Lancaster University Management School
Working Paper
2005/063

Skills and earnings revisited

Geraint Johnes

The Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
UK

SKILLS AND EARNINGS REVISITED

Geraint Johnes

Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
United Kingdom

Voice: +44 1524 594215
Fax: +44 1524 594244
E-mail:G.Johnes@lancaster.ac.uk

First version: September 2005

ABSTRACT

Regression and neural network models of wage determination are constructed where the explanatory variables include detailed information about skills. People skills, strategic skills, and IT skills all carry strong and significant wage premia; problem-solving skills (surprisingly) and physical skills (less surprisingly) do not. In contrast to the impact of school curriculum on subsequent earnings, the neural network modelling procedure does not pick up any significant nonlinearities in the relationship between skills and earnings.

JEL Classification: I21, J24, J31, C45
Keywords: skills, earnings, neural networks

The author is indebted to the Data Archive at the University of Essex for arranging access to the Skills Survey data used herein. For useful and constructive discussions, he also thanks, without implication, participants at the September 2005 Tartu conference on education economics.

Introduction

In a recent paper (Johnes, 2005), I have investigated the relationship between school curriculum and subsequent earnings. This is done in the context of a highly nonlinear model which allows for synergies across subjects, using methodology of neural networks. A key finding of this work is that the interaction between subjects within a curriculum is of considerable importance in determining future earnings.

A natural extension of this work is to investigate whether there exist nonlinear effects also in the relationship between skills and earnings.[1] Hence in the present paper, I use data from the Second Skills Survey of the Employed British Workforce to investigate this relationship; this survey was conducted in the early months of 2001. The work reported here builds on earlier work by Green (1998) and Dickerson and Green (2004), but differs from those contributions in that we here consider the interaction between a variety of skills.

The paper begins with a review of the relevant literature on both curriculum and skills. This is followed by a description of the methodology to be used and of the data employed in the present study. The results of the analysis are then reported and conclusions drawn.

1. Received literature

The skills agenda has gained in prominence over the last decade or so in the developed economies, with increasing awareness that the comparative advantage of such economies lies in high value added production. The quest to tilt comparative advantage in a favourable direction has led researchers to investigate the returns to skills of different kinds. Hence, for example, Murnane *et al*. (1995) have investigated the role played by cognitive skills (though these are rather crudely measured simply as mathematics test scores). In the UK, Machin (1996) investigated the change in the demand for skills, though this analysis was necessarily conducted at a high level of aggregation.

More detiled analysis became possible with the release of results from the first British Skills Survey, based on interviews conducted during the early part of 1997 (Green, 1998). There was, at this time, a suspicion that IT skills were being highly rewarded (Krueger, 1993) – though equally it has been shown that skills in using a pencil are also of value in the workplace (DiNardo and Pischke, 1997)!

In the parlance of educational researchers, the recent interest in skills has been termed the ‘search for the new literacies’, a desire to understand what are the basic skills that replace (or, probably more accurately, supplement) the three R’s of reading, ‘riting and ‘rithmetic. Candidates include IT skills, people skills, problem-solving skills, and so on. Green’s findings – based on an analysis of the first few principal components obtained by reducing data on some 38 distinct skill variables - suggest that IT skills

[1] I am indebted to Joseph Hassid who pointed this out to me during my presentation of Johnes (2005) at the Tartu conference on education economics, and who therefore provided the inspiration for the present paper.

earned a wage premium in 1997 of around 20 per cent. Communication skills and problem-solving were also highly valued in the labour market at this time.

There are several reasons why, even after a relatively short time, Green's results ought now to be revisited. First, the position in 1997 was not necessarily an equilibrium. In particular, the high premium paid to workers with IT skills at that time may not be sustainable in the long run, as we would expect the premium itself to entice more workers (especially the young) to train in this area, thereby increasing the supply of workers with these skills. Secondly, the availability of new data, from the Second Skills Survey, allows a snapshot to be taken at a more recent date (2001). Thirdly, the estimation technology introduced by Johnes (2005) in the area of curriculum offers much to our understanding of the way in which skills combine together – in a fashion reminiscent of the consumption technology literature pioneered by Lancaster (1966) – to offer employers a package, known as a worker, that is to a greater or lesser degree attractive. The opportunity exists now, therefore, to apply new estimation methodologies to new data on skills.

