A. E. Asfha, A. Vaish, Оценка рисков информационной безопасности в отраслевой информационной системе на основе теории нечетких множеств и искусственной нейронной сети, *Информатика и автоматизация*, 2024, выпуск 23, том 2, 542–571

DOI: 10.15622/ia.23.2.9

DOI 10.15622/ia.23.2.9

A.E. ASFHA, A. VAISH

INFORMATION SECURITY RISK ASSESSMENT IN INDUSTRY INFORMATION SYSTEM BASED ON FUZZY SET THEORY AND ARTIFICIAL NEURAL NETWORK

Asfha A.E., Vaish A. **Information Security Risk Assessment in Industry Information System Based on Fuzzy Set Theory and Artificial Neural Network.**

Abstract. Information security risk assessment is a crucial component of industrial management techniques that aids in identifying, quantifying, and evaluating risks in comparison to criteria for risk acceptance and organizationally pertinent objectives. Due to its capacity to combine several parameters to determine an overall risk, the traditional fuzzy-rule-based risk assessment technique has been used in numerous industries. The technique has a drawback because it is used in situations where there are several parameters that need to be evaluated, and each parameter is expressed by a different set of linguistic phrases. In this paper, fuzzy set theory and an artificial neural network (ANN) risk prediction model that can solve the issue at hand are provided. Also an algorithm that may change the risk-related factors and the overall risk level from a fuzzy property to a crisp-valued attribute is developed. The system was trained by using twelve samples representing 70%, 15%, and 15% of the dataset for training, testing, and validation, respectively. In addition, a stepwise regression model has also been designed, and its results are compared with the results of ANN. In terms of overall efficiency, the ANN model (R2= 0.99981, RMSE=0.00288, and MSE=0.00001,) performed better, though both models are satisfactory enough. It is concluded that a risk-predicting ANN model can produce accurate results as long as the training data accounts for all conceivable conditions.

Keywords: risk, risk assessment, artificial neural network, fuzzy set theory, industry information system, cement industry.

1. Introduction. Over the past few decades, industrial digitalization has altered conventional procedures and practices in virtually every industry, and numerous digitalization solutions have been included in manufacturing assets [1]. The facility and processing of industry is no exception, and since the early 2000s, it has undergone a rapid digitalization process. For example, the Cement industry infrastructure in particular is subject to large and growing cybersecurity threats in the form of threat actors, vulnerabilities, and potential consequences.

Cybercriminals and others could potentially conduct cyberattacks against the industrial infrastructure, and all industries are always targets of malicious attacks. Modern exploration and production industry techniques depend more and more on remotely connected operational equipment, which is frequently essential for security and susceptible to cyberattacks. Because its operational technologies may have fewer cybersecurity protective measures, older infrastructure is equally prone to attack [2]. Thus, a successful cyberattack on industry infrastructure could cause physical, environmental, and economic harm.

Therefore, over time, the complexity of information systems is increasing, and the issues of information security are becoming increasingly important for any industry information system. Information security is concerned with protecting data, particularly electronic data, from unwanted use [3]. The security of the information at their disposal must be evaluated by every industry that uses information. Consequently, information security analysis is required. The first step in the risk management process is to assess the potential for information security breaches. The assessment of a system's information security or the design phase typically involves the analysis of information security threats [4]. Assessing the capability and efficacy of control mechanisms used on information technology components and the architecture of information systems in general is the primary goal of an information security evaluation.

An information security assessment includes many **tasks**, such as evaluating the effectiveness of the information processing system, evaluating the security of the technologies used, the processing process, and the management of the automated system [5]. The overarching goal of an information security assessment is to ensure the confidentiality, integrity, and availability of an organization’s assets. There are numerous risk assessment tools, and they can be used in either of two ways. Therefore, approaches for analyzing information security threats can be either quantitative or qualitative, depending on the outcome of their assessment. The numerical value of risk is produced by the algorithm of a quantitative technique [6]. Information concerning unfavorable or unexpected events in the information security system that could endanger the protection of information (information security incidents) is often gathered using the input data for evaluation. However, the results are less accurate and relevant because there are frequently insufficient statistics.

The use of overly basic scales with three degrees of risk assessment (low, medium, and high) makes qualitative procedures more prevalent. Experts are interviewed for the assessment, but there is still limited use of intelligent methods [7]. It is clear that both of the aforementioned choices have a number of fundamental flaws. In order to overwhelm them, the latest research focused on identifying alternative techniques that would be both more accurate and more adaptive, as the constant emergence of new sources of threats often renders existing approaches inaccurate and ineffective. Among the promising approaches are models based on solving uncertainty problems, such as fuzzy logic models and artificial neural networks (ANN).

Finally, fuzzy logic and artificial neural network approaches have been recommended as the appropriate tools to improve the industry

information system and may help analyze complex conditions. Thus, the main purpose of this paper is to evaluate risk values in a more reliable, flexible, and objective manner by using this proposed method and prioritizing the level of risk value.

1.1. Problem Descriptions. Every processing industry performs a large number of operations and tasks on a daily basis. Each activity and procedure comes with its own set of hazards, which must be identified and ranked. The sector has numerous difficulties and costs as a result of its failure to identify accessible dangers, which can lead to a lack of competitiveness, a lack of greatness, a loss of representative trust, and, ultimately, a departure from the basic goal of adequacy. Thus, the aim of this section is to identify the existing problems and evaluate the efficiency and accuracy of information security risk analysis output in industry information system.

One of the primary research problems in information security risk analysis in the industrial processing system is the lack of appropriate and standardized methodologies for industry risk analysis in different stages of the risk management process, especially the shortcomings of qualitative and quantitative risk analysis methodologies, as well as the use of old techniques. In short, the criticism of the approaches is as follows.

By ensuring that the limitations of one form of data are balanced by the strengths of another. Thus, using or integrating both a fuzzy inference system (FIS) and an artificial neural network (ANN) will result in more accurate and efficient results in industry processing systems.

1.2. Research Goals. This research paper aims to increase the efficiency and accuracy of information security risk analysis result in industry information systems by developing an ANN model for determining the risk of critical information security incidents based on an ISO 27005 standard. To achieve this goal, the following research objectives are set:

1.3. Research Objective. The objectives are listed below:

Obj. #1: Analysis of the existing and most recent risk analysis methods and tools in industry information systems.

Obj. #2: The authors have identified the different information security risks that may exist during the early developmental phases of the industrial system. Experts' opinions have been collated for compiling this list. Then develop a solution to address the identified problem(s).

Obj. #3: To design and implement a fuzzy inference system and artificial neural network (ANN) technique to estimate the information security risk in industry information systems.

Obj. #4: Evaluating the efficiency and accuracy of the proposed ANN model. To validate the applicability and effectiveness of the proposed ANN

model in industry information systems through fuzzy multiple regression modeling (MRM).

The aim of this paper is to develop a novel method for conducting risk assessments in industry information systems. Thus, this paper presented a fuzzy inference system and artificial neural network (ANN) model for estimating, evaluating, and prioritizing a more accurate and efficient risk level that minimizes the limitations of the existing methods.

2. Literature review. Existing risk assessment approaches mostly differ in the applied risk assessment scales: quantitative or qualitative. The output of the algorithm of the quantitative approach is the numerical value of the risk [5]. The information on unforeseen occurrences and threats is typically used as assessment input. But the frequent absence of adequate statistics reduces the sufficiency of the outcomes. The most prevalent qualitative processes, however, employ too straightforward scales that typically have three degrees of risk assessment (low, medium, and high). The assessment is conducted through expert interviews and the use of clever techniques is still insufficient [8]. Furthermore, such outcomes are not reusable.

Due to the aforementioned flaws, experts are actively seeking a method that would produce high-quality results while being able to adapt to the threat landscape's ongoing changes, omit ineffective and irrelevant expert assessments, and allow for the reuse of earlier assessments [9]. Although it takes a lot of time and intellectual energy, the fuzzy logic and artificial neural network (ANN) approach is the most promising way in this research because it addresses the problems with current approaches, notably in terms of flexibility and adaptability [10]. Additionally, the ANN has cognitive features like self-learn, making it possible to identify the optimum solution while gathering knowledge of both internal and external processes.

Fuzzy logic is a valuable method for dealing with complexity and uncertainty, providing a way to model the systems by simulating human thinking without relying on quantitative and qualitative data in computation [11]. Due to the ambiguous and complex nature of the characteristics, evaluating industrial information systems utilizing sustainable decision-making processes is difficult. Due to a lack of knowledge and a high degree of domain-related uncertainty, it is challenging to quantify risks using standard mathematics. Simple risk assessment, ranking, and prioritization based on the expertise, experience, and opinions of experts are made possible by fuzzy logic-based methodologies [12].

The key to fuzzy logic is to find appropriate fuzzy rules. For example, fuzzy IF-THEN rules are IF-THEN statements. The amount of

rules needed varies depending on the particulars of the problem [13]. Membership functions are used to characterize specific linguistic labels in a problem. The complexity of the components or information, the cross-interaction and effect between various elements, and the subjectivity of some aspects all contribute to the fuzziness in the risk assessment of the industry information system, making it challenging to precisely quantify and characterize. Fuzzy logic offers a more adaptable technique of evaluation because it doesn't rely on exact mathematical models to define and process problems [14].

