International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 3 (2011) pp. 391-398
© MIR Labs, www.mirlabs.net/ijcisim/index.html

Implementation and Comparison of Machine Learning Classifiers for Information Security Risk Analysis of a Human Resources Department

Mete Eminagaoglu[1], Saban Eren[2]

[1]Department of Computer Programming, Yasar University,
Universite Caddesi, No:35-37, Bornova, Izmir, 35100, Turkey
mete.eminagaoglu@yasar.edu.tr

[2]Department of Statistics, Yasar University,
Universite Caddesi, No:35-37, Bornova, Izmir, 35100, Turkey
saban.eren@yasar.edu.tr

Abstract: **The aim of this study is threefold. First, a qualitative information security risk survey is implemented in human resources department of a logistics company. Second, a machine learning risk classification and prediction model with proper data set is established from the results obtained in this survey. Third, several classifier algorithms are tested where their training and test performances are compared using error rates, ROC curves, Kappa statistics and F-measures. The results show that some classifier algorithms can be used to estimate specific human based information security risks within acceptable error rates.**

Keywords: Information security, Risk assessment, Machine learning, binary classifiers, Human resources, security risk survey

I. Introduction

Proper and accurate assessment and management of the information security risks has become a critical issue in today's business world. However, information security risks can not always be estimated reliably because each organization might have different risks or the same risks with different levels due to divergent company environments, cultures, processes and organizations [1]. This yields to new methods and models for information security risk analysis.

In the recent years, some researches have been made which implement information security risk assessments using machine learning and similar computational intelligence, decision making and reasoning models such as fuzzy logic [2], [3], [4], [5], [6]. Most of these remarkable studies either focus on multi-classifier machine learning models for quantitative risks or technological aspects of information security such as IDS, Firewall, e-mail filter systems [7], [8], [9]. However, in today's business life; there exist some information security risks which cannot be properly quantified or which are based on human factors. In such situations, there is always a need for reliable and accurate automated qualitative risk assessment models which can be used easily by senior management without depending on the knowledge of information security experts. In addition, such new models must be implemented so as to minimize the drawbacks of qualitative risk methodologies such as subjectivity, uncertainty and false predictions [10], [11].

In this study, a proprietary qualitative information security risk assessment model is implemented in a logistics company by the aid of machine learning classifiers. First, an information security survey was conducted. Then, a machine learning model was generated using the results obtained from the survey. The model was tested among different selected learning algorithms and the results were analyzed.

II. Information Security Risk Survey

The survey was implemented in a logistics company in Turkey. During the implementation, there were six employees working in HR (human resources) department including HR Manager. The survey was conducted when the first author of this paper was working as information security manager in the same company. Due to privacy and confidentiality considerations of the company policies, the name of the company is not mentioned in this paper. HR Manager and information security manager made several meetings for the design of the survey. In these meetings the information security considerations that were mostly relevant to the company's human resources management were included in the survey questions. Hence only some specific assets, vulnerabilities and threats were taken into consideration. These are grouped as follows;

Asset List:

a) HR Manager
b) IT staff for HR department (giving IT services to HR department)
c) HR laptops
d) HR database
e) E-mails of HR department
f) HR documents
g) HR staff
h) Electrical infrastructure of HR office

i) Company's recruitment and employment strategies and verbal procedures
j) Employee termination strategies and verbal procedures
k) Electronic data stored on HR computers
l) IT network infrastructure of HR office
m) Fax and phone lines infrastructure of HR office

Vulnerability List:
a) People's tendency to make mistakes unintentionally
b) Lack of awareness and lack of compliance with company policies
c) Could be tempted to sell, give away, etc many critical information
d) Lack of technical knowledge and experience
e) Having disgruntled employee due to low wages, work conditions or possibility of being fired
f) Insufficient process or absence of; employee screening and monitoring controls
g) Insufficient process or absence of specific controls including change management, removal of user access rights and return of company assets
h) There's no inventory of HR assets in the company
i) Lack of business continuity plans and relevant controls

Threat List:
a) Mistakes, errors by people
b) Users' wrong data entry
c) Social engineering attacks (from outside)
d) Social engineering attacks (from inside)
e) Other rival companies
f) Technical hacker attacks (from outside)
g) Technical hacker attacks (from inside)
h) Physical damage by accident
i) Physical theft / lost
j) Unauthorized access to HR database
k) Malicious codes
l) Unavailability of employee due to health conditions
m) Unavailability of employee due to environmental hazards or disasters
n) Unavailability of employee due to kidnapping, sabotage, etc.
o) Physical damage intentionally (from inside)
p) Unavailability of HR data and systems in emergency response situations

