IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

CLASSIFYING POISONOUS AND EDIBLE MUSHROOMS IN THE AGARICUS AND LEPIOTA FAMILY USING MULTILAYER PERCEPTION

Mali H. Hakem Alameady*

DOI: 10.5281/zenodo.233441

ABSTRACT

Classification is one of the applications of feed-forward Artificial Neural Network (ANN). Classification can map data to predefined classes or groups. It is referred to as a supervised learning, because before examining data the classes are always determined. Multi-Layer Perception, is a supervised neutral networks model that is use to train and test data to build a model. In this experiment. Multi-Layer Perception is used to train the Data set to produce a model to make prediction of classifying .After preparing the Mushrooms data for training, only 8124 of dataset instances used to be train. Software used to mining data in this project is Neural Connection Version 2.0. This report, generally explaining the Classification, Multi-Layer Preceptor, Back propagation, Mushrooms, and details on the mining activity done to the selected datasets, to determine whether Mushroom's attribute is edible or Poison.

KEYWORDS: Classification, Multi-Layer Preceptor, Back propagation, Mushrooms.

INTRODUCTION

About every single year Data are doubled, but all the useful information are seems to be decreased. Area of data mining has arisen over decade in order to address the problem. It has not become only an important research areas, but it has also become one with a large potential in real world. The Multi-Layer Perception is a modeling and forecasting tool that uses a neural network to model your data. It can be used to classify patterns or to predict values from your data. Because it uses a supervised learning technique, it requires your data to contain targets for training the network. Classification is one of the tasks performed through data mining process. It is categorized as one of the prediction methods for a large hidden data. A model produced through data training should be able to distinguish the category of a new data of mushroom in these families. Mushroom, Agaricaceae, is belonged to members of a family of fungi with gills (Lentz). It is always categorized into to groups; edible and poisonous. The term mushroom is always used to refer to edible species, while the term toadstool is used for poisonous specious. However, the toadstool is always replaced by poisonous mushroom.

OBJECTIVES

The main objective of the study is determining if mushrooms are poisonous or edible. Specifically, the objectives are:

(a) Identify the target and the independent variables (attributes).
(b) Preprocess mushroom data so that it is suitable for training.

METHODOLOGY

Data Acquisition

This datasets obtained from UCI Repository of Machine Learning Database .Mushroom has recorded drawn from Audubon Society Fields, the main Guide to the North American Mushroom

Data description

This data set has include description of the hypothetical sample which is corresponding to the 23 species of the gilled mushroom in the Lepiota and Agarics Family, Each one of those species is identify as the definitely edible ,or definitely poisonous , or unknown edibility, and are not recommended at all. The latter class has been combined with a poisonous and edible based on 22 physical attributes as recorded in [11].

Table 1: Part of the normalization data

Attributes																						T
4	4	0	1	8	2	2	2	1	2	4	4	4	8	8	3	3	2	5	1	4	5	0
4	4	9	1	1	2	2	1	1	2	2	4	4	8	8	3	3	2	5	2	3	1	1
6	4	8	1	2	2	2	1	2	2	2	4	4	8	8	3	3	2	5	2	3	3	1
4	3	8	1	8	2	2	2	2	2	4	4	4	8	8	3	3	2	5	1	4	5	0
4	4	3	0	7	2	1	1	1	1	4	4	4	8	8	3	3	2	1	2	1	1	1
4	3	9	1	1	2	2	1	2	2	2	4	4	8	8	3	3	2	5	1	3	1	1
6	4	8	1	1	2	2	1	5	2	2	4	4	8	8	3	3	2	5	1	3	3	1
6	3	8	1	2	2	2	1	2	2	2	4	4	8	8	3	3	2	5	2	4	3	1
4	3	8	1	8	2	2	2	8	2	4	4	4	8	8	3	3	2	5	1	5	1	0
6	4	9	1	1	2	2	1	5	2	2	4	4	8	8	3	3	2	5	1	4	3	1
4	4	0	1	8	2	2	2	1	2	4	4	4	8	8	3	3	2	5	1	4	5	0

Data distribution:

The dataset was distributed into two different classes:

- Class 1 = Edible with the number of 4208 instances (51.28%)
- Class 2 = Poison with the number of 3916 instances (48.2%)

There are total of 8124 instances that were captured, however only 1000 are processed in the Neural Connections program.

