Multilayer Perceptron Neural Network Analysis: Prediction of School Outcome from the dimensions of School Process

Dr. M. Balamurugan
Assistant Professor in Education,
Government College of Education, Pudukkottai, Tamil Nadu, India 622001
Email - drbalamurugangcoe@gmail.com

Abstract
Quality education puts students at the center of the process; student achievement must be the school's first priority. Objective: To identify neural patterns of school related variables and research variables viz dimensions of school process, for the predictive classification of three levels of school outcome. Method and sample: School outcome and school process was assessed using ipsative method of research, 564 higher secondary teachers and 1189 higher secondary students served as the respondents to predict the outcome of 67 schools. Tools used: Nine tools namely, management and organisation scale, teaching and learning scale, student support and school ethos scale, self-concept scale, student values scale, student relationship scale, attitude towards learning scale, attitude towards school and studies scale student ambiance scale developed by the investigator and his supervisor are used. The first three tools responded by the teacher and rest of the six tools responded by the students. Findings: The prediction models are found to be highly significant and multilayer perceptron neural network analysis showed that the classroom interaction and student learning has the greatest impact on school outcome and these two sub-dimensions belong to 'teaching and learning' under the school process.

Key words
Area under Curve (AUC), Cumulative Gain Chart, Gini Coefficient, Lift chart, Multilayer Perceptron, Neural Network, Neural Network Analysis, Neural Network Architecture, ROC Curve, School Outcome, School Process

I. Introduction

Quality education puts students at the center and student achievement must be the school's first priority. The education in the 21st century presents challenges to quality assurance that were unimaginable just a quarter century ago. First, we must agree upon a set of "universal" attributes or standards of a quality educational experience—not the means to achieving the standards, but the standards themselves. Moreover, these standards must be applied independently of educational delivery method. Second, we must agree to evaluate educational programs and institutions in the context of the student experience, not the institutional experience. Schools should have standards, but the standards should not be standardized.

II. Theoretical framework

Lee, Rey, Mentele, & Garver, (2005) used an approach based on the concept of structured neural network in order to understand the patterns of variables predicting educational outcomes in higher education. Scott et al. (2004) have found that high school GPA was a very important predictor of academic achievement. Most of the predictive models developed for those key outcomes have been based on traditional methodological approaches. However, these models assume linear relationships between variables and do not always yield accurate predictive classifications. On the other hand, the use of machine-learning approaches such as artificial neural networks has been very effective in the classification of various educational outcomes, overcoming the limitations of traditional methodological approaches. The multilayer perceptron artificial neural network models, with a back propagation algorithm, were developed to classify levels of grade point average, academic retention, and degree completion outcomes. Among the predictors, learning strategies had the greatest contribution for the prediction of grade point average (Musso, Hernández & Cascallar, 2020).

2.1 School Process

In recent years, however, more attention has been paid to educational processes - how teachers and administrators use inputs to frame meaningful learning experiences for students. One of the major indicators of school quality is the school process, which can be well represented by Management & Organisation, Teaching & Learning and Student Support & School Ethos, hence these three major research variable with their dimensions selected for the study is given in Figure 1.

2.2 School Outcome

Outcome of the school can be stated in simple terms as outputs of students' cognitive achievement, completion rates, entrance to next level of education, certification, individual skills etc. Outcomes, on the other hand, refer to long term consequences of education such as employment, earnings, and changes in social attitudes / behaviour (Adams, 1997). Student Performances on the other hand is the outcome of the educational institutions in terms of student achievement in scoring marks in the examinations, their moral values, personality, positive attitude etc. to face the global competitive world with utmost confidence. In this study School Outcome consisting of Co-Scholastic and Scholastic Performances of the student is considered and given in the Figure 2.

