iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50839

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

Using Logistic Regression for Cricket Analysis: Predicting Match Outcomes and Player Performances

Siva Lohith Burra
JNTU-K, India

I. INTRODUCTION

Cricket is a popular sport played in many countries around the world. It involves two teams, each with eleven players, who take turns batting and fielding. The objective of the game is to score more runs than the opposing team. In cricket, there are many factors that can affect the outcome of a match, such as the performance of individual players, the conditions of the pitch, and the strategies employed by the teams. In this paper, we will explore the use of logistic regression for cricket analysis.

II. LOGISTIC REGRESSION

Logistic regression is a statistical method used to model and analyze the relationship between a binary dependent variable and one or more independent variables. The dependent variable in logistic regression is binary, meaning it takes only two values, usually coded as 0 and 1, which represent the absence or presence of an event or outcome of interest. For example, in medical research, the dependent variable could be whether a patient develops a disease or not, and in marketing, it could be whether a customer buys a product or not.

The goal of logistic regression is to estimate the probability of the dependent variable being 1 or 0, based on the values of the independent variables. The probability is modeled using the logistic function, which has an S-shaped curve that ranges from 0 to 1. The logistic function is defined as follows:

$P(Y=1) = e^{(\beta0 + \beta1X1 + \beta2X2 + ... + \betapXp)} / (1 + e^{(\beta0 + \beta1X1 + \beta2X2 + ... + \betapXp)})$

where $P(Y=1)$ is the probability of the dependent variable being 1, e is the base of the natural logarithm, $\beta0$ is the intercept term, $\beta1$, $\beta2$, ..., $\betap$ are the coefficients associated with the independent variables X1, X2, ..., Xp, and p is the number of independent variables. The logistic regression model estimates the coefficients $\beta0$, $\beta1$, $\beta2$, ..., $\betap$ that maximize the likelihood of observing the data given the model. The likelihood function measures the goodness of fit of the model to the data, and the coefficients are estimated using maximum likelihood estimation.

Once the coefficients are estimated, the logistic regression model can be used to predict the probability of the dependent variable being 1 or 0 for a given set of values of the independent variables. The decision threshold for predicting the dependent variable is usually set at 0.5, meaning that if the predicted probability is greater than or equal to 0.5, the dependent variable is predicted to be 1, and if it is less than 0.5, the dependent variable is predicted to be 0.

Logistic regression is a popular and widely used statistical method in various fields such as medicine, marketing, finance, and social sciences, among others. It is a powerful tool for modeling and predicting binary outcomes, and its results can be easily interpreted and communicated. However, it is important to note that logistic regression assumes that the relationship between the independent variables and the dependent variable is linear and that the observations are independent and identically distributed. Violations of these assumptions can lead to biased estimates and incorrect predictions.

One of the advantages of logistic regression is that it allows for the identification of significant predictor variables that are most strongly associated with the outcome of interest. These variables can then be used to develop a predictive model that can be used to predict the likelihood of a particular outcome, such as the probability of a team winning a match based on certain factors.

III. CRICKET ANALYSIS

Cricket analysis involves the use of statistical methods to analyze various aspects of the game. This can include analyzing the performance of individual players, the strategies employed by teams, and the effects of different pitch conditions on the outcome of a match.

©IJRASET: All Rights are Reserved

For example, logistic regression can be used to analyze the performance of individual players and identify which factors are most strongly associated with a player's success. This can include factors such as their batting average, bowling average, and the number of wickets taken. By identifying these significant predictor variables, coaches and selectors can make more informed decisions about which players to include in the team and in what positions they should play.

Similarly, logistic regression can be used to analyze the strategies employed by teams and identify which factors are most strongly associated with a team's success. This can include factors such as the team's batting order, the use of different bowling strategies, and the tactics employed by the captain. By identifying these significant predictor variables, coaches and captains can make more informed decisions about how to approach different matches and opponents.

