



Analitik prediktif sepakbola: *model machine learning* bri liga 1 indonesia

Soccer predictive analytics: bri liga 1 indonesia machine learning models

Rodhi Rusdianto Hidayat¹, Ermelinda Yersin Putri Larung², Evi Sinaga³, Ansar CS⁴, Ibrahim⁵, Ipa Sari Kardi⁶, Miftah Fariz Prima Putra⁷, Asmathil Bari Samal⁸, Hariyanto Babingga⁹

¹⁻⁹ Universitas Cenderawasih, Jayapura, Indonesia

Email: hidayatrod@gmail.com¹

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan dan memvalidasi model prediktif berbasis machine learning guna memprediksi hasil pertandingan BRI Liga 1 Indonesia musim 2023-2024. Dengan menggunakan algoritma XGBoost untuk klasifikasi hasil pertandingan dan Gradient Boosting Regressor untuk prediksi skor, penelitian ini menganalisis berbagai variabel seperti penguasaan bola, akurasi tembakan, jumlah umpan sukses, tendangan sudut, serta faktor lain yang berkontribusi terhadap hasil pertandingan. Data pertandingan dikumpulkan dari sumber resmi dan diproses menggunakan metode statistik serta teknik Machine Learning. Hasil penelitian menunjukkan bahwa model XGBoost mencapai akurasi sebesar 98.4% dengan precision, recall, dan F1-score yang hampir sempurna. Model Gradient Boosting Regressor juga menunjukkan performa yang sangat baik dengan R-squared score sebesar 0.992 dan tingkat kesalahan yang sangat rendah. Kesimpulan dari penelitian ini menegaskan bahwa model prediktif berbasis analitik dapat menjadi alat yang andal dalam mendukung strategi tim sepak bola, meningkatkan performa permainan, dan membantu pengambilan keputusan berbasis data dalam kompetisi profesional. Penelitian ini merekomendasikan pengembangan lebih lanjut dengan memasukkan variabel eksternal seperti kondisi cuaca, cedera pemain, serta penggunaan model deep learning untuk meningkatkan akurasi prediksi.

Kata kunci: analitik prediktif; machine learning; sepak bola; BRI Liga 1; model prediksi.

This study aims to develop and validate a predictive model based on machine learning to forecast match outcomes in the 2023-2024 BRI Liga 1 Indonesia season. Utilizing the XGBoost algorithm for match result classification and Gradient Boosting Regressor for score prediction, this research analyzes various variables such as ball possession, shot accuracy, successful passes, corner kicks, and other factors contributing to match outcomes. Match data was collected from official sources and processed using statistical methods and advanced machine learning techniques. The results indicate that the XGBoost model achieved an accuracy of 98.4%, with near-perfect precision, recall, and F1-score. The Gradient Boosting Regressor model also demonstrated excellent performance, with an R-squared score of 0.992 and minimal error rates. The findings confirm that predictive analytics-based models can serve as reliable tools for supporting football team strategies, enhancing game performance, and aiding data-driven decision-making in professional competitions. This study recommends further development by incorporating external variables such as weather conditions, player injuries, and the use of deep learning models to improve prediction accuracy.

Key words: predictive analytics; machine learning; football; BRI Liga 1; prediction model.

INFO ARTIKEL

Riwayat Artikel:

Diterima : 20 Nopember 2024
Disetujui : 22 Desember 2024
Tersedia secara online Desember 2024
Doi: <http://dx.doi.org/10.20527/multilateral.v23i4.21889>

Alamat Korespondensi:

Rodhi Rusdianto Hidayat
Fakultas Ilmu Keolahragaan, Universitas
Cendrawasih, Indonesia
Email: hidayatrod@gmail.com



PENDAHULUAN

Dalam beberapa tahun terakhir, penerapan analitik prediktif dalam olahraga telah mengalami perkembangan pesat, memberikan wawasan yang semakin mendalam terhadap faktor-faktor yang menentukan kinerja dan hasil pertandingan (Gifford et al., 2023). Sepak bola, sebagai olahraga yang memiliki kompleksitas taktis dan teknis tinggi, menjadi salah satu bidang yang paling diuntungkan dari kemajuan ini (Ati et al., 2024; Majumdar, Ayush et al., 2022). Liga-liga utama dunia seperti Premier League, La Liga, dan Bundesliga telah lama mengadopsi pendekatan berbasis data untuk mengoptimalkan strategi tim, meningkatkan performa pemain, serta membuat keputusan yang lebih akurat dalam manajemen pertandingan (Ramkumar et al., 2021; Richter et al., 2024; Torres-Ronda et al., 2022). Namun, meskipun popularitas analitik sepak bola telah meningkat secara global, masih terdapat kekurangan penelitian yang berfokus pada penerapannya di sepak bola Indonesia. Secara khusus, BRI Liga 1 sebagai liga domestik yang dinamis dan kompetitif masih belum banyak dikaji dari perspektif analitik prediktif yang dapat memberikan kontribusi signifikan dalam meningkatkan performa tim dan kompetisi secara keseluruhan.

