Predicting the Performance of Bowlers in IPL: An Application of Artificial Neural Network

Article *in* International Journal of Performance Analysis in Sport · April 2012
DOI: 10.1080/24748668.2012.11868584

Hemanta Saikia
Assam Agricultural University

Dibyojyoti Bhattacharjee
Assam University

Hoffie Lemmer
University of Johannesburg

International Journal of Performance Analysis in Sport
2012, 12, 75-89.

Predicting the Performance of Bowlers in IPL: An Application of Artificial Neural Network

Hemanta Saikia[1], Dibyojyoti Bhattacharjee[1] and Hermanus H. Lemmer[2]

[1] Department of Business Administration, Jawaharlal Nehru School of Management, Assam University, Silchar-788011, Assam, India.

[2] Department of Statistics, University of Johannesburg, Auckland Park, Republic of South Africa

Abstract

Application of data mining tools is often used in professional sports for evaluating players'/ teams' performance. Cricket is one of those sports where a large amount of numerical information is generated in every game. The game of cricket got a new dimension in April 2008, when Board of Control for Cricket in India (BCCI) initiated the Indian Premier League(IPL). It is a franchise based Twenty20 cricket tournament where teams are formed by competitive bidding from a collection of Indian and International players. Since, valuations of the players are determined through auction, so performance of individual player is always under scanner. The objective of this study is to analyze and predict the performance of bowlers in IPL, using artificial neural network. Based on the performance of bowlers in the first three seasons of IPL, the paper tries to predict the performances of those bowlers who entered in the league in its fourth season as their maiden IPL venture. The performances of these bowlers in IPL-IV are predicted, and the external validity of the model is tested using their actual performance in IPL-IV. This prediction can help the franchises to decide which bowler they should target for their team.

Keywords: Data Mining, Performance Measurement in Sports, Twenty20 Cricket

1. Introduction

Cricket is the sport that is immensely popular in India. The recent format of Twenty20 cricket which was first introduced in 2003 by the England and Wales Cricket Board (ECB), has gained huge recognition in India as well, like several other major cricket playing countries. With India winning the inaugural Twenty20 world cup in 2007 in South Africa, a massive interest in this format of cricket was generated in India.

Soon Subhash Chandra, the promoter of Essel group, started his own private Twenty20 cricket league for Zee TV in 2007 (Malcolm, Gemmell and Mehta, 2009). The name of that tournament was Indian Cricket League (ICL). The league was a six-team competition in its first year (2007) which had expanded to eight in its second year (2008). The tournament drew high profile names to play in its fixtures (Kitchin, 2008). However, the game of cricket got a new dimension in April 2008, when BCCI initiated the Indian Premier League (IPL). It is a Twenty20 cricket tournament being played among eight domestic teams, named after eight Indian cities/states, and owned by franchises (Mitra, 2010). The franchises formed their teams by competitive bidding from a collection of Indian and international players and the best of Indian upcoming talent (Saikia and Bhattacharjee, 2011). Each player has a base price fixed by the IPL authorities. However, there is no upper limit for the bid price. The valuation of players obtained through auction and the availability of players' performances have allowed researchers to infer on different aspects of this format of the game.

1.1. Performance measurement in cricket

Performance measurement is one of the imperative aspects in any format of the game of cricket, as each match generates a huge volume of performance related statistics. Relative to the other games, the contribution of individual players to the overall team performance is indispensable in cricket (Damodaran, 2006). In cricket, however, relevant literature pertaining to performance based classification of players is not plenty. Traditionally, different measures like batting average, economy rate, etc. are mostly used to compare the performances of batsmen and bowlers in cricket. Lewis (2005) stated that the existing measures for players' performances in cricket are unable to access the true abilities of the players. Also, since the different traditional measures are in different units of measurement so it is difficult to combine them. All these limitations to measure the performances of the cricketers are well discussed by Lewis (2005). Accordingly, an alternative performance measure is proposed by him, extending the Duckworth/Lewis method. The method can even take into consideration the situation at which a player performs. Lemmer (2004) developed a classification scheme for batsmen using performance data of one-day international (ODI) matches. Later Lemmer (2008) discussed the performances of cricketers in the first Twenty20 world cup. Van Staden (2009) used the terms ideal all-rounder, batting all-rounder and bowling all-rounder to classify the all-rounder's who participated in IPL-I. Brettenny (2010) reviewed the different existing measures of players' performances in cricket and used them to select players for a fantasy league using the binary integer programming model. Based on the strike rate and economy rate, Saikia and Bhattacharjee (2011) classified the performances of all-rounder's who participated in first three seasons of IPL. Then the factors responsible for classification of all-rounders are identified and naïve Bayesian classification model is used to predict the expected class of an incumbent all-rounder who had played only in the fourth season of IPL based on the responsible factors.

A team has a good chance of winning consistently, in any format of cricket, if the performances of individual players are reliable. Though both batting and bowling departments of a team need to perform significantly in determining the outcome of a cricket match, yet the present study gives emphasis on analyzing as well as predicting the performances of bowlers in IPL. In order to quantify the performances of bowlers, a measure called the combined bowling rate (CBR) developed by Lemmer (2002) is used. This measure is designed to quantify the performances of bowlers by combining the three traditional measures viz. *bowling average*, *economy rate* and *bowling strike rate*. Based on the performances of bowlers in the first three seasons of IPL, the paper tries to predict the performances of those bowlers who entered the IPL in its fourth season. The actual performances of these bowlers in IPL-IV are calculated, and the external validity of the neural network model is tested.

