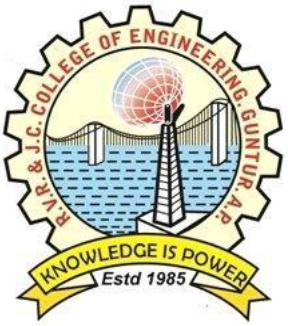
IPL SCORE PREDICTION

CB314 / CB351 - MACHINE LEARNING PROJECT
(III/IV B.Tech.)

Submitted in partial fulfilment of requirements for the award of the degree of
Computer Science and Business Systems Submitted By –

B Dileep (Y20CB010),
Md Abd AlRahman (Y20CB042),
T Praveen Kumar (Y20CB060)



Department of Computer Science and Business System

R.V.R. & J.C. COLLEGE OF ENGINEERING

(Autonomous) (NAAC 'A+' grade)

(Approved by AICTE, Affiliated to Acharya Nagarjuna University)

Chandramoulipuram, Chowdavaram, Guntur-522019,

Andhra Pradesh, India.

ABSTRACT

This project develops a machine learning model to predict scores in the Indian Premier League (IPL) through data collection, cleaning, and pre-processing. Various algorithms are applied and the best-performing one is selected for testing, with performance evaluated through metrics such as accuracy, precision, and recall. The results show that the model can effectively predict scores with high accuracy and provide valuable insights for fans and fantasy cricket players.

TABLE OF CONTENTS

1. Introduction

IPL (Indian Premier League) score prediction is a project that aims to use statistical and machine learning techniques to predict the outcome of IPL cricket matches

The purpose of this project is to generate accurate and reliable predictions that can be used by fans, analysts, and bookmakers to gain insights into the potential outcome of matches

2. Technologies Used

Python for Machine Learning Algorithms

HTML, CSS and JavaScript to develop client-side web pages

Flask and Pickle for implementing the python file in Web by creating links between pages

3. Procedure

3.1. Data Gathering

Dataset consists of data about IPL Matches from 2008 to 2017 with some useful fields for this project such as:

- Date
- Bat team
- Bowl team
- Runs
- Wickets
- Overs
- Runs in last 5 overs
- Wickets in last 5 overs
- Total score

3.2. Data Pre-processing

- Removing unwanted fields from Data Frame
- Considering data related to 8 main teams and removing the rest from Data Frame
- Removing Data of initial 5 overs from Data Frame
- Converting Date field from String to Date Datatype in Data Frame
- Converting the Bat team and Bowl team data to categorical values using OneHotEncoding method

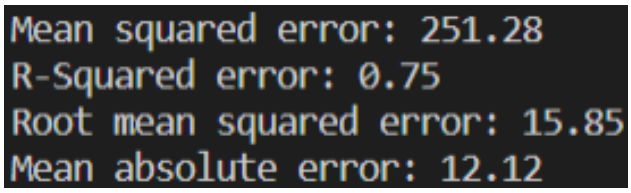
3.3. Model Building

Splitting Data Frame to train and test sets in which:

- Data of years less than 2017 in considered as training set
- Data of year 2017 is considered as testing set

3.4. Model Evaluation

After evaluating the built model, the results are as below:



3.5. Model Implementation

- Linear Regression is used to Predict the Range of Final Score of batting team based on inputs provided by user
- User inputs are taken from Web Application which uses Flask and Pickle to pass inputs from Client-side web page to Python file

3.6. Deployed Project Link

http://dpr106042.pythonanywhere.com/

4. Conclusion

Let us consider an IPL match SRH vs KKR held on April 15, 2022. Information related to it is as below:

- Batting team: SRH
- Bowling team: KKR
- Overs completed: 12
- Runs: 109
- Wickets: 2
- Runs scored in past 5 overs: 32
- Wickets taken in past 5 overs: 0

The Actual Final Score of SRH was 173/3

The Predicted Final Score of SRH by the above built model is "**The Final Score of Batting Team will be in Range 171 to 191**".