Artificial Neural Networks Predict Sustainable Development Goals Index

Seyed-Hadi Mirghaderi

Abstract The Sustainable Development Goals Index is an important index for measuring the movements toward sustainable goals. However, many indicators are needed for computing the index. This chapter aims to operationally show that for tackling the problem of the high number of indicators, artificial intelligence techniques may provide contributions. This chapter uses a combination of two famous techniques, including artificial neural networks and genetic algorithms. So, 288 indicators of 127 countries from 7 global reports were extracted, and the collinear and ineffective ones were removed. Finally, 90 indicators remained. A combination of genetic algorithms and artificial neural networks tried to find the best subset of remained indicators that provide a simple system for predicting Sustainable Development Goals Index. The results revealed that artificial neural networks with just four nodes and indicators include "Deaths from infectious diseases," "ICT use," "Expenditure on education," and "Assessment in reading, mathematics, and science" can predict sustainable development index with an accuracy rate of 97%. This chapter also validates the role of innovation in meeting Sustainable Development Goals (SDGs) and uncovers the insignificant role of environmental indicators in the Sustainable Development Goals Index.

Keywords Sustainable Development Goals Index · Artificial neural network · Genetic algorithm · Feature selection · Global reports

---

S.-H. Mirghaderi (✉)
Department of Management, School of Economics, Management, and Social Sciences, Shiraz University, Shiraz, Iran
e-mail: mirghaderi@shirazu.ac.ir

F. Mazzi, L. Floridi (eds.), *The Ethics of Artificial Intelligence for the Sustainable Development Goals*, Philosophical Studies Series 152,
https://doi.org/10.1007/978-3-031-21147-8_23

1 Introduction

Sustainable development (SD) refers to intergenerational equity and aims to optimize the consumption subject to support the needs of future generations (Keeble 1988). SD has three pillars, including environmental, social, and economic, which are interconnected (Brusseau 2019). SD has gradually received tremendous attention from academicians, politicians, business people, and economists (Omri 2020) due to the reveals of urgency in some global environmental issues (Elliott 2012), which lead to the international consensus on 17 Sustainable Development Goals (SDGs) for a better future. The agreement on SDGs was approved by all 193 members of Unite Nations (Sachs et al. 2017) and provides a basis for systematic and coordinated actions to shape a sustainable future in the global village (Costanza et al. 2016).

Global goal setting for tackling the world challenges in environmental, social, and economic aspects is the underlying reason for SDGs (Leal Filho 2020). Despite the excellent reason, the progress toward the SDGs is a problematical issue (Xu et al. 2020) that needs to be addressed. Although the UN Statistical Commission has proposed Sustainable Development Goals Index (SDGI), including 230 indicators for assessing the development toward the SDGs (Schmidt-Traub et al. 2017), there are many SDGI measuring problems, such as lack of systematic methods (Xu et al. 2020), lack of valid data (Schmidt-Traub et al. 2017), complicated interrelationship among SDGs (Costanza et al. 2016), and ignoring the uncertainty in SDGs (Ruiz-Morales et al. 2021). Therefore, proposing a simple alternative method for predicting SDGI is valuable for practitioners and academicians. For simplifying the SDGI prediction, we need a small number of suitable indicators selected from a pool of indicators (Hák et al. 2016) presented in global reports.

Global reports consist of indicators and indices which aim to pave the way for sustainable development (Shaker 2018). Although there are some indexes for sustainability, it is hard to draw an clear big picture of sustainability through them (Iddrisu and Bhattacharyya 2015). Furthermore, there is no single index that is widely adopted by scientists and politicians (Strezov et al. 2017). However, there is a wide range of indicators with collected data in global reports which attract researchers for reusing them to create sustainable measurement systems; examples of such approach were used by Iddrisu and Bhattacharyya (2015); Strezov et al. (2017); and Shaker (2018). Creating a SD measurement system using this approach needs to address a specific problem, that is, selecting a list of suitable indicators. The indicators must contribute to producing an efficient and noncomplicated SD measurement or prediction system.

The selection of indicators (or variables) is a well-known optimization problem in the artificial intelligence (AI) field (Alweshah et al. 2020), which encompass a wide range of proposed methods (George 2000) from statistical techniques (Borah et al. 2014) to heuristic search algorithms (Gnana et al. 2016) to neural networks (Chakraborty 1999). Also, prediction is applicable using several AI techniques

(Collins and Moons 2019), such as artificial neural networks (ANNs) and genetic algorithm (GA).

ANNs are one of the well-known techniques of AI that are inspired by the human brain (Okwu and Tartibu 2021), and GA is a metaheuristic algorithm inspired by the biological evolution of creatures (Mirjalili 2019). It seems that ANN and GA are useful for finding suitable indicators to create a system for predicting the SDGI values. In other words, the problem of too many indicators and hard-to-calculate SDGI may be tackled by using a combination of ANNs and GA.

The organization of the remaining parts is as follows. Sections 2, 3, and 4 provide a brief review of SDGI, ANN, and GA, respectively. Section 5 presents the research method and Sect. 6 provides the results of the research. Finally, the conclusion is presented in Sect. 7.

2 Sustainable Development Goals Index (SDGI)

In September 2000, 147 developing countries agreed on Millennium Development Goals (MDGs) to prove their commitment against global challenges such as hunger, poverty, disease, shelter-less people, and exclusion while enhancing environmental sustainability, gender equality, and education (Sachs and McArthur 2005). Based on the agreement, they set eight goals for the period between 2000 and 2015. The goals are (1) eradicate extreme poverty and hunger; (2) achieve universal primary education; (3) promote gender equality and empower women; (4) reduce child mortality; (5) improve maternal health; (6) combat HIV/AIDS, malaria, and other diseases; (7) ensure environmental sustainability; and (8) develop a global partnership for development (Kroll 2015).