1.2. Lemmer's bowling performance measure

Lemmer (2002) proposed a bowling performance measure called the combined bowling rate (CBR) which comprises of combining three traditional bowling statistics viz. *bowling average*, *economy rate* and *bowling strike rate*. According to Lemmer (2002) the harmonic mean can be used to find the average of ratios, provided the numerator is considered as fixed and the denominator as variable. Therefore, to measure the bowling performance of bowlers Lemmer (2002) used the harmonic mean to combine the above mentioned three traditional bowling statistics.

Let r be the total number of runs conceded by a bowler, w the total number of wickets taken by a bowler and b the total number of balls bowled by a bowler in a series of matches. Then the traditional bowling statistics can now be defined as

$$\text{Bowling average} = \frac{r}{w}$$

$$\text{Economy rate} = \frac{r}{b/6}$$

$$\text{Bowling strike rate} = \frac{b}{w}$$

It is observed that the bowling average and economy rate have the same numerator. Thus, in order to combine three traditional bowling statistics using the harmonic mean, it is necessary to adjust the bowling strike rate to have the same numerator as the bowling average and economy rate. For that purpose, Lemmer (2002) proposed the following adjustment to the bowling strike rate.

$$\text{Bowling strike rate} = \frac{b}{w} = \frac{b}{w} \times \frac{r}{r} = \frac{rb}{rw} = \frac{r}{rw/b}$$

This form of the bowling strike rate can now be used in the calculation of the harmonic mean to combine the three traditional bowling statistics. So, the combined bowling rate (CBR) defined by Lemmer (2002) was

$$CBR = \frac{3}{\dfrac{1}{bowling\,average} + \dfrac{1}{economy\,rate} + \dfrac{1}{bowling\,strike\,rate}} = \frac{3r}{w + b/6 + rw/b}$$

Thus, the values of CBR indicate the performances of the bowlers in a match or a series through the three prime skills of bowlers' viz. *bowling average*, *economy rate* and *bowling strike rate*. Since low values of these performance statistics indicate good bowling, hence low values of CBR indicate good performances of the bowlers and *vice-versa*.

2. Artificial neural networks in sports

Data mining is the technique which explores and analyzes data in order to predict or discover meaningful patterns and rules. The different data mining techniques commonly used in sports applications are decision trees, artificial neural networks, clustering, genetic algorithms, etc. An artificial neural network is the one of data mining technique that is commonly used in sports applications to predict the result of a tournament, performance modeling, classification, etc. The most important characteristic of an artificial neural network is its ability to learn. The purpose of the learning or training is to train the neural network to perform some task. One application of artificial neural network in different sports is modeling swimming performance by Silva et al (2007). The different concepts and approaches of artificial neural networks in different sports were discussed by Perl (2001). Perl and Weber (2004) also highlighted the importance of pattern recognition in sports using neural network. Young and Weckman (2008) used ANN to predict the performances of football players by translating players' rating values to National Football League (NFL) combined values. In cricket, by using the Duckworth-Lewis method, Bailey and Clarke (2006) used ANN to predict the outcome in one day international cricket matches while the game is in progress. Also, Choudhury et al (2007) used ANN for predicting the outcome of limited over cricket tournaments.

2.1 Functional form of artificial neural networks

Let us consider the following relation which describes the functional form of the basic neuron model.

$$y = f(x) \qquad \text{... (1)}$$

where, $f(x) = w_1x_1 + w_2x_2 + \ldots + w_nx_n + b$

More conveniently, it can be written as

$$y = \sum x_iw_i + b \qquad \ldots \text{(2)}$$

In the above model, containing a set of 'n' inputs x_i, each input signal is multiplied with an associated weight, w_i before it is applied to the processing and b is a bias term. Though the above functional form looks to be linear, it may or may not be true. However, if a linear relationship between response variable and predictors is appropriate, then the result of neural network is closely approximate to a linear regression model (Norusis, 2007). Usually in neural network, the form of relationship between response variable and predictors is determined during the learning process. Therefore, it would always be better, if we consider that we have some inputs $x_1, x_2, \ldots, x_n$ and the desired output is represented by the response variable 'y', instead of thinking about the functional form of the model structure. In literature, there are so many types of artificial neural networks but here we have used Multilayer Perceptron (MLP) neural network to predict the performances of bowlers in IPL.

2.2. Multilayer perceptron (MLP) neural network

A multilayer perceptron artificial neural network is an information processing system which consists of many nodes. Let us consider the following diagrammatic representation of MLP neural network (see Figure 1) where the x_1, x_2, x_3 elements are called neurons which process the information. The signals between neurons are transmitted by means of connection links (i.e. arrows). The links possess associated weights w_1, w_2, w_3, which are multiplied along with the incoming signals from inputs of the network. The output signal is obtained by applying activations to the net input.

