International Journal of Research in Engineering and Applied Sciences(IJREAS)
Available online at http://euroasiapub.org/journals.php
Vol. 12 Issue 01, January - 2022
ISSN(O): 2249-3905, ISSN(P) : 2349-6525 | Impact Factor: 6.573 |
Thomson Reuters ID: L-5236-2015

A STUDY ON FEATURE DISCRIMINATION FOR DISCERNING NOMINAL AND NUMERIC VARIABLES

Chandrasekharan Dinesh, Research Scholar, Department of Computer Science and Engineering, Kalinga University

Dr. Dev Ras Pandey, Professor, Department of Computer Science and Engineering, Kalinga University

ABSTRACT

The Internet's rapid expansion has made it much easier for individuals to share information globally, which has led to it being a vital tool in many fields, including biology, business, education, etc. Keeping the networks secure has grown essential and requires ongoing work as the Internet has become omnipresent in many facets of daily life. Numerous ML algorithms have the inherent ability to handle diverse types of data in various ways. Even though ML algorithms handle characteristics differently depending on their type, the algorithms have no way of knowing the type of the variable, hence it is the user's obligation to inform the learning algorithm of the variable type. We need to understand the different dynamics of the data, such as what forms the data can take, what the most common forms of data are, how to convert from one form to another, etc. before moving on and discussing the technique to predict the type for data variables. There is ample information in this regard in the few subsections that follow. The consideration of various factors' entropies will next proceed. However, the majority of ML techniques are made to function with numerical data. Some coding schemes are required for these algorithms to utilize the data from symbolic features in detection. The qualitative features were then replaced with numeric values using a coding technique of arbitrary assignment, which provides a correlation between each category of a symbolic feature and a sequence of integer. Any mathematical operation, other than a straightforward comparison, cannot be guaranteed to be valid by simply replacing the symbolic values with numbers.

KEY WORDS: Algorithms, Discrimination, Discerning, Nominal andNumeric Variables

International Journal of Research in Engineering and Applied Sciences(IJREAS)
Available online at http://euroasiapub.org/journals.php
Vol. 12 Issue 01, January - 2022
ISSN(O): 2249-3905, ISSN(P) : 2349-6525 | Impact Factor: 6.573 |
Thomson Reuters ID: L-5236-2015

INTRODUCTION

Given that the two have to be handled differently while calculating the similarity or difference of the entities constituted of such variables, we frequently find circumstances in mixed data analysis where it is essential to establish the variables as qualitative and quantitative. In these situations, it is frequently left to the skill of the adroit to manually classify the variables as nominal or numeric. An incorrect categorization of the variable will lead to incorrect (dis)similarity and have a detrimental effect on the final choice, whether it be classification or regression. In this study, we suggest a classification method based on ML for differentiating between the variables. The kind of variable is currently determined purely by the user's experience, as there is no clearly defined process for doing so. Therefore, a system with the exclusive responsibility of deciding the type of each variable should be present in any research effort that intends for the learning to be totally autonomous. The learning algorithm will then continue with its regular working process. In this approach, user-type assignment is less of a duty, and learning is totally automatic. The goal of this portion of the research is to identify a clear type assignment mechanism. Although there are many different forms of data, this work solely distinguishes between qualitative and quantitative data because they are the most common in the ML community.

TYPES OF DATA

Data may be considered of as a collection of different facts or pieces of information that are generally referred to as variables when it comes to statistical analysis. An identifiable piece of data with one or more values is called a variable. Those values may show up as text or a number (which could be converted into a number). Data can be collected in a variety of methods, of course, but a simple test can usually identify the result's typology with little difficulty. In the unlikely occasion that we need to quantify a sum associated with a certain occurrence, we compile data that takes quantitative elements into account. If we need to describe the quality of an observed phenomenon, we are gathering qualitative characteristics because we are unable to quantify it. As was already discussed, there are several different types of properties that can be utilized to represent real-life things. The following list includes the most common sorts of variables:

International Journal of Research in Engineering and Applied Sciences(IJREAS)
Available online at http://euroasiapub.org/journals.php
Vol. 12 Issue 01, January - 2022
ISSN(O): 2249-3905, ISSN(P) : 2349-6525 | Impact Factor: 6.573 |
Thomson Reuters ID: L-5236-2015

