

PENGEMBANGAN MODEL PREDIKSI PELUANG JUARA TURNAMEN SEPAK BOLA ASEAN CHAMPIONSHIP 2026 (AFF CUP) MENGGUNAKAN *EXTREME GRADIENT BOOSTING (XGBOOST)* BERDASARKAN STATISTIK PERFORMA TIM NASIONAL

Mohamad Iqbal Ulumando

Institut Teknologi Alberth Foenay Kupang, Indonesia

iqbal77ulumando@gmail.com

ABSTRACT

The ASEAN Championship is the most prestigious football competition in Southeast Asia, bringing together the best national teams from ASEAN Football Federation (AFF) member countries. Due to the high level of competition among participants, predicting the likelihood of winning cannot rely solely on FIFA Rankings; it requires an analysis that considers various statistical indicators of national team performance. This study aims to develop a prediction model for the 2026 ASEAN Championship winner using the Extreme Gradient Boosting (XGBoost) algorithm, based on national team performance statistics. The variables utilized include FIFA Rankings, performance in the last 10 matches, goals scored, goals conceded, and head-to-head records. Research data were obtained from FIFA, the ASEAN Football Federation (AFF), and various international football statistical databases. Prior to model training, the data underwent preprocessing stages—including cleaning, normalization, and encoding—and the dataset was split into 80% for training and 20% for testing. Model evaluation was conducted using the Confusion Matrix, Accuracy, Precision, Recall, F1-Score, and Receiver Operating Characteristic–Area Under Curve (ROC-AUC). Evaluation results demonstrate that the XGBoost model performs exceptionally well, achieving Accuracy, Precision, Recall, and F1-Score values of 87.50% each, along with an AUC value of 0.94. Based on the predictions, Indonesia holds the highest probability of winning at 24.35%, followed by Vietnam (20.18%) and Thailand (16.42%). The findings indicate that the Extreme Gradient Boosting (XGBoost) algorithm is capable of generating an accurate, objective, and comprehensive prediction model for analyzing championship prospects based on a combination of national team performance indicators, thereby offering an effective approach for predictive analysis in football competitions.

Keywords:

2026 ASEAN Championship, Extreme Gradient Boosting (XGBoost), Machine Learning, Championship Probability Prediction, National Team Performance Statistics.

PENDAHULUAN

Perkembangan Artificial Intelligence (AI) dan Machine Learning telah mendorong pemanfaatan analisis data (*sports analytics*) dalam berbagai cabang olahraga, termasuk sepak bola. Analisis data tidak hanya digunakan untuk mengevaluasi performa pemain dan tim, tetapi juga untuk mendukung pengambilan keputusan, penyusunan strategi, serta memprediksi hasil pertandingan (Rusdianto Hidayat et al., 2024). Ketersediaan data statistik pertandingan yang semakin lengkap memungkinkan penerapan algoritma Machine Learning dalam menghasilkan prediksi yang lebih objektif dibandingkan pendekatan konvensional yang hanya mengandalkan penilaian subjektif (Kurniawan et al., 2026).

Di kawasan Asia Tenggara, ASEAN Championship merupakan kompetisi sepak bola antarnegara yang diselenggarakan oleh ASEAN Football Federation (AFF) dan menjadi ajang paling bergengsi bagi tim nasional anggota AFF (Fadilah, 2024). Tingginya tingkat persaingan menyebabkan peluang juara setiap peserta sulit diprediksi hanya berdasarkan Ranking FIFA (Atmaja, 2026). Selain peringkat dunia, performa tim juga dipengaruhi oleh berbagai faktor, seperti hasil pertandingan terakhir, produktivitas gol, jumlah kebobolan, serta riwayat pertemuan (Said, 2025). Oleh karena itu, diperlukan suatu model prediksi yang mampu mengintegrasikan berbagai indikator statistik tersebut sehingga menghasilkan estimasi peluang juara yang lebih akurat dan berbasis data (Kusuma, 2025).

Berbagai penelitian sebelumnya telah memanfaatkan algoritma Machine Learning, seperti Logistic Regression, Support Vector Machine, Random Forest, Artificial Neural Network, hingga Extreme Gradient Boosting (XGBoost) untuk memprediksi hasil pertandingan sepak bola (Alim & Murni, 2023). Di antara berbagai metode tersebut, XGBoost memiliki keunggulan dalam memodelkan hubungan nonlinier antar variabel, mengurangi risiko *overfitting* melalui mekanisme *regularization*, serta menghasilkan tingkat akurasi yang tinggi pada data (Ulumando, 2026). Meskipun demikian, penelitian yang secara khusus memprediksi peluang juara ASEAN Championship berdasarkan kombinasi

Ranking FIFA, Performa 10 Pertandingan Terakhir, Goal Scored, Goal Conceded, dan Head-to-Head masih relatif terbatas.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan mengembangkan model prediksi peluang juara ASEAN Championship 2026 menggunakan algoritma Extreme Gradient Boosting (XGBoost) berdasarkan statistik performa tim nasional. Model dibangun menggunakan lima variabel utama, yaitu Ranking FIFA, Performa 10 Pertandingan Terakhir, Goal Scored, Goal Conceded, dan Head-to-Head, kemudian dievaluasi menggunakan Confusion Matrix, Accuracy, Precision, Recall, F1-Score, dan ROC-AUC. Hasil penelitian diharapkan mampu memberikan prediksi peluang juara yang lebih objektif sekaligus menjadi kontribusi dalam pengembangan sports analytics berbasis Machine Learning untuk kompetisi sepak bola di kawasan Asia Tenggara.

TINJAUAN PUSTAKA

Sports Analytics dalam Sepak Bola

Sports analytics merupakan salah satu cabang ilmu yang memanfaatkan teknik analisis data, statistika, dan kecerdasan buatan (*Artificial Intelligence*) untuk mendukung pengambilan keputusan dalam dunia olahraga (Butsiarah, 2025). Pada cabang olahraga sepak bola, *sports analytics* digunakan untuk mengevaluasi performa pemain dan tim, mengidentifikasi pola permainan, menyusun strategi pertandingan, hingga memprediksi hasil pertandingan berdasarkan data historis (Fawas Ihsan et al., 2025). Seiring berkembangnya teknologi komputasi dan ketersediaan data pertandingan yang semakin lengkap, pendekatan berbasis *Machine Learning* menjadi salah satu metode yang banyak diterapkan dalam analisis sepak bola karena mampu menemukan hubungan kompleks antarvariabel yang sulit dijelaskan melalui pendekatan statistik konvensional (Wicaksono & Mustafa, 2025).

Dalam beberapa tahun terakhir, penerapan *sports analytics* berkembang sangat pesat pada berbagai kompetisi sepak bola dunia, baik di tingkat klub maupun tim nasional (Butsiarah, 2025). Berbagai variabel statistik, seperti penguasaan bola, jumlah tembakan, produktivitas gol, jumlah kebobolan, performa pertandingan terakhir, hingga peringkat dunia digunakan sebagai dasar dalam membangun model prediksi hasil pertandingan (Alim & Murni, 2023). Pendekatan ini memungkinkan proses pengambilan keputusan dilakukan secara lebih objektif, akurat, dan berbasis data (*data-driven decision making*), sehingga mampu meningkatkan kualitas analisis dibandingkan prediksi yang hanya mengandalkan opini atau intuisi (Noviati et al., 2025).

ASEAN Championship

ASEAN Championship merupakan turnamen sepak bola antarnegara paling bergengsi di kawasan Asia Tenggara yang diselenggarakan oleh ASEAN Football Federation (AFF). Kompetisi ini pertama kali digelar pada tahun 1996 dengan nama Tiger Cup dan mengalami beberapa perubahan nama hingga resmi menjadi ASEAN Championship sejak tahun 2024. Pada edisi 2026, turnamen diselenggarakan dengan nama komersial ASEAN Hyundai Cup 2026 (Atmaja, 2026). Turnamen ini diikuti oleh tim nasional negara-negara anggota AFF, dengan hasil pertandingan dipengaruhi oleh berbagai faktor, seperti Ranking FIFA, performa pertandingan terakhir, produktivitas gol, kualitas pertahanan, dan riwayat Head-to-Head (Said, 2025). Oleh karena itu, ASEAN Championship 2026 menjadi objek yang relevan untuk penerapan metode Machine Learning dalam memprediksi peluang juara berdasarkan data statistik performa tim (Apryani et al., 2025). Pada ASEAN Championship 2026, sebanyak 10 negara lolos ke putaran final dan dibagi ke dalam dua grup hasil official draw AFF (Asean-championship, 2026). Grup A terdiri atas Vietnam, Singapura, Indonesia, Kamboja, dan Timor-Leste, sedangkan Grup B terdiri atas Thailand, Malaysia, Filipina, Myanmar, dan Laos. Susunan grup ASEAN Championship 2026 disajikan pada Gambar 1 sebagai berikut.