Fuzzy logic can handle fuzzy and uncertain situations by introducing membership functions to characterize the relationships between variables and mapping variables to the interval between 0 and 1 [15]. In the risk assessment process, fuzzy logic is divided into fuzzy inference, fuzzy clustering, and fuzzy decision-making. Fuzzy inference is the process of deducing one or more conclusions from fuzzy rules, and it can solve the problem of uncertainty and vagueness in decision systems [16]. For instance, when customs officers receive report information most of which are inaccurate linguistic information, fuzzy logic can be used to make the information fuzzy and analyze the risk by fuzzy IF-Then rules.

To help customs agents make wiser decisions, it is helpful to model each data point to each cluster using membership functions to represent similarity degrees between data and clusters, fuzzy clustering is used to group data based on comparable features and thoroughly conduct risk assessment [17]. The process of choosing the best course of action from a variety of alternatives is known as fuzzy decision-making, and it can take both the abruptness and smoothness of variables into account. It can be a useful sensitivity analysis method for determining how variables interact with one another and how this affects the output results [18].

A collection of neurons that are organized into layers and placed in a particular configuration makes up an artificial neural network (ANN). A multilayer network is one that has an input layer, one or more hidden layers, and an output layer. The number of parameters that are provided to the network as input in the input layer corresponds to the number of neurons in the output layer. The neurons in the buried layer increase dimensionality and are in charge of feature extraction. They support activities like classification and pattern recognition [19].

The structure of ANN depicts a schematic of a fully connected, three-layer neural network consisting of input neuron layers (or nodes, units), one or more hidden neuron layers, and a final layer which consists of the output neurons. There are several approaches to categorize neural networks, with the training method-based classification being the most

common. A neural network is trained when it has had its weights, biases, and maybe other parameters updated. Once trained, ANNs may implicitly identify novel patterns and generalize output based on previously learnt patterns [20].

The two main categories of training techniques are supervised and unsupervised. While the unsupervised training of neural networks, also known as self-organizing maps, primarily uses the classification and clustering algorithms, the supervised training method enables learning based on feedback [21]. Unsupervised networks are those that are not given any feedback and are typically requested to categorize the input vectors into groups and clusters. They are widely used in the industry for lithology identification and well log interpretation. The majority of neural network applications in the industry sector, however, are based on supervised training methods [22].

The methods for assessing risk have significantly advanced, and neural network techniques are now often used. Neural networks are able to automatically learn and extract nonlinear correlations between input data through extended training on vast volumes of data because of the numerous components and their complicated relationships in the risk assessment of import and export firms [23]. Because of the neural network's adaptive characteristics, it is possible to recognize these complicated relationships and constantly alter the model parameters until the best result is reached. Back-propagation (BP), multilayer perception (MLP), recurrent neural network (RNN), and radial basis function network (RBF) are a few of the regularly utilized artificial neural network architectures.

3. Methodology. This research methodology was implemented to evaluate the efficient and more reliable risk analysis in industry information systems. In order to collect data, a questionnaire was developed to identify different risks. This method offers sufficient results for all the research questions and objectives of the study to be addressed. The relevant areas of data collection were identified, and interviews were conducted with different management and expert staff of the cement industry to secure an accurate account of information about the risks. An opinion was also made by the researcher so as to obtain useful information that will yield results that can address the problem identified in the study.

The participants in this study were experts and staff from different sections of the cement industry information system ($N = 81$). The participants were executive management, regular staff, technical and asset operators, and third-party consulting companies.

Participants were asked to evaluate twelve different information assets based on a scale of five points (one, two... and five) to estimate the

likelihood and consequence of the threat and group them into a five-point Likert scale (very low, low, average, high, and very high), as shown in Table 1. The collected data was analyzed to calculate the likelihood of related threats and their consequences. Some specialists in the field of cement industry information systems confirmed the reliability of the questionnaires.

Table 1. Likert-scale questionnaires

Likelihood				
Very Low	Low	Average	High	Very High
1	2	3	4	5
Consequence				
Very Low	Low	Average	High	Very High
1	2	3	4	5

3.1. Risk factors identification. Identifying the industrial information system risk factors, and this is the process of identifying, assigning, and characterizing the types of risks. All aspects of the risk assessment process are included.

Asset identification. In the process of identifying assets and their value, we consider the value placed on assets (including information). What work was required to develop them, how much it costs to maintain, what damage would result if it were lost or destroyed and what benefit another party would gain if it were to obtain it.

Vulnerabilities identification. It is a weakness or absence in information systems, system security procedures, internal control, or implementation that could be exploited by a threat of sources. So this means in short control is absent, not efficient, and no longer relevant…etc.

Threat Identification. After identifying the assets that require protection, the threat to those assets must be identified and examined to determine the loss. Finally, the estimation of the likelihood and consequence of risk factors based on vulnerability and threat identification based on data collection.

3.2. Fuzzy Inference System (FIS) Model. Because of the uncertainty of the risk factors, the fuzzy logic method and a fuzzy inference system are used in this study. First, membership functions are determined for all likelihood and consequence. Hence, it could be deduced that the membership function is a curve showing a point mapping points of inputting data into membership values, whose interval is between zero and one. Figure 1 shows the FIS process.

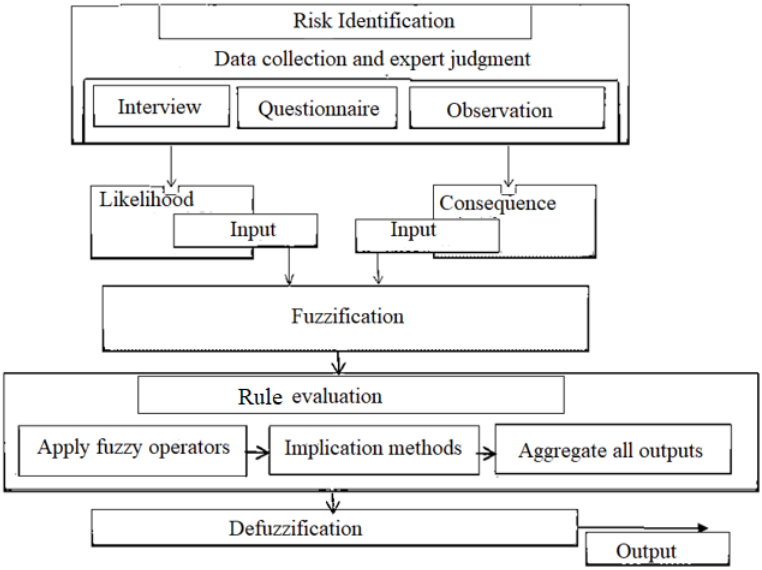


Fig. 1. Fuzzy inference system process

Fuzzification (Fuzzify Inputs). The first step is to take the inputs and determine the degree to which they belong to each of the appropriate fuzzy sets via membership functions (*fuzzification*) as noted in Table 2.

Table 2. Fuzzification table

Level	Linguistic Value	Fuzzy value
	Linguistic Variables (Likelihood of Security Risk Occurrence: 0-1)	
1	Very Low	(0.000, 0.125, 0.250)
2	Low	(0.200, 0.325, 0.450)
3	Averages	(0.350, 0.500, 0.650)
4	High	(0.550, 0.675, 0.800)
5	Very High	(0.750, 0.875, 1.000)
	Linguistic Variables (Consequence of Security Risk Occurrence: 0-10)	
1	Very Low	(0.000, 1.000, 2.000)
2	Low	(2.000, 3.250, 4.500)
3	Average	(3.500, 5.000, 6.000)
4	High	(5.500, 6.750, 8.000)
5	Very High	(7.500, 0.875, 10.000)
	Linguistic Variables (Security Risk Value: 0-1)	
1	Low	(0.000, 0.125, 0.250)
2	Very Low	(0.200, 0.325, 0.450)
3	Average	(0.350, 0.500, 0.650)
4	High	(0.550, 0.675, 0.800)
5	Very High	(0.750, 0.875, 1.000)

In this case, the likelihood (**L**) and consequence **(C) were** used as **crisp inputs (CI)** to the FIS (these values were taken from data collection and expert judgment).

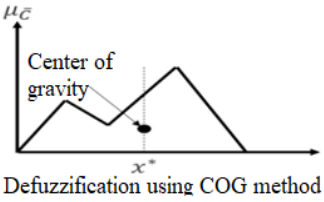
Fuzzy Rule. subsequently defining fuzzy membership functions, in this paper, Table 3 shows the 25 fuzzy rules constructed for the FIS.

Table 3. Risk matrix

Likelihood / Consequence	Very Low	Low	Average	High	Very High
Very Low	VL	VL	L	L	A
Low	VL	L	A	A	A
Average	L	A	A	H	H
High	L	A	H	H	VH
Very High	A	A	H	VH	VH
VL= Very Low, L=Low, A= Average, H= High, and VH=Very High					

Aggregation. It is the process of combining all of the fuzzy sets that symbolize each rule's outputs into a single fuzzy set. Interconversion occurs only once for each output variable, just prior to the final defuzzification phase.

Defuzzification. The last step in the fuzzy-molecular inference model is the defuzzification process, which is used to resolve a crisp value from the results of the inference process. Figure 2 indicates the defuzzification process using the center of gravity to finalize the FIS output.



