CHAPTER VII

DESIGN EVALUATION AND VALIDATION METHODS

7.1 Overview

In this chapter, NFRAM and the safe path selection model are verified and discussed. Then, the prototype is evaluated based on expert review and user evaluation. Two methods are used to validate the proposed architecture: the analytical method for validating NFRAM and the safe path selection model and the experimental method for evaluating the prototype.

7.2 Analysis and Verification of Models

In this section, we present the verification data for both NFRAM and the safe path selection model.

7.2.1 Validation of Neuro-Fuzzy Risk Assessment Model

The performance parameter values of the combination of FL and ANNs techniques are compared in this section. The proposed model is validated using a dataset that contains 30 records, which have been determined by the researcher and the Deminers (experts), as shown in Table 7.1.

Table 7. 1: Comparison Among the Results of NFRAM and the Mamdani , the Sugeno Models

Seq	**Input** (*S^**) Signal Strength (dBm)	**Input** (*P^**) Position (m)	**Input** (*LI^**) Landmine Intensity (Klm^2)	**Output** (*RS^*%*) Sugeno Model	**Output** (*RS^*%*) Mamdani Model	**Output** (*RS^*%*) NFRAM
1	−68	70	692	20	20	25.3
2	−77.7	69.28	801	21.9	23.1	28.9
3	−78.46	69.28	704.8	23	25	29
4	−99	69	367	27	30.2	33.3
5	−73	68	680	29.5	34	36.9
6	−55	66	741	32.9	36.8	40.8
7	−111.3	87	741	35.5	35.5	46.3
8	−51.2	63.25	837.3	38.1	38.1	48.3
9	−98.1	66	644	44.2	48.1	48.4
10	−82.1	63	849	38.8	38.8	49
11	−110.1	100	944	40	40	50
12	−106	94	825	40	40	50
13	−80	70	140	50	50	50
14	82.26	10.24	439.8	50	50	50
15	−92.89	63.25	837.3	42.8	45.6	51.6
16	−93.2	62	800	47.4	49.1	55
17	−62	60	728	49.7	50.1	57
18	−95.92	62	897	54	53.3	58.3
19	−76	59	777	50.3	48.6	59.9
20	−79	58	873	52.4	52.4	63.1
21	−98	60	861	60.5	57.6	64.2
22	−95	59	704	60	58.5	65.3
23	−95	59.64	331.3	60.7	59.4	65.8
24	−91.4	68.1	404	65	60	75
25	−95.16	35.54	271	73.5	72	81
26	−55	33	873	76	76	86.3
27	−83.02	27.11	656.6	80	80	88.4
28	−89.09	30.72	873.5	79	79	88.7
29	−99.5	29.3	355.2	80	80	88.9
30	−57	15	897	80	80	90

Several stages of training patterns are implemented in 4D input–output space to decrease errors, which are the difference between the simulation outputs and desired outputs, by setting and adjusting inner neuron weights (MFs and rules).

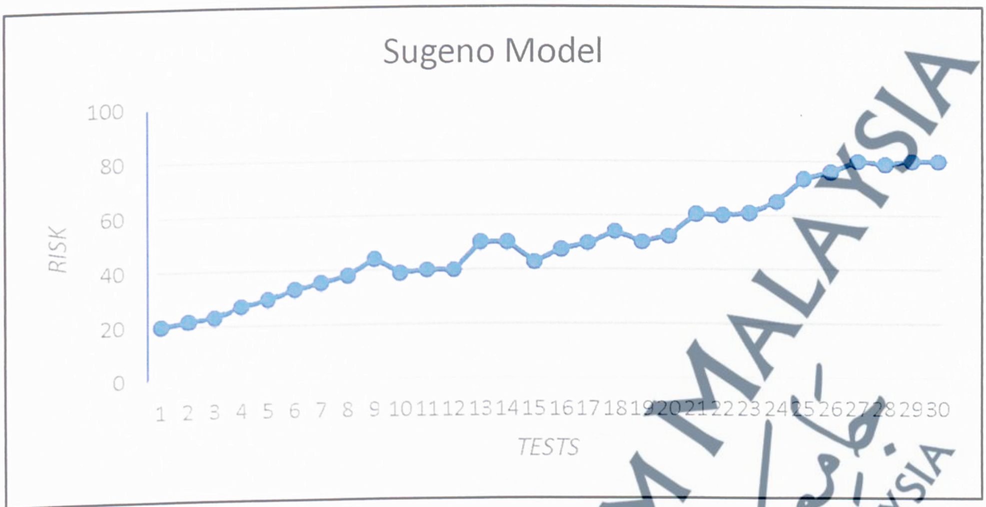
Each training pattern is determined by four variables: three input variables (signal strength, location and landmine intensity) and one output variable (risk). The input and output variables are represented by linguistic values, e.g. very low, low, moderate, high, very high). Through the training process, the weights of rules 2, 3, 6 and 7 are observed to approach 0, whereas the other weights remain high, thereby indicating that these rules are certainly false and can be removed from the neuro-fuzzy system.

Table 7.1 shows the training results. The first column presents the signal strength linguistic variable, the second the location linguistic variable, the third the landmine intensity linguistic variable, the fourth the output of the Sugeno model, the fifth the output of the Mamdani model and the last column the output of NFRAM.

The range of the output set is 0 to 100%. The FL technique is used in the validation procedure of NFRAM based on the Mamdani and Sugeno models. Figure 7.1 shows the risk value of landmines, which ranges from 20% to 80% using the Sugeno model on the dataset.

The results obtained from the model analysis are confirmed through the training test, which highlights that the landmine risk values are less scattered and the majority of the risk values are under moderate category.

Figure 7. 1 : Landmine Risk Values Generated Using the Sugeno Model



The results of the Mamdani model are extremely close to the output of the Sugeno model (Figures 7.2) because the two models use similar equations and structures through the fuzzy inference process.