At the expiration time of MDGs, in September 2015, all UN members for the period 2015–2030 agreed on 17 goals (Kroll 2015): (1) no poverty; (2) zero hunger; (3) good health and well-being; (4) quality education; (5) gender equality; (6) clean water and sanitation; (7) affordable and clean energy; (8) decent work and economic growth; (9) industry, innovation, and infrastructure; (10) reduced inequality; (11) sustainable cities and communities; (12) responsible consumption and production; (13) climate action; (14) life below water; (15) life on land; (16) Peace, justice and strong institutions; and (17) partnerships to achieve the goals (UN 2021).

SDGs are broader and more complex than MDGs. They are interrelated (Costanza et al. 2016), which cover the environmental, social, and economic aspects of SD (Allen et al. 2019). As Berglund and Gericke (2016) stated, SD as a complicated concept is not measurable unless it is broken down into specific global indicators. As Fig. 1 shows, SDGI has four layers. To measure the SDGI, 169 targets and 232 indicators were developed in 2019 (Barbier and Burgess 2019). But the number of indicators was decreased to 115 in 2020 (Sachs et al. 2020). Although the targets and indicators help monitor the status quo of countries (Alaimo et al. 2021), there are some critics regarding the high amount of indicators, the interrelationship between goals, missing values of indicators, etc.

Fig. 1 Pyramid of SDGI. (Source: Reyers et al. 2017)

In recent years, researchers have tried to resolve the critics and propose modifications in SDGI. For example, Xu et al. (2020) proposed a measurement system for quantifying the progress of china in SDGs. The system encompasses 119 indicators divided into 17 SDGs. Horan (2020) introduced a new version of SDGI based on interrelations between targets. It is argued that the new SDGI helps communicate with different stakeholders to undertake an integrated execution method for implementing SDG. Ruiz-Morales et al. (2021) proposed a new way for aggregating the value of each SDG using ordered weighted average (OWA) and prioritized OWA to encompass the uncertainty of SDGs. Bali Swain and Yang-Wallentin (2020) quantified and prioritized SDGs and their relations to SD to provide suggestions for countries to improve their SDGI by focusing on different aspects of SD.

3 Artificial Neural Networks (ANNs)

A significant part of artificial intelligence is ANNs (Wu and Feng 2018) which attract much attention from the 1980s (Wu and Feng 2018). The idea of ANNs was inspired by nervous system biology in the human body, which consists of a network of neurons named neural network. The network is an interconnected web of tremendous neurons which parallel process the collected data (Mishra and Srivastava 2014) to solve a specific problem (Abiodun et al. 2018), especially when the network is dense as in a human brain. In the brain, chemical reactions produce signals which play an essential role in controlling brain activities and creating a basis for learning (Russell and Norvig 2021). Based on a hypothesis, the learning process occurs at the connection points of two neurons when the connection intensity differs (Wu and Feng 2018).

Scientific attempts for modeling nervous system operation by mathematical formulation resulted in ANNs (Sivanandam and Deepa 2006). Although ANNs try to imitate the brain function, it has not been approached to capture the brain complexity. But there are two significant similarities between the brain and ANNs; both are constructed from highly interconnected simple computational elements (neurons), and the network function is determined by neurons connections (Hagan et al. 2016). In ANNs, each connection between neurons is denoted by a number named weight

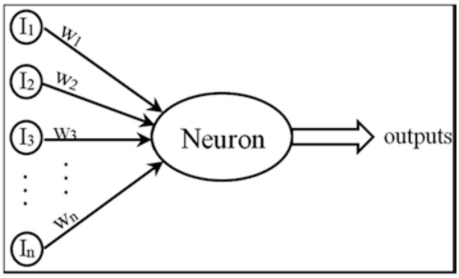

Fig. 2 Simple Neuron in ANNs. (Source: Aggarwal 2018)

(Wang 2003). The weight scales each input to a neuron and affects the function inside the neuron (Fig. 2) (Aggarwal 2018).

The weights are dynamically adjusted based on processing the specific inputs and the difference between actual and desired output (Floridi 2002). The weight updating process is the essence of learning (Ding et al. 2013) which can uncover the patterns in data and predict outputs often better than many statistical tools (Paliwal and Kumar 2009). Due to the capability of ANNs in solving the problems such as clustering, pattern recognition, and prediction in nonlinear and complex systems, the application of ANNs has expanded in many disciplines such as engineering, medicine, agriculture, mining, business, finance, arts, technology, etc. (Abiodun et al. 2018). In general, ANNs succeeded in providing high accuracy results for the problems in many disciplines (Gue et al. 2020).

Similar to other disciplines, sustainability has also taken advantage of ANNs. For example, Antanasijević et al. (2013) developed a model for predicting PM10 emissions at the national level. Gue et al. (2020) performed a critical review on utilizing ANNs in contributing SD. The study revealed that SDGs 6, 7, 11, and 12 have used more of ANNs. Also, the utilization includes modeling and predicting. Emmanuel et al. (2020) proposed a design of the neural network-based system for predicting the first six SDGs in less developed countries using patterns in big data.

4 Genetic Algorithm (GA)

GA was introduced by John Holland in the 1960s as an optimization algorithm. It was inspired by evolution in nature (Moriarity 2021). Evolution, as Charles Darwin (1859) discovered, is based on "survival of the fittest"; that means adapted creatures to the environment survive more rather than others. The fittest creature will have a higher chance to live and reproduce the next generation (Badar 2021), while the unfitted ones have less chance. The survival of the best is the principle of the evolution process (Sivanandam and Deepa 2008). As Kramer (2017) stated, evolution is a fruitful optimization process that can be seen in creatures. They utilize evolution-based strategies to produce near-optimal solutions for solving complicated problems (Moriarity 2021).

