American Journal of Science, Engineering and Technology
2023; 8(1): 12-21
http://www.sciencepublishinggroup.com/j/ajset
doi: 10.11648/j.ajset.20230801.12
ISSN: 2578-8345 (Print); ISSN: 2578-8353 (Online)

Application of the Backpropagation ANN to Assess the Adoption Level of Farmers to Integrated Pest Management in the Province of Soc Trang (Vietnam)

Nguyen Trung Dung[1, *], Bui Thi Thu Hoa[1], Nguyen Tuan Anh[2]

[1]Faculty of Economics and Management, Thuyloi University, Hanoi, Vietnam

[2]Institute of Water Economics and Management, Vietnam Academy for Water Resources, Hanoi, Vietnam

Email address:

dzung_nguyen_ktcs@tlu.edu.vn (Nguyen Trung Dung), anhnt.iwem@gmail.com (Nguyen Tuan Anh), thuhoa.ktcs@tlu.edu.vn (Bui Thi Thu Hoa)

[*]Corresponding author

To cite this article:

Nguyen Trung Dung, Bui Thi Thu Hoa, Nguyen Tuan Anh. Application of the Backpropagation ANN to Assess the Adoption Level of Farmers to Integrated Pest Management in the Province of Soc Trang (Vietnam). *American Journal of Science, Engineering and Technology*. Vol. 8, No. 1, 2023, pp. 12-21. doi: 10.11648/j.ajset.20230801.12

Received: December 14, 2022; **Accepted**: January 3, 2023; **Published**: January 13, 2023

Abstract: The integrated pest management (IPM) program was implemented in 2015 and 2016 in the province of Soc Trang. The research question is whether Artificial Neural Networks (ANNs) with pattern recognition can be useful for classifying farmers for a more realistic assessment of the performance of an IPM program. To evaluate the performance of the program, three datasets were collected, including dataset S1i with 450 farmers interviewed before conducting the IPM program, S2i with 250 farmers in the pilot area (communes/villages), and S3i with 50 farmers outside the pilot area. The conventional statistical assessment method (CAM) assumes that all farmers in each dataset behave similarly related to IPM concerning the seed, spray frequency, and dosage. This means that the original datasets were used to estimate the required statistical parameters. Thus, the traditional approach wastes information hidden in all surveyed data. Based on ANN, we can classify and determine the percentage of farmers in the six groups or the level of IPM adoption (3 neutral groups and 3 active groups) as well as the actual benefits of the IPM program. ANN-based assessment method (ANN-M) has been proven to be better than CAM in evaluating the performance of the project.

Keywords: Artificial Neural Networks, ANN-Based Classification, Farmers, IPM Adoption

1. Introduction

Artificial neural networks (ANNs) are one of the most important elements of machine learning and artificial intelligence (AI). They are modeled after the structure of the human brain and their function is based on nodes where simple processing takes place. The range of applications of ANN is very wide (medical, business, pharmacy, bankruptcy application, speech recognition, etc.) and also includes agriculture. ANNs are increasingly used by food manufacturers in all phases of agricultural production and efficient farm management [1]. There are some examples of their applications such as predicting production effects in agriculture, checking diseases and pests, intelligent weed control, and classifying crop quality. AI methods support decision-making systems, for example in the optimization of storage and transport processes, prediction of the costs incurred depending on the chosen direction of management, etc. Looking back at the development over the past four decades, the era of ANN began with a simplified application in many fields and remarkable success in pattern recognition (PR) [2, 21-22]. A pattern can be referred to as a set of items, objects, images, events, cases, situations, features, or abstractions where facets of a set are alike in an unequivocal sense. The statistical pattern approach has been the most widely studied and used in practice, and ANNs are increasingly attractive, effective, efficient, and successful in achieving PR on many problems [3-5]. Unlike conventional pattern approaches, ANN can easily model complex or multi-complex tasks [6]. The former conventional techniques applied to handle PR problems are classified into structural, statistical, and hybrid approaches [7]. However, both the statistical and structural approaches can produce unsatisfactory results if they are applied as a solution to complex PR problems only. Nowadays the ANN models are used because they can yield a better result in PR problems even in multi-complex tasks. In this article, we focus on investigating the application of ANN-based PR in integrated pest management (IPM).

IPM is an approach to pest control using a variety of technologies, selected from a menu of options, all of which are environmental and human health-friendly compared to traditional practices [8]. IPM was first introduced by the FAO in Vietnam in 1992 and has been applied to paddy, vegetable, and fruit tree production. IPM is the best combination of cultural, biological and chemical measures that provides the most cost-effective, environmentally sound and socially acceptable method of managing diseases, insects, weeds and other pests.

In the frame of the World Bank project "Mekong Delta water resources management for rural development" (WB6-project) executed in seven provinces from 2013-2016, the M&E team was responsible for monitoring and evaluating the sub-project IPM program [9, 10]. As part of the IPM program in Soc Trang, farmer training was conducted on the following topics: (1) Basic principles of IPM; (2) SRI (the System of Rice Intensification), "1 must, 5 reductions" (certified seed and 5 reductions (seed quantity, fertilizer, AWD (alternate wetting and drying, also water-saving technology in lowland rice), frequency of spraying pesticide, and post-harvest losses); (3) Biological and ecological measures such as "rice field and flower banks", green mushrooms, probiotics, etc. [10]; (4) Application of the correct principles (alternative technical measures to use pesticides, labor protection, and the ecological environment); (5) Implementation of four piloting fields as farmer field schools. One of the project targets was to reduce pesticide use by 50%. Therefore, the provincial Plant Protection Departments (PPD) conducted two household surveys in seven provinces; one at the start of the project (dataset S1i - no adoption) in the main districts of the province, and one at the end of the project (dataset S2i - full adoption) in pilot communes/villages. An additional survey has been executed outside the pilot communes/villages (dataset S3i - limited adoption).

To assess the performance of the IPM program, the conventional assessment method CAM was used. With this approach, the original datasets do not require classification or special consideration, then statistical values such as mean and variance are determined and various options – before and after projects, inside and outside the pilot communes – are compared. From a project evaluation point of view, this is sufficient to assess the project performance, but from a scientific point of view there are still research gaps, namely:

The first question is whether this traditional approach wastes information hidden in all the surveyed data because during the field investigation we found that the farmers had different KAP (knowledge, attitudes, and practices) or behavior in the use of pesticides. In fact, this behavior depends on the following factors: (1) Natural factors such as pest status in the fields; (2) User factors such as their knowledge, personal morality, farmers' will, and internal strength in applying IPM (from skepticism to enthusiasm), economic situation of households as well as pressure from outside such as commitment to apply IPM; (3) Market factors: prices and promotions of sellers, etc.

