

# Predicting Books' Overall Rating Using Artificial Neural Network

Ibrahim M. Nasser, Samy S. Abu-Naser

Department of Information Technology,  
Faculty of Engineering and Information Technology,  
Al-Azhar University, Gaza, Palestine  
Email: [Azhar.ibrahimn@gmail.com](mailto:Azhar.ibrahimn@gmail.com)

**Abstract:** We developed an Artificial Neural Network (ANN) model for predicting the overall rating of books. The prediction is based on some Factors (bookID, title, authors, isbn, language\_code, isbn13, # num\_pages, ratings\_count, text\_reviews\_count), which used as input variables and (average\_rating) as output for our ANN predictive model. Our model established, trained, and validated using data set, which its title is "Goodreads-books". Model evaluation showed that the ANN model is able to predict correctly 99.90% of the validation instances.

**Keywords:** Predictive Analysis, Data Mining, Machine Learning, Artificial Neural Networks, Books Rates, Goodreads

## 1. INTRODUCTION

In this study, we will analyze the dataset which contains various information of the books on the website of the world's largest book archive and book proposal site GoodReads. This dataset also includes information on the names, writers and spelling languages of books, as well as the rating and total score based on the votes given by various users. Artificial neural networks (ANNs) will be used for the analysis. Artificial neural networks are like biological neural networks and offer a technique, which solves the problem of prediction [3]. Neural networks contain input, hidden and output layers. Hidden layers

convert the input into usable thing to the output layer [5]. The ANN Model goes through training and validation on a dataset. Training in which that the network is trained is done on a dataset. Then a configuration is done to the weights of the connections between neurons. Validation in which that the network is validated to determine the prediction of a new dataset [6]. In this study, we used about 33% of the dataset instances for network validation, the remaining 67% for training. ANN Architecture is shown in figure 1.

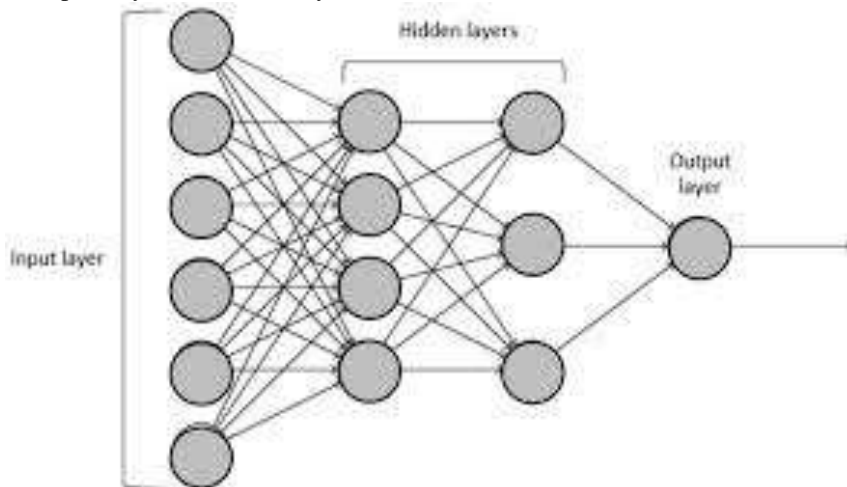


Figure 1: ANN Architecture

## 1. Literature Review

Nasser et.al built neural network based models for the purpose of classification, prediction, and diagnosis. They proposed an ANN to predict the category of movie's rate [1], predict the price range of mobile phone [48], predict the

category of animal [47], diagnosis the category of tumor [47], for diagnosis of Autism [49], and for diagnosis of lung cancer [40,46]. Authors in [12 – 26] developed ANN models that are very useful and valuable for classification, prediction, and diagnosis.

## 2. Methodology

We downloaded a data set from *kaggle* that contains books information from *goodreads* application/website. This dataset created by the user *Soumik* [4].

We did some preprocessing on the data, and then we trained our ANN model and validated it.

