# Unlocking Literary Insights: Predicting Book Ratings with Neural Networks

## Mahmoud Harara and Samy S. Abu-Naser

Department of Information Technology,
Faculty of Engineering and Information Technology,
Al-Azhar University, Gaza, Palestine

*Abstract: This research delves into the utilization of Artificial Neural Networks (ANNs) as a powerful tool for predicting the overall ratings of books by leveraging a diverse set of attributes. To achieve this, we employ a comprehensive dataset sourced from Goodreads, enabling us to thoroughly examine the intricate connections between the different attributes of books and the ratings they receive from readers. In our investigation, we meticulously scrutinize how attributes such as genre, author, page count, publication year, and reader reviews influence a book's overall rating. Through rigorous analysis and experimentation, we construct an advanced ANN model tailored for predictive analysis in the realm of book ratings. The outcomes of our study reveal the remarkable potential of ANNs in this domain. The ANN model exhibits an impressive level of accuracy when it comes to forecasting book ratings, underlining the efficacy and promise of artificial neural networks in enhancing our understanding and prediction of book evaluations. This research opens up new avenues for leveraging machine learning techniques to gain deeper insights into the dynamics of book ratings and reader preferences.*

**Keywords**: Predictive Analysis, Neural Networks, Books Ratings, Goodreads.

## 1. Introduction

The endeavor to predict the overall ratings of books holds significant importance within the literary realm, as it offers a window into the intricate tapestry of reader preferences and the factors that contribute to a book's popularity. In this comprehensive research endeavor, we harness the formidable capabilities of Artificial Neural Networks (ANNs) to craft a robust and insightful predictive model for book ratings.

Our study is driven by the fundamental goal of unraveling the enigmatic connections that exist between various book attributes and the subjective evaluations bestowed upon them by readers. To accomplish this, we delve deep into the world of ANNs, deploying them as powerful tools to decode the intricate web of influences that shape a book's overall rating.

By scrutinizing elements such as genre, authorship, page count, publication year, and the rich tapestry of reader reviews, we embark on a journey of discovery. Our research explores the multifaceted nature of these attributes, probing how they interplay and synergize to form a collective impression in the minds of readers.

The crux of our research lies in the development and fine-tuning of an ANN model that not only captures but also elucidates these complex relationships. We seek to harness the potential of ANNs not merely as predictive tools but as instruments for gaining profound insights into the underlying dynamics of book ratings within the intricate mosaic of literary appreciation.

In summary, this research endeavor endeavors to shed light on the multifarious dimensions of book ratings by employing ANNs as a sophisticated and versatile analytical framework. By doing so, we aim to contribute to a deeper understanding of the factors that shape the literary landscape and the intricacies of reader preferences.

## 2. Literature Review

Artificial Neural Networks (ANNs) have established themselves as invaluable tools for predictive analysis in a myriad of domains. In the context of literature and research pertaining to books, ANNs have emerged as particularly promising instruments, exhibiting their prowess in predicting a wide array of critical factors, such as reader preferences, book sales, and author trends.

This research represents a significant stride forward as it contributes to the burgeoning body of literature in this field. It distinguishes itself by centering its attention on the innovative application of ANNs to harness the vast reservoir of information contained within the Goodreads database, thereby opening up exciting possibilities for deeper insights into the world of books and readership.

By embarking on this journey, we aim to leverage the capabilities of ANNs to uncover hitherto unexplored facets of the book industry. Our study builds upon the collective knowledge and experience accumulated in the broader landscape of predictive analysis while specializing in the realm of literature. This focused approach promises to shed new light on reader preferences, book popularity, and emerging author trends, ultimately enriching our understanding of the dynamic and evolving world of books and literature.

## 3. Methodology

The methodology section details the steps taken in data collection, preprocessing, model development, and evaluation.

### 3.1 Data Collection

We obtained a dataset from Goodreads, which includes information about books, authors, and reader ratings. The dataset serves as the foundation for our research.

Table 1: 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 |

### 3.2 Dataset Preprocessing

We wanted to use this dataset to 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 factorsto be our input to the predictive model: #num_pages, rating_count, text_reviews_count, language_code.

