

ABSTRACT

Jason Kristanto (01082180038)

PREDICTION OF ULTIMATE FIGHTING CHAMPIONSHIP YEAR (1994-2022) USING XGBOOST AND RANDOM FOREST ALGORITHMS

(xiii + 66 pages: 8 figures; 2 tables, 0 appendices)

The Ultimate Fighting Championship (UFC) is the world's largest mixed martial arts (MMA) organization, featuring fighters from diverse martial arts disciplines. In today's data-driven landscape, predicting UFC match outcomes has become increasingly important—not only for coaches and analysts but also for the sports betting industry and fans. The main challenge in making such predictions lies in the numerous statistical variables that influence fight outcomes and the complexity of evolving fight data. With the growing availability of historical fight data, machine learning (ML)-based predictive methods offer a promising solution that can provide higher accuracy than traditional statistical or intuition-based approaches.

This study aims to develop a predictive model for UFC fight outcomes by comparing the performance of two popular ML algorithms: Random Forest and Extreme Gradient Boosting (XGBoost). The research methodology consists of data preprocessing, feature engineering, skill and form score calculation, model training, hyperparameter tuning, and evaluation using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC.

Experimental results show that the XGBoost model achieved a prediction accuracy of 69.1%, slightly outperforming the Random Forest model, which reached 68.64%. The XGBoost model also attained an F1-score of 0.68 and a recall of 0.69, indicating balanced performance. In contrast, Random Forest obtained an F1-score of 0.67. These findings suggest that XGBoost is more effective at handling the complexity of UFC fight data and provides more stable predictions. Therefore, XGBoost is recommended for developing data-driven UFC fight outcome prediction systems.

Reference: 15 (2001 - 2024)

ABSTRAK

Jason Kristanto (01082180038)

PREDIKSI HASIL PERTANDINGAN ULTIMATE FIGHTING CHAMPIONSHIP TAHUN (1994-2022) MENGGUNAKAN ALGORITMA XGBOOST DAN RANDOM FOREST

(xiii + 66 halaman: 8 gambar; 2 tabel; 0 lampiran)

Ultimate Fighting Championship (UFC) merupakan organisasi olahraga bela diri campuran (Mixed Martial Arts/MMA) terbesar di dunia, yang menyajikan pertarungan antara petarung dari berbagai disiplin seni bela diri. Dalam dunia yang semakin data-driven, prediksi hasil pertandingan UFC menjadi penting tidak hanya bagi pelatih dan analis olahraga, tetapi juga bagi industri taruhan olahraga dan penggemar. Permasalahan utama dalam prediksi ini adalah banyaknya variabel statistik yang saling mempengaruhi dan kompleksitas data pertarungan yang terus berkembang. Dengan meningkatnya volume data historis dan statistik pertandingan, metode prediktif berbasis pembelajaran mesin (machine learning) menjadi solusi yang potensial untuk menghasilkan prediksi yang lebih akurat dibandingkan metode konvensional.

Penelitian ini bertujuan untuk membangun model prediksi hasil pertandingan UFC dengan membandingkan performa dua algoritma machine learning populer, yaitu Random Forest dan Extreme Gradient Boosting (XGBoost). Tahapan penelitian meliputi preprocessing data, feature engineering, pembuatan fitur skor keterampilan dan performa, pelatihan model, tuning hyperparameter, serta evaluasi menggunakan metrik seperti akurasi, precision, recall, F1-score, dan ROC-AUC. Penelitian ini juga melakukan analisis fitur penting yang berkontribusi terhadap prediksi hasil pertandingan.

Hasil eksperimen menunjukkan bahwa model XGBoost mampu menghasilkan akurasi prediksi sebesar 69,1%, sedikit lebih tinggi dibandingkan Random Forest yang mencapai 68,64%. Selain itu, XGBoost memiliki skor F1 sebesar 0,68 dengan recall 0,69, yang menunjukkan kinerja model yang seimbang antara sensitivitas dan presisi. Sementara Random Forest menghasilkan skor F1 sebesar 0,67. Temuan ini menunjukkan bahwa XGBoost memiliki keunggulan dalam mengolah kompleksitas data pertarungan UFC dan mampu memberikan prediksi yang lebih stabil. Dengan demikian, model berbasis XGBoost direkomendasikan untuk digunakan dalam sistem prediksi hasil pertandingan UFC berbasis data statistik.

Referensi: 15 (2001 – 2024)