## 3. Original Dataset Description

**Table 1: Original Dataset Description**

#	Attribute	Description	Type
1.	bookID	A unique Identification number for each book.	Integer
2.	title	The name under which the book was published.	String
3.	authors	Names of the authors of the book. Multiple authors are delimited with -.	String
4.	average_rating	The average rating of the book received in total.	Real
5.	isbn	Another unique number to identify the book, the International Standard Book Number.	Long
6.	language_code	Helps understand what is the primary language of the book. For instance, eng is standard for English.	String
7.	isbn13	A 13-digit ISBN to identify the book, instead of the standard 11-digit ISBN.	Long
8.	# num_pages	Number of pages the book contains.	Integer
9.	ratings_count	Total number of ratings the book received.	Integer
10.	text_reviews_count	Total number of written text reviews the book received.	Integer

## 4. Dataset Preprocessing

We wanted to use this dataset to a build an ANN model to predict the overall rating of the books (attribute number 4).

The first thing we had to do, is choose a suitable factors for this prediction, and delete the unnecessary ones, we chose these factors to be our input to the predictive model: #num\_pages, rating\_count, text\_reviews\_count, language\_code.

Moreover, the dataset contain 13720 instances. After preprocessing it becomes 12241 which is a large a number to a neural network to deal with, so, we divided these samples to 8242 training instances, and 3999 validation instances.

In addition, because of the integer numbers of the inputs are too large comparing with the real rate values, we did a normalization to them so all the data are real.

Normalization formula was:

$$\text{Normalized value} = \frac{(\text{previous value} - \text{Min}(i_1 \dots i_n))}{(\text{Max}(i_1 \dots i_n) - \text{Min}(i_1 \dots i_n))}$$

While checking the instances, it has been noticed that there are a conflict between some instances; which means, there at least two books with the same input values but different rates, we excluded for the secondary ones. Moreover, there were validation instances that are out of range, we converted them to training.

Now, the dataset is ready for training and validation.

## 5. Our ANN Model

The resulted predictive ANN model is shown in Figure 2 and Figure 5.