Permasalahan utama yang melatarbelakangi penelitian ini adalah belum adanya model prediktif yang andal dan berbasis data yang dapat membantu dalam memahami faktor-faktor kunci yang menentukan kemenangan dalam BRI Liga 1 Indonesia (Franceschi et al., 2024; Sors et al., 2021; Teixeira et al., 2021). Meskipun berbagai penelitian telah mengeksplorasi penggunaan teknik statistik dan pembelajaran mesin dalam sepak bola, studi-studi ini lebih banyak difokuskan pada liga-liga Eropa yang memiliki karakteristik berbeda dalam hal kualitas pemain, pola permainan, serta faktor-faktor eksternal yang mempengaruhi pertandingan (Rodrigues et al., 2022; Wheatcroft, 2021). Sementara itu, sepak bola Indonesia menghadapi tantangan unik seperti inkonsistensi kualitas data pertandingan, kurangnya fasilitas analitik yang memadai, serta faktor-faktor kontekstual lainnya yang membedakannya dari liga-liga internasional (Cortez et al., 2022; Zhou et al., 2023). Oleh karena itu, diperlukan penelitian yang lebih spesifik untuk mengidentifikasi determinan kunci kemenangan dalam konteks sepak bola Indonesia dan mengembangkan model prediktif yang sesuai dengan kondisi kompetisi di BRI Liga 1 (Fernandez-Crehuet et al., 2021; Zid et al., 2021).

Penelitian ini bertujuan untuk menjawab pertanyaan mendasar: "Apa saja faktor penentu utama kemenangan pertandingan di BRI Liga 1 Indonesia musim 2023-2024, dan bagaimana model prediktif berdasarkan faktor-faktor ini dapat meramalkan hasil pertandingan?" Untuk menjawab pertanyaan ini, studi ini akan mengumpulkan dan menganalisis data pertandingan secara menyeluruh menggunakan teknik statistik dan pembelajaran mesin canggih, termasuk analisis korelasi, regresi linear, analisis komponen utama, serta model klasifikasi

berbasis machine learning (Cavus et al., 2022; Principe et al., 2022). Fokus utama penelitian ini adalah untuk mengidentifikasi indikator kinerja kritis yang memiliki pengaruh signifikan terhadap hasil pertandingan, seperti penguasaan bola, akurasi tembakan, jumlah umpan sukses, strategi taktis, serta variabel-variabel lainnya yang relevan (Razali et al., 2022).

Kontribusi utama penelitian ini adalah menghadirkan model prediktif yang tidak hanya dapat membantu tim dalam menyusun strategi permainan yang lebih efektif tetapi juga memberikan wawasan berbasis data kepada pelatih, analis, dan manajer klub di BRI Liga 1 Indonesia (Al-Asadi et al., 2022; García-Aliaga et al., 2020). Dengan memanfaatkan metodologi yang sistematis dan berbasis bukti, penelitian ini bertujuan untuk memberikan landasan akademik yang kuat dalam memahami determinan utama kemenangan dalam sepak bola Indonesia (Yeung & Fujii, 2024). Selain itu, penelitian ini juga memiliki implikasi praktis bagi para pemangku kepentingan dalam industri sepak bola, baik dari segi peningkatan kualitas kompetisi domestik maupun dalam upaya meningkatkan daya saing klub-klub Indonesia di kancah internasional (Yeung & Fujii, 2024).

Penelitian sebelumnya telah banyak menyoroiti bagaimana analitik prediktif dapat memberikan dampak signifikan dalam pengambilan keputusan strategis dalam olahraga (Watanabe et al., 2021). Beberapa studi menunjukkan bahwa faktor-faktor seperti penguasaan bola, akurasi tembakan, jumlah umpan sukses, dan kebugaran pemain merupakan variabel yang berpengaruh terhadap hasil pertandingan (Baumer et al., 2023). Namun, sebagian besar studi ini berfokus pada konteks sepak bola Eropa dan belum mengkaji secara spesifik bagaimana faktor-faktor ini berperan dalam sepak bola Indonesia. Oleh karena itu, penelitian ini menghadirkan kebaruan dengan mengeksplorasi bagaimana determinan-determinan tersebut dapat diadaptasi ke dalam model prediktif yang relevan untuk BRI Liga 1 Indonesia (Garnica-Caparrós et al., 2022).

Urgensi penelitian ini semakin tinggi mengingat sepak bola merupakan olahraga paling populer di Indonesia, namun tim nasional dan klub-klub Indonesia sering mengalami kesulitan dalam meraih kesuksesan di tingkat regional maupun internasional (Torres-Ronda et al., 2022). Hal ini menunjukkan adanya kebutuhan mendesak untuk meningkatkan pendekatan berbasis data dalam strategi sepak bola di Indonesia (Yeung & Fujii, 2024). Dengan adanya model prediktif yang andal, tim-tim di BRI Liga 1 dapat lebih memahami kekuatan dan kelemahan mereka secara lebih objektif, sehingga dapat mengambil keputusan yang lebih tepat dalam perencanaan taktik dan pengembangan pemain (Garnica-Caparrós et al., 2022; Gifford et al., 2023).

