

PENERAPAN ALGORITMA GAUSSIAN NAÏVE BAYES UNTUK KLASIFIKASI FORMASI SEPAK BOLA BERDASARKAN DATA ATRIBUT PEMAIN

Nama Mahasiswa : Sandy Mulia Kesuma
NIM : 6304221493
Dosen Pembimbing : Fajri Profesio Putra, M.Cs

ABSTRAK

Penelitian ini menelaah penggunaan algoritma Gaussian Naïve Bayes untuk mengklasifikasikan formasi sepak bola dengan memanfaatkan atribut statistik pemain sebagai dasar analisis. Permasalahan utama muncul karena pelatih kerap menentukan formasi berdasarkan intuisi tanpa dukungan evaluasi data yang objektif, sehingga keputusan taktis berpotensi kurang optimal. Dataset penelitian diperoleh dari platform Kaggle dan mencakup atribut numerik seperti *Pace*, *Shooting*, *Passing*, *Dribbling*, *Defending*, dan *Physical* yang merepresentasikan performa individu secara menyeluruh. Proses analisis mengikuti tahapan Knowledge Discovery in Databases (KDD), meliputi pemilihan data, preprocessing, transformasi, dan pembangunan model klasifikasi. Setiap atribut utama diolah menjadi nilai agregat agar mampu menggambarkan kemampuan pemain secara lebih komprehensif. Model Gaussian Naïve Bayes kemudian dilatih dengan membandingkan nilai agregat tersebut terhadap referensi formasi yang telah ditentukan. Hasil penelitian menunjukkan bahwa model mampu mengelompokkan formasi secara konsisten dan memberikan rekomendasi berbasis data yang lebih objektif dibandingkan pendekatan tradisional. Temuan ini diharapkan dapat menjadi dasar pengembangan sistem pendukung keputusan bagi pelatih dan analis dalam konteks analisis olahraga.

Kata kunci : Gaussian Naïve Bayes, Formasi Sepak Bola, Atribut Statistik Pemain

***APPLICATION OF THE GAUSSIAN NAÏVE BAYES
ALGORITHM FOR FOOTBALL FORMATION
CLASSIFICATION BASED ON PLAYER ATTRIBUTE DATA***

Name Of Student : Sandy Mulia Kesuma
Student ID Number : 6304221493
Supervisor : Fajri Profesio Putra, M.Cs

ABSTRACT

This study investigates the use of the Gaussian Naïve Bayes algorithm to classify football formations based on a set of player statistical attributes. The main issue addressed arises from the tendency of coaches to determine formations intuitively without relying on systematic and objective data analysis, which may lead to suboptimal tactical decisions. The dataset was obtained from the Kaggle platform and includes numerical attributes such as Pace, Shooting, Passing, Dribbling, Defending, and Physical that collectively describe individual player performance. The analytical workflow follows the Knowledge Discovery in Databases (KDD) framework, consisting of data selection, preprocessing, transformation, and classification model development. Each key attribute was processed into an aggregated value to more comprehensively represent the players' capabilities. The Gaussian Naïve Bayes model was trained by comparing these aggregated attributes with predefined formation references. The results indicate that the model is capable of grouping formations consistently and generating data-driven recommendations that are more objective than traditional, intuition-based approaches. These findings are expected to contribute to the development of decision-support systems for coaches and analysts, particularly within the broader context of sports analytics.

Keyword : *Gaussian Naïve Bayes, Football Formations, Player Statistical Attributes*