

ABSTRAK

Sepak bola merupakan olahraga paling populer yang ditonton dan dimainkan oleh jutaan orang di seluruh dunia. Berdasarkan hasil survei yang dilakukan oleh *Fédération Internationale de Football Association* (FIFA) pada tahun 2001, sepak bola ditemukan sebagai olahraga paling populer yang dimainkan sampai hari ini.

Manager dan pelatih klub sepak bola tentunya menginginkan timnya dapat meraih kemenangan untuk menjadi juara. Tetapi untuk mencapai itu semua dibutuhkan kekompakan antar pemain dalam satu tim. Kemenangan setiap pertandingan sepakbola juga tergantung pada pemilihan pemain yang tepat.

Adapun tahapan yang dilakukan pada penelitian ini yaitu mengambil data hasil segmentasi, lalu melakukan proses *preprocessing* yang meliputi data *cleaning*, data *selection*, dan data *transformation*. Metode yang digunakan adalah metode klasifikasi dengan menerapkan algoritma *naïve bayes*. Kemudian mengukur hasil klasifikasi dengan menggunakan *confusion matrix*.

Percobaan dilakukan pada 532 data dengan 16 atribut dan 1 label klasifikasi. Atribut tersebut yaitu *name*, *club*, *nationality*, *position*, *age*, *matches*, *starts*, *mins*, *goals*, *assists*, *passes_attempted*, *perc_passes_completed*, *penalty_goals*, *penalty_attempted*, *xG*, *xA* dan 1 label klasifikasi yaitu *player_recruitment*. Hasil percobaan dengan uji akurasi menggunakan *3-fold Cross Validation*, *5-Fold Cross Validation* dan *7-Fold Cross Validation* mendapatkan akurasi tertinggi sebesar 90.7895 % tanpa *preprocessing* dan 90.4144 % dengan *preprocessing*.

Kata kunci: Prediksi, Sepak Bola, *Naïve Bayes*, *Confusion Matrix*, *Cross Validation*.

ABSTRACT

Football is the most popular sport watched and played by millions of people around the world. Based on the results of a survey conducted by the Fédération Internationale de Football Association (FIFA) in 2001, football was found to be the most popular sport played to this day.

Football club managers and coaches certainly want their teams to win to become champions. But to achieve that all it takes cohesiveness between players in one team. The victory of every football match also depends on choosing the right players.

The stages carried out in this study were to collect data from segmentation results, then carry out the preprocessing process which includes data cleaning, data selection, and data transformation. The method used is a classification method by applying the Naïve Bayes algorithm. Then measure the results of the classification using the confusion matrix.

The experiment was conducted on 532 data with 16 attributes and 1 classification label. The attributes are name, club, nationality, position, age, matches, starts, mins, goals, assists, passes_attempted, perc_passes_completed, penalty_goals, penalty_attempted, xG, xA and 1 classification label, namely player_recruitment. Experimental results with accuracy tests using 3-fold Cross Validation, 5-Fold Cross Validation and 7-Fold Cross Validation obtained the highest accuracy of 90.7895% without preprocessing and 90.4144% with preprocessing.

Keywords: Prediction, Football, Naïve Bayes, Confusion Matrix, Cross Validation.