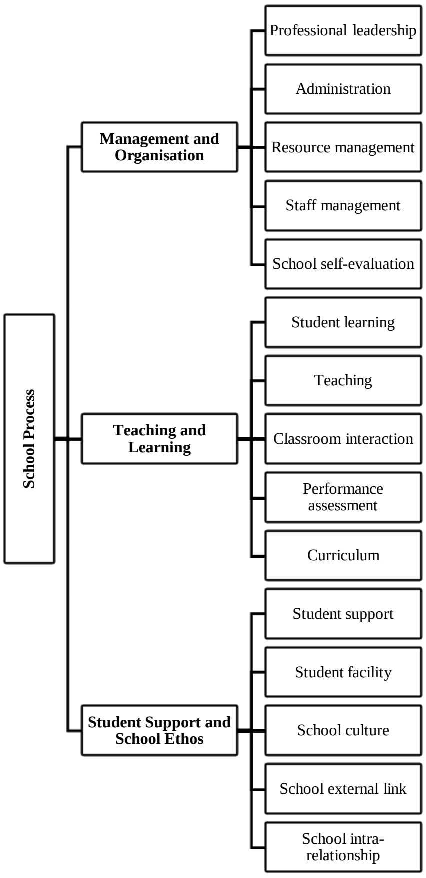

Figure. 1. Variables of school process

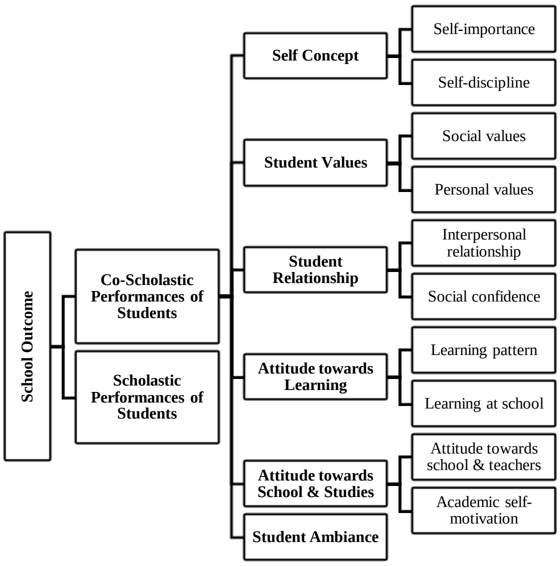

Figure. 2. Variables of school outcome

III. Method of the study

3.1 Objective of the study

To identify neural patterns of school related variables (District, Types of School, Boards of Affiliation, School Managements and Year of School Establishment) and research variables viz dimensions of School Process for the predictive classification of three levels of School Outcome.

3.2 Ipsative Method used

Ipsative method is employed in the present study, in which an attribute of an individual (institution) is measured relatively to his/her (its stakeholders) scores on other attributes (Cattell, 1944). Ipsative method is useful when a researcher wants to collect data on phenomena that cannot be directly observed. The study intends to collect data pertaining to the selected dimensions of school process through teachers, selected dimensions of the school outcome through students and establish relationship between these for the particular school.

3.3 Sample and sampling technique

To predict the Outcome of 67 schools (37 Chennai district and 30 Thiruvallur district) have been conveniently selected keeping in mind the Types of school, Managements of school, Boards of affiliation etc. In each 67 school minimum of five teacher and minimum of ten students were randomly chosen as the respondents. Thus totally 564 higher secondary teachers and 1189 higher secondary students served as the respondents to assess the school process and outcome of 67 schools. The distribution of sample was given in Table 1.

Table 1. Distribution of Sample – Schools

S. No	Variables	Categories and Codes used	Frequency	%
1	District	Chennai = 1	37	55.2
		Thiruvallur = 2	30	44.8
2	School Management	State Government =1	16	23.9
		Corporation = 2	5	7.5
		Private Aided = 3	15	22.4
		Central Government = 4	7	10.4
		Private Unaided = 5	24	35.8
3	Type of School	Boys = 1	12	17.9
		Girls = 2	9	13.4
		Co-Education = 3	46	68.7
4	Board of Affiliation	State Board =1	30	44.8
		Anglo-Indian = 2	6	9.0
		Matriculation = 3	21	31.3
		CBSE = 4	10	14.9
5	Year of School Establishment	Below 30 years = 1	20	29.9
		30 - 50 years = 2	28	41.8
		Above 50 years = 3	19	28.4
Total School			67	100

3.4 Tools used in the study

In the present study, nine tools namely, Management and Organisation Scale, Teaching and Learning Scale, Student Support and School Ethos Scale, Self-concept Scale, Student Values Scale, Student Relationship Scale, Attitude towards Learning Scale, Attitude towards School and Studies Scale Student Ambiance Scale developed by the Investigator and his Supervisor were used. The first three tools responded by teachers and rest of the six tools responded by students. The dimensions of these scales is presented in Figure 1 and Figure 2.