Figure 7.2: Landmine Risk Values Generated Using the Mamdani Model

Figure 7.3 shows the risk value of landmines, which ranges from 25% to 90% using NFRAM on the dataset. The curve presented in Figure 7.3 shows that the rules directly influence the prediction outcome because the training process in the values of the variables will increase the prediction outcome. Thus, these rules are certainly false and can be removed from the neuro-fuzzy system.

Figure 7.3: Landmine Risk Values Generated Using NFRAM

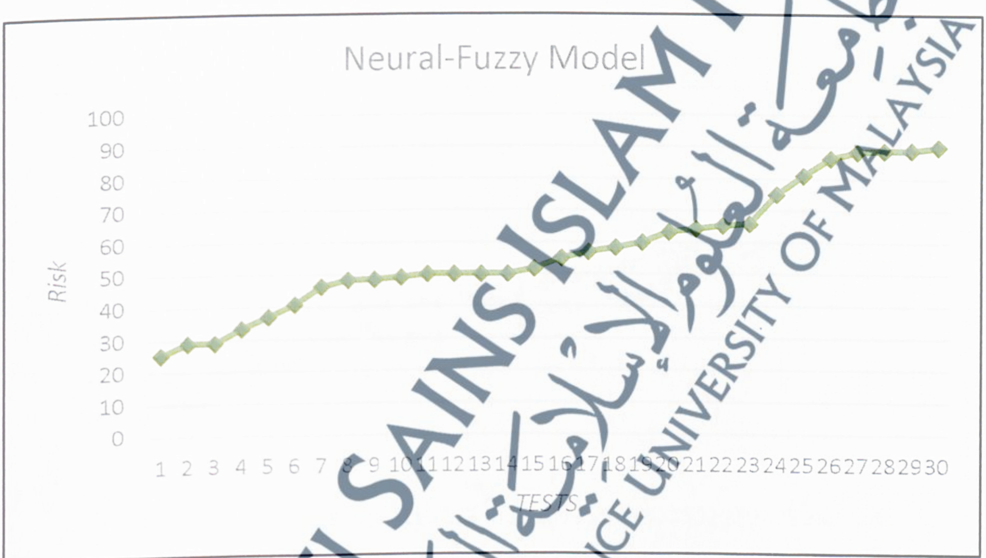
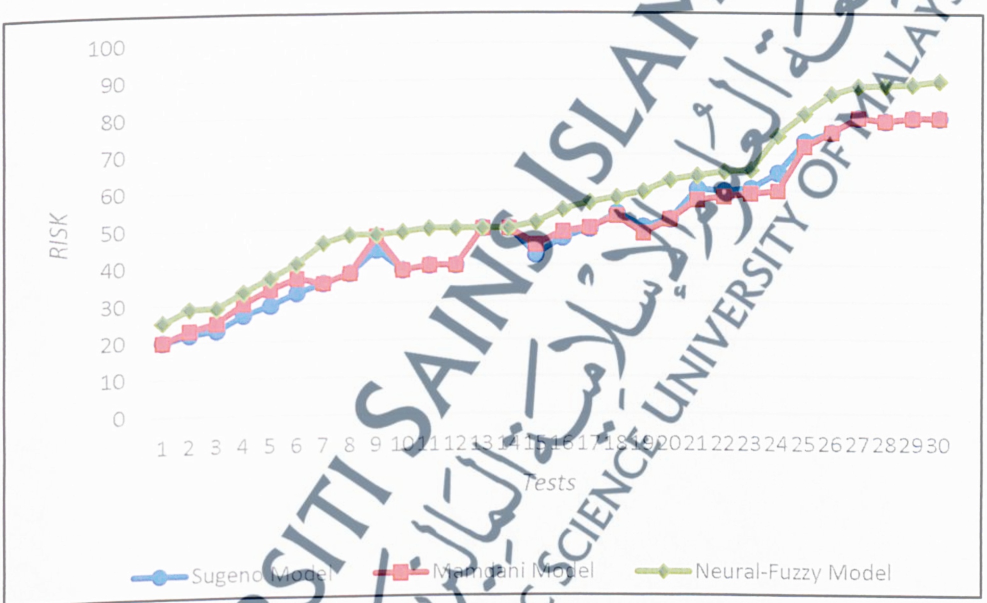


Figure 7.4 shows the comparison between the proposed NFRAM and the Mamdani and Sugeno models. As shown in Table 7.1 and Figure 7.4, an improvement in result is achieved by NFRAM compared with the results of the Mamdani and Sugeno models. NFRAM achieves a landmine risk classification accuracy of 10.8%. Such result suggests that the trained NFRAM system will most likely perform well on landmine datasets.

The landmine risk index, which is obtained through the latest aforementioned dataset (Table 7.1), is successively revised for a classification of landmine risk in linguistic terms (very low, low, moderate, high, very high) by using a cumulated frequency distribution curve (Figure 7.4) and selecting the points of the curve where a clear slope variation is observed as the limits of the classes.

Figure 7.4: Performance Comparison of the Proposed NFRAM Against the Mamdani and Sugeno Models



Surface viewer is used to generate and plot an output surface map for NFRAM and the Mamdani model. Figures 7.5 and 7.6 show the output surface viewer of NFRAM and the fuzzy Mamdani model. The output surface viewer shows a 3D curve that represents mapping from landmine intensity and position to risk degree.

In this section, the performance of the proposed NFRAM is investigated and compared with those of the Mamdani and Sugeno models based on the dataset selected by the researcher and the Deminers (experts). The validation results show that the proposed NFRAM provides better prediction results of landmine risks compared with the Mamdani and Sugeno models. We have created a more powerful and effective NFRAM by combining the advantages of FL and ANNs techniques, thereby improving the performance and operational effectiveness of tracking and ERA systems, which supports our claim in the problem statement.