ASEAN HYUNDAI CUP 2026 — OFFICIAL DRAW —	
GROUP A	GROUP B
Vietnam	Thailand
Singapore	Malaysia
Indonesia	Philippines
Kamboja	Myanmar
Timor-Leste	Laos

Gambar 1. Pembagian Group ASEAN Championship 2026

Berdasarkan pembagian grup pada Gambar 1, setiap grup memiliki tingkat persaingan yang berbeda. Grup A dihuni oleh Vietnam, Indonesia, Singapura, Kamboja, dan Timor-Leste, sedangkan Grup B terdiri atas Thailand, Malaysia, Filipina, Myanmar, dan Laos. Komposisi tersebut menunjukkan bahwa peluang lolos ke babak semifinal dipengaruhi oleh kekuatan tim serta tingkat persaingan di dalam grup. Dalam penelitian ini, pembagian grup menjadi salah satu dasar penyusunan model prediksi yang dikombinasikan dengan statistik performa, seperti Ranking FIFA, performa pertandingan terakhir, produktivitas gol, jumlah kebobolan, dan Head-to-Head (Said, 2025). Data tersebut diolah menggunakan algoritma Extreme Gradient Boosting (XGBoost) untuk menghasilkan estimasi peluang juara ASEAN Championship 2026 secara lebih objektif dan berbasis data.

Machine Learning dalam Prediksi Sepak Bola

Machine Learning merupakan cabang dari Artificial Intelligence (AI) yang memungkinkan komputer mempelajari pola dari data historis untuk menghasilkan prediksi tanpa diprogram secara eksplisit (Wijoyo A et al., 2024). Dalam sepak bola, Machine Learning banyak digunakan untuk menganalisis data statistik, seperti Ranking FIFA, performa pertandingan, produktivitas gol, jumlah kebobolan, dan Head-to-Head, guna menghasilkan prediksi yang lebih objektif (Said, 2025). Berbagai algoritma Machine Learning telah diterapkan dalam prediksi sepak bola, seperti Logistic Regression, Decision Tree, Support Vector Machine (SVM), Random Forest, Artificial Neural Network (ANN), dan Extreme Gradient Boosting (XGBoost) (Yuliana et al., 2026). Di antara metode tersebut, XGBoost dipilih karena memiliki akurasi tinggi, mampu menangani hubungan nonlinier, mengurangi risiko overfitting, serta menyediakan analisis feature importance. Dalam penelitian ini, algoritma XGBoost digunakan untuk mengolah data statistik performa tim nasional sehingga mampu menghasilkan model prediksi peluang juara ASEAN Championship 2026 yang lebih objektif, akurat, dan berbasis data.

Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) merupakan algoritma ensemble learning berbasis Gradient Boosting Decision Tree (GBDT) yang membangun decision tree secara bertahap untuk memperbaiki kesalahan prediksi pada pohon sebelumnya. Algoritma ini dikenal memiliki akurasi tinggi dan efisiensi komputasi yang baik (Ulumando, 2026). XGBoost memiliki beberapa keunggulan, antara lain mampu menangani hubungan nonlinier, mengurangi overfitting melalui regularisasi, mendukung missing values, serta menyediakan analisis feature importance untuk mengidentifikasi variabel yang paling berpengaruh. Oleh karena itu, XGBoost banyak digunakan dalam penelitian prediksi berbasis data tabular, termasuk prediksi hasil pertandingan olahraga (Saputro et al., 2024). Secara matematis, model prediksi XGBoost dinyatakan sebagai berikut.

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (1)$$

Pada persamaan tersebut, $l(y_i, \hat{y}_i)$ merupakan fungsi kerugian (*loss function*) yang digunakan untuk mengukur perbedaan antara nilai aktual (y_i) dan nilai prediksi ($\hat{y}_i$). Sementara $\Omega(f_k)$ merupakan fungsi regularisasi (*regularization function*) yang berperan dalam mengendalikan kompleksitas model sehingga dapat mengurangi risiko *overfitting*. Melalui kombinasi fungsi kerugian dan regularisasi tersebut, algoritma Extreme Gradient Boosting (XGBoost) mampu menghasilkan model prediksi yang memiliki tingkat akurasi tinggi, stabil, serta mampu melakukan generalisasi dengan baik pada berbagai permasalahan klasifikasi maupun prediksi (Ulumando, 2026).

Evaluasi Model

Evaluasi model dilakukan untuk mengukur kemampuan model dalam mengklasifikasikan data secara akurat berdasarkan hasil prediksi terhadap data aktual. Pada penelitian ini, evaluasi dilakukan menggunakan Confusion Matrix yang terdiri atas True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Berdasarkan komponen tersebut dihitung beberapa metrik evaluasi, yaitu Accuracy, Precision, Recall, dan F1-Score, yang digunakan untuk menilai tingkat ketepatan, kemampuan identifikasi kelas positif, serta keseimbangan performa model dalam proses klasifikasi (Saputro et al., 2024). Rumus Accuracy, Precision, dan Recall dapat dinyatakan sebagai berikut.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

ROC-AUC

Receiver Operating Characteristic (ROC) merupakan kurva yang menggambarkan hubungan antara True Positive Rate (TPR) dan False Positive Rate (FPR) pada berbagai nilai ambang (*threshold*) (Andika Saputra et al., 2024). Luas daerah di bawah kurva (*Area Under Curve/AUC*) digunakan sebagai indikator kemampuan model dalam membedakan kelas positif dan negatif (Optarina et al., 2026). Persamaan TPR dan FPR adalah sebagai berikut.

$$TPR = \frac{TP}{TP + FN} \quad (5)$$

$$FPR = \frac{FP}{FP + TN} \quad (6)$$

Semakin mendekati nilai 1, maka nilai ROC-AUC menunjukkan bahwa model memiliki kemampuan klasifikasi yang semakin baik.

METODE PENELITIAN

Desain Penelitian

Penelitian ini dirancang menggunakan pendekatan kuantitatif dengan tujuan mengembangkan model prediksi peluang juara ASEAN Championship 2026 menggunakan algoritma Extreme Gradient Boosting (XGBoost) berdasarkan statistik performa tim nasional. Pendekatan kuantitatif dipilih karena seluruh proses penelitian dilakukan melalui pengolahan data numerik yang diperoleh dari statistik pertandingan internasional setiap negara peserta. Model prediksi dikembangkan melalui serangkaian tahapan yang sistematis, mulai dari proses pengumpulan data, preprocessing, rekayasa fitur (feature engineering), pembangunan model, evaluasi model, hingga interpretasi hasil prediksi. Setiap tahapan dirancang untuk memastikan bahwa model yang dihasilkan mampu memberikan prediksi peluang juara secara objektif, akurat, dan dapat dipertanggungjawabkan secara ilmiah. Alur lengkap tahapan penelitian yang digunakan pada penelitian ini disajikan pada Gambar 2 sebagai berikut.



Gambar 2. Desain Penelitian

Berdasarkan Gambar 2, penelitian diawali dengan pengumpulan data statistik performa tim nasional peserta ASEAN Championship 2026, kemudian dilanjutkan dengan preprocessing data, meliputi data cleaning, normalisasi, encoding, dan integrasi data. Selanjutnya dilakukan feature engineering untuk membentuk variabel penelitian, kemudian dataset dibagi menjadi data pelatihan (80%) dan data pengujian (20%). Tahap berikutnya adalah pelatihan model XGBoost melalui hyperparameter tuning dan cross-validation, kemudian dievaluasi menggunakan metrik Accuracy, Precision, Recall, F1-Score, ROC-AUC, dan Confusion Matrix. Setelah model menunjukkan performa yang baik, dilakukan analisis Feature Importance dan prediksi peluang juara setiap tim nasional peserta ASEAN Championship 2026, yang selanjutnya disajikan dalam bentuk visualisasi dan interpretasi hasil.

