Using Fuzzy Logic to Predict Winners in Horseraces at the Champ de Mars

Manish Jogeeah
Faculty of Engineering
University of Mauritius
Reduit, Mauritius

Akshay Kumar Chandoo
Faculty of Engineering
University of Mauritius
Reduit, Mauritius

Pravin Selukoto Paupiah
Department of IT Amity
University Ebene,
Mauritius

Sameerchand Pudaruth
Faculty of Ocean Studies
University of Mauritius
Reduit, Mauritius

ABSTRACT

In this paper we have used a Fuzzy Logic approach to predict winners at the Champ de Mars race course. We have built a fuzzy logic system with 5 input variables and 1 output. The Fuzzy Inference System (FIS) also consists of 35 fuzzy rules based on expert experience. These rules are analysed turn by turn to evaluate the input values and enable the FIS to determine the output which is an aggregate of the individual results from the evaluation of each rule. The FIS is based on Mamdani algorithm which uses the centroid technique during the defuzzification process to produce a single value result. To our knowledge, this is the first real application of fuzzy logic in horse racing. The system was tested for 3 consecutive race meetings with a 77.7 % success of predicting horses finishing in top 3 places for 2 out of the 3 race meetings tested and an average of 23.1% horses winning for the 26 races used for testing. The system could be enhanced by adding more valuable input parameters and the addition of more rules which could provide greater flexibility and coverage. The knowledge gained from this research might be useful for the application of fuzzy logic to other areas like decision making in business, selection of models in construction firms, automating controllers and others.

KEYWORDS

Fuzzy Logic, horseracing, Champ de Mars, Fuzzy Inference System (FIS)

1 INTRODUCTION

Horse racing is organised in Mauritius at the Champ de Mars which is situated in Port Louis, the capital of the Republic of Mauritius. Mauritians are very fond of horse racing and follow horse racing events very closely. Usually a racing season last for about 8 months and consists of around 40 race meetings. Gambling on horse racing is very popular among Mauritians. Tips which circulate around the island as rumours tempts many people to bet. Several race magazines like L'Express Turf, Racetime, Defi Turf and many more are published some days before a race meeting and they contain a lot of information about the horses participating in the forthcoming race meeting. The magazines contain interviews, horses' current form, training updates and the analysis of professional tipsters.

These magazines serve as tools for the bettor to perform his analysis of the races and to come up with personal conclusions about the outcome of a race. However this does not seem to be a proper technique. This is because a large amount of data which could have been influential in the outcome of a particular race are ignored or are not available and also most racegoers are not experts in analysing information. Many of the race parameters involved influence the output in their own different ways, and moreover they do not always have a static effect on the output. Horse racing is quite unpredictable and keeps on evolving with different parameters and in the long run, an observable pattern or behaviour may surface.

The aim of this project is to investigate how successfully a fuzzy logic approach can be used to model horse races taking into consideration few parameters and then use this model to predict winners at the Champ de Mars. Fuzzy logic can adapt to inaccurate data and it relies on the experience of experts in the field. Every possible inputs to our system will have an influence on the outcome and their influence is always varying and is

rather unreliable, this is why the model should be based on what is approximate rather than on what is fixed because we have no such element in the field of horse racing that is guaranteed to produce the same output each time.
This paper proceeds as follows. In section 2, works previously done in similar fields are discussed. In section 3, the methodology and approach to solve the problem are described. Section 4 consists of evaluating the results of this project and make conclusions about the efficacy of the solution proposed. Finally section 5 concludes the paper with some ideas for further work.

