



## Efficiency Evaluation of State Transport Undertakings of India using DEA-NN approach

Dr. Punita Saxena (corresponding author)

Associate Professor, Department of Mathematics,  
Shaheed Rajguru College of Applied Sciences for Women  
(University of Delhi)

[punita.saxena@rajguru.du.ac.in](mailto:punita.saxena@rajguru.du.ac.in)

Dr. Amita Kapoor

Associate Professor, Department of Electronics,  
Shaheed Rajguru College of Applied Sciences for Women  
(University of Delhi)

### Abstract

The economy of any nation depends on the structure and functioning of its various sectors. Transport sector is one of the vital sectors for the financial system of any developing country. All other sectors are dependent on it either directly or indirectly. Thus improving the efficiency of this sector has become a major concern for the operators and the policy makers. The present paper presents an amalgamation of the two non-parametric techniques, Data Envelopment Analysis (DEA) and Neural Networks (NNs) to compute the efficiency scores of State Transport Undertakings of India. DEA is used to compute the efficiency scores of 27 DMUs. These scores are used to train a neural network model, namely the BPN model. The algorithm is developed and used for predicting the efficiency scores of other units of the data set. The results obtained are comparable and it has been shown that this approach helps in improving the discriminatory power of DEA.

---

# Council for Innovative Research

Peer Review Research Publishing System

**Journal:** International Journal Of Management & Information Technology

Vol . 10, No 5

[editorsijmit@gmail.com](mailto:editorsijmit@gmail.com)

[www.ijmit.com](http://www.ijmit.com)



## Introduction

An efficient transport system forms a backbone for any growing economy. The policy makers have realised this fact way back and are always on the lookout to make this backbone stronger. The managers and the operators of Public Transport system are always looking for solutions to improve the efficiency of their system. A lot has been done in this direction but yet a lot more needs to be achieved.

The major challenges that a developing nation including India faces in this area is bad management and a worse implementation of the policies. India has spent a lot in developing a good infrastructure in terms of roads, changing the technology of public buses, shifting to a sustainable source of energy and a lot more. In spite of investing so much in infrastructure development, the output achieved is much less in terms of the revenue generated from this sector. As a result, the operators end up with dissatisfied customers, overcrowded buses, chaos on the roads and a complete waste of fuel and energy. The focus, therefore, now has shifted to managerial solutions rather the technological ones.

The problem of inefficient public transport leads to many related problems such as increased use of personalised transport system, traffic congestion and extravagant use of energy resources. The policymakers are open to suggestions and are looking for alternatives to improve the services. Privatisation of the sector was looked upon as an alternative solution to this problem. It proved out to be good in the beginning, but then gradually it is proving to be a failure. The objective of the private operator is profit and not efficient service. Increasing the fleet size was considered as another alternative, but then it proved out to be an extra burden on the government.

Researchers have been working to devise methods that can improve the efficiency of Public Transport Services. Econometric methods using production function to forecast interstate bus transportation demand has been discussed by Gonclaves et.al [10], in Italy by Cambini and Filippini [3] and in India by Ramanathan [22]. Due to the technical problems faced by researchers in defining the productivity function, a non-parametric approach called as Data Envelopment Analysis has also been used to investigate the productivity of diverse transportation systems. A number of applications of DEA in transportation systems is available in the literature. Karlaftis [14] uses DEA as part of a methodology to assess the efficiency of transit operators. Pina and Torres [21] used DEA to compare the efficiency of public and private transit operators. Odeck and Alkadi [19] studied the efficiency in the Norwegian bus industry using DEA, Ji Han etal [13], have used DEA to evaluate the efficiencies of China's Public transport systems.

However, DEA has its own limitations and drawbacks namely-1) lack of discrimination among efficient decision making units.

2) unfitness of the weighing scheme that can be unreal at times.

3) multiple optimal solutions for the weighing scheme of extreme efficient Decision Making Units.

A lot of research has been done to improve the discriminatory power of DEA. Applying weight restrictions [12,23,27], using cone ratio model [4], introducing Assurance Region, finding super efficiencies and cross-efficiencies [1,7], using multiple objective approach [4] are few such techniques. These methods provide a framework for analysis and the decision maker has to select the suitability of various methods.