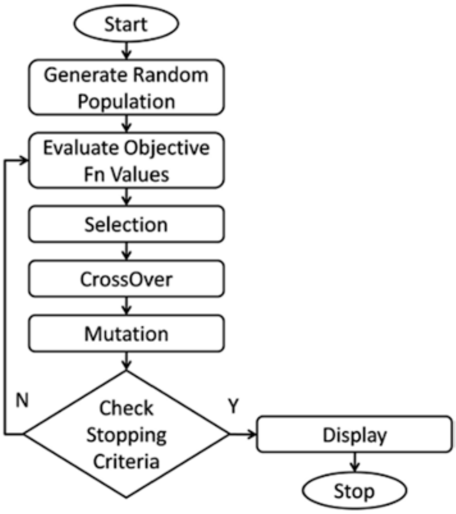

Fig. 3 GA procedure. (Source: Badar 2021)

GA uses a simulated evolution process to find near-optimal solutions (Badar 2021) in an iterative process through three biological-inspired operators named selection, crossover, and mutation (Katoch et al. 2021). Selection refers to choosing a certain number of current solutions for producing the next generation. Crossover means creating new solutions by combining existing solutions. The mutation is used to generate a different solution by manipulating the current solution. The selection operator has several methods, i.e., elite replacement (copy the best solution to the next generation as it is) and roulette wheel selection (selecting based on the probabilities related to the fitness function, i.e., the better solution has more chance to select) (Badar 2021). A technique for implementing crossover is the random respectful crossover which preserves the similarity of current solutions and randomly selects different points to create new solutions (Umbarkar and Sheth 2015). Mutation techniques try to explore the search space and increase the diversity of solutions (Moriarity 2021). It is implemented using methods such as randomly selecting a solution and changing a random point of it. The procedure of GA is presented in Fig. 3.

GA is a metaheuristic search algorithm that is flexible and attractive with many applications (Kramer 2017). Due to this capability, GA is the most implemented and researched metaheuristic with vast related published variants (Badar 2021). Nowadays, GA is a part of many applications in the artificial intelligence field (Moriarity 2021) to create methods that mimic and even do better than human intelligence (Kramer 2017).

5 Method

This chapter aims to create a simple model for predicting SDGI based on ANNs. To this end, a reverse pyramid method was used by following six steps include:

- Step 1: data gathering from the seven related global reports
- Step 2: data cleaning
- Step 3: handling missing values
- Step 4: handling collinear indicators
- Step 5: removing ineffective indicators
- Step 6: finding the best combination of indicators

By following the introduced steps, the research activities were conducted. The details of each step are presented in the following subsections.

- *Step 1: data gathering*

Some official and open-source reports are needed to create a pool of indicators. The best sources of indicators and their values are global reports. Table 1 shows the information of reports that are used in forming the required indicator pool.

The underlying logic of selecting reports is the relationship of the report to the triple bottom line of SD. It is expected that each report reflects at least one of the sustainable development pillars; for example, EPI is related to the environmental pillar, while HDI, PF, and SPI are more related to the social pillar and EF and DB refer to the economic pillar. It is assumed that GII can be related to all pillars. Due to the research process, if the abovementioned assumptions are not correct, it cannot negatively affect the research results. Also, the way for more research is open by selecting other or more reports.

- *Step 2: data cleaning*

The reports generally provide information based on a hierarchal structure of variables. They compact operational indicators to create high-level ones. Based on the goal of this research, the operational indicators were collected from each report. In

Table 1 Selected reports for data extraction

Report	Source
Environmental Performance Index (EPI)	https://epi.yale.edu/downloads/epi2020report20210112.pdf
Human Development Index (HDI)	http://hdr.undp.org/en/2020-report
Personal freedom (PF)	www.cato.org/human-freedom-index/2020
Social Progress Index (SPI)	www.socialprogress.org/index/global/results
Economic Freedom (EF)	www.heritage.org/index/download
Doing Business (DB)	www.doingbusiness.org/en/reports/global-reports/doing-business-2020
Global Innovation Index (GII)	www.globalinnovationindex.org/analysis-indicator

sum, 288 indicators were extracted from the reports. Table 2 shows the number of extracted indicators.

There is an operational indicator in GII which reflects the overall result of EPI. To have more homogenous indicators, this indicator was removed from the list. Also, only 127 countries were covered in all the mentioned reports; therefore, just their information was extracted from the publishing reports for the year 2020 and was organized in a database.

- *Step 3: handling missing values*

Approximately 1 percent of the database was not filled due to lacking information in the reports. In other words, there were missing values in the database. By using the global closest fit approach, the missing values of countries were replaced by the most similar country using Manhattan distance criteria:

$$d_{ij} = \sum_{k \in S} \left| c_{ik} - c_{jk} \right|$$

where i and j are denoted for two countries, S represents a set of non-missing indicators in country i and j, and c_k is denoted for k_{th} indicator.

All missing values are filled in using the mentioned method. Finding the most similar country for a country with missing value was a repetitive process. That is, after filling each missing value, the most similar country for the next missing value was found based on the sum of Manhattan distance between the country and other countries. The country with the minimum sum of distances is the similar one in which the missing value was filled by the indicator value of the similar country.

- *Step 4: removing collinear indicators*

The variance inflation factor (VIF) is a measure for finding collinear variables. Based on Algorithm 1, the indicators with higher VIF are iteratively and step-by-step removed. The remaining indicators have lower VIF and then are not collinear.

