Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 3, July 2026
https://doi.org/ 10.47709/cnahpc.v8i3.9162

Submitted : 2026-07-08
Accepted : 2026-07-21
Published : 2026-07-23

Comparative Evaluation of Naïve Bayes and Support Vector Machine for Human Development Index Classification

Dwi Asa Verano[1)], Eka Putri Apriliani[2)], Zahratul Aliah[3)], Sarah Nur Arna Cholifah[4)], Amanda Dwi Tio Agustin[5)], Thomas[6)], Shinta Puspasari[7)]
[1,2,3,4,5,6,7)]Universitas Indo Global Mandiri, Indonesia
[1)]dwiasa@uigm.ac.id, [2)]2023110121@students.uigm.ac.id, [3)]2023110112@students.uigm.ac.id,
[3)]2023110122@students.uigm.ac.id, [3)]2023110119@students.uigm.ac.id,
[3)]2023110036@students.uigm.ac.id, [3)]shinta@uigm.ac.id,

ABSTRACT

The Human Development Index (HDI) is one of the most widely used indicators for evaluating the quality of human development through health, education, and living standards. Accurate HDI classification is essential for supporting evidence-based public policy and regional development planning. Recent advances in machine learning provide opportunities to improve classification performance compared with conventional statistical approaches. This study aims to compare the performance of the Naïve Bayes and Support Vector Machine (SVM) algorithms for classifying the Human Development Index across Indonesian provinces and to identify the more effective classification model. The study employed a quantitative experimental approach using the 2024 Human Development Index dataset published by Statistics Indonesia (BPS), covering all 38 provinces. Four predictor variables were used: life expectancy at birth, expected years of schooling, mean years of schooling, and adjusted expenditure per capita. The dataset was preprocessed before being divided into training and testing sets using an 80:20 ratio. Both algorithms were implemented using Python's Scikit-learn library and evaluated through confusion matrix analysis, accuracy, precision, recall, and F1-score. Experimental results indicate that both algorithms successfully classified provincial HDI categories. However, SVM consistently outperformed Naïve Bayes, achieving an accuracy of 100% on the held-out test set (5-fold cross-validated mean accuracy of 86.4%), compared with 87.5% (cross-validated mean of 59.6%) obtained by Naïve Bayes. SVM also produced higher precision, recall, and F1-score, indicating stronger classification capability for multidimensional socioeconomic data. The findings demonstrate that Support Vector Machine provides superior performance for HDI classification compared with Naïve Bayes. This study contributes empirical evidence regarding the application of machine learning techniques in socioeconomic indicator classification and offers practical insights for decision support in regional development policy.

Keywords: Human Development Index; Machine Learning; Naïve Bayes; Support Vector Machine; Socioeconomic Classification

INTRODUCTION

Human development has become one of the most important indicators for measuring the success of national development because it reflects not only economic growth but also improvements in the quality of human life. The Human Development Index (HDI), introduced by the United Nations Development Programme (UNDP), evaluates human development through three fundamental dimensions: a long and healthy life, access to education, and decent standard of living. (United Nations Development Programme (UNDP), 2020).

In Indonesia, the HDI is officially published by Statistics Indonesia (BPS) and has become an important reference for evaluating regional development performance, identifying disparities among provinces, and formulating evidence-based public policies. Although Indonesia has experienced continuous improvement in the HDI over the past decade, significant disparities remain across provinces, indicating that regional socioeconomic development is still uneven. (Badan Pusat Statistik, 2024).

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 3, July 2026
https://doi.org/ 10.47709/cnahpc.v8i3.9162

Submitted : 2026-07-08
Accepted : 2026-07-21
Published : 2026-07-23

The rapid growth in data availability has encouraged researchers to employ machine learning techniques for socioeconomic analysis and decision support. Compared with conventional statistical approaches, machine learning algorithms can discover complex and nonlinear relationships among multiple variables while maintaining high predictive performance. Consequently, machine learning has been widely adopted in healthcare, education, economics, agriculture, and regional development. (Atmojo et al., 2024)(Fahmiyah & Ningrum, 2023)(Nurafidah et al., 2023).

Among supervised learning algorithms, Naïve Bayes and Support Vector Machine (SVM) are two of the most widely used classification methods because of their efficiency and effectiveness. Naïve Bayes is computationally simple, requires relatively small training data, and performs well for probabilistic classification problems. However, its assumption of conditional independence among predictor variables may reduce classification accuracy when strong correlations exist among socioeconomic indicators. (Anggara et al., 2024).

The Support Vector Machine (SVM) has demonstrated superior performance for multidimensional classification tasks because it constructs an optimal hyperplane that maximizes the separation margin between classes. When paired with a nonlinear kernel such as RBF or Polynomial, this capability can extend to datasets containing nonlinear relationships between predictors and classes; with a linear kernel, as implemented in the present study, the same margin-maximization principle instead provides robust separation for classes that are linearly related but overlapping. Therefore, SVM has become one of the most frequently applied machine learning algorithms in socioeconomic and public policy studies. (Pamungkas & Widiyanto, 2022).