However, DEA technique cannot be used to predict the performance of other DMUs. Recently, Artificial Neural Networks (ANNs) were introduced as alternatives to estimate efficiencies of DMUs [29]. This is another non-parametric technique that can be used for efficiency evaluations. A neural network (NN) is a massively parallel distributed processor made up of simple non-linear processing units called artificial neurons [11]. Inspired by biological brain, they are well known for their ability to generalize and learn from the environment. They have been explored in variety of applications like speech processing [18] face recognition [25], computer vision [11,24], prediction [11,15], function approximation [11], pattern classification [11] and forecasting [11,16]. The ability of NNs to reproduce the unknown relationship between a set of input variables of a system and its output variables makes NN a preferred choice in a variety of intelligent tasks.

Though DEA seeks a set of weights to maximize the technical efficiency of the decision making units, ANN on the other hand, seeks the set of weights to derive the best possible fit for the units under study. Further, the weights derived from ANN algorithms can also be used for predicting the efficiencies of the units that have not been included in the data set. Thus, DEA can be used as a technique to screen the data set for the training units and then ANN can be used as a tool to learn a nonlinear forecasting mode. It can then be used for the purpose of forecasting. Athanassopoulos and Curram [2] introduced the idea of combining DEA and Neural networks in the year 1996. It was later used by Costa and Markellos [6] and then by Pendharkar and Rodgers [20] in 2003. The DEA-NN approach has been used to evaluate the efficiencies of Bank branches by Wu et al in 2006 [30].

In this paper an effort is being made to project the usefulness and flexibility of the neural network algorithms in order to evaluate the efficiencies of Public Transport Undertakings of India. The efficiency scores obtained by DEA are computed. A random sample is selected from this data set and it is used as a training set for the ANN. The results obtained by using Neural Network algorithms are then used to revalidate the scores of the data set. A regression analysis between the scores of DEA and ANN show that there is a strong coefficient of determination between the two scores.

The objective of the paper is-

1. Compute the efficiencies of Decision making units (DMUs) under study.
2. To use DEA as a filter to obtain the units for the training set of ANN.
3. Train the ANN by using BPN model.



4. To predict the efficiency scores of the remaining units under study by applying ANN
5. To compare the two sets of efficiency scores and rank the units under study.

## 2. Models used

### 2.1 The basic DEA models

Data Envelopment analysis or DEA as it is commonly called, was put forth by Farrell in 1957 [9] and extended by Charnes, Cooper and Rhodes in 1978 [5]. It was initially used to evaluate and compare the efficiencies of non-profit organizations whose performance cannot be measured on the basis of profits.

The frequently used models of DEA are the CCR (Charnes, Cooper and Rhodes) and BCC (Banker, Charnes and Cooper). In the CCR model, the frontier is spanned by the linear combination of the units in the data set. The efficiency scores obtained from this model are known as technical efficiencies (TE). These scores reflect the radial distance from the estimated frontier to the unit under consideration. A score less than unity amounts to inefficiency in that unit. The CCR model is based on the assumption of constant returns to scale (CRS).

In the BCC model, the frontier is spanned by the convex hull of the units in the data set. The frontier in this model thus have piece-wise linear and concave characteristics. The efficiency scores of this model are known as pure technical efficiencies (PTE). It is based on the variable returns to scale (VRS) assumption.

Mathematically, the CCR model can be described as-

Consider a set of  $n$  units, each operating with  $m$  inputs and  $s$  outputs, let  $y_{rj}$  be the amount of the  $r$ th output from unit  $j$ , and  $x_{ij}$  be the amount of the  $i$ th input to the  $j$ th unit. According to the classical DEA model, the relative efficiency of a target unit  $j_0$  is obtained by maximizing the ratio of the virtual output to the ratio of the virtual input subject to the condition that this ratio is less than unity for all the units of the data set. Thus, the objective is to

$$\max h_{j_0}(u, v) = \frac{\sum_{r=1}^s u_r y_{rj_0}}{\sum_{i=1}^m v_i x_{ij_0}}$$

$$\frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \leq 1, \quad j = 1, 2, \dots, n$$

subject to

$$\frac{u_{rj_0}}{\sum_{i=1}^m v_i x_{ij}} \geq \varepsilon, \quad r = 1, 2, \dots, s \tag{1}$$

$$\frac{v_{ij_0}}{\sum_{i=1}^m v_i x_{ij}} \geq \varepsilon, \quad i = 1, 2, \dots, m$$

The decision variables  $u = (u_1, \dots, u_r, \dots, u_s)$  and  $v = (v_1, \dots, v_i, \dots, v_m)$  are respectively the weights given to the  $s$  outputs and to the  $m$  inputs. To obtain the relative efficiencies of all the units, the model is solved  $n$  times, for one unit at a time. Model (1) allows for great weight flexibility, as the weights are only restricted by the requirement that they should not be zero (the infinitesimal  $\varepsilon$  ensures that) and they should not make the efficiency of any unit greater than one.