IV. Multilayer perceptron neural network analysis

Basically, a neural network is a computational structure consisting of several highly interconnected computational elements, known as neurons, perceptrons, or nodes. Each “neuron” or unit carries out a very simple operation on its inputs and transfers the output to a subsequent node or nodes in the network topology (Specht, 1991). In general, an ANN consists of an input layer (which can be considered the independent variables), one or more hidden layers, and an output layer that is comparable to a categorical dependent variable (Cascallar, Boekaerts, & Costigan, 2006; Garson, 1998).

Artificial Neural Networks (ANN) are data structures that are capable of organising series of layers. In ANN analysis the input layer receives the value of the variables that can indicate school process in higher secondary schools, the hidden layer performs the mathematic operations to obtain the appropriate response that is shown by the three groups of School Outcome namely, low, average and high group. The neuronal output is predicted through the set of entries, its synaptic weights, activation function and the aggregation function.

Table 2. Case Processing Summary

		N	Percent
Sample	Training	50	74.6%
	Testing	17	25.4%
Valid		67	100.0%
Excluded		0	
Total		67	

The Table 2 revealed that in order to create a multilayer perceptron 50 (74.6 % of the cases) schools randomly assigned to the training and remaining 17 (25.4 % of the cases) schools used for testing the neural network. The five school related variables namely District, School Managements, Types of School, Boards of Affiliation and Year of School Establishment were loaded as factors of input layers as these are having categorical data. The research variables Professional Leadership, Administration, Resource Management, Staff Management, School Self Evaluation, Student Learning, Teaching, Classroom Interaction, Performance Assessment, Curriculum, Student Support, Student Facility, School Culture, School External Link and School Intra-relationship having scale data were loaded as covariates in the input layer. The dependent box was loaded with the School Outcome which contains three groups namely low, average and high. The three groups were classified using the procedure M $\pm$ 1σ, the School Outcome scores $\geq$ M $\pm$ 1σ were grouped under High Group and that of schools with School Outcome $\leq$ M $\pm$ 1σ were grouped under Low Group. The schools with School Outcome scores in between these two limits were grouped under Average Group. Each of these were connected to hidden layer using synaptic weights. In addition to the 32 input layers the bias which also added to the hidden layer. The activation function for the hidden layer was hyperbolic tangent and the same for output layer was chosen Softmax.