Defuzzification using COG method

If $\mu_{\bar{C}}$ is defined with **continuous MF**: If μC is defined with **discrete MF**:

$$x^* = \frac{\int \mu_{\bar{C}}(x) \cdot x \, dx}{\int \mu_{\bar{C}}(x) \, dx} \qquad x^* = \frac{\sum_{i=1}^{n} \mu_{\bar{C}}(x_i) \cdot x_i}{\sum_{i=1}^{n} \mu_{\bar{C}}(x_i)}$$

Fig. 2. Defuzzification processes using the Center of Gravity Method

Table 4 presents the likelihood and consequence given by Expert 1 for each security risk factor. The same procedure is then repeated for 80 experts, and the knowledge database is created. Here, the authors have assumed that the data is normally distributed. We know that if the data are

assumed to be normally distributed, Table 5 also presents the risk factors of all 81 experts in Raw Material Processing (RMP)".

Table 4. Likelihood and Consequence Given by Expert 1 for Each Security Risk Factor

	Risk factor (Asset)	Coded Linguistic Variable			Numerical Value		
		L	C	Risk	L	C	Risk
RMP	Raw Material Processing	3	5	4	0.500	8.750	0.675
HRS	Hardware and Software	3	3	3	0.500	5.000	0.500
NWF	Network and Firmware	2	2	2	0.325	3.250	0.325
HRM	Human Resource and Data	5	4	5	0.875	6.750	8.750
RPT	Reputation	4	2	3	0.675	3.250	0.500
RMP	Raw Material Processing	3	3	3	0.500	5.000	0.500
ST	Storage and Transportation	5	4	5	0.875	6.750	0.875
RMM	Raw Material Milling	2	1	1	0.325	1.250	0.125
CP	Clinker Production	4	3	4	0.675	5.000	0.675
CM	Cement Milling	1	2	1	0.125	3.250	0.125

Table 5. Available data for risk of «**Raw Material Processing (RMP)**» in Database

No.	Numerical value			Coded linguistic variable		
	Likelihood	Consequence	Risk Level	L	C	Risk Level
1.	3 (Average)	5 (Very High)	4 (High)	0.500	8.750	0.675
2.	4 (High)	3 (Average)	4 (High)	0.675	5.000	0.675
3.	1 (Very Low)	3 (Average)	2 (Low)	0.125	5.000	0.325
4.	4 (High)	5 (Very High)	5 (Very High)	0.675	8.750	0.875
5.	3 (Average)	4 (High)	4 (High)	0.500	6.750	0.675
6.	4 (High)	1 (Very Low)	2 (Low)	0.675	1.250	0.325
7.	5 (Very High)	3 (Average)	4 (High)	0.875	5.000	0.675
8.	3 (Average)	3 (Average)	3 (Average)	0.500	5.000	0.500
9.	2 (Low)	2 (Low)	2 (Low)	0.325	3.250	0.325
10.	4 (High)	4 (High)	4 (High)	0.675	6.750	0.675
11.	1 (Very High)	2 (Low)	1 (Very Low)	0.125	3.250	0.125
12.	4 (High)	4 (High)	4 (High)	0.675	6.750	0.675
13.	3 (Average)	4 (High)	4 (High)	0.500	6.750	0.675
14.	5 (Very High)	5 (Very High)	5 (Very High)	0.875	8.750	0.875
15.	3 (Average)	1 (Very High)	2 (Low)	0.500	1.250	0.325
16.	5 (Very High)	5 (Very High)	5 (Very High)	0.875	8.750	0.875
17.	4 (High)	3 (Average)	4 (High)	0.675	5.000	0.675
18.	1 (Very Low)	2 (Low)	1 (Very Low)	0.125	3.250	0.125
19.	4 (High)	4 (High)	4 (High)	0.675	6.750	0.675
20.	3 (Average)	5 (Very High)	4 (High)	0.500	8.750	0.675
21.	2 (Low)	1 (Very Low)	1 (Very Low)	0.325	1.250	0.125
22.	1 (Very Low)	3 (Average)	2 (Low)	0.125	5.000	0.325
23.	3 (Average)	3 (Average)	3 (Average)	0.500	5.000	0.500
24.	4 (High)	1 (Very Low)	2 (Low)	0.675	1.250	0.325
25.	3 (Average)	4 (High)	4 (High)	0.500	6.750	0.675
26.	3 (Average)	2 (Low)	3 (Average)	0.500	3.250	0.500
27.	4 (High)	5 (Very High)	5 (Very High)	0.675	8.750	0.875
28.	3 (Average)	1 (Very Low)	2 (Low)	0.500	1.250	0.325
29.	5 (Very High)	3 (Average)	4 (High)	0.875	5.000	0.675

Continuation of Table 5

30.	3 (Average)	3 (Average)	3 (Average)	0.500	5.000	0.500
31.	4 (High)	4 (High)	4 (High)	0.675	6.750	0.675
32.	1 (Very Low)	3 (Average)	2 (Low)	0.125	5.000	0.325
33.	4 (High)	4 (High)	4 (High)	0.675	6.750	0.675
34.	4 (High)	4 (High)	4 (High)	0.675	6.750	0.675
35.	5 (Very High)	5 (Very High)	5 (Very High)	0.875	8.750	0.875
36.	1 (Very Low)	2 (Low)	1 (Very Low)	0.125	3.250	0.125
37.	5 (Very High)	5 (Very High)	5 (Very High)	0.875	8.750	0.875
38.	4 (High)	3 (Average)	4 (High)	0.675	5.000	0.675
39.	3 (Average)	2 (Low)	3 (Average)	0.500	3.250	0.500
40.	4 (High)	4 (High)	4 (High)	0.675	6.750	0.675
41.	1 (Very Low)	2 (Low)	1 (Very Low)	0.125	3.250	0.125
42.	3 (Average)	5 (Very High)	4 (High)	0.500	8.750	0.675
43.	2 (Low)	1 (Very Low)	1 (Very Low)	0.325	1.250	0.125
44.	3 (Average)	3 (Average)	3 (Average)	0.500	5.000	0.500
45.	1 (Very Low)	3 (Average)	2 (Low)	0.125	5.000	0.325
46.	4 (High)	5 (Very High)	5 (Very High)	0.675	8.750	0.875
47.	4 (High)	2 (Low)	3 (Average)	0.675	3.250	0.500
48.	3 (Average)	4 (High)	4 (High)	0.500	6.750	0.675
49.	3 (Average)	5 (Very High)	4 (High)	0.500	8.750	0.675
50.	4 (High)	1 (Very Low)	2 (Low)	0.675	1.250	0.325
51.	3 (Average)	4 (High)	4 (High)	0.500	6.750	0.675
52.	3 (Average)	3 (Average)	3 (Average)	0.500	5.000	0.500
53.	2 (Low)	2 (Low)	2 (Low)	0.325	3.250	0.325
54.	4 (High)	4 (High)	4 (High)	0.675	6.750	0.675
55.	2 (Low)	1 (Very Low)	1 (Very Low)	0.325	1.250	0.125
56.	4 (High)	2 (Low)	3 (Average)	0.675	3.250	0.500
57.	4 (High)	3 (Average)	4 (High)	0.675	5.000	0.675
58.	4 (High)	4 (High)	4 (High)	0.675	6.750	0.675
59.	5 (Very High)	2 (Low)	3 (Average)	0.875	3.250	0.500
60.	4 (High)	4 (High)	4 (High)	0.675	6.750	0.675
61.	3 (Average)	1 (Very Low)	2 (Low)	0.500	1.250	0.325
62.	3 (Average)	4 (High)	4 (High)	0.500	6.750	0.675
63.	2 (Low)	1 (Very Low)	1 (Very Low)	0.325	1.250	0.125
64.	3 (Average)	3 (Average)	3 (Average)	0.500	5.000	0.500
65.	1 (Very Low)	3 (Average)	2 (Low)	0.125	5.000	0.325
66.	5 (Very High)	2 (Low)	3 (Average)	0.875	3.250	0.500
67.	3 (Average)	3(Average)	3 (Average)	0.500	5.000	0.500
68.	2 (Low)	2 (Low)	2 (Low)	0.325	3.250	0.325
69.	4 (High)	4 (High)	4 (High)	0.675	6.750	0.675
70.	5 (Very High)	3 (Average)	4 (High)	0.875	5.000	0.675
71.	3 (Average)	3 (Average)	3 (Average)	0.500	5.000	0.500
72.	4 (High)	1 (Very Low)	2 (Low)	0.675	1.250	0.325
73.	3 (Average)	5 (Very High)	4 (High)	0.500	8.750	0.675
74.	2 (Low)	2 (Low)	2 (Low)	0.325	3.250	0.325
75.	3 (Average)	1 (Very Low)	2 (Low)	0.500	1.250	0.325
76.	2 (Low)	5 (Very High)	3 (Average)	0.325	8.750	0.500
77.	4 (High)	4 (High)	4 (High)	0.675	6.750	0.675
78.	1 (Average)	1 (Low)	1 (Average)	0.125	1.250	0.125
79.	4 (High)	2 (Low)	3 (Average)	0.675	3.250	0.500
80.	5 (Very High)	4 (High)	5 (Very High)	0.875	6.750	0.875
81.	5 (Very High)	4 (High)	5 (Very High)	0.875	6.750	0.875

Figure 3 notes the number of 25 if-then rules in order to provide a better understanding of the proposed fuzzy inference system framework, and with the input of the likelihood of occurrence and consequence, the risk size can be calculated. For instance, with 0.125 and 3.25 for likelihood and consequence, respectively, the risk size would be 0.125. A likelihood of 0.125 is related to rules 1–5, and a consequence of 3.25 is related to rules 2, 7, 12, 17, and 22.