The fractional model (1) is solved as a linear program by setting the denominator in the objective function equal to some constant, say, 1 and then maximizing its numerator, as shown in the following model:



$$\max h_{j_0} = \sum_{r=1}^s u_r y_{rj_0}$$

subject to

(2)

$$\begin{aligned} \sum_{i=1}^m v_i x_{ij_0} &= 1 \\ \sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} &\leq 0, \quad j = 1, \dots, n \\ u_r, v_i &\geq \varepsilon \quad \forall r, i \end{aligned}$$

Thus, the objective is now to maximize the virtual output of the target unit subject to the condition that virtual output cannot exceed virtual input for every other unit. Technical Efficiencies (TE) are obtained from this model.

The BCC model is the dual of the CCR model along with an additional convexity constraint. Mathematically, the BCC model is-

$$\min z_{j_0} = \theta_{j_0} - \varepsilon \sum_{r=1}^s S_{rj_0}^+ - \varepsilon \sum_{i=1}^m S_{ij_0}^-$$

subject to

$$\begin{aligned} \sum_{j=1}^n \lambda_{jj_0} y_{rj} - S_{rj_0}^+ &= y_{rj_0}, \quad r = 1, 2, \dots, s \\ \sum \lambda_{jj_0} x_{ij} + S_{ij_0}^- &= \theta_{j_0} x_{ij_0}, \quad i = 1, 2, \dots, m \\ \lambda_{jj_0} &\geq 0, \quad j = 1, 2, \dots, n \\ \sum_{j=1}^n \lambda_{jj_0} &= 1 \\ \theta_{j_0}, &\text{unrestricted in sign} \\ S_{rj_0}^+, S_{ij_0}^- &\geq 0 \quad r = 1, 2, \dots, s; i = 1, 2, \dots, m \end{aligned} \tag{3}$$

where  $S_{rj_0}^+$  is slack in the  $r^{th}$  output of the target unit,  $S_{ij_0}^-$  is slack in the  $i^{th}$  input of the target unit,  $\lambda_{jj_0}$  are non negative dual variables and  $\theta_{j_0}$  is the reduction applied to all inputs of the target unit to improve efficiency. This reduction is applied simultaneously to all inputs and results in a radial movement towards the envelopment surface.

In the BCC model, the convexity constraint represents the returns to scale. Returns to scale reflects the extent to which a proportional increase in all inputs increases outputs. The efficiency scores thus obtained are called as the Pure Technical Efficiencies (PTE)

## 2.2 Multilayered Perceptron

The perceptron model was first proposed in the year 1958, by Frank Rosenblatt [24]. He demonstrated that using a method of trial and error (supervised learning) a computer can be made to learn logic functions. His perception convergence theorem increased the interest of many researchers in the field of neural networks. But, all the euphoria so created was short lived. The 1969 book by Minsky and Papert [17] demonstrated that there are fundamental limits on what single layer perceptrons can perform. Most specifically they proved mathematically that a simple perceptron can only solve linearly separable problems. They also argued that multilayered perceptrons should also suffer from the same problem.

Rumelhart, Hinton and Williams reported the development of Back propagation algorithm for the training of multilayered perceptron in the year 1986 [26]. They not only showed how to train multilayered perceptron NN, but also demonstrated that MLPs can solve non-linearly separable problems. Since then MLP have been used for function approximation and pattern classification tasks. Even today it is the most popular training algorithm for multilayered perceptrons.

Back propagation algorithm essentially is a least mean square error algorithm, where the weights are adjusted such that the least mean square error reduces. If  $d_i$  denotes the desired output of neuron  $i$  and  $y_i$  denotes the actual output of neuron  $i$  in the output layer, the mean square error (MSE) function is defined as:

$$E = \frac{1}{2N} \sum_{i=1}^N (d_i - y_i)^2 \tag{4}$$

Let  $W_{ij}$  be the weights connecting  $i^{th}$  neuron in layer  $l$  to  $j^{th}$  neuron in  $(l+1)$  layer, then the weights are updated using:

$$\Delta W_{ij}^{new} = \eta \delta^j x_i + \alpha \Delta W_{ij}^{old} \tag{5}$$

where  $x_i$  is the output of  $i^{th}$  neuron in the previous layer,  $\eta$  is the learning rate,  $\alpha$  is the momentum coefficient and  $\delta^j$  represents the propagating error for the neuron  $j$ . For neurons in the output layer  $\delta^j = (d_j - y_j)g'(h_j)$  while for the neurons in the hidden layer,  $\delta^j = \sum_k \delta^k W_{jk}g'(h_j)$ , here  $g'(h)$  represents the derivative of the activation function  $g(h)$ .