Several studies have successfully applied machine learning techniques to classify socioeconomic conditions and regional development indicators. Some studies have reported that Naïve Bayes provides satisfactory classification performance because of its computational efficiency, whereas other studies have concluded that SVM generally achieves higher accuracy for complex multidimensional datasets. Nevertheless, previous research has predominantly focused on educational, financial, medical, or demographic datasets, while comparative studies specifically addressing HDI classification using the latest official Indonesian provincial data are limited. (Dewi et al., 2025)(Syafitri et al., 2021).

Furthermore, several of the most closely related HDI classification studies reviewed below—particularly those comparing multiple algorithms on a single dataset—report classification performance primarily in terms of accuracy, with comparatively limited discussion of precision, recall, F1-score, and confusion matrix analysis. Consequently, policymakers receive insufficient empirical evidence when selecting the most appropriate machine learning algorithm to support regional development planning. (Arifat et al., 2023).

This limitation represents the research gap addressed in the present study. Unlike previous studies, the central novelty of this study is a comprehensive comparative evaluation between Naïve Bayes and support vector machines, built on the latest Human Development Index dataset published by Statistics Indonesia, covering all 38 Indonesian provinces. The comparison was conducted using multiple evaluation metrics, including accuracy, precision, recall, F1-score, and confusion matrix analysis.

Building on this core novelty, the study makes three contributions. First, it provides the comparative evaluation itself, using the official 2024 Indonesian HDI dataset and the full set of metrics described above. Second, because the complete experimental pipeline (data splitting, category thresholds, model configuration, and evaluation procedure) is fully specified, this same evaluation offers a transparent and reproducible basis for future replication. Third, the resulting comparison translates directly into a practical recommendation for policymakers on which algorithm is more suitable for classifying provincial HDI status.

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 3, July 2026
https://doi.org/ 10.47709/cnahpc.v8i3.9162

Submitted : 2026-07-08
Accepted : 2026-07-21
Published : 2026-07-23

LITERATURE REVIEW

Recent studies have demonstrated the growing adoption of machine learning techniques for socioeconomic classification and decision-support systems. Various supervised learning algorithms have been evaluated to improve classification accuracy, particularly for multidimensional datasets such as poverty, education, healthcare, and Human Development Index (HDI). Although these studies employed different datasets and evaluation strategies, they consistently reported that machine learning algorithms are capable of supporting evidence-based policy formulation more effectively than conventional analytical approaches. (Zuhdi et al., 2025).

Developed a comparative classification model for the Human Development Index in Indonesia using Decision Tree C4.5, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes (NB), and Extreme Gradient Boosting (XGBoost). Their study used 548 district-level observations and reported that SVM achieved an accuracy of 97%, outperforming Naïve Bayes (90%) and Decision Tree (86%). However, the primary objective of the study was to identify the best algorithm among several classifiers rather than conducting an in-depth comparative analysis between Naïve Bayes and SVM. Consequently, the characteristics, strengths, and weaknesses of these two algorithms were not discussed comprehensively. (Ramadhan et al., 2025).

Compared six machine learning algorithms for classifying Indonesian districts and cities based on the 2021 Human Development Index. Their experimental results indicated that SVM with a radial basis function (RBF) kernel produced the highest average classification accuracy among the evaluated algorithms. Although the study demonstrated the effectiveness of SVM for HDI classification, it focused on district-level data from 2021 and did not evaluate more recent provincial data. Furthermore, the study emphasized overall classification accuracy without providing a detailed comparison of precision, recall, and F1-score between Naïve Bayes and SVM. (Dewi et al., 2025).

Recent comparative studies conducted in other application domains also consistently reported that Support Vector Machine generally outperformed Naïve Bayes for complex classification problems because of its ability to construct optimal decision boundaries, whereas Naïve Bayes remained attractive due to its computational simplicity and shorter training time (Astrianka & Efendi, 2024). These findings suggest that algorithm selection should consider not only predictive accuracy but also dataset characteristics, computational efficiency, and model interpretability. (Almeyda et al., 2025).

Based on the reviewed studies, several research gaps can be identified**.** First, previous studies primarily focused on identifying the single best-performing algorithm instead of conducting a comprehensive comparison between Naïve Bayes and Support Vector Machine. Second, most studies employed datasets collected before 2024 or focused on district-level observations rather than the latest provincial Human Development Index dataset published by Statistics Indonesia. Third, as illustrated by Zuhdi et al. (2025) and Ramadhan et al. (2025) above, the studies most directly comparable to the present one reported overall classification accuracy while providing limited discussion of precision, recall, F1-score, and confusion matrix analysis. These limitations indicate the need for a more comprehensive comparative evaluation using the latest official HDI dataset to provide stronger empirical evidence regarding the performance of Naïve Bayes and Support Vector Machine for socioeconomic classification.