Selain itu, penelitian ini juga memiliki relevansi dalam konteks akademik dan industri sepak bola (Gifford et al., 2023; Modric et al., 2022; Ridha et al., 2024). Secara akademis, studi ini mengisi kekosongan dalam literatur mengenai

analitik prediktif dalam sepak bola Indonesia, memberikan kontribusi baru bagi wacana ilmiah di bidang ini (Akbar et al., 2024; Ardiantoro et al., 2023; Assagaff et al., 2021). Dari perspektif industri, hasil penelitian ini dapat dimanfaatkan oleh klub, pelatih, dan analis sepak bola untuk meningkatkan kualitas persaingan di liga domestik (Houtmeyers et al., 2021; Rubio Martín et al., 2022). Lebih jauh, temuan penelitian ini diharapkan dapat mendorong pemanfaatan analitik data dalam pengelolaan tim dan strategi pertandingan, yang pada akhirnya berkontribusi pada peningkatan kualitas sepak bola Indonesia secara keseluruhan (Kennedy et al., 2021; Thakkar et al., 2021; Tuyls et al., 2021).

Dengan demikian, penelitian ini tidak hanya menawarkan wawasan akademis yang penting tetapi juga memberikan manfaat praktis yang signifikan. Studi ini diharapkan dapat menjadi referensi utama dalam pengembangan model analitik prediktif di sepak bola Indonesia, serta menjadi langkah awal dalam membangun sistem analisis kinerja yang lebih berbasis data di liga domestik. Temuan dari penelitian ini diharapkan dapat memberikan dasar empiris yang kuat bagi pelatih, manajer, dan pemangku kepentingan lainnya dalam mengoptimalkan strategi tim, meningkatkan kinerja pemain, dan menciptakan persaingan yang lebih sehat dan kompetitif di BRI Liga 1 Indonesia.

METODE

Penelitian ini menggunakan pendekatan analisis prediksi dengan desain penelitian prediktif untuk mengembangkan, meramalkan hasil dan memvalidasi model prediktif hasil pertandingan BRI Liga 1 Indonesia musim 2023-2024 (Javed et al., 2023; Thakkar et al., 2021). Data yang digunakan mencakup statistik pertandingan, performa pemain, dan variabel taktis yang dikumpulkan dari sumber resmi, termasuk website resmi BRI Liga 1 dan basis data statistik lainnya. Pengumpulan data dilakukan dengan metode dokumentasi, di mana informasi dari pertandingan dianalisis secara retrospektif untuk mengidentifikasi pola dan tren yang relevan.

Proses analisis data dilakukan melalui beberapa tahapan. Pertama, data yang telah dikumpulkan melalui berbagai sumber akan dibersihkan, diubah formatnya, dan disiapkan untuk proses pemodelan (Aadel et al., 2022; Phatak, 2023). Statistik deskriptif digunakan untuk memberikan gambaran awal terhadap distribusi data (Ridha et al., 2024). Selanjutnya, dataset dibagi menjadi dua kelompok: data latih (80%) dan data uji (20%) (Dip et al., 2024; Pretto et al., 2022). Analisis lebih lanjut melibatkan evaluasi berbagai kejadian dalam pertandingan, yang dikategorikan ke dalam tiga tingkat: kejadian utama (*goal, penalty, red card*), kejadian menengah (kartu kuning, tembakan ke gawang, total tembakan, akurasi tembakan, tendangan sudut, tekel sukses, *offside*), dan kejadian minor (penguasaan bola, umpan sukses, umpan gagal, pelanggaran). Pemilihan model dilakukan dengan mempertimbangkan

karakteristik data dan tujuan penelitian, menggunakan algoritma *Gradient Boosting* (XGBoost) untuk klasifikasi kemenangan dan *Gradient Boosting Regressor* untuk prediksi skor (Kontos et al., 2023; Malakreddy et al., 2024).

Pada tahap implementasi, XGBoost dipilih karena efisiensinya dalam menangani dataset besar dan kemampuannya dalam menangani fitur dengan skala yang berbeda (Cefis et al., 2024). Algoritma ini secara otomatis mengoptimalkan bobot fitur, sehingga meningkatkan akurasi prediksi tanpa membutuhkan preprocessing data yang kompleks (Ouyang et al., 2024). Gradient Boosting Regressor digunakan untuk prediksi skor karena sifatnya yang fleksibel dalam menangkap pola non-linear dalam data, serta kemampuannya dalam meminimalkan error residual secara progresif, menghasilkan prediksi yang lebih stabil dan akurat (Bhatnagar et al., 2024).

