

A Proposed Artificial Neural Network for Predicting Movies Rates Category

Ibrahim M. Nasser, Mohammed O. Al-Shawwa, Samy S. Abu-Naser

Department of Information Technology, Faculty of Engineering and Information Technology
 Al-Azhar University- Gaza, Palestine
 Azhar.ibrahimn@gmail.com
 mohammed.o.alshawwa@gmail.com

Abstract: We proposed an Artificial Neural Network (ANN) in this paper for predicting the rate category of movies. A dataset used obtained from UCI repository created for research purposes. Our ANN prediction model was developed and validated; validation results showed that the ANN model is able to 92.19% accurately predict the category of movies' rate.

Keywords: Data Mining, Classification, Predictive Analysis, Artificial Neural Networks, movies classification

1. INTRODUCTION

Artificial neural networks (ANNs) are, similar to our neural networks and offer a relatively good technique, which solves the problem of classification and prediction. ANN is a collection of mathematical models, which can simulate characteristics of biological neural systems and have likenesses with adaptive human learning. ANNs made of connecting processing elements called neurons, connected by links, which contain weight coefficients that are, playing the role of synapses in our neural system. The neurons often come in three layers: input layer, one or more hidden layers and output layer, (ANN Architecture is shown in figure (1)). ANNs handle data as biological neural networks, in addition, ANN has the possibility of recalling, learning and eliminating errors, and high speed of getting the solution, [1] because of that, the neural networks can be used for solving complex problems, like classification and prediction [2]. ANNs were effectively applied in variety of applications for solving difficult and real problems [3].

ANN were found to be more efficient and more accurate than other classification techniques [4].

Classification by a neural network is done in two separate phases. First, the network is trained on a dataset. Then the weights of the connections between neurons are fixed so the network is validated to determine the classifications of a new dataset [5]. In this paper, we used about 70% of the total sample data for network training, and 30% for network validation.

While many models of ANNs have been proposed, the feed-forward neural networks (FNNs) are the most common and broadly used in many applications.

Mathematically, the problem of training an FNN is the minimization of an error function E ; In another word, to find a minimizer $w = (w_1, w_2, \dots, w_n)$ such that $w = \min E(w)$, where E is the batch error computed by the sum of square differences over all examples of the training dataset.

$$E(w) = \sum_{p=1}^p \sum_{j=1}^{N_L} ({}^L_{j,p}Y - t_{j,p})^2$$

${}^L_{j,p}Y$ is the output of the j -th neuron that belongs to the L -th (output) layer, N_L is the number of neurons of the output layer, $t_{j,p}$ is the anticipated response at the j -th neuron of the output layer at the input pattern p , and p represents the total number of patterns which used in the training dataset.

A traditional technique to solve this problem is by an iterative gradient-based training process, which produces a series of weights $\{w_k\}$ starting from an initial point $w_0 \in R^n$ Using the iterative formula

$$w_{k+1} = w_k + n_k d_k$$

where k is the current iteration, $n_k > 0$ is the learning rate and d_k is the decent search direction [5]. Our study main purpose is to develop a neural network as classification technique to predict the category of movies rate. A dataset from UCI repository [6,7] was used for this purpose.