Setting a Target

The main target of this experiment is determining if mushroom is edible (e) or poisonous (p)

Normalizing Attributes

Before any data train by the Neural Connection, the data normalization process is required. This to ensure all data value must be in number so it can be train by neural connection.

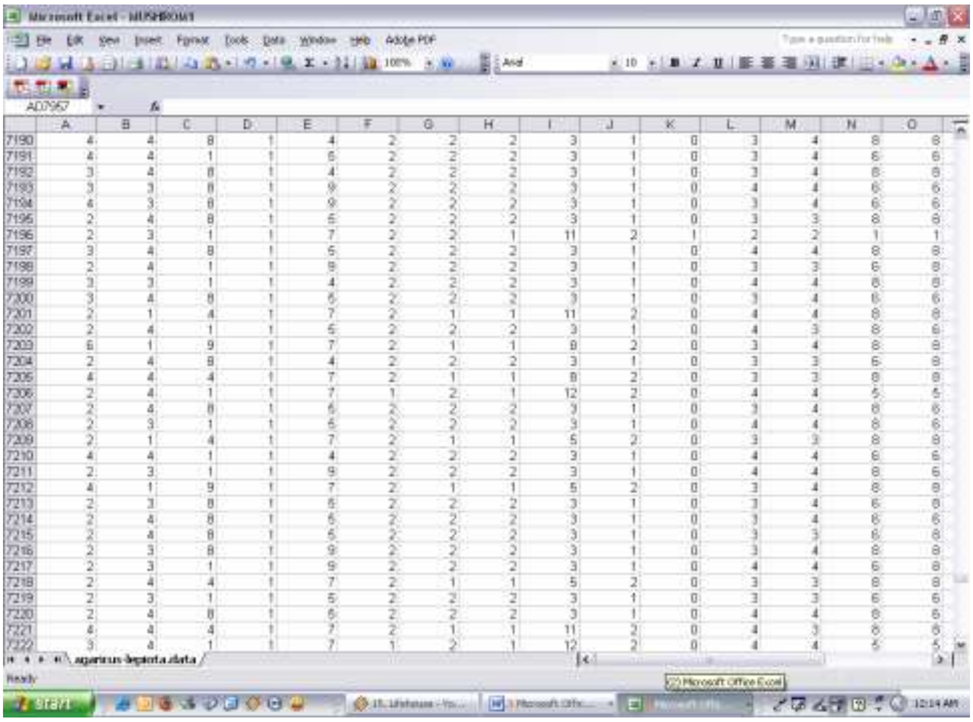

Figure 1: Sample of clean data

Neural Network Tool

The most common neural network model is the multi-layer preceptor (MLP). This type of neural network is known as a supervised network because it requires a desired output in order to learn. The goal of this type of network is

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

to create a model that correctly maps the input to the output using historical data so that the model can then be used to produce the output when the desired output is unknown. The MLP is a design that overcomes the shortcomings of the simple preceptor. The multi-layer preceptor can solve general nonlinear classification problems. A MLP is a hierarchical structure of several "simple" perceptions (with smooth transfer functions). MLP is a modeling and forecasting tool that uses Neural Network to model the data. It can be used to classify patterns or to predict values from data. Multi-layer preceptor is a supervised learning technique; it required that the data contain targets for training the network. The MLP with Back propagation consists of three layers: the input layers, where the data are introduced to the network; the hidden layer, where the data are processed and the output layer, where results for given are produced [4]. A graphical illustrate the MLP, is shown in the figure 2.

