International Journal of Advanced Science and Technology
Vol.124 (2018), pp. 51-59
http://dx.doi.org/10.14257/ijast.2018.124.05

Predicting Car Mileage per Gallon

Mohsen Afana, Jomana Ahmed, Bayan Harb, Bassem S. Abu-Nasser
Department of Information Technology
Faculty of Engineering & Information Technology
Al-Azhar University - Gaza, Palestine

Abstract

In this paper an Artificial Neural Network (ANN) model was used to help cars dealers recognize the many characteristics of cars, including manufacturers, their location and classification of cars according to several categories including: Make, Model, Type, Origin, DriveTrain, MSRP, Invoice, EngineSize, Cylinders, Horsepower, MPG_Highway, Weight, Wheelbase, Length. ANN was used in prediction of the number of miles per gallon when the car is driven in the city(MPG_City). The results showed that ANN model was able to predict MPG_City with 97.50 % accuracy. The factor of DriveTrain has the most influence on MPG_City evaluation. Similar studies can be carried out for the evaluation of other characteristics of cars.

Keywords: *Artificial Neural Networks, Dataset, Cars, ANN, Cars Types*

1. Introduction

This study evaluates the utilization of neural networks for predicting suitability of a car. It justifies the use of neural networks in this industry for the prediction process. Generally, car manufacturing industries include design, development, manufacturing, marketing and sale of different equipment for motor vehicles. The set of companies and factories involved in design, manufacturing, marketing, and sale of motor vehicles are a part of this industry.

In 2017, more than 97.3 million motor vehicles including regular cars and commercial cars were manufactured around the world. In 2017, a total number of 96.8 million cars were sold in the world, with 25.2 million sold in Europe, 32.6 million sold in Asia and Pacific Region, 25.8 million sold in the US and Canada, 4.6 million sold in Latin America, 7.4 million sold in the Middle East, and 1.2 million sold in Africa[7].

When the market was suffering a recession in the US and Japan, Asia and South America expressively grew and got sturdier. It appears that big markets in Russia, Brazil, India and China have practiced a rapid growth. The car industry, as one of the biggest industries in the world holding a great amount of persons, financial and time resources, is in grim need of precise predictions of its future and its contestants to reach great and sensitive decisions.

Maybe one of the major concerns of the managers and manufacturers in the car industry and the investors in this field is the prediction of cars sales and arrangement for the future manufacturing volume. If managers can have a more precise prediction regarding the future sales volume and car demand, they can unconditionally enhance the investment volume, used workforce and optimally use time to reach optimal decisions and convey their instruction plans.

ISSN: 2005-4238 IJAST

2. Objectives

The main aim of this study is to determine Cars which would be suitable and do well for the customer. There are many companies who make cars around the world each company makes a lot of cars for a different customer.

The quality of car delivered for any customer influences the response to the car dealer, so we must learn how to choose correctly the car because it affects the economy of society.

3. The Artificial Neural Networks

Artificial neural networks are one of the key tools used in machine learning. As the neural fragment of their name proposes, they are brain-inspired systems which are envisioned to imitate how persons learn. Neural networks made of input and output layers, besides (in most cases) a hidden layer comprising units that transmute the input into roughly that the output layer can practice. They are outstanding tools for discovering patterns which are far excessively complex or many for a human programmer to quote and teach the machine to identify [1-3].

While neural networks (also named perceptrons) have been around since the 1940s, it is only in the last several years where they have become a successful application of artificial intelligence. This is due to the algorithm backpropagation, which allows networks to adjust their hidden layers of neurons in situations where the outcome doesn't match what the creator is hoping for — like a network designed to recognize oranges, which misidentifies an apple, for example[3-5].

Another significant issue in ANN has been the appearance of deep learning neural networks, in which different layers of a multilayer network excerpt different features until it can recognize what it is looking for.

For a basic idea of how a deep learning neural network learns, imagine a factory line. After the raw materials (the data set) are input, they are then passed down the conveyer belt, with each subsequent stop or layer extracting a different set of high-level features. If the network should recognize an object, the first layer might examine the brightness of its pixels [6].