Figure 7.5: Surface View with Inputs Position and Landmine Intensity with Their Expected Risk Using NFRAM

Colour	Level of Risk
Yellow	Very high
Cyan and green	Moderate
Blue	Very low

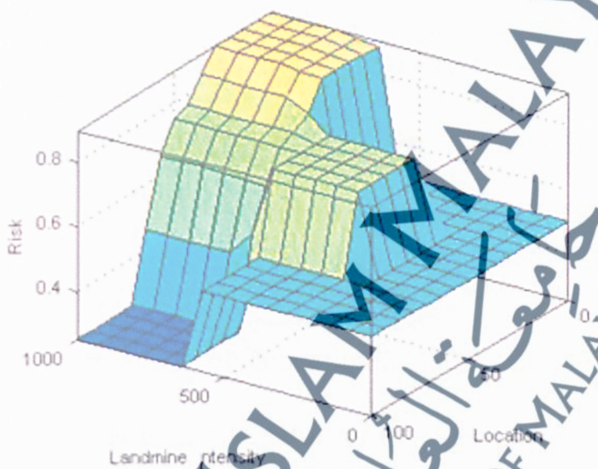
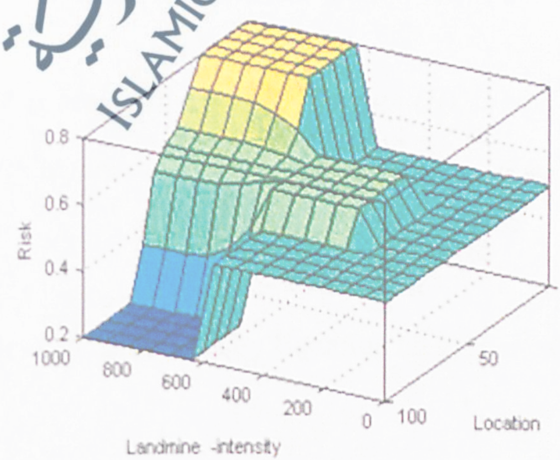


Figure 7.6: Surface View with Inputs Position and Landmine Intensity with Their Expected Risk Using the Mamdani Model

Colour	Level of Risk
Yellow	Very high
Cyan and green	Moderate
Blue	Very low



Mathematical Validation of the Model

The first step in the fuzzy inference process is fuzzification, which is the process of finding the membership value of a scalar (a number) in a fuzzy set. In this step, the crisp inputs *S*, P*, LI** (signal strength, location, landmine intensity) are taken and the degree to which these inputs belong to each appropriate fuzzy set is determined.

As an example, consider the current input (signal strength = −91.4), (location = 68.1) and (landmine intensity = 404). In this case, we rated signal strength as −91.4, which corresponds to $\mu = 0.76$ for an average MF given a graphical definition of poor; location is rated as 68.1, which corresponds to $\mu = 0.8$ for an average MF given a graphical definition of beside; landmine intensity is rated as 404, which corresponds to $\mu = 0.40117$ given a graphical definition of high.

To determine the control outputs of the fired rules, the rules that have been affected are *R1, R5, R8* based on MF.

R1: If S is average (Av), P* is inside (In) and LI* is high (H), then RS* is very high (VH).*

R5: If S is average (Av), P* is beside (Be) and LI* is high (H), then RS* is high (Hi).*

R8: If S is average (Av), P* is faraway (Farw) and LI* is high (H), then RS* is low (Lo).*

The strength of the fired rules can be defined by Equation (5.9). For example, for the first rule,

R1: If S is* average *(Av), P* is inside (In) and LI* is high (H), then RS* is very high (VH).*

$$\mu_{R^1(x,y)} = \mu AV(S^*) \wedge \mu In(p^*) \wedge \mu H(LI^*)$$

$$= min\,(0.76, 0.8544, 1)$$

$$= (0.76, \mu M(RS^*))$$

Consequently, all the control outputs can be calculated in the same way as that in the previous stage.

When Equations (5.10) and (5,11) are used, RS^* can be calculated as

$$RS^* = \{\,(0.76, \mu H(RS^*))\,, (0.8, \mu VH(RS^*))\,, (0.40117, \mu M(RS^*))\}.$$

The input of a fuzzy set (the aggregate output fuzzy set) and MFs are used to calculate a single output numerical value. This process is called defuzzification. When Equation (5.12) is used, RS^* can be calculated as follows:

$$\mathrm{RS}^* = \frac{-91.4 \times 0.76 + 68.1 \times 0.8 + 404 \times 0.40117}{0.76 \;\; + 0.8 + \;\; 0.40117} = 75\%.$$

7.2.2 Validation of the Safe Path Selection Model

To efficiently evaluate the safe path selection model, we introduce two pruning algorithms, namely, Dijkstra's algorithm and the Floyd–Warshall algorithm. We implemented these algorithms and compare the results with that of the proposed model. The objective of most algorithms is to find the shortest path from a given node to any other node on the graph and to minimize total time.

To produce an optimal itinerary, we will use an adjacency matrix *a* by defining a number of nodes that can represent a graph. The dimensions of the matrix will be equal to ($n \times n$), where n is number of vertices in the graph. The element of matrix *a[i][j]* is identified by an edge that connects the i-th and j-th vertices; the value in this case represents the weight of the corresponding edge. These edge (latitude and longitude) must be determined by experts (deminers) using special devices and then saved into the database. We have to consider that the value in *(a[i][j])* will be equal to infinity if no edge exists between vertices i and j.