METHOD

This study adopts a quantitative research design based on a data mining approach, employing the Naïve Bayes and Support Vector Machine (SVM) algorithms to classify the Human Development Index (HDI) status of Indonesian provinces. The overall research procedure follows six sequential stages: data collection, data preprocessing, data preparation, modelling using Naïve Bayes and SVM, model evaluation, and conclusion, as illustrated in Figure 1.

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 3, July 2026
https://doi.org/ 10.47709/cnahpc.v8i3.9162

Submitted : 2026-07-08
Accepted : 2026-07-21
Published : 2026-07-23

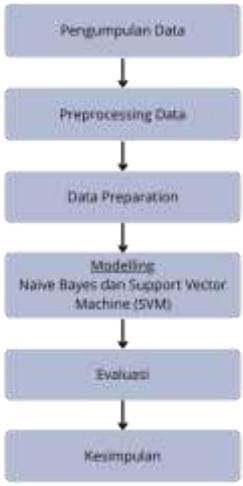

Figure 1. Research Flow Diagram

DATA COLLECTION

The dataset used in this study is secondary data obtained from the official website of Statistics Indonesia (Badan Pusat Statistik) for the year 2024, covering the survey results of all 38 Indonesian provinces (Badan Pusat Statistik, 2024). The dependent variable is the categorical Human Development Index (HDI) status, while the independent variables consist of life expectancy at birth, expected years of schooling, mean years of schooling, and adjusted expenditure per capita. The complete list of research variables is presented in Table 1.

Table 1. Research Variables

Variable	Description	Unit	Scale
Y	Human Development Index (HDI)	Index	Ordinal
X1	Life Expectancy at Birth	Years	Ratio
X2	Expected Years of Schooling	Years	Ratio
X3	Mean Years of Schooling	Years	Ratio
X4	Adjusted Real Expenditure per Capita	Rupiah	Ratio

DATA PREPROCESSING

Data preprocessing was carried out to prepare the raw dataset for analysis. The dataset was loaded into a pandas DataFrame, followed by an inspection of data structure, descriptive statistics, and the detection and handling of missing values and outliers. Categorical labels were subsequently transformed into numerical form using Label Encoding so that the data could be processed by the classification algorithms. The relationships among numerical features were also examined through pairwise visualization to support feature understanding prior to modelling. In addition, the four predictor variables were standardized (zero mean, unit variance) prior to model training, as described under Implementation Details below.

DATA SPLITTING

Prior to splitting, the continuous IPM score of each province was discretized into three classification categories (Low, Medium, High) using tercile thresholds, which yielded 13, 12, and 13 provinces respectively. Tercile-based thresholds were adopted instead of the official BPS status bands (e.g., IPM < 60 = Low) because the latter produced a highly imbalanced distribution (32 High, 5 Medium, 1 Low out of 38 provinces) that is unsuitable for multi-class classification with a dataset of this size. The preprocessed dataset of 38 provinces was then divided into training and testing subsets using an 80:20 ratio

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 3, July 2026
https://doi.org/ 10.47709/cnahpc.v8i3.9162

Submitted : 2026-07-08
Accepted : 2026-07-21
Published : 2026-07-23

with a fixed random state (random_state = 42), yielding 30 training samples and 8 testing samples, stratified by class to preserve the category proportions. To assess robustness beyond a single train-test split, 5-fold stratified cross-validation was additionally performed on the full dataset, and a paired t-test was used to evaluate whether the accuracy difference between the two algorithms across folds was statistically significant. This split allows the resulting models to be evaluated objectively and ensures adequate generalization capability on previously unseen data.

CLASSIFICATION ALGORITHMS

Two supervised classification algorithms were implemented using Python's Scikit-learn library: Gaussian Naïve Bayes and Support Vector Machine (SVM) with a linear kernel. Naïve Bayes is a probabilistic classifier derived from Bayes' theorem, which estimates the posterior probability of a class given a set of predictor variables under the assumption of conditional independence among features. Support Vector Machine, in contrast, constructs an optimal separating hyperplane that maximizes the margin between classes in an n-dimensional feature space, using a kernel function to transform the data when necessary (Cortes & Vapnik, 1995). Table 2 summarizes the kernel functions commonly used in SVM.

Table 2. SVM Kernel Functions

Kernel	Equation
Linear	K(xi, x) = xiT x
Polynomial	K(xi, x) = (γ · xiT x + r)P, γ > 0
RBF	K(xi, x) = exp(-γ |xi - x|2), γ > 0
Sigmoid	K(xi, x) = tanh(γ xiT x + r)

Among the kernel functions listed in Table 2, this study implements SVM with a linear kernel. This choice was made because the four predictor variables in this study exhibit an approximately linear relationship with the HDI target categories, so a linear decision boundary was expected to be sufficient without the added risk of overfitting that more flexible kernels (RBF, Polynomial, Sigmoid) can introduce on a small dataset of 38 observations. The linear kernel also keeps the decision boundary directly interpretable in terms of the original predictor variables, which is useful for a policy-oriented application such as HDI classification. A systematic comparison of kernel functions was outside the scope of this study and is identified as a direction for future research in the Discussion section.