13 assets, 9 vulnerabilities and 16 threats were included in the scope of the study. So, for each of the 13 assets, each of the 16 threats might impose a possible risk exploiting each of the 9 possible vulnerabilities. Regarding a many-to-many relationship, this would make up a total of (13 x 9 x 16) = 1872 possible combinations. However, in real life situations most of these possible combinations and relations are neither relevant nor sensible and their probability is 0. These were automatically discarded from the survey infrastructure. Some of the other possible combinations were not also taken into scope of the survey due to human resources' managerial strategies. This narrowed the scope of the survey to 57 distinct topics, or in other words, 57 possible combinations of assets, threats and vulnerabilities. These combinations and the relevant risks are based on / caused by human factors or directly related with the core business of human resources department of the company.

It should be mentioned that the assets used in the survey are also categorized into six different types in a sense that is similar to the international standards and best-practices [12] [13], [14]. These six asset types are as follows; Employees (human), Electronic data (including software), Hard copy documents, Infrastructure, IT Systems (computers, network switches, database systems, etc.), and Know-how (about business processes and managerial issues).

The evaluation criteria, ranks and related survey questions were also defined by HR Manager and information security manager. These questions were either to be answered on a (Yes / No) or (1 / 2 / 3 / 4 / 5), (0 / 1 / 2 / 3) ranked scale basis. These are denoted separately in Table 1, Table 2 and Table 3.

All the personnel in HR department including HR Manager answered these 9 questions for each of the 57 risk relations. The surveys were carried out independently and anonymously. In addition, all the respondents were given a 1 hour of training by information security manager before the survey sessions. However, in the actual survey materials which were provided both as Microsoft Excel spreadsheets and hard copy survey forms; these 9 questions were given in distinct columns and each distinct row was one of the 57 risk relations. The sample survey form with some of the collected data is denoted in Figure 1.

Asset Name	Vulnerability Name	Threat Name	Probability of Risk Occurance? (1 to 5.)	Negative Impact to business info Confidentiality (1 to 5)	Negative Impact to business data / info Integrity (1 to 5)	Negative Impact to business process A vailability (1 to 5)	How many times occured in previous year (0,1,2,3 0= None, 1= 1 ,2 = less	Has ever been trained in this subject? (Y/N)	A ny knowledge in this issue? (0,1,2,3)	Is there a policy / procedure in the company for this subject? (Y,N)	Overall Risk (1 to 5)
HR Database	People's tendency to make n	Users' wrong data entry	5	1	4	2	2	Yes	3	Yes	4
	Lack of awareness & lack of	Unauthorized access to HR I	3	5	4	1	1	Yes	2	Yes	4
E-mails of HR dep	Lack of awareness & lack of	malicious codes	5	2	2	3	1	No	2	Yes	4
	Lack of awareness & lack of	technical hacker attacks (fro	4	4	3	3	2	No	1	Yes	4
	Lack of awareness & lack of	technical hacker attacks (fro	2	4	3	2	0	No	1	Yes	2
	There's no inventory of HR a	unavailability of HR data & sy	3	1	2	4	0	No	1	No	5
HR documents	People's tendency to make n	Physical damage by acciden	4	1	2	5	2	Yes	2	No	4
	There's no inventory of HR a	unavailability of HR data & sy	5	1	2	4	1	No	1	No	5
	Lack of awareness & lack of	Physical theft / lost	3	4	2	5	2	Yes	2	No	4
	disgruntled employee due to	Physical damage intentionall	2	1	4	4	0	No	1	No	2
	People's tendency to make n	mistakes, errors	5	3	4	3	3	Yes	3	No	4
HR staff	Lack of business continuity	unavailability of employee du	3	1	2	4	3	No	3	No	4
	Lack of business continuity	unavailability of employee du	2	1	2	4	0	No	3	No	3
	Lack of business continuity	unavailability of employee du	1	3	2	4	0	No	0	No	2
	Lack of awareness & lack of	social engineering attacks (fr	2	4	4	1	0	No	0	No	3
	Lack of awareness & lack of	social engineering attacks (fr	3	4	4	1	1	No	0	No	4
	Could be tempted to sell, giv	other rival companies	2	5	1	4	2	No	1	No	3

Figure 1. A small excerpt from the information security risk survey

These nine questions plus three related parameters (asset, threat and vulnerability) make up a total of 12 criteria which are used as attributes for machine learning classifiers.