A. Overview of the Algorithms

First, a short review is performed on Dijkstra's algorithm and the Floyd–Warshall algorithm. Dijkstra's algorithm is one of the classic shortest path search algorithms that works sequentially; however, it is not well suited for shortest path searches in large graphs. It attempts to decrease the value of the label of the vertices in each step, and thus, it stops after visiting all the vertices.

The temporal complexity is $O(|E|+|V| \log |V|)$, where $|E|$ is the number of edges and $|V|$ is the number of vertices of the graph. Several authors have proposed various modifications of Dijkstra's algorithm. Some of these algorithms use heuristics to reduce the run time of the shortest path search.

Dijkstra's algorithm consists of n iterations. If all the vertices have been visited, then the algorithm is terminated; otherwise, we have to choose the vertex with the minimum (smallest) value at its label from the list of unvisited vertices (in the beginning, we will choose a starting point *s*). Dijkstra's algorithm can be written in Java language, as shown in Figure 7.7. The Java programming language is used to conduct all the tests in a simulation environment.

Figure 7.7: Part of Dijkstra's Algorithm is Written in Java Language

```
for (int i = 0; i < n; i++) {
    // vertex
    int v = -1;
    // finding minimum label among unvisited vertices
    for (int j = 0; j < n; j++) {
        if (u[j] == false && (v == -1 || d[j] < d[v])) {
            v = j;
        }
    }
    // set as visited.
    u[v] = true;

    for (int j = 0; j < n; j++) {
        // if there is exists path between v and j vertices
        if (a[v][j] > 0) {
            if (d[v] + a[v][j] < d[j]) {
                d[j] = d[v] + a[v][j];
            }
        }
    }
}
```

The Floyd–Warshall algorithm is an example of dynamic programming. It was published in its currently recognized form by Robert Floyd in 1962. The Floyd–Warshall algorithm is used to find the shortest paths in a weighted graph with positive or negative edge weights (but without negative cycles). A single execution of the algorithm will derive the lengths (summed weights) of the shortest paths among all pairs of vertices.

Although the Floyd–Warshall algorithm does not return details of the paths themselves, the paths can be reconstructed through simple modifications of the algorithm. Suppose that the vertices are numbered from 1 to n in graph G, where notation d_{ijk} is the shortest path from i to j, which also passes through vertex k. An edge exists between vertices i and j. This edge will be equal to d_{ij0}; otherwise, it will be set as infinity. For other values of d_{ij0}, however, the shortest path from i to j does not pass through vertex k .

d_{ij0} = length of edge between vertices i and j

$d_{ijk} = min\ (d_{ijk-1},\ d_{ikk-1} + d_{kjk-1})$

The Warshall algorithm can be written in Java language as shown in Figure 7.8.

Figure 7.8: Excerpt of The Floyd–Warshall Algorithm Written in Java

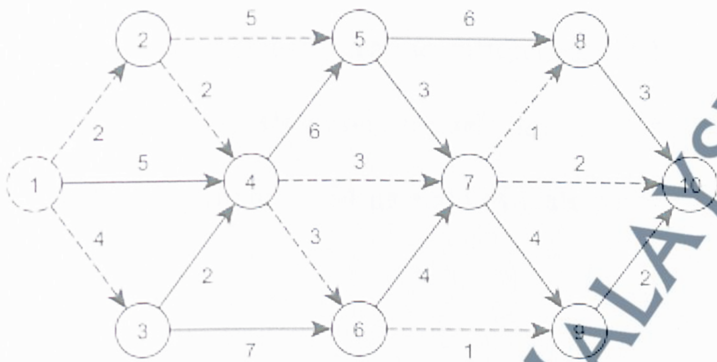
```
for (int k = 0; k < n; k++) {
    for (int i = 0; i < n; i++) {
        for (int j = 0; j < n; j++) {
            if (d[i][j] > d[i][k] + d[k][j]) {

                /*
                 * d[i][j] - is equal to
                 * the shortest path from i-th to j-th vertices.
                 */
                d[i][j] = d[i][k] + d[k][j];
            }
        }
    }
}
```


B. Graphs Used in the Simulation

To examine the efficiency and performance of the algorithms, manually directed graphs t, as shown in Figure 7.9, are used. A graph is created manually from 10 vertices that are joined together by 19 edges.

Figure 7.9: Custom Graph From 10 Vertices That are Joined Together by 19 Edges



C. Results of The Simulations of the Algorithms

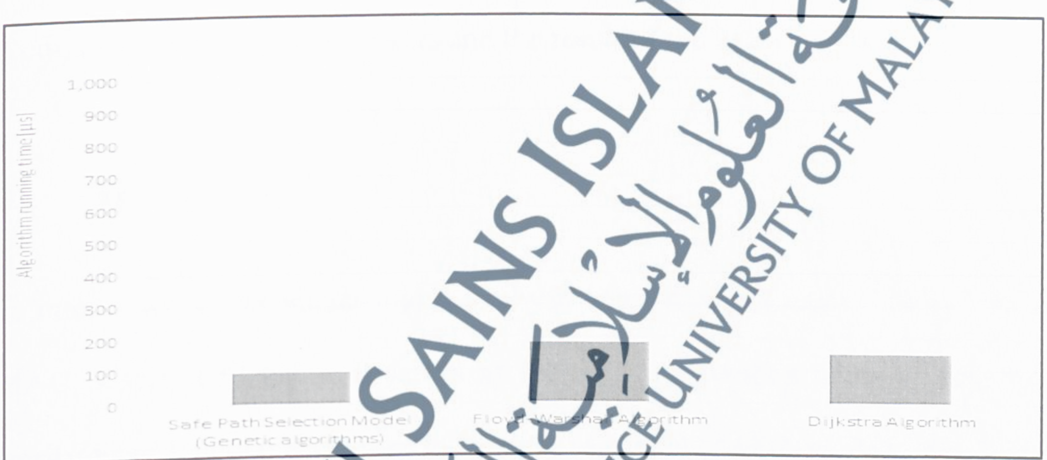
In the previous section, Dijkstra's algorithm and the Floyd–Warshall algorithm are examined to determine the efficacy of their operation in evaluating the safe path selection model. To obtain reliable results, we have tested each algorithm 100 times. The extreme results are rejected, and the average of the remaining results is calculated. The results of the performance of Dijkstra's algorithm, the Floyd–Warshall algorithm and the safe path selection model are presented in Table 7.2.