Variables using interval scales: Rough linear scale specifications exist for these variables. This kind of variable captures and depicts intervals or variations in values. These factors frequently include things like latitude, longitude, temperature, weight, and height. Nominal variables: These variables set one instance apart from another by having unique attribute names or values. Examples that are frequently used include employee ID, fingerprint, zip code, and gender. Binary Variables: Since computers can only store two values, binary variables are particularly common in the field of computer science. These variables can have one of the two possible values, 0 or 1, where the values signify the presence of a specific character. Category-specific variables Also known as qualitative variables, these variables can have a value selected from a limited and predetermined range. Categorical variables are distinguished from binary variables by the inclusion of more than two states, such as the protocol type, service, and intrusion classes in the case of the KDD99 data set. Ordinal Variables: Similar to categorical variables, these variables can have more than two states, but they also have a meaningful ordering of the classes. Examples include marks received on tests (such as A+, A, B+, B, C+, etc.) and height (e.g., tall, medium and short). Variables with a ratio scale are measurements that are positive on a non-linear scale, such as an exponential scale. Here, ratios and differences both have significance. Examples include Kelvin temperature, length, time, and counts.

NOMINAL TO NUMERIC CONVERSION

By now, it should be evident that the relevant data can be found in a wide variety of formats, including quantitative, qualitative, Date Time, etc. 1. As we've already established, the majority of ML algorithms are built so they can effectively handle data that is expressed as numbers. Therefore, the categorical symbols should be expressed as numbers before being fed any qualitative data to the ML system. A number of techniques have developed over time to address this issue. The most straightforward option is to simply ignore the qualitative factors, but doing so will result in the loss of relevant data. We provide a few of the methods that ML experts have employed over the years in the sections below.

International Journal of Research in Engineering and Applied Sciences(IJREAS)
Available online at http://euroasiapub.org/journals.php
Vol. 12 Issue 01, January - 2022
ISSN(O): 2249-3905, ISSN(P) : 2349-6525 | Impact Factor: 6.573 |
Thomson Reuters ID: L-5236-2015

DUMMY CODING

A common technique for transforming a categorical input variable into a continuous variable is dummy coding. Dummy is a duplicate variable that, as its name implies, represents one level of a categorical variable. A level's presence is indicated by 1 and its absence by 0, respectively. One dummy variable will be produced for each level that is present. Consider a data collection D with N dimensions and a categorical feature x that has n unique symbolic features. The category variable x's dummy coding with a binary string of length n will cause the data set to grow horizontally. The new data collection would therefore have N + n dimensions. The data set is sparse and storage space is wasted because there will only be one valid value for each location indicator for each record. This method was utilized by Yeung and Chow (2002) to translate 7 symbolic features from the KDD99 data set into numerical features. They did this to combine two data sets, one with 119 features and the other with 41 features.

LABEL ENCODER

It is employed to convert labels that are not numerical into ones that are (or nominal categorical variables). The range of a numerical label is always 0 to n classes 1. The label encoder has been used to carry out conversion. Nominal categorical variables frequently cause models to perform worse, which is a concern. For instance, we have the features "city" and "age" (which vary from 0 to 80). (81 different levels). Now that we've applied the label encoder to the "city" variable, "city" will be represented by a numeric value between 0 and 80. Since both variables will have comparable data points, the 'city' variable is now comparable to the 'age' variable, which is undoubtedly the wrong strategy.

USING ASCII

In reality, categorical values are just a series of characters, and ASCII is a character encoding scheme used to represent characters in computer systems. The character encoding standard used for electronic communication is called ASCII. In computers, telecommunications systems, and other devices, text is represented using ASCII codes. In the past, attempts have

International Journal of Research in Engineering and Applied Sciences(IJREAS)
Available online at http://euroasiapub.org/journals.php
Vol. 12 Issue 01, January - 2022
ISSN(O): 2249-3905, ISSN(P) : 2349-6525 | Impact Factor: 6.573 |
Thomson Reuters ID: L-5236-2015

been made to replace a category value by adding up its individual characters. According to Liu et al. (2004), categorical values were translated into their numerical counterpart by adding up the differences between each character's ASCII and that of uppercase A.