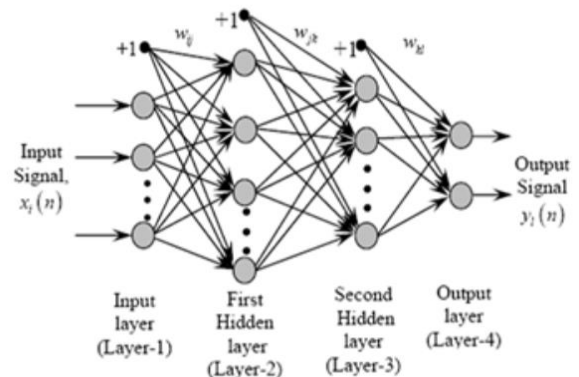


Figure 1: ANN Architecture

1. Literature Review

There are many studies involving Artificial Neural Network (ANN) for example : Artificial Neural Networks and expert systems were employed to obtain knowledge for the learner model in the Linear Programming Intelligent Tutoring System (LP-ITS) to be able to determine the academic performance level of the learners in order to offer him/her the proper difficulty level of linear programming problems to solve[8-12,15,18,21-23]; for predicting the performance of a sophomore student enrolled in engineering majors in the Faculty of Engineering and Information Technology in Al-Azhar University of Gaza was developed and tested [37,45]; ANN model was developed and tested to predict temperature in the surrounding environment [20]; for predicting critical cloud computing security issues by using Artificial Neural Network (ANNs) algorithms. However, they proposed the Levenberg–Marquardt based Back Propagation (LMBP) Algorithms to predict the performance for cloud security level [32]; for predicating the MPG rate for the forthcoming automobiles in the foremost relatively accurate evaluation for the approximated number which foresight the actual number to help through later design and manufacturing of later automobile [17,36]; to predict efficiency of antibiotics in treating various bacteria types [40]; to predict the rate of treatment expenditure on an individual or family in a country [46], for detecting early-stage non-small cell lung cancer (NSCLC) [38]; for the diagnosis of hepatitis virus [34,41]; for predicting the Letters from twenty dissimilar fonts for each letter [35], for Email Classification Using Artificial Neural Network [14]; Classification Prediction of SBRCTs Cancers Using Artificial Neural Network [16, 25]; for Diabetes Prediction Using Artificial Neural Network [29]; to predict Birth Weight [19]; to help cars dealers recognize the many characteristics of cars, including manufacturers, their location and classification of cars according to several categories including: Buying, Maint, Doors, Persons, Lug_boot, Safety, and Overall [13]; for Parkinson’s Disease Prediction Using Artificial Neural Network[39,42,44]; for desktop PC Troubleshooting[27]; for Tomato Leaves Diseases Detection Using Deep Learning[26]; Plant Seedlings Classification Using Deep Learning [24,43]; for predicating software analysis and risk management [30,31].

2. Methodology

We got a movie ranking data set that created by *Mehreen Ahmed*. We used this dataset which to build and validate our ANN model.

2.1 Dataset Description

Table 1: Original Dataset attributes description

#	Attribute	Type
1.	Movie	Text
2.	Year	Integer
3.	Genre	Integer
4.	Gross	Integer
5.	Budget	Integer
6.	Screens	Integer
7.	Sequel	Integer
8.	Sentiment	Integer
9.	Views	Integer
10.	Likes	Integer
11.	Dislikes	Integer
12.	Comments	Integer
13.	Aggregate Followers	Integer
14.	Ratings	Real

2.2 Dataset Preprocessing and Transformation

We did some preprocessing and transformation so the data is fit for predictive analysis. We used the first 13 attributes as inputs to our model except movie, and year attributes. In addition, the Ratings attribute was the used as the class to be predicted based on the input attributes. We normalized the values of the attributes: gross, budget, screens, views, likes, dislikes, comments, and aggregate followers, so they became real because they were large integer numbers. Normalization formula was

$$new\ value = \frac{(old\ value - Min(a_1 \dots a_n))}{(Max(a_1 \dots a_n) - Min(a_1 \dots a_n))}$$

We transform the class attribute (Ratings); we categorized rates based on the criteria showed below so it became nominal.

Table 2: Ratings Transformation

Interval	Value	Transformation
1 – 2.9	Bad	1
3 – 4.9	Good	2
5 – 7.9	Very good	3
8 - 10	Excellent	4

The resulted dataset description is shown in table (3).

Table 3: Description after preprocessing

#	Attribute	Type
1.	Genre	Integer
2.	Gross	Real
3.	Budget	Real
4.	Screens	Real
5.	Sequel	Integer
6.	Sentiment	Integer
7.	Views	Real
8.	Likes	Real
9.	Dislikes	Real
10.	Comments	Real
11.	Aggregate Followers	Real
12.	Ratings	Integer

2.3 The Neural Network

The resulted ANN Model is shown in figure (2).