Algorithm 1: Removing Collinear Indicators

```
1: Input data of 288 indicators
2: Calculate the VIF of each indicator
3: While max(VIF) ≥ 5
4:    Remove vector of the indicator with maximum VIF
5:    Recalculate the VIF of each indicator
6: End
7: Show remained indicators
```

VIF is computed using the following formula:

$$VIF_{\mathrm{i}} = \frac{1}{1 - R_{\mathrm{i}}^2}$$

Table 2 Number of indicators extracted from each report

Report	Number of operational indicators
Environmental Performance Index (EPI)	32
Human Development Index (HDI)	4
Personal Freedom (PF)	34
Social Progress Index (SPI)	50
Economic Freedom (EF)	42
Doing Business (DB)	47
Global Innovation Index (GII)	79
Total	**288**

where i is denoted for a selected indicator and R_i^2 represents the coefficient of determination for the indicator i. The higher the VIF value represents the more collinearity. As Larose (2015) acknowledged if $VIF_i \geq 5$ then the collinearity is moderate. Therefore, to avoid collinearity, we can remove the indicators with the VIF greater than 5 as mentioned in Algorithm 1. Applying the Algorithm caused to finding 135 collinear indicators, then the total number of remaining indicators decreased from 288 to 153.

- *Step 5: removing ineffective indicators*

Some indicators are not effective for participation in predicting SDGI. Therefore, just indicators must be used as input variables which can play an essential role in predicting SDGI by improving the performance of ANNs. The problem of finding the best subset of indicators in this research is an instance of a well-known typical problem in the literature named "feature selection" or "variable selection." There are several methods for producing solutions to the variable selection problem. But De et al. (1997) propose an ANN-based method that uses feature quality index (FQI) as a criterion for ranking variables. The underlying logic of the method is attractive and straightforward; if a variable is not essential, removing it must not harm the result of the network. In other words, if the presence of a variable does not result in better performance, the variable is ineffective and must be removed. Algorithm 2 was designed based on the mentioned logic. It compares the mean

Algorithm 2: Pseudocode of Removing Ineffective Indicators

```
1: Final_List = Ø
2: For i = 1 to 300
3:     List = {all remained indicators}
4:     While List has no change do:
5:         Randomly partition indicators to contain 20 indica-
           tors in each sub-set
6:         For each sub-set
```

```
              Run ANN and save MSE
              For j = 1 to 20
                  Put a vector of zero instead of indicator
                  j in the sub-set
                   Run ANN and save MSE_without_j
                   IF MSE_without_j ≤ MSE
                       Remove indicator j from the List
                  End
               End
           End
      End
      Add List to Final_List and make it unique
   End
   Remove duplicates from Final_List
```

square error (MSE) of an ANN output when a specific variable is present and when its values are replaced by a vector of zero.

To remove all ineffective variables, Algorithm 2 repetitively ran, while the input indicators were the remaining indicators of the previous run. Figure 4 shows the results of ten runs of the Algorithm. Finally, 63 ineffective indicators were found. Therefore, the number of final indicators decreased from 153 to 90.

- *Step 6: finding the best combination of indicators*

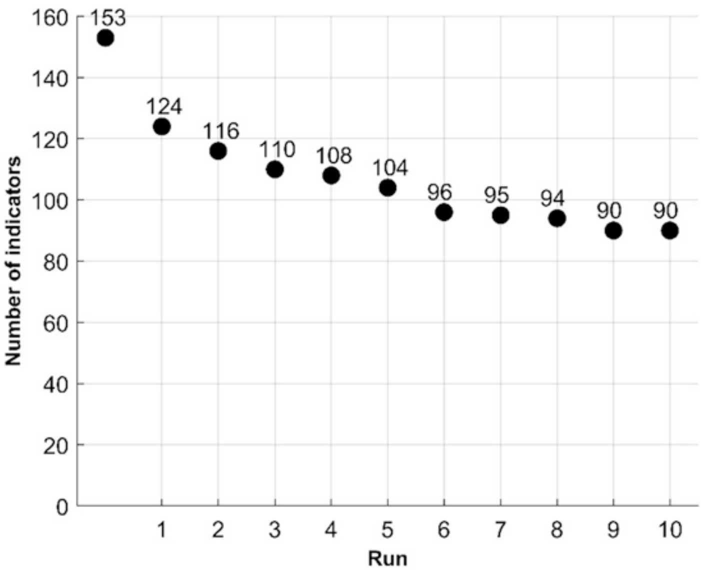

Fig. 4 Reduction of indicators using Algorithm 2

Although 90 indicators are effective in predicting SDGI, a simple predicting system must have a small number of input variables, while being capable of predicting the target values with a reasonable error. Therefore, it is necessary to select a subset of indicators that play the role of inputs for ANNs. It is expected that a simple ANNs design must have limited nodes. In this research, the limitation of nods is set to 20, that is, the number of nodes in ANNs is equal to or less than 20.

Testing from 1 to 20 nodes in ANNs may help to decide about the best number of nodes. It implies that combinations of 1 to 20 from 90 indicators must be tested. The total number of combinations is more than 7×10^{19}. The number of combinations is huge, and testing all of them is an energy- and time-consuming activity, while a good local solution may meet the need. Therefore, instead of testing all combinations, a genetic algorithm (GA) is used to find a reasonable solution. The GA is embedded in a repetitive ANN algorithm. Algorithm 3 shows this approach in more detail.