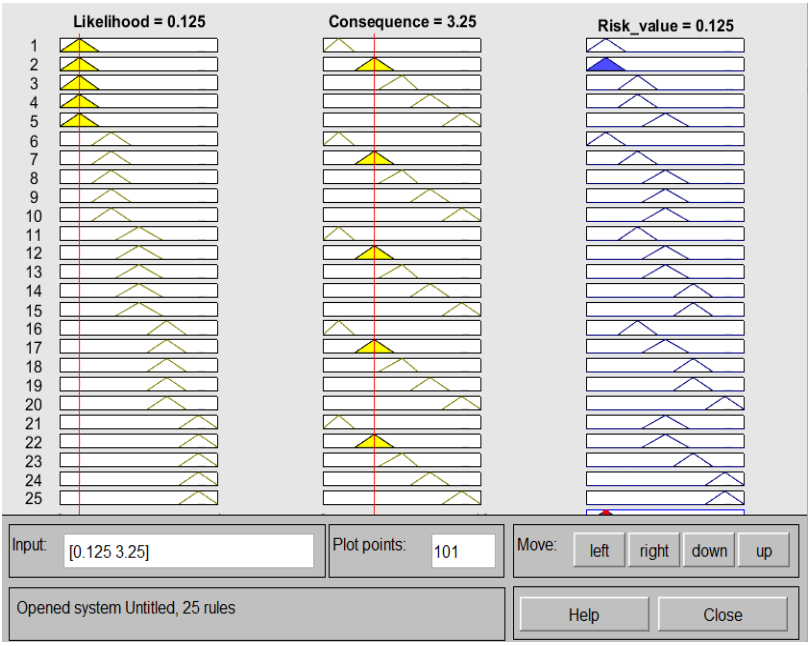


Fig. 3. Fuzzy rules according to Mamdani method

The fuzzy model designed by combining these rules estimates the risk value. The authors generated and plotted an output surface map for the industry information system fuzzy model using a surface viewer to visualize the dependence of one of the outputs on any one or two of the inputs. According to Mamdani, Figure 4 presents the processing industrial fuzzy model's output surface viewer.

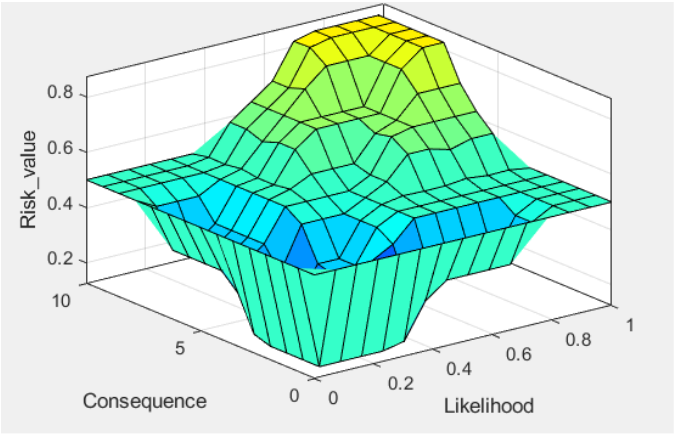


Fig. 4. 3D plots for 9 rules according to Mamdani method

Figure 5 indicates the normal probability illustration and the probability diagram of residuals for the criterion of "risk likelihood". Figure 6 indicates normal probability and residual illustrations for the criterion of "risk consequence" for the first factor "*Raw Material Processing (RMP)"*.

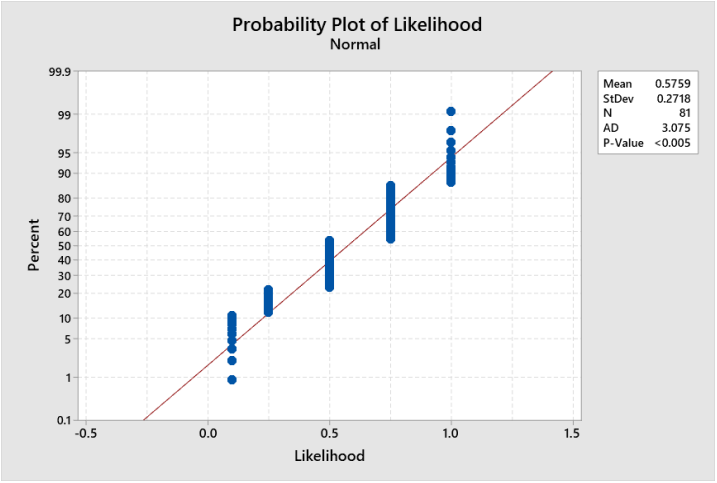


Fig. 5. Probability plot of likelihood

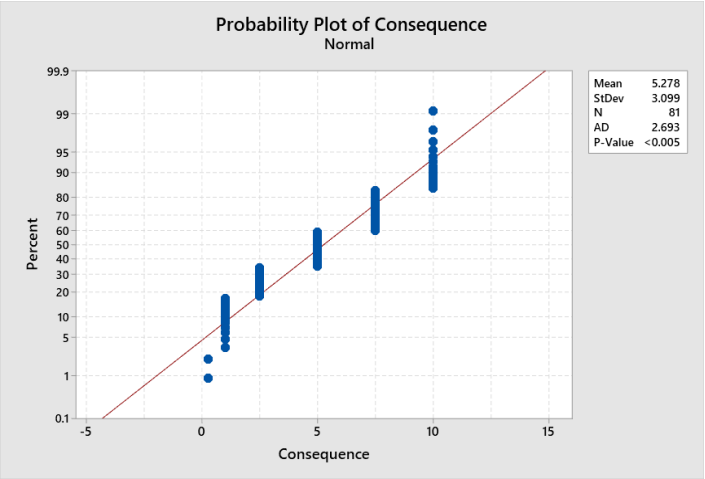


Fig. 6. Probability plot of consequence

4. Result and Discussion. This part uses a variety of statistical approaches to evaluate the quantitative data and provide the results of the data analysis in order to test the research hypotheses generated for the current study.

4.1. Data collection. Considering the chosen strategy of handing out the questionnaires to specific individuals one at a time, 95 were distributed. As a consequence, 81 of the 85 questionnaires received were complete and functional, yielding a response rate of 95.29%, which is regarded as excellent in research using a survey method and is displayed in Table 6. However, 10 employees failed to submit their surveys, and the remaining four representing 4.71% of the impractical forms were incomplete and contained inconsistent answers.

Table 6. The response rate of the participant

Questionnaire	Number	Percentage
Distributed	95	100 %
Received	85	89.47%
practical	81	95.29%
Impractical	4	4.71%

4.2. Performance evaluation. Minimum error occurrence has been considered as the basis for the selection of the best membership function. The performance of the designed fuzzy system has been evaluated on the basis of two types of errors, such as: – MSE (Mean Squared Error), and

RMSE (Root Mean Squared Error). According to the provided formulas, the correlation coefficient R between the data that were acquired and the data that ANN predicted has been determined (Equations (1) to (3)).

MSE (Mean Squared Error): it is the average squared difference between the value observed in a statistical study and the values predicted from a model.

$$\text{MSE}=\frac{1}{n}\sum_{t=1}^{n}(A_t - F_t)^2. \quad (1)$$

Root Mean Square Error (RMSE). It is a common method for calculating a model's error in predicting quantitative data. One of the most widely used indices in performance evaluations, the RMSE index, could explain the discrepancy between the model output and the real result. It is a non-negative number that has no upper bound and can be 0 when the projected and recorded outputs coincide exactly.

$$\text{RMSE}=\sqrt{\text{MSE}} \text{ (square root of MSE)}. \quad (2)$$

The correlation coefficient ($R^{2)}$ is a positive number that indicates how much of the variability in the dependent variable can be explained by the independent variable(s) and how well the model fits the data. $\mathbf{R^2}$ can take values between 0 and 1; 1 indicates the model can acquire all the variability of the output variable, while 0, which indicates a weak correlation between predicted and actual results, expresses this.

$$\text{R}=\frac{\sum_{t=1}^{n}(A_t-\bar{A})(F_t-\bar{F})}{\sqrt{\sum_{t=1}^{n}(A_t-\bar{A})^2 * \sum_{t=1}^{n}(F_t-\bar{F})^2}}, \quad (3)$$

$$\bar{A}=(\sum_{t=1}^{n} A_t)/n \text{ and } \bar{F}=(\sum_{t=1}^{n} F_t)/n,$$

where A_t, F_t, and n represent real data (Actual) data, estimate (Predicted) data, and the number of data, respectively.

4.3. Data prediction by ANN. In this research, a two-layer feed-forward with a backpropagation learning algorithm was used for the risk analysis model. Based on Figure 7, the input data consisted of 81 likelihood and consequence factors, and the output data from the FIS model was used as the target data to define the ANN output. To determine with ANN, the gray color (57= 70%) data points were selected for training, the green color (12=15%) for testing, and the remaining (12=15%) for data validation. The

number of hidden neurons was defined in different ways. The model was trained using Levenberg-Margardt with a backpropagation algorithm as noted in Table 7. In this paper, the authors used MATLAB software to evaluate the efficient results based on the ANN flowchart and FIS process. The outputs of the program which include the optimum membership functions for likelihood of occurrence, risk consequence, errors of training, test and validation, procedure of inference rules, and correlation between predicted data by network and training, test and validation data, are obtained.