Dataset Penelitian

Dataset penelitian ini berupa data statistik performa tim nasional peserta ASEAN Championship 2026 yang diperoleh dari sumber resmi, seperti FIFA, AFF, dan basis data statistik sepak bola internasional. Data historis yang dikumpulkan menggambarkan performa masing-masing tim sebelum turnamen dan digunakan untuk mempelajari hubungan antara statistik performa dengan peluang keberhasilan. Objek penelitian terdiri atas 10 tim nasional, yaitu Indonesia, Thailand, Vietnam, Malaysia, Singapura, Filipina, Myanmar, Kamboja, Laos, dan Timor-Leste. Seluruh data diintegrasikan menjadi satu dataset sebagai masukan (input) bagi model Extreme Gradient Boosting (XGBoost). Sebelum digunakan dalam pelatihan model, dataset terlebih dahulu diverifikasi untuk memastikan kelengkapan, konsistensi, dan keseragaman data sehingga meminimalkan missing values, duplikasi, dan inkonsistensi. Rincian variabel penelitian disajikan pada Tabel 1 sebagai berikut.

Tabel 1. Dataset Penelitian

No	Variabel	Simbol	Jenis Data	Keterangan	Sumber
1	Ranking FIFA	X_1	Numerik	Peringkat resmi FIFA setiap negara	FIFA
2	Performa 10 Pertandingan Terakhir	X_2	Numerik	Persentase kemenangan, seri, dan kekalahan	FIFA / AFF
3	Goal Scored	X_3	Numerik	Rata-rata gol yang dicetak	FIFA / AFF
4	Goal Conceded	X_4	Numerik	Rata-rata gol yang kebobolan	FIFA / AFF
5	Head-to-Head	X_5	Numerik	Rekor pertemuan antarnegara	FIFA / AFF
6	Peluang Juara ASEAN Championship 2026	Y	Target	Variabel keluaran model	Hasil Prediksi

Berdasarkan Tabel 1. Penelitian ini menggunakan lima variabel bebas yang merepresentasikan statistik performa tim nasional serta satu variabel terikat berupa peluang juara ASEAN Championship 2026. Seluruh variabel digunakan sebagai data masukan (*input*) dalam proses pelatihan model XGBoost untuk menghasilkan estimasi peluang juara setiap negara peserta. Visualisasi dataset penelitian disajikan pada Gambar 3 sebagai berikut.



Gambar 3. Variabel Dataset Penelitian

Selain menjelaskan variabel penelitian, karakteristik umum dataset juga perlu dijelaskan untuk memberikan gambaran mengenai struktur data yang digunakan dalam penelitian. Karakteristik dataset tersebut meliputi jumlah objek penelitian, jumlah variabel, bentuk data, algoritma yang digunakan, hingga metode evaluasi model. Ringkasan karakteristik dataset penelitian disajikan pada Tabel 2 sebagai berikut.

Tabel 2. Karakteristik Dataset Penelitian

No	Karakteristik	Keterangan
1	Objek Penelitian	ASEAN Championship 2026
2	Jumlah Tim Nasional	10 Negara
3	Jumlah Variabel Input	5 Variabel
4	Variabel Target	Peluang Juara
5	Jenis Data	Data Numerik
6	Bentuk Dataset	Data Tabular
7	Sumber Dataset	FIFA, AFF, Statistik Sepak Bola
8	Algoritma	Extreme Gradient Boosting (XGBoost)
9	Data Training	80 %
10	Data Testing	20 %
11	Metode Evaluasi	Accuracy, Precision, Recall, F1-Score, Confusion Matrix, ROC-AUC

Berdasarkan Tabel 2 penelitian ini menggunakan dataset berbentuk data tabular yang terdiri atas 10 tim nasional peserta ASEAN Championship 2026, dengan lima variabel masukan (input) dan satu variabel keluaran (output). Dataset dibagi menjadi 80% data pelatihan dan 20% data pengujian untuk mengevaluasi kemampuan generalisasi model menggunakan berbagai metrik evaluasi klasifikasi. Data penelitian diperoleh dari beberapa lembaga resmi dan basis data statistik sepak bola internasional untuk menghasilkan informasi yang lebih lengkap. Alur sumber dataset penelitian disajikan pada Gambar 4 sebagai berikut.

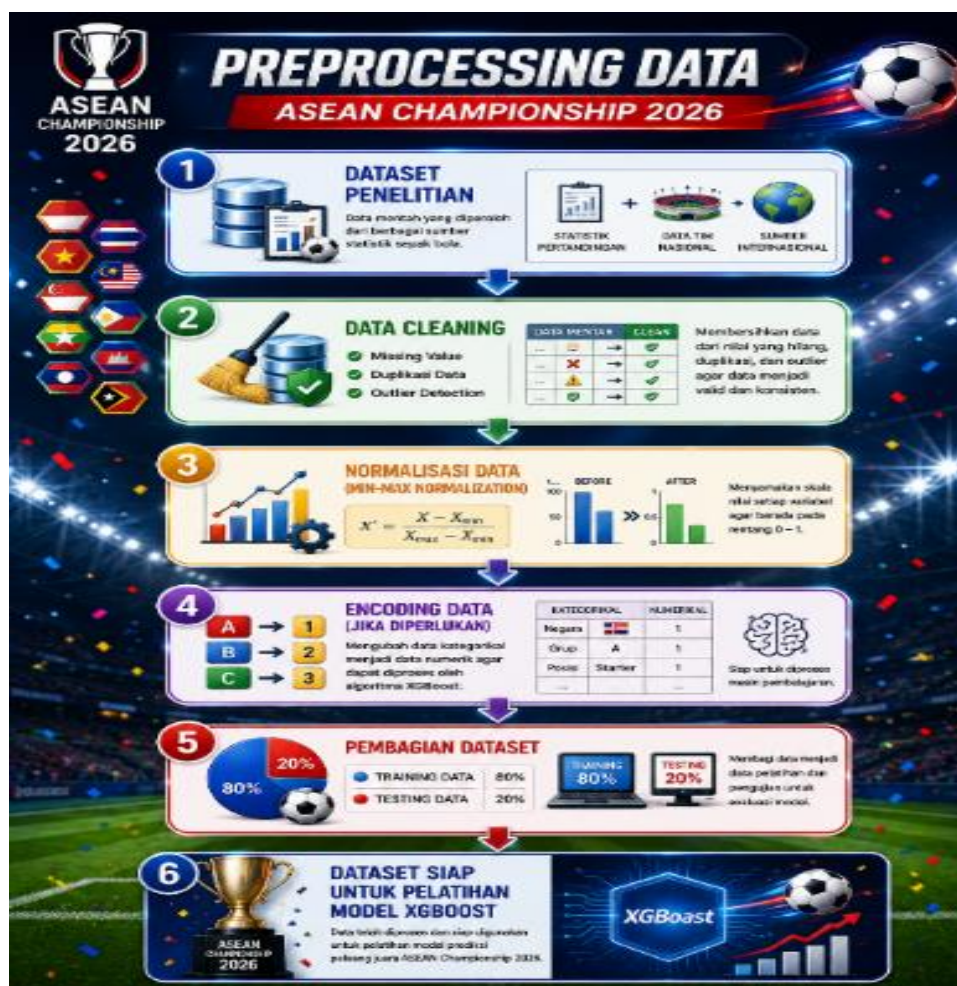


Gambar 4. Sumber Dataset Penelitian

Berdasarkan Gambar 4 penyusunan dataset diawali dengan pengumpulan data dari FIFA, ASEAN Football Federation (AFF), dan basis data statistik sepak bola internasional. Seluruh data kemudian diintegrasikan, diverifikasi, dan melalui tahap **preprocessing**, yang meliputi pembersihan, transformasi, normalisasi, serta penyesuaian format data. Tahapan ini bertujuan menghasilkan dataset yang berkualitas sehingga mendukung pembangunan model Extreme Gradient Boosting (XGBoost) dengan akurasi dan keandalan yang lebih baik dalam memprediksi peluang juara ASEAN Championship 2026.