Implementation details. To support replication, the complete configuration used in this study is reported here rather than in a separate code repository. Both models were implemented in Python 3 using scikit-learn: Naïve Bayes as GaussianNB() with default priors, and SVM as SVC(kernel='linear') with the default regularization parameter (C = 1.0). Prior to model fitting, all four predictor variables (X1-X4) were standardized to zero mean and unit variance using StandardScaler, fitted on the training set and applied unchanged to the test set to avoid data leakage; this step is particularly important for SVM, whose decision boundary is sensitive to feature scale. Data were split using train_test_split(test_size=0.2, random_state=42, stratify=y), and the supplementary cross-validation used StratifiedKFold(n_splits=5, shuffle=True, random_state=42). HDI categories were derived using tercile thresholds (pandas.qcut with 3 bins) applied to the 2024 IPM column described in Table 1. With this configuration fully specified, the reported results can be reproduced directly from the official BPS dataset without requiring supplementary pseudocode or an external repository.

MODEL EVALUATION

Model performance was evaluated and compared using a confusion matrix together with accuracy, precision, recall, and F1-score, computed for three HDI categories: Low, Medium, and High. The confusion matrix framework used in this study is presented in Table 3.

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 3, July 2026
https://doi.org/ 10.47709/cnahpc.v8i3.9162

Submitted : 2026-07-08
Accepted : 2026-07-21
Published : 2026-07-23

Table 3. Confusion Matrix Framework

	Predicted: True	Predicted: False	Total
Actual: True	TP	FN	P
Actual: False	FP	TN	N
Total	P'	N'	P+N

In this framework, True Positive (TP) and True Negative (TN) represent correctly classified positive and negative observations, respectively, while False Positive (FP) and False Negative (FN) represent misclassified negative and positive observations, respectively. These four values form the basis for calculating accuracy, precision, recall, and F1-score, which together provide a comprehensive assessment of classification performance for both algorithms.

To assess whether the observed performance difference between Naïve Bayes and SVM was robust rather than an artifact of a single 80:20 split, 5-fold stratified cross-validation was performed on the full 38-province dataset, and a paired t-test on the resulting fold-level accuracies was used to test whether the difference between the two algorithms was statistically significant. Receiver Operating Characteristic (ROC) curves and the Area Under the Curve (AUC) were not computed in this study: the target variable has three classes (Low, Medium, High), which requires a one-vs-rest ROC formulation, and with only 8 observations in the held-out test set, class-wise ROC curves would be highly unstable and not meaningfully interpretable. Accuracy, precision, recall, F1-score, the confusion matrix, cross-validated accuracy, and the paired t-test were therefore judged to provide a more reliable evaluation for a dataset of this size.

RESULT

DESCRIPTIVE STATISTICS

Prior to classification, the 38-province dataset was examined descriptively to characterize the distribution of each predictor and the target variable, as summarized in Table 4. Life expectancy at birth (X1) ranged from 67.39 to 75.99 years, with a mean of 72.88 years, reflecting variation in health outcomes across provinces. Expected years of schooling (X2) ranged from 9.63 to 15.70 years (mean = 13.21 years), while mean years of schooling (X3) ranged from 4.21 to 11.49 years (mean = 8.84 years), indicating persistent disparities in educational access and attainment. Adjusted expenditure per capita (X4) showed the widest relative disparity, ranging from Rp5,707,000 to Rp19,953,000 (mean = Rp11,591,974), reflecting substantial economic inequality among provinces. The HDI (Y) itself ranged from 54.43 to 84.15, with a mean of 73.50, indicating that most provinces fall within the medium-to-high human development category, while several provinces remain considerably below the national average.

Table 4. Descriptive Statistics of Predictor Variables

Statistic	X1	X2	X3	X4
Minimum	67.39	9.63	4.21	5,707,000
Maximum	75.99	15.70	11.49	19,953,000
Mean	72.88	13.21	8.84	11,591,974
Std. Dev.	2.30	1.02	1.23	2,463,450
Count	38	38	38	38

NAÏVE BAYES CLASSIFICATION RESULTS

The Gaussian Naïve Bayes model was trained on the training set and evaluated on the testing set for the three HDI categories: Low, Medium, and High. As shown in Table 5, the model achieved an overall accuracy of 0.88 on the 8-observation test set, with a weighted precision of 0.91, recall of 0.88, and F1-score of 0.86. Classification performance was perfect for the Low category (F1-score = 1.00) and weakest for the Medium category (F1-score = 0.67), where one of the two Medium-status test provinces was

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 3, July 2026
https://doi.org/ 10.47709/cnahpc.v8i3.9162

Submitted : 2026-07-08
Accepted : 2026-07-21
Published : 2026-07-23

misclassified as High, suggesting that the model had greater difficulty distinguishing provinces with borderline development status.