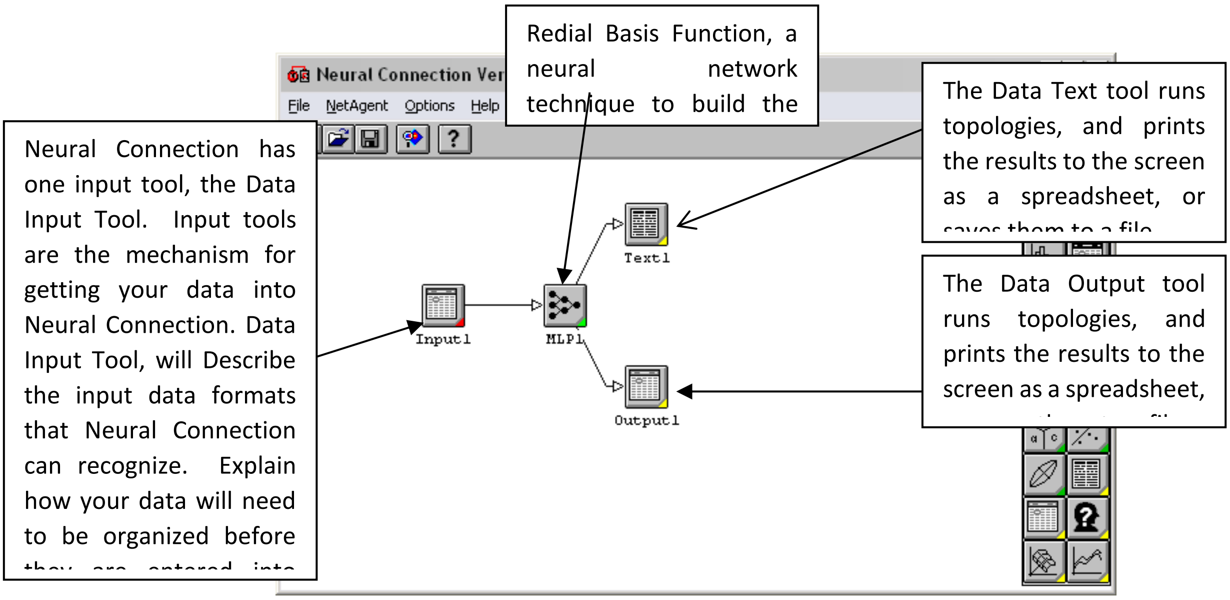

Figure 2: Multi-Layer Preceptor Model

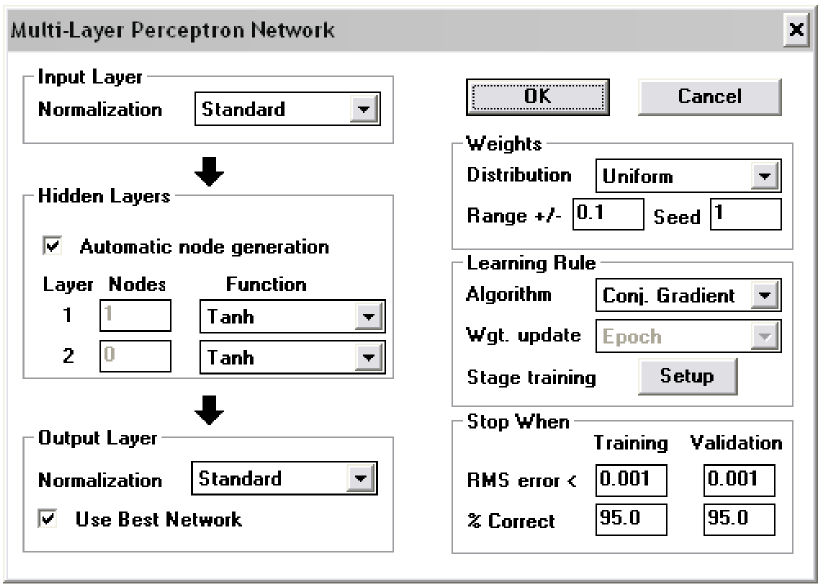

Figure 3: Multi-Layer Preceptor Network

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

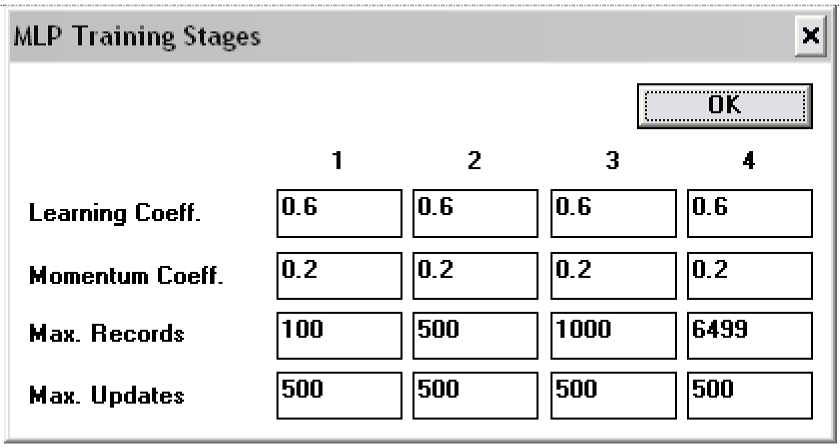

Figure 4: MLP Training Stages

The Experiments

Table 2: Data Distribution

Data Distribution	Percentage	Amount
Training data	80%	6499
Testing	10%	813