The second question is whether ANN using machine learning methods can divide the farmers of each survey into two subgroups: neutral and active with IPM. So in our study, the six subgroups or adoption levels (from no adoption to full adoption, presented in Figure 2) are suggested. Finally, whether ANN can determine the percentage of farmers in each adoption level and the actual benefits of the IPM program.

From the literature review related to the introduction of IPM, it is clear that everything stems from the research design and the determination of the level of behavior through the constructed survey. According to [11], the hypothesis is: "Do farmers not widely adopt environmentally friendly technologies?" and pointed out the following barriers to IPM adoption as lack of knowledge about IPM, lack of training facility, the inadequacy of IPM materials, availability of pesticides, lack of coordination between farmers and extension agent, fear about IPM program, etc. According to [12] four groups that influence the behavioral intention of farmers in China are studied, such as farmer's characteristics, knowledge, retailer, and authority. The authors showed that the inadequate perception behaviors of farmers in pesticide use, were mainly due to lack of knowledge, infective actions of government and pesticide retailers, and pursuit of high profits. Hadi [13] summarized the determinants of IPM behavior in rice farming systems in Iran as follows: (a) Exogenous factors: external factors (national and local policy and planning), access to information and inputs, and attitude of reference group; (b) Farm characteristics: farm size, soil quality (i.e. infiltration rate, the capacity of water retention, aggregate stability, soil structure, and organic matter), mechanization, number of plots, and labor use; (c) Farmer characteristics: age, education, farming experience, and knowledge, attitude, on-farm and off-farm income, quality of life, and spiritual and religious beliefs, and (d) Innovation.

2. Materials and Methods

2.1. Materials

Soc Trang is the primary fragrant rice-growing area in the Mekong River Delta (MD) (Figure 1), which has around 148,000 ha of paddy (44% for special varieties); and the acreage and paddy production of 9% of MD.

The main rice crops in the MD are: (1) winter-spring season (WS), sowing in November-December (end of the rainy season) and harvesting in early April (dry season); (2) summer-autumn season (SA), sowing in April and harvesting in mid-August; (3) the third season, sowing in May/June and harvesting in November. Rice varieties grown in this season need to be adapted to large amounts of water. It is not recommended to plant rice in this season.

To obtain performance evaluation data from the IPM program steps 4 and 6 of "Ten steps to a results-based monitoring and evaluation system" were performed [14]. Based on that, a household questionnaire on agricultural production and perception of pesticide use was drawn up. The questionnaire included: (1) General information about the interviewee and household, (2) Farmland information (number of plots, area, season, etc.), (3) Current status of pesticide use in the last two seasons in terms of frequency of spraying, plant growth stage, name of drug, the purpose of use, amount of pesticide and effectiveness of use, etc. (4) Farmers' pest control methods, (5) Participation in technical training, (6) Farmers' understanding of pest control.

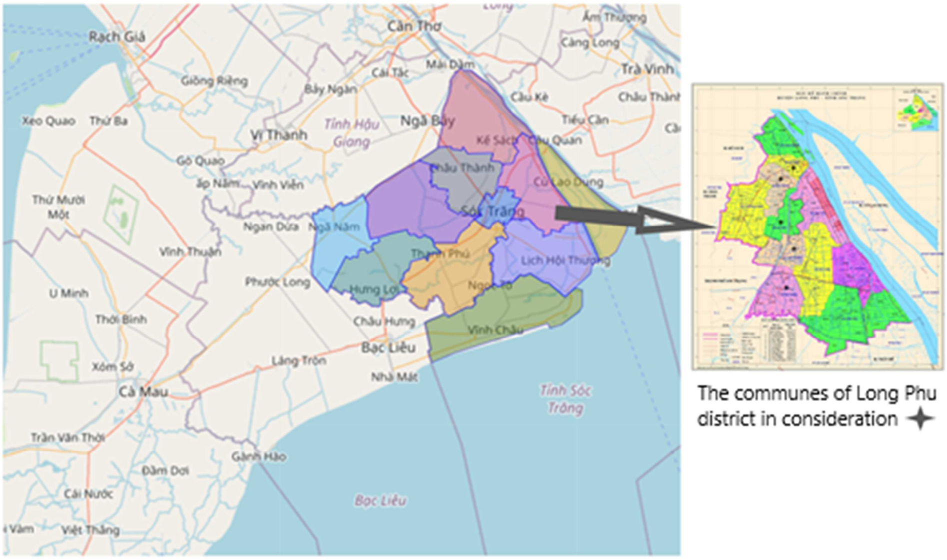

Figure 1. *The province of Soc Trang in the Mekong Delta (source: Google Map).*

The dataset of 750 farmer households has been collected in two stages: (1) Before the implementation of IPM program S1i: 450 farmers in 15 communes of 9 districts, nearly the whole province; (2) After the IPM program: (a) Inside the pilot communes/villages (PC) S2i with 250 households in 9 communes of 5 districts; (b) Outside the PC S3i with 50 farmers in 1 commune.

In Table 1, the first survey was conducted in all districts of province, and the second focused on 5 districts, but mainly on Long Phu. All data were collected using the same questionnaire. This study looked at two rice seasons. The first survey focused on the seasons WS 2014-15 and SA 2015, while the second survey focused on WS 2015-16 and SA 2016.

Table 1. *Distribution of the initial samples/datasets in Soc Trang Province.*

Districts	S1i (%)	S2i (%)	S3i (%)
Long Phu	33.3	60.0	100.0
Nga Nam	6.7	-	-
Thanh Tri	6.7	12.0	-
My Xuyen	6.7	-	-
My Tu	13.3	8.0	-
Ke Sach	6.7	-	-
Tran De	13.3	-	-
Chau Thanh	6.7	8.0	-
Soc Trang city	6.7	12.0	-
Total	100.0	100.0	100.0

In Table 2, there is a comparison between WS 2014-15 of S1i with WS 2015-16 of S2i and S3i (so-called season A), and between SA 2015 of S1i and SA 2016 of S2i and S3i (so-called season B). The results of the IPM program are summarized according to both seasons and datasets based on the original units. In the one-way ANOVA test on variables (seed density, paddy yield, spray frequency, spray dose) of three samples (S1i, S2i, and S3i) the Sig. F or $p < 0.05$, it means that the three samples are different. In contrast, the variable "farm size" has the Sig. F or $p > 0.05$, all three samples have no difference.