2 LITERATURE REVIEW

There are many areas and fields where fuzzy logic methods have been applied like robotics, pattern recognition, image processing, computer vision, forecasting, prediction, expert systems, decision making, modelling and data mining. In sports, fuzzy logic has been rarely used, but gradually its use has been tested and found to be promising.
Problems that are imprecise can be solved with fuzzy logic and may give more accurate results. The concept of fuzzy logic was first introduced by P. L. A. Zadeh [1]. New algorithms and decision analysis were developed using this theory. The implementation of fuzzy logic controllers by using genetic algorithms was discussed in [2]. Fuzzy logic, fuzzy sets and their different aspects which are important to create a fuzzy logic system are discussed in [3]. Fuzzy inference systems can be learned in different ways using neural network learning techniques which was presented by Abraham [4]. A loan risk calculator based on fuzzy logic was developed to help financial companies in making important decisions [5]. The impact on quality when developing software projects using fuzzy logic is discussed in [6].
One of the sports where fuzzy logic has been implemented and received good results is cricket. In [7] [8], the development of expert systems to analyse fast bowling based on vague conditions has been discussed. A training system incorporating Zadeh’s theory for classifying batting stroked have been suggested in [9] [10]. In [11], Singh *et al.* studied and evaluated the performances of cricket players on certain input parameters that were approximate. Different scenarios of the athletes were analysed based on the ranking of the athlete. Daily activities can be measured using modelling based on fuzzy techniques [12] [13]. Heller and Witte presented a control simulation for approximating dynamical behaviour of motor units [14]. Sports training can be modelled using fuzzy based machine learning algorithms in areas such as swimming [15].
Silverman used a Gibbs model in order to predict the speed of a horse. He assumed that the horse with the fastest speed would win most of the races. However, when tested with real data, the horse with the fastest speed won only 21.63% of the total races [16].
The study of Jannet Williams and Yan Li in Jamaica proved that taking relevant parameters into consideration can increase the probability to find the probable winners in a race. Artificial Neural Network was used to create a system that is trained from existing racing data [17]. A probabilistic approach was used in [18] to determine the winner of a horse race. Two hundred and forty races from the 2010 horse racing season from the Champs de Mars racecourse in Mauritius was used for testing. In [19], an experiment using fuzzy logic was carried out on 2012 horseracing season in Mauritius. In this case, 3.3 winners could be predicted per race meeting and this was slightly better than most tipsters in the island.

3 METHODOLOGY

The aim of this work is to be able to predict the winner or horses finishing in the first 3 places of a horse race by predicting the margin by which a horse loses the race. The horse with the least margin, is expected to be the winner. A fuzzy logic inference system is used to analyse each horse’s chances of winning. A Fuzzy Inference Systems (FIS) takes in inputs, analyse the inputs based on

some fuzzy rules and produces an output after the evaluation of each rule which when combined gives a final result value for predicted margin.
Two algorithms commonly used to determine output of a fuzzy inference system are Sugeno and Mamdani. Mamdani algorithm is chosen as the inference algorithm for this research as it widely recognised and used [20].

We have used the MATLAB R2014b Fuzzy Logic Designer to design the model the fuzzy logic inference system.

Horse races results for the 2014 season has been gathered in an excel sheet. The data consists of results for the first 41 race meetings. The sheet contains 363 races' details and also the different inputs, weight, draw, jockey worth, odds and previous performances, required by our FIS.

There are 3 main steps which are carried out in an FIS system before an output is produced. The steps are:

1. Fuzzification
   Inputs are analysed to see the degree of their membership based on the defined membership functions.

2. Inference
   The fuzzy operator takes as inputs membership values generated form the Step 1. The fuzzy rules are evaluated and an output is generated for each rule. Output from every rule are merged together to form a single fuzzy set called an aggregate.

3. Defuzzification
   Each rule is evaluated for a given set of inputs and each of these rules produce an output. These outputs are combined together into a single fuzzy set, and this is called the aggregation process. Defuzzification involves taking as input the output of the aggregation process, the single fuzzy set which contains multiple values and calculate a single value.

The fuzzy inference system has 5 input variables and 1 output variable. Each input variable will have 3 membership functions whereas the output has 4 membership functions. The membership functions classifies an input value in one or more than one category and based on this classification, rules are evaluated to come up with predictions.

3.1. Input Variables

1. Weight

Weight refers to weight a horse is allocated by the handicapper based on its rating. The weight also includes the jockey's weight. Usually maximum weight a horse is handicapped is 62kg and minimum weight a horse carries is around 49kg. The lesser the weight, the better chance a horse is considered to have. Weight can take any value from 49 to 62.