Algorithm 3: Pseudocode of Combination of ANN and GA

```
Input data of 90 indicators of 127 countries
Set parameters of GA such as number of generations, selec-
tion rate, crossover rate, and mutation rate
For N = 1 to 20     //N denote for the number of nodes
in ANN//
   Generate a population of set-indicators (each set-
   indicator consist of L indicators)
   For i = 1 to number of generation
       For k = 1 to number of population
           For r = 1 to 11
               Run ANN with N nodes using kth set-indicator
               in population as input
               Save RMSE, MAPD, and CorrelCoeff of each ANN
               in Performance(r)
           End
           P(k) = median(Performance)
       End
       Sort the population by RMSE in P and save the Best
       set-indicator
       Apply selection operator to form a part of
       new_population
       Apply crossover operator to form another part of
       new_ population
       Apply mutation operator to form the final part of
       new_ population
       population = new_ population
   End
   Show and save Best set-indicator and related Performance
   for the Node = N
End
```

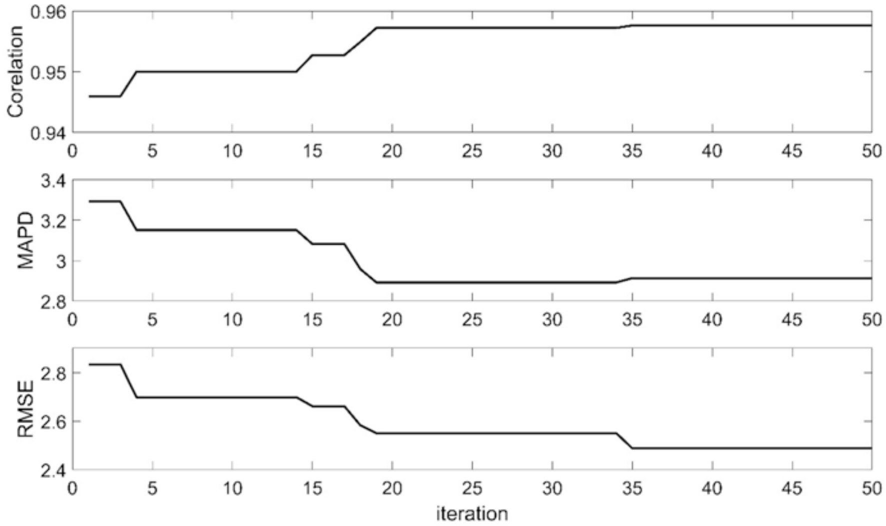

Fig. 5 Convergence plot of GA

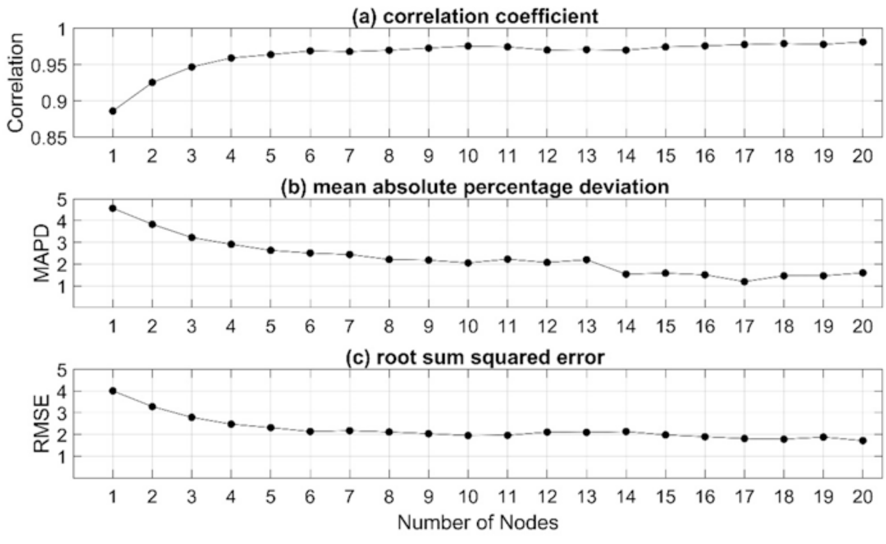

Fig. 6 Performance of ANN with different number of nodes

The GA used in this research encompasses 200 generations with 50 solutions in each generation. The elite replacement, crossover, and mutation rate are set to 0.1, 0.5, and 0.4, respectively. The fitness function is the root mean square error (RMSE) of related ANN. To ensure the robustness of the algorithm output, the ANN ran 11 times, and the median of the RMSEs was reckoned as the value of the fitness function. The selection operator was the roulette wheel, and the crossover method was the random respectful technique. For crossover, three ways were designed: (1) random selection from unused indicators in a selected solution, (2) random selection from indicators that have not emerged in the current solutions, and (3) randomly replacing an indicator in the current selection with a new one. Figure 5 shows the convergence plot of the GA for an ANN with four nodes (indicators). For simplicity, the iteration is limited to 50.

The result of running Algorithm 3 is shown in Fig. 6. The figure reveals that with only four nodes, the correlation between the predicted SDGI and real SDGI is more than 0.95, and on average, there is less than 3% error in predicting the SDGI of each country.

6 Results

The results revealed that among 288 indicators extracted from the selected global reports, just 90 indicators are helpful for predicting SDGI using ANNs. Although more indicators provide better prediction, to keep the simplicity, an ANN with four nodes in one hidden layer can predict SDGI with high accuracy. In the ANN, each node is related to one indicator. The most suitable indicators for predicting SDGI are “Deaths from infectious diseases,” “ICT use,” “Expenditure on education,” and “Assessment in reading, mathematics, and science.” Using these indicators, the ANN can forecast the SDGI with mean absolute percentage deviation (MAPD) equals 2.9126%, RMSE equals 2.4763, and the correlation between the predicted values and SDGI is 0.9592. The results show that designed ANN is a successful predictor for SDGI.

Other combinations of the indicators are also able to predict the SDGI. Table 3 represents some of the combinations. Although the higher the number of nodes produces better performance, the complication of ANN will also increase by adding more nodes.