The implementation and modeling is explained in more detail in the following section. All the data and values in the survey were scalable values which were suitable for a qualitative risk analysis and assessment [10], [11], [14].

How many times the similar event / risk has been in the previous year? (Notice: If none select 0, if only once select 1, if more than once and less than 5 select 2, if more than 5 select 3)	0	1	2	3
What is your level of experience or knowledge about this topic / subject? (Notice: If no knowledge select 0, if few knowledge select 1, if average knowledge select 2, if expert knowledge select 3)	0	1	2	3

Table 1. Survey questions with 0 to 3 scale.

Have you ever been trained in this subject?	Yes	No
Is there a policy / procedure in the company for this subject?	Yes	No

Table 2. Survey questions with Yes/No options.

What is the probability of this risk to occur?	1	2	3	4	5
If this risk occurs, what is the negative impact to confidentiality?	1	2	3	4	5
If this risk occurs, what is the negative impact to integrity?	1	2	3	4	5
If this risk occurs, what is the negative impact to availability?	1	2	3	4	5
Give an overall grade for this risk	1	2	3	4	5
General Notice: Very Low = 1 Low = 2 Medium = 3 High = 4 Very High = 5					

Table 3. Survey questions with 1 to 5 scale.

III. Modeling and Implementation for Machine Learning

The survey answers that were obtained from the respondents were analyzed in an MS Excel spreadsheet file. After the analysis, HR Manager and information security manager agreed upon the risk threshold value as 3. In other words, the overall risk values that were 1, 2 or 3 were to be treated as acceptable risks and were marked as “Risk = No”. Hence, all the other ones having overall risk values as 4 or 5 were marked as “Risk = Yes”. By this way, out of the 342 answers from the survey; 129 of them were categorized as non-risky (classified as “No”) and 213 of them were categorized as risky (classified as “Yes”). Thus, the basic model was to estimate whether an instance was risky or not. This approach also made it feasible for the binary classifiers in machine learning models where each instance coming from the data set is to be identified in any one of the two possible classes [15]. After this classification, the results were re-organized as a proper data set to be used as input for the machine learning classifiers. In the study, all the machine learning experiments were conducted by Weka software (version 3.6.0.). Eleven different built-in classifier algorithms within Weka software were chosen and this data set was used for the observations of learning performance among each of these classifiers. The names and the types of the classifier algorithms are given in Table 4.

It should be noted that; for each of the experiments among different classifiers used in this study, two phases were carried out and respective results were analyzed. Phase I was the training phase where the whole data set was used for training the classifier model. In Phase II, the same data set was used as test set by the aid of cross-validation methodology [15]. 10-folds stratified cross-validation was chosen as a best-practice option [15]. This was a crucial point in this study because no classifier algorithm can be evaluated reliably only by observing its performance values for training [16]. This is due to the fact that some classifier models algorithms have the danger of over-fitting which could be overcome by using test sets as well as train sets [15]. In this study, since the size of the data set was relatively small and it was obtained from survey answers, cross-validation was used for generating the test set.

During the experiments for all of the classifier algorithms, some specific parameter settings or initial values were used. These are listed in Table 5.

Name	Type
BayesNet	Bayes network learning
Bagging	Meta learner (REPTree as the base learner
LogitBoost	Meta learner (DecisionStump as the base learner)
Multiclass Classifier (2-class classifier)	Multinomial logistic regression model with a ridge estimator
Dagging	Meta learner (sequential minimal optimization algorithm as the base learner)
Lazy.LBR (Lazy Bayesian Rules)	Lazy Bayesian Rules learner
NB Tree (Naïve Bayes Tree)	Decision tree (builds a decision tree with Naïve Bayes classifiers at the leaves)
J48	Decision tree (C4.5 decision tree learner; implements C4.5 revision 8)
VFI (Voting Feature Intervals method)	Miscellaneous (classification by voting feature intervals)
SMO (Sequential Minimal Optimization)	Function (sequential minimal optimization algorithm for support vector classification)
DTNB (Decision Table Naïve Bayes)	Rule learner (decision table / Naïve Bayes hybrid classifier)

Table 4. List of classifier algorithms used in the experiments.