Table 2. *Comparison between three initial datasets in seasons A and B.*

Criteria	S1i	S2i	S3i
Sample N	450	250	50
Farm size (ha)	1.43	1.30	1.47
Seed density (kg/ha) in A	220.4	166.0	194.0
Seed density (kg/ha) in B	222.5	161.8	196.6
Paddy yield (tons/ha) in A	5.81	7.08	6.74
Paddy yield (tons/ha) in B	6.74	7.65	6.72
Spray frequency (times/ha) in A	9.88	7.12	9.24
Spray frequency (times/ha) in B	7.73	6.26	6.12
Spray dosage (litre/ha)* in A	5.0	3.3	5.7
Spray dosage (litre/ha)* in B	4.9	3.3	6.3
Cost (thousand VND/ha) in A	3,821	2,293	5,210
Cost (thousand VND/ha) in B	3,430	2,367	5,031

Note: * without snail insecticide because it is highly dependent on the particular field of households and behavior of farmers.
Exchange rate: 24.00 VND equivalent to 1 USD (status of June 2022).

In the IPM program, farmers are required to apply certified rice varieties. According to the survey results, nearly 100% of farmers use certified varieties provided by seed centers. The most commonly grown rice varieties in season WS were IR 50404 and OM6976; in season SA IR 50404 and OM 545. Therefore, this variable was not considered.

2.2. Methods

The methods used in this study were: (1) determining the IPM adoption levels used in classification, (2) the backpropagation (BP) method Feedforward neural networks (FNN).

2.2.1. Definition of the IPM Adoption Levels

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

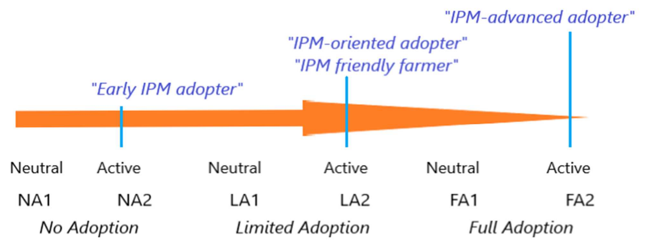

Figure 2. *The three IPM adoption groups and their six levels of adoption regarding spray dosage.*

Theoretically, adoption is a complex process that takes place in a person that consists of learning, making decisions and acting over a period of time. The adoption of a specific practice is not the result of a single decision to act but a series of actions through decisions. Rogers et al. have used the term "Innovation-decision process" in preference to the "adoption process" and have conceptualized the five stages: Knowledge K, Persuasion P, Decision D, Implementation I, and Confirmation C [15]. It is well known that the transition from K to I is complex. This is reflected in part in our study of people's awareness of pest control, for example, the use of work wear when spraying, and environmental protection when using pesticides. The transition from awareness to behavior can be shortened by many factors, the most prominent of which is the market. At the macro level, Vietnam is an agricultural country that aims to export safe and high-quality products, so it always has a policy to encourage farmers to integrate into the global supply chain with high-quality agricultural products. From this perception, when ANN classifies farmers based on input data, in each group two states are formed: normal/neutral/regular and irregular subgroups. The irregular subgroup can include either active or inactive farmers. So three initial datasets S1i, S2i, and S3i form six adoption levels and can be explained based on Figure 2 as follows:

(a) At the start of the IPM program: Without the intervention of the program, the neutral farmers are producing as normal (or BAU) and belonged to the NA1 group (no adoption). But there are always irregular farmers who can use less or more pesticides on their fields. They belong to NA2.

(b) In the pilot communes/villages the IPM program conducted different activities are: Most farmers who follow the general IPM requirements should be designated as neutral in FA1 (full adoption). However, some farmers have applied at a (probably) higher or lower level assessed as FA2.

(c) Outside the pilot communes/villages: Most neutral farmers with limited applications are in LA1 (limited adaptation) and a small group (LA2).

2.2.2. The Backpropagation ANN

Backpropagation (BP) is a widely used algorithm for FNN [16]. We'll keep it short here because there are many references that describe the problem in great mathematical detail. The multilayer perceptron (MLP) is a class of FNN. The term MLP is used ambiguously, sometimes loosely to mean any FNN, sometimes strictly to refer to networks composed of multiple layers of perceptron (with threshold activation). MLPs are sometimes colloquially referred to as "vanilla" neural networks, especially when they have a single hidden layer. The FNN in Figure 3 consists of one input, one hidden, and one output layer. Information moves in only one direction, forward from the input layer, through the hidden layer, and then to the output. For this relatively simple problem, a hidden layer is enough.

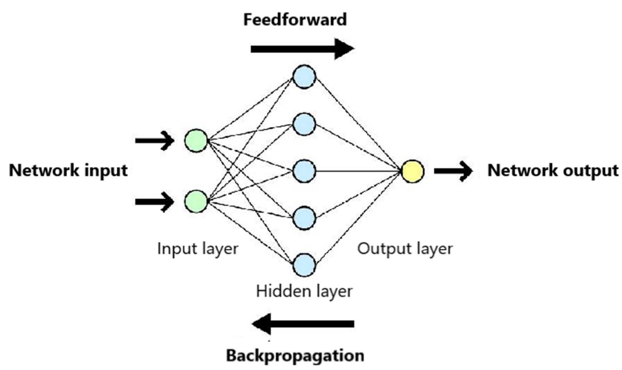

Figure 3. *Scheme of simple FNN.*

The goal of every training algorithm is to decrease this global error by adjusting the weights and biases. The BP algorithm works by computing the gradient of the loss function with respect to each weight by the chain rule, computing the gradient one layer at a time, iterating backward from the last layer to avoid redundant calculations of intermediate terms in the chain rule. Mathematically, it can be explained simply through the following formulas [3]:

The objective function of BP is

$$\frac{1}{2}SSE = \frac{1}{2}\sum_{t=1}^{T} SE_t \rightarrow Min$$

$$SE_t = [y_t - f(X_t, w)]^2$$

Where: SSE - the sum of squared error; SE_t - the squared error of observation t; T - the number of observations; w - weight; X_t - standardized input variable; y_t - standardized target variable.