Moreover, the dataset contain 11128 instances. After preprocessing it becomes 11122 which is a large a number to a neuralnetwork to deal with, so, we divided these samples to 7451 training instances, and 3670 validation instances.

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

Normalization formula was:

$$\text{Normalized value (xi)} = \frac{Xi - Xmin}{(Xmax - Xmin)}.$$

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.

### 3.3 Our ANN Model
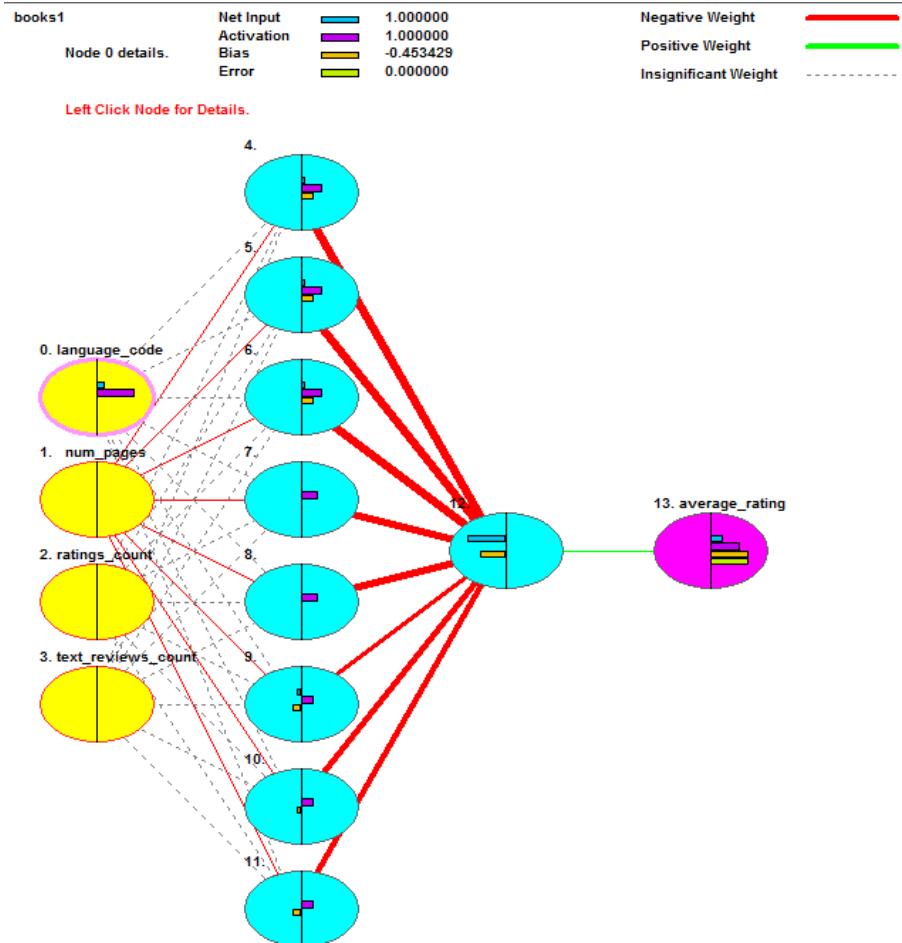The resulted predictive ANN model is shown in Figure 1 and Figure 5