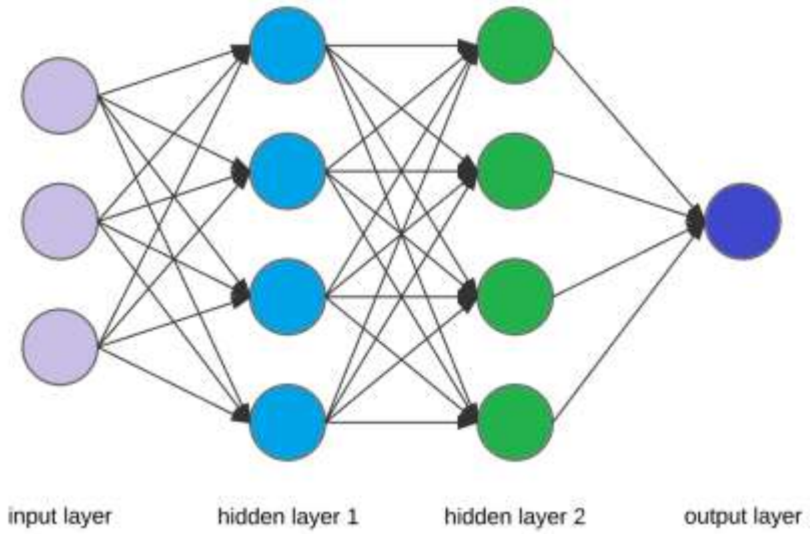

Figure 1: Typical ANN architecture

The next layer could then identify any edges in the image, based on lines of similar pixels. After this, another layer may recognize textures and shapes, and so on. By the time the fourth or fifth layer is reached, the deep learning net will have created complex feature detectors. It can figure out that certain image elements (such as a pair of eyes, a nose, and a mouth) are commonly found together.

Once this is done, the researchers who have trained the network can give labels to the output, and then use backpropagation to correct any mistakes which have been made. After a while, the network can carry out its own classification tasks with no need of humans to help every time.

Beyond this, there are different types of learning, such as supervised or unsupervised learning or reinforcement learning, in which the network learns for itself by trying to maximize its score — as memorably carried out by Google DeepMind's Atari game-playing bot [4].

3.1 Types of neural network

There are multiple types of neural network, each of which comes with their own specific use cases and levels of complexity. The most basic type of neural net is called a feedforward neural network, in which information travels in only one direction from input to output.

A more widely used type of network is the recurrent neural network, in which data can flow in multiple directions. These neural networks possess greater learning abilities and are widely used for more complex tasks such as learning handwriting or language recognition.

There are also convolutional neural networks, Boltzmann machine networks, Hopfield networks, and a variety of others. Picking the right network for your task depends on the data you have to train it with, and the specific application you have in mind. Sometimes, it may be desirable to use multiple approaches, such as would be the case with a challenging task like voice recognition.

3.2 The tasks that neural network can do:

From making cars drive autonomously on the roads, to generating shockingly realistic CGI faces, to machine translation, to fraud detection, to reading our minds, to recognizing when a cat is in the garden and turning on the sprinklers; neural nets are behind many of the biggest advances in Artificial Intelligence.

Broadly speaking, however, they are designed for spotting patterns in data. Specific tasks could include classification (classifying data sets into pre-defined classes), clustering (classifying data into different undefined categories), and prediction (using past events to guess future ones, like the stock market or movie box office).

3.3 How ANN learns:

In the same way that we learn from experience in our lives, neural networks require data to learn. In most cases, the more data that can be thrown at a neural network, the more accurate it will become. Think of it like any task you do over and over. Over time, you gradually get more efficient and make fewer mistakes.

When researchers or computer scientists set out to train a neural network, they typically divide their data into three sets. First is a training set, which helps the network establish the various weights between its nodes. After this, they fine-tune it using a validation data set. Finally, they'll use a test set to see if it can successfully turn the input into the desired output.

3.4 Limitations of neural networks

On a technical level, one of the bigger challenges is the amount of time it takes to train networks, which can require a considerable amount of compute power for more complex tasks. The biggest issue, however, is that neural networks are black boxes, in which the user feeds in data and receives answers. They can fine-tune the answers, but they don't have access to the exact decision making process.

This is a problem where many researchers are actively working on, but it will only become more pressing as artificial neural networks play a bigger and bigger role in our lives.

4. Methodology

By looking deeply through the cars types and soliciting the experience of human experts on cars' performance, several factors that are having influence customer response were outlined. These factors were cautiously studied and synchronized into a convenient number appropriate for computer coding within the Just Neural Network (JNN) environment.