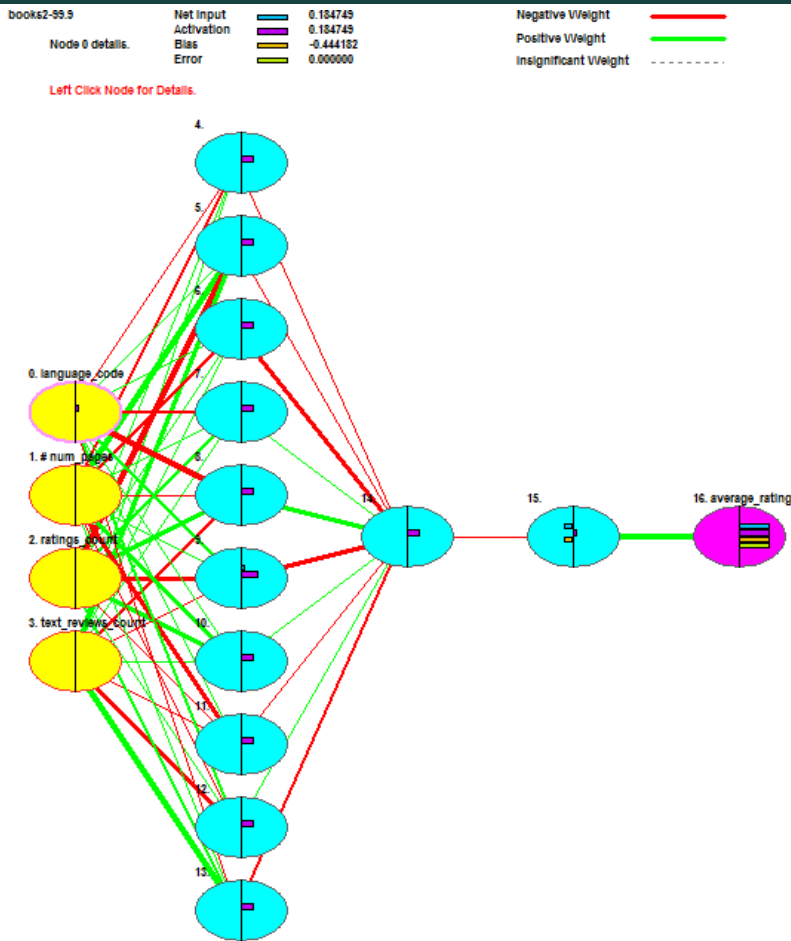


Figure 2: Our ANN Model

**Validation**

Our ANN model was able to predict the books' overall rate with 99.90% accuracy, with about 0.008 errors as seen in figure (3). Furthermore, The Model showed that the most

effective factor in a book's rate is the rating\_count. More details are shown in figure (4).

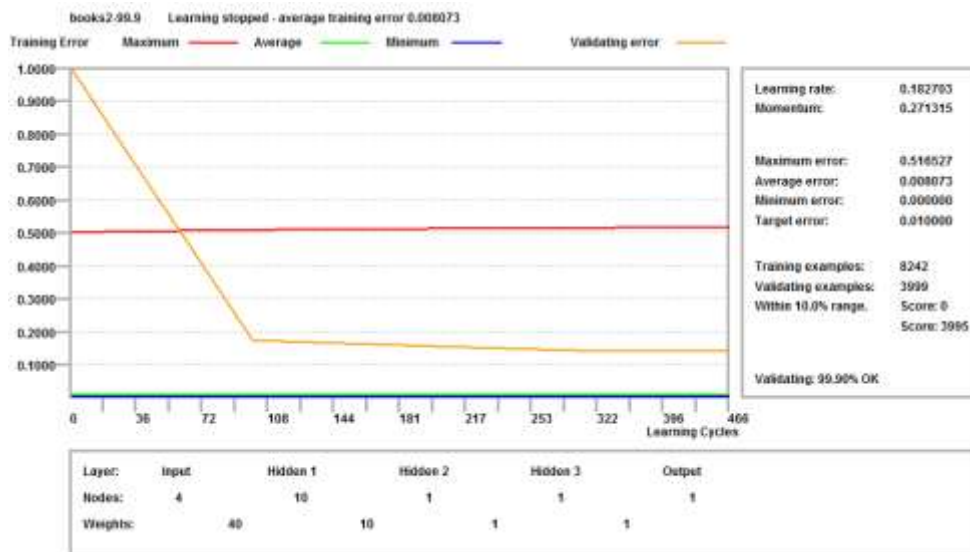


Figure 3: Validation and Errors

books2-99.9 469 cycles. Target error 0.0100 Average training error 0.008073  
 The first 4 of 4 Inputs in descending order.

Column	Input Name	Importance	Relative Importance
3	ratings_count	2.7584	
2	# num_pages	2.3979	
4	text_reviews_count	2.1959	
1	language_code	2.0857	