The primary success or accuracy criterion for any of the classifier algorithms was to achieve a maximum of 10%

error rate from the test set. In other words, the sum of TP (True Positive) and TN (True Negative) classifications should be at least 90% of the whole test set. This can be simply formulated as follows;

$$Let\ C = \frac{TP + TN}{(TP + TN + FP + FN)} * 100\,. \qquad (1)$$

$$If\ C \geq (90\%)\ then\ accept\ as\ accurate.$$

In (1), FP is denoted for False Positives and FN is denoted for False Negatives in the test set.

This implies that if a classifier algorithm distinguishes at least 90% of the risky (Risk = Yes) and non-risky (Risk = No) instances from the test set correctly; then it is accepted as a reliable risk learner and classifier. This threshold value was defined by the mutual agreement of HR Manager and information security manager. It should be noted that this value was also selected due to the information security risk assessment model in this study and senior management's objectives. Hence, for some other models or companies this value might be changed.

Name	Settings
BayesNet	search algorithm: hill climbing max. number of parents: 1 (this sets it to work as Naïve Bayes classifier) estimator: SimpleEstimator alpha: 0.5 use ADTree: No
Bagging	BagSize percentage: 100% (percentage of the training set data) calcOutofBag: False (out-of-bag error is not calculated) Base classifier: REPTree seed:1 (random seed number) number of iterations: 100
LogitBoost	likelihoodthreshold: -1.7976931348623157E ^{308} number of folds: 0 number of iterations: 10 number of runs: 1 seed :1 (random seed number) shrinkage: 1 re-sampling not used weight threshold: 100
Multiclass Classifier (2-class classifier)	Class: multinomial logistic regression model with a ridge estimator (This is used for 2-class, namely, Risk = Yes / No in this study) method: 1-against-all (This parameter sets the method to use for transforming the multi-class problem into several 2-class ones) random width factor: 2 (It sets the width multiplier when using random codes. The number of codes generated will be thus number multiplied by the number of classes) seed: 1 (random seed number) pair wise coupling not used
Dagging	classifier: John Platt's sequential minimal optimization algorithm for training a support vector classifier seed: 1 (random seed number) verbose is not set (If verbose is set and used; it outputs some additional information) number of folds: 10 (This parameter is used for splitting the training set into smaller chunks for the base classifier)
Lazy.LBR (Lazy Bayesian Rules)	Lazy Bayesian Rules Classifier. Lazy Bayesian Rules selectively relaxes the independence assumption, achieving lower error rates over a range of learning tasks. LBR defers processing to classification time, making it a highly efficient and accurate classification algorithm when small numbers of objects are to be classified. However, there are no flexible or easy-to-use parameters for this classifier.
NB Tree (Naïve Bayes Tree)	This is a decision tree algorithm with naive Bayes classifiers at the leaves. However, there are no flexible or easy-to-use parameters for this classifier.
J48	This classifier is used for generating a pruned or unpruned C4.5 decision tree confidence factor: 0.25 (The confidence factor is used for pruning where smaller values incur more pruning) pruned C4 is used reduced error pruning is not used binary splits on nominal attributes is not used number of folds: 3 (This parameter determines the amount of data used for reduced-error pruning. One fold is used for pruning, the rest for

	growing the tree) minimum number of objects: 2 (The minimum number of instances per leaf) seed: 1 (random seed number) Laplace based smoothing is not used (The counts at leaves are not smoothed using a Laplace based model) sub-tree raising is set to True (If this is set to True, it also covers the sub-tree raising operation when pruning)
VFI (Voting Feature Intervals method)	bias: 0.6 (This parameter determines the strength of bias towards more confident features) weight feature intervals by confidence option is set and used
SMO (Sequential Minimal Optimization)	Sequential minimal optimization algorithm for training a support vector classifier is used where it globally replaces all missing values and transforms nominal attributes into binary ones. complexity: 1 checks not used (If checks are used, it might increase computation time significantly) Epsilon: 1.0E $^{-12}$ (The epsilon value is used for round-off error and it is not changed) filter type: normalize training data seed: 1 (random seed number) tolerance parameter: 0.0010 (this is a specific parameter and it is not changed, used with its default initial value) Kernel: PolyKernel kernel cache size: 250007 exponent: 1 (this specific polynomial function kernel is used in the algorithm)
DTNB (Decision Table Naïve Bayes)	CrossVal: 1 (this parameter sets the number of folds for cross validation and if it is set to 1, it leaves one out) Evaluation Measure: Accuracy: Discrete Class RMSE: Numeric Class (this measure evaluates the performance of attribute combinations used in the decision table) Search method: DTNB BackwardsWithDelete (this is a specialized search method that performs a forward selection (naive Bayes) / backward elimination (decision table) and it also drops attributes entirely from the combined model)

Table 5. Some of the parameters used in the classifier algorithms.