Table 7. 2: Comparison of the running times of Dijkstra's algorithm, the Floyd–Warshall algorithm and the safe path selection model (in μs).

	Algorithm	Algorithm running time (in μs)
1	Dijkstra	150
2	Floyd–Warshall	184
3	Safe path selection model (GAs)	96

The chart in Figures 7.10 provides the running times of the two algorithms and the safe path selection model (GA). On the basis of the previous simulation, each algorithm exhibits particular features that eventually lead to differences in their properties and performance. A slight difference exists among the results, i.e. Dijkstra's algorithm = 150 μs, the Floyd–Warshall algorithm = 184 μs and GA (safe path selection model) = 96 μs.

Figure 7.10: Comparison among Dijkstra's algorithm, the Floyd–Warshall algorithm and the proposed safe path selection model in terms of running times.



The proposed safe path selection model (GA) outperforms the two global algorithms (Dijkstra's algorithm and the Floyd–Warshall algorithm). The previous comparison shows that the safe path selection model exhibits the best performance by applying the proposed approach (GA) and obtaining the optimal path within a similar period or even in less time than when using Dijkstra's and the Floyd–Warshall algorithms is possible. Furthermore, different optimal solutions may be produced because the result can vary each time GA is executed.

7.3 Expert Review and User Evaluation of The Prototype

This section presents the heuristic evaluation strategies used to examine the proposed prototype to determine whether it is suitable for saving people's lives and avoiding the risk of landmines; the evaluation includes ease of use, usefulness, user satisfaction and reliability based on Nielsen, J. (1994 a, b). To achieve the three objectives of this study, as mentioned in Chapter 1, two activities (expert review and user evaluation) are conducted in the review process to evaluate the ERAA prototype, particularly its usability and user satisfaction. The user evaluation stage includes the interview session we conducted with a group of users and the result of the usability test.

7.3.1 Expert Review

The interviews were conducted at the Libyan Mine Action Centre (LMAC) in Libya. LMAC is selected because it serves as the lead organization in addressing weapon security according to the Ministry of Defence of Libya. LMAC operates outside Tripoli and relies on approximately 20 employees. Although no final decision has yet been made, expanding regional branches into affected areas are being planned. To validate the proposed prototype, an experiment was conducted with two experts working at LMAC and specialising in mine clearing [Expert A and Expert B, refer to Appendix (7)]. These experts have sufficient knowledge about landmines in Libya and over 10 years of experience.

Four sessions were held between the researcher and the experts to complete the task.

Questionnaires were distributed to the two experts. The instruments used for the expert review are usefulness, ease of use, user satisfaction and reliability. During the interview, the mine clearers were asked to use the prototype for testing and to evaluate it while doing the test tasks. The experts were asked several questions related to the usability and benefits of ERAA. Both experts tended towards strongly agree and they agreed that the prototype was satisfactory.

7.3.2 User Evaluation

A. Data Collection

The test was conducted among Libyan citizens who live in areas surrounding the capital Tripoli and in areas affected by landmines. The sample comprised 30 participants. Each testing session is divided into three parts as follows:

1) Brief introduction to the study
2) Prototype testing
3) Interview (questionnaire)

The objective of the study was presented to the participants during the first session. The respondents were asked to use the prototype for testing during the second session. They were then asked to evaluate the prototype while doing the test tasks. Finally, an interview that required the respondents to fill in a questionnaire [see Appendix (1)] was conducted during the last session. All the data collected during the testing and interview sessions will be analysed to determine the findings of this study.

B. Data Analysis

SPSS software (version 11.5) is used to analyse the data collected through the questionnaire. During the analysis session, different statistics are used for data analysis. The results obtained through data analysis are presented in the following section.

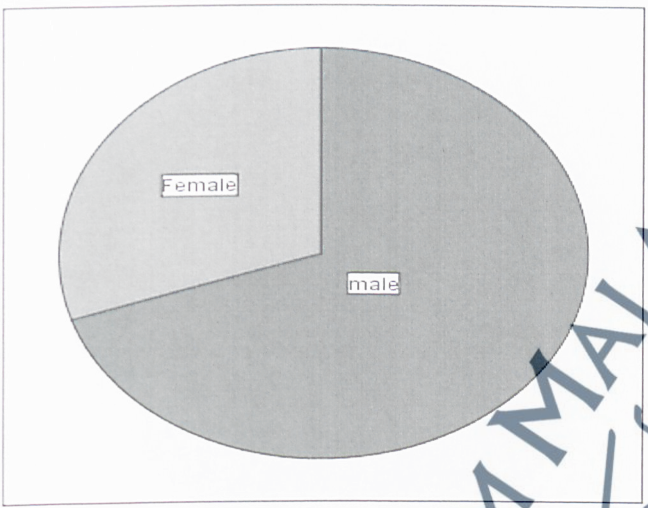
C. Demographic Profile of the Respondents

Before the questionnaire was disseminated to the participants, it was first evaluated and validated by several experts in the area. The demographic profile of the respondents is obtained from the first section of the questionnaire, which is composed of three items: (1) gender, (2) age and (3) educational background. The profile of the respondents is summarised in Table 7.3. The first column shows the total number of cases, whereas the second column presents the percentage of all the cases. The sample for the questionnaire comprises 30 randomly selected Libyan citizens, 70% are males and 30% are females.