Validation data	10%	812

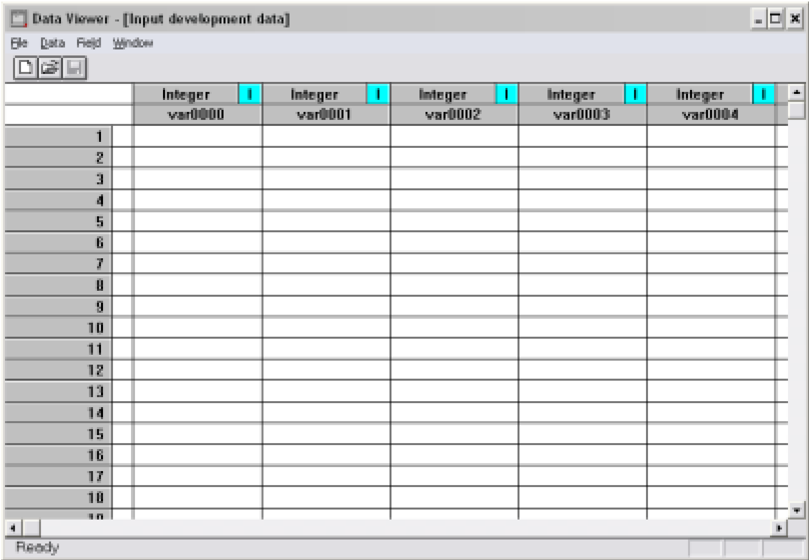

Figure 5: Data Viewer

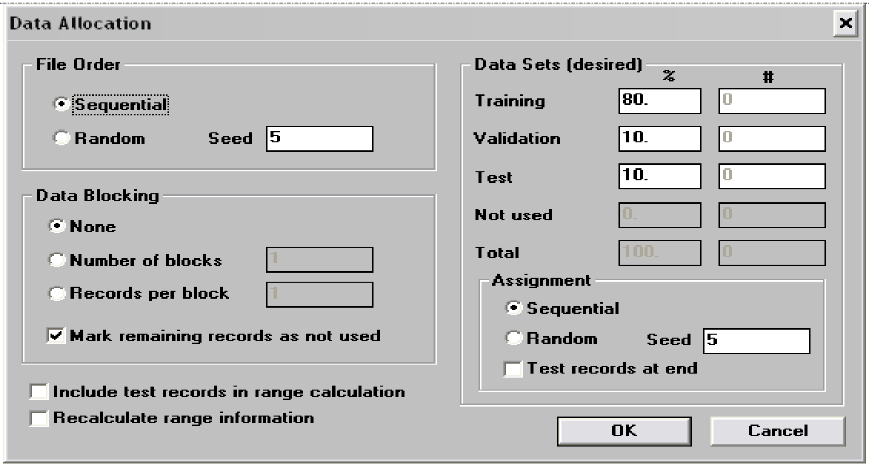

Figure 6: Data sets Allocation.

Before mining the mushrooms data, that is earlier clean the arrangement of the data is randomize to ensure fair distribution among all data. Since Neural Network can random the data automatically, therefore the instruction to random the data is stated in Figure 4.