As well as the error rates, all the algorithms were also compared and their performances were evaluated with respect to their Kappa Statistics, F-measures (weighted averages for “Risk = Yes” and “Risk=No”) and ROC (Receiver Operating Characteristic) curve area (weighted averages of the plot area beyond TP / FP curves for “Risk = Yes” and “Risk=No”) values. These metrics are also recommended for observing the performance of a machine learning classifiers as well as error rates [16], [17]. After the elimination of the low accurate or over-fitting classifier algorithms, the remaining classifier algorithms were also comparatively analyzed for different test set sizes to observe whether the total size of 342 instances in the study was sufficient or not.

IV. Results

The results showed that even some of the classifier algorithms seemed to be good learner for the training phase; they were over-fitting after the test phase. For instance, MultiClass classifier produced a correct classification rate in the training phase as 95.614%, however its performance degraded to 86.257 % in the test phase. Similarly, J48, Bagging and SMO classifiers seemed to be over-fitting. These results are also denoted in Table 6. On the other hand, even that the train and test set performances of LogitBoost and Dagging classifiers were not over-fitting, their learning performances were not accepted to be sufficient enough for the benchmark criteria (correct classification rate should be at least 90%) in this study. The other remaining five classifiers; Bayesian Network Learning (BayesNet), Lazy Bayesian Rules Learner (Lazy.LBR), Naïve Bayes Tree (NB Tree), Voting Feature Intervals Method (VFI) and Decision Table Naïve Bayes (DTNB) provided acceptable performance values regarding not only error rates but also Kappa Statistics, F-measures and ROC curves. However, it should be noticed that; NB Tree classifier might have a potential over-fitting threat due to the fact that; even it had produced a correct classification rate as 90.059%; there was a significant decrease when compared with its correct classification rate (95.614%) in the training phase. As mentioned in the previous sections, since there weren’t any alternative test data in this study; the over-fitting issue remained a question for NB Tree classifier.

Among all the eleven classifiers used in this study; Lazy.LBR and VFI classifiers generated the most accurate and satisfying results in terms of low error rates, high Kappa Statistics, ROC curve area and F-measures. These two classifiers also didn’t seem to be over-fitting where Lazy.LBR even provided higher performance values in its test phase with respect to its train phase. All these performance results are given in Table 6. Also, in the figures Figure 2 and Figure 3; ROC curves for Lazy.LBR and VFI classifiers derived from their 10-folds cross-validation test results are given. In these figures; y-axis denotes TP rate and x-axis denotes FP rate where the area beyond the curve gives the weighted average values of the plot area as a means of performance measure including cost of learning. In Figure 2, the ROC curve of Lazy.LBR is shown and this curve is derived from its TP / FP rate for "Risk = Yes" within the test phase with a ROC area value of 0.963. Similarly, In Figure 3, the ROC curve of VFI is given and this plot curve is derived from its TP / FP rate for "Risk = Yes" within the test phase with a ROC area value of 0.9651.

Also, a third experiment was also made with the five classifiers (BayesNet, Lazy.LBR, NB Tree, VFI and DTNB) that had produced acceptable performance values in the previous experiments. The aim of this experiment was to observe whether these classifiers provided a learning curve among different test set sizes. For each of the five classifiers, test set sizes of 57, 90, 114, 154, 205, 229, 256, 342 are used and their correct classification rates and F-measure rates are analyzed. The results showed that all of the five classifiers seemed to reach their maximum limits of their learning capacities when the test set size was increased up to 342 instances. All of these results are denoted within their learning curve progression in figures Figure 4, Figure 5, Figure 6, Figure 7 and Figure 8 for BayesNet, Lazy.LBR, NB Tree, VFI and DTNB, respectively. However, it should be mentioned that; due to the limited size of set samples in this study, these results do not assure whether these classifiers had reached to their maximum limits of their learning performances or not.