The synergies that might exist between skills are reminiscent of those that have been shown to exist between subjects studied in the curriculum at upper secondary school (Johnes, 2005). In the latter case, of course, it is no surprise to find that there exist economies of scope across subjects; this is, after all, why curricula composed of several subjects exist in the first place. If there was no synergy, there would be no reason to study subjects together rather than sequentially. This insight brought me, in my earlier paper, to investigate a flexible nonlinear model of earnings determination which allowed for interaction between subjects by way of a neural network (White, 1992). The results, based on data from the National Child Development Survey, strongly suggest that the impact of a subject on subsequent earnings depend crucially on the bundle of other subjects with which it is studied. This is illustrated in Table 1. Here we see that substituting mathematics for biology in a curriculum that also includes physics and chemistry has no effect on expected subsequent earnings. However, substituting mathematics for art in a curriculum that includes also history and physics has a substantial positive impact. The same applies to other subjects: in some curricula they impact on earnings, but in others they do not.

This work is important for two reasons. First, it suggests that claims made about the social value of a broad (or, for that matter, a narrow) curriculum are spurious inasmuch as they apply too broad a brush (Dolton and Vignoles, 2002a). Some broad curricula are good (in the sense that they enhance subsequent remuneration), some are bad; some narrow curricula are good, some are bad. It is the precise constellation of subjects contained within the curriculum that matters. Secondly, the results confirm that all subjects have an impact on subsequent earnings. This contradicts earlier findings by Altonji (1995), Dolton and Vignoles (2002b) and others which suggested that most subjects (with mathematics as an exception) do not impact on labour market performance. This is a key result, because Altonji's findings have been interpreted as evidence in favour of the signalling and screening model of wage determination, and against the human capital model.[2] The results of Johnes (2005) can therefore be interpreted as a blow against the signalling and screening model.

[2] His results suggest that years of education (which serve as a signal) influence subsequent earnings, whereas the curriculum content (which ought to determine human capital) does not.

It is clear that the flexibility afforded by a specification of the earnings function that permits interaction between key explanatory variables – be they the components of a curriculum or skills – can offer substantial new insight. In the remainder of this paper, we shall, within this type of framework, consider the specific role played by skills in wage determination.

2. Methodology

A simple model of wage determination can be defined by

$$\ln w = \mathbf{Z}\beta + f(\mathbf{x}) + u \qquad (1)$$

where w denotes the wage paid to an individual, **Z** is a vector of individual characteristics, **x** is a vector of variables each of which provides a measure of the worker's stock of an identifiable skill, and u is an error.[3] The simplest modelling procedure would be to impose a linearity constraint on f(**x**); results obtained from such a model are reported below, and in this case provide quite a good fit to the data.

But we can do better than *impose* such an assumption of linearity. The method of neural networks allows us to specify f(**x**) in a general manner that allows it to *approximate arbitrarily closely* whatever linear or nonlinear relationship truly maps the right hand side of (1) onto the left hand side. To see this intuitively, consider Figure 1. This shows how a simple feedforward neural network processes information. The large black dots represent neurodes. These are arranged in three layers. An input layer feeds into a second ('hidden') layer which in turn feeds into an output layer. In our application, the inputs are explanatory variables – in this case the x vector of skills. The output, f, is the impact that these skills has on (log) wages.

Information flows from each neurode in the input layer as a signal to each neurode in the hidden layer. There, the information is processed. This processing takes the following form: a weighted average of the signals is calculated; this is then transformed using some nonlinear mapping (which we may refer to as a 'squasher'). The squashed signal is then passed onto the output layer, where again a weighted average of incoming signals is squashed.

The combination of many nonlinear transformations within a network of this type allows the network to serve as a universal approximator. This explains the appeal of neural networks as an estimating technology used in contexts where the true functional form is unknown or extremely sophisticated. They are used in a large variety of pattern recognition softwares, including voice recognition and optical character recognition; they are also used in many economic forecasting contexts where the possibility of nonlinear processes might lead to chaotic dynamics that would otherwise be difficult to predict (see, for example, Swanson and White, 1997; Johnes, 1999, 2000).

---

[3] To keep the analysis relatively straightforward, I do not investigate interactions between the **x** and **Z** vectors. This is, of course, a feature of the present analysis that is common to all earlier, linear, studies of the effect of skills on earnings.