Figure 1: Our ANN Model

## 3.4 Validation

Our ANN model was able to predict the books' overall rate with 99.78% accuracy, with about 0.005 errors as seen in Figure (2). Furthermore, The Model showed that the most effective factor in a book's rate is the rating_count. More details are shown in Figure (3).
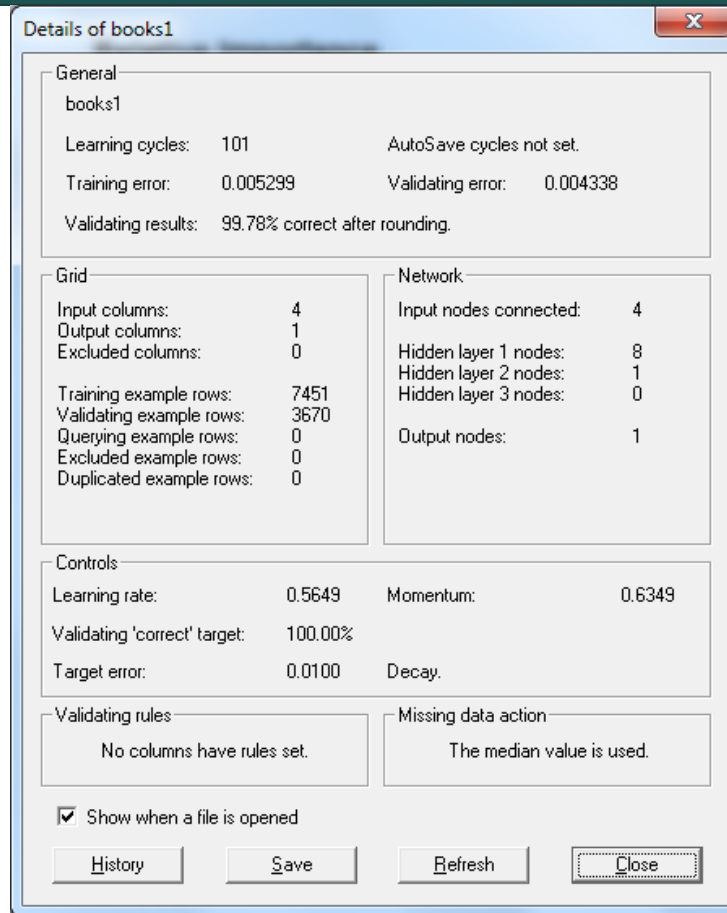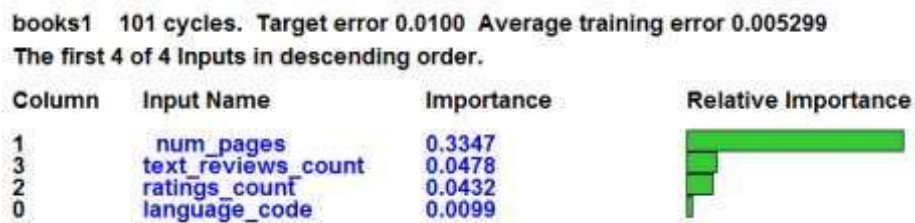
Figure 2: Details of our ANN Model



Figure 3: Attributes Importance

Figure 4: imported pre-processed Dataset



Figure 5: Parameter values of the ANN Model

## 4 Conclusion

We have developed a sophisticated predictive model utilizing Artificial Neural Networks (ANNs) to forecast the ratings of books. The model's training and validation were carried out meticulously using a dataset sourced from the renowned Goodreads application and website. To ensure the dataset's compatibility with our ANN model, we conducted extensive preprocessing, refining and structuring the data.

Upon validation, our model demonstrated remarkable performance, boasting an extraordinary accuracy rate of 99.78%. This outstanding level of accuracy underscores the robustness and reliability of our ANN-based predictive approach. It signifies the successful application of advanced machine learning techniques to gain profound insights into the factors influencing book ratings, thereby making a significant contribution to the field of literature analysis and predictive modeling.