MOTIVATION

In a mixed data analysis, the data records are made up of many kinds of variables. They could be qualitative or quantitative, or both. The qualitative variables' various values are represented symbolically. Using the Protocol variable from the KDD data set as an example. The TCP, Internet Control Message Protocol (ICMP), and Internet Gateway Protocol are the three possible symbolic values for the Protocol variable (IGP). However, the majority of ML algorithms are created in a way that necessitates the expression of data as numbers. There are numerous methods for converting nominal values to numerical values, but one that is widely used in the ML community is to replace various symbolic values with an integer constant. In that scenario, TCP, ICMP, and Internet Group Message Protocol (IGMP) can all be substituted for one another. Even though applying mathematical operations to numerical quantities may be possible, doing so is utterly irrational; for example, removing TCP from ICMP or multiplying ICMP by IGMP have no significance.

K-NN

Let's look at a case of the lazy classifier k-NN in order to better comprehend the ideas. The goal of k-NN is to locate the closest neighbors for a given point x by measuring the distances between the objects and calculating those distances. A simple substitution of TCP with 1 or IGMP with 3 does not warrant the subtraction of the two, i.e., (3-1), which in turn means the difference (IGMP - TCP), which has no meaning because it is abundantly evident from the fact that the Euclidean Distance measure is only practicable for numerical data. As a result, we noted that in order to handle this situation, we should have a measure that is appropriate for handling mixed data, and one such metric that we discovered in the literature was the Gower metric. According to the Gower, it is defined in a method that treats the two properties in separate ways. It employs straightforward Euclidean for the numerical attributes and a straightforward comparison for the nominal qualities. Even though Gower and other mixed

dis(similarity) metrics handle the two sorts of characteristics differently, they lack the ability to forecast the type of variable and are wholly dependent on the programmer to do so. There is potential for error since the wrong type of variable assignment will have an impact on the dis(similarity) computation and, in turn, the wrong neighbor set retrieval, which will have an adverse effect on the classification or regression results. It is necessary to have an automated method for predicting the type of variable in order to make a metric completely unsupervised and to eliminate the possibility of wrong assignment.

WORKING HYPOTHESIS

Assume that X is a random variable that can have any one of N possible values. Let's assume that a feature X can have n distinct values overall, out of the N possible values. One of the working hypotheses for this study is that if the feature is quantitative, n approaches towards infinity.

- If the characteristic is qualitative, n tends to a finite constant (the number of modalities). In reality, there are always circumstances in which N > n. N cannot be infinite since there is always a limit to N.
- •N must not be too tiny in order to distinguish between the two types of variables since for small values of N, both types of variables will behave similarly.
- N can never be too big: If N Measurement limit or if N Media limit, n = f's behavior as it increases (N).
- The value of n is greater for the quantitative variables than the qualitative variables if we are in the optimal range of N. Hypotheses

Calculating a variable's entropy is a logical way to measure the combination of N and n. The measure of uncertainty for variables with too big N and too small n is called entropy. Since there are fewer types of data and more instances overall, there is a larger likelihood and, hence, a smaller degree of uncertainty regarding the occurrence of any given value, which suggests that there should be less entropy.

International Journal of Research in Engineering and Applied Sciences(IJREAS)
Available online at http://euroasiapub.org/journals.php
Vol. 12 Issue 01, January - 2022
ISSN(O): 2249-3905, ISSN(P) : 2349-6525 | Impact Factor: 6.573 |
Thomson Reuters ID: L-5236-2015

ENTROPY

Entropy is a metric measuring the state's randomness or, more specifically, its average information content. The negative logarithm of the probability mass function for the value serves as the unit of information entropy for each potential data value. The event therefore conveys more "information" (surprise) when the source data is of lower probability (i.e., when a low probability event occurs) than when the source data is of higher likelihood. When each event is specified in this way, the quantity of information it contains is transformed into a random variable whose expected value is the information entropy. Entropy typically refers to disorder or uncertainty, and the notion of entropy used in information theory and statistical thermodynamics are directly comparable. These early attempts were rationalized into a cogent mathematical theory of communication by Shannon's key work, which was based on works by (Shannon, 2001) and (Bromiley et al., 2004). This work also launched the field of study known as information theory.