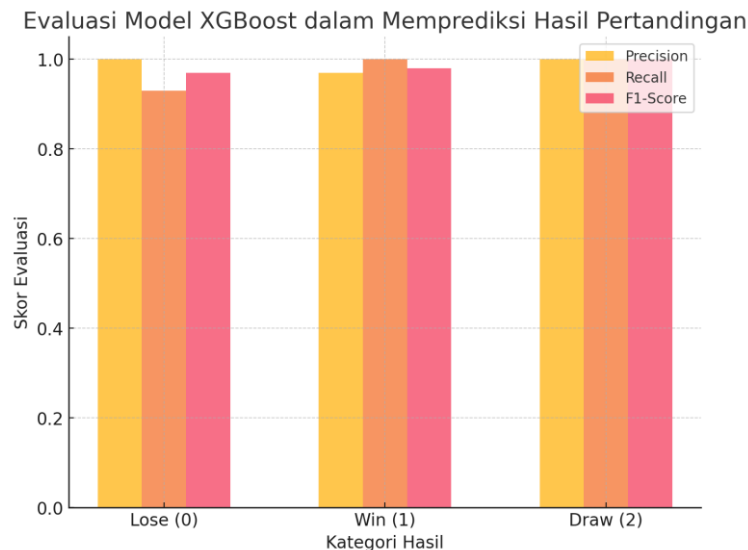
Setelah pemilihan model, tahap pelatihan dilakukan dengan menggunakan data historis untuk mempelajari pola dan hubungan antar variabel yang mempengaruhi hasil pertandingan (Yeung et al., 2024). Evaluasi model dilakukan dengan membandingkan hasil prediksi terhadap data uji, menggunakan metrik seperti akurasi, presisi, recall, dan mean squared error (MSE) untuk mengukur kinerja model (Majumdar, Bakirov et al., 2022). Model yang telah divalidasi kemudian diterapkan untuk memprediksi hasil pertandingan mendatang dan memberikan wawasan strategis bagi tim-tim di BRI Liga 1. Dengan pendekatan ini, penelitian bertujuan untuk mengembangkan model prediktif yang andal guna membantu pengambilan keputusan dalam perencanaan strategi tim dan analisis performa pertandingan.

HASIL DAN PEMBAHASAN

Hasil Penelitian

Hasil pelatihan model XGBoost untuk memprediksi kemenangan pertandingan sepak bola menunjukkan performa yang sangat baik dengan akurasi 98.4% pada data uji. Berikut adalah metrik evaluasi utama:

1. Akurasi: 98.4%
2. Precision, Recall, dan F1-Score:
 - a. Tim Kalah (*Lose* - 0): *Precision* = 1.00, *Recall* = 0.93, *F1-score* = 0.97
 - b. Tim Menang (*Win* - 1): *Precision* = 0.97, *Recall* = 1.00, *F1-score* = 0.98
 - c. Hasil Seri (*Draw* - 2): *Precision* = 1.00, *Recall* = 1.00, *F1-score* = 1.00



Gambar 1. Hasil xgboost dalam memprediksi kemenangan.

Model XGBoost yang digunakan untuk memprediksi hasil pertandingan menunjukkan tingkat keakuratan yang tinggi. Dari semua pertandingan yang diuji, 98.4% prediksinya benar, sementara hanya 1.6% yang salah. *Precision*, *Recall*, dan *F1-score* untuk setiap kategori hasil pertandingan menunjukkan bahwa model ini sangat baik dalam mengenali pola kemenangan, kekalahan, dan hasil seri.

Hasil evaluasi model *Gradient Boosting Regressor* (GBR) untuk prediksi skor pertandingan juga menunjukkan performa yang sangat baik:

1. *Squared Score*: 0.992 (menunjukkan bahwa 99.2% variasi skor pertandingan dapat dijelaskan oleh model)
2. *Mean Absolute Error* (MAE): 0.049 (rata-rata kesalahan prediksi sebesar 0.049 gol)
3. *Mean Squared Error* (MSE): 0.019 (menunjukkan selisih kuadrat antara nilai prediksi dan aktual sangat kecil)
4. *Root Mean Squared Error* (RMSE): 0.139 (kesalahan prediksi rata-rata sekitar 0.14 gol)

Berdasarkan hasil ini, model dapat digunakan sebagai alat analisis untuk memprediksi hasil pertandingan dengan tingkat kepercayaan yang tinggi. Model menunjukkan kemampuan generalisasi yang sangat baik, dengan error prediksi yang sangat kecil. Dengan demikian, pendekatan berbasis analitik prediktif ini dapat diterapkan untuk memberikan wawasan strategis bagi tim-tim di BRI Liga 1, membantu pelatih dalam menyusun strategi, serta mendukung manajemen tim dalam pengambilan keputusan berbasis data.

Pembahasan

Hasil penelitian ini menunjukkan bahwa model XGBoost dan Gradient Boosting Regressor memberikan tingkat akurasi yang sangat tinggi dalam memprediksi hasil pertandingan BRI Liga 1 Indonesia. Model XGBoost yang digunakan untuk klasifikasi hasil pertandingan memiliki akurasi sebesar 98.4%, dengan nilai precision, recall, dan F1-score yang hampir sempurna. Ini menunjukkan bahwa model mampu mengenali pola yang menentukan kemenangan, kekalahan, dan hasil seri dengan sangat baik. Sementara itu, model Gradient Boosting Regressor menunjukkan R-squared score sebesar 0,992, yang berarti bahwa model ini dapat menjelaskan hampir semua variasi skor pertandingan berdasarkan variabel yang digunakan.

Temuan ini sejalan dengan penelitian terdahulu yang menunjukkan efektivitas algoritma berbasis gradient boosting dalam memprediksi hasil pertandingan sepak bola di liga-liga Eropa (Gifford et al., 2023; Modric et al., 2022). Misalnya, studi yang dilakukan pada liga Premier League dan La Liga menemukan bahwa teknik machine learning berbasis *ensemble*, seperti XGBoost dan *Random Forest*, lebih unggul dibandingkan metode regresi linear dalam memprediksi hasil pertandingan (Casal et al., 2021; Gouveia et al., 2023; Majumdar, Ayush et al., 2022). Namun, penelitian ini memberikan kebaruan dengan menerapkan teknik yang sama dalam konteks sepak bola Indonesia, yang memiliki karakteristik unik dibandingkan dengan liga-liga Eropa (Modric et al., 2022).