Classifier name / Phase	Correctly classified instances	Kappa Statistic	TP Rate (Risk = Yes)	TN Rate (Risk = No)	F-measure (weighted average)	ROC Area (weighted average)
BayesNet / Train	91.521 %	0.8214	0.915	0.915	0.916	0.972
BayesNet / Test	90.936 %	0.8091	0.911	0.907	0.91	0.962
Bagging / Train	94.152 %	0.8755	0.953	0.922	0.942	0.986
Bagging / Test	88.304 %	0.7518	0.901	0.853	0.883	0.951
LogitBoost / Train	90.059 %	0.7864	0.934	0.845	0.9	0.971
LogitBoost / Test	88.889 %	0.7642	0.906	0.86	0.889	0.957
MultiClass / Train	95.614 %	0.9068	0.962	0.946	0.956	0.993
MultiClass / Test	86.257 %	0.7115	0.869	0.853	0.863	0.91
Dagging / Train	91.228 %	0.8121	0.939	0.868	0.912	0.969
Dagging / Test	87.134 %	0.7236	0.911	0.806	0.871	0.94
Lazy.LBR / Train	91.228 %	0.8156	0.911	0.915	0.913	0.97
Lazy.LBR / Test	91.521 %	0.822	0.911	0.922	0.916	0.963
NB Tree / Train	95.614 %	0.907	0.962	0.946	0.956	0.985
NB Tree / Test	90.059 %	0.789	0.915	0.876	0.901	0.953
J48 / Train	93.567 %	0.8631	0.948	0.915	0.936	0.982
J48 / Test	85.965 %	0.7058	0.864	0.853	0.861	0.935
VFI / Train	92.105 %	0.8342	0.915	0.93	0.922	0.975
VFI / Test	91.228 %	0.8156	0.911	0.915	0913	0.965
SMO / Train	94.152 %	0.8767	0.939	0.946	0.942	0.942
SMO / Test	88.596 %	0.7577	0.906	0.853	0.886	0.879
DTNB / Train	93.567 %	0.8639	0.939	0.93	0.936	0.976
DTNB / Test	90.936 %	0.8085	0.915	0.899	0.91	0.959

Table 6. Comparative results of eleven different classifier algorithms for 342 instances.