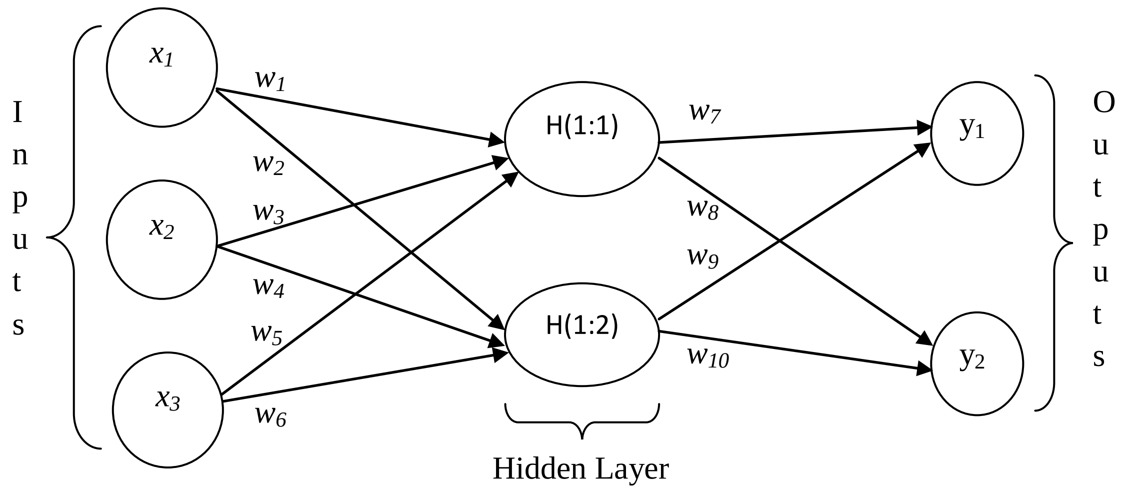


Figure 1. A general multilayer perception neural network diagram.

In the above diagram, H(1:1) is the first unit of the first hidden layer and H(1:2) is the second unit of the first hidden layer. However, though we have considered only a single hidden layer in the above diagram one can select a maximum of two hidden layers in the MLP neural network. The software used in this paper for performing a MLP neural network is Predictive Analytical Software (i.e. PASW) version 18.0.

3. Data and methodology

The data pertinent to the performances of the bowlers are collected from the website *www.cricinfo.org*. To quantify the performances of players it is necessary that players' statistics from a large number of games shall be considered. The actual quality of a player cannot be properly judged from one or two appearances. The effects of outstanding or poor, single performances are smoothed over the larger number of games (Lewis, 2005). Therefore, individual performances across a series of matches are required to provide a suitable frame of reference. The nature of the professional sport ensures that the majority of individuals will experience sufficient match-play to enable this type of methodology to be deployed (Bracewell and Ruggiero, 2009). Thus, some selection criterion needs to be set up which shall also help to identify the performance of players. The players who had satisfied all the following conditions were kept in the training sample.

- The bowler has played at least 7 matches in IPL.
- The bowler has bowled at least 100 balls in IPL.
- The bowler took at least 5 wickets in IPL.

There are only 83 players who satisfied all the above criteria. All these players are considered for the study and information about their performances in IPL, ODIs and Twenty20 internationals have been collected from the above mentioned sources.

3.1 Selection of variables

Like the procedure of regression, one has to choose dependent and independent variables for artificial neural network also. Several independent or input variables that are supposed to influence the performances of bowlers are considered. One is a nominal variable (e.g. bowling type: either fast or spin), while others are continuous (e.g. age, strike rate in Twenty20, etc.) in nature. Brief descriptions of the different input variables are given below.

Age: This is a demographic variable measuring the age of the bowler. It is the number of years completed by the player at the end of IPL-III.

ODI matches played (Mts_ODI): Measures the experience of the player in terms of number of one day international matches in which the player was in the playing eleven.

Bowling average in ODI (Bavg_ODI): Career bowling average of the bowlers in one day international cricket matches.

Economy rate in ODI (ER_ODI): Career economy rate of the bowler in one day international cricket matches.

Bowling strike rate in ODI (BSR_ODI): Career bowling strike rate of the bowler in one day international cricket matches.

Twenty20 matches played (Mts_T20): Measures the experience of the players in Twenty20 cricket. It is the number of international Twenty20 matches in which he was in the playing eleven.

Bowling average in Twenty20 (Bavg_T20): Career bowling average of the bowler in Twenty20 international cricket matches.

Economy rate in Twenty20 (ER_T20): Career economy rate of the bowler in Twenty20 international cricket matches.

Bowling strike rate in Twenty20 (BSR_T20): Career bowling strike rate of the bowler in Twenty20 international cricket matches.

Experience in international cricket (Exp_Int_cricket): This is another measure of experience which measures the international career of the player in terms of years.

Bowling Type (Bowl_type): This is a nominal variable (i.e. either Fast or Spin). The code '1' represents fast bowler and '2' represents spin bowler.

Bidding price in IPL (Bid_IPL): It is the amount of money in dollars for which a given player was auctioned in IPL.

In ANN, the response variable may be continuous (i.e. scale) or a categorical (i.e. nominal/ordinal) variable. Unlike regression, neural networks can handle more than one dependent variable. However, this paper used a categorical response variable to predict the performances of bowlers. This categorical variable has been generated from the CBR, a measure developed by Lemmer (2002) which is already stated in section 1.2. The three different categories of bowlers based on their performance generated from CBR are named *poor, average* and *good* performer. The detailed procedure of how the categorical dependent variable has been generated from CBR is described comprehensively in subsequent sections.