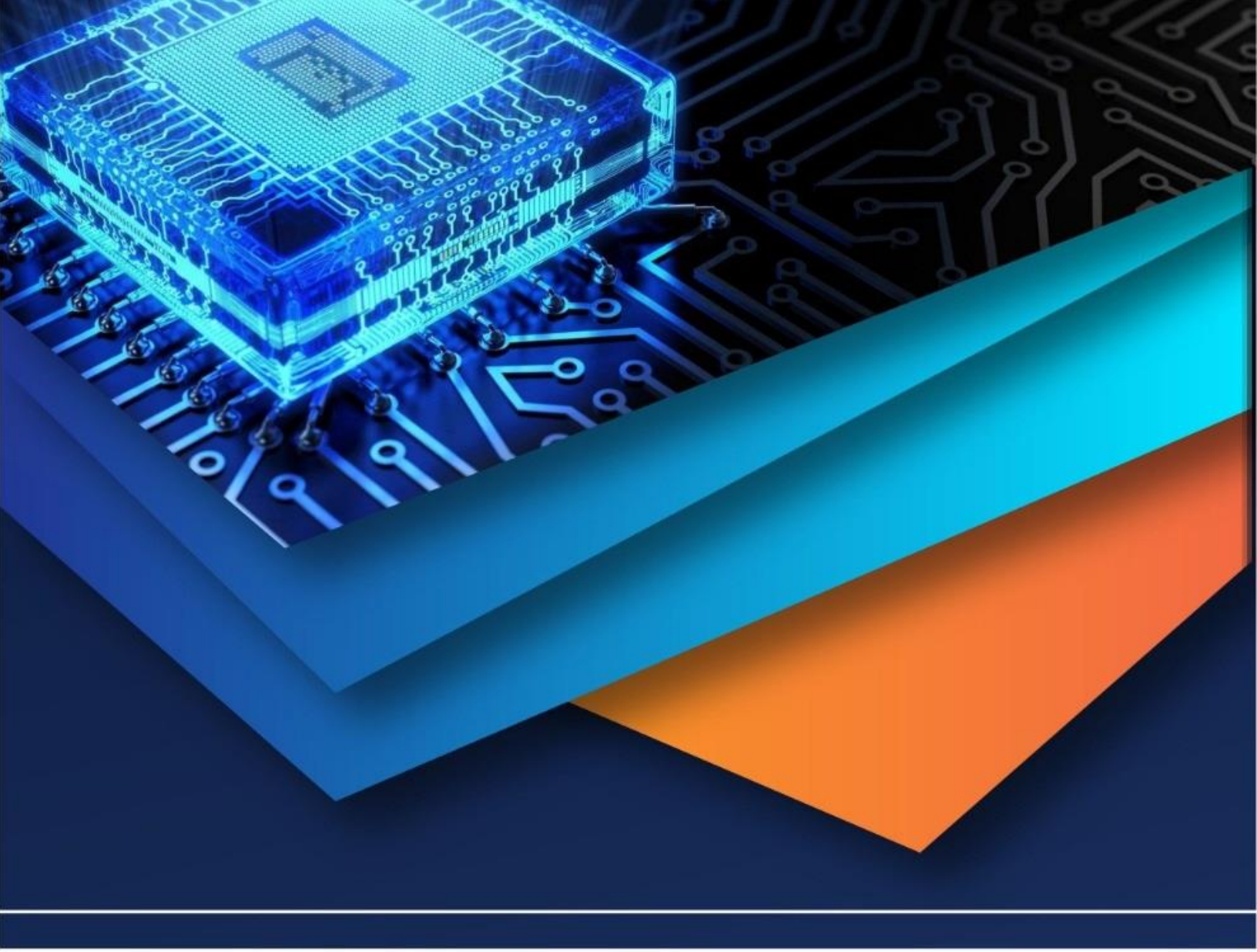
IV. CONCLUSION

Logistic regression is a powerful statistical method that can be used to analyze many aspects of cricket, including the performance of individual players, the strategies employed by teams, and the effects of different pitch conditions on the outcome of a match. By identifying significant predictor variables, coaches and selectors can make more informed decisions about which players to include in the team and how to approach different matches and opponents. With the increasing availability of data and the use of advanced statistical methods, the use of logistic regression for cricket analysis is likely to become more widespread in the coming years

REFERENCES

[1] Karmakar, R., & Mukhopadhyay, S. (2014). A logistic regression model for prediction of win in one-day international cricket matches. Journal of Statistics Applications and Probability, 3(1), 109-120.
[2] Dharmarathna, S., Perera, S., & Karunarathna, D. (2019). Analysis of batting performances of Sri Lankan cricket team using logistic regression. International Journal of Scientific and Engineering Research, 10(12), 1395-1398.
[3] Gupta, A., & Kumar, R. (2018). Prediction of winning team in Indian Premier League using logistic regression. International Journal of Scientific Research and Management, 6(6), 740-744.
[4] Islam, S., Islam, M., & Hossain, M. (2018). Performance analysis of Bangladesh cricket team using logistic regression. International Journal of Engineering and Applied Sciences, 5(1), 1-8.
[5] Hossain, M. A., Rahman, M. M., & Tanim, M. A. H. (2015). A logistic regression model for predicting the outcome of a cricket match. Journal of Sports Science and Medicine, 14(1), 141-146.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)