Table 7.3: Profile summary of the respondents based on gender

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	Male	21	70.0	70.0	70.0
	Female	9	30.0	30.0	100.0
	Total	30	100.0	100.0	

Figure 7.11: Gender of participants

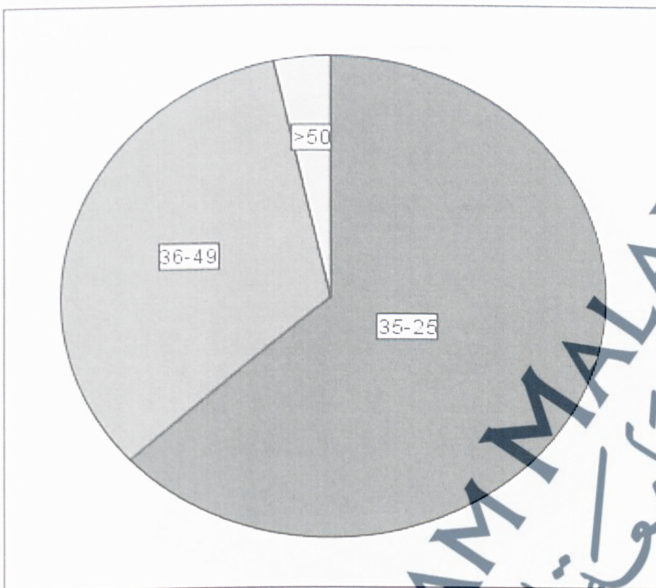


The majority of the participants are young, with the age group of 25–35 years comprising 63.3% of the participants, whereas the age group of 36–49 years comprises 33.3% of the participants, as shown in Table 7.4.

Table 7.4: Profile summary of the respondents based on age

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	35–25	19	63.3	63.3	63.3
	36–49	10	33.3	33.3	96.7
	>50	1	3.3	3.3	100.0
	Total	30	100.0	100.0	

Figure 7.12: Age of participants

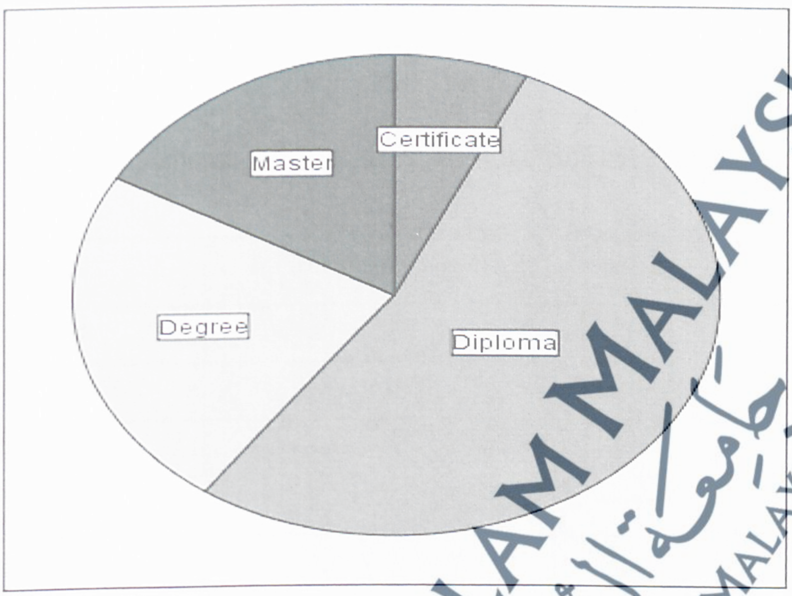


In terms of educational background, respondents with diploma degree comprise the largest number at 53.3%, those with degree comprise 23.3%, those with master's degree comprise 16.7% and those with certificate comprise 6.7%, as shown in Table 7.5.

Table 7.5: Profile summary of the respondents based on educational background

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	Certificate	2	6.7	6.7	6.7
	Diploma	16	53.3	53.3	60.0
	Degree	7	23.3	23.3	83.3
	Master	5	16.7	16.7	100.0
	Total	30	100.0	100.0	

Figure 7.13: Educational background



D. Findings

To investigate the hypothesis of the study, the researcher uses statistical methods, such as average, median, standard deviation and percentage. The researcher also uses statistical graphs to illustrate the findings.

Several methods can be used to test the validity and reliability of the questionnaire after confirming the questions and linguistic errors. The analysis of the validity and reliability of the questionnaire is an indication of the credibility of the test questions; thus, it will positively or negatively influence the results of the study. One of the methods used to test the validity and reliability of test scores is the Cronbach's alpha coefficient method.

Table 7.6 summarizes the Cronbach's alpha (a) value for each measure. The results show that the questionnaire has fulfilled all the required conditions to be valid for use when the Cronbach's alpha value is between 0.632 and 0.863. These values are acceptable and they satisfy the internal reliability criterion based on Nunnally (1978) and Pallant (2007).

Table 7.6: Cronbach's alpha values for all dimensions (n = 30)

Variable	Number of Items	Mean	Alpha (α)
Usefulness	11	5.594	0.821
Ease of Use	10	4.860	0.802
User Satisfaction	12	5.414	0.813
Reliability	5	5.400	0.864

The emergence of all measures with high means indicates that most of the participants have highly agreed on usefulness, ease of use, user satisfaction and reliability. Therefore, a conclusion can be drawn from the results that the participants have agreed that ERAA has appropriate usability level.

Descriptive statistic can be used to provide a formulation that can be easily understood by a project group. It has been used by the researcher to describe the usability evaluation of user perception of the ERAA prototype.

i. Usefulness Descriptive Statistics of ERAA

The first component of ERAA evaluation is usefulness. In general, the prototype should provide users with methods for escaping a minefield and assist them during emergency.

