps result of our model. IP address check and classification. Table 5.9 shows the combination of those results.

***Table 5.9: The overall results of our model***

| | Infected sites<br>(400) | | Clear sites<br>(1103) | |
|---|---|---|---|---|
| | True Positive Rate (TP) | False Negative Rate (FN) | True Negative Rate (TN) | False Positive Rate (FP) |
| IP address check<br>[First Phase] | 97.8% (391) | 2.2% (9) | 98.55% (1087) | 1.45% (16) |
| Classification (Naïve bayes)<br>[Second Phase] | 99% | - | - | 99.54 |
| Overall solution | 96.77% | 2.2% | 99.55% | 99.9% |

From Table 5.9 we can summarize the results of our method to defeat pharming and phishing attack. Our method shown in Figure 4.1 is divided into two phases: Phase one IP address check from two DNS servers (default DNS and alternate DNS, in which we chose GoogleDNS [32]). We chose Google DNS in our method because it's one of the best and most famous alternate DNS in the internet. But in spite of that we tested our method on other alternate DNS server (OpenDNS) and found that results were close, and somewhat the GoogleDNS were the best.

In our first phase in the model, we compared the two returned IP addresses from the default DNS server and from GoogleDNS server. If there was a match between those two IP addresses, then the web pages were considered as legitimate, otherwise the second phase on our model was applied.

In the second step of our model which is classification first we downloaded our data set from two sites: phishtank [28] and Phishing Working Group (APWG) [27], and we completed these datasets with phishing indicators and features which we collected by designing a program using Visual Studio program [31]. This program collects the features from any site by putting the URL of that site in the bar panel in running of that program as shown in Figure 4.4.

### 5.4 Conclusion:

By reviewing our results from two phases on our dataset (balanced an unbalanced datasets), we found that, during the first phase of our model, for infected sites we had 400 (26.6%) infected sites. Our first phase on the model detected 391 sites as infected and its real infected in 97.8%. As for the other 9 sites (2.2%) the model didn’t catch them on the first phase, this 2.2% percentage is a low percentage and we can minimize them in the future by doing the second phase of our model on this percentage. By doing that the percentage will be minimized to 0.03% (99% which is classification success percentage from 2.2% equal 0.03%) but in this case it will be time consuming. But as we mentioned before, in these type of attacks the accuracy is more important than performance.

For these 391 infected sites 97.8%, when it was tested on our second phase of our model (classification which has 99% success percentage) the success percentage of our model became 96.77% which is an excellent success ratio, especially since other researches haven’t reached them. On paper [6] S. Prevost and M. Laurent, got results reaching 95% success percentage, and other authors worked on defeating phishing and pharming attacks, and explained their techniques but didn’t mention the success ratio of their results.

# Chapter 6: Conclusion and Future work

This chapter concludes the work, its results and discussion. As well as further future work directions are remarked

## 6.1 Conclusion:

Phishing and Pharming are two of the most organized crimes of the twenty-first century [19] requiring very little skill on part of the fraudster. And the challenge of keeping sensitive information like bank accounts and passwords of the users safe from the hands of attackers becomes more important day after day.

Phishing attack is a fraudulent attempt to gain personal information from victims such as bank information, credit card information, social security, employment details, and online shopping account passwords and so on. Phishing attacks use fraudulent e-mails or websites designed to fool users into divulging personal financial data by stealing the trusted brands of well-known banks, e-commerce and credit card companies. Pharming is a sophisticated version of phishing attack working by poison or Manipulation of DNS server records to route users to fake sites for fraudulent purposes to steal the sensitive information.

Our model - based on a dual-step analysis and collaboration of multiple (default and reference) DNS servers- proposes an anti-pharming protection at the client-side for detection of DNS corruptions. We used GoogleDNS server as an alternate DNS server. Its implementation into the client's browser can be part of a global solution that combines both protections against phishing and pharming attacks. In our research, we demonstrate that the IP address check is a significant indicator of the legitimacy of a visited login website to discover the manipulation on DNS server. We tested our program to check IP addresses with 1503 sites. From those sites there were 400 sites identified as pharming sites, and the other 1103 were clean sites. The results from our testing on those sites were: for these pharming sites, the program can get 391 sites which is 97.8% percentage accurate. And for those clean sites our program could get 1087 sites which are 98.55% percentage accurate. This percentage is acceptable.