RESULTS

To determine the most suitable hidden units

The following parameters are fixed but number of hidden unit will vary:
Learning rate= 0.1 Momentum rate =0.1
Activation function = sigmoid Stopping criteria = 95%

Experiment 4.1a

Several training and test results have been obtain by using different number of hiding unit, and the result is shown in table 3

Table 3: Result to determine the best-hidden unit

No of hidden unit	Accuracy	
	Training	**Test**
2	**93.43%**	**92.49%**
4	92.18%	91.63%
5	92.65%	91.75%
7	92.03%	91.50%
10	92.23%	91.50%
12	92.57%	91.38%
14	92.69%	91.75%
18	92.66%	91.63%
20	**92.91%**	**92.00%**
22	92.78%	91.50%

The results shown in table 3 indicate that highest test result was achieved when the number of hidden unit is 2 (92.49%) and 20(92.00%). For this experiment, hidden unit 2 and 20 will be selected to be used in the next experiment.

Experiment 4.1b

Several training and test results have been obtain by using different number of weight seed, and the result is shown in table 4.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

Table 4 : Result to determine the best-hidden unit

Weight seed	Hidden unit			
	2		20	
	Train	Test	Train	Test
1	93.43%	92.49%	92.91%	92.00%
2	92.61%	91.63%	92.48%	91.38%
3	91.81%	91.38%	92.58%	91.75%
4	91.84%	91.38%	92.38%	91.50%
5	91.74%	91.13%	92.80%	91.75%
6	91.98%	91.38%	92.83%	92.00%
7	92.95%	91.87%	92.69%	91.75%
8	92.78%	91.87%	92.81%	91.87%
9	93.17%	92.12%	92.55%	91.50%
10	92.81%	91.87%	92.31%	91.26%
Average	**92.51%**	**91.71%**	**92.63%**	**91.68%**

Table 4: Comparison

Hidden unit	Train	Test
2	**92.46%**	**91.68%**
20	92.37%	91.42%

Based on the result displayed in the table 4, hidden unit 2 produces a higher average accuracy than hidden unit 20. Therefore, hidden unit 2 can be considered as the more suitable number of hidden unit to be selected to be used in the next experiment.

To determine the most suitable learning rate
The following parameters are fixed but number of learning rate will vary:
Hidden unit= 2 Momentum rate =0.1
Activation function = sigmoid Stopping criteria = 95%

Experiment 4.2a
Several training and test results have been obtain by using different number of learning rate, and the result is shown in table 5.

Table 5 : Result to determine the most suitable learning rate

Learning rate	Accuracy	
	Training	Test
0.1	93.43%	92.49%
0.2	94.05%	93.10%
0.3	94.61%	93.72%
0.4	**94.97%**	**94.21%**
0.5	94.97%	**94.21%**
0.6	**94.95%**	**94.21%**
0.7	94.95%	94.09%
0.8	434	
0.9	377	
1.0	333	

Through the result in the table 5 the accuracy of the test is not consistent. The trend increase when learning rate increase until learning rate 0.7, afterwards the accuracy is slightly drops. Learning rate 0.8 to 1.0 the accuracy null (stuck), it show that the training stop. The highest learning percentage of accuracy occurs through learning 0.4, 0.5 and 0.6 (test: 94.21%). For further experiment, learning rate 0.4 and 0.6 will be selected in the next experiment because they show the best learning rate between the other and the lowest train percentage. The experiment result is shows in table 6.

Experiment 4.2b

Several training and test results have been obtain by using different number of weight seed and the result is shown in table 6

Table 6 : Result to determine the most suitable learning rate

Weight seed	**Learning rate**			
	0.4		**0.6**	
	Train	**Test**	**Train**	**Test**
1	94.97%	94.21%	94.95%	94.21%
2	94.41%	93.84%	94.81%	93.97%
3	94.17%	93.47%	94.60%	93.84%
4	94.60%	93.84%	94.97%	94.21%
5	94.41%	93.84%	94.81%	93.97%
6	94.86%	93.97%	95.00%	94.21%
7	94.97%	94.21%	94.97%	94.21%
8	94.91%	94.21%	94.95%	94.21%
9	94.97	94.21	**474**	-
10	94.89%	94.21%	94.95%	94.09%
Average	**94.72**	**94.00**	**94.89**	**94.10**

Learn unit	**Train**	**Test**
0.4	94.72	94.00
0.6	**94.89**	**94.10**

Based on the result displayed in the table 6, learning rate 0.6 give the highest average accuracy than learning rate 0.4. Therefore, learning rate 0.6 can be considered as the more suitable number of learning rate to be used in the next experiment.

