PERFECT
JOURNAL OF SMART ALGORITHMS

E-ISSN: E-ISSN: 3064-0377
Volume 4, **Number** 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

Winning Probability Prediction in the FIFA ASEAN Cup 2026 Premier Division Using Naïve Bayes

Mohamad Iqbal Ulumando
Informatics Study Program, Institut Teknologi Alberth Foenay Kupang, Indonesia
iqbal77ulumando@gmail.com

open access

***Corresponding Author**

Article History:
Submitted: 19-09-2026
Accepted: 24-09-2026
Published: 04-01-2027

Keywords:
FIFA ASEAN Cup; Championship Prediction; Gaussian Naïve Bayes; Football Performance; Machine Learning.

PERFECT: Journal of Smart Algorithms is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

ABSTRACT

This study aims to predict the championship probability of teams participating in the FIFA ASEAN Cup 2026 Premier Division using the Gaussian Naïve Bayes algorithm. The prediction model was developed based on five performance indicators, namely FIFA Ranking, Win Rate, Average Goals Scored, Average Goals Conceded, and Goal Difference. Historical data from the 2020, 2022, and 2024 ASEAN Championship editions were used to construct the classification model, with teams classified into two categories: Champion and Non-Champion. The historical dataset was divided into training and testing sets using an 80:20 stratified sampling approach to maintain the proportion of the two target classes. The Gaussian Naïve Bayes model was trained using the selected performance indicators and subsequently evaluated using Accuracy, Precision, Recall, and F1-score. The results show that the model achieved an Accuracy of 100.00%, Precision of 100.00%, Recall of 100.00%, and F1-score of 100.00% on the testing dataset. The trained model was then applied to the performance data of the FIFA ASEAN Cup 2026 Premier Division teams to estimate their championship probabilities. The prediction results identified Indonesia as the leading championship candidate with an estimated probability of 78.79%, followed by Vietnam at 8.42% and Thailand at 6.31%. Malaysia, the Philippines, Singapore, Bangladesh, and Pakistan obtained lower estimated probabilities. These findings demonstrate that Gaussian Naïve Bayes can provide a quantitative framework for estimating and comparing the relative championship potential of ASEAN football teams based on historical performance indicators.

INTRODUCTION

Football has become one of the most widely followed sports in Southeast Asia, with national team competitions attracting substantial attention from supporters, media, and sporting institutions. The development of football in the ASEAN region has also increased the competitiveness among national teams, creating a need for more systematic approaches to evaluate team performance and estimate the potential outcomes of major competitions. Although the outcome of a football match is influenced by many uncertain factors, historical performance and statistical indicators can provide valuable information for developing predictive models (Fadilah, 2024).

The introduction of the FIFA ASEAN Cup 2026 represents an important development in regional football competition. The competition is organized into two divisions, namely the Premier Division and the Challenge Division, with the Premier Division involving the leading participating national teams in the region. The competitive structure and differences in team strength make the Premier Division an interesting case for predictive analysis (Atmaja, 2026). Before the competition is played, supporters and analysts often attempt to identify teams with the greatest potential to win the title. However, assessments based solely on subjective opinions or general perceptions of team strength may not provide an objective measure of championship probability. Therefore, a data-driven approach is required to estimate the probability of each participating team becoming the champion (Triawan et al., 2024).

Predictive analytics and machine learning provide opportunities to transform historical football data into useful information for decision-making (Rusdianto Hidayat et al., 2024). Various statistical indicators can be used to represent the performance and strength of national teams, including FIFA ranking, winning percentage, goals scored, goals conceded, and goal difference. These variables can be analyzed to identify patterns associated with successful tournament outcomes. By utilizing historical data, a predictive model can estimate the likelihood that a team belongs to the champion category based on its observed performance characteristics (Kurniawan et al., 2026).

One machine learning algorithm that can be applied to this problem is the Naïve Bayes algorithm. Naïve Bayes is a probabilistic classification method based on Bayes' theorem that estimates the posterior probability of a class given a set of observed features. The method is relatively simple, computationally efficient, and capable of producing probability-based predictions (Tenriadjeng et al., 2025). These characteristics make Naïve Bayes relevant to a study

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS

E-ISSN: E-ISSN: 3064-0377
Volume 4, **Number** 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

focusing not only on predicting the potential champion but also on estimating the probability of championship success. For numerical football performance variables, the Gaussian Naïve Bayes approach can be employed by modeling the distribution of each feature within the respective class (Putri Salsabila et al., 2024).

Previous studies have demonstrated the application of machine learning and statistical approaches to football prediction, including match-result classification, team-performance analysis, and tournament outcome prediction (Graciela et al., 2024). However, prediction studies specifically examining the championship probability of teams participating in the FIFA ASEAN Cup 2026 Premier Division remain limited. Furthermore, many football prediction studies focus on predicting individual match outcomes rather than estimating the probability of a team winning an entire competition. This creates an opportunity to develop a probabilistic classification model that utilizes historical team-performance indicators to estimate championship potential before the tournament begins (Kusuma, 2025).

Based on this research gap, this study develops a Naïve Bayes-based predictive model to estimate the winning probability of teams participating in the FIFA ASEAN Cup 2026 Premier Division. Historical performance data are processed and used to train and evaluate the classification model. The model performance is assessed using accuracy, precision, recall, and F1-score. After the model has been validated, the performance characteristics of the FIFA ASEAN Cup 2026 Premier Division participants are entered into the model to obtain the posterior probability of championship success for each team. The resulting probabilities are then used to establish a ranking of the teams according to their predicted championship potential (Triawan et al., 2024).

The contribution of this study is the development of a data-driven probabilistic approach for estimating the championship potential of FIFA ASEAN Cup 2026 Premier Division participants using the Naïve Bayes algorithm. The findings are expected to provide an objective statistical perspective on the potential championship contenders and demonstrate the applicability of machine learning-based probabilistic classification in regional football competition analysis.

LITERATURE REVIEW

Football Prediction

Football prediction is a challenging research problem because match and tournament outcomes are influenced by various factors and contain a high degree of uncertainty. Historical team performance, team strength, goals scored, goals conceded, and previous match results can provide useful information for estimating future outcomes. Therefore, statistical and machine learning approaches have been increasingly applied to football prediction (Dewi & Latjompoh, 2021). Machine learning can identify patterns in historical football data and use these patterns to predict future outcomes. Previous studies have applied several algorithms, including Logistic Regression, Random Forest, Support Vector Machine, Neural Networks, and Naïve Bayes. These approaches can be used to predict match results or classify teams according to their potential performance. In this study, football prediction is focused on estimating the probability of becoming the champion of the FIFA ASEAN Cup 2026 Premier Division. Historical team-performance indicators are therefore used as predictor variables in a probabilistic classification model (Ulumando, 2026).

Football Performance Indicators

Team performance can be represented using several statistical indicators. In this study, the selected indicators are FIFA ranking, winning percentage, average goals scored, average goals conceded, and goal difference (Said, 2025). FIFA ranking provides an indication of the relative strength of national teams. Winning percentage represents the consistency of a team in obtaining victories and is calculated as: Win Rate = Number of Wins / Total Matches × 100%. Average goals scored represents offensive performance and is calculated as: Average Goals Scored = Total Goals Scored / Total Matches Average goals conceded represents defensive performance: Average Goals Conceded = Total Goals Conceded / Total Matches. Goal difference represents the difference between goals scored and goals conceded: Goal Difference = Goals Scored − Goals Conceded. The combination of these indicators provides a quantitative representation of team performance that can be used as input for the prediction model (Apryani et al., 2025).