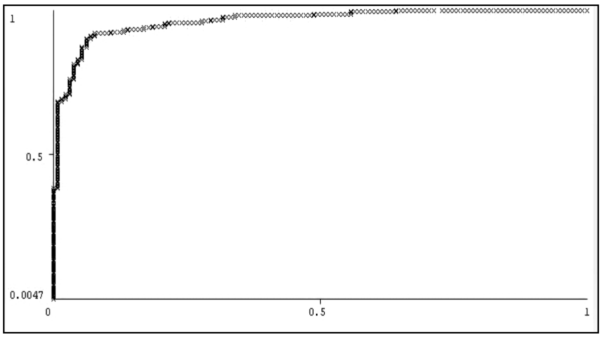

Figure 2. ROC curve for Lazy.LBR classifier

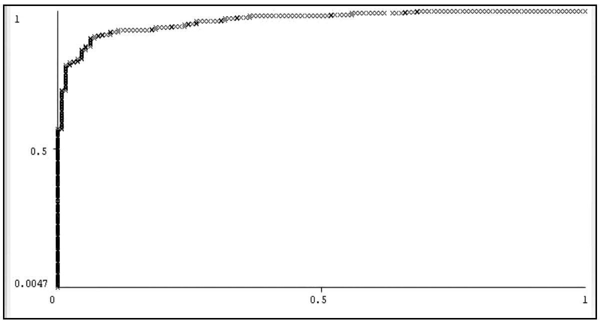

Figure 3. ROC curve for VFI classifier

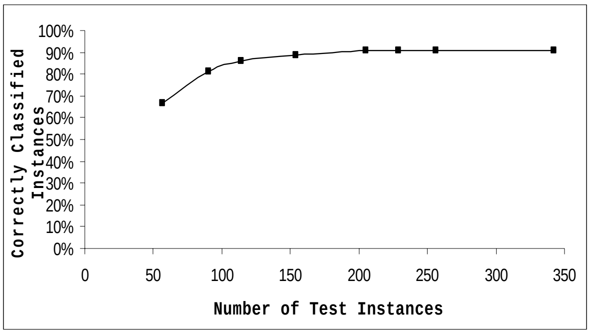

Figure 4. Learning performance of BayesNet

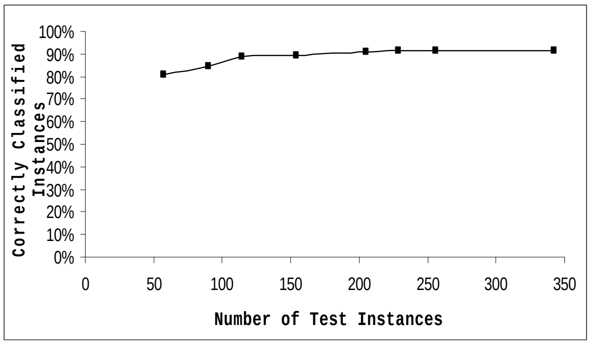

Figure 5. Learning performance of Lazy.LBR

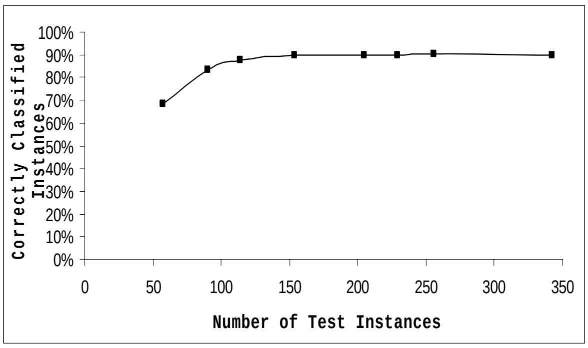

Figure 6. Learning performance of NB Tree

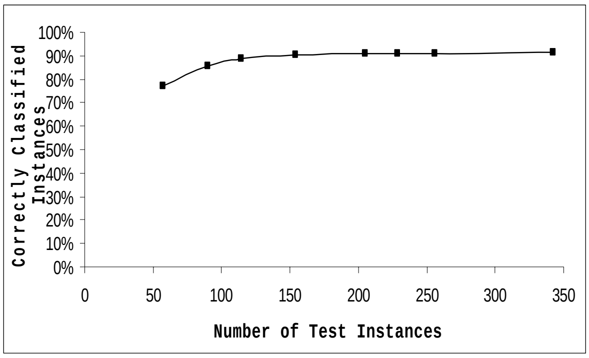

Figure 7. Learning performance of VFI

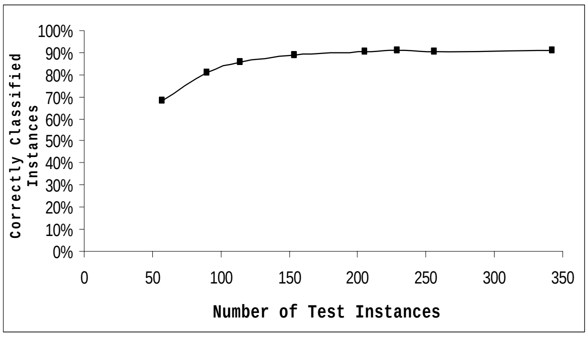

Figure 8. Learning performance of DTNB

V. Conclusions

In this paper, we developed a qualitative information security risk assessment methodology by the aid of machine learning classifier algorithms that was successfully implemented in the human resources department of a logistics company. Since the risk deduction is based on two parameters (risky and non-risky), binary classifier algorithms were proven to be a suitable model. Similar models and promising implementations can also be derived for other companies. In addition, based on this model; new models can be generated for other information security domains if risks are to be predicted by qualitative assessments.

However, it should be mentioned that; the data set size in this study was relatively small and could not be used for larger test sets. If it could be assured that over-fitting does not exist and learning curve has reached to its maximum level; then that learning algorithm might be used as a reliable and accurate information security risk assessment mechanism. Hence, if such models are to be generated from survey answers; then either the number of respondents must be increased or the same survey must be applied in several different companies or organizations. This would enable us to observe the performance of learning classifier algorithms with higher degrees of assurance using larger and more flexible data samples. Our research plan in the near future is to implement a similar study among several organizations with a much larger data set.

Another important issue in this study is the parameter selection and usage within the classifier algorithms. Some of the parameters regarding the five classifier algorithms (BayesNet, Lazy.LBR, NB Tree, VFI and DTNB) might be changed and additional results could be observed within the same data set. By this way, enhanced performance values for the learning capabilities of these algorithms might be obtained.

It should also be mentioned that there might be a subjectivity problem due to the values / scores provided by employees in qualitative risk assessments and relevant surveys. The probability and impact of this problem should be decreased as much as possible by means of additional mechanisms and methods in the risk assessment process. By this way, reliability and robustness of such qualitative risk assessment models might be improved.

Acknowledgement

Mete Eminagaoglu would like to thank Dr. Ajith Abraham for his contributions and recommendations during the discussion sessions regarding the preliminary work of this study which was presented in Proceedings of the 2010 International Conference on Computer Information Systems and Industrial Management Applications (CISIM 2010).