Table 7.7: Descriptive statistics of usefulness

	N	Minimum	Maximum	Mean	Standard Deviation
Q1	30	3	7	5.50	0.861
Q2	30	2	7	6.07	1.081
Q3	30	3	7	5.40	1.192
Q4	30	4	7	5.63	0.850
Q5	30	3	7	5.73	0.907
Q6	30	3	7	5.63	0.928
Q7	30	2	7	5.40	1.329
Q8	30	3	7	5.57	1.135
Q9	30	3	7	5.37	0.928
Q10	30	3	7	5.70	0.915
Q11	30	1	7	5.53	1.383
Valid *N* (listwise)	30				

Table 7.7 indicates that all the tasks meet over 70% of the predetermined success criterion. The mean ranges from 5.40 to 6.07 for each question from the questionnaire, whereas the standard deviation ranges from 0.861 to 1.383 for each question from the questionnaire. Seven questions have been used to measure system flexibility. After the corresponding values in the mean column are divided by its number, the results show that the mean of all the mean values, which corresponds to the flexibility of the questions, is 5.597. This value indicates that the total average number of users agrees that the system is useful.

ii. Ease of Use Descriptive Statistics of Environmental Risk Assessment Architecture

Q13 and Q14 have the lowest mean values because these questions have been formulated in reverse to ensure the credibility of the answer of the users. In general, however, Table 7.8 shows that the ease of use component of the ERAA prototype is considered high.

Table 7.8: Descriptive statistics of ease of use

	N	Minimum	Maximum	Mean	Standard Deviation
Q12	30	4	7	5.50	0.900
Q13	30	1	3	1.87	0.681
Q14	30	1	3	1.97	0.669
Q15	30	3	7	5.87	0.937
Q16	30	3	7	5.63	0.928
Q17	30	2	7	5.47	1.224
Q18	30	2	7	5.80	1.297
Q19	30	3	7	5.80	0.761
Q20	30	2	7	5.43	1.194
Q21	30	3	7	5.27	1.015
Valid *N* (listwise)	30				

iii. User Satisfaction Descriptive Statistics of Environmental Risk Assessment Architecture

The third component is user satisfaction. Table 7.9 shows that all the measures have high means (3.70 and 6.57), which indicates that most of the participants highly agree with the element of information quality (approximately 4.861).

Table 7.9: Descriptive statistics of user satisfaction

	N	Minimum	Maximum	Mean	Standard Deviation
Q22	30	3	7	4.93	0.944
Q23	30	3	7	5.77	0.898
Q24	30	4	7	6.57	0.728
Q25	30	4	7	5.70	0.702
Q26	30	3	7	5.50	0.861
Q27	30	3	7	5.43	0.817
Q28	30	3	7	5.50	0.938
Q29	30	3	7	5.03	1.066
Q30	30	3	7	5.43	0.935
Q31	30	4	7	5.77	0.774
Q32	30	3	7	5.63	0.964
Q33	30	1	7	3.70	2.136
Valid *N* (listwise)	30				

iv. Reliability Descriptive Statistics of Environmental Risk Assessment Architecture

The next component is reliability, which can be defined as a measure of the stability and overall performance of a system. Five questions are used to measure this component. Table 7.10 indicates high responses based on the overall results. The highest mean is achieved by Q37 ('The ERAA application provides the advice that I require to assist me during an emergency') and the lowest mean is obtained by Q35 ('The ERAA application can be relied upon to function properly').

Table 7.10: Descriptive statistics of reliability

	N	Minimum	Maximum	Mean	Standard Deviation
Q34	30	2	7	5.33	1.213
Q35	30	1	7	5.07	1.230
Q36	30	2	7	5.53	0.937
Q37	30	4	7	5.60	0.675
Q38	30	1	7	5.47	1.106
Valid *N* (listwise)	30				

A conclusion can be drawn based on the overall results (and as exhibited in Table 7.10) that the respondents agree that the prototype is capable of delivering stable and reliable overall performance.

v. Result of the Exploratory Factor Analysis

The factor analysis method aims to summarise multiple variables into a smaller number of factors, such that each factor has a function that will connect it with some or all of the variables. These factors can explain the largest possible proportion of variance in the original variables. The researcher used the factor analysis method to summarise multiple variables into fewer factors and use these factors to evaluate the prototype.

Table 7.11: Exploratory Factor Analysis

Code	Statement	Factor Loading
	A: Perceived Usefulness	
Q9	I can get sufficient information regarding the types of landmine.	0.894
Q10	ERAA provides guidance to avoid the risk of landmines.	0.818
Q7	ERAA will give me clarity on how to go forward with no sense of fear from the risk of landmine in my life.	0.792
Q2	Using ERAA will enable me to save my life and those of my family members.	0.702
Q3	Using ERAA will enable me to reduce tension from the risk of landmines.	0.551
Q4	Using ERAA will enable me to quickly obtain information about landmine.	0.551
Q11	Learning to utilise the ERAA app will help me avoid areas affected by landmine.	0.542
Q1	The methods for escaping a minefield provided by ERAA are sufficient to assist me during an emergency.	0.501
	B: Perceived Ease of Use	
Q18	Learning to operate the ERAA app is easy.	0.910
Q20	Interacting with the ERAA app by using a mobile phone is easy.	0.899
Q17	The ERAA interface is flexible.	0.876
Q16	Interaction within the ERAA app is understandable.	0.839
Q13	The ERAA interface is rigid and inflexible to interact with.	0.737
Q14	Interacting with the ERAA app requires considerable mental effort.	0.725
Q19	Interaction with the ERAA app is clear and understandable.	0.629
Q21	Avoiding minefields using the ERAA app is easy.	0.596
Q15	Overall, I find the ERAA app easy to use.	0.528
	C: User Satisfaction	
Q30	I am confident that I can make better decisions to avoid the risk of landmine during an emergency.	0.837
Q32	The ERAA app is capable of helping me make a choice during an emergency.	0.857
Q31	I am very pleased with my experience in using the ERAA app.	0.500
	D: Reliability	
Q37	The ERAA app provides the advice that I require to assist me during an emergency.	0.866
Q34	I can depend on the alert issued by the ERAA app to avoid minefields.	0.744
Q35	The ERAA app can be relied upon to function properly.	0.700
Q36	The ERAA app provides the help that I need to avoid minefields.	0.631
Q38	I can rely on the accuracy of ERAA maps to avoid minefields.	0.531