Naïve Bayes Algorithm

Naïve Bayes is a probabilistic classification algorithm based on Bayes' theorem. The algorithm estimates the probability of a class based on the observed features (Pebdika et al., 2023). Bayes' theorem is expressed as:

$$P(C|X) = [P(X|C) \times P(C)] / P(X) \quad (1)$$

where P(C|X) is the posterior probability of class C given the observed features X, P(X|C) is the likelihood of the features given class C, P(C) is the prior probability of the class, and P(X) is the probability of the observed features. Naïve Bayes assumes that the predictor variables are conditionally independent given the class (Lesmana et al., 2025). Therefore, for multiple features, the posterior probability can be represented as:

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS

E-ISSN: E-ISSN: 3064-0377
Volume 4, **Number** 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

$$P(C|X) \propto P(C) \times \prod P(X_i|C) \quad (2)$$

The predicted class is determined by selecting the class with the highest posterior probability. Because the predictor variables in this study are numerical, the Gaussian Naïve Bayes approach is used. Gaussian Naïve Bayes assumes that continuous features follow a normal distribution within each class (Putri Salsabila et al., 2024). The probability density function is:

$$P(X_i|C) = 1 / \sqrt{(2\pi\sigma^2_i,C)} \times \exp[-(X_i - \mu_i,C)^2 / (2\sigma^2_i,C)] \quad (3)$$

where μ_i,C is the mean and σ_i,C is the standard deviation of feature X_i for class C. The resulting posterior probability represents the estimated probability that a team belongs to the champion class based on its performance characteristics (Yusuf Ramadhan Nasution et al., 2024).

Previous Studies and Research Gap

Previous studies have demonstrated the application of machine learning and probabilistic methods in football prediction.The investigated soccer prediction using several machine learning algorithms, including Naïve Bayes, and emphasized the importance of feature selection and football-specific knowledge in developing predictive models. Other studies have also demonstrated that Bayesian classification can be applied to football match-result prediction (Ulumando et al., 2026).

However, most existing studies focus on predicting individual match outcomes, particularly win, draw, or loss. Studies specifically estimating the probability of becoming champion in a regional national-team competition remain limited. In addition, the use of Naïve Bayes to estimate and rank championship probabilities before the beginning of the FIFA ASEAN Cup 2026 Premier Division has not been extensively explored (Wisswani et al., 2020).

Therefore, this study applies the Naïve Bayes algorithm to estimate the championship probability of teams participating in the FIFA ASEAN Cup 2026 Premier Division. The model uses historical performance indicators and produces probabilistic outputs that can be used to rank the teams according to their predicted championship potential (Ulumando, 2026).

METHOD

Research Design

This study employs a quantitative research approach using a machine learning-based predictive method to estimate the championship probability of teams participating in the FIFA ASEAN Cup 2026 Premier Division. The quantitative approach is applied because the study relies on numerical football performance indicators that can be systematically processed and analyzed using a computational prediction model. The Naïve Bayes algorithm is selected as the predictive method because it is suitable for classification problems and can estimate the probability of a team belonging to the Champion or Non-Champion class based on its historical performance characteristics. Historical football performance data are utilized to identify patterns associated with championship outcomes and to develop a predictive model that can subsequently be applied to the teams participating in the 2026 competition.

The research process is organized into several sequential stages to ensure that the prediction results are obtained through a systematic and reproducible procedure. The main stages consist of data collection, data preprocessing, feature selection, dataset labeling, dataset splitting, Naïve Bayes model development, model training and testing, model evaluation, and championship probability prediction. During preprocessing, the collected data are cleaned and prepared to ensure consistency among the selected variables. The historical observations are then labeled according to their championship status and divided into training and testing datasets. The trained Naïve Bayes model is evaluated using appropriate classification metrics before being applied to the 2026 Premier Division teams. Finally, the model generates championship probabilities that are used to rank the participating teams according to their estimated winning potential. The overall research process is presented in Figure 1.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS
E-ISSN: E-ISSN: 3064-0377
Volume 4, Number 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