Formally, imposing a logistic squasher, a precise definition of f(**x**) may be given by

$$f(\mathbf{x}) = 1/\{1+\exp\{-\sum_{i}^{m} \rho_i / [1+\exp(-\sum_{j}^{n} \theta_{ij}\, x_j)]\}\} \tag{2}$$

where m is the number of neurodes in the (single) hidden layer and n is the number of different skills identified in the study.

The Second Skills Survey includes questions about a range of some 43 detailed skills that workers may or may not find important in carrying out their jobs. The degree to which each skill is deemed essential is evaluated on a 5-point Likhert scale. While it would be possible to include each of these skill variables in an earnings regression, there is inevitably a risk of multicollinearity. To counter this, we replicate the method of Green (1998) who reduces the dimensionality of the problem by using as regressors the first few principal components of the skills variables. In our case, ten principal components have eigenvalues above one, and so these have been retained for the analysis. Examination of the component matrix reveals that fairly unambiguous interpretations can be assigned to each of the retained principal components.

It should be clear from the above discussion that the architecture of a neural network is crucial in determining its success in fitting the data. A network which contains many neurodes in its hidden layer (or one which contains several hidden layers) could be constructed that fitted the data perfectly. This is not necessarily a desirable situation. Since any data contain a mix of signal and noise, a perfect fit would imply that the noise has been modelled as well as the signal, and that would mean that the model has poor out-of-sample forecasting properties. To guard against overfitting in this way, the model in the present paper has been kept very simple, with m=1. This has the further advantage that the coefficients of the model can be estimated statistically using standard nonlinear least squares, yielding the usual battery of statistical diagnostics. In this case, as we shall see later, the coefficients of the model are estimated with a reasonably high degree of precision. The full model is therefore given by

$$\ln w = \mathbf{Z}\beta + 1/\{1+\exp\{-\rho / [1+\exp(-\sum_{1}^{10} \theta_j\, x_j)]\}\} + u \tag{3}$$

It should be noted at this stage, that the neural network model as applied in this context performs a function that is in many respects analogous to a model in which the vector of regressors includes a full set of interaction terms between skills. The neural network approach is chosen in order to economise on degrees of freedom; with 10 skill variables, it would be necessary to include 1023 interaction terms.[4] As is clear from (3), the neural network allows us to replace these with just eleven terms, thus providing considerable economy in the estimation procedure.

[4] That is ${}_{10}C_1+{}_{10}C_2+{}_{10}C_3+\ldots+{}_{10}C_{10}$.

3. Data

The Second Skills Survey of the Employed British Workforce was conducted in February and March of 2001. It provides data on over 4000 employed individuals, with data collected on variables including age, schooling, experience, tenure, occupation, personal characteristics, and a large variety of skills variables.[5] Results of the survey have been analysed in a number of papers, the closest of which to the present study is that of Dickerson and Green (2004).

Descriptive statistics for key variables are presented in Table 2. The mean wage is £7.86 per hour; a useful benchmark for this is to compare it with the national minimum wage, which for adults at that time was £3.70.

The schooling variable used in this study is the age at which full-time schooling ended minus five. Experience is given by years of actual work experience – the Skills Survey is unusual in collecting this information directly from respondents – and is measured in years. Data on tenure with the current employer are also available, and are measured in months. Somewhat less than one third of respondents are members of a trade union, and a little over one half of all respondents are men. Just over one in twenty respondents are from ethnic minorities. Some 56 per cent of respondents were married at the time of the survey. Many respondents had responsibility for children under the age of 16; on average, respondents were each responsible for 0.78 children in this age group.

The demographic characteritsitcs of the sample – as measured by the variables discussed above and also by the regional distribution of respondents[6] – suggest that the sampling in the survey was representative of the population of employed workers in Britain in 2001.

4. Results

In Table 3, the results of a linear regression are reported. The signs on the coefficients on variables within Z are unsurprising, in line as they are with those obtained in many previous studies of this kind. The coefficient on years of schooling, at about 0.04, is somewhat lower than is typical, but one might surmise that the presence in the equation of skills variables accounts for this.[7] The coefficients on the linear and quadratic terms in experience suggest that earnings peak after 28 years. Tenure at the current employer further raises earnings (the quadratic term here being insignificant).

---

[5] The fact that all respondents are employed means that it is not possible to correct for sample selection biases due to nonrandom selection of individuals into employment. A small number of individuals are not included in the sample analysed in the present paper owing to incomplete data on wages, experience, or schooling.

[6] Not reported here for reasons of space.