Figure 4: Attributes Importance

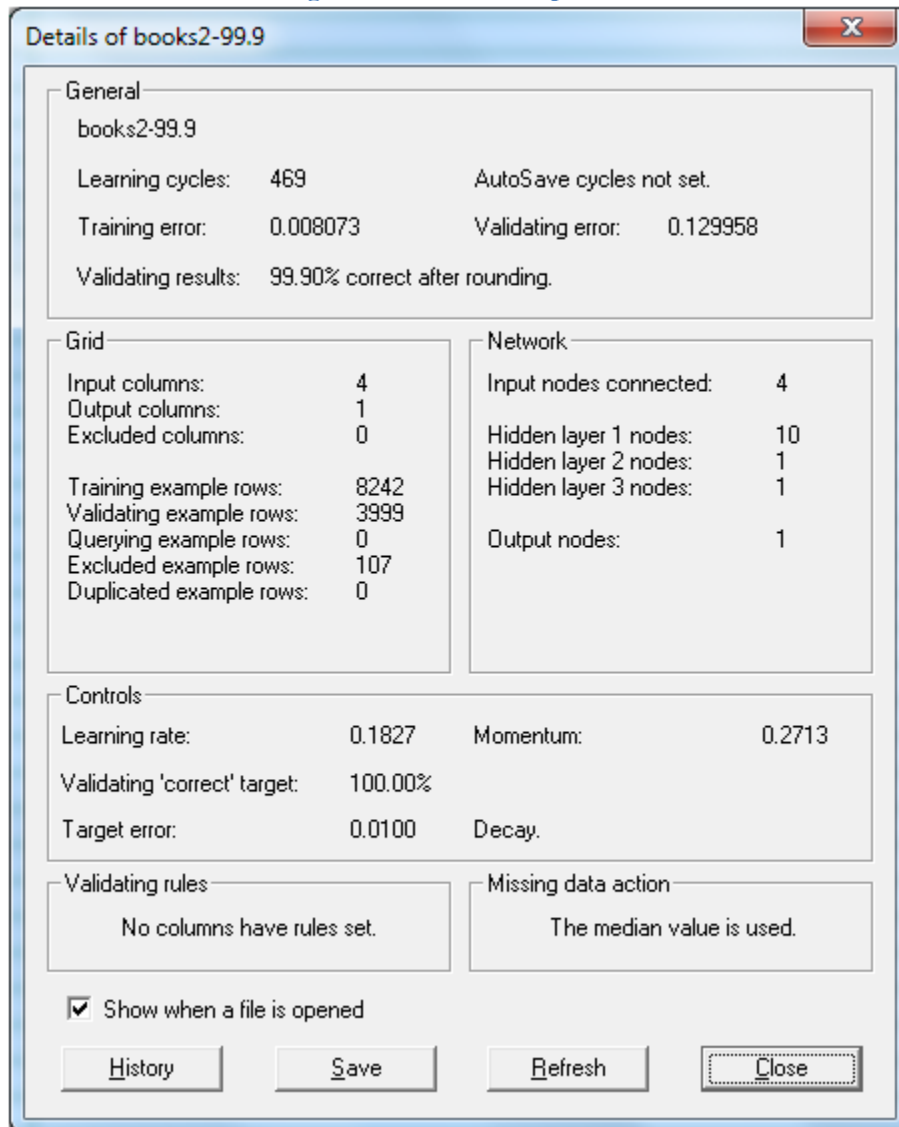


Figure 5: Details of our ANN Model

## 6. Conclusion

A predictive Artificial Neural Network Model for predicting books' rating was developed. The Model trained

and validated using a dataset from the goodreads application/ website. We did some preprocessing on the dataset to make

it suitable as input to our ANN model. Validation showed that the model is 99.90% accurate.