ENTROPY SHANNON

Entropy is a term used to describe the degree of disorder in physical systems or the amount of knowledge that can be learned through studying disordered systems. Shannon Entropy, a formal definition of Entropy, was developed by Claude Shannon. The quantity of information H(p) included in a series of occurrences p1, p2,..., pn is constrained to satisfy three conditions:

- With every pi equally probable, H should be a monotonically rising function of N. • H should be continuous in pi.
- H ought to be combined.

RENYI ENTROPY

Numerous more measures of information or entropy have been developed as a result of extensions to Shannon's initial study. As an illustration, Renyi was able to expand Shannon Entropy to a continuous family of Entropy measures by waiving the third criteria of Shannon,

International Journal of Research in Engineering and Applied Sciences(IJREAS)
Available online at http://euroasiapub.org/journals.php
Vol. 12 Issue 01, January - 2022
ISSN(O): 2249-3905, ISSN(P) : 2349-6525 | Impact Factor: 6.573 |
Thomson Reuters ID: L-5236-2015

that of additivity. As a measure of diversity, the Renyi Entropy is significant in ecology and statistics.

TSALLIS ENTROPY

Shanon Entropy serves as the foundation for both Reyni and Tsallis entropy. The probability event's Tsallis Entropy, P = p1, p2, and pn, is defined as, given that 0 and 1 are constants.

DATA-SET FORMATION

A data-set was created by using features from different publicly available data-sets that were located in the UCL library. Our objective was to create a data collection that combined qualitative, that is, data that is represented using a symbolic scale We choose the properties throughout the data set if the feature is both qualitative and quantitative, i.e., data may be measured using a numerical or interval scale.

We extract and concatenate the numerous features from multiple databanks to create a sizable mixed-type features data-set in order to achieve a representative data-set. The mixed-type features in this new benchmark data-set make up the dataset. They include those that are nominal and qualitative (2477 attributes) and others that are numerical and quantitative (1698 features). Finally, it displayed a data set of 4175 mixed-type features that ranged in dimensionality from 4 to 2450. While quantitative features display numerical values on a clearly defined interval scale that can be continuous or discrete, qualitative features utilize symbols to denote categories. Symbolic elements regularly show up in the stream of network traffic data.

Following that, we got a data collection with all the attributes having numerical values; however, qualitative and quantitative data should not be mixed. We then generated three alternative entropies for each variable and recorded them in a separate file, along with the label for each row that indicated whether an attribute was nominal or numeric.

After that, the data set was standardized to the range [0-1] in order to remove the possibility that any attribute with higher values might predominate over attributes with lower values.

International Journal of Research in Engineering and Applied Sciences(IJREAS)
Available online at http://euroasiapub.org/journals.php
Vol. 12 Issue 01, January - 2022
ISSN(O): 2249-3905, ISSN(P) : 2349-6525 | Impact Factor: 6.573 |
Thomson Reuters ID: L-5236-2015

CLASSIFIER

After a data set has been created, the classification process is what comes next. Since each instance of the data set can be either nominal or numeric, as we have already mentioned, there are only two labels in the data set, making this a binary classification problem. The following step is to choose a classifier that is suitable for binary classification. Through a thorough analysis of the literature, we discovered that SVM is more effective for binary classification. The goal of SVM, which is to maximize the separation distance between two classes of objects, is to belong to the family of hyper-plane based learning methods. We also used additional classifiers in addition to SVM to offer a base for comparison between different classifiers. The several classifiers that have been utilized in this work are listed in Table 6.1 along with their setups. Chapter 3 provides a thorough examination of classifiers. 10 fold cross-validations was employed to evaluate the results.

Table 1: Classifiers

S no	Classifier	Configuration
1	*k*-NN	n = 5, Distance = Euclidean, No distance weighting, batch size = 100
2	Artificial Neural Network (ANN)	Activation = Sigmoid, Layers = 3, Nodes = 3,5,2, learning rate = 0.3, momentum=02
3	SVM	kernel = RBF, Gamma = 0.001 to 1000, eps = 0.001
4	Decision Tree	Confidence Factor = 0.25, Prunning = False
5	NaiveBayes	Batch Size=100, use Supervised Discretization = False, use Kernel Estimator = False

International Journal of Research in Engineering and Applied Sciences(IJREAS)
Available online at http://euroasiapub.org/journals.php
Vol. 12 Issue 01, January - 2022
ISSN(O): 2249-3905, ISSN(P) : 2349-6525 | Impact Factor: 6.573 |
Thomson Reuters ID: L-5236-2015