The iteration rule for w of a network is:

$$w_{(n),t+1} = w_{(n),t} - \Delta w_{(n),t}$$

$$w_{(n),t+1} = w_{(n),t} - \gamma \frac{\partial SE_t(w)}{\partial w_{(n),t}} \qquad \frac{\partial SSE(w)}{\partial w_{(n)}} = \sum_{t=1}^{T} \frac{\partial SE_t(w)}{\partial w_{(n)}}$$

Where: $w_{(n),t+1} \equiv w_{(n),T+1}$; $t = 1, \ldots, T$; n - iteration step. The variables in this study are presented in Table 3.

Table 3. *The statistical description of the variables (display only for sample S1i in season A).*

Variable name	Category	Measurement, unit	Mean	St. De.
Dependent variable				
Adoption level	Nominal	6 levels	-	-
Independent variable				
Gender	Ordinal	binary	-	-
Age	Scale	-	45.62	11.412
National minorities (Kinh, Khmer, Chinese	Nominal	1, 2, 3	-	-
Education level	Scale	year	7.04	2.910
Farming experiences	Scale	year	21.67	11.597
Farm size	Scale	year	1.479	1.3658
Number of plots	Scale	number	1.12	0.343
Ownership of land	Ordinal	binary	-	-
Seeding density	Scale	Kg/ha	350	220.42
Paddy yield	Scale	Tons/ha	7.9	5.814
Rice Price	Scale	10^3 VND/ton	6400	4554.4
Spray frequency	Scale	Times/ha	8.8	-
Pesticide dosage	Scale	Litres/ha	18.82	-
Treatment cost	Scale	10^3 VND/ha	3,821	-

2.2.3. The FNN Structure

The ANN nodes in the input layer consist of three fixed node groups (farmer & farm characteristics, and rice cultivation) and a variable group such as frequency, dosage, and cost of pesticide use (Figure 4). In this study we temporarily omitted the fertilization factor as the main goal of the WB6-project is to reduce the use of pesticides by 50%. The frequency and the cost of spraying must be reduced accordingly. To converge results and avoid over-scaling of the ANN, we did not include all three subgroups at once but rather included each group separately. Architecturally, we declared a moderate number of input variables, the minimum number of nodes in the hidden layer, and a single output node, also IPM adoption.

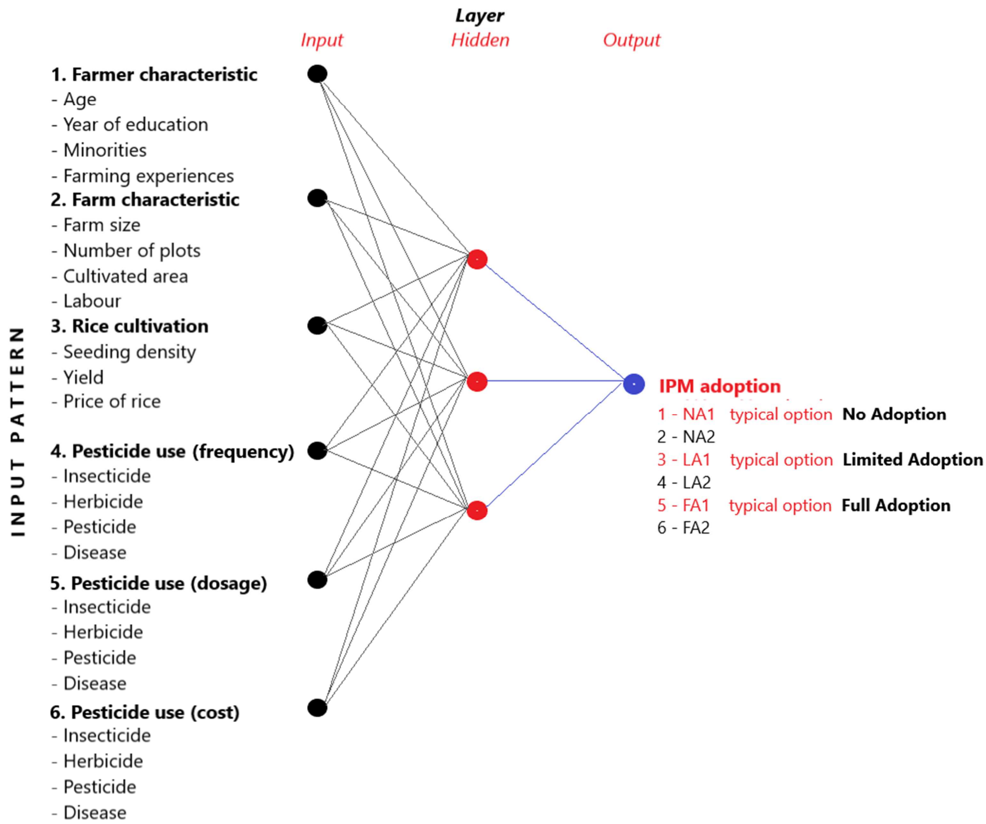

Figure 4. *The theoretical ANN structure for the farmer classification.*

3. Results

3.1. The Output of Network Training

The principle of division in BP is that 70% of the data is randomly selected for training and the remaining 30% for testing. Figure 5 is an example of a random network training session with the results, for example, the percentage correctness of S1i is 97.5% in training and 96.2% in testing. The "overall correctness percent" in training and testing is quite good, 95.7% and 92.7%, respectively.

Figure 6 shows the normalized importance of variables. For example, in season A yield, seed density, herbicide dose, and disease dose are very important (>60%). Ethnicity, education, and gender are of least importance. It is worth noting that pesticide dose is of little importance compared to herbicide dose and disease dose.

The main intermediate results of the data classification are summarized in Table 4. The meaning of the numbers is explained below using the examples:

1) According to the frequency in season A, the initial dataset S1i with 450 cases was divided into three subgroups: 313 in NA1 (i.e. 69.6% of S1i), 136 in NA2 (30.2%), and 1 in LA1 (0.2%). The same applies to S2i and S3i for the three criteria spray frequency, dosage, and costs.
2) Relocating of farmers because they did not match the original group: For example, according to dosage 17 farmers in A (6.8% of S2i, see columns 5 & 6) and 20 farmers in B (8.0% of S2i, see columns 11 & 12) were moved to group NA1.