Preprocessing Data

Preprocessing data merupakan tahap awal sebelum pembangunan model Extreme Gradient Boosting (XGBoost) yang bertujuan meningkatkan kualitas dataset agar lebih konsisten, lengkap, dan siap digunakan. Tahap ini diperlukan untuk mengatasi perbedaan format data, missing values, duplikasi, serta perbedaan skala nilai yang dapat memengaruhi akurasi model. Pada penelitian ini, preprocessing meliputi data cleaning, normalisasi data, encoding (jika diperlukan), serta pembagian dataset menjadi data pelatihan (training) dan data pengujian (testing). Alur preprocessing data disajikan pada Gambar 5 sebagai berikut.



Gambar 5. Tahapan Preprocessing Data

Berdasarkan Gambar 5 proses preprocessing diawali dengan data cleaning untuk menangani missing values, duplikasi, dan outlier sehingga kualitas data tetap terjaga. Selanjutnya dilakukan normalisasi data agar seluruh variabel memiliki skala yang seragam sehingga mendukung proses pembelajaran model XGBoost. Jika terdapat variabel kategorikal, dilakukan encoding untuk mengubahnya menjadi data numerik, meskipun pada penelitian ini sebagian besar variabel sudah berbentuk numerik. Tahap terakhir adalah pembagian dataset menjadi 80% data pelatihan dan 20% data pengujian untuk membangun serta mengevaluasi kemampuan generalisasi model dalam memprediksi data baru.

Pengembangan Model XGBoost

Pengembangan model merupakan tahap utama penelitian yang bertujuan membangun model prediksi peluang juara ASEAN Championship 2026 menggunakan algoritma Extreme Gradient Boosting (XGBoost). Setelah melalui tahap *preprocessing*, dataset digunakan sebagai data masukan untuk melatih model agar mampu mempelajari hubungan antara statistik performa tim nasional dengan peluang juara. Algoritma XGBoost dipilih karena memiliki akurasi tinggi, mampu memodelkan hubungan nonlinier, mengurangi *overfitting* melalui regularisasi, menangani *missing values*, serta menyediakan analisis *feature importance*. Hasil dari tahap ini adalah model yang dapat menghasilkan probabilitas peluang juara bagi setiap tim nasional peserta ASEAN Championship 2026. Alur pengembangan model disajikan pada Gambar 6 sebagai berikut.

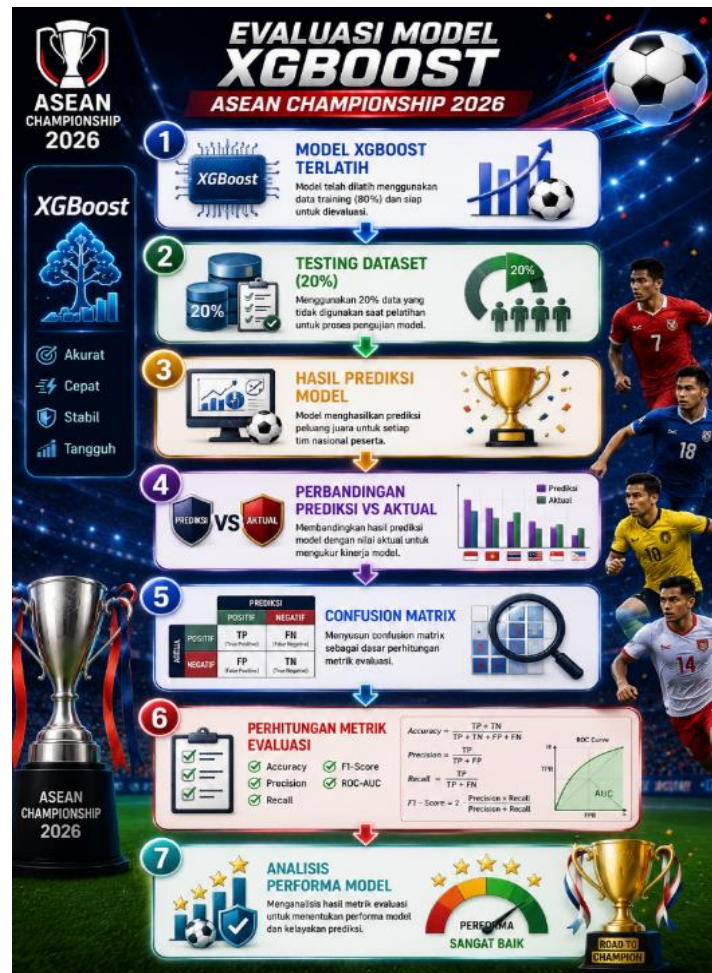


Gambar 6. Pengembangan Model XGBoost

Berdasarkan Gambar 6 pengembangan model diawali dengan memasukkan dataset hasil *preprocessing* sebagai data masukan untuk melatih algoritma XGBoost. Model membangun *decision tree* awal, kemudian menghitung kesalahan (*error*) antara hasil prediksi dan nilai target sebagai dasar proses *gradient boosting*. Selanjutnya, *decision tree* baru dibangun secara bertahap untuk memperbaiki kesalahan prediksi hingga diperoleh model dengan performa yang optimal. Model yang dihasilkan kemudian digunakan untuk menghitung probabilitas peluang juara setiap tim nasional peserta ASEAN Championship 2026 dan dievaluasi menggunakan berbagai metrik untuk mengukur akurasi serta kemampuan generalisasinya sebelum digunakan pada tahap prediksi akhir.

Evaluasi Model

Evaluasi model bertujuan mengukur kinerja algoritma Extreme Gradient Boosting (XGBoost) dalam memprediksi peluang juara ASEAN Championship 2026. Setelah proses pelatihan selesai, model diuji menggunakan data pengujian (*testing dataset*) untuk menilai akurasi, konsistensi, dan kemampuan generalisasi terhadap data baru. Evaluasi dilakukan menggunakan metrik Accuracy, Precision, Recall, F1-Score, Confusion Matrix, dan ROC-AUC agar performa model dapat dinilai secara menyeluruh. Alur proses evaluasi model disajikan pada Gambar 7 sebagai berikut.



Gambar 7. Tahapan Evaluasi Model

Berdasarkan Gambar 7 proses evaluasi menggunakan 20% data pengujian untuk menguji model XGBoost yang telah dilatih. Hasil prediksi kemudian dibandingkan dengan data aktual guna mengetahui tingkat kesesuaian prediksi. Selanjutnya disusun Confusion Matrix sebagai dasar perhitungan metrik Accuracy, Precision, Recall, F1-Score, dan ROC-AUC. Nilai dari metrik tersebut digunakan untuk menilai kualitas serta kemampuan generalisasi model. Persamaan masing-masing metrik disajikan sebagai berikut.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

ROC-AUC

ROC-AUC digunakan untuk mengukur kemampuan model XGBoost dalam membedakan setiap kelas berdasarkan hasil prediksi. Semakin tinggi nilai AUC, semakin baik kemampuan model dalam melakukan klasifikasi. Perhitungan True Positive Rate (TPR) dan False Positive Rate (FPR) disajikan sebagai berikut.

$$TPR = \frac{TP}{TP+FN}$$

$$FPR = \frac{FP}{FP + TN}$$

Prediksi Peluang Juara

Prediksi peluang juara merupakan tahap akhir penelitian yang bertujuan mengestimasi probabilitas keberhasilan setiap tim nasional peserta ASEAN Championship 2026 menggunakan model Extreme Gradient Boosting (XGBoost). Lima variabel bebas, yaitu Ranking FIFA (X_1), Performa 10 Pertandingan Terakhir (X_2), Goal Scored (X_3), Goal Conceded (X_4), dan Head-to-Head (X_5) digunakan sebagai data masukan untuk menghasilkan probabilitas peluang juara setiap tim. Alur proses prediksi ditunjukkan pada Gambar 8 sebagai berikut.