[7] With some sense of irony, I note that the strongly significant coefficient on schooling might be interpreted as evidence of signalling – and that this contrasts with the results of the companion paper (Johnes, 2005). If human capital is captured by the skills variables, and if schooling (observable by firms) merely acts as a signal of the presence of skills (that are observed by us but not necessarily by firms), then the fact that schooling is significant reflects signalling. Of course, the significance of the skill variables suggests that human capital too has a part to play in wage determination.

The premia associated with being male and with being white are in line with those found in other studies. For men, there is a premium associated with marriage; the presence of dependent children does not affect the earnings of men or women.[8] Union membership is associated with a significant wage premium of around 6 per cent.

Regional dummies (not reported in the table for reasons of space) indicate that wages are, other things being equal, highest in London, followed by the rest of the south east and the east. This result, too, is consistent with the findings of many earlier studies of wage determination in the UK.

Now turn to the coefficients on the skill variables. Three of these are negative: unsurprisingly jobs requiring physical skills are less remunerative than others; more surprising is the fact that jobs requiring problem-solving skills are relatively unremunerative, though this is a finding that concurs with the results of Dickerson and Green (2004). The third negative coefficient attaches to basic literacy.

Strongly positive and significant coefficients attach to the other skills identified in the study. Unsurprisingly perhaps, the highest coefficient (and the most significant) is associated with a skill variable that encompasses human resource management and strategic planning skills. These skills, all other things being equal, carry an earnings premium of around 11 per cent.[9] A premium of 9 per cent is associated with IT skills, and a premium of 6 per cent with quantitative skills. All of the skill variables are statistically significant at conventional levels.

Table 4 reports the results of the neural network specification given by equation (3). The coefficients on the linear terms are similar to the corresponding measures reported in Table 3. Meanwhile the coefficients on the nonlinear terms are difficult to interpret. For this reason we report the results of an exercise designed to illustrate how skills may combine with each other to have an effect on wages that is independent of the impact that each skill, on its own, can have. A male located in Wales, given mean values of all other variables in the **Z** vector, and given a value of -½ for all skill measures,[10] would – according to the model in Table 3 – be expected to earn £8.43 per hour (measured, of course, in 2001 UK£). Each row of Table 4 investigates the effect of simultaneously increasing to ½ the value of three of the skill measures, holding everything else constant. It is readily observed that some groups of skills are considerably more highly rewarded than others. The combination of skills in the table that is most highly remunerated is strategy and HR, IT and advanced analytical skills.

In common with the findings reported in Table 1, the results of the neural network analysis suggests that the premium attached to any one specific skill might depend upon the skills with which it is combined. For example, as is evident from Table 5, replacing oral communication skills by quantitative skills has a negligible effect on expected earnings if these skills are accompanied by written communication and physical skills, but a much larger effect if accompanied by written communication

[8] Of course, only employed people are included in the survey; it is likely that the presence of children has an impact on participation, especially of women, but this is something upon which the present dataset cannot throw light.

[9] $100(e^{\beta}-1)$.

[10] Recall that the skill measures are principal components. These therefore have a mean of zero and a standard deviation of 1.

and teamwork skills. However, for reasons that will become apparent in a moment, not too much should be made of this finding.

Having reported the results of both the linear and the neural network specifications of the earnings function, we are in a position to consider the key question: which performs better? The answer to this question may be obtained by use of White's neural network test for nonlinearity (White, 1989; Lee *et al*., 1993; Granger and Teräsvirta, 1993). The diagnostic statistic of 0.24 unambiguously fails the $\chi^2(1)$ test at conventional levels of significance. Hence it can be concluded that, in this case (in contrast with the case of curriculum considered by Johnes, 2005), the neural network model adds nothing to our understanding of wage determination beyond what we already know from linear models.

5. Conclusion

In one respect, the positive contribution of the present paper might therfore be regarded as limited. The results above confirm that nonlinearity is not an issue in the mapping between skills and earnings – at least in Britain at the turn of the millennium. This is reassuring in that it provides considerable backing to the results obtained by Dickerson and Green (2004).

Physical skills are remunerated relatively poorly. Amongst skills that are well remunerated, personal and strategic skills and IT skills stand out as having strong impact. Surprisingly (although this confirms the finding of Dickerson and Green), and in contrast with results from the first skills survey, problem-solving skills do not appear to carry a positive premium.