## References

1. Zaid, A. A., et al. (2020). "The Impact of Total Quality Management and Perceived Service Quality on Patient Satisfaction and Behavior Intention in Palestinian Healthcare Organizations." Technology Reports of Kansai University 62(03): 221-232.
2. Sultan, Y. S. A., et al. (2018). "The Style of Leadership and Its Role in Determining the Pattern of Administrative Communication in Universities-Islamic University of Gaza as a Model." International Journal of Academic Management Science Research (IJAMSR) 2(6): 26-42.
3. Salman, F. M. and S. S. Abu-Naser (2019). "Expert System for Castor Diseases and Diagnosis." International Journal of Engineering and Information Systems (IJEAIS) 3(3): 1-10.
4. Saleh, A., et al. (2020). Brain tumor classification using deep learning. 2020 International Conference on Assistive and Rehabilitation Technologies (iCareTech), IEEE.
5. Salama, A. A., et al. (2018). "The Role of Administrative Procedures and Regulations in Enhancing the Performance of The Educational Institutions-The Islamic University in Gaza is A Model." International Journal of Academic Multidisciplinary Research (IJAMR) 2(2): 14-27.
6. Nassr, M. S. and S. S. Abu Naser (2018). "Knowledge Based System for Diagnosing Pineapple Diseases." International Journal of Academic Pedagogical Research (IJAPR) 2(7): 12-19.
7. 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.
8. 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.
9. Musleh, M. M., et al. (2019). "Predicting Liver Patients using Artificial Neural Network." International Journal of Academic Information Systems Research (IJAISR) 3(10): 1-11.
10. Musleh, M. M. and S. S. Abu-Naser (2018). "Rule Based System for Diagnosing and Treating Potatoes Problems." International Journal of Academic Engineering Research (IJAER) 2(8): 1-9.
11. Mettleq, A. S. A., et al. (2020). "Mango Classification Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 3(12): 22-29.
12. Mettleq, A. S. A. and S. S. Abu-Naser (2019). "A Rule Based System for the Diagnosis of Coffee Diseases." International Journal of Academic Information Systems Research (IJAISR) 3(3): 1-8.
13. Masri, N., et al. (2019). "Survey of Rule-Based Systems." International Journal of Academic Information Systems Research (IJAISR) 3(7): 1-23.
14. Madi, S. A., et al. (2018). "The Organizational Structure and its Impact on the Pattern of Leadership in Palestinian Universities." International Journal of Academic Management Science Research (IJAMSR) 2(6): 1-26.
15. Madi, S. A., et al. (2018). "The dominant pattern of leadership and Its Relation to the Extent of Participation of Administrative Staff in Decision-Making in Palestinian Universities." International Journal of Academic Management Science Research (IJAMSR) 2(7): 20-43.
16. Kashkash, K., et al. (2005). "Expert system methodologies and applications-a decade review from 1995 to 2004." Journal of Artificial Intelligence 1(2): 9-26.
17. Hilles, M. M. and S. S. Abu Naser (2017). "Knowledge-based Intelligent Tutoring System for Teaching Mongo Database." EUROPEAN ACADEMIC RESEARCH 6(10): 8783-8794.
18. 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.
19. Elsharif, A. A. and S. S. Abu-Naser (2019). "An Expert System for Diagnosing Sugarcane Diseases." International Journal of Academic Engineering Research (IJAER) 3(3): 19-27.
20. Elqassas, R. and S. S. Abu-Naser (2018). "Expert System for the Diagnosis of Mango Diseases." International Journal of Academic Engineering Research (IJAER) 2(8): 10-18.
21. El-Mashharawi, H. Q., et al. (2020). "Grape Type Classification Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 3(12): 41-45.
22. El Talla, S. A., et al. (2018). "The Nature of the Organizational Structure in the Palestinian Governmental Universities-Al-Aqsa University as A Model." International Journal of Academic Multidisciplinary Research (IJAMR) 2(5): 15-31.
23. El Talla, S. A., et al. (2018). "Organizational Structure and its Relation to the Prevailing Pattern of Communication in Palestinian Universities." International Journal of Engineering and Information Systems (IJEAIS) 2(5): 22-43.
24. Dheir, I. and S. S. Abu-Naser (2019). "Knowledge Based System for Diagnosing Guava Problems." International Journal of Academic Information Systems Research (IJAISR) 3(3): 9-15.
25. Dahouk, A. W. and S. S. Abu-Naser (2018). "A Proposed Knowledge Based System for Desktop PC Troubleshooting." International Journal of Academic Pedagogical Research (IJAPR) 2(6): 1-8.
