



ISSN Online: 2821-1936

Transactions on Data Analysis in Social Science

Journal Homepage: <https://transoscience.ir>

Forecasting the position of science, technology, and innovation in higher education institutions of the world in the global ranking system using artificial neural networks

S. Fatahi^{1,*}, F. Amiri²

¹ Assistant Professor, Information Technology Research Institute, Iran Science and Information Technology Research Institute (Irandoct), Tehran, Iran

² Assistant Professor, Department of Computer Engineering, Hamedan University of Technology, Hamedan, Iran

ARTICLE INFO	ABSTRACT
<p>Article History: Received 8 April 2022 Received in revised form 12 June 2022 Accepted 24 August 2022 Available online 1 September 2022</p>	<p>In recent years, evaluating the performance quality of universities and higher education institutions has become a global priority. To meet this need, numerous ranking systems have been developed, each employing different indicators to assess institutional performance. In Iran, the "Positioning System of Science, Technology, and Innovation of Iran in the World" (NAMA), implemented by the Iranian Research Institute for Science and Technology (IRANDOC), provides regular reports on the status of national universities and higher education institutions according to key global benchmarks. Beyond assessment, however, the ability to predict future performance based on past and current data is a vital component of strategic planning and decision-making. Institutions that can forecast trends with minimal error are better positioned to identify effective strategies and improve their competitiveness in global ranking systems. This study applies artificial neural networks (ANN) to predict the rankings of universities and higher education institutions within the Times Higher Education (THE) framework. The results demonstrate that the proposed ANN model is capable of predicting THE ranking indicators, overall scores, and institutional standings with satisfactory accuracy. These findings highlight the potential of ANN as a decision-support tool for improving the global visibility and strategic planning of universities.</p>
<p>Keywords: Times Higher Education Ranking System, Time Series Data, Prediction, NAMA (Positioning System of Science, Technology, And Innovation of Iran in The World)</p>	

1. INTRODUCTION

Today, one of the most effective mechanisms for achieving sustainable development in society is education, and the importance of education is felt more than ever [1, 2]. Therefore, choosing a university or higher education institution from among the countless universities worldwide has become a major concern for students and their families, as a wrong choice can have negative consequences for individuals' futures [3]. Thus, ranking universities and institutions based on performance quality is a fundamental issue from the perspective of education experts,

* Corresponding Author: Email: fatahi@irandoc.ac.ir

Corresponding author: Assistant Professor, Information Technology Research Institute, Iran Science and Information Technology Research Institute (Irandoct), Tehran, Iran



especially in higher education. On the other hand, from the perspective of university administrators and higher education institutions, paying attention to ranking systems especially at the global level is critical for competing to attract top students, commercializing research, and securing research funding. Independent and specialized organizations in the field of ranking and quality assurance in global higher education, such as Times Higher Education in the UK, QS, U.S. News in the U.S., and Shanghai in China, publish annual reports on university rankings based on specific indicators. One of the most well-known systems is the Times Higher Education (THE) ranking system. The aim of this study is to predict the rankings and scores of universities worldwide based on the Times Higher Education ranking using artificial neural networks.

2. THEORETICAL FRAMEWORK

2.1. Time Series

Time series data refers to a set of data points that are collected at regular time intervals. These data points are represented with time stamps. The process of extracting or discovering patterns from time series data is known as time series data mining. The most important step in the process of time series data mining is identifying useful patterns in the data and analyzing them. Time series analysis and forecasting are key aspects of time series data mining. Many definitions of time series can be found in academic papers and texts [14]. A series of observations arranged by time or another quantity is referred to as a time series [5, 6]. Based on these definitions, the data related to university rankings and higher education institutions fall into this category of data, as they are reported annually by various organizations. The data used in this study comes from the Times Higher Education publication, specifically the data from 2011 to 2019.

2.2. Times Higher Education Ranking System

Times Higher Education, a London-based publication dedicated to higher education news and issues, uses data tables from Thomson Reuters to evaluate universities across all of their key missions: teaching, research, knowledge transfer, and international outlook. It ranks universities based on 13 performance indicators, which are considered the most comprehensive and balanced indicators, trusted by students, universities, academic administrators, industry, and government. In the Times Higher Education ranking system, these 13 performance indicators are divided into five categories, as shown in Table 1.

Table 1. The indicators of the Times ranking system

Score	explanation	Index group
30%	learning environment	Education
30%	Volume, revenue and credit	Research (including 3 indicators)
30%	The impact of research	
2.5%	Innovation	Industrial income
7.5%	Staff, students and research	International perspective

3. PREDICTION

Prediction is both the science and art of forecasting future events. In prediction, the aim is to estimate future events by analyzing and examining past data. To do this, past data is systematically and regularly combined using a predefined method to estimate the future. Predicting future values of a time series is a critical issue in many fields, such as economics, production planning, forecasting, sales, and control. However, there is no one-size-fits-all prediction method that works in every situation. Instead, we need to find the method that is most suitable for the current set of conditions. Various methods for forecasting time series data are discussed in the literature, with one of the most reliable methods being the use of artificial neural networks (ANN), which is used in this study.