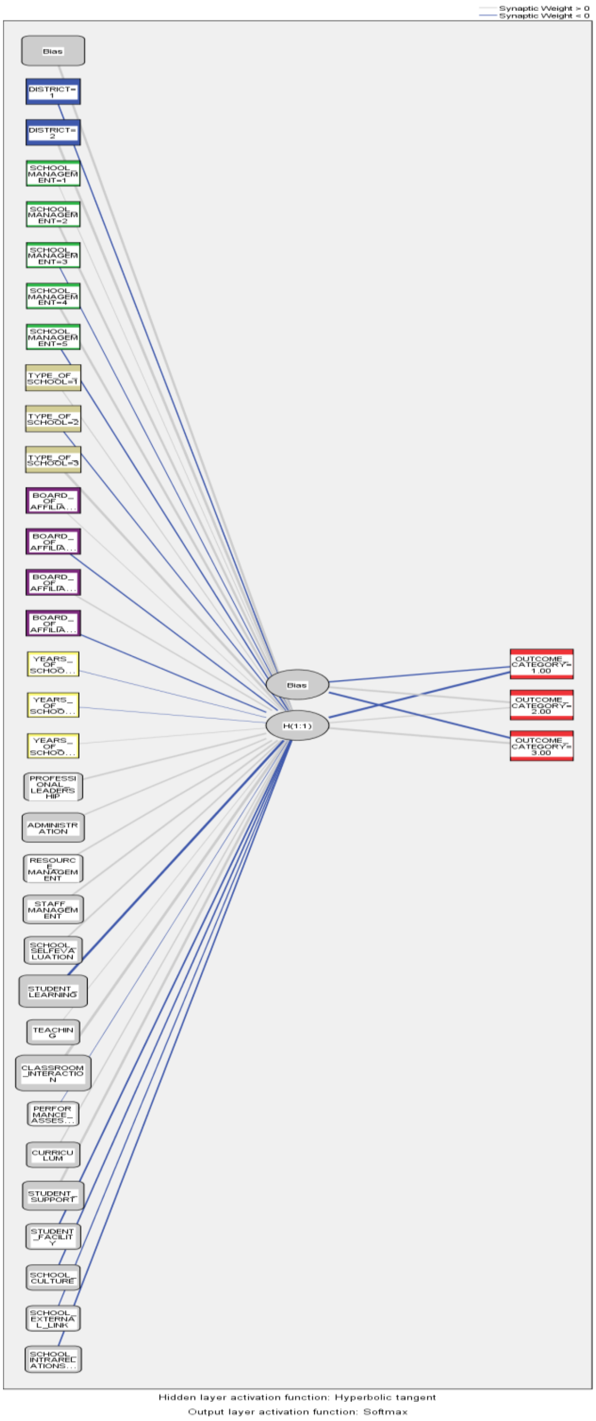

Figure 3. Neural Network Architecture

Table 3. Parameter Estimates

Predictor		Predicted			
		Hidden Layer 1	Output Layer		
		H(1:1)	[OUTCOME_ CATEGORY= 1.00]	[OUTCOME_ CATEGORY =2.00]	[OUTCOME_ CATEGORY= 3.00]
Input Layer	(Bias)	.657			
	[DISTRICT=1]	-.258			
	[DISTRICT=2]	.737			
	[SCHOOL_MANAGEMENT=1]	.100			
	[SCHOOL_MANAGEMENT=2]	.599			
	[SCHOOL_MANAGEMENT=3]	-.153			
	[SCHOOL_MANAGEMENT=4]	.662			
	[SCHOOL_MANAGEMENT=5]	-.272			
	[TYPES_OF_SCHOOL=1]	.157			
	[TYPES_OF_SCHOOL=2]	-.200			
	[TYPES_OF_SCHOOL=3]	.921			
	[BOARDS_OF_AFFILIATION=1]	.144			
	[BOARDS_OF_AFFILIATION=2]	-.342			
	[BOARDS_OF_AFFILIATION=3]	.495			
	[BOARDS_OF_AFFILIATION=4]	-.409			
	[YEAR_OF_SCHOOL_EST=1]	-.082			
	[YEAR_OF_SCHOOL_EST=2]	-.061			
	[YEAR_OF_SCHOOL_EST=3]	.187			
	PROFESSIONAL_LEADERSHIP	.677			
	ADMINISTRATION	.547			
	RESOURCE_MANAGEMENT	.575			
	STAFF_MANAGEMENT	.585			
	SCHOOL_SELF_EVALUATION	.520			
	STUDENT_LEARNING	-.898			
	TEACHING	.123			
	CLASSROOM_INTERACTION	1.220			
	PERFORMANCE_ASSESSMENT	-.060			
	CURRICULUM	.207			
	STUDENT_SUPPORT	.705			
	STUDENT_FACILITY	-.366			
	SCHOOL_CULTURE	-.290			
	SCHOOL_EXTERNAL_LINK	-.197			
	SCHOOL_INTRA_RELATIONSHIP	-.332			
Hidden Layer 1	(Bias)		-.585	1.202	-.915
	H(1:1)		-2.811	.539	1.713

Synaptic weights for each network path was shown the Table 3. As evident from the neural network architecture as it contain one hidden layer and that too with only one hidden node H (1:1) along with the bias. The highest synaptic weight was found to be 1.220 for Classroom Interaction and hidden node H (1:1). Again this one node of hidden layer was connected to three output layers. Similarly the highest synaptic weight was found to be -2.811 for hidden node H (1:1) and output layer 1 (Low School Outcome).