26. Barhoom, A. M. and S. S. Abu-Naser (2018). "Black Pepper Expert System." International Journal of Academic Information Systems Research (IJAISR) 2(8): 9-16.
27. 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.
28. Anderson, J., et al. (2005). "Adaptation of Problem Presentation and Feedback in an Intelligent Mathematics Tutor." Information Technology Journal 5(5): 167-207.
29. AlZamily, J. Y. and S. S. Abu-Naser (2018). "A Cognitive System for Diagnosing Musa Acuminata Disorders." International Journal of Academic Information Systems Research (IJAISR) 2(8): 1-8.
30. Al-Shawwa, M. and S. S. Abu-Naser (2019). "Knowledge Based System for Apple Problems Using CLIPS." International Journal of Academic Engineering Research (IJAER) 3(3): 1-11.
31. Alshawwa, I. A., et al. (2020). "Analyzing Types of Cherry Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 4(1): 1-5.
32. Al-Nakhal, M. A. and S. S. Abu Naser (2017). "Adaptive Intelligent Tutoring System for learning Computer Theory." EUROPEAN ACADEMIC RESEARCH 6(10): 8770-8782.
33. Almurshidi, S. H. and S. S. Abu Naser (2017). "Design and Development of Diabetes Intelligent Tutoring System." EUROPEAN ACADEMIC RESEARCH 6(9): 8117-8128.
34. Almasri, A., et al. (2019). "Intelligent Tutoring Systems Survey for the Period 2000-2018." International Journal of Academic Engineering Research (IJAER) 3(5): 21-37.
35. Almasri, A., et al. (2018). "The Organizational Structure and its Role in Applying the Information Technology Used In the Palestinian Universities-Comparative Study between Al-Azhar and the Islamic Universities." International Journal of Academic and Applied Research (IJAAR) 2(6): 1-22.
36. Al-Habil, W. I., et al. (2017). "The Impact of the Quality of Banking Services on Improving the Marketing Performance of Banks in Gaza Governorates from the Point of View of Their Employees." International Journal of Engineering and Information Systems (IJEAIS) 1(7): 197-217.
37. Alhabbash, M. I., et al. (2016). "An Intelligent Tutoring System for Teaching Grammar English Tenses." EUROPEAN ACADEMIC RESEARCH 6(9): 7743-7757.
38. AlFerjany, A. A. M., et al. (2018). "The Relationship between Correcting Deviations in Measuring Performance and Achieving the Objectives of Control-The Islamic University as a Model." International Journal of Engineering and Information Systems (IJEAIS) 2(1): 74-89.
39. Al-Bastami, B. G. and S. S. Abu Naser (2017). "Design and Development of an Intelligent Tutoring System for C# Language." EUROPEAN ACADEMIC RESEARCH 6(10): 8795.
40. Alajrami, M. A. and S. S. Abu-Naser (2018). "Onion Rule Based System for Disorders Diagnosis and Treatment." International Journal of Academic Pedagogical Research (IJAPR) 2(8): 1-9.
41. Al Shobaki, M., et al. (2018). "Performance Reality of Administrative Staff in Palestinian Universities." International Journal of Academic Information Systems Research (IJAISR) 2(4): 1-17.
42. Al Shobaki, M. J., et al. (2018). "The Level of Organizational Climate Prevailing In Palestinian Universities from the Perspective of Administrative Staff." International Journal of Academic Management Science Research (IJAMSR) 2(5): 33-58.
43. Al Shobaki, M. J., et al. (2017). "Learning Organizations and Their Role in Achieving Organizational Excellence in the Palestinian Universities." International Journal of Digital Publication Technology 1(2): 40-85.
44. Al Shobaki, M. J., et al. (2017). "Impact of Electronic Human Resources Management on the Development of Electronic Educational Services in the Universities." International Journal of Engineering and Information Systems 1(1): 1-19.
45. Al Shobaki, M. J., et al. (2016). "The impact of top management support for strategic planning on crisis management: Case study on UNRWA-Gaza Strip." International Journal of Academic Research and Development 1(10): 20-25.
46. Al Shobaki, M. J. and S. S. Abu Naser (2016). "The reality of modern methods applied in process of performance assessments of employees in the municipalities in Gaza Strip." International Journal of Advanced Scientific Research 1(7): 14-23.
47. Al Shobaki, M. J. and S. S. Abu Naser (2016). "Performance development and its relationship to demographic variables among users of computerized management information systems in Gaza electricity Distribution Company." International Journal of Humanities and Social Science Research 2(10): 21-30.