## References

1. Nasser, I. M., Al-Shawwa, M. O., & Abu-Naser, S. S. (2019). A Proposed Artificial Neural Network for Predicting Movies Rates Category. *International Journal of Academic Engineering Research (IJAER)*, 3(2), 21-25.
2. L. Dormehl, "Digital Trends," 5 1 2019. [Online]. Available: <https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/>. [Accessed 15 3 2019].
3. I. E. Livieris, K. Drakopoulou and P. Pintelas, "Predicting students' performance using artificial neural networks," in 8th PanHellenic Conference with International Participation Information and Communication Technologies in Education, Volos, Greece, 2012.
4. <https://www.kaggle.com/jealousleopard/goodreadsbooks>
5. Abu Naser, S., Baker, J., Cruz, I., Liotta, G., Tamassia, R., & Cooper, M. & Heller, R.(1996). Information Visualization. *Information Technology Journal*, 7(2).
6. Abu Naser, S., Sulisel, O., Alexandrescu, A., Anderson, J., Skwarecki, E., Anderson, J., . . . Brusilovsky, P. (2005). Growth and Maturity of Intelligent Tutoring Systems. *Information Technology Journal*, 7(7), 9-37.
7. Abu Naser, S., Zaqout, I., Ghosh, M. A., Atallah, R., & Alajrami, E. (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.
8. 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.
9. Abu Naser, S. S., Anderson, J., Corbett, A., Koedinger, K., Pelletier, R., Beal, C., . . . Soh, L. (2005). Adaptation of Problem Presentation and Feedback in an Intelligent Mathematics Tutor. *Information Technology Journal*, 5(5), 167-207.
10. Abu Naser, S. S., Baker, J., Cruz, I., Liotta, G., Tamassia, R., Cooper, M., . . . Feldman, M. (1996). Information Visualization. *Information Technology Journal*, 7(2), pp: 403-404.
11. Abu-Naser, S., Hissi, H. E.-, Rass, M. A.-, khozondar, N. E.-, Kashkash, K., Fayyad, M., . . . Fallat, R. (2008). Medical Informatics: Computer Applications in Health Care and Biomedicine. *Journal of Artificial Intelligence*, 3(4), 78-85.
12. Abu-Naser, S., Kashkash, K., Fayyad, M., Azaab, S., Naser, S., & Sulisel, O. (1995). & Beattie, GA (2000). Expert system methodologies and applications-a decade review from, 9-26.
13. Abu-Naser, S., Kashkash, K., Fayyad, M., Azaab, S., Riley, M., Williamson, M., . . . Maloy, O. (2005). Expert system methodologies and applications-a decade review from 1995 to 2004. *Journal of Artificial Intelligence*, 1(2), 9-26.
14. Afana, M., Ahmed, J., Harb, B., Abu-Nasser, B. S., & Abu-Naser, S. S. (2018). Artificial Neural Network for Forecasting Car Mileage per Gallon in the City. *International Journal of Advanced Science and Technology*, 124, 51-59.
15. Ahmed, A., Masri, N., Sultan, Y. A., Akkila, A. N., Almasri, A., Mahmoud, A. Y., . . . Abu-Naser, S. S. (2019). Knowledge-Based Systems Survey. *International Journal of Academic Engineering Research (IJAER)*, 3(7), 1-22.
16. Akkila, A. N., Almasri, A., Ahmed, A., Al-Masri, N., Sultan, Y. A., Mahmoud, A. Y., . . . Abu-Naser, S. S. (2019). Survey of Intelligent Tutoring Systems up to the end of 2017. *International Journal of Academic Information Systems Research (IJAISR)*, 3(4), 36-49.
17. Al-Ani, I. A. R., Sidek, L. M., Desa, M. M., Basri, N. A., Burns, J., Bhutani, J., . . . Fashokun, A. (2007). Water pollution and its effects on human health in rural areas of Faisalabad. *Journal of Environmental Science and Technology*, 5(5), 1-17.
18. Alghoul, A., Al Ajrami, S., Al Jarousha, G., Harb, G., & Abu-Naser, S. S. (2018). Email Classification Using Artificial Neural Network. *International Journal of Academic Engineering Research (IJAER)*, 2(11), 8-14.
19. Alkronz, E. S., Moghayer, K. A., Meimeh, M., Gazzaz, M., Abu-Nasser, B. S., & Abu-Naser, S. S. (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.
20. Almasri, A., Ahmed, A., Al-Masri, N., Sultan, Y. A., Mahmoud, A. Y., Zaqout, I., . . . Abu-Naser, S. S. (2019). Intelligent Tutoring Systems Survey for