In the present problem of efficiency evaluation of the various transports systems, there exists a nonlinear relation between the input output of DMUs and their respective efficiency. Therefore to model such a non-linear relationship a multilayered Neural Network(NN) was employed.

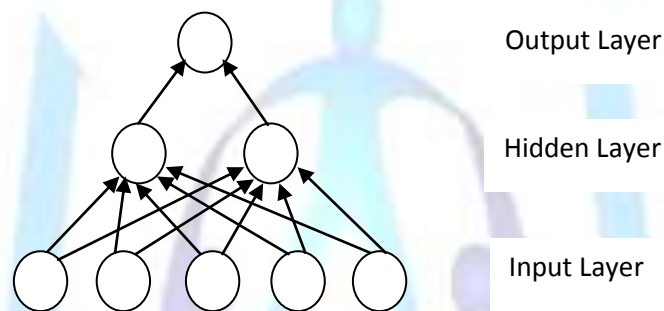


Figure 1 Architecture of the 5-2-1 BPN

In this paper a BPN with structure 5-2-1 is used, it consists of three layers, an input layer with five neurons, a hidden layer with two neurons and an output layer with only one neuron. Neurons of input layer function only as buffer units and neurons of hidden layer and output layer have sigmoid activation function of the form:

$$g(h) = \frac{1}{1 + e^{-ah+b}} \tag{6}$$

where  $h$ , activity is the weighted sum of all inputs,  $a$  and  $b$  are two constants for our network both are set to value 0.5. To train the network, a data set of 27 DMUs was taken. Using the BCC model of DEA, Pure Technical Efficiencies (PTE) were calculated for all the decision making units (DMUs) under study. Randomly 37% of the data set was selected for training the Neural Network. The standard BPN Algorithm was used to update the weights according to the equations. The learned network was then used to predict the efficiency scores of the remaining data set.

### Data and Variables

The State Transport Undertakings in India are run either by government departments or Municipal undertakings or corporations or by private operators. In order to have a homogeneous data set, private operators have been excluded from the study. Also some of the undertakings did not have the complete data, so only 26 Public Transport Undertakings have been considered in this study. In order to have meaningful interpretations of the efficiency scores, an average unit has been constructed by taking the average of all the input and output variables and then studying it as an individual decision making unit along with the rest of the data set. The data has been collected for the year 2009-2010 [8].

Three variables have been identified as the input variables namely, the Fleet Size (FS), Total Staff (TS) and Fuel Consumption (FC). Technically, the inputs should include both capital and labour elements. Fleet size and fuel consumption represent the capital elements and total staff represents the labour element in the study.

Two variables, namely the passenger-kilometers (PK) and seat-kilometers(SK) have been taken as the output variables. The passenger-kilometer measures the annual production of the units under study.

The descriptive statistics of these variables are given in Table-1

Table-1 Descriptive Statistics of the Variables

Variables	Maximum	Minimum	Average	Standard Deviation
Fleet size	21255	62	4045.5	4921.022



<b>Total staff</b>	115898	344	22422.08	27967.58
<b>Fuel consumption</b>	5249.2424	6.153	923.9977	1200.7277
<b>Passenger Kilometers</b>	974778	543	158696.4	132.4242
<b>Seat Kilometers</b>	1417673	913	233374.6	314639

A relationship amongst the input and the output variables was measured. Table-2 shows that the output and the input variables are strongly correlated. Thus, the cause and effect relationship of the variables has not been violated during the period of study.