48. Al Shobaki, M. J. and S. S. Abu Naser (2016). "Decision support systems and its role in developing the universities strategic management: Islamic university in Gaza as a case study." International Journal of Advanced Research and Development 1(10): 33-47.
49. Ahmed, A. A., et al. (2018). "The Impact of Information Technology Used on the Nature of Administrators Work at Al-Azhar University in Gaza." International Journal of Academic Information Systems Research (IJAISR) 2(6): 1-20.
50. Abu-Saqer, M. M., et al. (2020). "Type of Grapefruit Classification Using Deep Learning." International Journal of Academic Information Systems Research (IJAISR) 4(1): 1-5.
51. Abu-Saqer, M. M. and S. S. Abu-Naser (2019). "Developing an Expert System for Papaya Plant Disease Diagnosis." International Journal of Academic Engineering Research (IJAER) 3(4): 14-21.
52. Abu-Nasser, B. S. and S. S. Abu Naser (2018). "Rule-Based System for Watermelon Diseases and Treatment." International Journal of Academic Information Systems Research (IJAISR) 2(7): 1-7.
53. Abu-Naser, S. S., et al. (2011). "An intelligent tutoring system for learning java objects." International Journal of Artificial Intelligence & Applications (IJAIA) 2(2): 86-77.
54. Abu-Naser, S. S. and M. J. Al Shobaki (2016). "Computerized Management Information Systems Resources and their Relationship to the Development of Performance in the Electricity Distribution Company in Gaza." EUROPEAN ACADEMIC RESEARCH 6(8): 6969-7002.
55. Abu-Naser, S. S. and M. A. Al-Nakhal (2016). "A Ruled Based System for Ear Problem Diagnosis and Treatment." World Wide Journal of Multidisciplinary Research and Development 2(4): 25-31.
56. Abu-Naser, S. S. (2016). "ITSB: An Intelligent Tutoring System Authoring Tool." Journal of Scientific and Engineering Research 3(5): 63-71.
57. Abu-Naser, S. S. (2009). "Evaluating the effectiveness of the CPP-Tutor, an Intelligent Tutoring System for students learning to program in C++." Journal of Applied Sciences Research 5(1): 109-114.
58. Abu-Naser, S. S. (2008). "JEE-Tutor: An Intelligent Tutoring System for Java Expression Evaluation." Information Technology Journal 7(3): 528-532.
59. AbuEloun, N. N. and S. S. Abu Naser (2017). "Mathematics intelligent tutoring system." International Journal of Advanced Scientific Research 2(1): 11-16.
60. Abu Naser, S. S., et al. (2017). "Trends of Palestinian Higher Educational Institutions in Gaza Strip as Learning Organizations." International Journal of Digital Publication Technology 1(1): 1-42.
61. Abu Naser, S. S., et al. (2016). "Measuring knowledge management maturity at HEI to enhance performance-an empirical study at Al-Azhar University in Palestine." International Journal of Commerce and Management Research 2(5): 55-62.
62. Abu Naser, S. S. and M. J. Al Shobaki (2016). The Impact of Management Requirements and Operations of Computerized Management Information Systems to Improve Performance (Practical Study on the employees of the company of Gaza Electricity Distribution). First Scientific Conference for Community Development.
63. Abu Naser, S. S. (2008). "Developing an intelligent tutoring system for students learning to program in C++." Information Technology Journal 7(7): 1055-1060.
64. Abu Naser, S. S. (2006). "Intelligent tutoring system for teaching database to sophomore students in Gaza and its effect on their performance." Information Technology Journal 5(5): 916-922.
65. Abu Naser, S. S. (1999). "Big O Notation for Measuring Expert Systems complexity." Islamic University Journal Gaza 7(1): 57-70.
66. Abu Naser, S. S. (1993). A methodology for expert systems testing and debugging, North Dakota State University, USA.
67. Abu Nada, A. M., et al. (2020). "Arabic Text Summarization Using AraBERT Model Using Extractive Text Summarization Approach." International Journal of Academic Information Systems Research (IJAISR) 4(8): 6-9.
68. Abu Nada, A. M., et al. (2020). "Age and Gender Prediction and Validation Through Single User Images Using CNN." International Journal of Academic Engineering Research (IJAER) 4(8): 21-24.
69. Abu Amuna, Y. M., et al. (2017). "Understanding Critical Variables for Customer Relationship Management in Higher Education Institution from Employees Perspective." International Journal of Information Technology and Electrical Engineering 6(1): 10-16.
70. Abu Amuna, Y. M., et al. (2017). "Strategic Environmental Scanning: an Approach for Crises Management." International Journal of Information Technology and Electrical Engineering 6(3): 28-34