These factors were classified as input variables as shown in table 1.

The output variables embody some likely levels of performance of cars in terms of Excel file with a study summary performance of cars as shown in table 2.

4.1 The Input Variables:

This car dataset has attributes relating to cars such Make, model, type, organ, invoice, Cylinders, etc.

Table1: Input Attributes

Attributes:	Usage:	Type:	Range:
Make	Input	Alphabetic	
Model	Input	Alphabetic	
Type	Input	Alphabetic	
Origin	Input	Alphabetic	
DriveTrain	Input	Alphabetic	
MSRP	Input	Numeric	$10,280 to $192,465
Invoice	Input	Numeric	$9,875 to $173,560
EngineSize	Input	Numeric	1 to 8
Cylinders	Input	Numeric	3 to 12
Horsepower	Input	Numeric	73 to 500
MPG_Highway	Input	Numeric	12 to 66
Weight	Input	Numeric	1,850 to 7,190
Wheelbase	Input	Numeric	89 to 143
Length	Input	Numeric	144 to 238

4.2 The output Variable:

The attribute which represents the output variable is called MPG_City (Number of miles per gallon when the car is driven in the city).

Table2: shows the output variable

Attributes:	**Usage:**	**Type:**	**Values:**	**Range:**
MPG_City	Output	Numeric	Number of miles per gallon	10 to 60

4.3 Network design and optimization

The main goal of the work presented in this section was done to find which one of the various feed-forward backpropagation ANN combinations and architectures would be the best for predicting car mileage per gallon when driven in the city and also the network architecture that would be suitable for the small dataset available for the training.

The design and optimization of the ANN network was implemented out in three steps. In the first step the fundamental parameters for the network, like the transfer functions and the training functions have been examined.

In the second step the pre-processing of the data that is proper for the network has been examined. Finally,

In the third step of the network optimization has been implemented to find the most proper number of hidden layers and the number of proper neurons in each layer. It has been proposed by many researchers that the selection of all the factors that have been defined above is a trial and error process [1].

4.4 Evaluation of the ANN Model

There are 628 instances in the car data set. The data set then was divided into a training set (428) with a percentage of (69%) and a validation set (200) with a percentage of (31%).

The final optimized ANN model comprised five layers:

- One input layer which has 14 input (14 attributes),
- Three hidden layers, the first hidden layer has three neurons, the second hidden layer has one neuron, and the third hidden layer has 7 neurons.
- One output layer with one neuron.

The final optimized ANN model was trained with the training data set and then validated with the validation data set.

The final optimized ANN model architecture is shown in Figure 3. The most important factors contributing in ANN model was determined and the most influential factor was DriveTrain as shown in Figure 4.

The training and validation summary is shown in Figure 2 and Figure 5. The accuracy of predication of MPG_City was 99.5% and the number of cycles was 27,646 as shown in Figure 5.