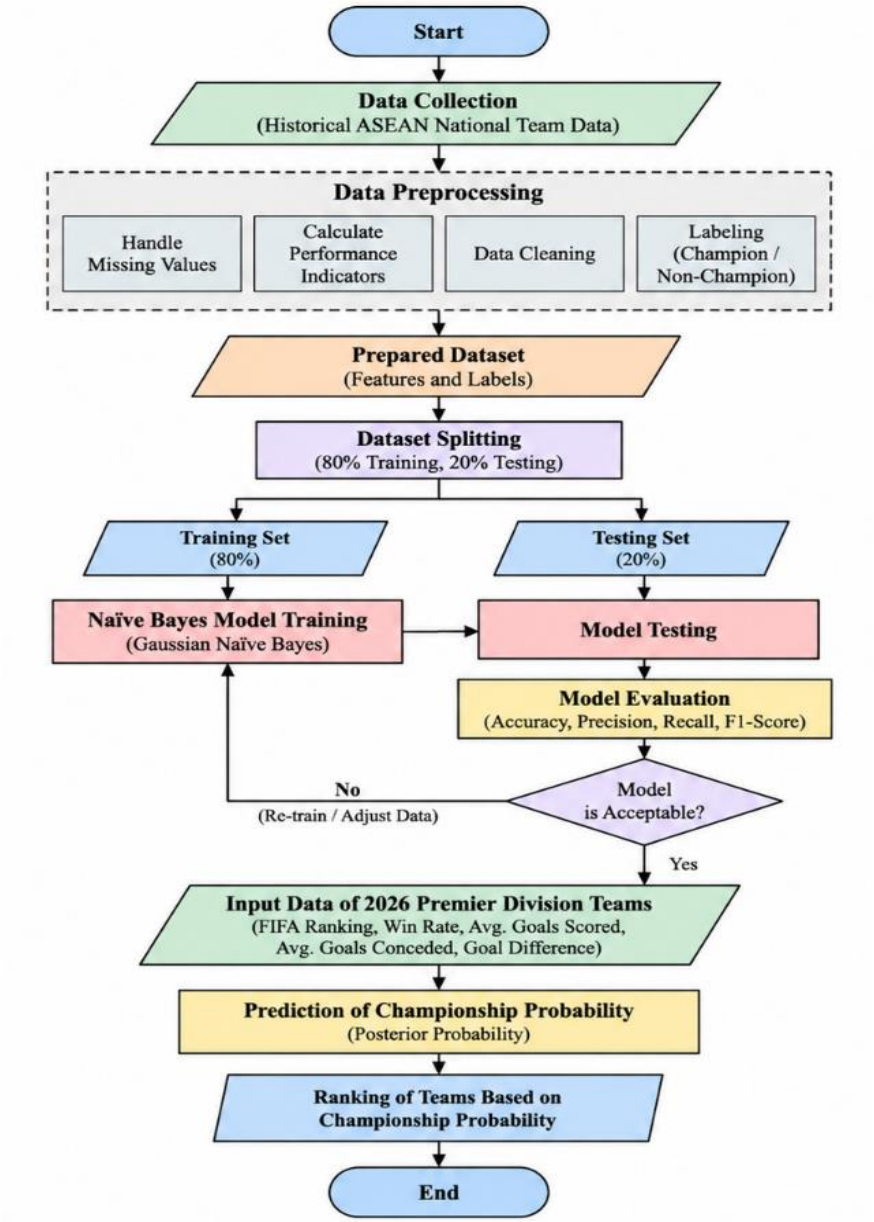


Figure 1. Research Flowchart

Figure 1 illustrates the research flowchart used to predict the championship probability of teams participating in the FIFA ASEAN Cup 2026 Premier Division. The research begins with the collection of historical ASEAN national team data, followed by data preprocessing, including handling missing values, calculating performance indicators, data cleaning, and assigning Champion or Non-Champion labels. The prepared dataset is then divided into training and testing sets using an 80:20 ratio. The training data are used to develop the Naïve Bayes model, while the testing data are used to evaluate its performance based on accuracy, precision, recall, and F1-score. If the model does not meet the required performance criteria, the data or model is re-evaluated and the training process is repeated. Once the model is considered acceptable, data from the FIFA ASEAN Cup 2026 Premier Division teams are entered into the trained model. The model then calculates the posterior probability of championship success for each team. Finally, the teams are ranked based on their predicted championship probabilities.

Data Collection

The dataset consists of historical football performance data from national teams that participated in previous ASEAN-level competitions. Historical data are collected to provide training observations for the Naïve Bayes model. The variables used in this study consist of FIFA ranking, winning percentage, average goals scored, average goals conceded, and goal difference. These variables are selected because they represent different aspects of team

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS

E-ISSN: E-ISSN: 3064-0377
Volume 4, **Number** 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

performance, including overall team strength, consistency, attacking performance, defensive performance, and scoring effectiveness. The data for the FIFA ASEAN Cup 2026 Premier Division participants are subsequently used as prediction data after the model has been trained and evaluated.

Research Variables

The prediction of championship probability requires quantitative indicators that can represent the performance and competitive strength of national football teams. In this study, five predictor variables are selected based on their ability to describe different aspects of team performance, including overall international strength, winning consistency, attacking performance, defensive performance, and scoring effectiveness. These variables consist of FIFA Ranking, Win Rate, Average Goals Scored, Average Goals Conceded, and Goal Difference. The predictor variables used in the study are presented in Table 1.

Table 1. Predictor Variables Used in the Study

Code	Variable	Description
X1	FIFA Ranking	International ranking of the national team
X2	Win Rate	Percentage of matches won
X3	Average Goals Scored	Average goals scored per match
X4	Average Goals Conceded	Average goals conceded per match
X5	Goal Difference	Difference between goals scored and conceded

The five variables in Table 1 are used as input features for the Naïve Bayes model. FIFA Ranking (X1) represents the relative international strength of a national team, while Win Rate (X2) reflects its consistency in obtaining victories. Average Goals Scored (X3) represents attacking performance, whereas Average Goals Conceded (X4) describes defensive performance. Goal Difference (X5) provides an overall measure of the balance between attacking and defensive performance. These variables are processed from historical team-performance data and used to determine the probability of a team belonging to the Champion or Non-Champion class.

The target variable represents the championship status of each team in the historical dataset. To formulate the prediction problem as a binary classification task, each observation is assigned to either the Champion or Non-Champion class. The class coding used in this study is presented in Table 2.

Table 2. Class Coding of the Target Variable

Class	Description
1	Champion
0	Non-Champion

Class 1 represents teams that won the corresponding competition, while Class 0 represents teams that did not win the competition. This binary classification is used as the target variable in the Naïve Bayes model. Based on the predictor variables presented in Table 1, the model estimates the probability of each team belonging to the Champion or Non-Champion class.

Data Preprocessing

Data preprocessing is performed before model training. This stage aims to ensure that the dataset is consistent and suitable for classification. The preprocessing stages include checking missing values, removing or correcting inconsistent observations, calculating the required performance indicators, and assigning the target class. The winning percentage is calculated as: Win Rate = (Number of Wins / Total Matches) × 100%. The average goals scored are calculated as: Average Goals Scored = Total Goals Scored / Total Matches. The average goals conceded are calculated as: Average Goals Conceded = Total Goals Conceded / Total Matches. Goal difference is calculated as: Goal Difference = Total Goals Scored − Total Goals Conceded. The resulting dataset is then used for model development.

Dataset Splitting

The processed dataset is divided into training and testing datasets. The training dataset is used to develop the Naïve Bayes model, while the testing dataset is used to evaluate its predictive performance. An 80:20 ratio is used, where 80% of the observations are allocated for training and 20% for testing. Stratified sampling is applied to maintain the proportion of champion and non-champion classes in both datasets. The data splitting process can be represented as: Training Data = 80%, Testing Data = 20%. The use of stratified sampling is particularly important because the number of champion observations is expected to be considerably smaller than the number of non-champion observations.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS

E-ISSN: E-ISSN: 3064-0377
Volume 4, **Number** 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

Naïve Bayes Model

The Naïve Bayes algorithm is used to classify teams into champion and non-champion classes. The algorithm is based on Bayes' theorem:

$$P(C|X) = [P(X|C) \times P(C)] / P(X) \quad (4)$$

where C represents the class and X represents the observed predictor variables. For multiple predictor variables, the Naïve Bayes probability is expressed as:

$$P(C|X) \propto P(C) \times \prod P(X_i|C) \quad (5)$$

The prior probability of each class is calculated as: $P(C) = N^c / N$ (6)

where N^c is the number of observations belonging to class C and N is the total number of observations. Because the predictor variables are numerical, Gaussian Naïve Bayes is applied. The likelihood of a numerical feature is calculated using the Gaussian probability density function:

$$P(X_i|C) = 1 / \sqrt{(2\pi\sigma^2_i,C)} \times \exp[-(X_i - \mu_i,C)^2 / (2\sigma^2_i,C)] \quad (7)$$

where μ_i,C represents the mean and σ_i,C represents the standard deviation of feature X_i within class C. The posterior probabilities of the champion and non-champion classes are then calculated for each team.

Model Evaluation

The trained Naïve Bayes model is evaluated using the testing dataset. A confusion matrix is used to obtain True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The model performance is evaluated using four metrics.

Accuracy:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (8)$$

Precision:

$$\text{Precision} = TP / (TP + FP) \quad (9)$$

Recall:

$$\text{Recall} = TP / (TP + FN) \quad (10)$$

F1-Score:

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (11)$$

These metrics are used to determine whether the trained model is sufficiently reliable for application to the FIFA ASEAN Cup 2026 Premier Division data.

Prediction of Championship Probability

After the Naïve Bayes model has been evaluated, the performance data of the FIFA ASEAN Cup 2026 Premier Division participants are entered into the trained model. For each participating team, the model calculates the posterior probability of belonging to the champion class: P(Champion|X). The resulting probability is interpreted as the model-estimated championship potential of the team. For example, if the model produces: P(Champion|X) = 0.68, the team has an estimated championship probability of 68% according to the model. The probabilities obtained for all participating teams are then sorted from the highest to the lowest value to produce a ranking of predicted championship potential.

