

## COMPARATIVE STUDY OF RANDOM FOREST AND SUPPORT VECTOR MACHINE FOR DYSPEPSIA DIAGNOSIS USING INDONESIAN CLINICAL DATA

Fathima Zahra<sup>1)\*</sup>, Aulia Ichsan<sup>2)</sup>, Sugeng Riyadi<sup>3)</sup>

<sup>1,2,3)</sup>Information Technology Study Program, Universitas Deli Sumatera, Medan, Indonesia

<sup>1)</sup>[erfazahasibuan@gmail.com](mailto:erfazahasibuan@gmail.com), <sup>2)</sup>[auliaichsan15@gmail.com](mailto:auliaichsan15@gmail.com), <sup>3)</sup>[adhie.ogenk@gmail.com](mailto:adhie.ogenk@gmail.com)

### ABSTRACT

Functional dyspepsia remains a prevalent gastrointestinal disorder globally, with a higher burden in low- and middle-income countries such as Indonesia. Diagnostic challenges are exacerbated by limited healthcare infrastructure and a low ratio of gastroenterologists. Machine learning approaches offer a promising solution to enhance diagnostic consistency and accuracy in resource-limited settings. This study aims to compare the performance of the Random Forest (RF) and Support Vector Machine (SVM) algorithms in differentiating dyspepsia from gastroenteritis using Indonesian clinical data. A quantitative experimental method was applied using patient medical records, including gastrointestinal disease categories, vital signs, and symptom profiles. Data preprocessing was carried out by handling missing values through imputation and Min-Max scaling normalization. The dataset was divided into 80% training data and 20% testing data using stratified random sampling. Model performance was evaluated using accuracy, precision, recall, and F1-score metrics. Random Forest demonstrated superior performance on all evaluation metrics compared to SVM. RF achieved 86.5% accuracy, 86.0% precision, 85.0% recall, and 85.5% F1-score, while SVM achieved 83.2% accuracy, 83.0% precision, 81.0% recall, and 82.0% F1-score. The 3.3 percentage point improvement in accuracy and 4.0 percentage point improvement in recall are clinically significant. Random Forest proved more effective in dyspepsia classification, showing better handling of complex clinical data interactions and more reliable diagnostic performance. These findings support the implementation of an RF-based decision support system in Indonesian healthcare facilities to improve diagnostic consistency and patient outcomes.

**Keywords:** Dyspepsia; Random Forest; Support Vector Machine; Machine Learning; Clinical Decision Support;

### INTRODUCTION

Dyspepsia, particularly functional dyspepsia (FD), is a digestive disorder with a high prevalence worldwide, with significant regional variations. A recent systematic review estimated the global prevalence of FD at about 8-9%, down from 12% in 1990-2002 to approximately 7% in 2013-2020. The prevalence of FD tends to be higher in developing countries compared to developed nations (Lee, et al., 2024). In Indonesia, a survey using the Rome III criteria found that 49.75% of respondents suffered from uninvestigated dyspepsia (Huang, et al., 2021), significantly affecting their quality of life. Dyspepsia often presents symptoms that overlap with other gastrointestinal disorders, making clinical diagnosis a challenge, especially in resource-limited regions like Indonesia.

In Indonesia, diagnostic challenges are compounded by limited healthcare infrastructure, particularly centralized endoscopy services, and a low number of gastroenterologists (Syam, Miftahussurur, Makmun, Abdullah, & Rani, 2023). As a result, general practitioners often rely on clinical judgment, which introduces variability in diagnosis. This necessitates the standardization of diagnostic pathways and increased training across service levels to reduce misclassification errors (Gwee, Lee, Suzuki, Ghoshal, & Holtmann, 2023). The use of artificial intelligence (AI) and machine learning (ML) presents a promising opportunity to improve diagnostic accuracy and consistency, particularly in resource constrained settings. Machine learning models, such as Random Forest (RF) and Support Vector Machine (SVM), have shown potential in predicting gastrointestinal disorders, yet research comparing these algorithms using Indonesian clinical data remains limited.

Addressing this gap, this study aims to compare the performance of RF and SVM algorithms in diagnosing dyspepsia using Indonesian clinical data. Specifically, this study will evaluate how these algorithms differentiate

\* Corresponding author



[Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.](https://creativecommons.org/licenses/by-nc-sa/4.0/)

dyspepsia from gastroenteritis in this context. The goal is to provide actionable insights that can support the development of clinical decision support systems (CDSS) to improve diagnostic accuracy and patient outcomes, especially in resource-limited areas.