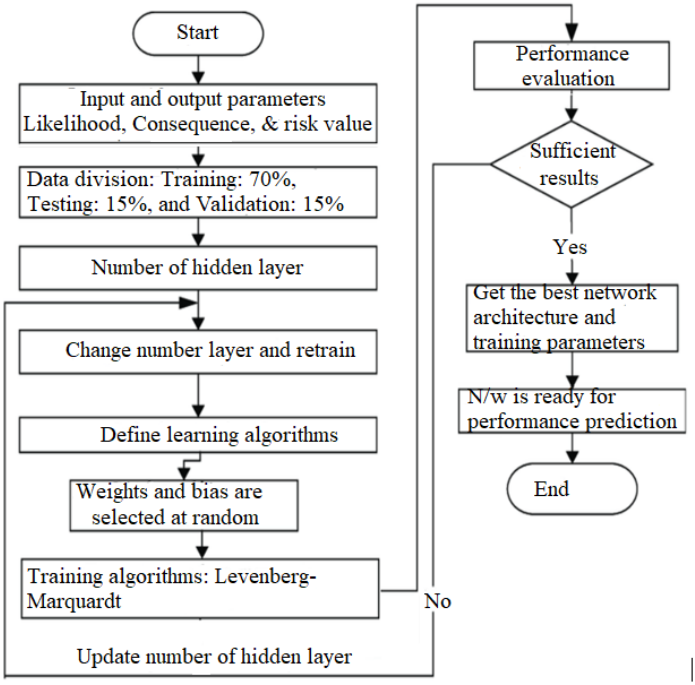


Fig. 7. ANN flowchart

Figure 8 indicates the function-fitting neural network. It is the process of training a neural network on a set of inputs in order to produce an associated set of target outputs. After you build the network with the preferred hidden layers and the training algorithm, you must train it using a set of training data. This research risk analysis was applied with different hidden layers of ANN (n = 10, 15, 25, and 50) and then the authors have selected the lowest error and best fit with the data.

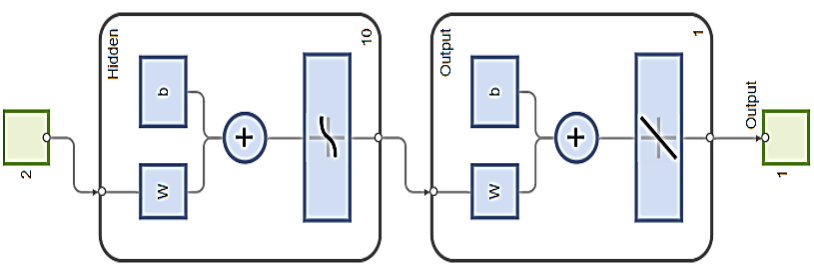


Fig. 8. Function fitting neural network (view)

Table 7. The specifications of the proposed ANFIS model

Parameters	Description/Value
Number of layers	3 (Input, output, and hidden layer)
Number of inputs(Predicators)	(2*81 double)
Number of outputs (Responses)	1 (1*81 double)
Hidden layer	10
Number of iteration	1000
Training Algorithm	Levenberg-Marquardt
Data Division	Random

The neural network regression has been shown in Figure 9, which demonstrates the interaction of the network with the training, test, and validation data. The correlation coefficient was found to be 1.00000, 0.99991, and 1.00000 for training, test, and validation data, respectively. Moreover, the straight line illustrates the linear relationship between the model-predicted and target output data. These results imply that there is a good match between the observed and model-predicted data. As a result, the model is adequate to forecast the data with high precision. The overall correlation coefficient (0.99998) confirms the outstanding prediction performance of the developed ANN model.

The plot for the best validation performance against the training data has been 6.9154e-18 at epoch 5 as shown in Figure 10. The circle in the plot clearly depicts that the validation plot lies exactly between the actual data plot and the observed data plot. Therefore, the research work is said to be validated.

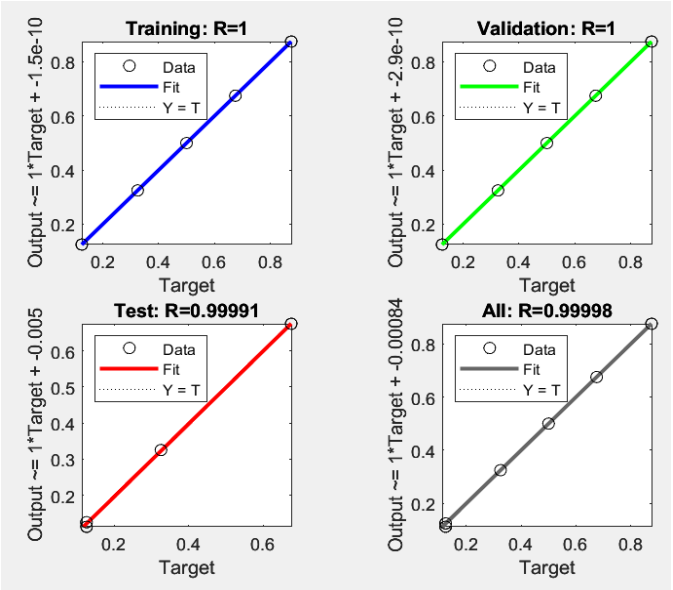


Fig. 9. ANN regression plot

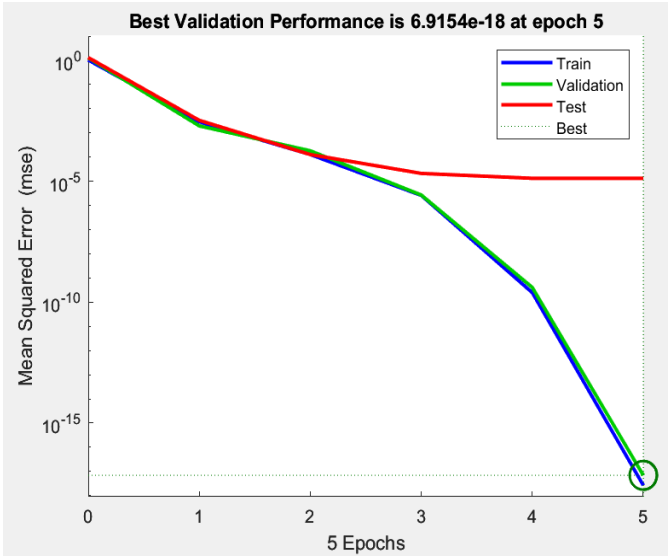


Fig. 10. ANN validation performance

Plotting the gradient values, mu, and validation fail has been shown in Figure 11. Gradient represents the slope of the tangent of a graph of a function. It points to the direction in which there is a high rate of increase for the considering function. The momentum constant or momentum parameter **(mu)** is the control parameter for the back-propagation neural network that we modeled, and the choice of mu directly affects the error convergence. **A validation check** is used to terminate the learning of the neural network. The number of validation checks will depend on the number of successive iterations of the neural network. Thus, gradient, mu, and validation check are 4.1263e-09, 1e-10, and 0 respectively at epoch 31 as shown in Figure 11.

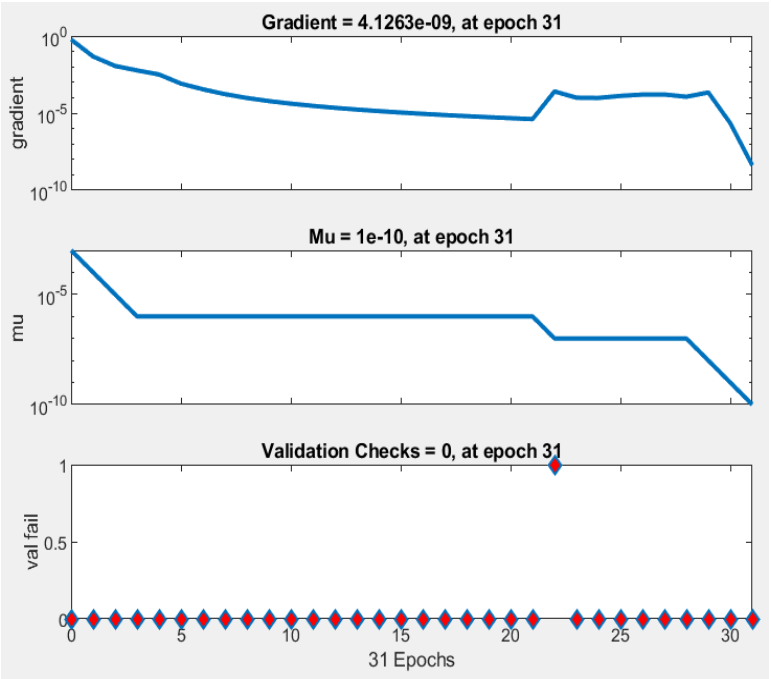


Fig. 11. ANN training state

Table 8 noted the information security risk prediction of "**Raw Material Processing (RMP)"**, and the coefficient of determination (R^2), RMSE, and MSE were also found. The results imply a good fit between the model-predicted data and the experimental data, indicating the models' aptness and coherence.