Table 4. Classification Results of the Sample

Sample	Observed	No of Cases	Predicted			
			Low Group	Average Group	High Group	Percent Correct
Training	Low Group	7	7	0	0	100.00%
	Average Group	33	1	32	0	97.00%
	High Group	10	0	10	0	0.00%
	Overall Percent	(50)	16.00%	84.00%	0.00%	78.00%
Testing	Low Group	3	3	0	0	100.00%
	Average Group	14	1	13	0	92.90%
	High Group	0	0	0	0	0.00%
	Overall Percent	(17)	23.50%	76.50%	0.00%	94.10%

Of the 50 cases 39 cases of training have been correctly classified (78.00 %). All 7 cases (Low Group) have been correctly classified. One out of 33 cases (Average Group) have been wrongly classified (Negatively classified). All the 10 cases (High Group) have been wrongly classified (Negatively classified). Of the 17 cases 16 cases of testing have been correctly classified (94.10 %). All 3 cases (Low Group) have been correctly classified. One out of 13 cases (Average Group) have been wrongly classified (Negatively classified). There was no case found in the High Group for testing. The classification results have been found to be statistically significant.

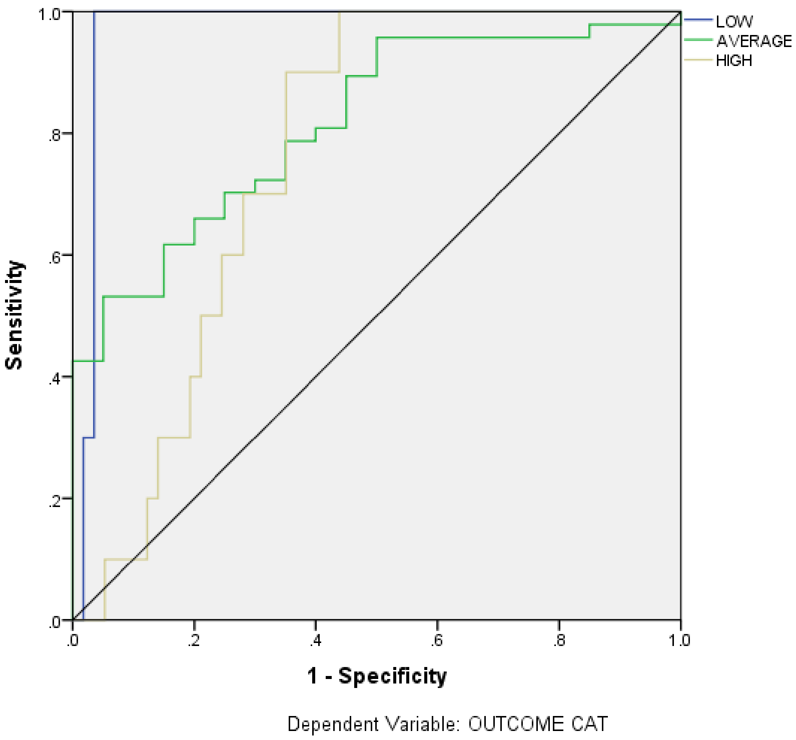

Figure 4. Receiver Operating Characteristic (ROC) Curve

Table 5. Area under Curve and Gini coefficient

Outcome Category	Area Under Curve (AUC)	Gini coefficient = 2*AUC – 1
Low	0.97	0.94
Average	0.816	0.632
High	0.761	0.522

The ROC Figure 4 and the Table 5 shows that the model predictions were good and further the Low category displays excellent model of prediction.

