

ABSTRAK

Penempatan posisi pemain yang tepat merupakan faktor kritis dalam kesuksesan tim sepak bola, namun ketidaktepatan sering terjadi akibat kurangnya pendekatan berbasis data. Penelitian ini bertujuan mengoptimalkan klasifikasi posisi pemain La Liga menggunakan algoritma *Logistic Regression* dengan hyperparameter tuning. Dataset diambil dari EA FC SPORTS 25, mencakup 529 pemain dan 29 atribut statistik (kecepatan, tembakan, bertahan) sebagai fitur. Metode penelitian meliputi: (1) preprocessing data (standarisasi fitur numerik), (2) pembagian data dengan rasio 80:20, 70:30, 60:40, dan 90:10, (3) pelatihan model dengan Grid Search untuk optimasi parameter (C: 0.01–100, solver: lbfgs/liblinear, max_iter: 200–1000), serta (4) evaluasi menggunakan akurasi, presisi, recall, dan F1-score. Hasil menunjukkan kombinasi terbaik pada rasio 80:20 dengan C=1, solver=liblinear, dan max_iter=200, menghasilkan akurasi 92,45%, presisi 92,64%, recall 92,45%, dan F1-score 92,44%. Optimasi hyperparameter meningkatkan akurasi 2-4% dibanding konfigurasi default, sementara rasio 60:40 memberikan akurasi terendah (88,68%) akibat keterbatasan data latih. Penelitian membuktikan efektivitas Logistic Regression dalam klasifikasi multikelas berbasis atribut teknis pemain, sekaligus menegaskan pentingnya tuning dan proporsi data yang seimbang. Implikasi praktisnya, model ini dapat menjadi alat bantu objektif bagi pelatih dalam strategi penempatan pemain.

Kata Kunci: Klasifikasi, Logistic Regression, Hyperparameter tuning.

ABSTRACT

Optimal player positioning is a critical factor in football team success, yet inaccuracies often arise due to the lack of data-driven approaches. This study aims to optimize the classification of La Liga player positions using the Logistic Regression algorithm with hyperparameter tuning. The dataset, sourced from EA FC SPORTS 25, includes 529 players and 29 statistical attributes (e.g., speed, shooting, defense) as features. The methodology involves: (1) data preprocessing (standardizing numerical features), (2) data splitting with ratios of 80:20, 70:30, 60:40, and 90:10, (3) model training using Grid Search to optimize parameters (C : 0.01–100, solver: lbfgs/liblinear, max_iter: 200–1000), and (4) evaluation using accuracy, precision, recall, and F1-score. Results show the best combination at an 80:20 split with $C=1$, solver=liblinear, and max_iter=200, achieving 92.45% accuracy, 92.64% precision, 92.45% recall, and 92.44% F1-score. Hyperparameter tuning improved accuracy by 2-4% compared to default configurations, while the 60:40 ratio yielded the lowest accuracy (88.68%) due to limited training data. This study proves the effectiveness of Logistic Regression in multiclass classification based on technical player attributes and emphasizes the importance of tuning and balanced data proportions. Practically, this model can serve as an objective decision-support tool for coaches in strategic player positioning.

Keywords: Classification, Logistic Regression, Hyperparameter Tuning.