In addition, the classification of the webpage results which was the second phase of our approach indicated that the classification techniques helped significantly to differentiate legitimate from fraudulent websites for up to 99% of more than 1500 sets of tested WebPages. The measures of evaluating the second phase of our model "the classification of the web pages" are confusion matrices. Also to perform the comparisons of the tested algorithms through the results, each classifier was evaluated using the detection rate, classification error (misclassification) rate, accuracy, and F-measure.

We used the efficiency of the data mining techniques to strengthen our model. We tested our dataset using three different classification techniques, decision tree, naïve bayes, and K-nearest neighbor techniques. We found that naïve bayes classification had the best accuracy when we tested it on our dataset; naïve bayes had 99% accuracy which was the best rate from those three classification techniques. We extracted our results on two scenarios, first: when our dataset contains 400 records as attacks, and 1103 records as clean sites (balanced dataset). Second: when our dataset contains 10 records as attacks, and 1103 records as clean sites (unbalanced dataset).

Our model proposes an anti-pharming protection for the user on how to deal with secretive information like bank accounts, in order to detect pharming and phishing attacks at the client-side, so the evaluation of our model was based on accuracy. When we tested our model on our own laptop, we found that our model had accuracy with more than 96.7% rate. And Table 6.1 shows the overall percentages of our model.

***Table 6.1: Overall percentages of our model***

| | Infected sites | | Clear sites | |
|---|---|---|---|---|
| | True Positive Rate (TP) | False Negative Rate (FN) | True Negative Rate (TN) | False Positive Rate (FP) |
| IP address check<br>[First Phase] | 97.8% | 2.2% | 98.55% | 1.45% |
| Classification (Naïve bayes)<br>[Second Phase] | 99% | - | - | 99.54 |
| Overall solution | 96.77% | 2.2% | 99.55% | 99.9% |

Finally, we hope that the proposed model to be integrated into global solutions that combine protection against both phishing – such as an anti-phishing toolbar – and pharming attacks.

## 6.2 Future works:

We can summarize the future work of our research as follow:

- Search for a solution to detect pharming attack on server side, in addition to a solution on client side.
- For classification we can use other classification theories, like Neural Network for example to increase success rate and accuracy for the model.
- Using clustering methods instead of classification methods to predict pharming and phishing attacks.
- Search for more characteristics and indicators for phishing and pharming web sites to promote and strengthen the solution to detect phishing and pharming attacks.
- Testing the validation of DNS server by checking IP addresses on two levels, first with local DNS server on the LAN, then with alternate DNS server like GoogleDNS server.

# References

[1] S. Stamm, Z. Ramzan, and J. Markus, “Drive-By pharming,” in Proceedings of the 9th international conference on Information and communications security, China, 2007, pp. 495-506.

[2] P. Sengar and V.Kumar, "Client-Side Defense against Phishing with PageSafe", International Journal of Computer Applications, July 2010, P6-10,

[3] B. Aslam, L. Wu and Cliff C. Zou, "PwdIP-Hash - A Lightweight Solution to Phishing and Pharming Attacks", Ninth IEEE International Symposium on Network Computing and Applications, 2010, pp.198-203.

[4] A. Martin, Na. Anutthamaa, M. Sathyavathy, M. Francois, Dr.P. Venkatesan, "A Framework for Predicting Phishing Websites Using Neural Networks", International Journal of Computer Science Issues "IJCSI", March 2011, pp.330-336.

[5] S. Prevost, G. Granadillo, and M. Laurent, "A dual Approach To Detect Pharming Attacks At The Client-Side", in Proceeding of the 4th IFIP International Conference on New Technologies, Mobility and Security (NTMS), Paris, France, Feb. 2011

[6] S. Prevost and M. Laurent, "Defeating Pharming Attacks At Client-Side", Network and System Security (NSS), Sep. 2011.

[7] M. Aburrous, M. Hossain, K. Dahal, and F. Thabtah "Predicting Phishing Websites Using Classification Mining Techniques with Experimental Case Studies" Seventh International Conference on Information Technology, 2010.

[8] G. Ollman, “The Pharming Guide,” Jul. 2005.