2. Draw

Draw indicates the barrier number from which a horse will be starting the race. A high barrier draw is considered a drawback at the Champ de Mars. A horse with a high barrier draw has higher possibility of racing wide, hence covers more distance and get tired earlier. Draw can take any value from 1 to 12.

3. Odds

The favourite horse will generally offer lowest return. The odds of a horse carries a lot of information and this is an important factor to consider in order to identify winners.

Odds will be calculated as follows:

$$\text{Odds} = \frac{\text{Odds}}{\text{Highest Odds}} * 10$$

The horse with the highest odds that is the one least expected to win will have an Odds value of 10.

4. JockeyWorth

Different jockeys have different abilities. Some jockeys with better riding style and judgement perform better than others.

JockeyWorth will give an insight about the ability of a jockey. A horse having a good rider on board will definitely have a better chance in the race. Jockeys are rated on a scale of 0-10. JockeyWorth takes a value between 0 and 10.

JockeyWorth is calculated as:

a: jockey's position in jockey's log
b: worst jockey position in jockey's log from this race

$$\text{JockeyWorth} = 10 - \frac{a}{b} \times 10$$

5. PreviousPerf

The last performances of the horse are considered as this will give an idea of the fitness and form of the horse. A horse with excellent previous performances is considered to have a better chance compared to a horse with poor previous performances.

Previous Performance is calculated as follows:

$$\text{PreviousPerf} = \frac{\text{Sum of last N Performances}}{\text{N}}$$

A horse finishing first will have the least value for PreviousPerf and will be equal to one.

3.2 Output Variable

1. Margin

The margin is the distance between the winning horse and horses behind. The winner will have a margin value of zero. The horse finishing last will have the highest margin. Because some horses can finish very far behind, we limit the margin to 20 lengths.

3.3 Fuzzy Rules

Considering the inputs and output, 35 rules have been created for mapping purposes. Different rules have different weights based on expert experience and observations. Some of the rules defined are described below:

Rule 1. If (Weight is Light) and (Draw is Excellent) and (JockeyWorth is Top) and (Odds is Favourite) and (PreviousPerf is Good) then (Margin is VeryClose) (1)

Rule 2. If (Weight is Light) and (Draw is Excellent) and (JockeyWorth is Top) and (Odds is FairChance) and (PreviousPerf is Good) then (Margin is VeryClose) (1)

Rule 5. If (Weight is Light) and (Draw is Good) and (JockeyWorth is Top) and (Odds is Favourite) and (PreviousPerf is Good) then (Margin is VeryClose) (1)

Rule 6. If (Weight is Light) and (Draw is Good) and (JockeyWorth is Top) and (Odds is Underdog) and (PreviousPerf is Good) then (Margin is Close) (0.5)

Rule 8. If (Weight is Light) and (Draw is Bad) and (JockeyWorth is Top) and (Odds is Favourite) and (PreviousPerf is Good) then (Margin is VeryClose) (0.8)

Rule 12. If (Weight is Light) and (Draw is Bad) and (JockeyWorth is Bad) and (Odds is Underdog) and (PreviousPerf is Bad) then (Margin is VeryFar) (1)

Rule 14. If (Weight is Average) and (Draw is Excellent) and (JockeyWorth is Good) and (Odds is Favourite) and (PreviousPerf is Good) then (Margin is VeryClose) (0.9)

Rule 15. If (Weight is Average) and (Draw is Good) and (JockeyWorth is Good) and (Odds is FairChance) and (PreviousPerf is Good) then (Margin is Close) (1)

Rule 17. If (Weight is Average) and (Draw is Bad) and (JockeyWorth is Top) and (Odds is FairChance) and (PreviousPerf is Good) then (Margin is Close) (1)

Rule 19. If (Weight is Average) and (Draw is Good) and (JockeyWorth is Top) and (Odds is Underdog) and (PreviousPerf is Average) then (Margin is Close) (0.5)