The Exploratory Factor Analysis results in Table 7.11 show that 4 factors exist, namely, usefulness, ease of use, satisfaction and reliability. Each item of a variable has been arranged in ascending order from the highest value of the factor loading. The researcher deletes 13 vocabularies from the final analysis because of the following reasons: **1)** the low saturation value of some factors, **2)** some factors are less than 0.40 and **3)** some factors within the number of items are less than 3 items.

7.3.3 System Testing

Testing is a critical tool that is used to ensure the quality of mobile applications. The purpose of testing is to discover errors. Testing is the process of attempting to discover every conceivable fault or weakness in a product. It provides a means to check the functionality of components. It is the process of training software with the intent of ensuring that the software system meets its requirements and the expectations of users.

The researcher uses two types of tools to test the proposed application. The first tools are emulators and mobile devices; the second tool is the HP App Pulse Mobile. Firstly, the researcher uses mobiles phone and a PC emulator to test the application. Samsung GT-18552, Sony Xperia, Sony Ericsson K800 and Nokia N95 mobile phones are used for the test process. These phones support a screen with resolution of 176 × 208 pixels and 1 262 144 colours, along with 3G networks for high data transfer rates. Moreover, some of these phones support the Java location API, whereas some have built-in GPS receivers. Table 7.12 shows the results of this test.

Table 7.12: Mobile phone capabilities and system performance

Mobile model	Capabilities	System Performance
Samsung GT-18552	• **NETWORK**: GSM/HSPA • **Wi-Fi**: Wi-Fi 802.11 b/g/n • **GPS**: A-GPS	System worked well
Sony Xperia	• **NETWORK**: GSM/HSPA • **Wi-Fi**: Wi-Fi 802.11 a/b/g/n, dual-band, Wi-Fi Direct, DLNA, hotspot • **GPS**: A-GPS, GLONASS	System worked well
Nokia N95	• **NETWORK**: GSM/HSPA • **Wi-Fi**: Wi-Fi 802.11 b/g, UPnP technology • **GPS**: A-GPS, Nokia Maps	System did not work
Sony Ericsson K800	• **NETWORK**: GSM/UMTS • **Wi-Fi**: No • **GPS**: No	System did not work

The next tool used to test the app is HP App Pulse Mobile. Firstly, we have to create an account in the HP website, as shown in Appendix 19. Several steps must be followed. Firstly, we have to add the HP App Pulse Mobile to our Android app, as shown in Appendix 20. Then, we have to download SDK and extract the App Pulse Mobile SDK, as shown in Appendix 21. We open a Windows command line (Start > Run > cmd), and enter the following: <Unzipped SDK directory>\AppPulse_mobile.bat -appkey <application key> <path and name of your APK>, as shown in Appendices 22a and 22b.

A total of 30 participants are recruited for this test. The results of the test are presented in Figure 7.14. The average score obtained is 98, which indicates that the system works well.

Figure 7.14: Results of the performance measurement system

The results of the test show that the system has strong and weak features.

- The system can run on any mobile phone that supports the Android platform, except for mobile phones that are not equipped with GPS capabilities.
- The system can run on mobile phones that support Google Maps.
- The application size is small because all the maps for landmine-affected areas and other data are stored in the server.
- The system does not work indoor because signal strength is too low to penetrate most buildings.

- The system provides the location of a user in real time and the locations of landmine-affected areas whenever the user moves from his/her location. This feature will help the user avoid areas affected by landmines.

7.4 Summary

Two verification methods were used to examine the efficacy and performance of NFRAM. The same dataset, which is considered the same input for the three models (Mamdani model, Sugeno model and NFRAM), was used in this research, along with three parameters (signal strength, position and landmine intensity). The validation results show that the proposed NFRAM provides better prediction results for landmine risks compared with the Mamdani and Sugeno models. An improvement in the results with a difference rate of 0% ≅ 10.8%, as shown in Table 7.1 and Figure 7.4, is observed based on the results obtained before and after training. In conclusion, NFRAM is highly efficient for predicting the risk of landmines.

For the save path selection model and based on the previous simulation and the computed time complexity for Dijkstra's algorithm and the Floyd–Warshall algorithm, these algorithms are determined to be acceptable in terms of their overall performance in solving the shortest path problem. However, all of these algorithms produce only one solution.

The comparison results of the save path selection model with the previous algorithms show that the save path selection model yields different results of the optimal solutions each time because the result can vary every time the save path selection model is executed, which can be considered the main advantage of this model.

The final section (7.3) presented heuristics evaluation strategies that were used in this study to examine the prototype. Two activities (expert review and user evaluation) were conducted in the review process to evaluate the ERAA prototype. In particular, usability and user satisfaction tests were performed on the ERAA prototype. The first activity, i.e. expert review, was conducted by involving mine clearers from LMAC. All the experts who participated in the study found that the proposed prototype is well accepted. In the second activity, evaluation was performed by Libyan citizens who live in the areas surrounding the capital (Tripoli). Overall, the evaluation sessions were successfully conducted and positive results were derived.