3.2 Classification based on CBR

The CBR values are calculated for all the bowlers who were selected in the training sample. A meaningful classification based on players' bowling performance would be in terms of a suitable interval from an assumed probability distribution of CBR. Thus, the probability distribution of CBR should be examined to facilitate the classification of the bowlers on the basis of their performance. For testing the hypothetical distribution of the CBR, the Kolmogorov-Smirnov (K-S) test is used.

However, for performing the Kolmogorov-Smirnov test, the theoretical probability distribution needs to be completely specified (i.e. the values of the parameters should be known). In this case, the parameters are estimated from the data. Now after finding the distribution of CBR, one can find two real numbers x_1 and x_2 to divide the range $[0, \infty]$ into three linear intervals namely $(0, x_1)$, (x_1, x_2) and (x_2, ∞) with the same probability weight of 33.33 percent. As already mentioned, CBR can be termed an inverse measure; therefore, the following intervals are used to categorize the various stages of performance of the bowlers.

i) Good if $0 \leq \text{CBR} < x_1$
ii) Average if $x_1 \leq \text{CBR} < x_2$
iii) Poor if $x_2 \leq \text{CBR} < \infty$

As of the above classification, the study has generated the categorical dependent variable based on CBR values and used it to train the neural network. This dependent variable is also used to predict the appropriate class of the bowlers in the fourth season of IPL. The cut off points acquired from CBR can be seen in Table 1.

4. Data analysis and result

One probable distribution of CBR may be the normal distribution and the probability density function is given by,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2} \qquad -\infty < x < \infty \qquad \dots (3)$$

It has been observed that the normal distribution did not fit properly to the calculated CBR values. Therefore, the original CBR values were transformed using $\log_{10}$(CBR) to fit the normal distribution. The estimated values of the parameters μ and σ^2 are 1.158 and 0.0063 respectively obtained from the transformed CBR values. The corresponding result of the Kolmogorov-Smirnov test is provided in Appendix-B.

The *p*-value of the Kolmogorov-Smirnov test statistic at the 5 percent level of significance is given by 0.401, providing sufficient evidence that the transformed CBR (i.e. $\log_{10}$(CBR)) values can be considered to follow the normal distribution as in equation (3). One can also visualize it through a histogram along with the normal probability curve in Figure 2.

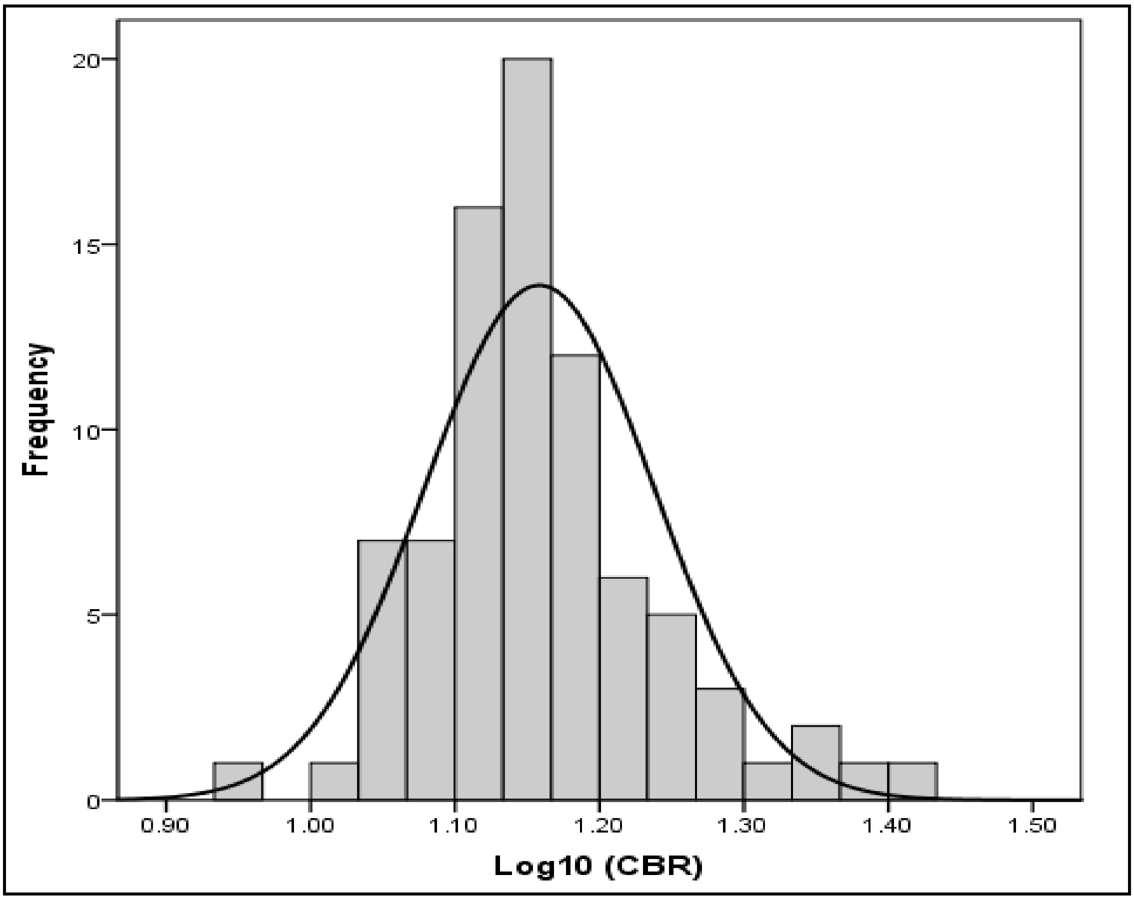


Figure 2. Histogram for transformed CBR values with normal curve.