RESEARCH METHODOLOGY

The establishment of the data set marked the start of this work. There is currently no benchmark data-set for this problem because it is new. The data was manually taken from the UCL library. Each data collection in the UCL library includes a range of traits or properties, including qualitative, quantitative, binary, etc. Only nominal and numerical features were used. Since most ML algorithms only work with numeric data, the symbolic values were substituted with integer constants when taking into account the nominal properties. The range of a numeric label is always 0 to nclasses 1. After that, we determined the Shanon, Reyni, and Tsallis entropies for each variable. We placed a binary label indicating whether a variable was nominal or numeric in addition to storing the calculated Entropy values in a separate Excel sheet. The same is true for each and every variable. After the dataset has been prepared, we create a classification model utilizing different classification algorithms. In terms of the configuration of several classifiers, we chose the accepted and typical one. The data-set preparation phase includes feature extraction, nominal to numeric conversion, and normalization. The classification phase makes up the second phase. We employed 10-fold cross-validation to train and test the system, in which the data set is divided into ten subsets, nine of which are used for training and one is left over for testing. Ten times total are spent repeating the procedure. The average of the outcomes over all iterations is then calculated.

PERFORMANCE MEASURE

Any system should have some performance metrics that measure the effectiveness of the classification or clustering process. We employ the metrics that have been widely used for classification systems because our work is essentially a classification process where the goal is to determine if the characteristics are nominal or numeric.

RESULTS AND DISCUSSION

The features from the UCL library were manually extracted to create a labeled data set. Following that, the Shannon, Reyn, and Cialis entropies were determined for each variable. As a result, we used numerous classifiers on the data set. Five different classifiers in all were

International Journal of Research in Engineering and Applied Sciences(IJREAS)
Available online at http://euroasiapub.org/journals.php
Vol. 12 Issue 01, January - 2022
ISSN(O): 2249-3905, ISSN(P) : 2349-6525 | Impact Factor: 6.573 |
Thomson Reuters ID: L-5236-2015

used to categorize the data set shown. An procedure called 10-fold cross-validation is used to test the model. The classification model's efficacy was assessed using the performance measures listed. The classification outcomes obtained by different algorithms are shown. How precisely the model categorizes the quantitative and qualitative data is measured by the model's accuracy. SVM and ANN both have accuracy values of 99.951, with SVM having the best accuracy at 99.951. The results are the average of all the runs after applying SVM with RBF kernels to various gamma values. k-NN provides the lowest accuracy across all the data. The fact that k-NN is the simplest classifier and doesn't include a learning phase accounts for the reason for the poor performance. The success of k-NN depends critically on the size of the neighborhood and the choice of a distance metric.

Table 2: Classification Results

Classifier	Accuracy	Precision	Recall	f-Measure	ROC	MAE
k-NN	56.837	0.323	0.568	0.412	0.61	0.0514
NB	73.537	0.677	0.785	0.713	0.937	0.0203
DT	92.740	0.989	0.927	0.8802	0.963	0.0063
ANN	97.859	0.962	0.979	0.971	0.991	0.0047
SVM	**99.951**	**0.999**	**0.999**	**0.999**	**1**	**0.0001**

SVM performs better than other classifier models across the board, not just in terms of accuracy. SVM reports precision 0.999 and recall 0.999 better than all the other classification models, as seen. The two most common metrics for gauging the effectiveness of any classification model are recall and precision. The f-measure combines recall and precision into a single quantified inequality that should be as small as possible. SVM produces a ROC of 1, which is far higher than that of any other classifier. SVM has a Mean Absolute Error

(MAE) of 0.0001, which is significantly less than all other classifiers, with k-NN having the highest MAE at 0.0514. By now, it should be obvious that SVM is better able to reliably identify the data set than other models. SVM still misclassified a small number of examples, but overall, it performs much better than all of its rivals. The ability of SVM to handle large amounts of data is most likely why it performs better than other methods for the problem.

CONCLUSION

The two most common categories of data are nominal and numeric, and we discussed all of the different types of variables in length in this chapter. After that, we spoke about how to convert a nominal number to a numeric one. As a result, we talked about the need to distinguish between the two sorts of variables before proposing a strategy for automatically categorizing attributes. The results of the suggested model on the 1000 instances of produced data, 630 of which tend to be nominal and 370 of which tend to be numeric, demonstrate that we were able to accurately categorize 93% of the occurrences.