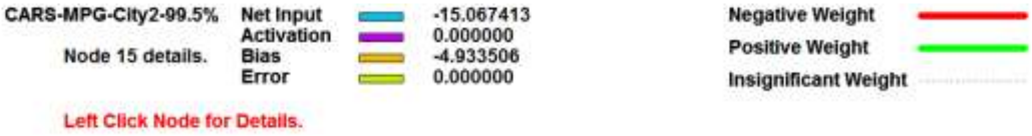

Figure 2: Final ANN model summary

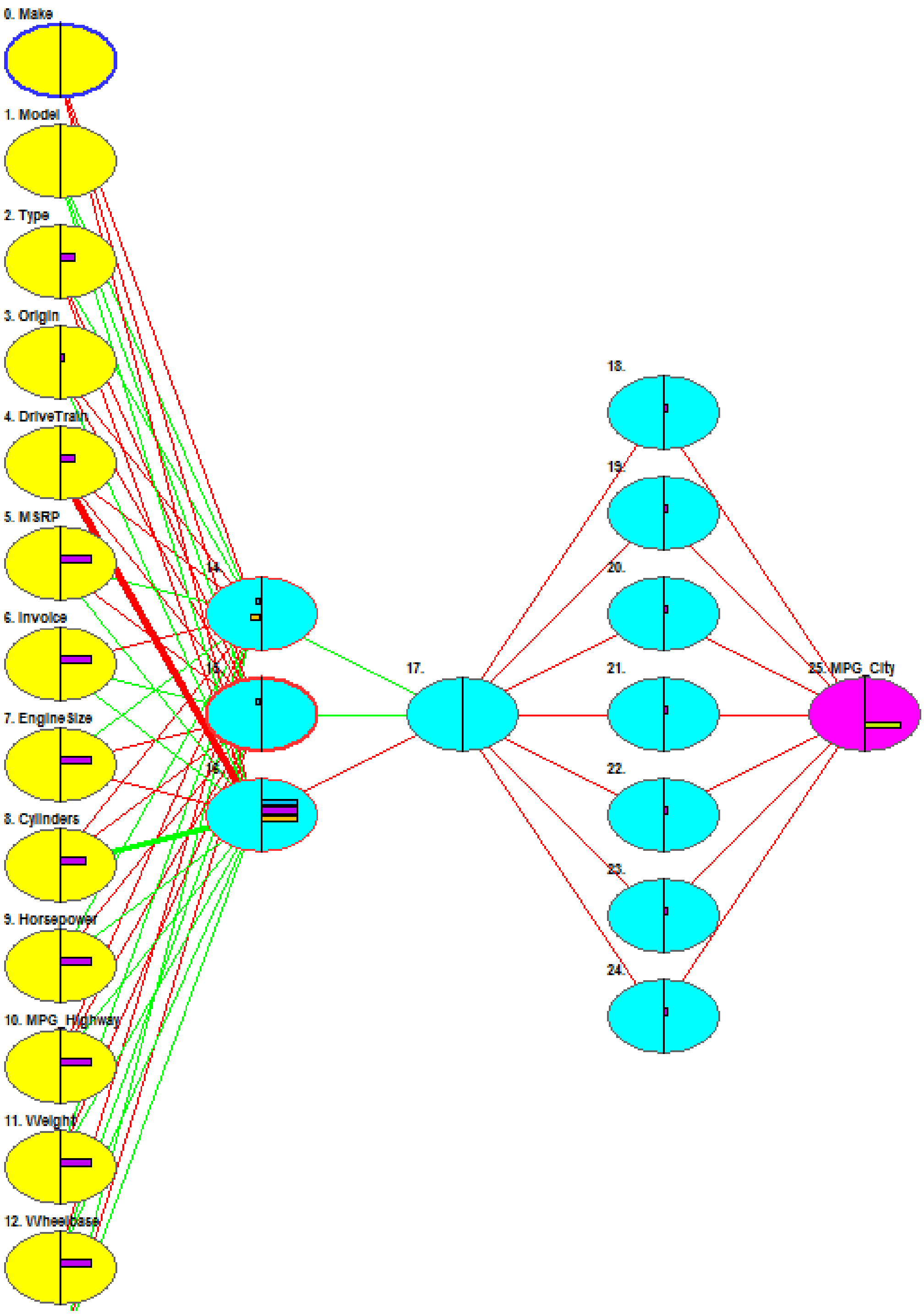

Figure 3: Final ANN model Architecture

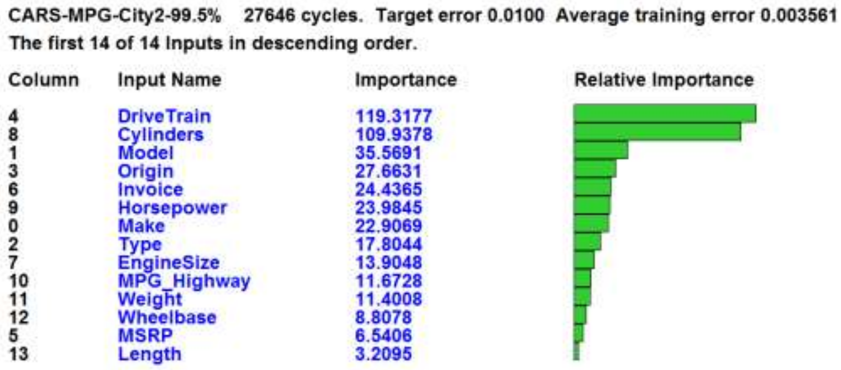

Figure 4: Most influencing factories on ANN model

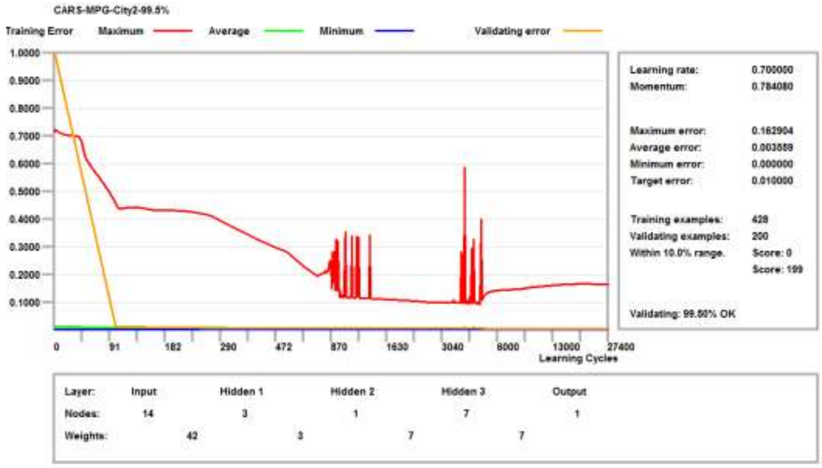

Figure 5: Training and validation Summary

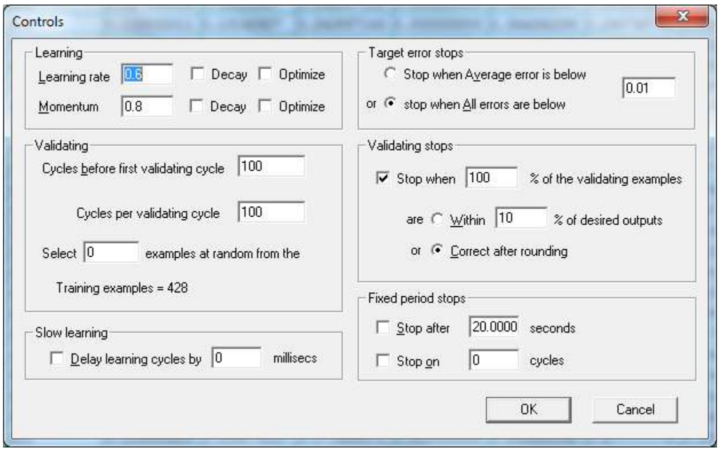

Figure 6: Control of the parameter of the ANN Model

5. Conclusion

Acceptance of the product of increasing concern, the Car makers must know which factors influencing customers' buying decisions. Recently, deep evaluation of the product acceptability has been done, unfortunately, manufacturers repeatedly misinterpret the real needs of consumers, and how to test the acceptability of the product is the vital matter of product development.

In this paper, for the evaluation of car lines empirical research, using ANN for the evaluation of automotive forecasting a comparative study, the experimental results show that, ANN can assess car evaluation prediction well. In the evaluation of the car: the most significant factors of the ANN model were identified, and the accuracy of predication of the MPG_City was 99.5%.