The cut off points that are obtained from the assumed normal distribution of $\log_{10}$(CBR) are given below.

Table 1. Interval for classification of log10(CBR).

Classification	log10(CBR)
Good performance	below 1.125
Average performance	1.125 – 1.193
Poor performance	1.193 & above

4.1 Training or Learning of MLP neural network

The process of neural network training or learning refers to find out the best set of weights for the network (Francis, 2001). These weights are also known as parameters of the neural network model. Like in a regression model, these parameters need to be estimated before the network can be adopted for further use. The weights of different input variables from input layer to hidden layer and hidden layer to output layer can be seen in Appendix-C. Also, how many input and hidden units are used for the study in the input and hidden layer of the network diagram can be seen in Appendix-A.

The difficulty arises in an MLP neural network is that the estimated weights or parameters are not easily interpretable. It is because a large input to hidden layer weight doesn't necessarily mean that the input has a huge effect on the output and *vice-versa*. Instead, it may bring around abrupt changes in the output layer. Thus, incorporating both the layers of weights various measures are proposed to measure the importance of input variables. Garson (1991) proposed a measure which is a sum of products of normalized weights. But this measure doesn't take into account the signs of the weights. Milne (1995) modified Garson's formula by taking the absolute values of certain terms. Yet, the absolute values are not applied correctly to avoid cancellation of weights of opposite signs. However, based on the training sample, PASW 18.0 performs a sensitivity analysis, which computes the importance of each predictor in determining the performances of the bowlers. The importance of the independent variables for the study can be seen in Figure 3.

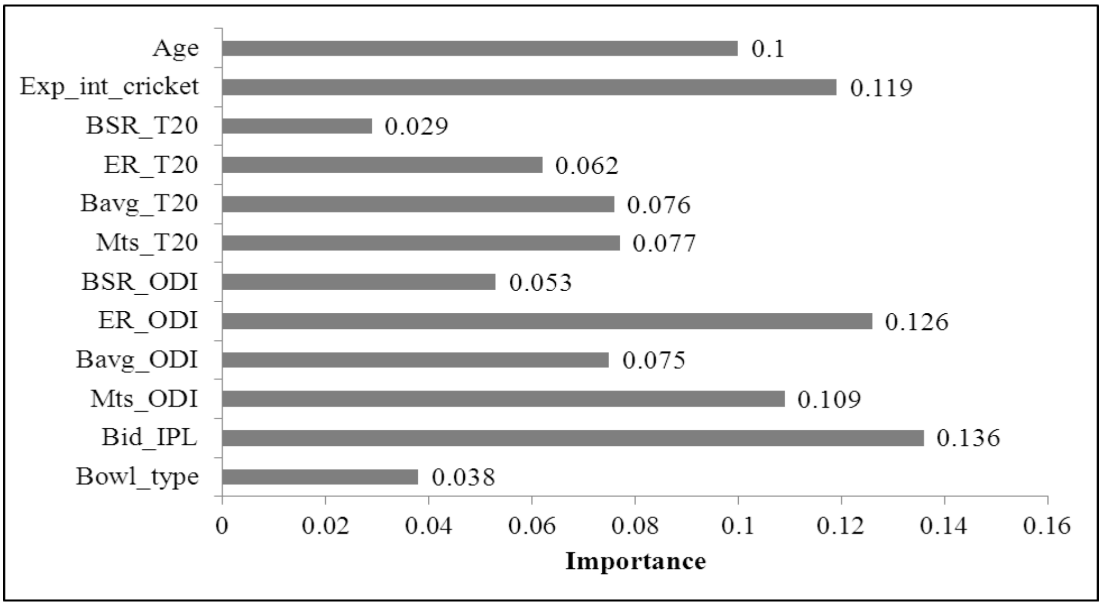


Figure 3. Bar diagram for importance of independent variables

In the above graphical representation, it can be revealed that the variable bid price in IPL has the largest effect on predicting the performance of bowlers, followed by the economy rate in one-day international cricket matches. Also, it is seen that amongst the different input variables, bowling strike rate in Twenty20 cricket has the least importance in predicting the performances of bowlers. Besides, the classification table for training data shows that overall 70.7 percent cases are correctly classified during training against 29.3 percent incorrect prediction (see Appendix-D).

4.2 External validation of neural network model

To test the external validity of the model a different sample of bowlers was considered. The players considered for the validity exercise are those who had not participated in the first three seasons of IPL but had played in the fourth season only. To get sufficient information about the bowling performances of bowlers in IPL-IV, the bowlers who played at least 5 matches and bowled at least 60 balls are considered. Seventy seven (77) bowlers are found to satisfy the above mentioned criteria and seven (7) of these players had played in the fourth season only. The information of three imperative variables viz. *bowling average*, *economy rate* and *bowling strike rate* that are required to calculate the CBR for the bowlers, who had satisfied the above mentioned criteria was collected from *www.cricinfo.org* on 5^{th} April, 2011. The different classes of bowling performance are identified by the same procedure as discussed in section 3.2. Here also the CBR values are transformed (i.e. $\log_{10}$(CBR)) to determine actual performances of the bowlers in IPL-IV. Again the Kolmogorov-Smirnov test statistic's value of 0.405 has confirmed that the values of $\log_{10}$(CBR) for IPL-IV follow normal distribution. It can also be visualized through a histogram along with the normal probability curve in Figure 4. The cut off points obtained from the $\log_{10}$(CBR) to measure the actual bowling performance of the players in IPL-IV are given in Table 2.