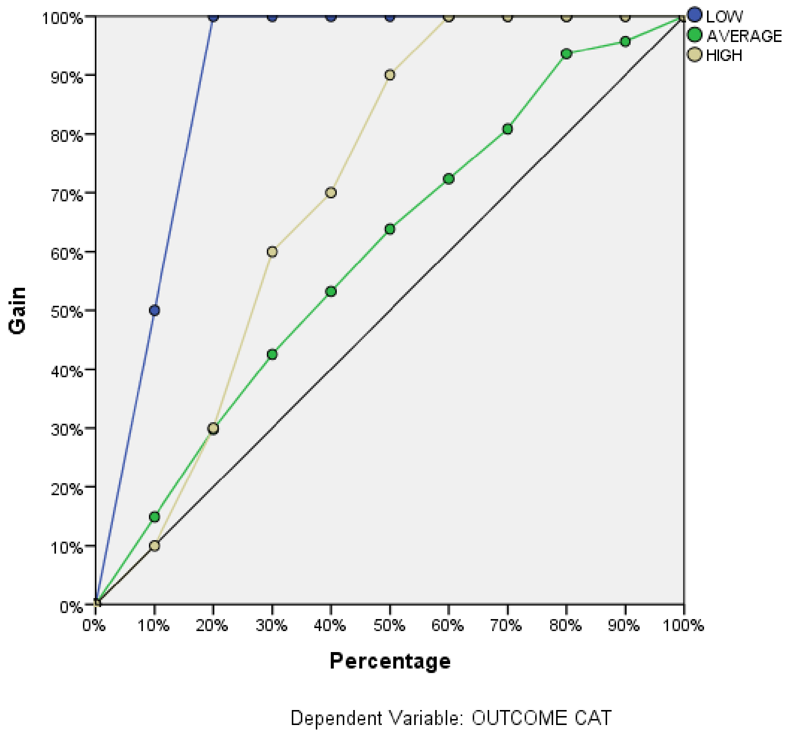

Figure 5. Cumulative Gain Chart

Figure 5 displays the cumulative gain for three categories of School Outcome and a baseline. Cumulative gains and lift charts are graphical supports for measuring the model performance. The cumulative gains chart displays the percentage of the overall number of cases in a given category "gained" by targeting a percentage of the total number of cases. For example, the first point on the curve for the Low category is at (x, y = 10%, 50%), which means that if we score a dataset with the network and sort all of the cases by predicted pseudo-probability of Low, we would expect that the top 10% to contain approximately 50% of all of the cases that actually take the category Low.

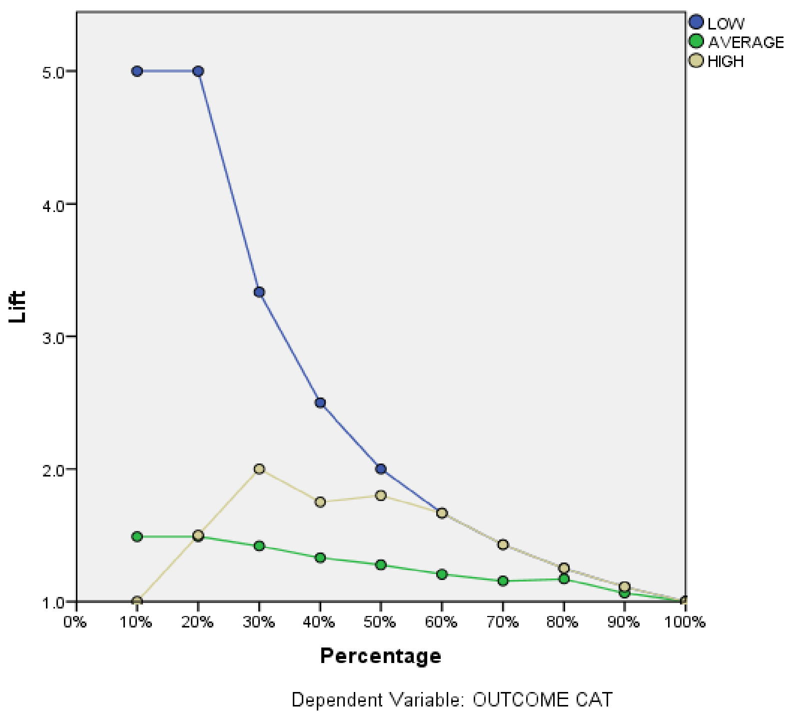

Figure 6. Lift Chart

Lift is a measure of the effectiveness of a predictive model calculated as the ratio between the results obtained with and without the predictive model. Figure 6 displays a lift chart for three categories of School Outcome. This chart is derived from the cumulative gains chart and the values on the y axis correspond to the ratio of the cumulative gain for each curve to the baseline. Thus, the lift at 10% for the category Low is approx. 50% / 10% = 5. Lift chart has been based on the combined training and testing samples.