References

1. Abu-Naser, S. S. (2012). "Predicting learners performance using artificial neural networks in linear programming intelligent tutoring system." International Journal of Artificial Intelligence & Applications 3(2): 65.
2. Afana, M., et al. (2018). "Artificial Neural Network for Forecasting Car Mileage per Gallon in the City." International Journal of Advanced Science and Technology 124: 51-59.
3. Ahmed, A., et al. (2019). "Knowledge-Based Systems Survey." International Journal of Academic Engineering Research (IJAER) 3(7): 1-22.
4. Alajrami, E., et al. (2019). "Blood Donation Prediction using Artificial Neural Network." Blood 3(10): 1-7.
5. Alghoul, A., et al. (2018). "Email Classification Using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 2(11): 8-14.
6. Alkronz, E. S., et al. (2019). "Prediction of Whether Mushroom is Edible or Poisonous Using Back-propagation Neural Network." International Journal of Academic and Applied Research (IJAAR) 3(2): 1-8.
7. Al-Massri, R., et al. (2018). "Classification Prediction of SBRCTs Cancers Using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 2(11): 1-7.
8. Al-Mubayyed, O. M., et al. (2019). "Predicting Overall Car Performance Using Artificial Neural Network." International Journal of Academic and Applied Research (IJAAR) 3(1): 1-5.
9. Al-Shawwa, M. and S. S. Abu-Naser (2019). "Predicting Birth Weight Using Artificial Neural Network." International Journal of Academic Health and Medical Research (IJAHMR) 3(1): 9-14.
10. Al-Shawwa, M. and S. S. Abu-Naser (2019). "Predicting Effect of Oxygen Consumption of Thylakoid Membranes (Chloroplasts) from Spinach after Inhibition Using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 3(2): 15-20.
11. Al-Shawwa, M., et al. (2018). "Predicting Temperature and Humidity in the Surrounding Environment Using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 2(9): 1-6.
12. Anderson, J., et al. (2005). "Adaptation of Problem Presentation and Feedback in an Intelligent Mathematics Tutor." Information Technology Journal 5(5): 167-207.
13. Ashqar, B. A. M. and S. S. Abu-Naser (2019). "Identifying Images of Invasive Hydrangea Using Pre-Trained Deep Convolutional Neural Networks." International Journal of Academic Engineering Research (IJAER) 3(3): 28-36.
14. Ashqar, B. A. M. and S. S. Abu-Naser (2019). "Image-Based Tomato Leaves Diseases Detection Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 2(12): 10-16.
15. Ashqar, B. A., et al. (2019). "Plant Seedlings Classification Using Deep Learning." International Journal of Academic Information Systems Research (IJAISR) 3(1): 7-14.
16. Baker, J., et al. "& Heller, R.(1996)." Information Visualization. Information Technology Journal 7(2).
17. Baker, J., et al. (1996). "Information Visualization." Information Technology Journal 7(2): pp: 403-404.
18. Baraka, M. H., et al. (2008). "A Proposed Expert System For Guiding Freshman Students In Selecting A Major In Al-Azhar University, Gaza." Journal of Theoretical & Applied Information Technology 4(9).
19. Barhoom, A. M., et al. (2019). "Predicting Titanic Survivors using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 3(9): 8-12.
20. Chand, P. S., et al. (2008). "MADAMS: Mining and Acquisition of Data by ANT-MINER Samples." Journal of Theoretical & Applied Information Technology 4(10).
21. Chen, R.-S., et al. (2008). "Evaluating structural equation models with unobservable variables and measurement error." Information Technology Journal 10(2): 1055-1060.
22. Dalffa, M. A., et al. (2019). "Tic-Tac-Toe Learning Using Artificial Neural Networks." International Journal of Engineering and Information Systems (IJEAIS) 3(2): 9-19.
23. El_Jerjawi, N. S. and S. S. Abu-Naser (2018). "Diabetes Prediction Using Artificial Neural Network." International Journal of Advanced Science and Technology 121: 55-64.
24. El-Khatib, M. J., et al. (2019). "Glass Classification Using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 3(2): 25-31.
25. Elzamly, A., et al. (2015). "Classification of Software Risks with Discriminant Analysis Techniques in Software planning Development Process." International Journal of Advanced Science and Technology 81: 35-48.
26. Elzamly, A., et al. (2015). "Predicting Software Analysis Process Risks Using Linear Stepwise Discriminant Analysis: Statistical Methods." Int. J. Adv. Inf. Sci. Technol 38(38): 108-115.
27. Elzamly, A., et al. (2016). "A New Conceptual Framework Modelling for Cloud Computing Risk Management in Banking Organizations." International Journal of Grid and Distributed Computing 9(9): 137-154.
28. Elzamly, A., et al. (2017). "Predicting Critical Cloud Computing Security Issues using Artificial Neural Network (ANNs) Algorithms in Banking Organizations." International Journal of Information Technology and Electrical Engineering 6(2): 40-45.
29. Elzamly, A., et al. (2019). "Critical Cloud Computing Risks for Banking Organizations: Issues and Challenges." Religación. Revista de Ciencias Sociales y Humanidades 4(18): 673-682.