Table 5. Classification Report – Naïve Bayes

Class	Precision	Recall	F1-Score	Support
Low	1.00	1.00	1.00	3
Medium	1.00	0.50	0.67	2
High	0.75	1.00	0.86	3
Accuracy			0.88	8
Macro Avg	0.92	0.83	0.84	8
Weighted Avg	0.91	0.88	0.86	8

The confusion matrix for the Naïve Bayes model, presented in Figure 3, shows that the model's only misclassification occurred within the Medium category, where one Medium-status province was predicted as High, which is consistent with the lower recall (0.50) obtained for the Medium category. The Low and High categories were each classified with perfect precision and recall.

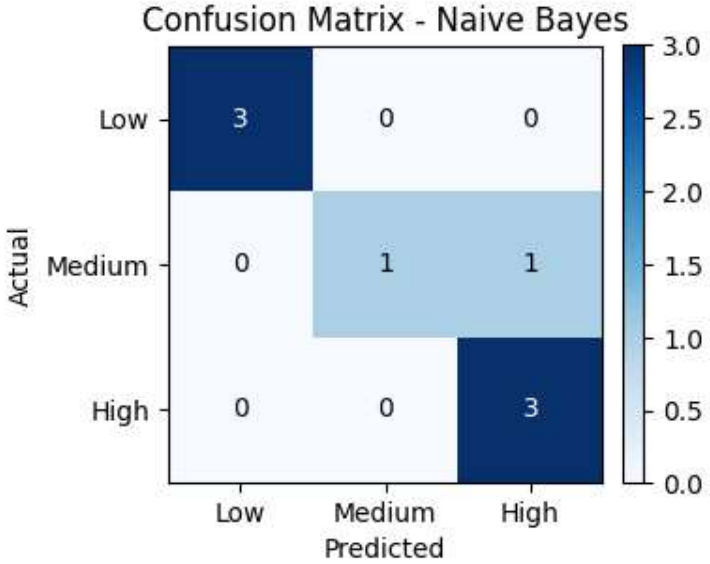

Figure 3. Confusion Matrix – Naïve Bayes

SUPPORT VECTOR MACHINE CLASSIFICATION RESULTS

The linear-kernel Support Vector Machine (SVM) model was trained and evaluated using the same training and testing configuration. As presented in Table 6, SVM achieved a higher overall accuracy of 1.00 on the same 8-observation test set, with a weighted precision, recall, and F1-score of 1.00, correctly classifying every test province across all three categories, including the Medium-status province that Naïve Bayes misclassified.

Table 6. Classification Report – Support Vector Machine

Class	Precision	Recall	F1-Score	Support
Low	1.00	1.00	1.00	3
Medium	1.00	1.00	1.00	2
High	1.00	1.00	1.00	3
Accuracy			1.00	8
Macro Avg	1.00	1.00	1.00	8
Weighted Avg	1.00	1.00	1.00	8

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 3, July 2026
https://doi.org/ 10.47709/cnahpc.v8i3.9162

Submitted : 2026-07-08
Accepted : 2026-07-21
Published : 2026-07-23

The confusion matrix for the SVM model, shown in Figure 4, indicates zero misclassified observations, compared with one misclassification for Naïve Bayes, in which the sole error occurred within the Medium category: SVM correctly classified 2 out of 2 actual Medium-status test provinces, compared with 1 out of 2 for Naïve Bayes, which misclassified the remaining province as High.

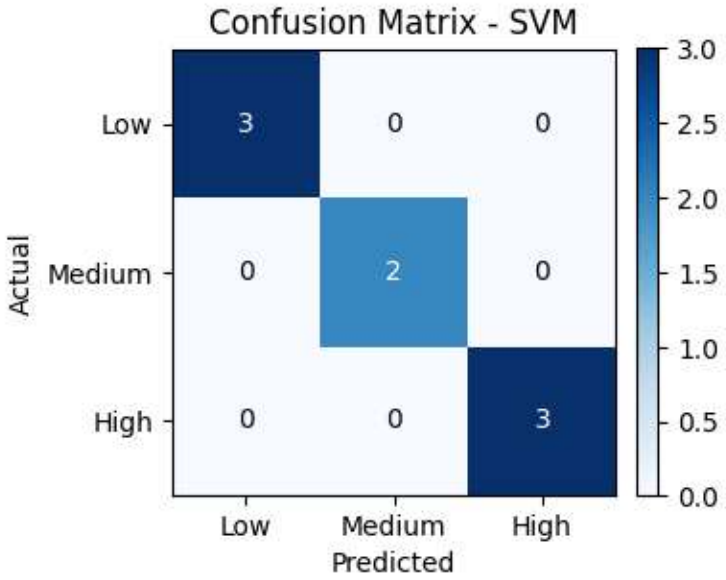

Figure 4. Confusion Matrix – Support Vector Machine

COMPARISON BETWEEN NAÏVE BAYES AND SVM

Figure 5 compares the number of correct (Positive) and incorrect (Negative) predictions produced by both algorithms, showing that SVM produced more correct predictions and fewer errors than Naïve Bayes.