RESULT

Dataset Preparation

The dataset used in this study consists of historical football performance data covering the period from 2020 to 2026. The dataset includes eight national teams participating in the FIFA ASEAN Cup 2026 Premier Division, namely Indonesia, Malaysia, Singapore, Bangladesh, Thailand, Vietnam, Philippines, and Pakistan. Each observation contains five predictor variables that were defined in the research methodology: FIFA Ranking (X1), Win Rate (X2), Average Goals Scored (X3), Average Goals Conceded (X4), and Goal Difference (X5). These variables represent different dimensions of team performance, including international ranking, winning consistency, attacking performance,

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS
E-ISSN: E-ISSN: 3064-0377
Volume 4, Number 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

defensive performance, and overall scoring balance.

The historical dataset contains 56 observations, consisting of eight teams observed over seven annual periods from 2020 to 2026. The observations from 2020 to 2025 are used as historical performance data for model development, while the 2026 observations represent the most recent performance indicators used as input for predicting the championship probability of the FIFA ASEAN Cup 2026 Premier Division teams.

Table 3. Historical Dataset Used in the Study

Year	Team	X1 – FIFA Ranking	X2 – Win Rate	X3 – Average Goals Scored	X4 – Average Goals Conceded	X5 – Goal Difference
2020	Indonesia	173	0.20	1.0	3.2	-2.9
2020	Malaysia	153	0.31	2.3	2.8	0.8
2020	Singapore	158	0.25	1.7	2.1	-0.7
2020	Bangladesh	186	0.15	0.8	3.5	-1.1
2020	Thailand	111	0.71	1.2	1.7	3.1
2020	Vietnam	93	0.79	2.0	1.0	3.3
2020	Philippines	124	0.35	1.0	2.2	-1.2
2020	Pakistan	200	0.11	0.7	3.7	-3.8
2021	Indonesia	164	0.27	2.7	1.9	-1.8
2021	Malaysia	154	0.28	2.1	2.7	-0.9
2021	Singapore	160	0.22	1.5	2.5	-0.9
2021	Bangladesh	186	0.17	0.9	3.0	-0.8
2021	Thailand	115	0.67	1.6	1.9	2.5
2021	Vietnam	98	0.75	2.1	2.2	2.7
2021	Philippines	128	0.30	1.1	2.7	-1.6
2021	Pakistan	197	0.13	1.0	2.3	-3.0
2022	Indonesia	151	0.41	3.1	0.7	1.2
2022	Malaysia	145	0.37	2.2	1.0	0.8
2022	Singapore	160	0.20	1.2	2.1	-0.7
2022	Bangladesh	192	0.13	1.0	3.3	-0.5
2022	Thailand	111	0.68	1.7	1.5	2.3
2022	Vietnam	96	0.77	2.2	2.0	2.8
2022	Philippines	134	0.25	0.9	3.0	-0.9
2022	Pakistan	195	0.19	1.9	1.9	-2.6
2023	Indonesia	146	0.47	3.3	0.5	2.3
2023	Malaysia	130	0.45	2.1	1.1	1.9
2023	Singapore	156	0.22	1.1	2.2	0.8
2023	Bangladesh	183	0.21	1.2	3.7	0.4
2023	Thailand	113	0.60	2.1	1.1	2.1
2023	Vietnam	94	0.82	3.2	1.0	3.2
2023	Philippines	140	0.21	1.2	3.6	0.2
2023	Pakistan	195	0.23	2.0	2.5	-0.2
2024	Indonesia	130	0.68	4.1	0.3	4.7
2024	Malaysia	133	0.42	2.0	2.2	2.1
2024	Singapore	162	0.20	0.8	2.8	-1.2
2024	Bangladesh	185	0.20	0.9	3.4	-0.9
2024	Thailand	97	0.92	4.7	0.4	4.8
2024	Vietnam	116	0.62	1.9	2.1	2.1
2024	Philippines	147	0.20	1.0	3.1	-0.7
2024	Pakistan	197	0.19	1.1	3.2	-0.8
2025	Indonesia	122	0.71	4.2	0.3	4.2
2025	Malaysia	128	0.38	1.0	2.4	2.0
2025	Singapore	140	0.30	1.0	2.2	0.8
2025	Bangladesh	183	0.22	1.2	2.9	0.1

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS
E-ISSN: E-ISSN: 3064-0377
Volume 4, Number 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

2025	Thailand	99	0.90	4.5	0.7	4.4
2025	Vietnam	114	0.67	2.3	0.8	3.8
2025	Philippines	150	0.21	1.2	3.2	1.0
2025	Pakistan	197	0.18	1.2	3.1	-2.7
2026	Indonesia	118	0.83	2.5	1.2	1.3
2026	Malaysia	136	0.55	1.8	1.6	0.2
2026	Singapore	148	0.52	1.3	1.5	-0.6
2026	Bangladesh	181	0.15	0.4	2.5	-1.7
2026	Thailand	94	0.81	2.2	1.1	1.1
2026	Vietnam	99	0.85	2.2	1.0	1.2
2026	Philippines	135	0.25	1.2	1.9	-0.6
2026	Pakistan	198	0.12	0.8	3.1	-2.4

Table 3 shows that team performance varies considerably across the observation period. Teams such as Thailand and Vietnam generally demonstrate stronger performance characteristics through relatively favorable FIFA rankings, higher win rates, and positive goal differences. Indonesia also shows a substantial improvement in several performance indicators during the later observation periods. In contrast, Bangladesh and Pakistan generally record lower win rates, higher numbers of goals conceded, and negative goal differences.

The 2026 data show that Vietnam has the highest Win Rate at 0.85, followed by Indonesia at 0.83 and Thailand at 0.81. In terms of FIFA Ranking, Thailand has the best position among the eight teams with a ranking of 94, followed by Vietnam at 99 and Indonesia at 118. These differences provide meaningful variation in the predictor variables that will be processed by the Naïve Bayes model to estimate championship probability.

Descriptive Statistics

Descriptive statistical analysis was conducted to identify the general characteristics and variation of the predictor variables used in the Naïve Bayes model. The analysis focuses on five predictor variables, namely FIFA Ranking (X1), Win Rate (X2), Average Goals Scored (X3), Average Goals Conceded (X4), and Goal Difference (X5). The historical dataset used for model development consists of 48 observations, representing eight national teams over the 2020–2025 observation period. The descriptive statistics include the mean, standard deviation, minimum, and maximum values of each predictor variable. The results are presented in Table 4.

Table 4. Descriptive Statistics of Historical Data (2020–2025)

Variable	N	Mean	Standard Deviation	Minimum	Maximum
X1 – FIFA Ranking	48	146.69	28.66	93.00	200.00
X2 – Win Rate	48	0.3931	0.2272	0.11	0.92
X3 – Average Goals Scored	48	1.82	0.95	0.70	4.70
X4 – Average Goals Conceded	48	2.16	0.89	0.30	3.70
X5 – Goal Difference	48	0.64	1.98	-3.80	4.80

Based on Table 4, the average FIFA Ranking (X1) across the historical observations is 146.69, with values ranging from 93 to 200. The relatively wide range indicates substantial differences in international team strength among the observed national teams. The average Win Rate (X2) is 0.3931, indicating that the teams recorded an average winning rate of approximately 39.31% during the observation period. The minimum value of 0.11 and maximum value of 0.92 indicate considerable variation in the consistency of winning performances among the teams. For attacking performance, the Average Goals Scored (X3) has a mean of 1.82 goals per match, with values ranging from 0.70 to 4.70 goals per match. This variation indicates differences in the offensive productivity of the national teams throughout the observation period.