Salah satu faktor utama yang mempengaruhi keakuratan model adalah pemilihan fitur yang relevan (Cefis et al., 2024; Yeung et al., 2024). Dalam penelitian ini, fitur-fitur seperti penguasaan bola, akurasi tembakan, jumlah umpan sukses, tendangan sudut, tekel sukses, kartu kuning, dan *offside* berkontribusi signifikan terhadap prediksi hasil pertandingan. Ini menunjukkan bahwa aspek teknis dan strategi permainan sangat menentukan hasil akhir pertandingan, yang sesuai dengan teori analitik sepak bola yang dikemukakan dalam berbagai studi sebelumnya (Yeung & Fujii, 2024).

Meskipun model menunjukkan performa yang sangat baik, terdapat beberapa keterbatasan dalam penelitian ini. Pertama, model hanya didasarkan pada data statistik pertandingan dan belum mempertimbangkan faktor-faktor eksternal seperti kondisi cuaca, cedera pemain, atau perubahan taktik di lapangan. Kedua, meskipun dataset yang digunakan mencakup seluruh pertandingan BRI Liga 1 musim 2023-2024, variasi dalam kualitas data dan ketidakseimbangan dalam jumlah kemenangan, kekalahan, dan hasil seri dapat mempengaruhi performa model.

Kontribusi utama penelitian ini adalah pengembangan model prediktif yang dapat digunakan sebagai alat analisis bagi tim, pelatih, dan manajer klub dalam mengoptimalkan strategi permainan. Dengan memanfaatkan wawasan yang

dihasilkan oleh model ini, tim dapat mengidentifikasi faktor-faktor kunci yang meningkatkan peluang kemenangan, serta mengantisipasi pola permainan lawan dengan lebih efektif (Akbar et al., 2024; Assagaff et al., 2021; Gifford et al., 2023; Ouyang et al., 2024). Selain itu, penelitian ini membuka peluang bagi pengembangan model yang lebih kompleks dengan mempertimbangkan variabel tambahan yang dapat meningkatkan akurasi prediksi.

Dampak dari penelitian ini tidak hanya terbatas pada ranah akademik, tetapi juga dapat diterapkan dalam praktik olahraga profesional (Herberger et al., 2021; Teixeira et al., 2021). Model prediktif seperti ini dapat digunakan oleh tim pelatih dalam merancang strategi permainan, mengoptimalkan susunan pemain, dan bahkan mendukung pengambilan keputusan dalam perekrutan pemain. Selain itu, model ini juga dapat dimanfaatkan oleh industri taruhan olahraga untuk memberikan prediksi yang lebih akurat kepada para analis dan penggemar sepak bola (Kontos et al., 2023; Majumdar, Bakirov et al., 2022; Rodrigues et al., 2022).

Ke depan, penelitian ini dapat dikembangkan lebih lanjut dengan beberapa perbaikan. Salah satunya adalah penggunaan teknik deep learning yang lebih kompleks, seperti *Recurrent Neural Networks* (RNN) atau *Long Short-Term Memory* (LSTM), untuk menangkap pola sekuensial dalam data pertandingan. Selain itu, pengayaan dataset dengan informasi tambahan seperti statistik individu pemain, data pelacakan pergerakan pemain, dan faktor psikologis tim dapat meningkatkan kualitas prediksi yang dihasilkan.

Secara keseluruhan, penelitian ini memberikan kontribusi signifikan dalam penerapan analitik prediktif dalam sepak bola Indonesia. Dengan menggunakan pendekatan berbasis machine learning, penelitian ini berhasil mengidentifikasi faktor-faktor kunci yang menentukan hasil pertandingan dan mengembangkan model yang dapat diandalkan untuk mendukung pengambilan keputusan strategis dalam kompetisi sepak bola profesional. Dengan adanya model ini, diharapkan kualitas permainan di BRI Liga 1 dapat terus meningkat, dan tim-tim Indonesia dapat bersaing lebih kompetitif di tingkat regional maupun internasional.

SIMPULAN

Penelitian ini telah berhasil mengembangkan dan memvalidasi model prediktif berbasis XGBoost dan *Gradient Boosting Regressor* untuk memprediksi hasil pertandingan BRI Liga 1 Indonesia musim 2023-2024. Model yang digunakan menunjukkan akurasi tinggi dalam mengidentifikasi pola kemenangan, kekalahan, dan hasil seri, serta memprediksi skor pertandingan dengan tingkat kesalahan yang sangat rendah. Hal ini menegaskan bahwa analitik prediktif berbasis *machine learning* dapat menjadi alat yang andal dalam mendukung strategi dan pengambilan keputusan dalam sepak bola profesional.

Dari hasil penelitian ini, dapat disimpulkan bahwa faktor-faktor seperti penguasaan bola, akurasi tembakan, jumlah umpan sukses, dan strategi permainan memiliki pengaruh signifikan terhadap hasil pertandingan. Model yang dikembangkan dapat digunakan oleh pelatih dan manajemen tim untuk mengoptimalkan kinerja tim, merancang strategi berbasis data, serta meningkatkan daya saing dalam kompetisi.