In another respect, however, the analysis reported here, and in the companion paper (Johnes, 2005) provide a challenge to applied microeconometricians working in the economics of education and in other fields alike. Tools for estimating sophisticated nonlinearities and interactions between variables, with economy in terms of degrees of freedom, have been available for some time now. As my work on curriculum has shown, these nonlinearities can sometimes be important – both in a quantitative sense and in dimensions that are more conceptual or policy-related. Yet they remain undetected by traditional linear econometric estimation, and this suggests that there is little room for complacency in the rigour of our estimation methods. Subjecting the analysis of earnings as a function of skills to a battering by nonlinear analysis, and finding that the results from more conventional investigations in the received literature pass unscathed, is comforting news – but it *is* still news.

The changes in premia attached to some skills between 1997 and 2001 are noteworthy, but should not surprise economists trained to view the market as ensuring that, in the long run, rates of return to investments in training will equalise. This being the case, it is likely that the returns to various skills will change further over the years ahead. A re-evaluation on a regular basis is likely to prove of value.

References

Altonji, Joseph (1995) "The effects of high school curriculum on education and labor market outcomes." *Journal of Human Resources*, 30, 409-438.

Dickerson, Andy and Green, Francis (2004) "The growth and valuation of computing and other generic skills." *Oxford Economic Papers*, 56, 371-406.

DiNardo, John E. and Pischke, Jörn-Steffen (1997) "The returns to computer use revisited: have pencils changed the wage structure too?" *Quarterly Journal of Economics*, 112, 291-303.

Dolton, Peter and Vignoles, Anna (2002a) "Is a Broader Curriculum Better?" *Economics of Education Review*, 21, 415-430.

Dolton, Peter and Vignoles, Anna (2002b) "The returns on post-compulsory school mathematics study." *Economica*, 69, 113-141.

Green, Francis (1998) "The value of skills." mimeo, University of Kent.

Granger, Clive W.J. and Teräsvirta, Timo (1993) "*Modelling nonlinear economic relationships.*", Oxford: Oxford University Press.

Johnes, Geraint (1999) "Forecasting unemployment." *Applied Economics Letters*, 6, 605-607.

Johnes, Geraint (2000) "Up around the bend: linear and nonlinear models of the UK economy compared." *International Review of Applied Economics*, 14, 485-493.

Johnes, Geraint (2005) "Don't know much about history...: Revisiting the impact of curriculum on subsequent labour market outcomes." *Bulletin of Economic Research*, 57, 249-272.

Krueger, Alan B. (1993) "How computers have changed the wage structure: evidence from microdata 1984-1989." *Quarterly Journal of Economics*, 108, 33-60.

Lancaster, Kelvin J. (1966) "Change and innovation in the technology of consumption." *American Economic Review*, 56, S14-S23.

Lee, Tae-Hwy, White, Halbert and Granger, Clive W.J. (1993) "Testing for neglected nonlinearity in time series models." *Journal of Econometrics*, 56, 269-290.

Machin, Steve (1996) "Changes in the relative demand for skills." in Alison L. Booth and Dennis J. Snower (eds) *Acquiring Skills*, Cambridge: Cambridge University Press.

Murnane, Richard J., Willet, John B. and Levy, Frank (1995) "The growing importance of cognitive skills in wage determination." *Review of Economics and Statistics*, 77, 251-266.

Swanson, Norman and White, Halbert (1997) "Forecasting economic time series using flexible versus fixed specifications and linear versus nonlinear econometric models." *International Journal of Forecasting*, 13, 439-462.

White, Halbert (1989b) "An additional hidden unit test for neglected nonlinearity in multilayer feedforward networks." in Proceedings of the international joint conference on neural networks, New York: IEEE Press.

White, Halbert (1992) "*Artificial neural networks: approximation and learning theory.*" Oxford: Blackwell.

Table 1 The impact of curriculum on expected earnings: selected subject combinations

Curriculum			Predicted wage (£/hr)
History	Art	Physics	7.13
History	Maths	Physics	7.61
Biology	Chemistry	Physics	6.53
Chemistry	Maths	Physics	6.53
French	English	History	6.63
Chemistry	Social Science	Statistics	6.98
French	Art	History	6.53
French	Academic Art	History	7.18

Note: the three most commonly studied combinations of subjects are: biology, chemistry and physics; mathematics, chemistry and physics; social science, academic art and history. The predicted wages are based on mean values of a vector of control variables, and are measured in 1991 UK£.
Source: Johnes (2005).