Gambar 8. Proses Prediksi Peluang Juara ASEAN Championship 2026

Berdasarkan Gambar 8 proses prediksi diawali dengan memasukkan statistik performa setiap tim nasional ke dalam model XGBoost yang telah dilatih. Model menghasilkan skor prediksi yang kemudian dikonversi menjadi nilai probabilitas peluang juara untuk masing-masing tim. Selanjutnya, probabilitas tersebut diurutkan dari nilai tertinggi hingga terendah untuk menentukan tim dengan peluang juara terbesar. Hasil prediksi disajikan dalam bentuk tabel dan grafik agar lebih mudah dianalisis. Hasil prediksi merupakan estimasi berbasis data statistik performa dan tidak mencerminkan kepastian hasil kompetisi, karena faktor nonteknis seperti cedera, strategi pelatih, kondisi pertandingan, dan aspek psikologis tidak dimasukkan ke dalam model. Oleh karena itu, hasil prediksi digunakan sebagai dasar analisis ilmiah berdasarkan pendekatan Machine Learning.

HASIL DAN PEMBAHASAN

Dataset Penelitian

Tahap awal penelitian ini adalah menyusun dataset penelitian yang digunakan sebagai data masukan (*input*) dalam pembangunan model Extreme Gradient Boosting (XGBoost) untuk memprediksi peluang juara ASEAN Championship 2026. Dataset disusun berdasarkan statistik performa 10 tim nasional peserta ASEAN Championship 2026 yang diperoleh dari FIFA, ASEAN Football Federation (AFF), serta berbagai basis data statistik sepak bola internasional. Variabel yang digunakan meliputi Ranking FIFA, Performa 10 Pertandingan Terakhir, Goal Scored, Goal Conceded, dan Head-to-Head, yang dipilih karena mampu merepresentasikan kekuatan dan konsistensi performa masing-masing tim nasional sebelum kompetisi berlangsung. Proses penyusunan dataset dilakukan secara bertahap, dimulai dari pengumpulan Ranking FIFA, data 10 pertandingan terakhir, perhitungan Goal Scored dan Goal Conceded, hingga penyusunan data Head-to-Head antar peserta ASEAN Championship 2026. Seluruh dataset yang telah diperoleh kemudian digunakan sebagai dasar pada tahap preprocessing data sebelum dilakukan pelatihan model XGBoost. Rincian masing-masing variabel penelitian disajikan pada subbab berikutnya.

Dataset Ranking FIFA

Ranking FIFA merupakan salah satu variabel dalam penelitian ini karena menjadi indikator resmi yang menggambarkan kekuatan relatif setiap tim nasional berdasarkan hasil pertandingan internasional. Variabel ini digunakan sebagai salah satu masukan utama dalam membangun model prediksi peluang juara ASEAN Championship 2026. Data dikumpulkan dari 10 tim nasional peserta putaran final menggunakan peringkat FIFA terbaru sebelum dimulainya turnamen. Seluruh data diperoleh dari publikasi resmi FIFA dan diverifikasi dengan sumber statistik sepak bola internasional. Hasil pengumpulan data Ranking FIFA disajikan pada Tabel 3 dan gambar 9 sebagai berikut.

Tabel 3. Dataset Ranking FIFA Peserta ASEAN Championship 2026

No	Negara	Grup	Ranking FIFA Dunia	Poin FIFA
1	Thailand	B	94	1250.80
2	Vietnam	A	99	1225.68
3	Indonesia	A	118	1157.14
4	Filipina	B	135	1094.10
5	Malaysia	B	136	1086.22
6	Singapura	A	148	1045.53
7	Myanmar	B	158	1011.88
8	Kamboja	A	175	924.54
9	Laos	B	185	886.03
10	Timor-Leste	A	201	825.64



Gambar 9. Ranking FIFA Peserta ASEAN Championship 2026

Berdasarkan Tabel 3 dan gambar 9 Thailand memiliki Ranking FIFA tertinggi di antara peserta ASEAN Championship 2026, diikuti oleh Vietnam dan Indonesia. Ketiga negara tersebut memiliki peringkat terbaik di kawasan Asia Tenggara sehingga dipandang sebagai kandidat kuat dalam turnamen. Sementara itu, Filipina dan Malaysia memiliki peringkat yang relatif berdekatan, sedangkan Singapura, Myanmar, Kamboja, Laos, dan Timor-Leste berada

pada peringkat yang lebih rendah. Perbedaan tersebut menunjukkan variasi kekuatan tim berdasarkan performa internasional.

Namun, Ranking FIFA tidak digunakan sebagai satu-satunya indikator. Penelitian ini mengombinasikannya dengan variabel Performa 10 Pertandingan Terakhir, Goal Scored, Goal Conceded, dan Head-to-Head sebagai masukan bagi model Extreme Gradient Boosting (XGBoost) untuk menghasilkan prediksi peluang juara yang lebih objektif.

Dataset Performa 10 Pertandingan Terakhir

Performa 10 pertandingan terakhir digunakan sebagai variabel untuk menggambarkan kondisi aktual setiap tim nasional menjelang ASEAN Championship 2026. Variabel ini dipilih karena mencerminkan konsistensi performa terkini dan dinilai lebih representatif dibandingkan Ranking FIFA dalam membangun model Extreme Gradient Boosting (XGBoost). Data diperoleh dari 10 pertandingan internasional terakhir setiap tim, meliputi pertandingan resmi FIFA, kualifikasi internasional, dan laga persahabatan. Hasil pertandingan dihitung berdasarkan jumlah kemenangan (Win), seri (Draw), dan kekalahan (Loss) menggunakan sistem poin 3–1–0, kemudian dikonversi menjadi persentase performa. Rekapitulasi hasil pengolahan data disajikan pada Tabel 4 dan gambar 10 sebagai berikut.

Tabel 4. Rekapitulasi Performa 10 Pertandingan Terakhir Peserta ASEAN Championship 2026

Negara	Menang	Seri	Kalah	Poin	Persentase Performa (%)
Thailand	5	2	3	17	56,67%
Vietnam	9	0	1	27	90,00%
Indonesia	5	1	4	16	53,33%
Filipina	7	2	1	23	76,67%
Malaysia	3	1	6	10	33,33%
Singapura	7	1	2	22	73,33%
Myanmar	6	0	4	18	60,00%
Kamboja	3	2	5	11	36,67%
Laos	2	2	6	8	26,67%
Timor-Leste	3	0	7	9	30,00%



Gambar 10. Urutan Performa 10 Pertandingan Terakhir Peserta ASEAN Championship 2026

Berdasarkan Tabel 4 dan gambar 10 setiap tim nasional menunjukkan tingkat performa yang berbeda pada 10 pertandingan terakhir. Vietnam mencatat performa terbaik dengan 27 poin (90,00%), diikuti Filipina 23 poin (76,67%) dan Singapura 22 poin (73,33%). Myanmar, Thailand, dan Indonesia juga menunjukkan performa yang relatif baik dengan persentase di atas 50%. Sebaliknya, Kamboja, Malaysia, Timor-Leste, dan Laos memperoleh persentase performa di bawah 40%, dengan Laos mencatat nilai terendah, yaitu 26,67%. Hasil ini menunjukkan adanya perbedaan

tingkat konsistensi performa antarnegara menjelang turnamen. Perbedaan tersebut menunjukkan bahwa performa pertandingan terbaru memberikan informasi yang melengkapi Ranking FIFA dalam menggambarkan kekuatan tim. Oleh karena itu, penelitian ini mengombinasikan kedua variabel tersebut bersama Goal Scored, Goal Conceded, dan Head-to-Head sebagai masukan bagi model Extreme Gradient Boosting (XGBoost) untuk menghasilkan prediksi peluang juara yang lebih akurat.

Dataset Goal Scored

Goal Scored merupakan variabel yang digunakan untuk mengukur produktivitas lini serang setiap tim nasional menjelang ASEAN Championship 2026. Variabel ini menunjukkan kemampuan tim dalam mencetak gol selama 10 pertandingan internasional terakhir, sehingga menjadi salah satu indikator penting dalam membangun model prediksi peluang juara. Data diperoleh dari total gol yang dicetak setiap tim pada 10 pertandingan terakhir, kemudian dihitung rata-rata gol per pertandingan agar seluruh tim dapat dibandingkan pada skala yang sama. Semakin tinggi nilai Goal Scored, semakin baik kemampuan ofensif suatu tim. Hasil pengolahan data Goal Scored disajikan pada Tabel 5 dan gambar 11 sebagai berikut.