Table 3 shows that many indicators belong to the Global Innovation Index report. It implies the role of innovation in facilitating the movement toward SDGs and increasing the value of SDGI for countries. Another astonishing fact in the table is the poor emergence of indicators from the EPI, which reports the environmental status. When we can predict SDGI without indicators from the environmental aspect, it means that maybe there is a bias in SDGI. The bias may be occurred due to the insufficient attention to environmental goals in calculating SDGI or undermining the environmental issues in profit of social and or economic issues. This is an interesting topic for further research.

Table 3 Input(s) and performance of ANN

Number of nodes	Indicators	Source	RMSI	MAPD	Correlation
1	ICT use	GII	4.006	4.5694	0.8860
2	Deaths from infectious diseases GERD performed by business enterprise	SPI GII	3.2868	3.8320	0.9253
3	Deaths from infectious diseases ICT use Assessment in reading, mathematics, and science	SPI GII GII	2.7884	3.2222	0.9469
4	Deaths from infectious diseases ICT use Expenditure on education Assessment in reading, mathematics, and science	SPI GII GII GII	2.4761	2.9126	0.9592
5	Deaths from infectious diseases ICT use Expenditure on education Assessment in reading, mathematics, and science Judicial independence	SPI GII GII GII EF	2.3176	2.6337	0.9638
6	Deaths from infectious diseases ICT use Expenditure on education Assessment in reading, mathematics, and science Venture capital deals Utility model applications by origin	SPI GII GII GII GII GII	2.1393	2.5092	0.9689
7	Deaths from infectious diseases ICT use Expenditure on education Assessment in reading, mathematics, and science Venture capital deals SNM.new Hiring and firing regulations	SPI GII GII GII GII EPI EF	2.1757	2.4433	0.9680
8	Deaths from infectious diseases ICT use Child stunting Expenditure on education Assessment in reading, mathematics, and science Patent applications by origin ISO 14001 environmental certificates ICT services imports	SPI GII SPI GII GII GII GII GII	2.1170	2.2160	0.9699

(continued)

Table 3 (continued)

Number of nodes	Indicators	Source	RMSI	MAPD	Correlation
9	Deaths from infectious diseases ICT use Child stunting Expenditure on education Assessment in reading, mathematics, and science Government investment Women's Movement Political and operational stability FGT.new	SPI GII SPI GII GII EF PF GII EPI	2.0407	2.1835	0.9729
10	Deaths from infectious diseases ICT use Expenditure on education Assessment in reading, mathematics, and science Women with advanced education ISO 9001 quality certificates ICT services imports Access to foreign newspapers Paying taxes-time (hours) Employment in knowledge-intensive services	SPI GII GII GII SPI GII GII PF DB GII	1.9536	2.0590	0.9756

7 Conclusions

This chapter explored seven global indexes, including Environmental Performance Index (EPI), Doing Business (DB), Global Innovation Index (GII), Economic Freedom (EF), Personal Freedom (PF), Social Progress Index (SPI), and Human Development Index (HDI). The indexes provide 288 operational indicators from the social, economic, and environmental aspects of 127 countries. The collinear and ineffective indicators were removed in two separate steps. From the 90 remaining indicators, artificial neural networks (ANNs) could yield outstanding results using just a combination of four indicators include "Deaths from infectious diseases," "ICT use," "Expenditure on education," and "Assessment in reading, mathematics, and science." The designed ANN creates a simple model for predicting Sustainable Development Goals Index (SDGI) and avoids the complicated computation of many indicators.

This research also uncovered two facts behind SDGI. First, GII indicators play a prominent role in predicting SDGI. This finding can validate the role of innovation in meeting SDGs and propose to search for solutions to sustainable development problems through innovation. Second, the role of environmental indicators in calculating SDGI is neglectable. Because we succeed in predicting SDGI while ignoring environmental indicators, the SDGI is not relying on environmental indicators, or maybe the role of other aspects is bolder than the environmental aspect. Clarifying the bias in SDGI needs more research. This research also opens the door for using other global reports and indicators to develop another prediction system for SDGI to measure the progress toward SGDs.

References

Abiodun, Oludare Isaac, Aman Jantan, Abiodun Esther Omolara, Kemi Victoria Dada, Nachaat AbdElatif Mohamed, and Humaira Arshad. 2018. State-of-the-Art in Artificial Neural Network Applications: A Survey. *Heliyon* 4 (11): e00938. https://doi.org/10.1016/j.heliyon.2018.e00938.

Aggarwal, Charu C. 2018. Neural Networks and Deep Learning. *Springer* 10: 978–973.

Alaimo, Leonardo Salvatore, Alberto Arcagni, Marco Fattore, and Filomena Maggino. 2021. Synthesis of Multi-indicator System Over Time: A Poset-based Approach. *Social Indicators Research* 157 (1): 77–99. https://doi.org/10.1007/s11205-020-02398-5.

Allen, Cameron, Graciela Metternicht, and Thomas Wiedmann. 2019. Prioritising SDG Targets: Assessing Baselines, Gaps and Interlinkages. *Sustainability Science* 14 (2): 421–438. https://doi.org/10.1007/s11625-018-0596-8.

Alweshah, Mohammed, Saleh Al Khalaileh, Brij B. Gupta, Ammar Almomani, Abdelaziz I. Hammouri, and Mohammed Azmi Al-Betar. 2020. The Monarch Butterfly Optimization Algorithm for Solving Feature Selection Problems. *Neural Computing and Applications*. https://doi.org/10.1007/s00521-020-05210-0.

Antanasijević, Davor Z., Viktor V. Pocajt, Dragan S. Povrenović, Mirjana Đ. Ristić, and Aleksandra A. Perić-Grujić. 2013. PM10 Emission Forecasting Using Artificial Neural Networks and Genetic Algorithm Input Variable Optimization. *Science of the Total Environment* 443: 511–519. https://doi.org/10.1016/j.scitotenv.2012.10.110.