4. PREVIOUS STUDIES

Numerous studies have been conducted on ranking systems both inside and outside of Iran. In 2011, a comprehensive model for ranking and evaluating universities and higher education institutions was developed by Jafari et al. (2011) by extracting 28 factors. In the same year, Niazi and Abounouri (2011) attempted to rank several universities based on their knowledge management infrastructure [7, 8]. Ahmad and colleagues (2013) collected and presented the latest ranking indicators in the field of education. In a study by Khosrojerdi and Zareatkhar (2014), data from seven ranking systems (Shanghai, QS, Friant International, Webometrics, HEC, and Leiden University) were examined, and the results showed a significant correlation between these systems and the rankings of the top 50 universities globally, with the strongest correlation found between the Shanghai and Times rankings. In another study conducted in the same year by Sadeghnia (2014), based on the ISC ranking system and a questionnaire on the scientific ranking of universities and research institutions in Qom, the findings revealed that research was the major weakness of institutions, while student facilities were their strength. In the same year, a study by Akbari Khosroshahi (2014) proposed a method for evaluating and ranking e-learning institutions and virtual universities [9-12]. Akbari Khosroshahi used two questionnaires based on effective criteria and indicators for evaluating e-learning institutions. The results from the questionnaires showed that indicators such as interactive content, bandwidth, and ease of use were important evaluation factors. The proposed model was applied to three universities: Ferdowsi University of Mashhad, Sahand University of Technology, and Shahid Madani University of Azerbaijan, and these universities were ranked accordingly.

In 2017, Vahdat Zadeh and colleagues (2017) examined 285,526 articles from 76 universities and higher education institutions within the country, including citation impact and the number of citations, and ranked Iranian universities. The results showed that Tarbiat Modares University and Sharif University of Technology were among the top universities. Finally, the results of these rankings were compared with those from the ISC and URAP systems, and the findings indicated that the ranking of Iranian universities, based on article-related indicators, showed a correlation with the ISC and URAP rankings [13].

In 2013, Jones et al. compared the Shanghai and QS ranking systems and criticized them. Also, Lin et al. (2013) investigated the impact of calculation methods on university rankings based on the number of publications and citations. In a 2017 study by Tabasiom and colleagues, a prediction system for university rankings was proposed. They used a dataset of past Times Higher Education (THE) ranking indicators. The data was divided into training and testing sets, and their proposed model was tested. Finally, the prediction system was evaluated using ROC curves, recall, and the number of correct predictions. It is worth mentioning that their proposed method did not use artificial neural networks [14-16].

A review of past research shows that while the issue of predicting university rankings is of great importance, there has not been sufficient research conducted on this topic. The aim of this study is to predict the rankings and scores of universities based on the Times Higher Education ranking system using artificial neural networks.

5. PROPOSED MODEL

Neural networks are modern computational systems and methods that are used for machine learning, knowledge representation, and applying knowledge for predicting the outputs of complex systems. The fundamental idea behind these networks is inspired by the functioning of the human brain's nervous system to process data for learning. The most significant attraction of neural networks is their flexibility in estimating a wide range of relationships and functions between data. Neural networks that are sufficiently complex can provide an accurate estimate of any arbitrary function.

To predict university rankings and scores based on the Times Higher Education ranking system, a neural network was designed for each of the indicators in this system. The network inputs include the most recent score obtained for the specific indicator and the difference between that score and the score from the previous year. The network's output is the indicator score for the year in question.

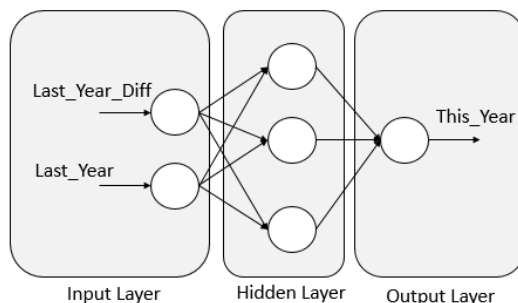


Fig. 1. Artificial neural network designed to predict university scores

6. DATA COLLECTION

Another tool used in this research was the Times Higher Education ranking data. In 2004, Times Higher Education, in collaboration with the QS Quacquarelli Symonds Institute, published a list of the world's top universities. However, since 2009, this collaboration ended, and the publication has been working with Thomson Reuters. Data from the years 2011 to 2018 were used as training data for the neural network, while data from 2019 was used as test data.

7. FINDINGS

As mentioned, the data from the years 2011 to 2018 for each indicator were used as training data, and the data from 2019 were used as test data. The accuracy of the neural networks designed for each indicator for the test data is reported in Table 2.