Table 8. ANN InfoSec risk prediction of «**Raw Material Processing (RMP)**»

				Risk Prediction
Risk factor (Asset)	**Likelihood**	**Consequence**	**InfoSec Risk**	**ANN**
Raw Material Processing (RMP)	0.675	6.750	0.675	0.674
	0.875	3.250	0.500	0.500
	0.675	6.750	0.675	0.675
	0.500	1.250	0.325	0.327
	0.500	6.750	0.675	0.675
	0.325	1.250	0.125	0.126
	0.500	5.000	0.500	0.500
	0.125	5.000	0.325	0.325
	0.875	3.250	0.500	0.500
	0.500	5.000	0.500	0.500
	0.325	3.250	0.325	0.325
	0.675	6.750	0.675	0.682
	0.875	5.000	0.675	0.675
	0.500	5.000	0.500	0.500
	0.675	1.250	0.325	0.325
	0.500	8.750	0.675	0.675
	0.325	3.250	0.325	0.325
	0.500	1.250	0.325	0.328
	0.325	8.750	0.500	0.500
	0.675	6.750	0.675	0.675
	0.125	1.250	0.125	0.113
	0.675	3.250	0.500	0.500
	0.875	6.750	0.875	0.875
	0.875	6.750	0.875	0.875
			RMSE	**0.00288**
			MSE	**0.00001**
			R	**0.99991**
			R^2	**0.99981**

4.4. Validation of InfoSec Risk analysis via Fuzzy Multiple Regression Modeling (MRM). A comparison of the findings acquired is necessary for confirming and validating the efficacy of the technique being used to solve any problem. The current method and the alternative procedures that were previously applied in the prior research investigations must be compared in this comparison. The authors used the ANN to evaluate the security risk in the aforementioned case study.

Multiple Regression Analysis (MRA) is a statistical technique that predicts the outcome of a response variable using a variety of explanatory variables. This technique will be heavily employed to represent the causal relationships between inputs and outputs. Equation 4 serves as a presentation of the multiple regression approach.

$$\boldsymbol{Risk = X_0 + X_1 * Likelihood + X_2 * Consequence}, \quad (4)$$

where X_0 is a fixed and X_1 and X_2 are regression coefficients.

The stepwise regression method has been applied for the first risk factor of *"Raw Material Processing (RMP)*" by using MINITAB 19 software to choose the best regression method for the prediction of risk size. Stepwise regression models have been presented in this paper. These models are shown in Tables 9 – 12.

Table 9. Correlation Coefficient among Input and Output Factors

	Likelihood	Consequence/Impact	**Security Risk**
Likelihood	1.0000		
Consequence	0.219	1.0000	
Security Risk	0.709	0.793	1.0000

Table 10. Consists of the Multiple Regression Equation for security risk through the hierarchy

Multiple Regression Equation		R^2
Security Risk Evaluation Model	**Regression Equation** Risk Level = -0.0419+ 0.5033 Likelihood + 0.05699 Consequence	0.93104 %

Table 11. Multiple Regression (MRL) Equations for each identified security risk factor

Risk Factor	**Multiple Regression Equation**	**RMSE**	**MSE**	R^2
RMP	-0.0715 + 0.5191 * Likelihood + 0.05997 * Consequence	**0.05250**	**0.00276**	**0.93104**
HRS	-0.0386 + 0.5843 * L + 0.05286 * C	0.05672	0.00322	0.91832
NWF	-0.0281 + 0.4858 * L + 0.05749 * C	0.03993	0.00159	0.97120
HRM	-0.0109 + 0.6597 * L + 0.05594 * C	0.04738	0.00224	0.96112
RPT	-0.0535 + 0.4693 * L + 0.06453 * C	0.04164	0.00173	0.96216
RMP	-0.0178 + 0.4941 * L + 0.05398 * C	0.06825	0.00466	0.90721
ST	-0.1065 + 0.5837 * L + 0.05915 * C	0.04015	0.00161	0.97082
RMM	-0.1356 + 0.6471 * L + 0.06000 * C	0.06106	0.00373	0.91012
CP	-0.0904 + 0.5987 * L + 0.05971 * C	0.04738	0.00225	0.95919
CM	-0.0130 + 0.5530 * L + 0.05104 * C	0.05168	0.00267	0.94401

Table 12. Risk prediction using the MRM model

				Risk Prediction
Risk factor (Asset)	**Likelihood**	**Consequence**	**InfoSec Risk**	**MRM**
Raw Material Processing (RMP)	0.675	6.750	0.675	0.683
	0.875	3.250	0.500	0.584
	0.675	6.750	0.675	0.683
	0.500	1.250	0.325	0.281
	0.500	6.750	0.675	0.594
	0.325	1.250	0.125	0.193
	0.500	5.000	0.500	0.495
	0.125	5.000	0.325	0.306
	0.875	3.250	0.500	0.584
	0.500	5.000	0.500	0.495
	0.325	3.250	0.325	0.307
	0.675	6.750	0.675	0.683
	0.875	5.000	0.675	0.683
	0.500	5.000	0.500	0.495
	0.675	1.250	0.325	0.369
	0.500	8.750	0.675	0.708
	0.325	3.250	0.325	0.307
	0.500	1.250	0.325	0.281
	0.325	8.750	0.500	0.620
	0.675	6.750	0.675	0.683
	0.125	1.250	0.125	0.092
	0.675	3.250	0.500	0.483
	0.875	6.750	0.875	0.783
	0.875	6.750	0.875	0.783
			RMSE	**0.05250**
			MSE	**0.00276**
			R	**0.96491**
			R^2	**0.93104**

4.5. Comparison between actual and model-predicted results. In this section, to prove the effectiveness of the proposed method, we compare our proposed algorithm with different methods. The authors compare our proposed ANN classifier with fuzzy regression modeling (MRM). The comparison and statistical analysis of the actual values and the model-predicted values of risk analysis in industry information systems are presented in Table 13. It was found that both models have sufficient capability to predict the properties of the industry.

Table 13. Comparison between actual and model-predicted results

				Risk Prediction	
Risk factor (Asset)	**Likelihood**	**Consequence**	**InfoSec Risk**	ANN model predicted	MRM model predicted
Raw Material Processing (RMP)	0.675	6.750	0.675	0.674	0.683
	0.875	3.250	0.500	0.500	0.584
	0.675	6.750	0.675	0.675	0.683
	0.500	1.250	0.325	0.327	0.281
	0.500	6.750	0.675	0.675	0.594
	0.325	1.250	0.125	0.126	0.193
	0.500	5.000	0.500	0.500	0.495
	0.125	5.000	0.325	0.325	0.306
	0.875	3.250	0.500	0.500	0.584
	0.500	5.000	0.500	0.500	0.495
	0.325	3.250	0.325	0.325	0.307
	0.675	6.750	0.675	0.682	0.683
	0.875	5.000	0.675	0.675	0.683
	0.500	5.000	0.500	0.500	0.495
	0.675	1.250	0.325	0.325	0.369
	0.500	8.750	0.675	0.675	0.708
	0.325	3.250	0.325	0.325	0.307
	0.500	1.250	0.325	0.328	0.281
	0.325	8.750	0.500	0.500	0.620
	0.675	6.750	0.675	0.675	0.683
	0.125	1.250	0.125	0.113	0.092
	0.675	3.250	0.500	0.500	0.483
	0.875	6.750	0.875	0.875	0.783
	0.875	6.750	0.875	0.875	0.783
			RMSE	**0.00288**	**0.05250**
			MSE	**0.00001**	**0.00276**
			R	**0.99991**	**0.96491**
			R^2	**0.99981**	**0.93104**

As represented in Figures 12, 13, 14 in terms of overall efficiency, the ANN model ($R^2 = 0.99981$, RMSE = 0.00288, MSE = 0.00001) performed better than the MRM model ($R^2 = 0.93104$, RMSE = 0.05250, MSE = 0.00276), though both are satisfactory enough. Figure 15 shows the time series plot of actual observed values versus the values predicted by the ANN and MRM models on the test dataset.