Table 2. Interval for classification of $\log_{10}$(CBR) in IPL-IV

Classification	$\log_{10}$(CBR)
Good performance	below 1.456
Average performance	1.456 – 1.189
Poor performance	1.189 & above

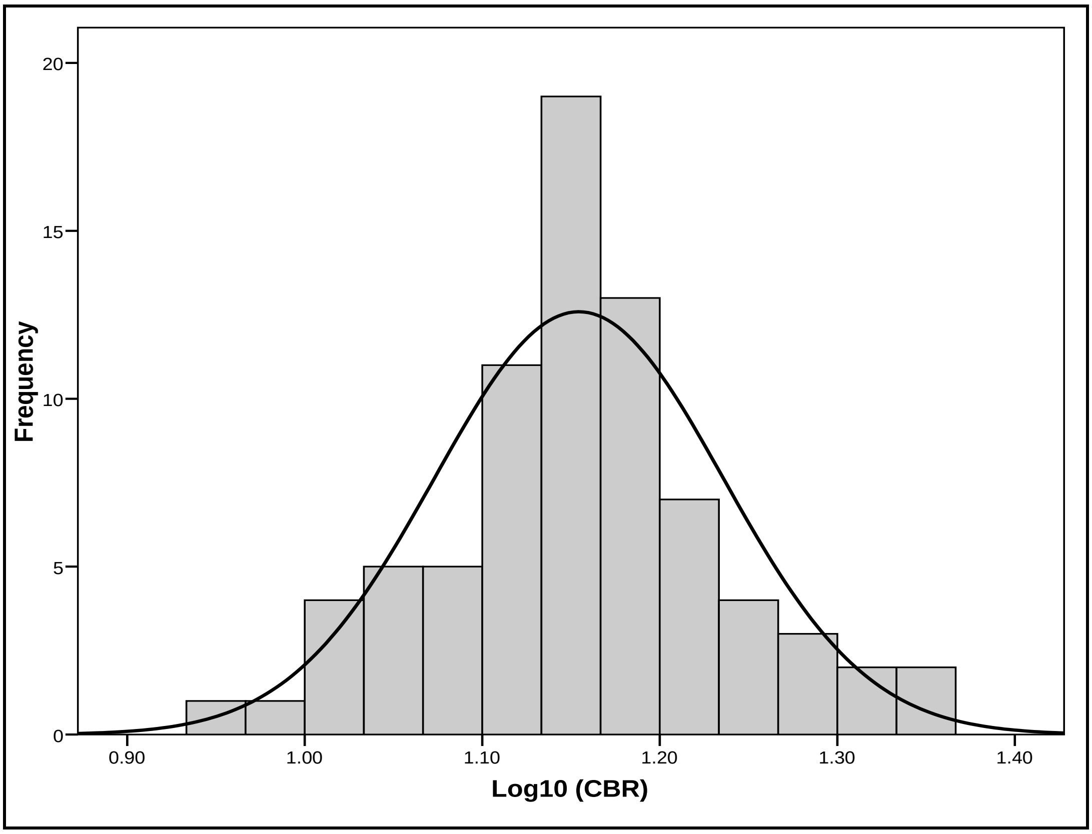


Figure 4. Histogram for CBR in IPL-IV with normal curve

Finally, the predicted bowling performances of seven bowlers who had participated only in IPL-IV are recorded and their actual bowling performance is also determined to check the validity of the neural network model. The predicted and actual bowling performances of the selected bowlers in IPL-IV are provided in Table 3.

Table 3. Predicted and actual bowling performance of seven bowlers in IPL-IV

Sl. No.	Players name	Predicted class of the performance in IPL-IV	Actual class of the performance in IPL-IV	Bid price in IPL-IV
1	J Franklin	Good	Poor	$100000
2	J Ryder*	Average	Average	$160000
3	AC Thomas	Good	Average	$100000
4	SA Hasan*	Good	Good	$425000
5	W Parnell*	Good	Good	$610000
6	R Harris*	Average	Average	$325000
7	JE Taylor*	Average	Average	$100000

* The predicted and actual performance/categories of these bowlers are correctly predicted.
Source: Compute based on data collected from *www.cricinfo.org*

In table 3, it has been observed that artificial neural network model is able to predict the exact class of performance of 5 bowlers out of 7 bowlers who had played in IPL-IV only. Thus, a franchisee depending on the team's requirement can decide which bowlers to bid for and who shall not be considered. The fresh auction of IPL-IV was held in 8th and 9th of January 2011, with two new teams joining the league. In the new auction, enormous raise in the players' price compared to the auctions of IPL-I. The average bid price of the IPL-IV auction is $495,605 while in the previous auction (before IPL-I) was $448,217. As evident from the above table, the player W Parnell was bided precisely compared to his performance as his bid price was above the average in IPL IV. This prediction would be helpful for the franchisee to decide which bowlers they should target for their team and at what cost.