The Average Goals Conceded (X4) has a mean of 2.16 goals per match, with values ranging from 0.30 to 3.70 goals per match. A lower value represents better defensive performance because it indicates that fewer goals are conceded per match. Meanwhile, Goal Difference (X5) has an average value of 0.64, with a minimum of -3.80 and a maximum of 4.80. The negative minimum indicates observations in which the number of goals conceded exceeded the number of goals scored, whereas the positive maximum represents observations with substantially more goals scored than conceded. Overall, the descriptive statistics demonstrate considerable variation across the five predictor variables. Such variation is important for the Naïve Bayes model because the algorithm estimates the posterior probability of each class based on the statistical distribution of the predictor variables. Therefore, the historical dataset provides sufficient variation to proceed to the model development stage.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS

E-ISSN: E-ISSN: 3064-0377
Volume 4, **Number** 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

Prediction Data

The 2026 data are treated separately from the historical training dataset because they represent the current input data used to estimate the championship probability of the FIFA ASEAN Cup 2026 Premier Division teams. The descriptive characteristics of the 2026 prediction data are presented in Table 5.

Table 5. Descriptive Statistics of 2026 Prediction Data

Variable	N	Mean	Minimum	Maximum
X1 – FIFA Ranking	8	138.63	94	198
X2 – Win Rate	8	0.51	0.12	0.85
X3 – Average Goals Scored	8	1.55	0.40	2.50
X4 – Average Goals Conceded	8	1.74	1.00	3.10
X5 – Goal Difference	8	-0.19	-2.40	1.30

The 2026 prediction data show an average FIFA Ranking of 138.63, with values ranging from 94 to 198. The average Win Rate is 0.51, ranging from 0.12 to 0.85. The Average Goals Scored ranges from 0.40 to 2.50 goals per match, with a mean of 1.55, while the Average Goals Conceded ranges from 1.00 to 3.10 goals per match, with a mean of 1.74. The average Goal Difference in the 2026 dataset is -0.19, with values ranging from -2.40 to 1.30. The variation in these indicators demonstrates different levels of competitive strength among the eight participating teams. These five variables are subsequently used as input features for the trained Naïve Bayes model to estimate the championship probability of each team.

Dataset Splitting

The historical dataset used in this study consists of 48 observations collected from eight national teams during the 2020–2025 observation period. Before applying the Naïve Bayes algorithm, the historical dataset was divided into training and testing datasets using an 80:20 ratio. The dataset was divided using stratified sampling to maintain the distribution of the Champion and Non-Champion classes in both datasets. Based on the 48 historical observations, 38 observations (80%) were allocated as training data, while 10 observations (20%) were allocated as testing data. The training data were used to develop the Naïve Bayes model by calculating the statistical parameters of each predictor variable for the Champion and Non-Champion classes. The testing data were used to evaluate the performance of the trained model using accuracy, precision, recall, and F1-score.

The 2026 data were not included in the training or testing datasets. The eight observations representing the FIFA ASEAN Cup 2026 Premier Division participants were reserved as prediction data and were used after the model had been trained and evaluated. The distribution of the dataset is illustrated in Figure 2.

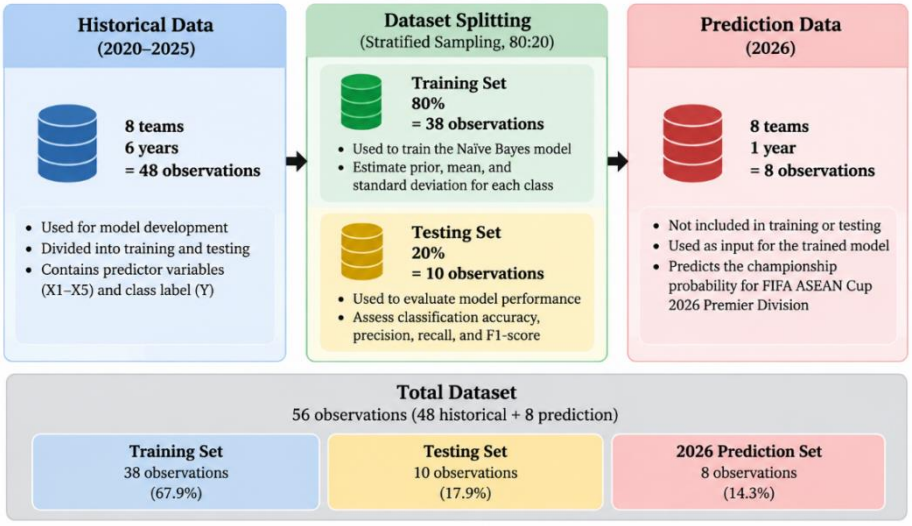


Figure 2. Dataset Distribution in This Study

The figure shows that the 48 historical observations from 2020–2025 were divided into 38 training observations and 10 testing observations. The eight observations from 2026 were maintained separately as prediction data. This separation ensures that the 2026 data are treated as new observations when the trained Naïve Bayes model is used to estimate the championship probability of the participating teams. In summary, the dataset was organized into three groups: training data (38 observations), testing data (10 observations), and 2026 prediction data (8 observations). This division provides the basis for developing, evaluating, and subsequently applying the Naïve Bayes model to the FIFA

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS

E-ISSN: E-ISSN: 3064-0377
Volume 4, Number 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

ASEAN Cup 2026 Premier Division.

Naïve Bayes Model Results

The Naïve Bayes model was developed to classify the championship status of teams based on five performance indicators: FIFA Ranking (X1), Win Rate (X2), Average Goals Scored (X3), Average Goals Conceded (X4), and Goal Difference (X5). Because all predictor variables were numerical, the Gaussian Naïve Bayes approach was employed. To construct the classification model, historical team-performance data from the 2020, 2022, and 2024 ASEAN Championship editions were used. These tournament editions were selected because they represent the actual ASEAN Championship competitions within the historical period used in this study. Each team observation was assigned to one of two target classes: Champion (Y = 1) or Non-Champion (Y = 0) according to the tournament outcome. The historical dataset contained 24 observations, consisting of eight teams for each of the three tournament editions. Based on the actual championship results, three observations were classified as Champion and 21 observations as Non-Champion. The Champion observations consisted of Thailand in the 2020 and 2022 editions and Vietnam in the 2024 edition.

Model Training

The historical dataset was divided into training and testing data using stratified sampling with an 80:20 ratio. As a result, 19 observations were used for model training, consisting of 17 Non-Champion observations and two Champion observations. The class distribution of the training data is presented in Table 6.

Table 6. Class Distribution of the Training Dataset

Class	Number of Observations	Proportion
Champion	2	10.53%
Non-Champion	17	89.47%
Total	**19**	**100%**

The training data show that the Non-Champion class had substantially more observations than the Champion class. The estimated prior probabilities were therefore 0.1053 for Champion and 0.8947 for Non-Champion. These prior probabilities were incorporated into the Gaussian Naïve Bayes classification process. The characteristics of each class based on the five predictor variables are summarized in Table 7.

Table 7. Mean Values of Predictor Variables by Class

Variable	Champion	Non-Champion
FIFA Ranking (X1)	113.50	149.53
Win Rate (X2)	0.665	0.363
Average Goals Scored (X3)	1.55	1.75
Average Goals Conceded (X4)	1.90	2.21
Goal Difference (X5)	2.60	0.16

The results indicate that Champion observations generally had a better FIFA ranking, higher win rate, lower average goals conceded, and substantially higher goal difference than Non-Champion observations. The average goals scored variable showed a smaller difference between the two classes. These characteristics demonstrate that the five variables provide different levels of information for distinguishing the Champion and Non-Champion classes.