Ke depan, penelitian ini dapat ditingkatkan dengan mempertimbangkan faktor eksternal seperti kondisi cuaca, cedera pemain, dan aspek psikologis tim dalam model prediksi. Selain itu, pengembangan metode *deep learning* dan integrasi data *real-time* dapat lebih meningkatkan akurasi dan manfaat praktis model ini dalam dunia sepak bola profesional.

DAFTAR PUSTAKA

- Aadel, M., & Al-Asadi, M. (2022). *A Comparative Study for Football Analytics With Data Mining and Artificial Intelligence Techniques*. Selçuk Üniversitesi Fen Bilimleri Enstitüsü. <https://hdl.handle.net/20.500.12395/50417>
- Akbar, A., Karim, Z. A., Zakaria, J., Imami, M. K. W., Rahman, M. A., Ron Mahayunan, G., & Yonghong, G. (2024). Formulating a Development Path Model for Under-17 Football Players in Indonesia and Malaysia: A Step Towards Excellence. *Retos*, 55(55), 1010–1017. <https://doi.org/10.47197/retos.v55.105891>
- Al-Asadi, M. A., & Tasdemir, S. (2022). Predict the Value of Football Players Using FIFA Video Game Data and Machine Learning Techniques. *IEEE Access*, 10, 22631–22645. <https://doi.org/10.1109/ACCESS.2022.3154767>
- Ardiantoro, L., M. Akbar, R., & Nur S., Y. (2023). Tactical Analysis of Indonesian National Football Team Using Fp-Tree Algorithm. *Proceeding of International Conference on Science, Health, And Technology*, 124–130. <https://doi.org/10.47701/icohetech.v4i1.3381>
- Assagaff, D. Z., Priya, N., Santoso, B., & Sulistyono, J. (2021). The Development of Football Model to Improve Short Pass Skills in Scoring Tactics. *Journal Of Indonesia Sport Education and Adapted Physical Education (JISEAPE)*, 2(1), 16–22. <http://ejournal.utp.ac.id/index.php/JISEAPE/article/view/1509>
- Ati, A., Bouchet, P., & Ben Jeddou, R. (2024). Using Multi-Criteria Decision-Making and Machine Learning for Football Player Selection and Performance Prediction: A Systematic Review. *Data Science and Management*, 7(2), 79–88. <https://doi.org/10.1016/J.DSM.2023.11.001>
- Baumer, B. S., Matthews, G. J., & Nguyen, Q. (2023). Big ideas in sports analytics and statistical tools for their investigation. *Wiley Interdisciplinary*

Reviews: Computational Statistics, 15(6), e1612.
<https://doi.org/10.1002/wics.1612>

Bhatnagar, P., Lokesh, G. H., Shreyas, J., & Flammini, F. (2024). Rating Prediction of Football Players using Machine Learning. *ACM International Conference Proceeding Series*, 121–126.
<https://doi.org/10.1145/3674029.3674049>

Casal, C. A., Losada, J. L., Barreira, D., & Maneiro, R. (2021). Multivariate Exploratory Comparative Analysis of LaLiga Teams: Principal Component Analysis. *International Journal of Environmental Research and Public Health* 2021, Vol. 18, Page 3176, 18(6), 3176.
<https://doi.org/10.3390/IJERPH18063176>

Cavus, M., & Biecek, P. (2022). Explainable Expected Goal Models for Performance Analysis in Football Analytics. *Proceedings - 2022 IEEE 9th International Conference on Data Science and Advanced Analytics, DSAA 2022*. <https://doi.org/10.1109/DSAA54385.2022.10032440>

Cefis, M., & Carpita, M. (2024). Accuracy and explainability of statistical and machine learning xG models in football. *Statistics*.
<https://doi.org/10.1080/02331888.2024.2445305>

Cortez, A., Trigo, A., & Loureiro, N. (2022). Football Match Line-Up Prediction Based on Physiological Variables: A Machine Learning Approach. *Computers* 2022, Vol. 11, Page 40, 11(3), 40
<https://doi.org/10.3390/COMPUTERS11030040>

Dip, A. Das, Rahman, N., & Ahmed, M. (2024). Predicting Football Match Results: An Analysis of Feature Selection and Machine Learning Techniques Using a Curated Dataset. *PEEIACON 2024 - International Conference on Power, Electrical, Electronics and Industrial Applications*, 927–932.
<https://doi.org/10.1109/PEEIACON63629.2024.10800599>

Fernandez-Crehuet, J. M., Rosales-Salas, J., & Navarro, P. (2021). A performance index of football teams: the Spanish case. *Sport in Society*, 24(2), 131–146. <https://doi.org/10.1080/17430437.2019.1621841>

Franceschi, M., Brocard, J. F., Follert, F., & Gouguet, J. J. (2024). Determinants of football players' valuation: A systematic review. *Journal of Economic Surveys*, 38(3), 577–600. <https://doi.org/10.1111/JOES.12552>

García-Aliaga, A., Marquina, M., Coterón, J., Rodríguez-González, A., & Luengo-Sánchez, S. (2020). In-Game Behaviour Analysis of Football Players Using Machine Learning Techniques Based on Player Statistics. *International Journal of Sports Science & Coaching*, 16(1), 148–157.
<https://doi.org/10.1177/1747954120959762>