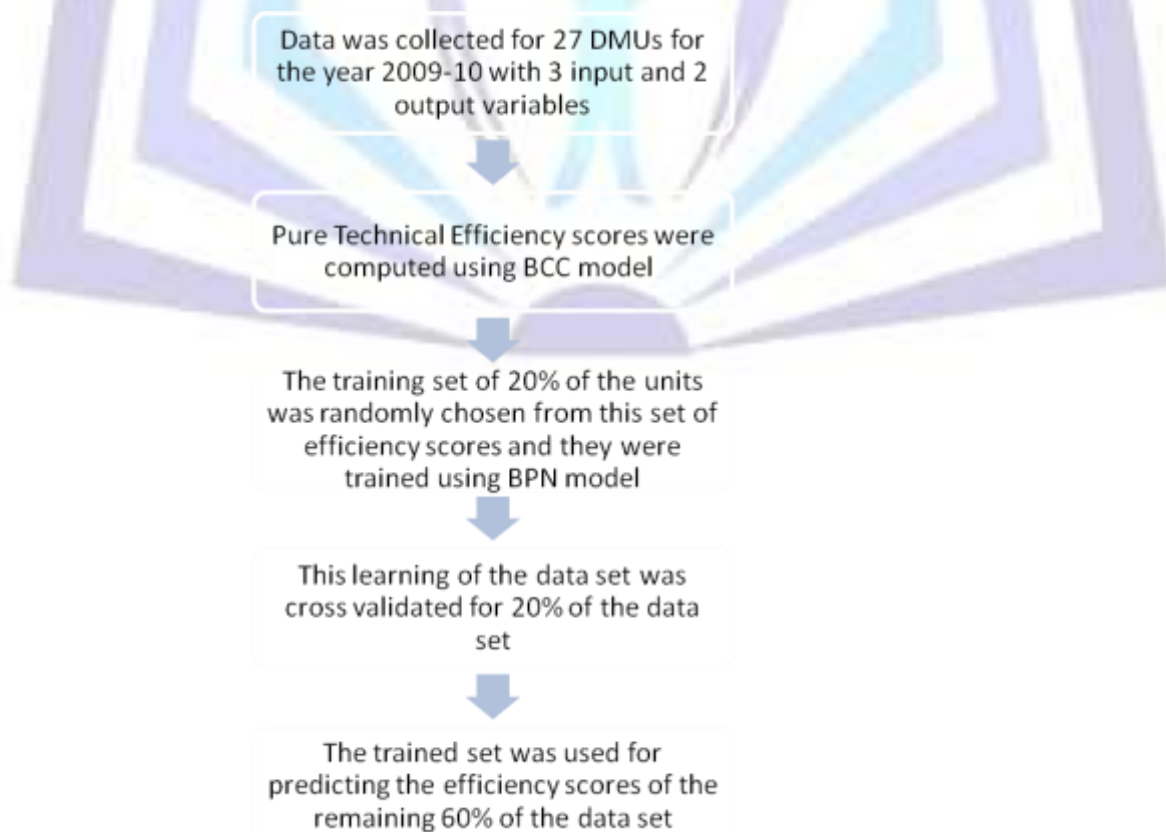
**Table 2: Correlation between the input and the output variables.**

<b>Output</b>	<b>Inputs</b>	<b>Fleet size</b>	<b>Total staff</b>	<b>Fuel consumption</b>
<b>Passenger kilometers</b>		0.9817	0.9512	0.9896
<b>Seat kilometers</b>		0.99	0.9601	0.995

Since, the paper deals with the analysis of efficiency scores between the Public Transport Undertakings, and the two output variables are passenger kilometers and seat kilometers, the output maximizing models of DEA are used for efficiency evaluations. The technical efficiencies (TE) using CCR model, the pure technical efficiencies (PTE) using BCC model, have been evaluated for the data set of 27 units including an average unit.

## Methodology

In this paper DEA-ANN models have been used for evaluating the efficiencies of decision making units (DMUs) under study. To summarise





## Results and Discussions

### 5.1 Technical Efficiencies and Pure Technical Efficiencies

Out of the 27 units under study, only 7 of them were seen to be technically efficient with their PTE scores equal to 1. However, this group of the data set contributes to only 42% of the total annual production. The inefficient units are responsible for the remaining 58% production. Out of these efficient units, two are observed to have their TE scores less than 1. This means that these two units need to improve their scales of operation.

Also, 7 of the units had their PTE score lying in between 0.62 to 0.74. They are the set of the least efficient units. Five of these seven transport undertakings are of the major metropolitan cities of India. Their contribution to the annual production is observed to be only 12%. An efficient public transport system in these cities would solve a lot more related problems as well. The summary statistics of these efficiency scores is given in Table-3.