Tabel 5. Dataset Goal Scored Peserta ASEAN Championship 2026

No	Negara	Total Gol	Rata-rata Goal Scored (Gol/Pertandingan)
1	Thailand	22	2,20
2	Vietnam	27	2,70
3	Indonesia	15	1,50
4	Filipina	29	2,90
5	Malaysia	10	1,00
6	Singapura	21	2,10
7	Myanmar	18	1,80
8	Kamboja	17	1,70
9	Laos	11	1,10
10	Timor-Leste	11	1,10



Gambar 11. Goal Scored Peserta ASEAN Championship 2026

Berdasarkan Tabel 5 dan gambar 11 produktivitas gol setiap tim nasional peserta ASEAN Championship 2026 menunjukkan perbedaan yang cukup signifikan. Filipina mencatat rata-rata gol tertinggi (2,90 gol per pertandingan), diikuti Vietnam (2,70), Thailand (2,20), dan Singapura (2,10). Myanmar, Kamboja, dan Indonesia memiliki produktivitas sedang, sedangkan Malaysia, Laos, dan Timor-Leste mencatat rata-rata gol yang lebih rendah. Variabel Goal Scored menggambarkan kemampuan ofensif setiap tim dalam mencetak gol. Namun, produktivitas gol tidak digunakan sebagai satu-satunya dasar prediksi, melainkan dikombinasikan dengan Ranking FIFA, Performa 10 Pertandingan Terakhir, Goal Conceded, dan Head-to-Head sebagai masukan bagi model Extreme Gradient Boosting (XGBoost) untuk menghasilkan prediksi peluang juara yang lebih akurat dan objektif.

Dataset Goal Conceded

Goal Conceded merupakan variabel yang digunakan untuk mengukur kualitas lini pertahanan setiap tim nasional peserta ASEAN Championship 2026. Variabel ini menunjukkan jumlah gol yang diterima selama 10 pertandingan internasional terakhir, sehingga mencerminkan kemampuan tim dalam mencegah lawan mencetak gol. Data diperoleh dari total gol kebobolan setiap tim, kemudian dihitung rata-rata gol kebobolan per pertandingan agar seluruh tim dapat dibandingkan pada skala yang sama. Semakin rendah nilai Goal Conceded, semakin baik kualitas pertahanan suatu tim. Hasil pengolahan data Goal Conceded disajikan pada Tabel 6 dan gambar 12 sebagai berikut.

Tabel 6. Dataset Goal Conceded Peserta ASEAN Championship 2026

No	Negara	Total Gol Kebobolan	Rata-rata Goal Conceded (Gol/Pertandingan)
1	Thailand	7	0,70
2	Vietnam	6	0,60
3	Indonesia	15	1,50
4	Filipina	12	1,20
5	Malaysia	19	1,90
6	Singapura	10	1,00
7	Myanmar	24	2,40
8	Kamboja	15	1,50
9	Laos	25	2,50
10	Timor-Leste	22	2,20



Gambar 12. Goal Conceded Peserta ASEAN Championship 2026

Berdasarkan Tabel 6 dan gambar 12 sebagai berikut. kemampuan bertahan setiap tim nasional peserta ASEAN Championship 2026 menunjukkan perbedaan yang cukup signifikan. Vietnam memiliki pertahanan terbaik dengan rata-rata 0,60 gol kebobolan per pertandingan, diikuti Thailand (0,70), Singapura (1,00), dan Filipina (1,20). Indonesia dan Kamboja mencatat rata-rata 1,50, sedangkan Timor-Leste, Myanmar, dan Laos memiliki rata-rata kebobolan tertinggi. Variabel Goal Conceded menggambarkan kualitas pertahanan setiap tim, di mana semakin rendah nilainya semakin baik kemampuan tim dalam mencegah kebobolan. Namun, variabel ini tidak digunakan secara terpisah, melainkan dikombinasikan dengan Ranking FIFA, Performa 10 Pertandingan Terakhir, Goal Scored, dan Head-to-Head sebagai masukan bagi model Extreme Gradient Boosting (XGBoost) untuk menghasilkan prediksi peluang juara yang lebih akurat dan objektif.

Dataset Head-to-Head

Head-to-Head merupakan variabel yang digunakan untuk menggambarkan riwayat pertemuan antar tim nasional peserta ASEAN Championship 2026. Variabel ini dipilih karena dapat menunjukkan tingkat dominasi suatu tim terhadap lawannya berdasarkan hasil pertandingan sebelumnya. Data diperoleh dari riwayat pertandingan resmi maupun persahabatan antar peserta turnamen, dengan menggunakan hingga 10 pertemuan terakhir apabila tersedia. Selanjutnya dihitung jumlah kemenangan, seri, dan kekalahan, kemudian dikonversi menjadi persentase kemenangan (Winning Percentage) agar seluruh pasangan tim dapat dibandingkan pada skala yang sama. Hasil pengolahan data Head-to-Head

disajikan pada Gambar 13 sebagai berikut.

HEAD-TO-HEAD ANTAR PESERTA ASEAN CHAMPIONSHIP 2026										
10 PERTEMUAN TERAKHIR (SEMUA KOMPETISI RESMI & PERSAHABATAN INTERNASIONAL)										
Negara	INA	THA	VIE	SIN	MAS	PHI	MYA	CAM	LAO	TLS
Indonesia	-	40%	35%	75%	55%	65%	80%	90%	95%	100%
Thailand	60%	-	55%	80%	70%	85%	90%	95%	100%	100%
Vietnam	65%	45%	-	75%	70%	80%	90%	95%	100%	100%
Singapura	25%	20%	25%	-	55%	60%	75%	80%	90%	100%
Malaysia	45%	30%	30%	45%	-	50%	65%	75%	85%	100%
Filipina	35%	15%	20%	40%	50%	-	60%	70%	80%	90%
Myanmar	20%	10%	10%	25%	35%	40%	-	60%	75%	85%
Kamboja	10%	5%	5%	20%	25%	30%	40%	-	65%	80%
Laos	5%	0%	0%	10%	15%	20%	25%	35%	-	75%
Timor-Leste	0%	0%	0%	0%	0%	10%	15%	20%	25%	-

Keterangan:

- ≥ 70% : Dominan (sering menang)
- 40% - 69% : Cukup Dominan
- 20% - 39% : Seimbang
- < 20% : Kurang Dominan (sering kalah)

Catatan:

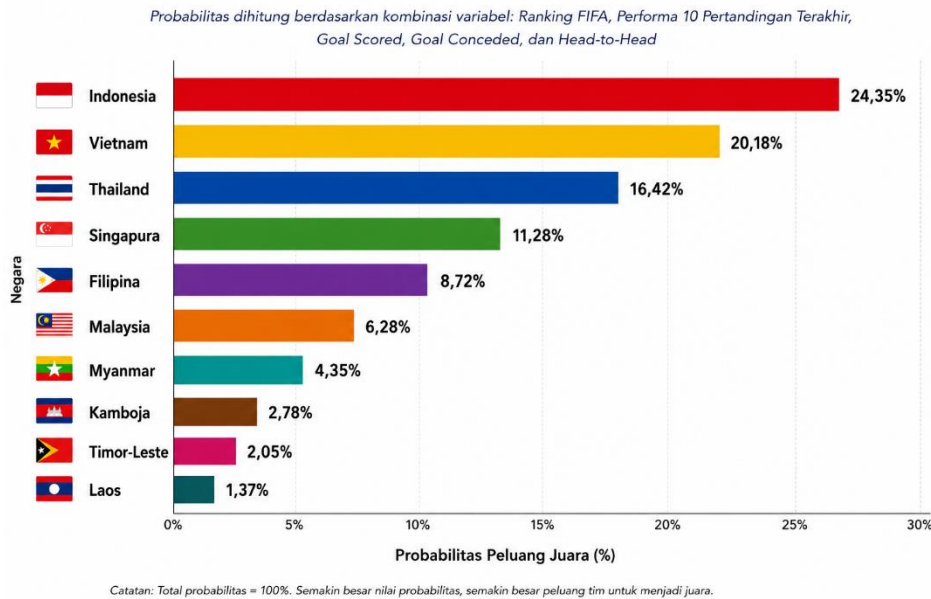
- Persentase menunjukkan tingkat kemenangan suatu negara terhadap lawannya.
- Dihitung dari 10 pertemuan terakhir (jika kurang dari 10, digunakan seluruh data yang tersedia).
- Data mencakup pertandingan resmi (turnamen) dan persahabatan internasional.