ISBN: 978-1-941968-14-7 ©2015 SDIWC

Rule 34. If (Weight is Heavy) and (Draw is Excellent) and (JockeyWorth is Top) and (Odds is FairChance) and (PreviousPerf is Bad) then (Margin is Close) (1)

Rule 35. If (Weight is Heavy) and (Draw is Good) and (JockeyWorth is Good) and (Odds is FairChance) and (PreviousPerf is Average) then (Margin is VeryClose) (0.5)

3.4. Design of the FIS in MATLAB Fuzzy Logic Designer

The figure below depicts the FIS with its 5 input variables: Weight, Draw, JockeyWorth, Odds and PreviousPerf and 1 output variable: Margin.

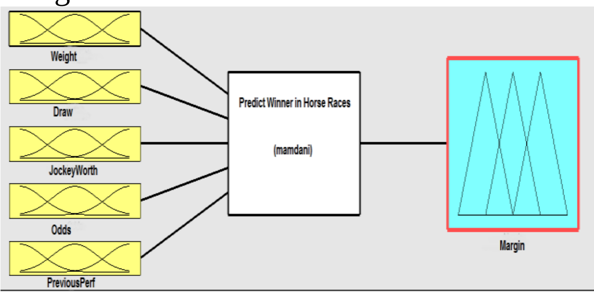

Figure 3.1: FIS with 5 Inputs and 1 Output

The Gaussian curve membership function (gaussmf) has been used to model the degree of membership for all the input variables and it takes two parameters.

The figures below show the membership functions for the 5 inputs.

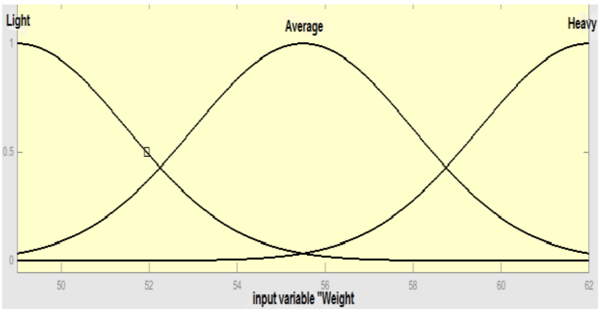

Figure 3.2: Membership Functions for Input Variable "Weight"

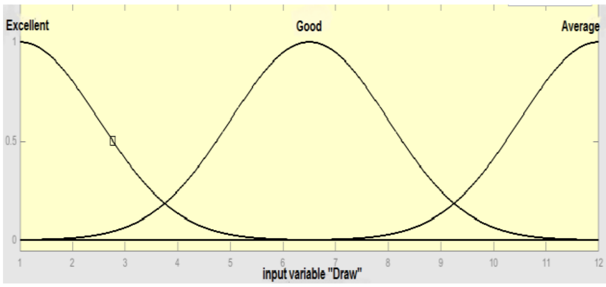

Figure 3.3: Membership Functions for Input Variable "Draw"

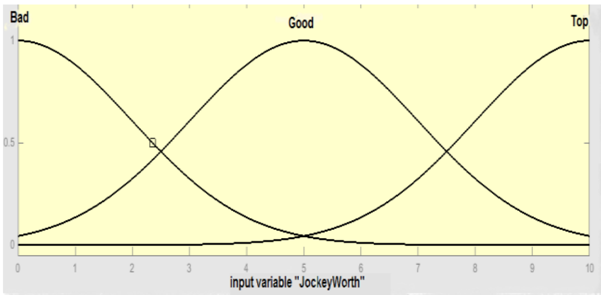

Figure 3.4: Membership Functions for Input Variable "JockeyWorth"

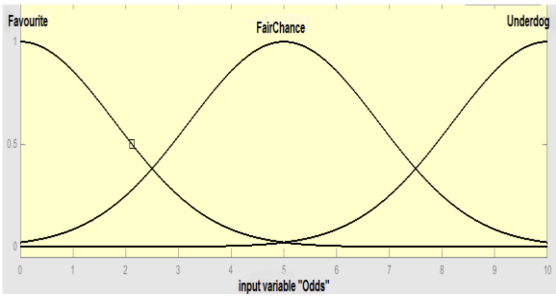

Figure 3.5: Membership Functions for Input Variable "Odds"