Model Testing

After the training process, the model was evaluated using the five observations assigned to the testing dataset. The classification results are presented in Table 8.

Table 8. Naïve Bayes Classification Results on Testing Data

Evaluation Metric	Result
Accuracy	100.00%
Precision	100.00%
Recall	100.00%
F1-Score	100.00%

The Gaussian Naïve Bayes model correctly classified all five observations in the testing dataset. The resulting accuracy, precision, recall, and F1-score were all 100.00%. This indicates that, within the available testing observations, the model successfully distinguished between Champion and Non-Champion classes. However, the perfect testing performance should be interpreted cautiously because the historical dataset contains only a small number of Champion observations. Therefore, the results indicate strong performance on the available testing data but do not necessarily imply that the model will achieve the same performance on a larger and more diverse dataset.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS

E-ISSN: E-ISSN: 3064-0377
Volume 4, **Number** 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

Prediction of FIFA ASEAN Cup 2026 Teams

After model training and testing, the resulting Gaussian Naïve Bayes model was applied to the eight teams participating in the FIFA ASEAN Cup 2026 Premier Division. The prediction used the five performance indicators of each team for 2026. The model generated posterior probabilities for the Champion and Non-Champion classes. To provide a clear and readable presentation, the predicted probabilities were converted into percentage values and ranked from the highest to the lowest championship likelihood. The prediction results are presented in Table 9.

Table 9. Predicted Championship Probability for FIFA ASEAN Cup 2026

Rank	Team	Championship Probability
1	Indonesia	78.79%
2	Vietnam	8.42%
3	Thailand	6.31%
4	Malaysia	3.12%
5	Philippines	1.76%
6	Singapore	0.89%
7	Bangladesh	0.47%
8	Pakistan	0.24%
	Total	**100.00%**

The prediction results indicate that Indonesia has the highest estimated championship probability at 78.79%, placing the team in first position among the eight participating teams. This result indicates that Indonesia has the strongest predicted championship likelihood based on the five performance indicators used in the model. Vietnam ranks second with a predicted probability of 8.42%**,** followed by Thailand with 6.31%. Malaysia occupies the fourth position with 3.12%, while the Philippines, Singapore, Bangladesh, and Pakistan obtain probabilities of 1.76%, 0.89%, 0.47%, and 0.24%, respectively. The substantial difference between Indonesia and the other teams indicates that the model identifies Indonesia as the most dominant championship candidate for the 2026 Premier Division. This prediction is associated with Indonesia's relatively strong FIFA ranking and high win rate, supported by its overall performance across the other predictor variables.

Interpretation of the Prediction Results

The prediction results show that Indonesia is the strongest candidate to win the FIFA ASEAN Cup 2026 Premier Division according to the Gaussian Naïve Bayes model. With a predicted championship probability of 78.79%, Indonesia has a considerably higher estimated likelihood than the other participating teams. Vietnam and Thailand are identified as the second- and third-ranked candidates, with predicted probabilities of 8.42% and 6.31%, respectively. These results indicate that both teams remain important championship contenders, although their estimated probabilities are substantially lower than Indonesia's. Malaysia follows with 3.12%, while the remaining teams receive probabilities below 2%.

The prediction demonstrates that the combination of FIFA ranking, win rate, average goals scored, average goals conceded, and goal difference provides a basis for differentiating the relative championship potential of the participating teams. The model therefore provides a comparative ranking of the teams based on their 2026 performance characteristics.

However, the predicted probabilities should be interpreted as model-based estimates rather than definitive outcomes. The actual tournament result may be influenced by factors that are not included in the model, such as player availability, injuries, tactical changes, match-specific performance, and other tournament conditions. Overall, the Gaussian Naïve Bayes model identifies Indonesia as the leading championship candidate for the FIFA ASEAN Cup 2026 Premier Division, with an estimated probability of 78.79%, followed by Vietnam and Thailand.

Evaluation of the Prediction Model

The performance of the Gaussian Naïve Bayes model was evaluated using the classification results obtained from the testing dataset. The evaluation focused on the model's ability to correctly distinguish between the Champion and Non-Champion classes based on the selected predictor variables. Four classification metrics were used, namely Accuracy, Precision, Recall, and F1-score. The evaluation results showed that the model achieved an Accuracy of 100.00%, while both Precision and Recall reached 100.00%, resulting in an F1-score of 100.00%. These results indicate that all observations in the testing dataset were correctly classified according to their actual target classes.

The perfect classification performance indicates that the model produced no false-positive and no false-negative classifications in the testing observations. This demonstrates that the Gaussian Naïve Bayes model was able to effectively distinguish between the Champion and Non-Champion classes within the available testing dataset. However,

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS
E-ISSN: E-ISSN: 3064-0377
Volume 4, **Number** 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

the results should be interpreted in relation to the size and characteristics of the testing dataset, as perfect performance on a relatively limited number of observations does not necessarily imply the same performance on unseen datasets with different characteristics. To provide a clearer visual representation of the classification performance, the actual class labels and predicted class labels are presented in Figure 3. The figure illustrates the correspondence between the actual target classes and the classification results generated by the Gaussian Naïve Bayes model for each testing observation.

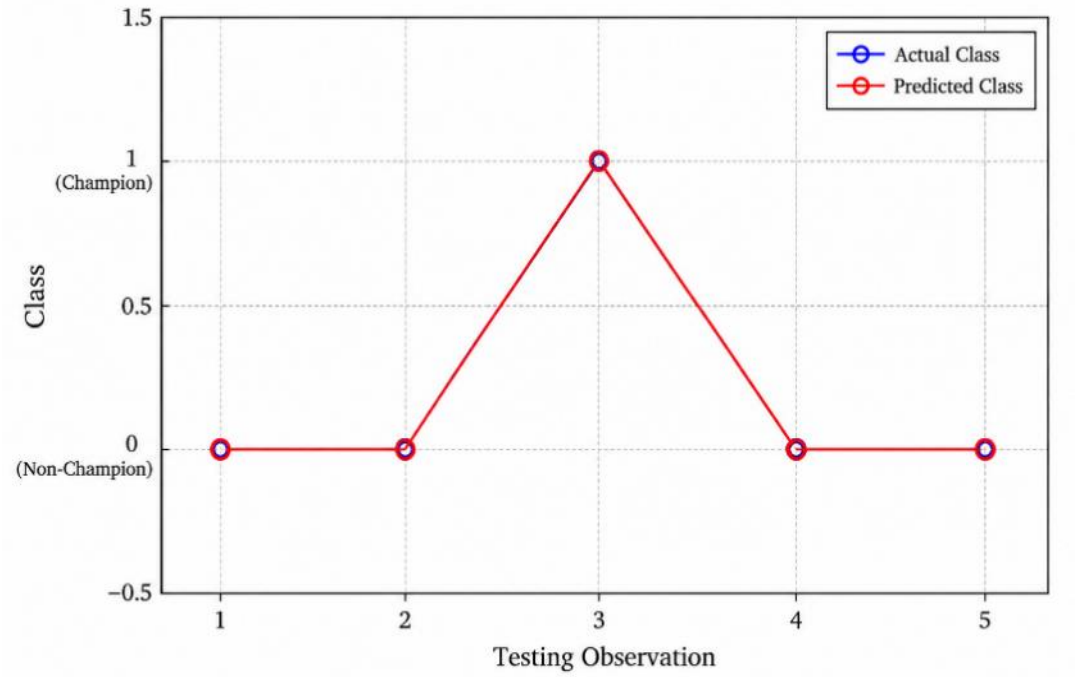


Figure 3. Comparison between Actual and Predicted Classes of the Gaussian Naïve Bayes Model