References

[1] S. L. Pfleeger, R. K. Cunningham, "Why Measuring Security is hard", *IEEE Security & Privacy*, IEEE, vol.8 (4), pp. 46-54, 2010.

[2] W. Qu, D. Z. Zhang, "Security Metrics, Models and Application with SVM in Information Security Management". In *IEEE Proceedings of the Sixth International Conference on Machine Learning and Cybernetics*, Hong Kong, 2007.

[3] J. J. Lv, W. H. Qiu, Y. Z. Wang, N. Zou, "A Study on Information Security Optimization Based on MFDSM". In *IEEE Proceedings of the Fifth International Conference on Machine Learning and Cybernetics*, Dalian, 2006.

[4] G. Giacinto, F. Roli, L. Didaci, "Fusion of Multiple Classifiers for Intrusion Detection in Computer Networks", *Pattern Recognition Letters*, Elsevier Science B.V., pp. 1795-1803, 2003.

[5] X. Long, Q. Yong, L. Qianmu, "Information Security Risk Assessment Based on Analytic Hierarchy Process and Fuzzy Comprehensive". In *IEEE Proceedings of the 2008 International Conference on Risk Management & Engineering Management*, 2008.

[6] D-M. Zhao, J-H. Wang, J. Wu, J-F. Ma, "Using Fuzzy Logic and Entropy Theory to Risk Assessment of the Information Security". In *IEEE Proceedings of the Fourth International Conference on Machine Learning and Cybernetics*, Guangzhou, 2005.

[7] A. Abraham, K. Smith, R. Jain, L. Jain, "Network and Information Security: A Computational Intelligence Approach", *Journal of Network and Computer Applications*, Elsevier Science, vol. 30 (1), pp. 1-3, 2007.

[8] T. Pietraszek, A. Tanner, "Data Mining and Machine Learning - Towards Reducing False Positives in

Intrusion Detection", *Information Security Technical Report*, Elsevier Ltd., v. 10, pp. 169-183, 2005.

[9] A. Abraham, C. Grosan, Y. Chen, "Cyber Security and the Evolution in Intrusion Detection Systems", *Journal of Engineering and Technology*, I-Manager Publications, vol. 1 (1), pp. 74-81, 2005.
[10] T. R. Peltier, Information Security Risk Analysis, Auerbach Publications, U.S.A., 2001.
[11] D. J. Landoll, The Security Risk Assessment Handbook - A Complete Guide for Performing Security Risk Assessments, Auerbach Publications, U.S.A., 2006.
[12] ISO, Information Security Management Systems - Requirements ISO/IEC 27001:2005, 2005.
[13] H. F. Tipton, M. Krause, Information Security Management Handbook, Auerbach Publications, 2007.
[14] ISO, *Information Security Risk Management ISO/IEC 27005:2008*, 2008.
[15] I. H. Witten, E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques 2nd ed.*, Morgan Kaufmann, U.S.A., 2005.
[16] G.M. Weiss, F. Provost, "Learning when Training Data are Costly: The Effect of Class Distribution on Tree Induction", *Journal of Artificial Intelligence Research*, AI Access Foundation and Morgan Kaufmann Publishers, U.S.A., pp. 315-354, 2003.
[17] M. Vuk, T. Curk, "ROC Curve, Lift Chart and Calibration Plot", *Metodoloski Zvezki*, vol. 3 (1), pp. 89-108, 2006.

Author Biographies

Mete Eminagaoglu graduated from the department of Computer Engineering, Ege University, Turkey in 1996. He received his MS degree in 1999 at the department of Computer Engineering, Izmir Institute of Technology, Turkey. Mete worked in private sector for fourteen years as a manager, auditor and consultant. He has CISSP and CISA certifications. He currently works as a lecturer at Yasar University, Turkey and he continues his PhD education at the Department of Computer Engineering, Trakya University in Turkey. His main research areas are information security risk modeling and assessment, machine learning, data mining, information security management strategies and cryptosystems.

Saban Eren is the Dean of Faculty of Science and Letters of Yasar University, Izmir, Turkey. He worked in the Computer Engineering Department of Ege University, Izmir in Turkey for more than fifteen years and he was also the chairman of the same department for three years. His current research interests include IT security, statistics and software reliability. He has authored some national and international publications in several areas such as statistics, software applications, computer programming and computational methodologies.