Badar, Altaf QH. 2021. *Evolutionary Optimization Algorithms*.

Bali Swain, R., and F. Yang-Wallentin. 2020. Achieving Sustainable Development Goals: Predicaments and Strategies. *International Journal of Sustainable Development & World Ecology* 27 (2): 96–106. https://doi.org/10.1080/13504509.2019.1692316.

Barbier, Edward B., and Joanne C. Burgess. 2019. Sustainable Development Goal Indicators: Analyzing Trade-Offs and Complementarities. *World Development* 122: 295–305. https://doi.org/10.1016/j.worlddev.2019.05.026.

Berglund, Teresa, and Niklas Gericke. 2016. Separated and Integrated Perspectives on Environmental, Economic, and Social Dimensions – An Investigation of Student Views on Sustainable Development. *Environmental Education Research* 22 (8): 1115–1138. https://doi.org/10.1080/13504622.2015.1063589.

Borah, Pallabi, Hasin A. Ahmed, and Dhruba K. Bhattacharyya. 2014. A Statistical Feature Selection Technique. *Network Modeling Analysis in Health Informatics and Bioinformatics* 3 (1): 55. https://doi.org/10.1007/s13721-014-0055-0.

Brusseau, M.L. 2019. Chapter 32 – Sustainable Development and Other Solutions to Pollution and Global Change. In *Environmental and Pollution Science*, ed. Mark L. Brusseau, Ian L. Pepper, and Charles P. Gerba, 3rd ed., 585–603. Academic Press.

Chakraborty, Basabi. 1999. Feature Selection by Artificial Neural Network for Pattern Classification. In *Pattern Recognition in Soft Computing Paradigm*, 95–109.

Collins, Gary S., and Karel G.M. Moons. 2019. Reporting of Artificial Intelligence Prediction Models. *The Lancet* 393 (10181): 1577–1579. https://doi.org/10.1016/S0140-6736(19)30037-6.

Costanza, Robert, Lew Daly, Lorenzo Fioramonti, Enrico Giovannini, Ida Kubiszewski, Lars Fogh Mortensen, Kate E. Pickett, Kristin Vala Ragnarsdottir, Roberto De Vogli, and Richard Wilkinson. 2016. Modelling and Measuring Sustainable Wellbeing in Connection With the UN Sustainable Development Goals. *Ecological Economics* 130: 350–355. https://doi.org/10.1016/j.ecolecon.2016.07.009.

De, Rajat K., Nikhil R. Pal, and Sankar K. Pal. 1997. Feature Analysis: Neural Network and Fuzzy Set Theoretic Approaches. *Pattern Recognition* 30 (10): 1579–1590. https://doi.org/10.1016/S0031-3203(96)00190-2.

Ding, Shifei, Hui Li, Su Chunyang, Yu Junzhao, and Fengxiang Jin. 2013. Evolutionary Artificial Neural Networks: A Review. *Artificial Intelligence Review* 39 (3): 251–260. https://doi.org/10.1007/s10462-011-9270-6.

Elliott, Jennifer. 2012. *An Introduction to Sustainable Development*. Routledge.

Emmanuel, Okewu, M. Ananya, Sanjay Misra, and Murat Koyuncu. 2020. A Deep Neural Network-Based Advisory Framework for Attainment of Sustainable Development Goals 1-6. *Sustainability* 12 (24): 10524. https://doi.org/10.3390/su122410524.

Floridi, Luciano. 2002. *Philosophy and computing: An introduction*. Routledge.

George, Edward I. 2000. The Variable Selection Problem. *Journal of the American Statistical Association* 95 (452): 1304–1308. https://doi.org/10.1080/01621459.2000.10474336.

Gnana, D. Asir, S. Appavu Antony, Alias Balamurugan, E. Jebamalar, and Leavline. 2016. Literature Review on Feature Selection Methods for High-Dimensional Data. *International Journal of Computer Applications* 136 (1): 9–17.

Gue, Ivan Henderson V., Aristotle T. Ubando, Ming-Lang Tseng, and Raymond R. Tan. 2020. Artificial Neural Networks for Sustainable Development: A Critical Review. *Clean Technologies and Environmental Policy* 22 (7): 1449–1465. https://doi.org/10.1007/s10098-020-01883-2.

Hagan, M.T., H. Demuth, M. Beale, and O. De Jesus. 2016. *Neural Network Design*. 2nd ed. Lexington.

Hák, Tomáš, Svatava Janoušková, and Bedřich Moldan. 2016. Sustainable Development Goals: A Need for Relevant Indicators. *Ecological Indicators* 60: 565–573. https://doi.org/10.1016/j.ecolind.2015.08.003.

Horan, David. 2020. National Baselines for Integrated Implementation of an Environmental Sustainable Development Goal Assessed in a New Integrated SDG Index. *Sustainability* 12 (17): 6955. https://doi.org/10.3390/su12176955.

Iddrisu, Insah, and Subhes C. Bhattacharyya. 2015. Sustainable Energy Development Index: A Multi-Dimensional Indicator For Measuring Sustainable Energy Development. *Renewable and Sustainable Energy Reviews* 50: 513–530. https://doi.org/10.1016/j.rser.2015.05.032.

Katoch, Sourabh, Sumit Singh Chauhan, and Vijay Kumar. 2021. A Review on Genetic Algorithm: Past, Present, and Future. *Multimedia Tools and Applications* 80 (5): 8091–8126. https://doi.org/10.1007/s11042-020-10139-6.

Keeble, Brian R. 1988. The Brundtland Report: 'Our Common Future'. *Medicine and War* 4 (1): 17–25. https://doi.org/10.1080/07488008808408783.