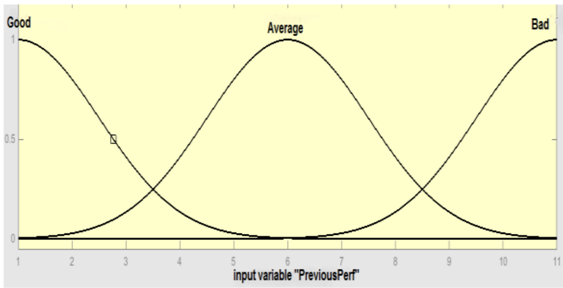

Figure 3.6: Membership Functions for Input Variable "PreviousPerf"

The generalized bell membership function (gbellmf) has been used to model the degree of membership for the output variable and it takes three parameters.

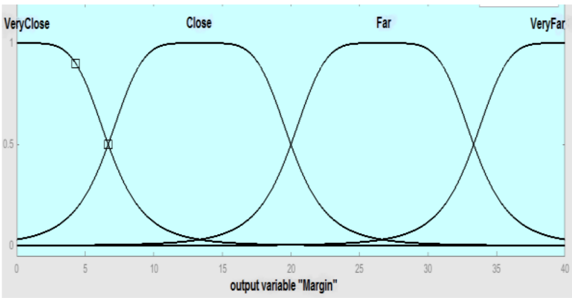

Figure 3.7: Membership Functions for Output Variable "Margin"

4 EXPERIMENTS, RESULTS AND EVALUATION

In this research the system has been built using MATLAB R2014b app 'Fuzzy Logic Designer'. Testing was performed for the three meetings of 2014 horse racing season.

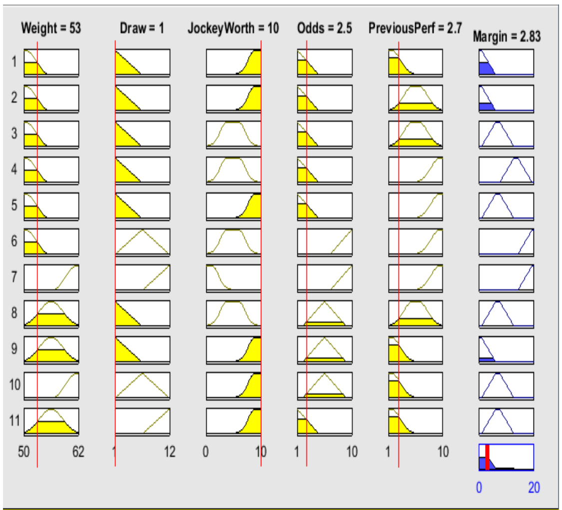

Figure 4.1: Test and observation using some specific input values

The figure above depicts a particular scenario whereby we can see the input value of each input variable and the outcome produced from the evaluation of each of these rules. In the lower right corner, the diagram represents the outcome of the aggregation process whereby outcomes from each rule are combined together into a single fuzzy set.

The final output during the defuzzification process was determined using the centroid method. This method is the most commonly used in Mamdani algorithm and it very precise. Using this method the centre of area under the curve is returned.

TABLE 1. PREDICTED RESULTS FOR MEETING 31

	Race Meeting 31								
	Predicted Results					**Actual Results**			
Race	1st	2nd	3rd	4th		1st	2nd	3rd	4th
1	7	1	3	2		2	1	7	3
2	9	5	3	7		7	3	9	2
3	5	6	7	1		4	5	6	8
4	8	2	4	1		2	1	4	8
5	7	2	5	4		7	6	1	5
6	2	3	4	7		2	6	3	5
7	7	1	8	6		1	7	5	3
8	8	7	5	1		6	7	1	5
9	4	5	3	8		5	4	9	8

Table 1 shows the predictions and results for the 31st race meeting. Out of 9 races, in only 2 races, the winners were correctly predicted. Three of the predicted winners finished in the second place, 2 finished third, 1 finished fourth and 1 did not make it in the quartet. Thus, 77.7 % of the horses predicted to come first came into the first three places.