To determine the most suitable momentum rate

Experiment 4.3a

Several training and test results have been obtain by using different number of momentum rate and the result is shown in table 7**.**

Table 7 : Result to determine the most suitable momentum rate

Momentum coffee	**Accuracy**	
	Training	**Test**
0.1	**94.95%**	**94.21%**
0.2	**94.95%**	**94.09%**
0.3	471	471
0.4	405	405
0.5	338	338
0.6	271	271

0.7	204	204
0.8	135	135
0.9	16	16
1.0	8	8

Through the result in the table 7, the accuracy of the test is not consistent. The trend increase when learning rate increase until Momentum coffee 0.2, afterwards the accuracy is slightly drops. Momentum coffee 0.3 to 1.0 the accuracy null (stuck), it show that the training stop. Therefore, Momentum coffee 0.1 and 0.2 will be selected in the next experiment.

Experiment 4.3b
Several training and test results have been obtain by using different number of weight seed and the result is shown in table 8.

Table 8 : Result to determine the most suitable momentum rate

Weight seed	**Momentum rate**			
	0.1		**0.2**	
	Train	**Test**	**Train**	**Test**
1	94.95%	94.21%	94.95%	94.21%
2	94.81%	93.97%	94.81%	93.97%
3	94.60%	93.84%	94.60%	93.84%
4	94.97%	94.21%	94.97%	94.21%
5	94.81%	93.97%	94.81%	93.97%
6	95.00%	94.21%	95.00%	94.21%
7	94.97%	94.21%	94.97%	94.21%
8	94.95%	94.21%	94.95%	94.21%
9	-	-	412	412
10	94.95%	94.09%	94.95%	94.09%
Average	94.89%	94.10%	94.93%	94.17%
Momentum rate	**Train**	**Test**		
0.1	94.89%	94.10%		
0.2	**94.93%**	**94.17%**		

Based on the result displayed in the table 8, momentum rate 0.2 give the highest average accuracy than learning rate 0.1. Therefore, learning rate 0.2 can be considered as the more suitable number of learning rate to be used in the next experiment.

To determine the best Activation Function
The following parameters are fixed but number of momentum rate will vary:
Hidden unit = 2 Momentum rate = 0.2
Learning rate = 0.6 Stopping criteria = 95%

Experiment 4.4a
Several training and test results have been obtain by using different number of weight seed and the result is shown in table 9.

Table 9 : Result to determine the best Activation Function

Weight seed	**Linear**		**Sigmoid**		**Tanh**	
	Train	**Test**	**Train**	**Test**	**Train**	**Test**
1	94.95%	94.21%	94.95%	94.09%	94.95%	94.09%

2	94.81%	93.97%	94.95%	94.21%	94.95%	94.21%
3	94.60%	93.84%	94.71%	93.97%	94.71%	93.97%
4	94.97%	94.21%	94.97%	94.21%	94.97%	94.21%
5	94.81%	93.97%	94.94%	94.21%	94.94%	94.21%
6	95.00%	94.21%	95.00%	94.33%	95.00%	94.33%
7	94.97%	94.21%	94.97%	94.21%	94.97%	94.21%
8	94.95%	94.21%	94.95%	94.21%	94.95%	94.21%
9	-	-	**412**	**412**	**412**	**412**
10	94.95%	94.09%	94.95%	94.09%	94.95%	94.09%
Average	**94.89%**	**94.10%**	**94.93%**	**94.17%**	**94.93%**	**94.17%**

From Table 9, the result of the experiment shows that both sigmoid and Tanh show a decreasing on the percentage of test accuracy, when the Weight Seed is decreasing. Therefore sigmoid will be chosen as the best activation function for this model, because it is the most one used.

To determine the best number of Epoch
The following parameters are fixed but number of momentum rate will vary:
Hidden unit = 2 Momentum rate = 0.2
Activation function = Sigmoid Learning rate = 0.6

Experiment 4.5a
Several training and test results have been obtain by using different number of Epoch and the result is shown in table 10.