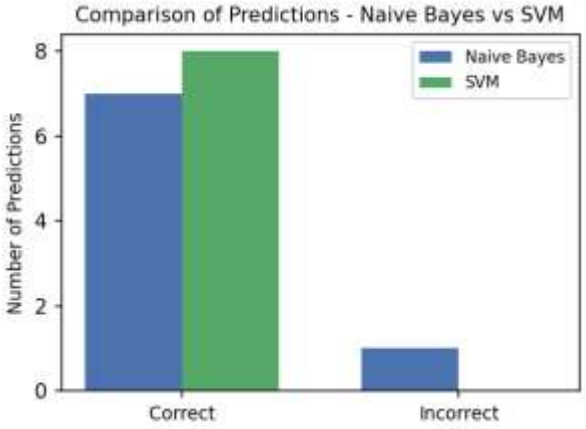

Figure 5. Comparison of Predictions – Naïve Bayes vs SVM

Table 7 summarizes the overall evaluation metrics for both algorithms based on the held-out test set. SVM consistently outperformed Naïve Bayes across all four metrics, with the largest gap observed in F1-score (a difference of 0.14), while the smallest gap was observed in precision (a difference of 0.09). Because a single 8-observation test split can be sensitive to which provinces happen to fall into the test set, these results were additionally verified using 5-fold stratified cross-validation on the full 38-province

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 3, July 2026
https://doi.org/ 10.47709/cnahpc.v8i3.9162

Submitted : 2026-07-08
Accepted : 2026-07-21
Published : 2026-07-23

dataset, which yielded a mean accuracy of 0.86 (± 0.09) for SVM and 0.60 (± 0.14) for Naïve Bayes. A paired t-test across the five folds confirmed that this difference was statistically significant ($t = 4.28$, $p = 0.013$). These results indicate that, for this dataset, SVM provides more reliable and consistent classification of provincial HDI status than Naïve Bayes.

Table 7. Comparison of Evaluation Metrics between Naïve Bayes and SVM

Metric	Naïve Bayes	SVM	Difference
Accuracy	0.88	1.00	0.12
Precision	0.91	1.00	0.09
Recall	0.88	1.00	0.12
F1-Score	0.86	1.00	0.14

Table 8 lists the condensed Python implementation used to generate the results reported in Tables 4-7 and Figures 3-5, consistent with the configuration described under Implementation Details above

Table 8. Python Implementation Code (Reproducibility)

```
import pandas as pd
from sklearn.model_selection import train_test_split,
    StratifiedKFold, cross_val_score
from sklearn.naive_bayes import GaussianNB
from sklearn.svm import SVC
from sklearn.pipeline import Pipeline
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score, classification_report
from scipy import stats

# 1. Categorize IPM into 3 balanced classes (tercile thresholds)
df['label'] = pd.qcut(df['Y_IPM'], 3, labels=['Low','Medium','High'])
X = df[['X1_UHH','X2_HLS','X3_RLS','X4_Pengeluaran']].values
y = df['label'].values

# 2. 80:20 stratified train/test split
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42, stratify=y)

# 3. Standardize features (fit on train, apply to test)
scaler = StandardScaler()
X_train_s = scaler.fit_transform(X_train)
X_test_s = scaler.transform(X_test)

# 4. Train and evaluate Naive Bayes
nb = GaussianNB().fit(X_train_s, y_train)
print(classification_report(y_test, nb.predict(X_test_s)))

# 5. Train and evaluate SVM (linear kernel)
svm = SVC(kernel='linear').fit(X_train_s, y_train)
print(classification_report(y_test, svm.predict(X_test_s)))

# 6. 5-fold stratified cross-validation (robustness check)
skf = StratifiedKFold(n_splits=5, shuffle=True, random_state=42)
nb_pipe = Pipeline([('scale', StandardScaler()), ('clf', GaussianNB())])
svm_pipe = Pipeline([('scale', StandardScaler()), ('clf', SVC(kernel='linear'))])
nb_scores = cross_val_score(nb_pipe, X, y, cv=skf, scoring='accuracy')
svm_scores = cross_val_score(svm_pipe, X, y, cv=skf, scoring='accuracy')

# 7. Paired t-test on fold-level accuracies (statistical significance)
t_stat, p_value = stats.ttest_rel(svm_scores, nb_scores)
```

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 3, July 2026
https://doi.org/ 10.47709/cnahpc.v8i3.9162

Submitted : 2026-07-08
Accepted : 2026-07-21
Published : 2026-07-23

DISCUSSIONS

The experimental results presented in Figure 5 and Table 7 demonstrate that Support Vector Machine (SVM) consistently outperformed Naïve Bayes across all evaluation metrics, including accuracy, precision, recall, and F1-score. This finding indicates that SVM provides better classification capability for Human Development Index (HDI) data, which consist of multidimensional socioeconomic indicators. The superior performance of SVM can be explained by its ability to construct an optimal separating hyperplane that maximizes the margin between different classes. In contrast, Naïve Bayes relies on the assumption of conditional independence among predictor variables, whereas HDI indicators such as life expectancy, expected years of schooling, mean years of schooling, and adjusted expenditure per capita are naturally correlated. As a result, the probabilistic assumption adopted by Naïve Bayes may reduce its effectiveness in distinguishing provinces with similar socioeconomic characteristics (Cortes & Vapnik, 1995)(Zhang, 2004).