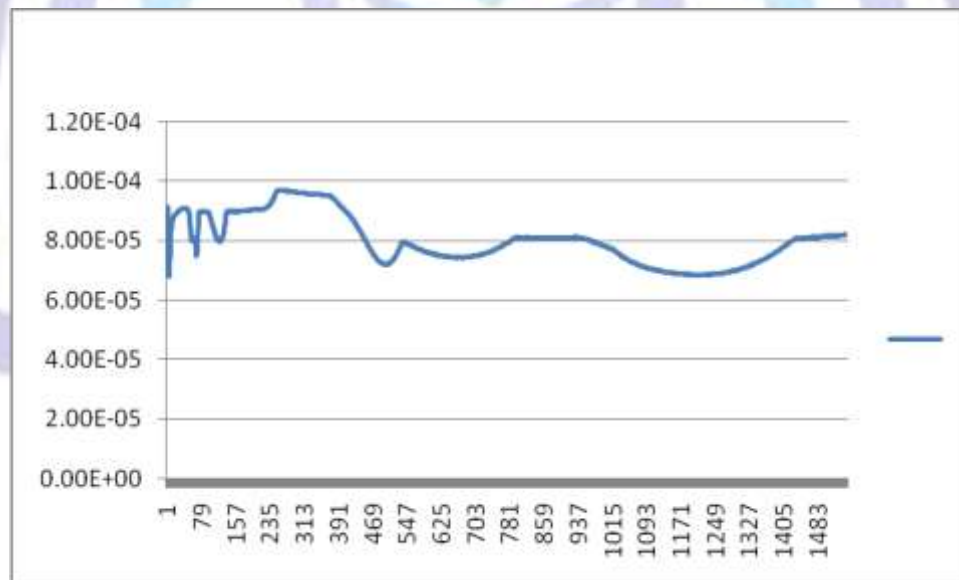
**Table 3: Summary statistics of the TE and PTE scores.**

Scores	No. of efficient units	Minimum	Maximum	Average	Standard deviation	Average Unit
TE	5	0.456	1	0.819	0.15	0.8413
PTE	7	0.62	1	0.873	0.13	0.99

The average unit was observed to be operating on increasing returns to scale. Five of the units were observed to be operating on decreasing returns to scale.

### 5.2 Neural Network efficiency scores

The BPN model of Neural Networks was used to compute the efficiency scores. For the purpose of this paper the efficiency scores for 27 DMUs was computed using DEA. Randomly 10 DMUs were selected to train the Neural Network. The BPN network was trained for this data set with maximum mean square errors (MMSE) of 0.002. The learning rate parameter for the network was chosen to be 0.08 and to avoid the network getting stuck in a local minima a momentum coefficient of 0.03 was chosen. The presence of constant  $a$  in the activation function (4) resulted in an effective learning rate of 0.04. Each unit from the data set was presented one by one in random order for every epoch of training. Figure 2 shows the variation of MSE with the number of epochs, as expected the MSE settles to a minimum value after 47587 epoch for MSE=0.000073. From the figure it can be clearly seen that the network also met few such local minima's before settling to the minimum MSE value. After the training process the network was used to predict the efficiency scores of the remaining DMUs.



**Figure 2 MSE vs. Number of Epochs as the BPN learns training data set 1**

The Table 4 summarises the Neural Network efficiency scores of all the 27 units under study.