[9] D. Davidowicz, "Domain Name System (DNS) Security", 1999.

[10] T. Phyu, "Survey of Classification Techniques In Data Mining", in Proceedings of the International MultiConference of Engineers and Computer Scientists, 2009.

[11] R. Layton and P. Watters, "Determining Provenance In Phishing Websites Using Automated Conceptual Analysis", in proceeding of 4th Annual APWG eCrime, 2009.

[12] O. Mahmood, "Three Phase Checking Against Phishing And Pharming Attacks", in Proceedings of the 11th Annual Conference of Asia Pacific Decision Sciences Institute Hong Kong, 2006, pp.399-402.

[13] Y. Cao, W. Han, Y. Le "Anti-phishing Based on Automated Individual White-List", Proceedings of the 4th ACM workshop on Digital identity management, Alexandria, Viriginia, USA: ACM, 2008, pp.51-60.

[14] R.Sumathi, R.Vidhya Prakash, "Prediction of Phishing Websites Using Optimization Techniques", International Journal of Modern Engineering Research (IJMER), 2012, pp.341-348.

[15] M. Alkhozae, O. Batarfi, "Phishing Websites Detection based on Phishing Characteristics in the Webpage Source Code", International Journal of Information and Communication Technology Research, 2011, pp.285-291.

[16] M. Aburrous, M.A. Hossain, F. Thabatah and K. Dahal, "Intelligent phishing website detection system using fuzzy techniques", in 3rd International Conference on Information and Communication Technologies: From Theory to Applications (ICTTA), 2008. pp.7913-7921.

[17] L. Witten, and E. Frank," Data Mining: Practical Machine Learning Tools and Techniques", the Morgan Kaufmann Series in Data Management Systems, 2nd edition, 2005.

[18] J. Han and M. Kamber, "Data Mining Concepts and techniques" 2nd edition, the Morgan Kaufmann Series in Data Management Systems, 2006.

[19] S. Afroz and R. Greenstadt, " PhishZoo: An Automated Web Phishing Detection Approach Based on Profiling and Fuzzy Matching", 2009.

[20] D. Olson, and D. Delen, "Advanced data mining techniques", Springer-Verlag Berlin Heidelberg, 2008.

[21] N. Ye, "The Handbook of Data Mining", Lawrence Erlbaum Associates, Inc., 2003.

[22] L. Jiang, H. Zhang, and J. Su, "Learning k-Nearest Neighbor Naive Bayes for Ranking", Excellent Youth Foundation of China University of Geosciences, 2005.

[23] T. Srivastava, "Phishing and Pharming - The Evil Twins", SANS Institute Reading Room site, 2007.

[24] F. Carli, "Security Issues with DNS", e SANS Institute Reading Room site, 2003.

[25] M. Aburrous, M. Hossain, K. Dahal, and F. Thabtah "Associative Classification Techniques for predicting e-Banking Phishing Websites", MCIT2010, pp.9-12.

[26] T. Moore and R. Clayton, "An empirical analysis of the current state of phishing attack and defence", In Proceedings of the Workshop on the Economics of Information Security (WEIS2007).

[27] APWG-Anti-Phishing working group. Available: http://www.apwg.org/, [last accessed, 26/8/2013].

[28] “PhishTank.” . Available: http://www.phishtank.com/, [last accessed, 26/8/2013].

[29] The future hacker's site, Available: http://thefuturehackers.com/2011/11/pharming.html , last accessed 7/2/2012.

[30] Rapid Miner 5.1, Available: http://www.rapidminer.com, (2013, March), [Online], [last accessed, 26/8/2013].

[31] Visual Studio 2012, Available: http://www.microsoft.com/visualstudio/eng/downloads, [last accessed, 26/8/2013].

[32] Google DNS server, Available: http://code.google.com/intl/fr/speed/public-dns/index.html [last accessed, 26/8/2013].

[33] InfoSec Institute, Available: http://resources.infosecinstitute.com/pharming-attack/ [last accessed, 23/10/2013].

[34] Domesticating IT, Available: http://www.domesticatingit.com/domain-web-and-email-hosting-explained/dns/ [last accessed, 23/10/2013].

[35] Symantec website, Available: http://www.symantec.com/connect/blogs/symantec-intelligence-report-february-2012 [last accessed, 29/10/2013].