30. Heriz, H. H., et al. (2018). "English Alphabet Prediction Using Artificial Neural Networks." International Journal of Academic Pedagogical Research (IJAPR) 2(11): 8-14.
31. Jamala, M. N. and S. S. Abu-Naser (2018). "Predicting MPG for Automobile Using Artificial Neural Network Analysis." International Journal of Academic Information Systems Research (IJAISR) 2(10): 5-21.
32. Kashf, D. W. A., et al. (2018). "Predicting DNA Lung Cancer using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 2(10): 6-13.
33. Khalil, A. J., et al. (2019). "Energy Efficiency Predicting using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 3(9): 1-8.
34. Li, L., et al. (2011). "Hybrid Quantum-inspired genetic algorithm for extracting association rule in data mining." Information Technology Journal 12(4): 1437-1441.
35. Abu Naser, S. S., et al. (2016). "Design and Development of Mobile University Student Guide." Journal of Multidisciplinary Engineering Science Studies (JMESS) 2(1): 193-197.
36. Abu Naser, S. S., et al. (2016). "Design and Development of Mobile Blood Donor Tracker." Journal of
37. Marouf, A. and S. S. Abu-Naser (2018). "Predicting Antibiotic Susceptibility Using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 2(10): 1-5.
38. Masri, N., et al. (2019). "Survey of Rule-Based Systems." International Journal of Academic Information Systems Research (IJAISR) 3(7): 1-23.
39. Naser, S. S. A. "TOP 10 NEURAL NETWORK PAPERS: RECOMMENDED READING– ARTIFICIAL INTELLIGENCE RESEARCH."
40. Nasser, I. M. and S. S. Abu-Naser (2019). "Artificial Neural Network for Predicting Animals Category." International Journal of Academic and Applied Research (IJAAR) 3(2): 18-24.
41. Nasser, I. M. and S. S. Abu-Naser (2019). "Lung Cancer Detection Using Artificial Neural Network." International Journal of Engineering and Information Systems (IJEAIS) 3(3): 17-23.
42. Nasser, I. M. and S. S. Abu-Naser (2019). "Predicting Books' Overall Rating Using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 3(8): 11-17.
43. Nasser, I. M. and S. S. Abu-Naser (2019). "Predicting Tumor Category Using Artificial Neural Networks." International Journal of Academic Health and Medical Research (IJAHMR) 3(2): 1-7.
44. Nasser, I. M., et al. (2019). "A Proposed Artificial Neural Network for Predicting Movies Rates Category." International Journal of Academic Engineering Research (IJAER) 3(2): 21-25.
45. Nasser, I. M., et al. (2019). "Artificial Neural Network for Diagnose Autism Spectrum Disorder." International Journal of Academic Information Systems Research (IJAISR) 3(2): 27-32.
46. Nasser, I. M., et al. (2019). "Developing Artificial Neural Network for Predicting Mobile Phone Price Range." International Journal of Academic Information Systems Research (IJAISR) 3(2): 1-6.
47. Ng, S., et al. (2010). "Ad hoc networks based on rough set distance learning method." Information Technology Journal 10(9): 239-251.
48. Owaied, H. H., et al. (2009). "Using rules to support case-based reasoning for harmonizing melodies." Journal of Applied Sciences 11(14): pp: 31-41.
49. Sadek, R. M., et al. (2019). "Parkinson's Disease Prediction Using Artificial Neural Network." International Journal of Academic Health and Medical Research (IJAHMR) 3(1): 1-8.
50. Salah, M., et al. (2018). "Predicting Medical Expenses Using Artificial Neural Network." International Journal of Engineering and Information Systems (IJEAIS) 2(20): 11-17.
51. Sulisel, O., et al. (2005). "Growth and Maturity of Intelligent Tutoring Systems." Information Technology Journal 7(7): 9-37.
52. Zaqout, I., et al. (2015). "Predicting Student Performance Using Artificial Neural Network: in the Faculty of Engineering and Information Technology." International Journal of Hybrid Information Technology 8(2): 221-228.
53. Abu-Nasser, Bassem. "Medical Expert Systems Survey." *International Journal of Engineering and Information Systems (IJEAIS)* 1, no. 7 (2017): 218-224.
54. Abu-Nasser, Bassem S., and Samy S. Abu-Naser. "Cognitive System for Helping Farmers in Diagnosing Watermelon Diseases." *International Journal of Academic Information Systems Research (IJAISR)* 2, no. 7 (2018): 1-7.
55. Abu-Nasser, Bassem S., and Samy S. Abu Naser. "Rule-Based System for Watermelon Diseases and Treatment." *International Journal of Academic Information Systems Research (IJAISR)* 2, no. 7 (2018): 1-7.