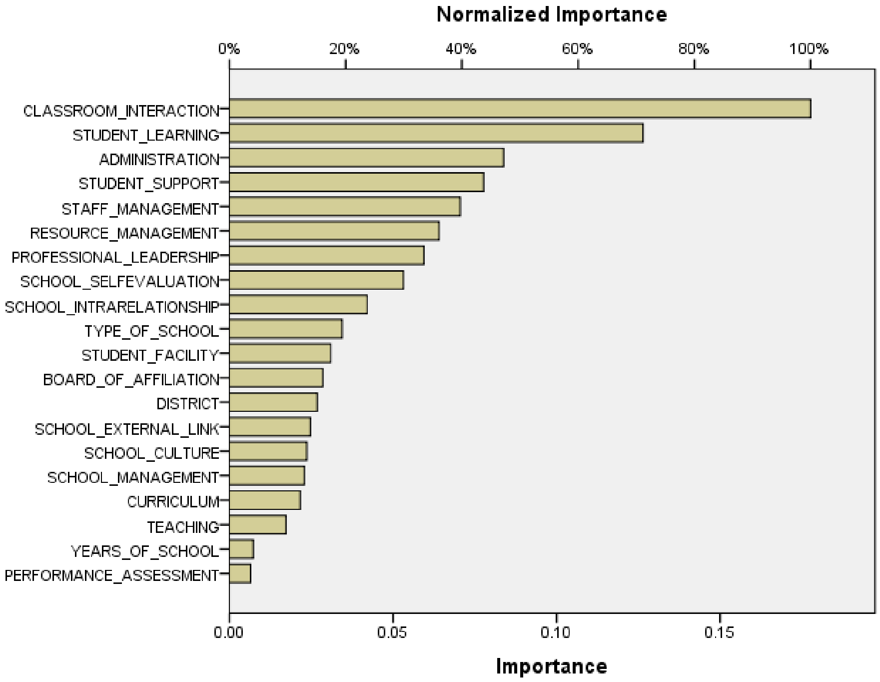

Figure 7. Normalized Importance of the Predictive Variables on School Outcome

Table 6. Independent Variable Importance

Variables	Importance	Normalized Importance
Classroom Interaction	.178	100.0 %
Student Learning	.126	71.2 %
Administration	.084	47.1 %
Student Support	.078	43.7 %
Staff Management	.071	39.7 %
Resource Management	.064	36.1 %
Professional Leadership	.059	33.5 %
School Self Evaluation	.053	29.9 %
School Intra-relationship	.042	23.7 %
Types of School	.034	19.4 %
Student Facility	.031	17.4 %
Boards of Affiliation	.029	16.1 %
District	.027	15.1 %
School External Link	.025	13.9 %
School Culture	.024	13.3 %
School Managements	.023	12.9 %
Curriculum	.022	12.2 %
Teaching	.017	9.8 %
Year of School Establishment	.007	4.1 %
Performance Assessment	.007	3.7 %

The importance of an independent variable is a measure of how much the predictive value of the network model changes for different values of the independent variable. It was given in the Table 6 and Figure. 7. The Classroom

Interaction and Student Learning has the greatest impact on School Outcome as their normalized importance were above 50%. Both these dimensions belong to Overall Teaching and Learning. These two were followed by the variable Administration, which is the dimension of Overall Management and Organisation.