Kramer, Oliver. 2017. Genetic Algorithms. In *Genetic Algorithm Essentials*, 11–19. Springer.

Kroll, Christian. 2015. *Sustainable Development Goals: Are the rich countries ready?* Bertelsmann Stiftung Gütersloh, Germany.

Larose, Daniel T. 2015. *Data Mining and Predictive Analytics*. Wiley.

Leal Filho, Walter. 2020. Viewpoint: Accelerating the Implementation of the SDGs. *International Journal of Sustainability in Higher Education* 21 (3): 507–511. https://doi.org/10.1108/IJSHE-01-2020-0011.

Mirjalili, Seyedali. 2019. Genetic Algorithm. In *Evolutionary Algorithms and Neural Networks: Theory and Applications*, 43–55. Cham: Springer.

Mishra, M., and M. Srivastava. 2014. *A View of Artificial Neural Network*. 2014 International Conference on Advances in Engineering & Technology Research (ICAETR – 2014), 1–2 Aug. 2014.

Moriarity, Sean. 2021. *Genetic Algorithms in Elixir: Solve Problems Using Evolution*. Pragmatic Bookshelf.

Okwu, Modestus O., and Lagouge K. Tartibu. 2021. Artificial Neural Network. In *Metaheuristic Optimization: Nature-Inspired Algorithms Swarm and Computational Intelligence, Theory and Applications*, 133–145. Cham: Springer International Publishing.

Omri, Anis. 2020. Technological Innovation and Sustainable Development : Does the Stage of Development Matter? *Environmental Impact Assessment Review* 83: 106398. https://doi.org/10.1016/j.eiar.2020.106398.

Paliwal, Mukta, and Usha A. Kumar. 2009. Neural Networks and Statistical Techniques: A Review of Applications. *Expert Systems with Applications* 36 (1): 2–17. https://doi.org/10.1016/j.eswa.2007.10.005.

Reyers, Belinda, Mark Stafford-Smith, Karl-Heinz Erb, Robert J. Scholes, and Odirilwe Selomane. 2017. Essential Variables help to focus Sustainable Development Goals monitoring. *Current Opinion in Environmental Sustainability* 26-27: 97–105. https://doi.org/10.1016/j.cosust.2017.05.003.

Ruiz-Morales, Betzabe, Irma Cristina Espitia-Moreno, Victor G. Alfaro-Garcia, and Ernesto Leon-Castro. 2021. Sustainable Development Goals Analysis with Ordered Weighted Average Operators. *Sustainability* 13 (9): 5240. https://doi.org/10.3390/su13095240.

Russell, Stuart, and Peter Norvig. 2021. *Artificial Intelligence: A Modern Approach*. 4th ed. Pearson.

Sachs, J.D., and J.W. McArthur. 2005. The Millennium Project: A Plan For Meeting the Millennium Development Goals. *The Lancet* 365 (9456): 347–353. https://doi.org/10.1016/S0140-6736(05)17791-5.

Sachs, Jeffrey D, Guido Schmidt-Traub, Christian Kroll, Guillaume Lafortune, Grayson Fuller, and Finn Woelm. 2020. *Sustainable Development Report 2020*.

Sachs, Jeffrey, Guido Schmidt-Traub, Christian Kroll, David Durand-Delacre, and Katerina Teksoz. 2017. *SDG Index and Dashboards Report 2017*. New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN).

Schmidt-Traub, Guido, Christian Kroll, Katerina Teksoz, David Durand-Delacre, and Jeffrey D. Sachs. 2017. National baselines for the Sustainable Development Goals assessed in the SDG Index and Dashboards. *Nature Geoscience* 10 (8): 547–555. https://doi.org/10.1038/ngeo2985.

Shaker, Richard Ross. 2018. A Mega-Index for the Americas and Its Underlying Sustainable Development Correlations. *Ecological Indicators* 89: 466–479. https://doi.org/10.1016/j.ecolind.2018.01.050.

Sivanandam, S.N., and S.N. Deepa. 2008. Genetic Algorithms. In *Introduction to Genetic Algorithms*, 15–37. Berlin, Heidelberg: Springer Berlin Heidelberg.

———. 2006. *Introduction to Neural Networks Using Matlab 6.0*. Tata McGraw-Hill Education.

Strezov, Vladimir, Annette Evans, and Tim J. Evans. 2017. Assessment of the Economic, Social and Environmental Dimensions of the Indicators for Sustainable Development. *Sustainable Development* 25 (3): 242–253. https://doi.org/10.1002/sd.1649.

Umbarkar, Anant J., and Pranali D. Sheth. 2015. Crossover Operators in Genetic Algorithms: A Review. *ICTACT Journal on Soft Computing* 6 (1).

UN. 2021. *Envision2030: 17 Goals to Transform the World for Persons with Disabilities*. Accessed 13 Oct 2021. https://www.un.org/development/desa/disabilities/envision2030.html.

Wu, Yu-chen, and Jun-wen Feng. 2018. Development and Application of Artificial Neural Network. *Wireless Personal Communications* 102 (2): 1645–1656. https://doi.org/10.1007/s11277-017-5224-x.

Xu, Zhenci, Sophia N. Chau, Xiuzhi Chen, Jian Zhang, Yingjie Li, Thomas Dietz, Jinyan Wang, Julie A. Winkler, Fan Fan, Baorong Huang, Shuxin Li, Wu Shaohua, Anna Herzberger, Ying Tang, Dequ Hong, Yunkai Li, and Jianguo Liu. 2020. Assessing Progress Towards Sustainable Development Over Space and Time. *Nature* 577 (7788): 74–78. https://doi.org/10.1038/s41586-019-1846-3.