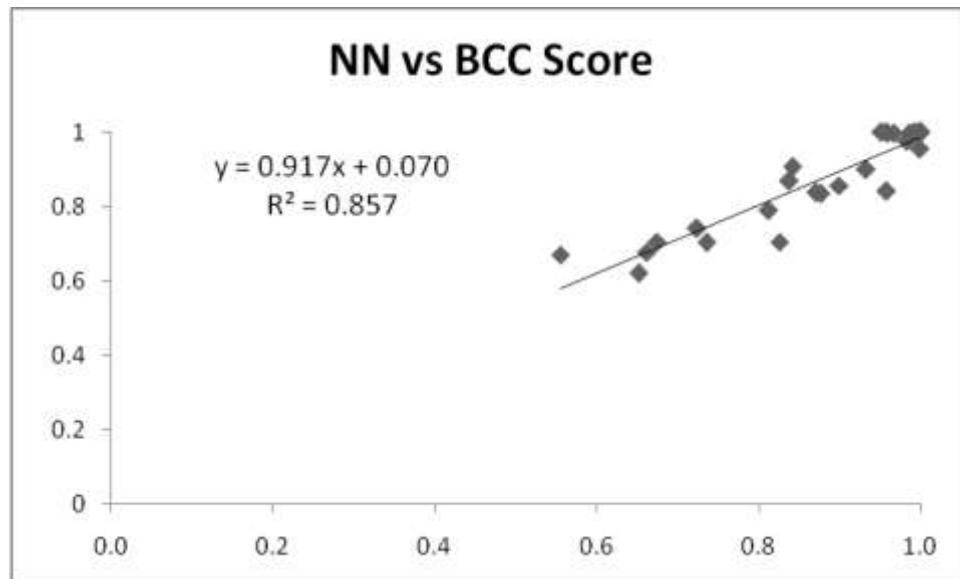


**Table 4: Summary statistics of the Neural Network Efficiency Scores.**

S. No.	Units	CCR Scores	BCC scores	NN score	Error Sq
1	BEST	0.619	0.62	0.587849	0.001034
2	Calcutta	0.639	0.671	0.556202	0.013179
3	N Ben STC	0.649	0.678	0.720394	0.001797
4	Bihar	0.596	0.704	0.815429	0.012416
5	Kerala	0.695	0.705	0.744429	0.001555
6	Pune	0.698	0.706	0.609683	0.009277
7	DTC	0.739	0.741	0.591033	0.02249
8	Himachal	0.781	0.791	0.774122	0.000285
9	S Ben STC	0.76	0.834	0.889282	0.003056
10	Maha	0.745	0.838	0.95196	0.012987
11	Average	0.84	0.8416	0.848584	4.88E-05
12	Kadamba	0.781	0.858	0.908403	0.002541
13	Bangalore	0.853	0.869	0.84312	0.00067
14	N w Kn	0.899	0.901	0.88502	0.000255
15	N E KnRT	0.904	0.907	0.779616	0.016227
16	Karnatka	0.935	0.956	0.977987	0.000483
17	Guj	0.943	0.976	0.95808	0.000321
18	PRTC	0.969	0.997	0.957785	0.001538
19	Chandigarh	0.909	0.998	0.949769	0.002326
20	Kolhapur	0.665	0.999	0.955422	0.001899
21	Andhra	0.946	1	0.956395	0.001901
22	UP	1	1	0.984284	0.000247
23	Raj	1	1	0.996251	1.41E-05
24	SETC(TN)	1	1	0.994096	3.49E-05
25	Orissa	1	1	0.955899	0.001945
26	MEGTC	0.456	1	0.952481	0.002258
27	MTC(CNI)	1	1	0.964285	0.001276
					0.11206
				MSqErr	0.004482

As can be seen from the table above, the units that had the TE and PTE scores as 1 have also been ranked amongst themselves. Thus this approach helps in improving the discriminatory power of DEA. Further, a regression analysis was carried out between the PTE and Neural Network scores. Figure 3 gives the scatter plot and the regression equation between the two scores. The coefficient of determination was observed to be 0.8574 which proves the strong relationship between the two efficiency measures.





**Figure 3 Regression equation between BCC and NN scores**

### Conclusions

In this paper a DEA and NN approach has been used to compute the efficiency scores of Indian bus companies. This approach is more robust than the conventional DEA approach as it clearly gives a proper ranking structure amongst the data sets who turn out to be efficient as per the DEA model. This is an improvement in the discriminatory power of DEA. Further, this approach can also be used for prediction purposes. Though there are a lot of similarities in the approaches but the latter one is more flexible.

Though the study in this paper is extensive yet it is not exhaustive. The number of units in the data set can be increased to train the network more efficiently. Further, the network can also be used to predict the efficiency scores of those similar DMUs that are not a part of the original data set.