Classification

Sample	Observed	Predicted			
		S1i	S2i	S3i	Percent Correct
Training	S1i	311	8	0	97.5%
	S2i	8	117	1	92.9%
	S3i	3	0	21	87.5%
	Overall Percent	68.7%	26.7%	4.7%	95.7%
Testing	S1i	125	3	2	96.2%
	S2i	1	48	0	98.0%
	S3i	7	2	17	65.4%
	Overall Percent	64.9%	25.9%	9.3%	92.7%

Dependent Variable: Adoption levels

Figure 5. *ANN-based classification - Number of farmers observed and predicted with the correctness (this table taken from SPSS protocol).*

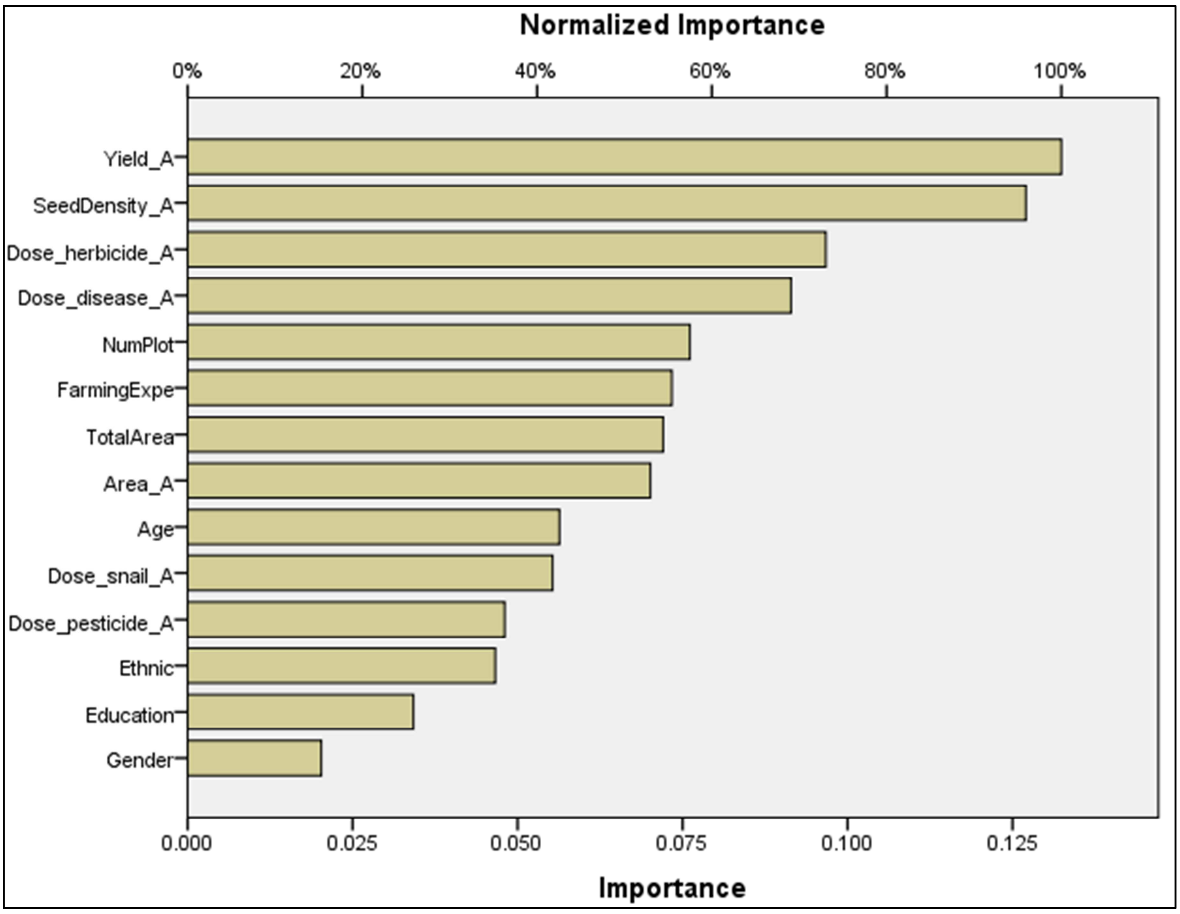

Figure 6. *The independent variable importance in ANN-based classification.*

3.2. Analysis of Situations Before and After IPM Program Implementation

a) The situation of Soc Trang before the implementation of the IPM program

The general situation before the IPM program was that to ensure high paddy yield, farmers applied dense sowing density, sprayed many times (for prevention and treatment), and used many pesticides. The following figures in columns 2 & 3 of Table 5 relating to the node groups can explain this situation:

1) "spray frequency" in season A, 70.3% of farmers in the NA1 sowed 216.8 kg/ha and 29.7% of the NA2 226.4 kg/ha. Due to thick sowing, the NA2 group sprayed an average of 10.2 times/season compared to 9.71 of the NA1. As a result, the NA2 group had a higher yield of 5.97 tons/ha than the 5.74 of NA1;
2) "spray dosage" in season A, 99.6% of NA1 with a seed density of 221.6 kg/ha sprayed 5.06 litres/ha, whereas 0.4% NA2 sowed less (210 kg/ha) and thus sprayed less (4.4 litres/ha), and achieved the paddy yield of 5.9 tons/ha compared with 4.85 of NA1. These two households used very little pesticide and therefore had low paddy yields. These may be exceptions or these farmers may have trade-offs between spraying and yield.
3) "treatment cost", 99.8% belongs to NA1 and the other figures can be seen in the table.

The situation of season B are similar in column 8 & 9.

b) After implementation of the IPM program

In the pilot area in seasons A and B, the situation in adoption stages FA1 and FA2 (columns 4 & 5, 10 & 11) is significantly different. The following figures present the situation by node groups:

1) "spray frequency" in A: 17.9% of FA1 had a seed density of 156.4 kg/ha, sprayed 7.55 times/season; in contrast, 82.1% of FA2 had sown 168.2 kg/ha and sprayed 5.97 times/season (i.e. reduction of 1.58 times/season) and had a paddy yield of 7.11 tons/ha (higher than 6.9 tons/ha of FA1). Compared to outside the pilot area in the same season, this result is really impressive (columns 6 & 7, and 12 & 13 respectively).
2) "spray dosage" in A: 72.8% of FA1 had a seed density of 165.4 kg/ha, sprayed 3.41 litres/ha and had a paddy yield of 6.81 tons/ha. In contrast, 27.2% of FA2 with a seed density of 160.9, consumed 2.91 (i.e. reduction of 0.5 litres/ha) and achieved a higher rice yield of 7.22. In season B, the results are even better.
3) "treatment cost" in A: 80.4% FA1 have a seed density of 165.3 kg/ha vs. 19.6% of FA2 174, have the cost of 1.97 Mill. VND/ha vs. 1.69, and a paddy yield of 6.98 tons/ha vs. 7.27. In season B, the results are even better.