- the Period 2000-2018. *International Journal of Academic Engineering Research (IJAER)*, 3(5), 21-37.
21. Al-Massri, R., Al-Astel, Y., Ziadia, H., Mousa, D. K., & Abu-Naser, S. S. (2018). Classification Prediction of SBRCTs Cancers Using Artificial Neural Network. *International Journal of Academic Engineering Research (IJAER)*, 2(11), 1-7.
  22. Al-Mubayyed, O. M., Abu-Nasser, B. S., & Abu-Naser, S. S. (2019). Predicting Overall Car Performance Using Artificial Neural Network. *International Journal of Academic and Applied Research (IJAAR)*, 3(1), 1-5.
  23. Al-Shawwa, M., & Abu-Naser, S. S. (2019). Predicting Birth Weight Using Artificial Neural Network. *International Journal of Academic Health and Medical Research (IJAHMR)*, 3(1), 9-14.
  24. Al-Shawwa, M., & Abu-Naser, S. S. (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.
  25. Al-Shawwa, M., Al-Absi, A. A.-R., Hassanein, S. A., Baraka, K. A., & Abu-Naser, S. S. (2018). Predicting Temperature and Humidity in the Surrounding Environment Using Artificial Neural Network. *International Journal of Academic Pedagogical Research (IJAPR)*, 2(9), 1-6.
  26. Ashqar, B. A., Abu-Nasser, B. S., & Abu-Naser, S. S. (2019). Plant Seedlings Classification Using Deep Learning. *International Journal of Academic Information Systems Research (IJASIR)*, 3(1), 7-14.
  27. Ashqar, B. A. M., & Abu-Naser, S. S. (2019). Identifying Images of Invasive Hydrangea Using Pre-Trained Deep Convolutional Neural Networks. *International Journal of Academic Engineering Research (IJAER)*, 3(3), 28-36.
  28. Ashqar, B. A. M., & Abu-Naser, S. S. (2019). Image-Based Tomato Leaves Diseases Detection Using Deep Learning. *International Journal of Academic Engineering Research (IJAER)*, 2(12), 10-16.
  29. Atallah, R. R., & Abu Naser, S. S. (2014). Data mining techniques in higher education an empirical study for the university of Palestine. *International Journal Of Modern Engineering Research (IJMER)*, 4(4), 48-52.
  30. Chen, R.-S., Tsai, C.-H., Abu-Naser, S., Bishop, A., Bishop, C., Arbaugh, J., . . . Trent, B. (2008). Evaluating structural equation models with unobservable variables and measurement error. *Information Technology Journal*, 10(2), 1055-1060.
  31. Dalffa, M. A., Abu-Nasser, B. S., & Abu-Naser, S. S. (2019). Tic-Tac-Toe Learning Using Artificial Neural Networks. *International Journal of Engineering and Information Systems (IJEAIS)*, 3(2), 9-19.
  32. El\_Jerjawi, N. S., & Abu-Naser, S. S. (2018). Diabetes Prediction Using Artificial Neural Network. *International Journal of Advanced Science and Technology*, 121, 55-64.
  33. El-Khatib, M. J., Abu-Nasser, B. S., & Abu-Naser, S. S. (2019). Glass Classification Using Artificial Neural Network. *International Journal of Academic Pedagogical Research (IJAPR)*, 3(2), 25-31.
  34. Elzamly, A., Abu Naser, S. S., Hussin, B., & Doheir, M. (2015). Predicting Software Analysis Process Risks Using Linear Stepwise Discriminant Analysis: Statistical Methods. *Int. J. Adv. Inf. Sci. Technol*, 38(38), 108-115.
  35. Elzamly, A., Hussin, B., Abu Naser, S., Khanfar, K., Doheir, M., Selamat, A., & Rashed, A. (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.
  36. Elzamly, A., Hussin, B., Abu Naser, S. S., & Doheir, M. (2015). Classification of Software Risks with Discriminant Analysis Techniques in Software planning Development Process. *International Journal of Advanced Science and Technology*, 81, 35-48.
  37. Elzamly, A., Hussin, B., Abu Naser, S. S., Shibutani, T., & Doheir, M. (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.
  38. Heriz, H. H., Salah, H. M., Abdu, S. B. A., El Sbihi, M. M., & Abu-Naser, S. S. (2018). English Alphabet Prediction Using Artificial Neural Networks. *International Journal of Academic Pedagogical Research (IJAPR)*, 2(11), 8-14.
  39. Jamala, M. N., & Abu-Naser, S. S. (2018). Predicting MPG for Automobile Using Artificial