5. Conclusion

The salaries of players in IPL that are decided through auction are a way of quantifying players' performance in monetary terms. Thus, it is a matter of decision making from the part of the franchise about which player is to be bided for and up to what cost. Therefore, the artificial neural network model used in this paper can help a franchisee to take a decision. The salaries offered to the players prior to the first season of IPL were valid for three years. At the beginning of the fourth season of IPL, new agreements are made through fresh bidding. If the league continues, the fresh bidding shall take place after every three years, in such a context, this study may be used to predict the probable category of bowlers on the basis of their previous performances. This prediction can help the franchisee to decide which bowlers they should target for their team and who should not be considered at all.

The artificial neural network model when applied for external validity was found to be 71.43 percent accurate. This has resulted after a training sample used to develop the model was loaded with only 83 bowlers. As the age of the league increases, a larger number of bowlers can be considered in the training sample and a much better prediction is expected from the model. However, the form of a bowler when he actually plays in the tournament is a deterministic factor for the actual classification. But the model, once rich in data, is supposed to work well provided the performance of the bowler is not much variant to what he had been performing in the previous matches.

6. References

Bailey, M. and Clarke, S. R. (2006). Predicting the match outcome in one day international cricket matches, while the game is in progress. **Journal of Sports Science and Medicine**, 5(4), 480-487.

Bracewell, P. J. and Ruggiero, K. (2009). A parametric control chart for monitoring individual batting performances in cricket. **Journal of Quantitative Analysis in Sports**, 1-19.

Brettenny, W. (2010). Integer optimization for the selection of a fantasy league cricket team. Dissertation Report, South Africa: Nelson Mandela Metropolitan University.

Choudhury, D. R., Bhargav, P., Reena and Kain, S. (2007). Use of artificial neural networks for predicting the outcome of cricket tournaments. **International Journal of Sports Science and Engineering**, 1(2), 87-96.

Damodaran, U. (2006). Stochastic dominance and analysis of ODI batting performance: The Indian cricket team, 1989-2005. **Journal of Sports Science and Medicine**, 5, 503-508.

Francis, L. (2001). Neural network demystified. Available online at http://citeseerx.ist.psu.edu/viwedoc/download?doi=10.1.1.116.1597&rep=rep1&type=pdf (Retrived on 5 January 2010).

Garson, G. D. (1991). Interpreting neural network connection weights. AI Expert, April, 47-51.

Kitchin, P. (2008). The development of limited overs cricket: London's loss of power. **London Journal of Tourism, Sports and Creative Industries**, 1(2), 70-75.

Lemmer, H. (2002). The combined bowling rate as a measure of bowling performance in cricket. **South African Journal for Research in Sport, Physical Education and Recreation**, 24(2), 37-44.

Lemmer, H. (2004). A measure for the batting performance of cricket players. **South African Journal for Research in Sport, Physical Education and Recreation**, 26(1), 55-64.

Lemmer, H. (2008). An analysis of players' performances in the first cricket Twenty20 world cup series. **South African Journal for Research in Sport, Physical Education and Recreation**, 30, 71-77.

Lewis, A. J. (2005). Towards fairer measures of player performance in one day cricket. **Journal of the Operational Research Society**, 56, 804-815.

Malcolm, D., Gemmell, J. and Mehta, N. (2009). Cricket and modernity: International and interdisciplinary perspectives on the study of the imperial game, **Sport in Society**, 12(4/5), 431?446.

Milne, L. K. (1995). Feature selection with neural networks with contribution measures. Artificial Intelligence AI'95 (pp. 1-8). Canberra: Isaac Council.

Mitra, S. (2010). The IPL: India's foray into world sports business. **Sport in Society**, 13(9), 1314-1333.

Norusis, M. (2007). **The SPSS statistical procedure companion**. New Delhi: Prentice Hall.

Perl, J. (2001). Artificial neural networks in sports: New concepts and approaches. **International Journal of Performance Analysis in Sports**, 1(1), 106-121.

Perl, J. and Weber, K. (2004). A neural network approach to pattern leraning in sports. **International Journal of Computer Science in Sports**, 3(1), 1-4.