In the LA group (columns 6 & 7 and 12 & 13), almost 100% of the households are in LA1 due to the small sample size.

c) Implementation of independent-samples T-tests for NA, FA, and LA

T-test was performed for each group of NA, FA and LA. In general, there is a statistical difference (Sig. 2-tailed < 0.05) between NA1 and NA2, FA1 and FA2, and LA1 and LA2 in both paddy seasons regarding two criteria of spray frequency and dosage. The test results are marked with characters ᴥ or ‡ in Table 5, for example in terms of dosage in season B, in the NA (NA1) 5.06 litres/ha of NA1 > 4.40 of NA2, in FA 3.41 litres/ha of FA1 > 2.91 of FA2. Thus, six stages of adoption are shown in Figure 2 and NA 2 are considered "Early IPM adopter", LA2 "IPM-oriented adopter" or IPM friendly farmer", and FA2 as "IPM-advanced adopter".

3.3. Comparison Between CAM and ANN-M (ANN-Based Assessment Method)

The real benefits of the IPM program in the province of Soc Trang are expressed in initial units (kg, tons, litres...) as shown in Table 5. These are increased paddy yield, reduced seed density and pesticide use in comparison to "before IPM" or "outside the pilot area" for all six adoption levels and both seasons. Table 6 shows the percentage increase or decrease when comparing the two groups "before vs. after" and "inside vs. outside" or in the detailed comparisons:

1) in CAM: "FA vs. NA" or "FA vs. LA", and
2) in ANN-M: "FA1 vs. NA1" and "FA2 vs. NA1" and "FA1 vs. LA1".

Here are some relevant comments regarding two assessment methods:

a) Method CAM

When comparing "before vs. after" or "FA vs. NA", seed density decreased by 24.7% in A and 27.3% in B, spray frequency decreased from 19% to 27.9%, and dose decreased by 34.2% and 42.2%, respectively, and costs decreased by 31.1% and 41.8%, respectively. However, paddy yield increased by 21.9% and 13.5%, respectively.

When comparing "inside vs. outside" or "FA vs. LA", seed density decreased by 14.4% in A and 14.7% in B, and pesticide dose decreased by 33.5% and 48.5% respectively (the target of WB6-project is pesticide reduction by 50%), the frequency decreased by 22.9% in A, but increased by 2.3% in B or SA 2016. According to PPD, due to the complicated pesticide and disease situation in SA 2016, the spraying strategy in the pilot communes/villages was: more frequently, but with significantly lower dosages.

b) Method ANN-M

In order to have an overall picture, it is recommended to first analyse the whole including all three groups of ANN nodes.

1) Comparison "before vs. after", seed density decreased from 20.6% to 32.6%, yield increased from 9.4% to 24.1%, spray frequency decreased from 18.3% to 54.5%, dosage reduction from 32.6% to 42.5%, cost reduction from 31.9% to 50%.
2) Comparison "inside vs. outside", seed density decreased from 15.6% to 20.2%, yield increased from 2.2% to 14%, spray frequency decreased by 17.2% in A, but increased by 2.4% in B, dosage decreased from 34.6% to 50.8%, cost decreased from 55.3% to 60.7%.

Table 4. *Results of ANN-based classification from the original group into adoption levels.*

Initial data & sample	Adoption level	Season A in node group						Season B in node group					
		Frequency		Dosage		Cost		Frequency		Dosage		Cost	
		N	%	N	%	N	%	N	%	N	%	N	%
Col. 1	2	3	4	5	6	7	8	9	10	11	12	13	14
S1i	NA1	313	69.6	421	93.6	439	97.6	429	95.3	421	93.6	435	96.7
(450)	NA2	136	30.2	2	0.4	1	0.2	2	0.4	9	2.0	2	0.4
	FA1	-	-	23	5.1	4	0.9	8	1.8	18	4.0	8	1.8
	LA1	1	0.2	4	0.9	6	1.3	11	2.4	3	0.4	5	1.1
	Total	*450*	*100*	*450*	*100*	*450*	*100*	*450*	*100*	*450*	*100*	*450*	*100*
S2i	NA1	2	0.8	17	6.8	9	3.6	12	4.8	20	8	14	5.6
(250)	FA1	44	17.6	161	64.4	193	77.2	234	93.6	139	55.6	235	94.0
	FA2	202	80.8	70	28.0	48	19.2	2	0.8	91	36.4	1	0.4
	LA1	2	0.8	2	0.8	-	-	2	0.8	-	-	-	-
	Total	*250*	*100*	*250*	*100*	*250*	*100*	*250*	*100*	*250*	*100*	*250*	*100*
S3i	NA1	7	14.0	7	14.0	4	8.0	13	26.0	13	26.0	4	8.0
(50)	FA1	-	-	3	6.0	-	-	7	14.0	1	2.0	-	-
	LA1	41	82.0	40	80.0	46	92.0	30	60.0	36	72.0	46	92.0
	LA2	2	4.0	-	-	-	-	-	-	-	-	-	-
	Total	*50*	*100*	*50*	*100*	*50*	*100*	*50*	*100*	*50*	*100*	*50*	*100*

Unit in table: Frequency (times/season), Dosage (liters/ha), and Cost (Mill. VND/ha).

Table 5. *Comparison of mean values of six adoption levels in seasons (with t-test for equality of means within the group of NA, FA, and LA).*