### References:

1. Anderson, P., Peterson, N. C. A Procedure for Ranking Efficient Units in DEA, *Management Science*, Vol. 39, No 10, 1993, pp. 1261–1264.
2. Asthanassopoulos, A.D., & Curram, S.P (1996), "A comparison of data envelopment analysis and artificial neural networks as tool for assessing the efficiency of decision making units." *Journal of the Operational research Society*, 47(8), 1000-1016.
3. Cambini C., Filippini M. (2003). "Competitive Tendering and Optimal Size in the Regional Bus Transportation Industry." *Annals of Public and Cooperative Economics*, 74, 163-182.
4. Charnes, A., Cooper, W. W., Wei, Q. L., Huang, Z. M. Cone Ratio DEA and Multi-Objective Programming, *Int. J. of System Science*, Vol. 20, No 7, 1989, pp. 1099–1118.
5. Charnes, A; Cooper, W.W and Rhodes, E.(1978), "Measuring the Efficiency of Decision making Units." *European Journal of Operational Research* 2, 429-444.
6. Costa, A., & Markellos, R.N. (1997), "Evaluating public transport efficiency with neural network models." *Transportation Research C*, 5(5), 301-312.
7. Doyle, J. R., Green, R. H. Efficiency and Cross-Efficiency in DEA: Derivatives Meaning and Uses, *J. of Operational Research Society*, Vol. 45, No 5, 1994, pp. 567–578.
8. Editor, *Indian Journal of Transport Management*, (2010), Vol.34, Number 2.
9. Farrell, M.J.(1957), "The measurement of productive efficiency." *Journal of the Royal Statistical Society*, 12093, 253-281.
10. Gonclaves, M.B., E.T Bez, & A.G.Novaes, 92007), "Econometric models applied to forecast interstate bus transportation demand in Brazil, *Electronic Transactions of the XXVI ANPET- Brazilian Annual Conference on Transportation research and Teaching*, Rio de Janeiro.
11. Haykin, Simon S., et al. *Neural networks and learning machines*. Vol. 3. New York: Prentice Hall, 2009.
12. Hosseinzadeh, L., F., Jahanshahloo, G. and Memariani, A., (2000), "A method for finding common set of weights by multiple objective programming in data envelopment analysis, *Southwest Journal of Pure and Applied Mathematics*, 1, 44-54.
13. Ji Han and Yoshitsugu Hayashi (2008), "A Data Envelopment Analysis for Evaluating the Performance of China's Urban Public Transport Systems," *International Journal of Urban Sciences*, 12(12), pp.173-183.
14. Karlaftis, M.G (2003), "Investigating Transit Production and Performance: A Programming Approach". *Transportation Research A*, 37(3), 225-240.



15. Maier, Holger R., et al. "Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions." *Environmental Modelling & Software* 25.8 (2010): 891-909.
16. Mellit, Adel, and Alessandro Massi Pavan. "A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy." *Solar Energy* 84.5 (2010): 807-821.
17. Minsky M. L. and Papert S. A. 1969. *Perceptrons*. Cambridge, MA: MIT Press.
18. Morgan, David P., and Christopher L. Scofield. *Neural networks and speech processing*. Springer US, 1991.
19. Odeck, J. and Alkadi, A. (2001), "Evaluating efficiency in the Norwegian bus industry using data envelopment analysis." *Transportation*, 28, 211-232.
20. Pendharkar, P.C. & Rodger, J.A. (2003), "Technical-efficiency based selection of learning cases to improve forecasting accuracy of neural networks under monotonicity assumption." *Decision Support Systems*, 36(1), 117-136.
21. Pina V. and Torres L. (2001), "Analysis of the efficiency of local government services delivery. An application to urban public transport," *Transportation Research Part A*, Vol.35, no.10, pp.929-944.
22. Ramanathan, R. (2004), "Indian Transport towards the New Millennium: Performance, Analysis and Policy, New Delhi, Concept Publishing Company.
23. Roll, Y. and Golany, B., (1993), "Alternate methods of treating factor weights in DEA", *Omega*, 21, 99-109.
24. Rosenblatt, Frank (1958), *The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain*, Cornell Aeronautical Laboratory, Psychological Review, v65, No. 6, pp. 386-408
25. Rowley, Henry A., Shumeet Baluja, and Takeo Kanade. "Neural network-based face detection." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 20.1 (1998): 23-38.
26. Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. *Learning representations by back-propagating errors*. MIT Press, Cambridge, MA, USA, 1988.
27. Saati, S., (2008), "Determining a common set of weights in DEA by solving a linear programming", *Journal of Industrial Engineering International*, Vol.4, No.6, 51-56.
28. Schuster, Mike, and John Wiley. "Neural networks for speech processing." *Encyclopedia of Electrical and Electronic Engineering, John Wiley and Sons* (1999).
29. Wang, S. (2003). Adaptive non-parametric efficiency frontier analysis: A neural-network-based model. *Computers and Operations Research*, 30, 279-295.
30. Wu Desheng (Dash), Yang Zijiang, Liang Liang (2006). "Using DEA-neural network approach to evaluate branch efficiency of a large Canadian Bank", *Expert Systems with Applications*, 31, 108-115.