- Garnica-Caparrós, M., Memmert, D., & Wunderlich, F. (2022). Artificial data in sports forecasting: a simulation framework for analysing predictive models in sports. *Information Systems and E-Business Management*, 20(3), 551–580. <https://doi.org/10.1007/S10257-022-00560-9/FIGURES/8>
- Gifford, M., & Bayrak, T. (2023). A predictive analytics model for forecasting outcomes in the National Football League games using decision tree and logistic regression. *Decision Analytics Journal*, 8, 100296. <https://doi.org/10.1016/J.DAJOUR.2023.100296>
- Gouveia, V., Duarte, J. P., Nóbrega, A., Sarmiento, H., Pimenta, E., Domingos, F., Antunes, R., Monteiro, D., Matos, R., Amaro, N., & Araújo, I. (2023). Notational Analysis on Goal Scoring and Comparison in Two of the Most Important Soccer Leagues: Spanish La Liga and English Premier League. *Applied Sciences* 2023, Vol. 13, Page 6903, 13(12), 6903. <https://doi.org/10.3390/APP13126903>
- Herberger, T. A., & Litke, C. (2021). The Impact of Big Data and Sports Analytics on Professional Football: A Systematic Literature Review. *Springer Proceedings in Business and Economics*, 147–171. https://doi.org/10.1007/978-3-030-77340-3_12
- Houtmeyers, K. C., Vanrenterghem, J., Jaspers, A., Ruf, L., Brink, M. S., & Helsen, W. F. (2021). Load Monitoring Practice in European Elite Football and the Impact of Club Culture and Financial Resources. *Frontiers in Sports and Active Living*, 3, 679824. <https://doi.org/10.3389/FSPOR.2021.679824/BIBTEX>
- Javed, D., Jhanjhi, N. Z., & Khan, N. A. (2023). Football Analytics for Goal Prediction to Assess Player Performance. *Lecture Notes in Bioengineering*, 245–257. https://doi.org/10.1007/978-981-99-0297-2_20
- Kennedy, H., Kunkel, T., & Funk, D. C. (2021). Using Predictive Analytics to Measure Effectiveness of Social Media Engagement: A Digital Measurement Perspective. *Sport Marketing Quarterly*, 30(4), 265–277. <https://doi.org/10.32731/SMQ.304.1221.02>
- Kontos, C., & Karlis, D. (2023). Football analytics based on player tracking data using interpolation techniques for the prediction of missing coordinates. *Statistica Applicata - Italian Journal of Applied Statistics*, 35(2). <https://doi.org/10.26398/IJAS.0035-010>
- Majumdar, A., Bakirov, R., Hodges, D., Scott, S., & Rees, T. (2022). Machine Learning for Understanding and Predicting Injuries in Football. *Sports Medicine - Open*, 8(1), 1–10. <https://doi.org/10.1186/s40798-022-00465-4>

- Majumdar, Ayush, Kaur, R., Kulkarni, T., Jiruwala, M., Shah, S., & Pise, N. (2022). Football Match Prediction using Exploratory Data Analysis & Multi-Output Regression. *2022 IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation, IATMSI 2022*. <https://doi.org/10.1109/iatmsi56455.2022.10119340>
- Malakreddy, B., Venkataraman, S., Khan, M. S., Nidhi, Padmanabhuni, S., & Natarajan, S. (2024). Optimizing Semantic Segmentation for Enhanced Football Analytics: A Pixel-level Approach. *Procedia Computer Science*, 235, 2662–2673. <https://doi.org/10.1016/j.procs.2024.04.251>
- Modric, T., Versic, S., Stojanovic, M., Chmura, P., Andrzejewski, M., Konefał, M., & Sekulic, D. (2022). Factors affecting match running performance in elite soccer: Analysis of UEFA Champions League matches. *Biology of Sport*, 40(2), 409–416. <https://doi.org/10.5114/biolsport.2023.116453>
- Ouyang, Y., Li, X., Zhou, W., Hong, W., Zheng, W., Qi, F., & Peng, L. (2024). Integration of machine learning XGBoost and SHAP models for NBA game outcome prediction and quantitative analysis methodology. *PLOS ONE*, 19(7), e0307478. <https://doi.org/10.1371/journal.pone.0307478>
- Phatak, A. (2023). *Data-driven performance analysis in soccer: A compilation of data science and machine learning techniques for pre-processing and knowledge discovery*. Deutsche Sporthochschule Köln. <https://fis.dshs-koeln.de/de/publications/data-driven-performance-analysis-in-soccer-a-compilation-of-data->
- Pretto, & Fabricio. (2022). *Development of a Football Analytics Web Application for Player Scouting*. <https://repositorio.utdt.edu/handle/20.500.13098/11892>
- Principe, V. A., de Souza Vale, R. G., de Castro, J. B. P., Carvano, L. M., Henriques, R. A. P., de Almeida e Sousa Lobo, V. J., & de Alkmim Moreira Nunes, R. (2022). A computational literature review of football performance analysis through probabilistic topic modeling. *Artificial Intelligence Review*, 55(2), 1351–1371. <https://doi.org/10.1007/s10462-021-09998-8>
- Ramkumar, P. N., Luu, B. C., Haeberle, H. S., Karnuta, J. M., Nwachukwu, B. U., & Williams, R. J. (2021). Sports Medicine and Artificial Intelligence: A Primer. *https://Doi.Org/10.1177/03635465211008648*, 50(4), 1166–1174. <https://doi.org/10.1177/03635465211008648>
- Razali, N., Mustapha, A., Arbaiy, N., & Lin, P. C. (2022). Deep Learning for Football Outcomes Prediction based on Football Rating System. *AIP Conference Proceedings*, 2644(1). <https://doi.org/10.1063/5.0104587/2831433>