Table 10: for determining best number of Epoch

Epoch	**Sigmoid**	
	Training	**Test**
100	93.55%	92.36%
200	**94.20%**	**93.35%**
300	**94.91%**	**94.21%**
400	94.95%	94.21%
500	513	513
600	**513**	**513**
700	-	-
800	-	-
900	-	-
1000	-	-

Through result in table 10, the accuracy of the test is not consistent. The trend increase when sigmoid increase until Epoch 400, afterwards the accuracy is slightly drops. Epoch 500 to 1000, the accuracy null (stuck), it show that the training stop. The highest sigmoid percentage of accuracy occurs through Epoch 200,300 and 400 (test: 93.35% and 94.21%). For further experiment, Epoch 200 and 300 will be used in the next experiment because they show the best Epoch result between the other and the lowest train percentage. The experiment result is shows in table 11.

Experiment 4.5b
Several training and test results have been obtain by using different number of weight seed and the result is shown in table 11.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

Table 11 : for determining best number of Epoch

Weight seed	**Epoch**			
	200		**300**	
	Train	**Test**	**Train**	**Test**
1	95.17%	94.33%	94.95%	94.21%
2	94.08%	93.47%	94.41%	93.84%
3	93.94%	93.10%	94.17%	93.47%
4	94.03%	93.23%	94.58%	93.84%
5	94.03%	93.47%	94.41%	93.84%
6	94.20%	93.35%	94.92%	94.09%
7	94.49%	93.84%	94.95%	94.21%
8	94.23%	93.35%	94.94%	94.21%
9	94.52%	93.23%	95.14%	93.97%
10	94.20%	93.35%	94.91%	94.21%
Average	94.29%	93.47%	**94.74%**	**93.99%**

Epoch	**Train**	**Test**
200	94.29%	93.47%
300	**94.74%**	**93.99%**

Based on result displayed in table 4.5b, Epoch 300 gives the highest average accuracy than Epoch 200. Therefore, Epoch 300 can be considered as the more suitable number of Epoch.

The Network Architecture

The architecture of the model for the species of the mushroom in Lepiota and Agarics data after this experiment is shown below:

