Chapter 28
Artificial Intelligence Using Distribution Free Data (90 Patients)

1 General Purpose

Artificial intelligence is an engineering method that simulates the structures and operating principles of the human brain. The artificial neural network is a distribution-free based on layers of artificial neurons that transduce imputed information.

2 Schematic Overview of Type of Data File

Outcome	predictor	predictor	predictor...
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.

3　Primary Scientific Question

Does artificial intelligence better predict nonlinear outcomes from multiple predictors than other models, like mathematical equations obtained from regression models.

4　Data Example

Gender	Age	Weight	Height	Surfacemeas	Surfacecomp
1,00	13,00	30,50	138,50	10072,90	10770,00
,00	5,00	15,00	101,00	6189,00	6490,00
,00	,00	2,50	51,50	1906,20	1890,00
1,00	11,00	30,00	141,00	10290,60	10750,00
1,00	15,00	40,50	154,00	13221,60	13080,00
,00	11,00	27,00	136,00	9654,50	10000,00
,00	5,00	15,00	106,00	6768,20	6610,00
1,00	5,00	15,00	103,00	6194,10	6540,00
1,00	3,00	13,50	96,00	5830,20	6010,00
,00	13,00	36,00	150,00	11759,00	12150,00

Gender 1 male, 0 female
age years
weight kg
height meters (m)
surfacemeas = surface measured m^2
surfacecomp = surface computed from Hancock equation (J Pediatr 1978).

　　We will use neural network instead of the Hancock equation for predicting the body surface from the body height and weight. The above data file consists of a row for the first 10 patients from a 90 patient study with different factors (left four columns) and one dependent variable, the photometrically measured body surface (variable 5). The entire data file is in extras.springer.com, and is entitled "chapter28neuralnetwork". Using SPSS with the neural network add-on module, we will assess whether a neural network with two hidden layers of neurons is able to adequately predict the measured body surfaces, and whether it outperforms the mathematical model of Haycock (* = sign of multiplication):

$$\text{body surface} = 0.024265 * \text{height}^{0.3964} * \text{weight}^{0.5378}.$$

Start by opening the data file in SPSS.

5 Neural Network Analysis

For analysis the statistical model Multilayer Perceptron in the module Neural Networks is required.

Command:

Neural Networks. . .. Multilayer Perceptron. . ..Select Dependent Variable: the measured body surface. . .. Factors: body height and weight, and covariates, age and gender....main dialog box....click Partitioning: set the Training Sample (70), Test Sample (20)....click Architecture: set the Numbers of Hidden Layers (2)....click Activation Function: click Hyperbolic Tangens....click Output: click Diagrams, Descriptions, Synaptic Weights....click Training: Maximal Time for Calculations 15 min, Maximal Numbers of Iterations 2000....click OK.

The synaptic weights and body surfaces predicted by the neural network are displayed in the main screen. The results are in the 7th column of the data file.

Gender	Age	Weight	Height	Surfacemeas	Surfacecomp	Surfacepred
1,00	13,00	30,50	138,50	10072,90	10770,00	10129,64
,00	5,00	15,00	101,00	6189,00	6490,00	6307,14
,00	,00	2,50	51,50	1906,20	1890,00	2565,16
1,00	11,00	30,00	141,00	10290,60	10750,00	10598,32
1,00	15,00	40,50	154,00	13221,60	13080,00	13688,06
,00	11,00	27,00	136,00	9654,50	10000,00	9682,47
,00	5,00	15,00	106,00	6768,20	6610,00	6758,45
1,00	5,00	15,00	103,00	6194,10	6540,00	6533,28
1,00	3,00	13,50	96,00	5830,20	6010,00	6096,53
,00	13,00	36,00	150,00	11759,00	12150,00	11788,01

Gender 1 male, 0 female
age years
weight kg
height meters (m)
surfacemeas = surface measured m^2
surfacecomp = surface computed from Hancock equation (J Pediatr 1978)
surfacepred = surface predicted from neural network

Both the predicted values from the neural network and from the Haycock equation are close to the measured values. When performing a linear regression with neural network as predictor, the r square value was 0,983, while the Haycock produced an r square value of 0,995. Although the Hancock equation performed slightly better, the neural network method produced adequate accuracy defined as an r-square value larger than 0,95.

6   Conclusion

We conclude that neural network is a very sensitive data modeling program, particularly suitable for making predictions from non-Gaussian data. Like Monte Carlo methods it is a distribution-free methodology, which is based on layers of artificial neurons that transduce imputed information. It is available in the SPSS add-on module Neural Network. Artificial intelligence, otherwise called neural network, is a data producing methodology that simulates the structures and operating principles of the human brain. It can be used for modeling purposes, and is, particularly, suitable for modeling distribution-free and nonnormal data patterns.

7   Note

More background, theoretical, and mathematical information of Artificial intelligence is given in Statistics applied to clinical studies 5th edition, Chap. 58, Springer Heidelberg Germany, 2012, from the same authors.