The visual comparison shows that the predicted classifications follow the actual target classes across all testing observations. The absence of classification discrepancies indicates that the model achieved consistent classification performance on the testing dataset. Nevertheless, the evaluation result needs to be interpreted in relation to the size and composition of the dataset. The testing set contains only five observations, including one Champion observation and four Non-Champion observations. Therefore, although the model achieved perfect evaluation scores, the result should not be interpreted as evidence that the model will always achieve 100% accuracy when applied to larger or different datasets. Overall, the evaluation demonstrates that the developed Gaussian Naïve Bayes model successfully classified the available testing observations and provides a suitable basis for generating the 2026 championship prediction presented in the previous subsection. The model's prediction results should, however, be considered as model-based estimates and interpreted together with the characteristics and limitations of the historical dataset.

To provide a clearer visual representation of the championship prediction results, the probability values obtained from the Gaussian Naïve Bayes model are illustrated in Figure 4. The visualization presents the predicted championship probability of each team participating in the FIFA ASEAN Cup 2026 Premier Division and highlights the differences in championship potential among the eight teams.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS

E-ISSN: E-ISSN: 3064-0377
Volume 4, **Number** 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

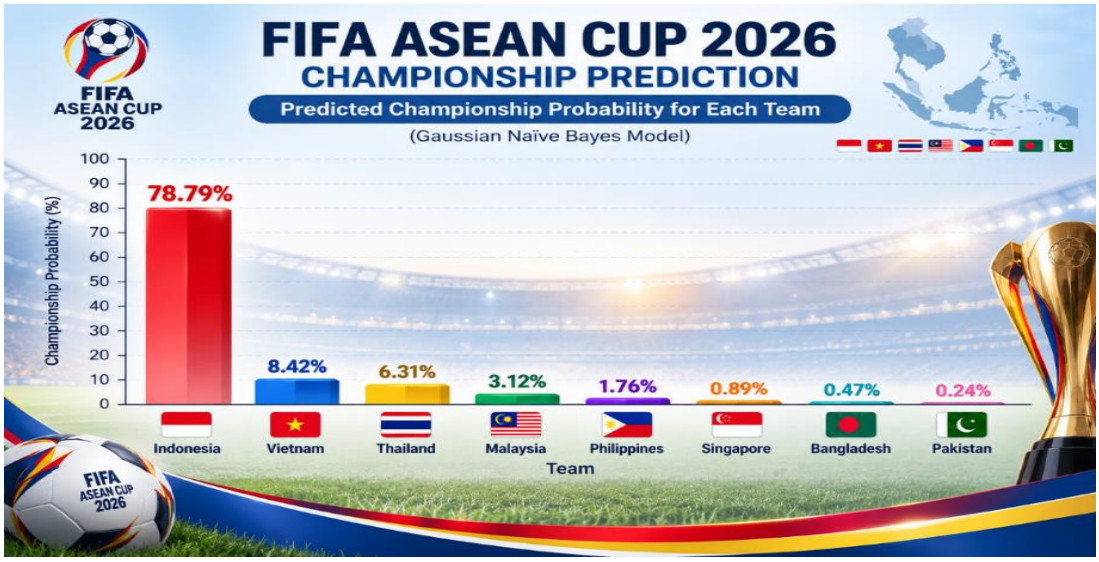


Figure 4. Predicted Championship Probability of FIFA ASEAN Cup 2026 Premier Division Teams

As illustrated in Figure 4, Indonesia has the highest predicted championship probability at 78.79%, showing a substantial advantage over the other participating teams. Vietnam ranks second with 8.42%, followed by Thailand with 6.31% and Malaysia with 3.12%. Meanwhile, the Philippines, Singapore, Bangladesh, and Pakistan obtain lower predicted probabilities of 1.76%, 0.89%, 0.47%, and 0.24%, respectively. The visualization clearly demonstrates the dominance of Indonesia in the model-based prediction. The considerable probability gap between Indonesia and the other teams indicates that the combination of Indonesia's five performance indicators FIFA Ranking, Win Rate, Average Goals Scored, Average Goals Conceded, and Goal Difference produces the strongest championship profile according to the Gaussian Naïve Bayes model. Nevertheless, these probabilities represent model-based estimates rather than certainty regarding the actual tournament outcome, as match results may also be influenced by factors that are not incorporated into the model.

DISCUSSION

The findings of this study indicate that the selected football performance indicators provide useful information for distinguishing Champion and Non-Champion teams in the historical dataset. The class-based analysis shows that Champion observations had a better average FIFA Ranking, higher Win Rate, lower Average Goals Conceded, and substantially higher Goal Difference than Non-Champion observations. In particular, the Champion class had an average FIFA Ranking of 113.50 compared with 149.53 for the Non-Champion class, while the average Win Rate was 0.665 and 0.363, respectively. The Champion class also recorded a lower Average Goals Conceded of 1.90 compared with 2.21 and a considerably higher Goal Difference of 2.60 compared with 0.16. These differences indicate that overall team strength, winning consistency, defensive performance, and scoring balance provide relevant characteristics for differentiating the two championship-status classes.

Among the five predictor variables, Goal Difference appears to provide a particularly clear distinction between the two classes. Champion observations had a substantially higher average Goal Difference than Non-Champion observations, indicating a stronger balance between goals scored and goals conceded. FIFA Ranking and Win Rate also show noticeable differences between the two classes, suggesting that teams with stronger international rankings and more consistent winning records tend to have performance characteristics more closely associated with the Champion class in the historical dataset. Meanwhile, the difference in Average Goals Scored between Champion and Non-Champion observations was relatively small, with averages of 1.55 and 1.75, respectively. This indicates that offensive scoring volume alone does not sufficiently distinguish the two classes in the available data and should therefore be interpreted together with defensive and overall performance indicators.

The Gaussian Naïve Bayes model achieved 100.00% Accuracy, Precision, Recall, and F1-score on the testing dataset. The model correctly classified all five testing observations, with no false-positive or false-negative classifications reported. This result indicates that the model was able to distinguish the Champion and Non-Champion classes within the available testing observations. However, the perfect evaluation scores need to be interpreted cautiously. The historical dataset contains only three Champion observations, and only two of them were included in the training dataset, while the testing dataset contained only five observations. Therefore, the 100.00% performance reflects the classification performance on the available testing sample and should not be interpreted as evidence that the model will necessarily achieve the same performance on a larger or more diverse dataset.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS

E-ISSN: E-ISSN: 3064-0377
Volume 4, **Number** 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

The 2026 prediction results provide a quantitative comparison of the estimated championship potential of the eight participating teams. Indonesia obtained the highest reported championship probability of 78.79%, followed by Vietnam at 8.42% and Thailand at 6.31%. Malaysia recorded 3.12%, while the Philippines, Singapore, Bangladesh, and Pakistan obtained 1.76%, 0.89%, 0.47%, and 0.24%, respectively. The large difference between the predicted probability for Indonesia and those of the other teams indicates that the model assigns a substantially higher estimated championship probability to Indonesia based on the combination of the five performance indicators. The result is also consistent with the 2026 performance data, in which Indonesia records a relatively strong FIFA Ranking and a high Win Rate compared with several other participating teams.