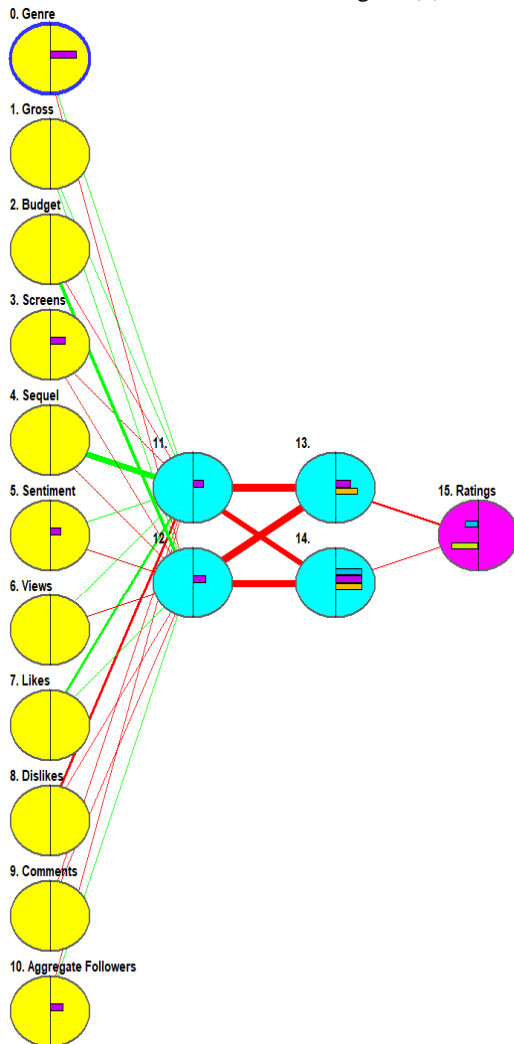


Figure 2: Our ANN Model

2.4 Results

Our ANN model was able to predict the rate class with 92.19% accuracy, after 11776860 learning cycles with about 1% training error rate as seen in figure (3). In addition, Our Model showed that the most attribute that has effect on the movie rate was “sequel”. More details are shown in figure (4).

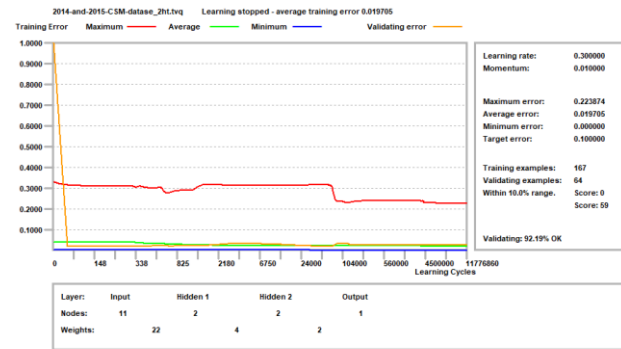


Figure 3: Validation and errors rates

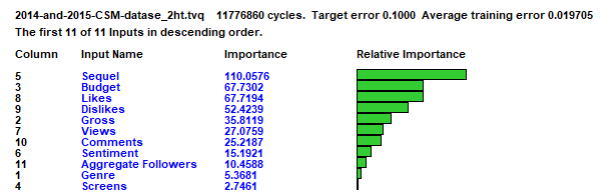


Figure 4: Attributes Importance

3. Conclusion

An artificial Neural Network for predicting the rate category of a movie was developed. The model was validated; it was 92.19 accurate in predict the rate category. This study showed that the neural network is able to predict movies rate category, so it can used for rating movies in the future.

References

1. M. S. Sonja Isljamovic, "Predicting Students' Academic Performance Using Artificial Neural Network : A Case Study From Faculty Of Organizational Sciences," The Eurasia Proceedings Of Educational & Social Sciences (Epress), Pp. 68-72, 1 2014.
2. I.C. Yeh, Application of artificial neural network model and implementation. Taiwan: Scholars Books, 1999. J. Thomas, M. Hass, Data Mining in Higher Education: University Student Declaration of Major, Information Systems, 2011.
3. S.H. Liao, C.H. Wen, "Artificial neural networks classification and clustering of methodologies and applications – literature analysis form 1995 to 2005",