TABLE 2. PREDICTED RESULTS FOR MEETING 32

	Race Meeting 32								
	Predicted Results					**Actual Results**			
Race	1st	2nd	3rd	4th		1st	2nd	3rd	4th
1	5	4	8	2		4	1	7	5
2	6	1	8	4		6	1	3	4
3	7	6	9	1|4		1	2	3	6
4	1	4	3	6		2	3	7	1
5	3	5	2	4		5	1	4	6
6	8	7|9		4		6	5	9	1
7	1	2	3	6		3	5	6	2
8	8	7	4	9		4	5	1	8
9	6	3	4	7		4	6	7	1

Table 2 shows the results and predictions made for the 32nd race meeting. 1 of the 9 predicted winners was successful and 1 finished in the places. The other 7 predicted winners finished outside the top 3 places. Prediction for this race meeting was quite unsuccessful with only 22.2% of the predicted winners finishing among the top 3.

ISBN: 978-1-941968-14-7 ©2015 SDIWC

TABLE 3. PREDICTED RESULTS FOR MEETING 33

	Race Meeting 33								
	Predicted Results					Actual Results			
Race	1st	2nd	3rd	4th		1st	2nd	3rd	4th
1	6	3	10	9		6	2	4	3
2	1	4	8	5		7	1	4	9
3	4	1	5	2		4	2	5	1
4	4	1	7	5		4	7	5	1
5	5	1	9	2		2	4	1	7
6	8	3	1	2		5	1	6	7
7	4	5	7	6		2	4	5	6
8	6	3	1	7		3	8	6	5
9	3	8	4	5|10		8	3	1	5

Table 3 shows the results and predictions for the 33rd race meeting. 3 winners were correctly predicted out of 9. 3 finished second and 1 finished third. 2 of the predicted winners did not make it in the trifecta or quartet. The percentage of predicted winners finishing placed amounted to 77.7%. The percentage of success is quite low for the 32nd race meeting. However a minority of race meetings with less than 50% successful predictions is acceptable. More tests should be performed to find out whether the majority of predictions for several race meetings are above 70%.

5 CONCLUSIONS

The aim of the paper was to assess the appropriateness of Fuzzy Logic to predict winners at the Champ de Mars. Five parameters which were believed to be influential to the outcome of a race were considered. The system performs quite well even with a small set of 35 fuzzy rules. The system has proved to be very efficient when it comes to predicting horses finishing in the first 3 places. A score of 77.7 % for 2 race meetings out of 3 makes the system quite reliable for place bets. The research puts more emphasis on catching horses to place rather than win. Finding place bets is more easy and less risky especially in horseracing, where so many variables exist, they keep changing every time and horseracing has the characteristic of being unpredictable most of the time. Future works will involve the addition and improvement of fuzzy rules and investigation on parameters which could have major influence on the outcome of a race and hence be considered as inputs to the system. The risk in making the system more complex might be that its performance deviates from the reasonable one it produces right now because more parameters involves a larger space where a larger margin of error might exist.

The Fuzzy Logic technique could also be improved by combining it with Artificial Neural Network and rules could be automatically generated and tuned as large data sets are used to train the neural networks. The knowledge gained from this research could be easily adapted to other areas like decision-making in business, automating control of machines in industries, selection of construction models and many more areas where expert advice and experience is very important.

REFERENCES

[1] L. A. Zadeh, 1965. Fuzzy sets. *Information and Control*, Vol. 8, No. 3, pp. 338-353.

[2] A. Homaifar and E. McCormick, 1995. Simultaneous design of membership functions and rule sets for fuzzy controllers using genetic algorithms. *IEEE Transaction on Fuzzy Systems*, Vol. 3, No. 2, pp. 129-139.