- Richter, C., O'Reilly, M., & Delahunt, E. (2024). Machine learning in sports science: challenges and opportunities. *Sports Biomechanics*, 23(8), 961–967. <https://doi.org/10.1080/14763141.2021.1910334>
- Ridha, S., Mashud, M., Warni, H., & Arifin, S. (2024). Indonesia Student Fitness: A Quantittive Descriptive Study SDN 5 Sungai Ulin, Banjarbaru. *Multilateral: Jurnal Pendidikan Jasmani dan Olahraga*, 23(4), 68–80. <https://doi.org/10.20527/multilateral.v23i4.20878>
- Rodrigues, F., & Pinto, Â. (2022). Prediction of football match results with Machine Learning. *Procedia Computer Science*, 204, 463–470. <https://doi.org/10.1016/j.procs.2022.08.057>
- Rubio Martín, G., Manuel García, C. M., Rodríguez-López, Á., & Gonzalez Sanchez, F. J. (2022). Measuring football clubs' human capital: analytical and dynamic models based on footballers' life cycles. *Journal of Intellectual Capital*, 23(5), 1107–1137. <https://doi.org/10.1108/JIC-06-2020-021>
- Sors, F., Grassi, M., Agostini, T., & Murgia, M. (2021). The sound of silence in association football: Home advantage and referee bias decrease in matches played without spectators. *European Journal of Sport Science*, 21(12), 1597–1605. <https://doi.org/10.1080/17461391.2020.1845814>
- Teixeira, J. E., Leal, M., Ferraz, R., Ribeiro, J., Cachada, J. M., Barbosa, T. M., Monteiro, A. M., & Forte, P. (2021). Effects of Match Location, Quality of Opposition and Match Outcome on Match Running Performance in a Portuguese Professional Football Team. *Entropy*, 23, 23(8), 973. <https://doi.org/10.3390/E23080973>
- Thakkar, P., & Shah, M. (2021). An Assessment of Football Through the Lens of Data Science. *Annals of Data Science*, 8(4), 823–836. <https://doi.org/10.1007/S40745-021-00323-2>
- Torres-Ronda, L., Beanland, E., Whitehead, S., Sweeting, A., & Clubb, J. (2022). Tracking Systems in Team Sports: A Narrative Review of Applications of the Data and Sport Specific Analysis. *Sports Medicine - Open*, 8(1), 1–22. <https://doi.org/10.1186/s40798-022-00408-z>
- Tuyls, K., Omidshafiei, S., Muller, P., Wang, Z., Connor, J., Hennes, D., Graham, I., Spearman, W., Waskett, T., Steele, D., Luc, P., Recasens, A., Galashov, A., Thornton, G., Elie, R., Sprechmann, P., Moreno, P., Cao, K., Garnelo, M., ... Hassabis, D. (2021). Game Plan: What AI can do for Football, and What Football can do for AI. *Journal of Artificial Intelligence Research*, 71, 41–88. <https://doi.org/10.1613/JAIR.1.12505>

- Watanabe, N. M., Shapiro, S., & Drayer, J. (2021). Big Data and Analytics in Sport Management. *Journal of Sport Management*, 35(3), 197–202. <https://doi.org/10.1123/JSM.2021-0067>
- Wheatcroft, E. (2021). Forecasting football matches by predicting match statistics. *Journal of Sports Analytics*, 7(2), 77–97. <https://doi.org/10.3233/JSA-200462>
- Yeung, C., Bunker, R., Umemoto, R., & Fujii, K. (2024). Evaluating soccer match prediction models: a deep learning approach and feature optimization for gradient-boosted trees. *Machine Learning*, 113(10), 7541–7564. <https://doi.org/10.1007/s10994-024-06608-w>
- Yeung, C., & Fujii, K. (2024). A strategic framework for optimal decisions in football 1-vs-1 shot-taking situations: an integrated approach of machine learning, theory-based modeling, and game theory. *Complex and Intelligent Systems*, 10(5), 5989–6008. <https://doi.org/10.1007/s40747-024-01466-4>
- Zhou, W., Yu, G., You, S., & Wang, Z. (2023). An Improved Passing Network for Evaluating Football Team Performance. *Applied Sciences* 2023, Vol. 13, Page 845, 13(2), 845. <https://doi.org/10.3390/APP13020845>
- Zid, A., Amiruddin, R., Auzair, S. M., & Janor, H. (2021). Comprehensive performance measurement system (CPMS) and satisfaction of Malaysian football players. *Current Psychology* 2021 42:9, 42(9), 7070–7080. <https://doi.org/10.1007/S12144-021-02004-Y>