Node groups	Season A at adoption levels						Season B at adoption levels					
	NA1	NA2	FA1	FA2	LA1	LA2	NA1	NA2	FA1	FA2	LA1	LA2
Column 1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
a) Spray frequency												
Sample N	322	136	44	202	44	2	454	2	249	2	43	-
Percent (%)	70.3	29.7	17.9	82.1	95.7	4.3	99.6	0.4	99.2	0.8	100	0
Yield (tons/ha)	5.74*	5.97*	6.90‡	7.11‡	6.75‡	7.0‡	6.78‡	6.75‡	7.58‡	7.45‡	6.65	-
Seed density (kg/ha)	216*	226*	156*	168*	195*	165*	222‡	250‡	160‡	150‡	201	-
Spraying (times/crop)	9.7*	10.2*	7.6*	5.9*	9.1*	10.5*	7.7‡	7.5‡	6.3*	3.5*	6.1	-
b) Spray dosage												
Sample N	445	2	187	70	46	-	453	10	158	92	37	-
Percent (%)	99.6	0.4	77.8	27.2	100.0	0	97.8	2.2	63.2	36.8	100.0	0
Yield (tons/ha)	5.90*	4.85*	6.81*	7.22*	6.66	-	6.81‡	6.71‡	7.54‡	7.45‡	6.79	-
Seed density (kg/ha)	221‡	210‡	165‡	161‡	199	-	223‡	221‡	159‡	155‡	198	-
Amount (litres/ha)												
- Snail insecticide*	13.79	10.00	16.85	13.10	38.87	-	13.72	26.07	9.19	12.91	46.32	-
- Herbicide	1.13	0.90	1.12	0.94	1.43	-	1.16	0.90	1.05	0.99	1.47	-
- Pesticide	1.02	0.80	0.60	0.45	0.62	-	0.94	0.64	0.51	0.62	0.82	-
- Disease	2.91	2.70	1.69	1.52	3.62	-	2.84	2.59	1.69	1.62	4.32	-
Total without snail	*5.06**	*4.40**	*3.41**	*2.91**	*5.67*	-	*4.93**	*4.13**	*3.25*‡	*3.23*‡	*6.61*	-
c) Treatment cost												
Sample N	452	1	197	48	52	-	453	2	243	1	51	-
Percent (%)	99.8	0.2	80.4	19.6	100.0	0	99.6	0.4	99.6	0.4	100.0	0
Yield (tons/ha)	5.86	5.30	6.98	7.27	6.57	-	6.78	6.75	7.59	7.40	6.69	-
Seed density (kg/ha)	219	200	165	174	196	-	222	250	160	160	197	-
Cost (Mill. VND/ha)												
- Snail insecticide*	0.41	0.90	0.31	0.19	0.44	-	0.41	0.70	0.28	0.0	0.43	-
- Herbicide	0.32	0.24	0.26	0.29	0.45	-	0.32	0.29	0.25	0.0	0.47	-
- Pesticide	1.61	1.65	0.39	0.28	1.35	-	1.03	1.44	0.48	0.18	1.40	-
- Disease	1.45	1.90	1.32	1.12	3.21	-	1.66	1.95	1.32	1.64	2.72	-
Total without snail	*3.38*‡	*3.79*‡	*1.97**	*1.69**	*5.01*	-	*3.01*‡	*3.68*‡	*2.05*‡	*1.82*‡	*4.59*	-

Note: * The use of snail insecticides depends on the disease situation in the respective field and is therefore very difficult to take into account. t-test for equality of means: * Sig. 2-tailed <0.05 and ‡ Sig. 2-tailed > 0.05.

Table 6. *ANN-M and CAM in comparison in three ANN node groups.*

Node groups & criteria	Increase or decrease (%) based on the method									
	CAM				**ANN-M**					
	Before vs. After		**Inside vs. Outside**		**Before vs. After**		**Inside vs. Outside**			
	FA vs. NA		***FA vs. LA***		***FA1 vs. NA1***		***FA2 vs. NA1***		***FA2 vs. LA1***	
	A	**B**	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**
a) Spray frequency										
Yield	21.9	13.5	5.1	13.8	20.2	11.8	23.9	9.9	2.2	14.0
Seed density	-24.7	-27.3	-14.4	-14.7	-27.9	-27.9	-22.4	-32.6	-20.0	-20.2
Frequency	-27.9	-19.0	-22.9	2.3	-22.2	-18.3	-38.5	-54.5	-17.2	2.4
b) Spray dosage										
Yield	*	*	*	*	15.4	10.7	22.4	9.4	2.3	11.0
Seed density	*	*	*	*	-25.4	-28.6	-27.4	-30.4	-17.1	-19.4
Frequency	-34.2	-42.2	-33.5	-48.5	-32.6	-34.2	-42.5	-34.6	-39.9	-50.8
c) Spray cost										
Yield	*	*	*	*	19.1	11.9	24.1	9.1	6.2	13.5
Seed density	*	*	*	*	-24.6	-28.0	-20.6	-28.1	-15.6	-18.6
Frequency	-41.8	-31.1	-58.3	-54.6	-41.7	-31.9	-50.0	-39.5	-60.7	-55.3

Note: * Same data as in item a) because CAM does not differ the data according to frequency, dosage, and cost

4. Conclusion

Vietnam is also a relatively heavy pesticide user, although there have been many programs over the years to promote the widespread adoption of IPM [17]. Many authors have pointed to three main factors contributing to slow demand-side adoption: farmers' awareness and knowledge, perception of low returns of IPM technologies, risk, and uncertainty. However, according to WB [20], economic factors are the main reason for the slow adoption of IPM in developing countries.

With the available data from the WB6-project, we wanted to test a different assessment method than the traditional method still used. The traditional method CAM assumes that all farmers have the same attitude and practice towards the use of pesticides. This would include active and skeptical farmers. A mean value for the entire group does not reflect the reality of the IPM application. If we split it into two groups: positive and neutral or skeptical, it's great. Information such as pesticide use (mean values) along with the percentage of farmers in each group and in each paddy season is very valuable to the project management board and organizations related to the IPM program. This approach has great significance for improving post-project sustainability. Therefore, in this study, we apply the experimental approach to overcome the weakness of CAM. It focuses on data mining using ANN [18, 19].

Over the past two or three decades, there have been many studies that have used BP of ANN, for example, to classify a bank's potential customers - remain or leave prediction, or the simplest dual classification (DC) - based on the bank's raw customer data. From that example came the idea of applying DC to the problem of farmer classification or adoption level classification in the evaluation of the IPM program. Also, the input is the original three initial heterogeneous datasets (S1i, S2i, and S3i). According to DC, in principle, each dataset is classified into two groups: neutral and active. The neutral group is the majority group in each dataset or the typical group of the dataset, and the active group is the minority group or the outlier group. So, the three original datasets (S1i, S2i, and S3i) were transformed into six adoption levels (NA1 & NA2, LA1 & LA2, FA1 & FA2) in which there are three neutral groups (NA1, LA1, and FA1) and three active groups: NA2 or "early IPM adopters", LA2 or "IPM-oriented adopter" and FA2 or "IPM-advanced adopter". Because the simultaneous inclusion of three different datasets in the calculation results in a data shift: (1) upgrade from NA to LA or FA or (2) downgrade from FA and LA to NA. Of course, there's not much variation between these adoption levels.

The architecture of ANN applied here includes two fixed node groups (farm & farmer characteristics, paddy cultivation) and a variable group. The variable group consists of one of three subgroups (spray frequency, dose, and costs) in seasons A and B. The reason for not including all three groups at the same time is that the network size is large, many datasets are excluded due to missing data, and the calculation results are scattered. Specifically, the focus of the study was to first reduce the amount of pesticides, then the number of sprays, and finally the cost of treatment. The results show that BP of ANN is an effective tool and can solve many problems in machine learning.