[3] G. Hayward and V. Davidson, 2003. Fuzzy logic applications. *Analyst*, Vol. 128, No. 11, pp. 1304–1306.

[4] N. Nedjah and L. de M. Mourelle, 2005. *Fuzzy System Engineering: Theory and Practice*. 1st ed. Berlin: Springer-Verlag Berlin Heidelberg.

[5] S. Kumar, N. Bhatia and N. Kapoor, 2011. Fuzzy logic based tool for loan risk prediction. In: ed. 2nd *International Conference on Computer and Communication Technology*, 15-17 September 2011 India. New York: IEEE, pp. 180-183.

[6] S. Kumar, N. Bhatia and N. Kapoor, 2011. Software risk analysis using fuzzy logic. *International Journal of Computer Information Systems*, Vol. 2, No. 2, pp. 7-12.

[7] R. Bartlett, 2003. "The science and medicine of cricket: an overview and update". *Journal of Sports Sciences*, Vol. 21, No. 9, pp. 733-752.

[8] R. Bartlett, 2006. "Artificial intelligence in sports biomechanics: New dawn or false hope?" *Journal of Sports Science and Medicine*, Vol. 5, No. 4, pp. 474-479.

[9] K. Curtis, 2009. Cricket batting technique analyzer/trainer using fuzzy logic. In: ed. 16th *International Conference on Digital Signal Processing*, 5-7 July 2009 Greece. New York: IEEE, pp. 1056-1061.

[10] K. Curtis, 2010. Cricket Batting Technique Analyzer/Trainer: A Proposed Solution using Fuzzy Set Theory to Assist West Indies Cricket. In: ed. 9th *WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases*, 20 February 2010 United Kingdom. Wisconsin: WSEAS, pp. 71-76.

[11] G. Singh, N. Bhatia, and S. Singh, 2011. Fuzzy Logic Based Cricket Player Performance Evaluator. *IJCA Special Issue on Artificial Intelligence Techniques - Novel Approaches & Practical Applications*, Vo1.1, No. 3, pp. 11-16.

[12] D. Olaru, and B. Smith, 2002. Fuzzy logic models for activity schedules. In: ed. 24th *Conference of Australian Institutes of Transportation Research* pp. 1-9, 4-6 December 2002 Sydney. NSW: Oceania Publications.

[13] D. Olaru, and B. Smith, 2003. Modeling daily activity schedules with fuzzy logic. *Transportation*, Vol. 32, No. 4, pp. 423-441

[14] M. Heller, and K. Witte, 2006. A Dynamic Approach for Modeling and Simulation of Motor Unit Discharge Behavior Using Recurrent Fuzzy-Techniques. *International Journal of Computer Science in Sport*, Vol. 5, No. 1, pp. 30-40.

[15] E. Mezyk and O. Unold, 2011. Machine learning approach to model sport training. *Computers in Human Behavior*, Vol. 27, No. 5, pp. 1499-1506.

[16] N. Silverman, 2013. *Optimal Decisions with Multiple Agents of Varying Performance*. Thesis (PhD). University of California

[17] J. William and Y. Li, 2008. A Case Study Using Neural Network Algorithms: Horse Racing Predictions in Jamaica. In: ed. 8th *International Conference on Artificial Intelligence*, 14-17 July 2008 Las Vegas. : Aalborg: CSREA Press, pp. 14-17.

[18] S. Pudaruth, N. Medard and Z. B. Dookhun, 2013. Horse Racing Prediction at the Champs de Mars using a Weighted Probabilistic Approach, International Journal of Computer Applications, Vol. 72, No. 5, pp. 39-42.

[19] S. Pudaruth, R. Seesaha and L. Rambacussing, 2013. Generating Horse Racing Tips at the Champs de Mars using Fuzzy Logic, *International Journal of Computer Science and Technology*, Vol. 4, No. 3, pp. 7-11.

[20] S. N. Sivanandam, S. Sumathi and S. N. Deepa, 2007. *Introduction to Fuzzy Logic using MATLAB*. 1st ed. Berlin: Springer-Verlag Berlin Heidelberg.