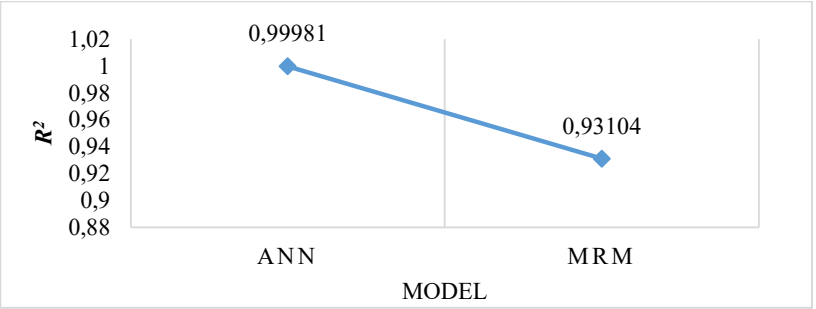


Fig. 12. Correlation Coefficient of ANN and Stepwise Regression

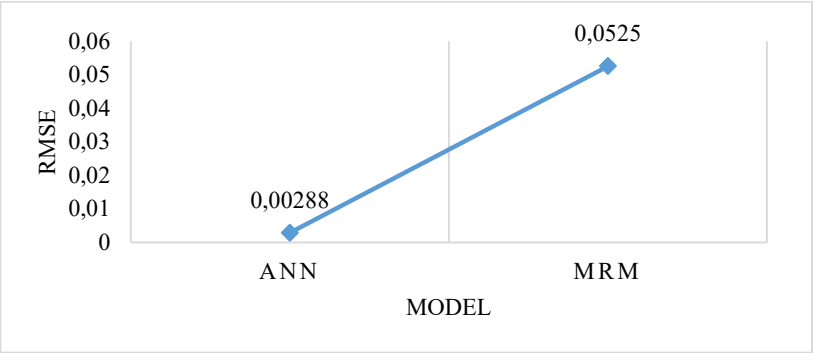


Fig. 13. RMSE of ANN and Stepwise Regression

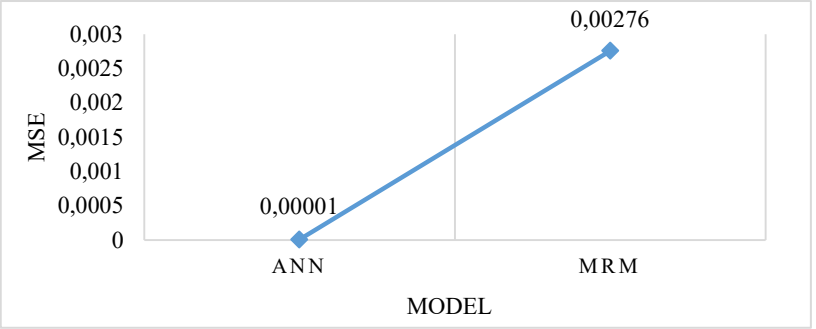


Fig. 14. MSE of ANN and Stepwise Regression

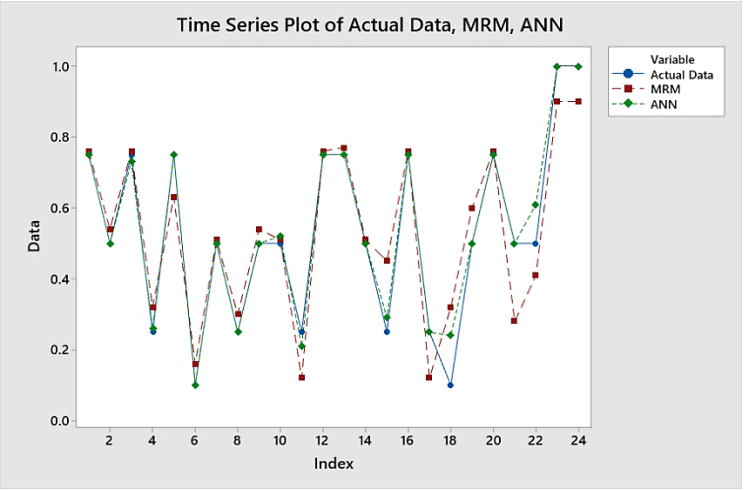


Fig. 15. Time Series Plot of ANN and MRM based on Actual data

5. Conclusion. Due to their shortcomings, both qualitative and quantitative methods are considered non-complete, subjective, including an element of randomness, and difficult to update or reuse. At this time many papers provide the necessary horizon scanning, focusing on AI-based methods, fuzzy logic, adaptive neural fuzzy inference system (ANFIS), and artificial neural networks (ANNs) and their usage for a more effective calculation of risk, considering the mix of qualitative input parameters such as likelihood and consequence. Thus, in this study, an information security risk assessment model based on fuzzy logic and an artificial neural network (ANN) is proposed to evaluate and calculate both qualitative and quantitative risks in a more reliable, flexible, and objective manner. The application of an artificial neural network can be used to assess information security risk since they have self-learn ability, can solve uncertain problems, and are appropriate for quantity data processing.

After fuzzy membership, functions are constructed for likelihood, consequence, and risk value. In order to obtain a more reliable and less subjective approach to the risk assessment process, an ANN has been used in this new model. Finally, in terms of overall efficiency, the ANN model (R^2= 0.99981, RMSE=0.00288, and MSE=0.00001,) performed better performance, though both models are satisfactory enough.

References

1. Verhoef P.C., Broekhuizen T., Bart Y., Bhattacharya A., Dong J.Q., Fabian N., Haenlein M. Digital transformation: A multidisciplinary reflection and research

agenda. Journal of business research. 2021. vol. 122. pp. 889–901. DOI: 10.1016/j.jbusres.2019.09.022.
2. Mazhar T., Irfan H.M., Khan S., Haq I., Ullah I., Iqbal M., Hamam H. Analysis of Cyber Security Attacks and Its Solutions for the Smart grid Using Machine Learning and Blockchain Methods. Future Internet. 2023. vol. 15(2). no. 83. DOI: 10.3390/fi15020083.
3. Alhassan M.M., Adjei-Quaye A. Information Security in an Organization. International Journal of Computer. 2017. T. 24. № 1. C. 100–116. [Online]. URL: https://ijcjournal.org/index.php/InternationalJournalOfComputer/article/view/820.
4. Shaikh F.A., Siponen M. Information security risk assessments following cybersecurity breaches: The mediating role of top management attention to cybersecurity. Comput. Secur. 2023. vol. 124. no. 102974. DOI: 10.1016/j.cose.2022.102974.
5. Cruz S.T. Information security risk assessment. Information Security Management Handbook. 2007. pp. 243–250. DOI: 10.3390/encyclopedia1030050.
6. Yevseiev S., Shmatko O., Romashchenko N. Algorithm of Information Security Risk Assessment Based on Fuzzy-Multiple Approach. Adv. Inf. Syst. 2019. vol. 3. no. 2. pp. 73–79. DOI: 10.20998/2522-9052.2019.2.13.
7. By I. et al. Implementing of qualitative risk assessment procedures. 2021. pp. 1–275.
8. Aven T. Risk assessment and risk management: Review of recent advances on their foundation. European Journal of Operational Research. 2016. vol. 253. no. 1. pp. 1–13. DOI: 10.1016/j.ejor.2015.12.023.
9. Tariq U., Ahmed I., Bashir A.K., Shaukat K. A Critical Cybersecurity Analysis and Future Research Directions for the Internet of Things: A Comprehensive Review. Sensors. 2023. vol. 23(8). no. 4117. DOI: 10.3390/s23084117.
10. de Campos Souza P.V., Lughofer E. Evolving fuzzy neural classifier that integrates uncertainty from human-expert feedback. 2023. vol. 14. pp. 319–341.
11. Bozic V. Fuzzy Approach to Risk Management: Enhancing Decision-Making Under Uncertainty. 2023. DOI: 10.13140/RG.2.2.13517.82405.
12. Kaka S., Hussin H., Khan R., Akbar A., Sarwar U., Ansari J. Fuzzy Logic-Based Quantitative Risk Assessment Model for Hse in Oil and Gas Industry. Journal of Tianjin University Science and Technology. 2022. pp. 93–109. DOI: 10.17605/OSF.IO/WVG2H.
13. Nikmanesh M., Feili A., Sorooshian S. Employee Productivity Assessment Using Fuzzy Inference System. Information. 2023. vol. 14(7). no. 423. DOI: 10.3390/info14070423.
14. Crnogorac L., Tokalic R., Gutic K., Jovanovic S., Dukanovic D. Fuzzy logic model for stability assessment of underground facilities. Podzemni radovi. 2020. no. 36. pp. 29–48. DOI: 10.5937/podrad2036029c.
15. Parra-Dominguez J., Alonso-Garcia M., Corchado J.M. Fuzzy Logic to Measure the Degree of Compliance with a Target in an SDG –The Case of SDG 11. Mathematics. 2023. vol. 11(13). no. 2967. DOI: 10.3390/math11132967.
16. Madanda V.C., Sengani F., Mulenga F. Applications of Fuzzy Theory-Based Approaches in Tunnelling Geomechanics: a State-of-the-Art Review. Mining, Metallurgy and Exploration. 2023. vol. 40. no. 3. pp. 819–837. DOI: 10.1007/s42461-023-00767-5.
17. Xie J., Deng Q., Xia S., Zhao Y., Wang G., Gao X. Research on Efficient Fuzzy Clustering Method Based on Local Fuzzy Granular balls. 2023. pp. 1–10. [Online]. URL: http://arxiv.org/abs/2303.03590.
18. Aliyeva K., Aliyeva A., Aliyev R., Ozdeser M. Application of Fuzzy Simple Additive Weighting Method in Group Decision-Making for Capital Investment. Axioms. 2023. vol. 12(8). no. 797. DOI: 10.3390/axioms12080797.

19. Alaloul W., Qureshi A.H. Data Processing Using Artificial Neural Networks. IntechOpen. 2020. 26 p. DOI: 10.5772/intechopen.91935.
20. Yang G.R., Wang X.J. Artificial Neural Networks for Neuroscientists: A Primer. Neuron. 2020. vol. 107. no. 6. pp. 1048–1070. DOI: 10.1016/j.neuron.2020.09.005.
21. Sarker I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Computer Science. 2021. vol. 2(3). no. 160. DOI: 10.1007/s42979-021-00592-x.
22. Zhang J., He Y., Zhang Y., Li W., Zhang J. Well-Logging-Based Lithology Classification Using Machine Learning Methods for High-Quality Reservoir Identification: A Case Study of Baikouquan Formation in Mahu Area of Junggar Basin, NW China. Energies. 2022. vol. 15. no. 10. DOI: 10.3390/en15103675.
23. Sarker I.H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. SN Computer Science. 2021. vol. 2(6). no. 420. DOI: 10.1007/s42979-021-00815-1.