The confusion matrix analysis further supports these findings by showing that SVM produced fewer classification errors than Naïve Bayes, particularly for provinces belonging to adjacent HDI categories. This suggests that SVM provides stronger generalization capability when handling observations with overlapping characteristics. Meanwhile, Naïve Bayes correctly classified most observations but showed greater difficulty in separating provinces with similar development profiles. These results indicate that evaluating classification models solely based on overall accuracy may not be sufficient, as confusion matrix analysis provides additional insight into the distribution and pattern of classification errors.

The findings of this study are consistent with several previous studies. (Ramadhan et al., 2025) reported that Support Vector Machine achieved better classification performance than Naïve Bayes for HDI classification in Indonesia, although K-Nearest Neighbor obtained the highest overall accuracy in their comparative evaluation. Similarly, (Dewi et al., 2025) concluded that SVM was among the most reliable algorithms for classifying Indonesian districts and cities based on HDI indicators because of its ability to model complex decision boundaries. Furthermore.

From a practical perspective, the findings suggest that Support Vector Machine can serve as an effective decision-support tool for regional socioeconomic analysis and HDI classification. More accurate classification of provincial development status may assist policymakers in identifying regional disparities and prioritizing development programs based on reliable analytical evidence. Nevertheless, Naïve Bayes remains a valuable alternative for applications requiring faster computation and lower computational resources, particularly when dataset size is relatively limited.

Despite the encouraging results, this study has several limitations. The analysis was conducted using only the official Human Development Index dataset for 38 Indonesian provinces in 2024, which limits the size of the held-out test set (n = 8) and, despite the supporting 5-fold cross-validation, still constrains statistical power. In addition, the SVM model was evaluated using only a linear kernel; nonlinear kernels such as RBF and Polynomial (summarized in Table 2) were not tested, so it remains unclear whether a nonlinear decision boundary would further improve performance on this dataset. No hyperparameter optimization was performed for either algorithm. Consequently, the findings should be interpreted within the scope of the experimental setting used in this study. Future research is recommended to utilize multi-year datasets, incorporate additional socioeconomic indicators, evaluate more advanced machine learning algorithms such as Random Forest, Extreme Gradient Boosting, and Artificial Neural Networks, systematically compare SVM kernel functions, and apply hyperparameter optimization to further improve model robustness and predictive performance.

CONCLUSION

This study compared the performance of the Naïve Bayes and Support Vector Machine (SVM) algorithms for classifying the Human Development Index (HDI) in Indonesia using the official 2024 dataset published by Statistics Indonesia. Experimental results demonstrated that SVM achieved superior performance across all evaluation metrics, obtaining a held-out test accuracy of 100% (5-fold cross-

* Corresponding author

Submitted : 2026-07-08
Accepted : 2026-07-21
Published : 2026-07-23

validated mean of 86.4%**),** weighted precision, recall, and F1-score of 100%, while Naïve Bayes achieved a held-out test accuracy of 87.5% (cross-validated mean of 59.6%), weighted precision of 91%, recall of 88%, and F1-score of 86%. These findings, further supported by a statistically significant paired t-test across cross-validation folds ($p = 0.013$), indicate that SVM is more effective in classifying multidimensional socioeconomic data because it can construct an optimal decision boundary, whereas Naïve Bayes remains a computationally efficient alternative with satisfactory classification performance. Therefore, within the experimental setting of this study, SVM is considered the more suitable algorithm for HDI classification in Indonesia. Nevertheless, this study was limited to the official 2024 provincial HDI dataset and two classification algorithms. Future research should consider incorporating multi-year datasets, additional socioeconomic indicators, advanced machine learning algorithms, and hyperparameter optimization techniques to further improve classification performance and model generalization.

REFERENCES

Almeyda, H. P., Fathoni, M. Y., & Afrad, M. (2025). Comparison of Naïve Bayes Algorithm and Support Vector Machine in Sentiment Analysis on Social Media X for KIP College Program. *2025 8th International Conference on Information and Communications Technology (ICOIACT)*, 113–118. https://doi.org/10.1109/ICOIACT67584.2025.11344927

Anggara, R., Mukhti, T. O., Kurniawati, Y., & Fitria, D. (2024). Comparison of Naïve Bayes and K-Nearest Neighbors Methods in Classifying Human Development Index by Districts/City Indonesia in 2022. *UNP Journal of Statistics and Data Science*, *2*(4), 483–488. https://doi.org/https://doi.org/10.24036/ujsds/vol2-iss4/319