Finally, the question arises whether ANN-M makes calculations and comparisons more difficult, and a waste of time. We can say that ANN-M is more specific and precise compared to CAM. Under CAM, all farmers participating in a study are automatically combined into a common group. With this subjective aggregation in a basket, we started making a mistake and wasting information right away, because a lot of the information hidden in the data couldn't be used. Therefore, the result is not as expected and is roughly evaluated. ANN-M helps identify farmers active in the IPM program. These are the "early IPM adopter" before the IPM program, "IPM-oriented adopter" or "IPM friendly adopter" outside the pilot area, and "IPM-advanced adopter" in the PA. Perhaps this number of farmers is not many, but most importantly, they will be role models in using pesticides. Their level of use (frequency and dosage) is considered a target value for other farmers from which they can learn; their field is considered a model field for visiting in form of farmer field schools, etc. These farmers can also present their experience in the IPM training. These groups are therefore used as leverage to increase the effectiveness and sustainability of the IPM program.

The province of Soc Trang achieved the main target of the WB6-project to reduce the use of pesticides by 50%. In fact, in the best-case scenario, it has decreased to 42.5% (compared to 2015) and 50.8% (compared to outside PA in 2016). This is a major effort by the provincial authorities and farmers. The significant reduction in the amount of pesticides used in rice production has helped the province to export many of the world's best rice varieties ST24 and ST25, especially to markets requiring high quality.

From the perspective of donors such as the World Bank, ADB, and an international organization or MARD, this is very important, as it allows us to make a more comprehensive and realistic assessment, identify the types of groups that are neutral or active in applying IPM in each paddy season and see the trend of shifting farmers' behavior from neutral to positive over time.

Acknowledgements

Sincere thanks to PPD of the province of Soc Trang and all farmers inside and outside the pilot area for supporting our M&E team to carry out the project task with their creativity and farming experience.

References

[1] Kujawa, S. & Niedbała, G., 2021. Artificial Neural Networks in Agriculture. *MDPI,* 11 (Artificial Neural Networks in Agriculture), pp. 2-6.

[2] Abodun, O. I. et al., 2017. *Comprehensive Review of Artificial Neural Network Application to Pattern Recognition.* s. l., IEEE.

[3] Anders, U., 1997. *Statistische neuronale Netze.* München: Verlag Vahlen.

[4] Bala, R. & Kumar, D., 2017. Classification Using ANN: A Review. *International Journal of Computational Intelligence Research,* Volume 13 (7), pp. 1811-1820.

[5] Jain, A., 2000. Statistical Pattern Recognition: A Review. *IEEE transactions on pattern analysis and machine intelligence, vol. 22, No. 1. January,* pp. 4-37.

[6] Nikzad, M., Movagharnejad, K. & Talebnia, F., 2012. Comparative study between neural network model and mathematical models for prediction of glucose concentration during enzymatic hydrolysis. *International Journal of Computer Applications,* Volume 56 (1), pp. 43-48.

[7] McAvoy, T. J. et al., 1992. A comparison of neural networks and partial least squares for deconvoluting fluorescence spectra. *Biotechnology and bioengineering,* Volume 40, pp. 53-62.

[8] Alwang, J., Norton, G. & Larochelle, C., 2019. Obstacles to widespread diffusion of IPM in developing countries: Lessons from the field. *Journal of Integrated Pest Management,* (2019) 10 (1): 10 (doi: 10.1093/jipm/pmz008), pp. 1-8.

[9] Dung, N. T., 2018. Economic and environmental effects of Integrated Pest Management program: A case study of Hau Giang province (Mekong Delta). *J. Viet. Env. 2018, 9 (2),* pp. 77-85.

[10] Dung, N. T., Anh, N. T. N. & Kien, P. D., 2021. IPM Program Combined with "Rice Fields, Flower Banks" or BIO-IPM Program in Soc Trang Province (Vietnam). *Journal Sustainability in Environment,* pp. 1-18.

[11] Kabir, M. & Rainis, R., 2015. Do Farmers Not Widely Adopt Environmentally Friendly Technologies? Lesson from Integrated Pest Management (IPM). *Journal Modern Applied Science Vol. 9, No. 3,* pp. 208-215.

[12] Fan, L., et al., 2015. Factors affecting farmers' behaviour in pesticide use: Insights from a field study in northern China. *Journal of science total environment,* Volume 537, pp. 360-368.

[13] Hadi, V., 2012. Exploring the determinants of adoption behaviour of clean technologies in agriculture: a case of integrated pest management. *Asian Journal of Technology Innovation 20 (1),* pp. 67-82.

[14] Kusek, J. & Rist, R., 2005. *Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners.* 1. ed. Washington DC: The International Bank for Reconstruction and Development.

[15] Rogers, E. M., Singhal, A. & Quinland, M. M., 2009. Diffusion of innovations. In: *An integrated approach to communication theory and research.* NY: Routledge, pp. 418-434.

[16] Gullu, M., Yilmaz, M. & Yilmaz, I., 2011. *Application of Back Propagation Artificial Neural Network for Modelling Local GPS/Levelling Geoid Undulations: A Comparative Study.* Marocco, FIG Working Week 2011.

[17] WB, 2016. *Transforming Vietnamese Agriculture: Gaining More from Less (Vietnam Development Report 2016),* Washington: The World Bank.

[18] Wiedmann, K.-P., Buckler, F. & Buxel, H., 2001. Data Mining - ein führendes Überblick. In: F. B. Klaus-Peter Wiedmann, ed. *Neuronale Netze im Marketing-Management: Praxisorientierte Einführung in modernes Data-mining.* Wiesbaden: Gabler, pp. 15-34.

[19] Arockiaraj, M., 2013. Applications of Neural Networks in data mining. *Research Inventy: International Journal Of Engineering And Science, Vol. 3, Issue 1 (May 2013),* pp. 08-11.

[20] WB, 2011. *Pest Management Framework - (PMF).* Washington: The World Bank.

[21] Saravan, K. & Sasithra, S., 2014. Review on classification based on artificial neural networks. *International Journal of Ambient Systems and Applications (IJASA) Vol. 2, No. 4,* pp. 11-18.

[22] Yang, L. & Xuewei, C., 2020. ANN-Based Continual Classification in Agriculture. *Journal Agriculture 2020, 10, 178.*