

**ENSEMBLE APPROACH TO PREDICT CRICKET MATCH  
SCORE BEFORE A MATCH USING MACHINE  
LEARNING TECHNIQUES**

**Subodha Didulaka Pathirana**

A dissertation submitted in partial fulfilment of the requirement for Bachelor of  
Engineering (Honours) degree in Software Engineering

**Department of Computer Science and Engineering  
Informatics Institute of Technology, Sri Lanka  
in collaboration with  
University of Westminster, UK**

**2020**

## Abstract

Sport activity is a competitive activity in which individuals or teams aim to use, maintain, or improve physical ability and skills providing entertainment to participants and spectators. Cricket is the most famous game in South Asia and the second most famous game in the world.

The main problem this research address is how to predict the final score of a cricket match as early as possible. The identification is based on factors such as team, opposition, venue, players' performance etc. Since lot of statics needs to be done manually for each player and team, it is very time consuming. Existing systems lack in producing score predictions and predicting the scores before the matches. Therefore, to predict the scores before the match an approach was proposed which considers player performance and team related factors.

The problem was divided into two parts as identifying factors that affect match score and determining the suitable algorithm for the prediction. Identification of the factors was done using existing systems, surveys, and feature selection techniques. Ensemble Learning was selected as the machine learning approach to address this problem.

This research identified team, opposition team, player batting index, player bowling index, ground, and run rate as the factors which will impact on the final scores. The regression ensemble model was created using a weight average technique for this purpose. The model gave an accuracy of 82%. The overall feedback of the experts' evaluation was positive and few suggested having a better ensemble approach as future enhancements.

*Keywords – Cricket, Regression, Ensemble Learning, Machine Learning, Batting Index, Bowling Index*