Table 2 Descriptive statistics

Variable	Mean	Standard Deviation
ln hourly wage	2.06	0.55
years of schooling	12.55	2.77
years of experience	19.90	10.74
tenure (months)	94.70	97.20
union membership	0.32	0.47
male	0.52	0.50
white	0.94	0.23
currently married	0.56	0.50
number of dependent children	0.78	1.04

Table 3 Regression results

variable	coefficient	variable	coefficient
constant	0.9558 (15.35)	physical	-0.1871 (27.36)
schooling	0.0380 (13.59)	strategy & HR	0.1013 (15.13)
experience	0.0223 (8.82)	IT	0.0876 (13.11)
experience squared	-0.0004 (6.75)	quantitative	0.0577 (8.56)
tenure	0.0006 (3.19)	teamwork	0.0300 (4.51)
tenure squared	-0.338×10^{-6} (0.62)	oral communication	0.0238 (3.58)
union member	0.0602 (3.97)	self-motivation	0.0434 (6.69)
male	0.1551 (7.40)	written communication	-0.0405 (6.20)
white	0.1333 (4.59)	advanced analytical	0.0537 (8.22)
married	0.0541 (2.73)	problem-solving	-0.0252 (3.91)
married x female	-0.0567 (2.11)		
dependent children	0.0060 (0.67)	regional dummies (11)	yes
dependent children x female	0.0014 (0.11)	observations R squared log-likelihood	4052 0.4617 -2082.785

Note: t statistics in parentheses

Table 4 Neural network results

variable	coefficient	variable	coefficient
constant	0.5966 (9.42)	ρ	-4.0183 (9.27)
		θ associated with:	
schooling	0.0463 (17.09)	physical	2.1916 (10.81)
experience	0.0245 (9.62)	strategy & HR	-1.4871 (8.73)
experience squared	-0.0004 (7.75)	IT	-1.3485 (9.70)
tenure	0.0009 (4.31)	quantitative	-0.6087 (5.24)
tenure squared	-0.712 x 10-6 (1.30)	teamwork	-0.4328 (3.70)
union member	0.0576 (3.82)	oral communication	-0.3772 (3.43)
male	0.1619 (7.72)	self-motivation	-0.5004 (3.97)
white	0.1465 (4.99)	written communication	0.5421 (4.64)
married	0.0678 (3.38)	advanced analytical	-0.7537 (5.77)
married x female	-0.0660 (2.42)	problem-solving	0.4031 (3.99)
dependent children	0.0055 (0.61)	regional dummies (11)	yes
dependent children x female	-0.0040 (0.30)	observations R squared log-likelihood	4052 0.4432 -2151.531

Table 5 Examples of the neural network results

skill mix			predicted wage (£/hr)
physical	written communication	problem-solving	8.22
strategy & HR	IT	advanced analytical	12.25
physical	quantitative	written communication	8.24
physical	oral communication	written communication	8.23
teamwork	quantitative	written communication	8.61
teamwork	oral communication	written communication	8.51

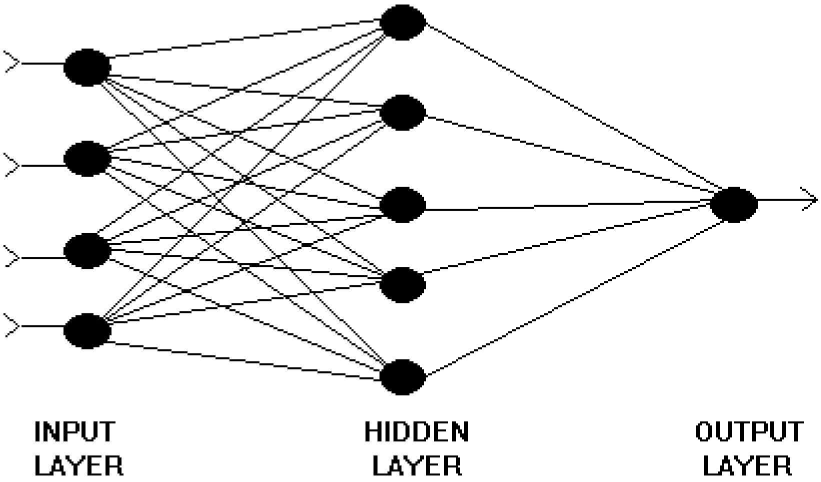

Figure 1: Diagrammatic representation of a feedforward neural network