Table 2. Prediction accuracy results for Times indicators

International perspective	Industrial income	Citations	Research	Education	index name
%96.67	%69.8	%95.16	%94.24	%93.61	Prediction accuracy

The accuracy reported in Table 2 for each of the performance indicators in the Times Higher Education ranking system is presented separately. The overall score for the ranking of each university is obtained using the coefficients listed in Table 1. The predicted results, along with their corresponding coefficients, were used to calculate the overall score for each university. Subsequently, the ranking was determined based on the overall score. The results showed that the predicted scores had an accuracy of 98.40%. The ranking results indicated that the rank of 13 universities out of the top 97 universities (3 universities were excluded due to missing data) was predicted with high accuracy. Universities ranked 24 and below were predicted with a one rank difference either before or after. Ten universities had a prediction difference of two ranks, and eleven universities had a prediction difference of three ranks. In general, it can be concluded that approximately 60% of the universities were predicted with a deviation of only three ranks. It is worth mentioning that for the industrial income indicator, a significant error was observed in the overall predictions, which is attributed to the lack of data for this indicator.

8. CONCLUSION AND FUTURE WORK

The aim of this research was to predict the ranking of universities and higher education institutions based on the global Times Higher Education ranking system. The method used in this research was the application of artificial neural networks. One of the major limitations of this study was the small volume of training data. In this research, only the data of universities that had a ranking in the Times Higher Education system for at least 7 years were used for the neural network. The results obtained were generally acceptable, and with an accuracy of 98.4%, the overall scores of universities and higher education institutions worldwide can be predicted based on the Times Higher Education ranking system.

For future work, we plan to apply other machine learning methods, such as support vector machines, to the Times Higher Education dataset.

Transparency Statement

The data supporting this study are available upon reasonable request to the corresponding author, subject to ethical and confidentiality considerations.

Acknowledgments

We would like to express our gratitude to all individuals who contributed to this project.

Declaration of Interest

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

REFERENCES

- [1] Fatahi, S. (2019). An experimental study on an adaptive e-learning environment based on learner's personality and emotion. *Education and Information Technologies*, 24, 1–17. <https://doi.org/10.1007/s10639-019-09868-5>
- [2] Fatahi, S., & Moradian, S. (2018). An empirical study on the impact of using an adaptive e-learning environment based on learner's personality and emotion. *International Association for Development of the Information Society Conference Proceedings*.
- [3] Mortazavi, A. (2013). Designing a decision support system for custom university ranking (Master's thesis). Tarbiat Modares University, Tehran.
- [4] Mahalakshmi, G., Sridevi, S., & Rajaram, S. (2016). A survey on forecasting of time series data. In 2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16) (pp. 1–7). IEEE. <https://doi.org/10.1109/ICCTIDE.2016.7725358>
- [5] Sultani Gardfarmazi, S., Saberi, A., & Ghesori, M. (2017). Determining the best time series model for forecasting annual precipitation at selected stations in West Azerbaijan Province. *Journal of Applied Geographic Sciences*, 17(44), 87–105.
- [6] Adhikari, R., & Agrawal, R. K. (2013). An introductory study on time series modeling and forecasting. arXiv preprint, arXiv:1302.6613.
- [7] Jafari, M., Noori, S., & Talebi, D. (2011). Factor analysis of key indicators in the comprehensive evaluation system of higher education in the country. *Human Resource Management Research*, 1(1), 43–64.
- [8] Niazi, A., & Abounoori, E. (2011). Ranking selected universities based on knowledge management infrastructure. *Daneshvar-e-Ravabet*, 50(18), 263–286.
- [9] Ahmad, S., Aini Elahi, B., & Akbari Lakeh, M. (2013). Indicators of ranking the universities of Iran and the world. *Teb and Tazkiyeh*, 1(22), 9–16.
- [10] Akbari Khosroshahi, M. (2014). A method for evaluating and ranking e-learning institutions and virtual

universities using a fuzzy multi-criteria decision-making approach (Master's thesis). Payame Noor University of Tehran.

- [11] Khosroojerd, M., & Zarratkar, N. (2012). A review of the results of seven university ranking systems worldwide. *Information Processing and Management Research Journal*, 28(1), 1–20.
- [12] Sadeghi Nia, M. (2014). Scientific ranking of universities and higher education institutions in Qom (Master's thesis). Tarbiat Modares University, Tehran.
- [13] Vahdat Zadeh, M., Zarei Benadkouki, M., Oliaei, M., & Lotfi, M. (2017). Analysis of university rankings in Iran using scientometric indicators. *Information Processing and Management Research Journal*, 33(1), 1–15.
- [14] Jöns, H., & Hoyler, M. (2013). Global geographies of higher education: The perspective of world university rankings. *Geoforum*, 46, 45–59. <https://doi.org/10.1016/j.geoforum.2012.12.014>
- [15] Lin, C. S., Huang, M. H., & Chen, D. Z. (2013). The influences of counting methods on university rankings based on paper count and citation count. *Journal of Informetrics*, 7(3), 611–621. <https://doi.org/10.1016/j.joi.2013.03.007>
- [16] Tabassum, A., Hasan, M., Ahmed, S., Tasmin, R., Abdullah, D. M., & Musharrat, T. (2017). University ranking prediction system by analyzing influential global performance indicators. In 2017 9th International Conference on Knowledge and Smart Technology (KST) (pp. 1–6). IEEE. <https://doi.org/10.1109/KST.2017.7886119>