Cricket Prediction Trainer

**Problem Statement:**

The problem statement was to predict weather the team will win the match or not.

**Methodology:**

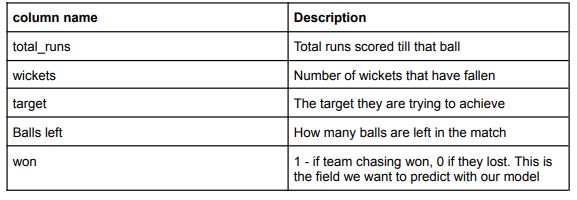
I use crisp dm methodology because this approach fits according to the given requirement.

* Business Understanding
* Data Understanding
* Data Preparation
* Modeling
* Evaluation

**Task 1:**

Task 1 was to develop a model and evaluate a model that can predict the match will won by the specific team or not.

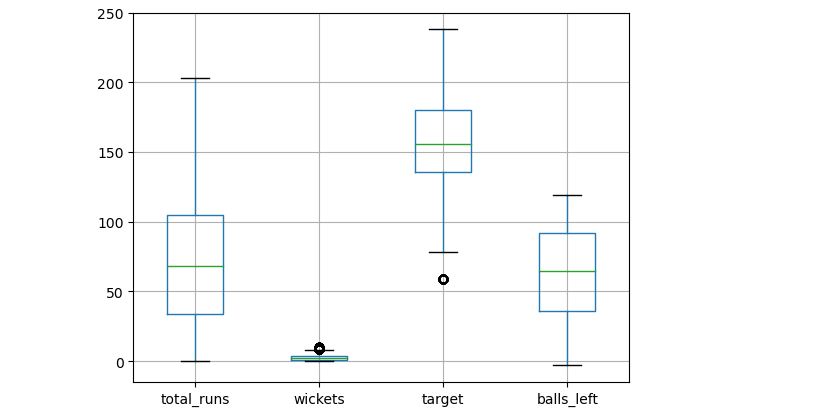
**Columns:**



**Model Building and Evaluation :**

**What Model is Use Why I use it ?**

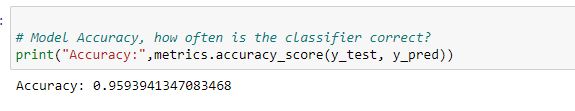
* I use Random forest as we see we have two columns contain outliers one is **wicket** and the other one is **target** that why I decided to use random forest because they **are robust to outliers** and they also do not over fit the data that two condition make our results better and generalize.

****

**Model Evaluation :**

**Model Accuracy:**

As model accuracy was good but we do not want to relay on that because that classes are not equally distributed. we also have to check other evaluation techniques.

****

**Confusion Matrix :**

I use confusion matrix to check how m model is performing it tells us quickly that our values are falling into right matrix or not these matrix are **tp,fp,tn,fn**.

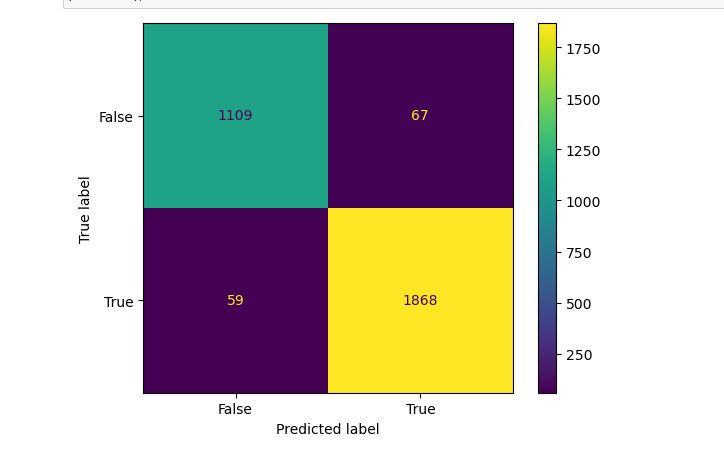
**As in my case my model is giving good results.**

**Tp defines : Match will won by team will actually won .**

**Fp defines : Match will not won by team but our model predict it will won by the team.**

**Tn defines : Match will not won by the team and it will not .**

**Fn defines : Match will won by team but our model predict it will not won by the team.**

****

**Roc And Auc :**

Roc and Auc difference between fpr and tpr rate it actually defines how well our model differentiate between 0 class and 1 class if it value is near to 1 we can score it as a good model .