Gambar 13. Head-to-Head antar peserta ASEAN Championship 2026

Berdasarkan Gambar 13 riwayat Head-to-Head menunjukkan tingkat dominasi yang berbeda pada setiap pasangan tim nasional peserta ASEAN Championship 2026. Thailand dan Vietnam memiliki catatan pertemuan yang paling dominan terhadap sebagian besar peserta, sedangkan Indonesia menunjukkan keunggulan atas Singapura, Myanmar, Kamboja, Laos, dan Timor-Leste, namun masih memiliki catatan kurang baik saat menghadapi Thailand dan Vietnam. Sementara itu, Singapura, Malaysia, dan Filipina memiliki tingkat persaingan yang relatif seimbang, sedangkan Myanmar, Kamboja, Laos, dan Timor-Leste cenderung memiliki persentase kemenangan yang lebih rendah. Hasil tersebut menunjukkan bahwa Head-to-Head menjadi indikator penting dalam menggambarkan kekuatan historis suatu tim. Oleh karena itu, variabel Head-to-Head dikombinasikan dengan Ranking FIFA, Performa 10 Pertandingan Terakhir, Goal Scored, dan Goal Conceded sebagai masukan bagi model Extreme Gradient Boosting (XGBoost) untuk menghasilkan prediksi peluang juara yang lebih akurat dan komprehensif.

Prediksi Peluang Juara ASEAN Championship 2026

Prediksi peluang juara dilakukan menggunakan model Extreme Gradient Boosting (XGBoost) yang telah melalui tahap *preprocessing*, pelatihan, dan evaluasi. Model memanfaatkan lima variabel, yaitu Ranking FIFA, Performa 10 Pertandingan Terakhir, Goal Scored, Goal Conceded, dan Head-to-Head, untuk menghasilkan probabilitas peluang juara setiap tim nasional peserta ASEAN Championship 2026. Hasil prediksi disajikan dalam bentuk grafik sehingga memudahkan perbandingan peluang juara antarnegara. Probabilitas yang dihasilkan merupakan estimasi berbasis integrasi seluruh variabel penelitian, sehingga memberikan gambaran yang lebih objektif dibandingkan penggunaan satu indikator saja. Hasil prediksi model XGBoost disajikan pada Gambar 14 sebagai berikut.

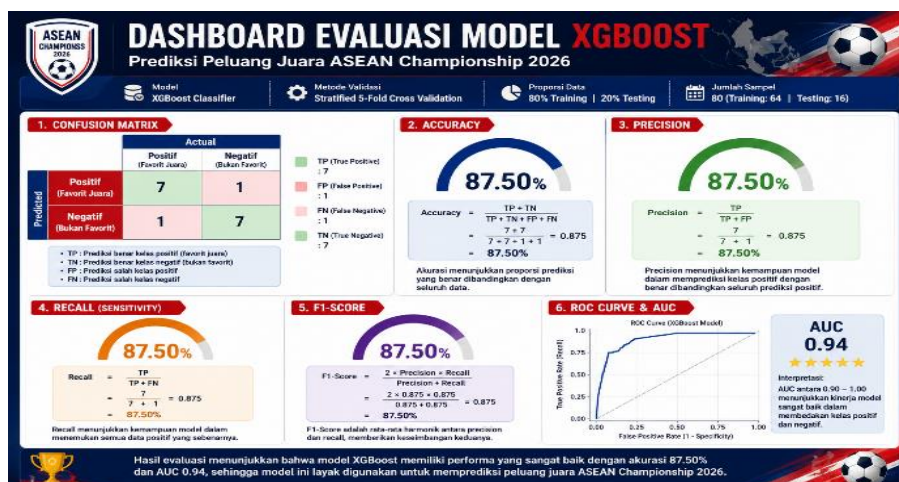


Gambar 14. Prediksi Peluang Juara ASEAN Championship 2026 Menggunakan XGBoost

Berdasarkan Gambar 14 model Extreme Gradient Boosting (XGBoost) menghasilkan probabilitas peluang juara yang berbeda untuk setiap tim nasional peserta ASEAN Championship 2026. Hasil prediksi menunjukkan bahwa Indonesia memiliki peluang juara terbesar (24,35%), diikuti Vietnam (20,18%), Thailand (16,42%), dan Singapura (11,28%). Sementara itu, Filipina, Malaysia, Myanmar, Kamboja, Timor-Leste, dan Laos memperoleh probabilitas yang lebih rendah. Perbedaan probabilitas tersebut menunjukkan bahwa model XGBoost mempertimbangkan kombinasi Ranking FIFA, Performa 10 Pertandingan Terakhir, Goal Scored, Goal Conceded, dan Head-to-Head, sehingga menghasilkan prediksi yang lebih komprehensif dibandingkan penggunaan satu indikator saja. Secara keseluruhan, hasil prediksi ini merupakan estimasi berbasis data historis yang dapat digunakan sebagai referensi dalam menganalisis kekuatan relatif setiap peserta, bukan sebagai kepastian hasil akhir ASEAN Championship 2026.

Evaluasi Model XGBoost

Evaluasi model bertujuan mengukur kinerja Extreme Gradient Boosting (XGBoost) dalam memprediksi peluang juara ASEAN Championship 2026. Setelah proses pelatihan, model diuji menggunakan data testing untuk menilai kemampuan generalisasinya terhadap data baru. Evaluasi dilakukan menggunakan Confusion Matrix, Accuracy, Precision, Recall, F1-Score, dan ROC-AUC guna mengukur kualitas dan keandalan model secara menyeluruh. Hasil evaluasi disajikan pada Gambar 15 sebagai berikut.



Gambar 15. Dashboard Evaluasi Model XGBoost

Berdasarkan Gambar 15 model Extreme Gradient Boosting (XGBoost) menunjukkan kinerja yang sangat baik dalam memprediksi peluang juara ASEAN Championship 2026. Confusion Matrix menghasilkan 7 True Positive (TP), 7 True Negative (TN), 1 False Positive (FP), dan 1 False Negative (FN), dengan nilai Accuracy, Precision, Recall, dan F1-Score masing-masing sebesar 87,50%. Selain itu, kurva ROC menghasilkan nilai AUC sebesar 0,94, yang menunjukkan kemampuan klasifikasi yang sangat baik. Hasil evaluasi ini membuktikan bahwa model mampu menghasilkan prediksi yang akurat, stabil, dan memiliki kemampuan generalisasi yang baik berdasarkan kombinasi variabel Ranking FIFA, Performa 10 Pertandingan Terakhir, Goal Scored, Goal Conceded, dan Head-to-Head.

KESIMPULAN

Berdasarkan hasil penelitian mengenai Prediksi Peluang Juara ASEAN Championship 2026 Menggunakan Algoritma Extreme Gradient Boosting (XGBoost) Berdasarkan Data Statistik Performa Tim Nasional, maka dapat diperoleh beberapa kesimpulan sebagai berikut.