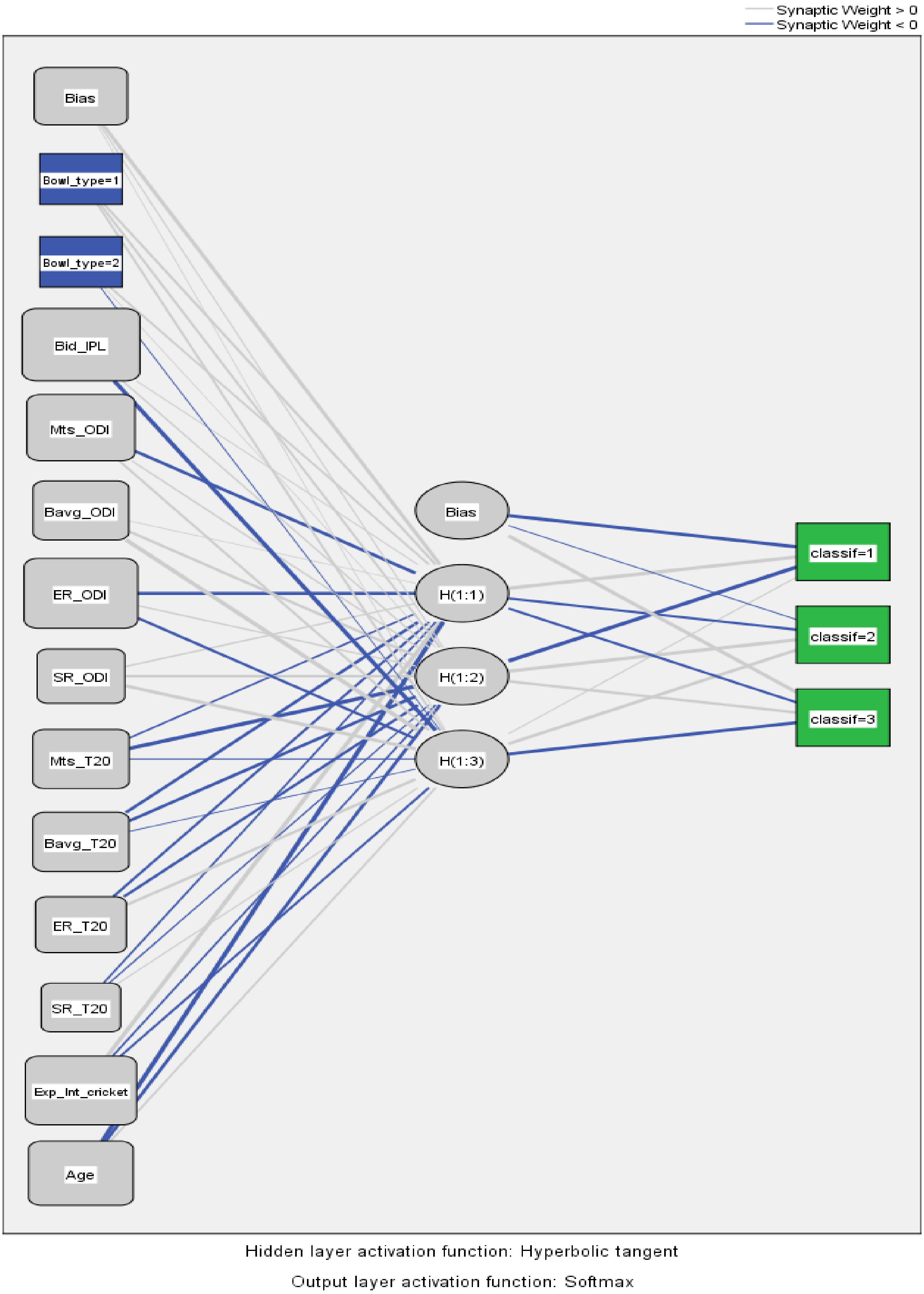
Saikia, H. and Bhattacharjee, D. (2011). On classification of all-rounders of the Indian Premier League (IPL): A Bayesian approach. **Vikalpa**, 36(4), 25-40.

Silva, A. J., Costa, A. M., Oliveira, P. M., Reis, V. M., Saavedra, J., Perl, J., Rouboa, A. and Marinho, D. A. (2007). The use ofneural network technology to model swimming performance. **Journal of Sports Science and Medicine**, 6(1), 117-125.

Van Staden, P. J. (2009). Comparison of cricketers' bowling and batting performances using graphical displays. **Current Science**, 96, 764-766.

Young, W. A. and Weckman, G. R. (2008). Evaluating the effects of aging for professional football players in combine events using performance-aging curves. **International Journal of Sports Science and Engineering**, 2(3), 131-143.

Appendix-A: The multilayer network diagram for the current problem



Data Source: The graph is drawn based on the data from the training sample and is drawn using the package PASW version 18.0.

Appendix-B: One-sample Kolmogorov-Smirnov test

		log10(CBR)
N		83
Normal Parameters	Mean	1.1583
	Std. Deviation	.07944
Most Extreme Differences	Absolute	.098
	Positive	.098
	Negative	-.085
Kolmogorov-Smirnov Z		.894
Asymp. Sig. (2-tailed)		.401
Test distribution		Normal

Appendix-C: Estimation of weights from MLP neural network

Predictor		Predicted					
		Hidden Layer 1			Output Layer		
		H(1:1)	H(1:2)	H(1:3)	[classif=1]	[classif=2]	[classif=3]
Input Layer	(Bias)	2.273	.288	.030			
	[Bowl_type=1]	.811	.698	.633			
	[Bowl_type=2]	.597	.026	-.125			
	Bid_IPL	.212	.478	-2.874			
	Mts_ODI	-1.851	.488	.801			
	Bavg_ODI	.105	1.367	4.520			
	ER_ODI	-1.810	.492	-1.314			
	SR_ODI	.542	.950	2.372			
	Mts_T20	-.580	-3.791	-.472			
	Bavg_T20	-1.529	-1.754	-.126			
	ER_T20	-.888	-1.344	1.423			
	SR_T20	-.578	-.426	.371			
	Exp_Int_cricket	2.718	-.547	-.750			
	Age	-3.004	-1.681	.596			
Hidden Layer 1	(Bias)				-2.458	-.144	2.559
	H(1:1)				2.692	-1.149	-1.114
	H(1:2)				-2.962	2.617	1.510
	H(1:3)				.303	2.118	-2.347

Appendix-D: Classification information during training

Sample	Observed	Predicted			
		poor	average	good	Percent Correct
Training	Poor	20	7	0	74.1%
	Average	3	32	0	91.4%
	Good	4	10	6	30.0%
	Overall Percent	32.9%	59.8%	7.3%	70.7%