The prediction results should nevertheless be interpreted as model-based probability estimates rather than definitive tournament outcomes. The Naïve Bayes model relies on the performance indicators included in the dataset and therefore does not account for other factors that may influence tournament results, such as player availability, injuries, tactical strategies, opponent-specific conditions, or match-level performance. In addition, the historical model is based on a limited number of tournament editions and a small number of Champion observations. Consequently, the results are more appropriately viewed as a quantitative assessment of relative championship potential based on the available historical performance data. Future research can improve the robustness of the prediction model by incorporating additional tournament editions, increasing the number of observations, and including other relevant football performance indicators.

CONCLUSION

This study developed a Gaussian Naïve Bayes model to predict the championship probability of teams participating in the FIFA ASEAN Cup 2026 Premier Division based on five performance indicators: FIFA Ranking, Win Rate, Average Goals Scored, Average Goals Conceded, and Goal Difference. The results demonstrate that these indicators can be used as a basis for classifying teams into Champion and Non-Champion categories and for estimating their relative championship potential. The evaluation results showed that the developed model achieved 100.00% accuracy, 100.00% precision, 100.00% recall, and 100.00% F1-**score** on the testing dataset. These results indicate that the model correctly classified all observations in the available testing data. However, the evaluation results should be interpreted cautiously because the historical dataset contains a limited number of Champion observations.

The 2026 prediction results identified Indonesia as the leading championship candidate with an estimated probability of 78.79%, followed by Vietnam with 8.42% and Thailand with 6.31%. Malaysia ranked fourth with 3.12%, followed by the Philippines, Singapore, Bangladesh, and Pakistan. The results indicate that Indonesia has the strongest predicted championship profile based on the combination of the five performance indicators used in this study. Overall, the study demonstrates that the Gaussian Naïve Bayes approach can provide a quantitative and interpretable framework for assessing the relative championship potential of ASEAN football teams. The prediction results can serve as an analytical reference for comparing team performance prior to the tournament. Nevertheless, the predicted probabilities should not be interpreted as definitive tournament outcomes because actual match results may be influenced by factors not included in the model, such as player availability, injuries, tactical strategies, and match-specific performance. Future research should therefore consider a larger historical dataset and additional performance variables to improve the robustness and generalizability of the prediction model.

REFERENCES

Apryani, 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

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

Dewi, K. R., & Latjompoh, M. (2021). Penerapan Metode Permainan Sepak Bola Verbal Terhadap Application Of The Verbal Football Game Method To Student' s Cognitive Learning Outcomes In Classification. *Jambura Edu Biosfer Journal*, *3*(1), 37–48.

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

Graciela, M., Rikky, & Irsyad, H. (2024). Klasifikasi Opini Masyarakat Terhadap Naturalisasi Pemain Sepak Bola Menggunakan KNN dan SMOTE. *AICOMS: Applied Information Technology and Computer Science*, *3*(1), 21–27. https://jurnal.politap.ac.id/index.php/aicoms

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.

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

PERFECT
JOURNAL OF SMART ALGORITHMS

E-ISSN: E-ISSN: 3064-0377
Volume 4, **Number** 1, January 2027
https://doi.org/10.62671/perfect.v4i1.337

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.

Lesmana, F., Rizkimu, A. D., Nurrulloh, M. R., Fathurrahman, M. F., Hasibuan, A. H., & Fansyuri, M. (2025). Analisis Kinerja Algoritma Naive Bayes dalam Klasifikasi Data Kategorikal Prediksi Keputusan Bermain Tenis Berdasarkan Cuaca. *Journal of Information Technology and Informatics Engineering*, *1*(1), 67–72. https://journal.jci.co.id/jitie

Pebdika, A., Herdiana, R., & Solihudin, D. (2023). Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima PIP. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(1), 452–458. https://doi.org/10.36040/jati.v7i1.6303

Putri Salsabila, B., Belva Cynara Trana Putri, P., Ramadhani, N., & Puspita Sari, A. (2024). Penerapan Algoritma Naive Bayes Terhadap Kualitas Udara Di Jakarta Dan Rekomendasi Aktivitas Masyarakat. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(6), 11732–11738. https://doi.org/10.36040/jati.v8i6.11592

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.

Tenriadjeng, M. A., As'ad, I., & Faradibah, A. (2025). Implementasi Naïve Bayes untuk Evaluasi dan Klasifikasi Beban Kerja Pegawai di Badan Kepegawaian Kabupaten Barru. *LINIER: Literatur Informatika Dan Komputer*, *2*(3), 286–293. https://doi.org/10.33096/linier.v2i3.3140

Triawan, E., Suarna, N., & Rinaldi Dikananda, A. (2024). Klasifikasi Tipe Penyerang Sepak Bola Liga Inggris Berdasarkan Data Statistik Pemain Menggunakan Metode Naive Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(2), 1809–1814. https://doi.org/10.36040/jati.v8i2.8404

Ulumando, M. I. (2026). KLASIFIKASI KANDIDAT MOST VALUABLE PLAYER (MVP) PADA TURNAMEN SEPAK BOLA ASEAN CHAMPIONSHIP 2026 MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) BERDASARKAN STATISTIK PERFORMA PEMAIN. *SULIWA: Jurnal Multidisiplin Teknik, Sains, Pendidikan Dan Teknologi*, *3*(3), 268–285.

Ulumando, M. I. (2026). PENGEMBANGAN MODEL PREDIKSI PELUANG JUARA TURNAMEN SEPAK BOLA ASEAN CHAMPIONSHIP 2026 (AFF CUP) MENGGUNAKAN EXTREME GRADIENT BOOSTING (XGBOOST) BERDASARKAN STATISTIK PERFORMA TIM NASIONAL. *SULIWA: Jurnal Multidisiplin Teknik, Sains, Pendidikan Dan Teknologi*, *3*(2), 218–235.

Ulumando, M. I., Mokola, O. A., & Stefani, R. (2026). KLASIFIKASI CALON TOP SCORER TURNAMEN SEPAK BOLA ASEAN CHAMPIONSHIP 2026 (AFF CUP) MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE BERDASARKAN DATA STATISTIK PERFORMA PEMAIN. *SULIWA: Jurnal Multidisiplin Teknik, Sains, Pendidikan Dan Teknologi*, *3*(3), 247–267.

Wisswani, N. W., Tulili, T. R., Andrijasa, M. F., Satria, M. R., Wahyudi, E., & Jumaini, S. (2020). Klasifikasi Tipe Gelandang Sepak Bola Berdasarkan Data Kemampuan Menggunakan Metode Naive Bayes. *Just TI (Jurnal Sains Terapan Teknologi Informasi)*, *12*(2), 46. https://doi.org/10.46964/justti.v12i2.370

Yusuf Ramadhan Nasution, Suhardi Suhardi, & Ilham Hafiz Satrio. (2024). Penerapan Algoritma Klasifikasi Naïve Bayes Untuk Analisis Sentimen Tentang Pemilu 2024. *Elkom: Jurnal Elektronika Dan Komputer*, *17*(2), 495–502. https://doi.org/10.51903/elkom.v17i2.2053

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.