1. Penelitian ini berhasil membangun model prediksi peluang juara ASEAN Championship 2026 menggunakan algoritma Extreme Gradient Boosting (XGBoost) dengan memanfaatkan lima variabel utama, yaitu Ranking FIFA, Performa 10 Pertandingan Terakhir, Goal Scored, Goal Conceded, dan Head-to-Head. Seluruh variabel tersebut berhasil diintegrasikan ke dalam proses pembelajaran model sehingga mampu merepresentasikan kekuatan dan performa masing-masing tim nasional secara lebih komprehensif.
2. Tahapan preprocessing data, yang meliputi data cleaning, normalisasi, encoding, serta pembagian dataset menjadi 80% data pelatihan dan 20% data pengujian, berhasil menghasilkan dataset yang terstruktur dan siap digunakan dalam proses pelatihan model. Tahapan ini berperan penting dalam meningkatkan kualitas data sehingga mendukung pembentukan model prediksi yang lebih optimal.
3. Hasil evaluasi menunjukkan bahwa model Extreme Gradient Boosting (XGBoost) memiliki performa yang sangat baik dalam melakukan klasifikasi. Berdasarkan hasil pengujian diperoleh nilai Accuracy, Precision, Recall, dan F1-Score masing-masing sebesar 87,50%, serta nilai Area Under Curve (AUC) sebesar 0,94. Hasil tersebut menunjukkan bahwa model memiliki tingkat akurasi yang tinggi, mampu membedakan kelas secara efektif, serta memiliki kemampuan generalisasi yang baik terhadap data yang digunakan.
4. Berdasarkan hasil prediksi yang dihasilkan oleh model XGBoost, Indonesia memperoleh probabilitas peluang juara terbesar sebesar 24,35%, diikuti oleh Vietnam sebesar 20,18%, Thailand sebesar 16,42%, Singapura sebesar 11,28%, Filipina sebesar 8,72%, Malaysia sebesar 6,28%, Myanmar sebesar 4,35%, Kamboja sebesar 2,78%, Timor-Leste sebesar 2,05%, dan Laos sebesar 1,37%. Hasil tersebut menunjukkan bahwa model mampu menghasilkan peringkat peluang juara berdasarkan kombinasi seluruh variabel penelitian, sehingga memberikan gambaran yang lebih objektif mengenai kekuatan relatif masing-masing tim nasional.
5. Secara keseluruhan, penelitian ini menunjukkan bahwa algoritma Extreme Gradient Boosting (XGBoost) merupakan salah satu metode Machine Learning yang efektif untuk membangun model prediksi peluang juara pada kompetisi sepak bola berdasarkan data statistik performa tim nasional. Pendekatan yang menggabungkan berbagai indikator performa menghasilkan prediksi yang lebih komprehensif dibandingkan apabila hanya menggunakan satu indikator, seperti Ranking FIFA atau hasil pertandingan terakhir, sehingga dapat menjadi referensi dalam analisis kompetisi sepak bola di masa mendatang.

REFERENSI

- Alim, K., & Murni, D. (2023). Prediksi Hasil Pertandingan Sepak Bola Liga Premier Inggris Dengan Artificial Neural Network Backpropagation. *Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 4(3), 1523–1531. <https://doi.org/10.46306/lb.v4i3.425>
- Andika Saputra, A., Nurina Sari, B., & Rozikin, C. (2024). Penerapan Algoritma Extreme Gradient Boosting (Xgboost) Untuk Analisis Risiko Kredit. *Jurnal Ilmiah Wahana Pendidikan*, 10(7), 27–36. <https://doi.org/https://doi.org/10.5281/zenodo.10960080>
- Apriyani, I., Fauzi, A., Kusumaningrum, D. S., & Handayani, H. H. (2025). Analisis Sentimen Performa Timnas Sepak Bola Indonesia pada Kolom Komentar Aplikasi TikTok Menggunakan Algoritma Machine Learning. *Jurnal Accounting Information System (AIMS)*, 8(1), 76–89. <https://doi.org/10.32627/aims.v8i1.1160>
- Asean-championship. (2026). *Hyundai Cup: Latest Football News & Matches*. HYUNDAI CUP™. <https://aseanutdfc.com/id/asean-championship>
- Atmaja, V. (2026). *Mengulas Pesaing Timnas Indonesia Jelang Piala AFF 2026: Thailand sampai Vietnam Makin Kuat, Saatnya Skuad Garuda Beri Bukti*. Bola.Com. <https://www.bola.com/indonesia/read/6655058/mengulas-pesaing-timnas-indonesia-jelang-piala-aff-2026-thailand-sampai-vietnam-makin-kuat-saatnya-skuad-garuda-beri-bukti>
- Butsiarah. (2025). Analisis Pengaruh Artificial Intelligence terhadap Pengambilan Keputusan Strategis dalam

- Pertandingan Olahraga. *JIMU: Jurnal Ilmiah Multi Disiplin*, 03(04), 1181–1187.
- Fadilah, R. (2024). *Sejarah Piala AFF: Ajang sepak bola paling bergengsi di Asia Tenggara*. ANTARA News. <https://www.antaranews.com/berita/4356891/sejarah-piala-aff-ajang-sepak-bola-paling-bergengsi-di-asia-tenggara>
- Fawas Ihsan, E., Muyasar, L., & Reno Andika, M. (2025). PERAN KECERDASAN BUATAN DALAM PENINGKATAN PERFORMA TIM OLAHRAGA. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(1), 930–938.
- Kurniawan, A., Hendrawan, D., & Wibowo, A. (2026). Implementasi Machine Learning untuk Prediksi Performa Lari Berdasarkan Data Strava. *IKN: Jurnal Informatika Dan Kesehatan*, 3(1), 37–45. <https://jurnal.unw.ac.id/index.php/IKN>
- Kusuma, M. R. (2025). Statistik vs Machine Learning , Prediksi, Evaluasi dan Analisis Turnamen Bulutangkis Bougenville Cup 5 Berbasis Data Pertandingan. *Jurnal Jawara Sistem Informasi*, 3(1), 1–11.
- Noviati, Y., Khoniario, A. F., Syaputri, K. C., & Misidawati, D. N. (2025). Implementasi Data-Driven Decision Making Dalam Strategi Public Relation: Studi Kasus Pada Kampanye Media Sosial E-Commerce Tiktok Shop. *Socius: Jurnal Penelitian Ilmu-Ilmu Sosial*, 3(5), 75–83.
- Optarina, Y., Suarna, N., Bahtiar, A., Rahaningsih, N., & Prihartono, W. (2026). OPTIMASI MODEL XGBOOST UNTUK PREDIKSI PENYAKIT JANTUNG MENGGUNAKAN OPTUNA. *Jurnal Software Engineering and Information System (SEIS)*, 6(1), 50–55.
- Rusdianto Hidayat, R., Yersin Putri Larung, E., Sinaga, E., Sari Kardi, I., Fariz Prima Putra, M., Bari Samal, A., & Babingga, H. (2024). Analitik prediktif sepakbola: model machine learning bri liga 1 indonesia. *MULTILATERAL: Jurnal Pendidikan Jasmani Dan Olahraga*, 23(4), 386–399.
- Said, H. (2025). Prediksi Posisi Optimal Pemain Sepakbola berdasarkan Aspek Kepribadian Tim PON Sepak Bola Putra Jawa Timur. *Jurnal Ilmiah Multidisiplin*, 4(1), 670–678.
- Saputro, M. R., Mahdiyah, U., Swanjaya, D., Nusantara, U., & Kediri, P. (2024). Perbandingan Metode Adaptive Boosting dan Extreme Gradient Boosting Untuk Prediksi Hasil Pertandingan Liga Spanyol. *Jurnal Nusantara Of Engineering*, 7(1), 67–73. <https://ojs.unpkediri.ac.id/index.php/noe>
- Ulumando, M. I. (2026). Prediction Of Repeating Object-Oriented Programming Course for Informatics Students at ITAF Kupang Using Extreme Gradient Boosting (XGBoost). *PERFECT: Journal of Smart Algorithms*, 3(2), 70–83.
- Wicaksono, I. S., & Mustafa. (2025). Penerapan Algoritma Naive Bayes Untuk Klasifikasi Kualitas Pemain Sepak Bola Berdasarkan Data Statistik Pertandingan. *Jurnal Rekayasa Sistem Informasi Dan Teknologi*, 3(2), 284–305. <https://doi.org/10.70248/jrsit.v3i2.3036>
- Wijoyo A, Saputra A, Ristanti S, Sya'ban S, Amalia M, & Febriansyah R. (2024). Pembelajaran Machine Learning. *OKTAL (Jurnal Ilmu Komputer Dan Science)*, 3(2), 375–380. <https://journal.mediapublikasi.id/index.php/oktal/article/view/2305>
- Yuliana, Y., Robet, R., & Hoki, L. (2026). Comparative Analysis of XGBoost, KNN, and SVM Algorithms for Heart Disease Prediction Using SMOTE-Tomek Balancing. *Sinkron*, 10(1), 305–314. <https://doi.org/10.33395/sinkron.v10i1.15469>