

ABSTRAK

Nama : M Sukron Martua

NIM : 1157010036

Judul : Penerapan Metode *Random Forest* untuk Prediksi Hasil Pertandingan Sepakbola Dengan Menggunakan Data Asli dan Data *Oversampling*

Dalam pertandingan sepakbola, ada beberapa faktor yang mempengaruhi kemenangan suatu tim dalam memenangkan pertandingan. Maka dengan hal itu perlu adanya klasifikasi untuk memudahkan melihat faktor yang berpengaruh terhadap kemenangan suatu tim. Ada beberapa metode yang dapat digunakan untuk klasifikasi, salah satunya adalah dengan menggunakan metode *random forest*. Metode ini dapat mengurutkan mana variabel yang berpengaruh dibandingkan dengan variabel yang lain berdasarkan dari data yang digunakan, serta menentukan seberapa besar nilai variabel tersebut untuk membuat peluang kemenangan tim menjadi lebih besar. Beberapa faktor yang mempengaruhi tim dalam memenangkan pertandingan adalah total tembakan, tembakan tepat sasaran, total serangan, tendangan sudut, penguasaan bola, serangan berbahaya, kartu kuning dan kartu merah. Dalam proses klasifikasi diperlukan adanya perbandingan untuk melihat hasil yang terbaik dalam proses analisis. Data variabel yang tidak seimbang dilakukan metode penyeimbangan data menggunakan metode *oversampling*. Sehingga dari perbandingan tersebut di dapatlah metode yang lebih baik untuk proses klasifikasi dalam pertandingan sepakbola.

Kata kunci : Sepakbola, Klasifikasi, *Random Forest*, *Oversampling*

ABSTRACT

Name : M Sukron Martua

NIM : 1157010036

Title : ***Application of the Random Forest Method for Predicting the Results of Football Matches Using Original Data and Oversampling Data***

In a football match, there are several factors that influence a team's victory in winning the match. So with that it is necessary to have a classification to make it easier to see the factors that influence the victory of a team. There are several methods that can be used for classification, one of which is using the random forest method. This method can sort which influential variables are compared to other variables based on the data used, and determine how big the value of these variables is to make the team's chances of winning bigger. Some of the factors that influence the team in winning the match are total shots, shots on target, total attacks, corners, ball possession, dangerous attacks, yellow cards and red cards. In the classification process, a comparison is needed to see the best results in the analysis process. Variable data that is not balanced is carried out by the data balancing method using the oversampling method. So that from this comparison a better method can be obtained for the classification process in football matches.

Keywords : Soccer, Classification, Random Forest, Oversampling