Asfha Amanuel — Post-graduate student, Department of information technology security (FBI), ITMO University; Eritrea Institute of Technology. Research interests: information security methods and systems, information and cyber security, risk management. The number of publications — 4. baquesti2003@gmail.com; 49, Kronverksky Av., 197101, St. Petersburg, Russia; office phone: +7(952)378-2147.

Vaish Abhishek — Assistant professor, It department, Indian Institute of Information Technology, Allahabad. Research interests: information security, information security laws and regulations, cyber diplomacy, network security, IT Governance, enterprise recourses planning. The number of publications — 67. abhishek@iiita.ac.in; Uttar Pradesh, 211015, Deghat Jhalwa, India; office phone: +91(790)535-6150.

УДК 004.056 **DOI** 10.15622/ia.23.2.9

А.Э. Асфха, А. Вайш

ОЦЕНКА РИСКОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В ОТРАСЛЕВОЙ ИНФОРМАЦИОННОЙ СИСТЕМЕ НА ОСНОВЕ ТЕОРИИ НЕЧЕТКИХ МНОЖЕСТВ И ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ

Асфха А.Э., Вайш А. **Оценка рисков информационной безопасности в отраслевой информационной системе на основе теории нечетких множеств и искусственной нейронной сети.**

Аннотация. Оценка рисков информационной безопасности является важнейшим компонентом методов промышленного менеджмента, который помогает выявлять, количественно определять и оценивать риски в сравнении с критериями принятия рисков и целями, относящимися к организации. Благодаря своей способности комбинировать несколько параметров для определения общего риска традиционный метод оценки рисков, основанный на нечетких правилах, используется во многих отраслях промышленности. Этот метод имеет недостаток, поскольку он используется в ситуациях, когда необходимо оценить несколько параметров, и каждый параметр выражается различным набором лингвистических фраз. В этой статье представлены теория нечетких множеств и модель прогнозирования рисков с использованием искусственной нейронной сети (ANN), которые могут решить рассматриваемую проблему. Также разработан алгоритм, который может изменять факторы, связанные с риском, и общий уровень риска с нечеткого свойства на атрибут с четким значением. Система была обучена с использованием двенадцати выборок, представляющих 70%, 15% и 15% набора данных для обучения, тестирования и валидации соответственно. Кроме того, также была разработана пошаговая регрессионная модель, и ее результаты сравниваются с результатами ANN. С точки зрения общей эффективности, модель ANN (R2= 0,99981, RMSE=0,00288 и MSE=0,00001) показала лучшую производительность, хотя обе модели достаточно удовлетворительны. Делается вывод, что модель ANN, прогнозирующая риск, может давать точные результаты до тех пор, пока обучающие данные учитывают все мыслимые условия.

Ключевые слова: риск, оценка риска, искусственная нейронная сеть, теория нечетких множеств, отраслевая информационная система, цементная промышленность.

Литература

1. Verhoef P.C., Broekhuizen T., Bart Y., Bhattacharya A., Dong J.Q., Fabian N., Haenlein M. Digital transformation: A multidisciplinary reflection and research agenda. Journal of business research. 2021. vol. 122. pp. 889–901. DOI: 10.1016/j.jbusres.2019.09.022.
2. Mazhar T., Irfan H.M., Khan S., Haq I., Ullah I., Iqbal M., Hamam H. Analysis of Cyber Security Attacks and Its Solutions for the Smart grid Using Machine Learning and Blockchain Methods. Future Internet. 2023. vol. 15(2). no. 83. DOI: 10.3390/fi15020083.
3. Alhassan M.M., Adjei-Quaye A. Information Security in an Organization. International Journal of Computer. 2017. Т. 24. № 1. С. 100–116. [Online]. URL: https://ijcjournal.org/index.php/InternationalJournalOfComputer/article/view/820.
4. Shaikh F.A., Siponen M. Information security risk assessments following cybersecurity breaches: The mediating role of top management attention to

cybersecurity. Comput. Secur. 2023. vol. 124. no. 102974. DOI: 10.1016/j.cose.2022.102974.
5. Cruz S.T. Information security risk assessment. Information Security Management Handbook. 2007. pp. 243–250. DOI: 10.3390/encyclopedia1030050.
6. Yevseiev S., Shmatko O., Romashchenko N. Algorithm of Information Security Risk Assessment Based on Fuzzy-Multiple Approach. Adv. Inf. Syst. 2019. vol. 3. no. 2. pp. 73–79. DOI: 10.20998/2522-9052.2019.2.13.
7. By I. et al. Implementing of qualitative risk assessment procedures. 2021. pp. 1–275.
8. Aven T. Risk assessment and risk management: Review of recent advances on their foundation. European Journal of Operational Research. 2016. vol. 253. no. 1. pp. 1–13. DOI: 10.1016/j.ejor.2015.12.023.
9. Tariq U., Ahmed I., Bashir A.K., Shaukat K. A Critical Cybersecurity Analysis and Future Research Directions for the Internet of Things: A Comprehensive Review. Sensors. 2023. vol. 23(8). no. 4117. DOI: 10.3390/s23084117.
10. de Campos Souza P.V., Lughofer E. Evolving fuzzy neural classifier that integrates uncertainty from human-expert feedback. 2023. vol. 14. pp. 319–341.
11. Bozic V. Fuzzy Approach to Risk Management: Enhancing Decision-Making Under Uncertainty. 2023. DOI: 10.13140/RG.2.2.13517.82405.
12. Kaka S., Hussin H., Khan R., Akbar A., Sarwar U., Ansari J. Fuzzy Logic-Based Quantitative Risk Assessment Model for Hse in Oil and Gas Industry. Journal of Tianjin University Science and Technology. 2022. pp. 93–109. DOI: 10.17605/OSF.IO/WVG2H.
13. Nikmanesh M., Feili A., Sorooshian S. Employee Productivity Assessment Using Fuzzy Inference System. Information. 2023. vol. 14(7). no. 423. DOI: 10.3390/info14070423.
14. Crnogorac L., Tokalic R., Gutic K., Jovanovic S., Dukanovic D. Fuzzy logic model for stability assessment of underground facilities. Podzemni radovi. 2020. no. 36. pp. 29–48. DOI: 10.5937/podrad2036029c.
15. Parra-Dominguez J., Alonso-Garcia M., Corchado J.M. Fuzzy Logic to Measure the Degree of Compliance with a Target in an SDG –The Case of SDG 11. Mathematics. 2023. vol. 11(13). no. 2967. DOI: 10.3390/math11132967.
16. Madanda V.C., Sengani F., Mulenga F. Applications of Fuzzy Theory-Based Approaches in Tunnelling Geomechanics: a State-of-the-Art Review. Mining, Metallurgy and Exploration. 2023. vol. 40. no. 3. pp. 819–837. DOI: 10.1007/s42461-023-00767-5.
17. Xie J., Deng Q., Xia S., Zhao Y., Wang G., Gao X. Research on Efficient Fuzzy Clustering Method Based on Local Fuzzy Granular balls. 2023. pp. 1–10. [Online]. URL: http://arxiv.org/abs/2303.03590.
18. Aliyeva K., Aliyeva A., Aliyev R., Ozdeser M. Application of Fuzzy Simple Additive Weighting Method in Group Decision-Making for Capital Investment. Axioms. 2023. vol. 12(8). no. 797. DOI: 10.3390/axioms12080797.
19. Alaloul W., Qureshi A.H. Data Processing Using Artificial Neural Networks. IntechOpen. 2020. 26 p. DOI: 10.5772/intechopen.91935.
20. Yang G.R., Wang X.J. Artificial Neural Networks for Neuroscientists: A Primer. Neuron. 2020. vol. 107. no. 6. pp. 1048–1070. DOI: 10.1016/j.neuron.2020.09.005.
21. Sarker I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Computer Science. 2021. vol. 2(3). no. 160. DOI: 10.1007/s42979-021-00592-x.
22. Zhang J., He Y., Zhang Y., Li W., Zhang J. Well-Logging-Based Lithology Classification Using Machine Learning Methods for High-Quality Reservoir Identification: A Case Study of Baikouquan Formation in Mahu Area of Junggar Basin, NW China. Energies. 2022. vol. 15. no. 10. DOI: 10.3390/en15103675.

23. Sarker I.H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. SN Computer Science. 2021. vol. 2(6). no. 420. DOI: 10.1007/s42979-021-00815-1.

Асфха Амануэль Эстифанос — аспирант, факультет безопасности информационных технологий (ФБИТ), Университет ИТМО; Эритрейский технологический институт. Область научных интересов: методы и системы защиты информации, информационная и кибербезопасность, управление рисками. Число научных публикаций — 4. baquesti2003@gmail.com; Кронверкский проспект, 49, 197101, Санкт-Петербург, Россия; р.т.: +7(952)378-2147.

Вайш Абхишек — доцент, факультет информационных технологий, Индийский институт информационных технологий, Аллахабад. Область научных интересов: информационная безопасность, законы и нормативные акты в области информационной безопасности, кибердипломатия, сетевая безопасность, управление ИТ, планирование ресурсов предприятия. Число научных публикаций — 67. abhishek@iiita.ac.in; Уттар-Прадеш, 211015, Дегхат Джалва, Индия; р.т.: +91(790)535-6150.