Arifat, M., Putri, W. A., & Mufida, A. S. (2023). Penerapan Metode Naive Bayes Classifier Untuk Klasifikasi Indeks Pembangunan Manusia Di Provinsi Jawa Timur. *Jurnal Statistika Dan Komputasi (STATKOM)*, *2*(1), 31–43. https://doi.org/https://doi.org/10.32665/statkom.v2i1.1661

Astrianka, O. O., & Efendi, A. (2024). Classification of Java Cities/Regencies Based on Human Development Index Using Discriminant Analysis and Naïve Bayes Classifier. *Asian Journal of Probability and Statistics*, *26*(1), 87–103. https://doi.org/10.9734/ajpas/2024/v26i1586

Atmojo, F. W., Nurlita, C. I., & Nurchim, N. (2024). ANALISIS PEMANFAATAN MACHINE LEARNING GUNA PREDIKSI INDEKS PEMBANGUNAN MANUSIA. *Jurnal Sistem Informasi Dan Teknik Komputer*, *9*(2), 89–96. https://doi.org/https://doi.org/10.51876/simtek.v9i2.390

Badan Pusat Statistik. (2024). *Indeks Pembangunan Manusia 2024*. https://www.bps.go.id

Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1023/A:1022627411411

Dewi, N. K. A. P. S., Wijayanto, A. W., & Nursiyono, J. A. (2025). Comparison of Machine Learning Algorithms in Classifying Districts/Cities in Indonesia According to the Human Development Index (HDI) in 2021. *Jurnal Sains, Nalar, Dan Aplikasi Teknologi Informasi*, *4*(1), 26–33. https://doi.org/https://doi.org/10.20885/snati.v4.i1.4

Fahmiyah, I., & Ningrum, R. A. (2023). Human Development Clustering in Indonesia: Using K-Means Method and Based on Human Development Index Categories. *Journal of Advanced Technology and Multidiscipline (JATM)*, *2*(1), 27–33. https://doi.org/https://doi.org/10.20473/jatm.v2i1.45070

Nurafidah, Suryowati, K., & Jatipaningrum, M. T. (2023). PERBANDINGAN METODE K-NEAREST NEIGHBOR DAN RANDOM FOREST PADA KLASIFIKASI INDEKS PEMBANGUNAN MANUSIA DI KABUPATEN/KOTA SELURUH INDONESIA. *Jurnal Statistika Industri Dan Komputasi*, *8*(1), 58–67. https://doi.org/https://doi.org/10.34151/statistika.v8i1.4419

Pamungkas, C. A., & Widiyanto, W. W. (2022). KLASIFIKASI INDEKS PEMBANGUNAN MANUSIA DI INDONESIA TAHUN 2022 DENGAN SUPPORT VECTOR MACHINE. *Jurnal Ilmiah Sistem Informasi Dan Ilmu Komputer*, *2*(3), 139–145. https://doi.org/https://doi.org/10.55606/juisik.v2i3

Purnami Sari Dewi, N. K. A., Wijayanto, A. W., & Nursiyono, J. A. (2025). Comparison of Machine Learning Algorithms in Classifying Districts/Cities in Indonesia According to the Human

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 3, July 2026
https://doi.org/ 10.47709/cnahpc.v8i3.9162

Submitted : 2026-07-08
Accepted : 2026-07-21
Published : 2026-07-23

Development Index (HDI) in 2021. *Jurnal Sains, Nalar, Dan Aplikasi Teknologi Informasi*, *4*(1), 26–33. https://doi.org/https://doi.org/10.20885/snati.v4.i1.4

Ramadhan, I., Budiman, & Alamsyah, N. (2025). Optimization of Human Development Index in Indonesia Using Decision Tree C4.5, Support Vector Machine Algorithm, K-Nearest Neighbors, Naïve Bayes, and Extreme Gradient Boosting. *Jurnal Informatika*, *12*(1), 15–26. https://doi.org/https://doi.org/10.31294/informatika.v12i1.12253

Syafitri, F., Goejantoro, R., & Wasono. (2021). Regresi Logistik dengan Metode Bayes untuk Pemodelan Indeks Pembangunan Manusia Kabupaten/Kota di Pulau Kalimantan. *Jurnal Eksponensial*, *12*(2), 103–110. https://doi.org/https://doi.org/10.30872/eksponensial.v12i2.802

United Nations Development Programme (UNDP). (2020). *Technical Notes: Human Development Indices and Indicators*. undp.org

Zhang, H. (2004). *The Optimality of Naive Bayes*. 562–567. https://ojs.aaai.org/index.php/FLAIRS/article/view/5969

Zuhdi, S., Fatatik, I. N., Prihasno, I. N. F. H., & Rozaq, H. A. A. (2025). Geographically Weighted Random Forests for Human Development Index of Central Java Prediction. *Jurnal Teknik Informatika (JUTIF)*, *6*(4), 2756–2768. https://doi.org/https://doi.org/10.52436/1.jutif.2025.6.4.5204

* Corresponding author