- Expert Systems with Applications, 32(1), pp. 1–11, 2007.
4. Lerner B., Guterman H., Aladjem M., and Dinstein I. A comparative study of neural network based feature extraction paradigms. *Pattern Recognition Letters*, 20(1), p.p. 7–14, 1999.
 5. K. D. P. P. Ioannis E. Livieris "Predicting students' performance using artificial neural networks".
 6. Ahmed M, Jahangir M, Afzal H, Majeed A, Siddiqi I. Using Crowd-source based features from social media and Conventional features to predict the movies popularity. In *Smart City/SocialCom/SustainCom (SmartCity)*, 2015 IEEE International Conference on 2015 Dec 19 (pp. 273-278). IEEE.
 7. Dua, D. and Karra Taniskidou, E. (2017). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.
 8. AbuEl-Reesh, J. Y., & Abu Naser, S. S. (2017). A Knowledge Based System for Diagnosing Shortness of Breath in Infants and Children. *International Journal of Engineering and Information Systems (IJEAIS)*, 1(4), 102-115.
 9. Abu-Naser, S., Al-Masri, A., Sultan, Y. A., & Zaqout, I. (2011). A prototype decision support system for optimizing the effectiveness of elearning in educational institutions. *International Journal of Data Mining & Knowledge Management Process (IJDMP)*, 1, 1-13.
 10. Abu-Naser, S., Hissi, H. E.-, Rass, M. A.-, khozondar, N. E.-, Abu-Naser, S., Kashkash, K., . . . Shortliffe, E. (2008). Medical Informatics: Computer Applications in Health Care and Biomedicine. *Journal of Artificial Intelligence*, 3(4), 78-85.
 11. Abu-Nasser, B. S. (2017). Medical Expert Systems Survey. *International Journal of Engineering and Information Systems*, 1(7), 218-224.
 12. Abu-Nasser, B. S., & Abu-Naser, S. S. (2018). Cognitive System for Helping Farmers in Diagnosing Watermelon Diseases. *International Journal of Academic Information Systems Research (IJASIR)*, 2(7), 1-7.
 13. 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.
 14. 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.
 15. Almadhoun, H. R., & Abu Naser, S. S. (2018). Banana Knowledge Based System Diagnosis and Treatment. *International Journal of Academic Pedagogical Research (IJAPR)*, 2(7), 1-11.
 16. 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.
 17. 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.
 18. Almurshidi, S. H., & Abu-Naser, S. S. (2018). EXPERT SYSTEM FOR DIAGNOSING BREAST CANCER. Al-Azhar University, Gaza, Palestine.
 19. 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.
 20. 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.
 21. AlZamily, J. Y., & Abu-Naser, S. S. (2018). A Cognitive System for Diagnosing Musa Acuminata Disorders. *International Journal of Academic Information Systems Research (IJASIR)*, 2(8), 1-8.
 22. 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.
 23. Abu Naser, S. S., & Ola, A. Z. A. (2008). AN EXPERT SYSTEM FOR DIAGNOSING EYE DISEASES USING CLIPS. *Journal of Theoretical & Applied Information Technology*, 4(10).
 24. Ashqar, B. A., Bassem S, A.-N., & Abu-Naser, S. S. (2019). Plant Seedlings Classification Using Deep Learning. *International Journal of Academic Information Systems Research (IJASIR)*, 3(1), 7-14.
 25. Abu Naser, S. S., & Bastami, B. G. (2016). A proposed rule based system for breasts cancer diagnosis. *World Wide Journal of Multidisciplinary Research and Development*, 2(5), 27-33.
 26. Ashqar, B. AM, & 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.
 27. Dahouk, A. W., & Abu-Naser, S. S. (2018). A Proposed Knowledge Based System for Desktop PC Troubleshooting. *International Journal of Academic Pedagogical Research (IJAPR)*, 2(6), 1-8.
 28. 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.
 29. El_Jerjawi, N. S., & Abu-Naser, S. S. (2018). Diabetes Prediction Using Artificial Neural Network.
-

- International Journal of Advanced Science and Technology, 124, 1-10.
30. 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.
31. 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.
32. Elzamly, A., Hussin, B., Abu Naser, S. S., Shibusani, 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.
33. 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.
34. Abu Naser, S., Al-Dahdooh, R., Mushtaha, A., & El-Naffar, M. (2010). Knowledge management in ESMEDA: expert system for medical diagnostic assistance. *AIML Journal*, 10(1), 31-40.
35. 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.
36. 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.
37. Abu Naser, S. S., & Zaqout, I. S. (2016). Knowledge-based systems that determine the appropriate students major: In the faculty of engineering and information technology. *World Wide Journal of Multidisciplinary Research and Development*, 2(10), 26-34.
38. 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.
39. Abu Naser, S. S., & Al-Bayed, M. H. (2016). Detecting Health Problems Related to Addiction of Video Game Playing Using an Expert System. *World Wide Journal of Multidisciplinary Research and Development*, 2(9), 7-12.
40. 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.
41. 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.
42. Mrouf, A., Albatish, I., Mosa, M., & Abu Naser, S. S. (2017). Knowledge Based System for Long-term Abdominal Pain (Stomach Pain) Diagnosis and Treatment. *International Journal of Engineering and Information Systems (IJEAIS)*, 1(4), 71-88.
43. Nassr, M. S., & Abu Naser, S. S. (2018). Knowledge Based System for Diagnosing Pineapple Diseases. *International Journal of Academic Pedagogical Research (IJAPR)*, 2(7), 12-19.
44. 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.
45. 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.
46. 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.
47. 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).
48. 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).
49. 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).