V. Finding of the study

Multilayer Perceptron Neural Network Analysis made with 20 variables School related variables and Research variables viz dimensions of School Process as input with three categories of School Outcome as output, yielded the neural network architecture contain one hidden layer and that too with only one hidden node H (1:1) along with the bias. The prediction models are highly significant from the fact that out of 50 cases 39 cases of training have been correctly classified (78.00 %) and out of 17 cases 16 cases of testing have been correctly classified (94.10 %). Moreover the Classroom Interaction and the Student Learning have the greatest impact on School Outcome.

VI. Conclusion

The higher secondary school stage is the crucial in the life of the any individual. After this stage individual acquire confidence to face the global competition with the skills especially developed during these days. Life skills, laboratory skills, critical thinking, problem solving, decision making are name a few skills, which will be of great help in the profession that any individual going to choose in the near future. Finally, only scholastic abilities are just not enough in the global scenario. It must be accompanied with that of co-curricular or co-scholastic skills. Thus higher secondary schools must focus on the scholastics as well as co-scholastics in order to make their benchmark in the school outcome in turn leads to school quality.

References

Adams, D. (1997). Defining Educational Quality. *Educational Planning*, *9*(3), 3-18

Cascallar, E. C., Boekaerts, M., & Costigan, T. E. (2006) Assessment in the Evaluation of Self- Regulation as a Process, *Educational Psychology Review*, *18*(3), 297-306.

Cattell, R. B. (1944). Psychological measurement: Normative, ipsative, interactive. *Psychological Review*, *51*(5), 292-303.

Darling-Hammond, L., & Friedlaender, D. (2008). Creating Excellent and Equitable Schools. *Educational Leadership*, *65*(8), 14-21.

De Grauwe, A. (2005). Improving the Quality of Education through School-Based Management: Learning from International Experiences. *International Review of Education*, *51*(4), 269-287.

Delors, J., Al Mufti, I., Amagi, I., Carneiro, R., Chung, F., Geremek, B., Gorham, W., Kornhauser, A., Manley, M., Padrón, Q. M., Savané, M. A., Singh, K., Stavenhagen, R., Myong, W. S., & Zhou, N. (1996). *Learning: The Treasure Within: Report to UNESCO of the International Commission on Education for the Twenty-first Century*. Paris: UNESCO. Retrieved from www.unesco.org/delors/

Fitz-Gibbon, C.T., & Kochan, S. (2000). School Effectiveness and Education Indicators. In C. Teddlie and D. Reynolds (Eds). *The International Handbook of School Effectiveness Research*. (pp. 257-282). London: Falmer Press.

Garson, G. D. (1998). *Neural Networks. An Introductory Guide for Social Scientists*. London: Sage Publications Ltd.

Hanushek, E. A. (1997). Incentives Are Key to Improved Schools. *Forum for Applied Research and Public Policy*, 62-67.

Mahapatra, S.S., & Khan, M.S. (2007). A neural network approach for assessing quality in technical education: an empirical study, *Int. J. Productivity and Quality Management*, *2*(3), 287–306.

Musso, M.F., Hernández, C.F.R. & Cascallar, E.C. Predicting key educational outcomes in academic trajectories: a machine-learning approach. *High Educ* 80, 875–894 (2020).https://doi.org/10.1007/s10734-020-00520-7

Reddy, S. (2007). *School quality: Perspectives from the developed and developing countries.* Azim Premji Foundation. Retrieved from http://www.publicarchives.org/pdf/ConsolidatedSchoolQualityreport.pdf

Specht, D. (1991). A general regression neural network. *IEEE transactions on neural networks*, *2*(6), 568-576.

Srivastava, T. (2016). *7 Important Model Evaluation Error Metrics, Everyone should know.* Retrieved from https://www.analyticsvidhya.com/blog/2016/02/7-important-model-evaluation-error-metrics

Willm, J. D. (1992). *Monitoring School Performance: A Guide for Educators.* Washington, DC: Falmer

About the Author

M. Balamurugan is a Faculty at Government College of Education, Pudukkottai, Tamil Nadu, India. He acquired Ph.D. in Education from University of Madras, India and UGC NET in Education. His experience in the field of academics at various levels of teacher education extends over a period of 15 years. His research areas include Management. Pedagogy, Assessment and Psychology in the field of Education.