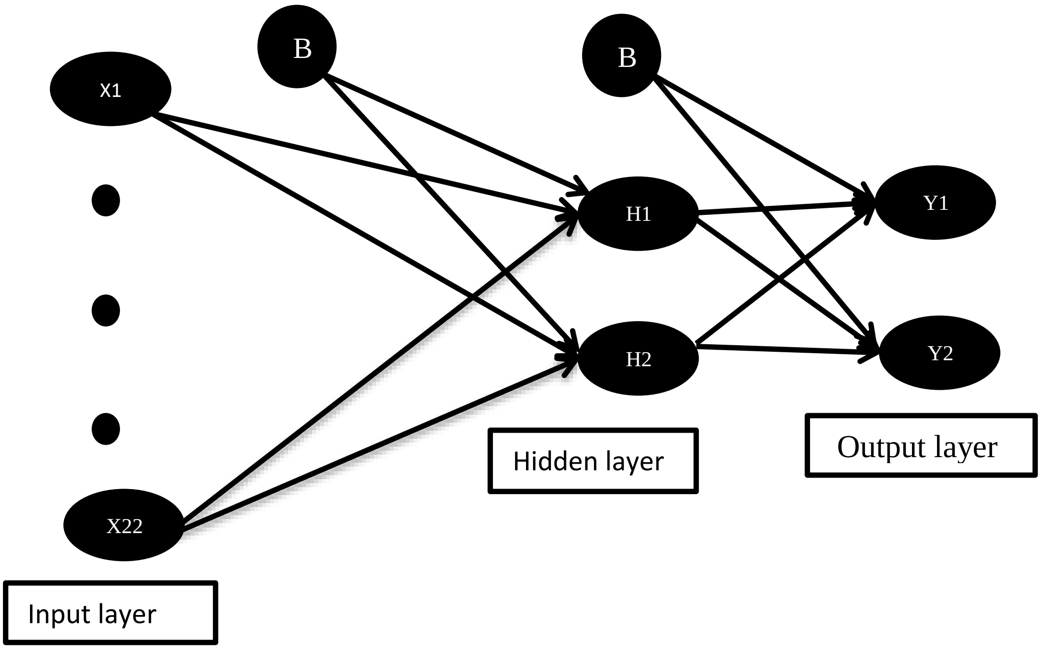

Figure 7 : the network activation

CONCLUSION

Based to all experiments that have been done for all test and train required to get for predicting whether the mushrooms definitely edible or poisonous. The result showed that the best-hidden unit is 2, the best learning rate is 0.6, the best moment rate is 0.2, the best activation function is sigmoid and best result of epoch is 300.

REFERENCES

[1] Hongmei Yan, Jun Zheng, Yingtao Jiang, Chenglin Peng, and Qinghui Li, "Development Of A Decision Support System For Heart Disease Diagnosis Using Multilayer Perceptron", Bioengineering Institute, Chongqing University, China, 2003.

[2] Kyung-Joong Kim, Sung-Bae Cho, "Prediction of colon cancer using an evolutionary neural network ", Department of Computer Science, Yonsei University, 134 Shinchon- dong, Sudaemoon- ku, Seoul 120-749, South Korea, 2003.

[3] Shigeaki Ogose, Shindoh Sasaki, Kazumasa Itoh, "Mobile Communication Systems with Multi-layered Wireless Network using ad *hoc* Network ", Faculty of Engineering, Kagawa University, 2217-20 Hayashicho, Takamatsu, 761-0396 Japan, 2003.

[4] Konstantinos Nikellis1, Yorgos Koutsoyannopoulos', Sotiris Bantas' and Nikolaos K. Uzunoglu, "A Fast Full-Wave Modeling Methodology for Stripline Structures with Vertical Interconnects in Multi-Layer Dielectrics", Helic S.A Argyroupolis, Athens, 164 52, Greece 'National Technical University of Athens Zografos, Athens, 157 73, Greece, 2004.

[5] Zakalyk L. I., Korzh RO., Danchyshyn I. V., Ivanyk R. V. "Tense- deformation Condition in Multi-layered Structure "– Radio Engineering Faculty, Lviv Polytechnic National University, S. Bandery Str., 12, Lviv, 79013, UKRAINE, 2002.

[6] E. Oki, D. Shimazaki, K. Shiomoto, N. Matsuura, W. Imajuku, and N. Yamanaka, " Performance of Distributed- Controlled Dynamic Wavelength- Conversion GMPLS Networks," Optical Communications Networks, No. 1, pp. 355 - 358, Nov, 2002.

[7] Chih- Hsiu Lin and An- Yeu (Andy) Wu, " ROBUST DECISION FEEDBACK EQUALIZER DESIGN USING SOFTTHRESHOLD- BASED MULTI- LAYER DETECTION SCHEME" Graduate Institute of Electronics Engineering, and Department of Electrical Engineering, National Taiwan University Taipei, 106, Taiwan, R. O. C., 2004.

[8] Lalitha Agnihotri, Nevenka Dimitrova, Mario Soletic, " MULTI-LAYERED VIDEOTEXT EXTRACTION METHOD" Philips research USA, briarcliff Manor, NY 10510, USA., 2002.

[9] Y. Pei and J. W. Modestino, " Multilayered H. 263+ Video Encoding and Delivery over an AWGN Channel: a Joint Source- Channel Coding Approach," Submitted to *IEEE Trans. On Image Pmc.*, 2000.

[10] *Chung-Yu Wu*[1] *and Jen-Chieh Wang*, "OPTIMAL STRUCTURE OF INTERCONNECTION LINES FOR GHz GAIGA-SCALE NANO-CMOS SYSTEM-ON-CHIP DESIGN" Nanoelectronics and Giga-Scale Syatems Lab., National Chiao Tung University, Taiwan [2]National Chip Implementation Center, Taiwan, 2004.

[11] UCI Machine Learning Repository, (1987, April 27) 'Mushroom Database'.(Center for Machine Learning and Intelligent Systems),Available: https://archive.ics.uci.edu/ml/datasets/Mushroom (Accessed: 2016, February 20).