# Score Prediction of IPL Dataset

# Karthik P S(1NT18CS068)

# Nikhil Raj Koti L(1NT18CS107)

**Abstract**

# A decision tree is a tree with core nodes that may be used as tests (on data patterns) and leaf nodes that can be used as categories (of these patterns). To acquire the proper output to the input pattern, these checks are filtered down through the tree. Decision Tree methods are applicable and useful in a wide range of industries. It can be used to replace statistical processes for finding data, extracting text, finding missing data in a class, and improving search engines, as well as in medical disciplines. There has been a slew of decision tree algorithms proposed. They differ in terms of precision and cost-effectiveness. It's also critical for us to understand which algorithm is the best to apply.

**Introduction**

Machine learning (ML) is **a type of artificial intelligence (AI)** that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values. The proposal of our project is to predict the scores of a team batting first based on the previous year observations using various models such as Decision Tree Regression, Linear Regression, Support Vector Regression.

# Data Set

The dataset is chosen from Kaggle website which consisting of set of attributes such as:

* Id: Unique match id.

• Sepal\_length: Length of the petal.

• Sepal\_width: Width of the petal.

• Petal length: Length of the leaf.

• Petal width : Width of the leaf.

• species: Type of flower.

The challenge is lot of information is available which may not be necessary to predict our model we need to identify such attributes and trim them and pre-processing of data is required.

# Machine Learning Methods

# 1)Decision Tree Algorithm:

A Decision Tree has many analogies in real life and turns out, it has influenced a wide area of Machine Learning, covering both Classification and Regression. In decision analysis, a decision tree can be used to visually and explicitly represent decisions and decision making. Based on the fitting data we are going to calculate the accuracy and error in the model.

**2)Linear Regression:-**

linear regression is a linear approach for modelling the relationship between a scalar response and one or more explanatory variables. We are going to choose the dependent variable and predict the accuracy and error in the model.

3)**Support Vector Machine Regression:**

Support Vector Machines (SVMs) are well known in classification problems. We are going to explore more on support vectors and calculate the accuracy and error in the model

Finally, we choose a model with less error and more accuracy.

# Assessment:-

Train and test split

k-Fold Cross-Validation

# Presentation and Visualization

We will be showing both graph and prediction output.

**Roles**

Each one of us are dividing the work equally.my role to collect the dataset trim the un-required attribute and perform some pre-processing of data, train the decision tree regression model calculate the accuracy and errors in the model. Team-mate role is to train the other two model and calculate the accuracy and error in the model. Finally, we both can come to an conclusion that which model is has better accuracy and less error.

**Schedule:**

**Date Task to be completed**

23/12/2021 Download and observe the dataset.

28/12/2021 Learn about the model.

05/01/2022 Execution of the model and predicting the output.

11/01/2022 Based on the accuracy, choosing the best model for

Prediction.

**Bibliography**

* **https://www.kaggle.com/uciml/iris**
* [**https://towardsdatascience.com/an-introduction-to-support-vector-regression-svr-a3ebc1672c2**](https://towardsdatascience.com/an-introduction-to-support-vector-regression-svr-a3ebc1672c2)