- Neural Network Analysis. *International Journal of Academic Information Systems Research (IJAIRS)*, 2(10), 5-21.
40. Kashf, D. W. A., Okasha, A. N., Sahyoun, N. A., El-Rabi, R. E., & Abu-Naser, S. S. (2018). Predicting DNA Lung Cancer using Artificial Neural Network. *International Journal of Academic Pedagogical Research (IJAPR)*, 2(10), 6-13.
41. Li, L., Chen, N., He, C., Lang, F., Li, H., Wang, H., . . . Gong, P. (2011). Hybrid Quantum-inspired genetic algorithm for extracting association rule in data mining. *Information Technology Journal*, 12(4), 1437-1441.
42. Marouf, A., & Abu-Naser, S. S. (2018). Predicting Antibiotic Susceptibility Using Artificial Neural Network. *International Journal of Academic Pedagogical Research (IJAPR)*, 2(10), 1-5.
43. Masri, N., Sultan, Y. A., Akkila, A. N., Almasri, A., Ahmed, A., Mahmoud, A. Y., . . . Abu-Naser, S. S. (2019). Survey of Rule-Based Systems. *International Journal of Academic Information Systems Research (IJAIRS)*, 3(7), 1-23.
44. Metwally, N. F., AbuSharekh, E. K., & Abu-Naser, S. S. (2018). Diagnosis of Hepatitis Virus Using Artificial Neural Network. *International Journal of Academic Pedagogical Research (IJAPR)*, 2(11), 1-7.
45. Nasser, I. M., & Abu-Naser, S. S. (2019). Artificial Neural Network for Predicting Animals Category. *International Journal of Academic and Applied Research (IJAAR)*, 3(2), 18-24.
46. Nasser, I. M., & Abu-Naser, S. S. (2019). Lung Cancer Detection Using Artificial Neural Network. *International Journal of Engineering and Information Systems (IJEAIS)*, 3(3), 17-23.
47. Nasser, I. M., & Abu-Naser, S. S. (2019). Predicting Tumor Category Using Artificial Neural Networks. *International Journal of Academic Health and Medical Research (IJAHMR)*, 3(2), 1-7.
48. Nasser, I. M., Al-Shawwa, M., & Abu-Naser, S. S. (2019). Developing Artificial Neural Network for Predicting Mobile Phone Price Range. *International Journal of Academic Information Systems Research (IJAIRS)*, 3(2), 1-6.
49. Nasser, I. M., Al-Shawwa, M. O., & Abu-Naser, S. S. (2019). Artificial Neural Network for Diagnose Autism Spectrum Disorder. *International Journal of Academic Information Systems Research (IJAIRS)*, 3(2), 27-32.
50. Ng, S., Wong, C., Lee, T., Lee, F., Abu-Naser, S., El-Hissi, H., . . . James, A. (2010). Ad hoc networks based on rough set distance learning method. *Information Technology Journal*, 10(9), 239-251.
51. Owaied, H. H., Abu-Ara, M. M., Qasem, M. M., Fahmy, H., Douligeris, C., Aha, D., . . . Dillon, T. (2009). Using rules to support case-based reasoning for harmonizing melodies. *Journal of Applied Sciences*, 11(14), pp: 31-41.
52. Sadek, R. M., Mohammed, S. A., Abunbehan, A. R. K., Ghattas, A. K. H. A., Badawi, M. R., Mortaja, M. N., . . . Abu-Naser, S. S. (2019). Parkinson's Disease Prediction Using Artificial Neural Network. *International Journal of Academic Health and Medical Research (IJAHMR)*, 3(1), 1-8.
53. Salah, M., Altalla, K., Salah, A., & Abu-Naser, S. S. (2018). Predicting Medical Expenses Using Artificial Neural Network. *International Journal of Engineering and Information Systems (IJEAIS)*, 2(20), 11-17.
-