REFERENCES

1. Abbes, T., Bouhoula, A., and Rusinowitch, M. (2010). Efficient Decision Tree for Protocol Analysis in Intrusion Detection. International Journal of Security and Networks, 5(4):220–235.
2. Abduvaliyev, A., Pathan, A.-S. K., Zhou, J., Roman, R., and Wong, W.-C. (2013). On the Vital areas of Intrusion Detection Systems in Wireless Sensor Networks. IEEE Communications Surveys & Tutorials, 15(3):1223–1237.
3. Abubakar, A. I., Chiroma, H., Muaz, S. A., and Ila, L. B. (2015). A Review of the Advances in Cyber Security Benchmark Datasets for Evaluating Data-driven based Intrusion Detection Systems. Procedia Computer Science, 62:221–227.
4. Aburomman, A. A. and Reaz, M. B. I. (2016a). A novel SVM-kNN-PSO ensemble method for Intrusion Detection System. Applied Soft Computing, 38:360–372.
5. Aburomman, A. A. and Reaz, M. B. I. (2016b). Ensemble of Binary SVM classifiers based on PCA and LDA Feature Extraction for Intrusion Detection. In

Proceedings of IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference. 636–640.

6. Ahmed, M., Mahmood, A. N., and Hu, J. (2016). A Survey of Network Anomaly Detection Techniques. Journal of Network and Computer Applications, 60:19–31.
7. Akhbardeh, A. and Jacobs, M. A. (2012). Comparative Analysis of Nonlinear Dimensionality Reduction Techniques for Breast MRI Segmentation. MedicalPhysics, 39(4):2275–2289.
8. Alarifi, S. S. and Wolthusen, S. D. (2012). Detecting Anomalies in IaaS Environments through Virtual Machine Host System Call Analysis. In Proceedings of IEEE International Conference for Internet Technology And Secured Transactions. 211– 218.
9. Almusallam, N. Y., Tari, Z., Bertok, P., and Zomaya, A. Y. (2017). Dimensionality Reduction for Intrusion Detection Systems in Multi-data Streams — A Review and Proposal of Unsupervised Feature Selection Scheme. In Proceedings of Springer Emergent Computation. 467–487.
10. Amaral, J. P., Oliveira, L. M., Rodrigues, J. J., Han, G., and Shu, L. (2014). Policy and Network-based Intrusion Detection System for IPv6-enabled Wireless Sensor Networks. In Proceedings of IEEE International Conference on Communications. 1796–1801.
11. Ambusaidi, M. A., He, X., Nanda, P., and Tan, Z. (2016). Building an Intrusion Detection System using a Filter-based Feature Selection Algorithm. IEEE Transactions on Computers, 65(10):2986–2998.
12. Ariu, D., Tronci, R., and Giacinto, G. (2011). HMMPayl: An Intrusion Detection System based on Hidden Markov Models. Computers & Security, 30(4):221–241.
13. Belavagi, M. C. and Muniyal, B. (2016). Performance Evaluation of Supervised Machine Learning Algorithms for Intrusion Detection. Procedia Computer Science, 89:117–123.
14. Belkin, M. and Niyogi, P. (2003). Laplacian Eigenmaps for Dimensionality Reduction and Data Representation. Neural Computation, 15(6):1373–1396.

International Journal of Research in Engineering and Applied Sciences(IJREAS)
Available online at http://euroasiapub.org/journals.php
Vol. 12 Issue 01, January - 2022
ISSN(O): 2249-3905, ISSN(P) : 2349-6525 | Impact Factor: 6.573 |
Thomson Reuters ID: L-5236-2015

15. Bhattacharyya, D. K. and Kalita, J. K. (2013). Network Anomaly Detection: A MachineLearning Perspective. CRC Press.

16. Bhuyan, M. H., Bhattacharyya, D., and Kalita, J. K. (2011). NADO: Network Anomaly Detection using Outlier Approach. In Proceedings of ACM International Conferenceon Communication, Computing & Security. 531–536.

17. Bhuyan, M. H., Bhattacharyya, D., and Kalita, J. K. (2012). An Effective Unsupervised Network Anomaly Detection Method. In Proceedings of ACM International Conference