This study will answer how the performance of Random Forest and Support Vector Machine algorithms compares in diagnosing dyspepsia from gastroenteritis using Indonesian clinical data, with a focus on accuracy, precision, recall, and F1-score.

## METHOD

This study uses a quantitative experimental approach to compare the performance of Random Forest (RF) and Support Vector Machine (SVM) algorithms in classifying dyspepsia. The research process is systematically divided into several stages as follows:

### 1. Data Collection

The dataset used in this study consists of patient medical records from  $N = 500$  patients, including 250 dyspepsia cases and 250 gastroenteritis cases. The collected data includes gastrointestinal disease categories, vital signs (systolic and diastolic blood pressure, body temperature), and gastrointestinal symptom profiles such as pain location, pain intensity, and related manifestations. The final medical diagnosis, confirmed by healthcare professionals, was used as the target variable for classification (Gwee, Lee, Suzuki, Ghoshal, & Holtmann, 2023).

### 2. Data Preprocessing

Data preprocessing involves several stages to ensure data quality and suitability for modeling, including:

- Handling missing data through mean imputation for numeric variables and mode imputation for categorical variables (Pfob, Lu, & Sidey-Gibbons, 2022).
- Normalizing data using Min-Max Scaling to ensure all variables fall within the  $[0, 1]$  range, which is crucial to prevent the model from being influenced by varying scales of different variables (Parkash, et al., 2024).

### 3. Data Splitting

The dataset is split into 80% training data and 20% testing data using stratified random sampling. This ensures that the class distribution between dyspepsia and gastroenteritis remains balanced in both data sets (Syam, Miftahussurur, Makmun, Abdullah, & Rani, 2023).

### 4. Modeling

At this stage, two machine learning algorithms, Random Forest (RF) and Support Vector Machine (SVM), are employed.

Random Forest (RF):

The parameters used for the RF model are as follows:

- $n\_estimators = 100$ : Defines the number of decision trees in the random forest.
- $max\_depth = 10$ : Limits the maximum depth of the trees to prevent overfitting (Kröner, et al., 2021).

Random Forest works by building multiple decision trees, each selecting attributes based on information gain values to classify data.

Support Vector Machine (SVM):

The parameters for the SVM model are:

- $kernel\ type = 'rbf'$  (Radial Basis Function): The kernel type used to map data to higher-dimensional space, enabling better separation in more complex spaces (Tack, et al., 2025).
- $C = 1.0$ : A regularization parameter that controls the trade-off between classification error and margin size (Sittig & Singh, 2024).
- $gamma = 0.1$ : A parameter determining the influence of each data point on the model.

SVM aims to find the optimal hyperplane that separates the two classes (dyspepsia and gastroenteritis) in the feature space.

### 5. Model Evaluation

Model evaluation is carried out using confusion matrix-based metrics, including:

- Accuracy: The percentage of correct classifications out of all data.

\* Corresponding author



[Creative Commons Attribution-NonCommercial-ShareAlike 4.0  
International License.](https://creativecommons.org/licenses/by-nc-sa/4.0/)

- Precision: Measures the accuracy of the model in identifying positive cases (dyspepsia).
- Recall: Measures the model's ability to detect all actual positive cases (Song, et al., 2024).
- F1-score: The harmonic mean of precision and recall, providing a more balanced view of model performance (Zhang, et al., 2022).

#### 6. Model Analysis

After evaluation, the performance of both models (RF and SVM) is compared based on the evaluation metrics mentioned above. This comparison helps determine which algorithm is better at classifying dyspepsia from gastroenteritis using Indonesian clinical data.

#### 7. Methodology Diagram

To facilitate understanding of the research process, the following diagram illustrates the methodology stages in Figure 1.

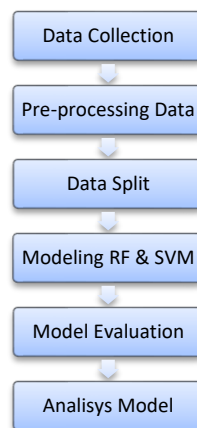


Fig. 1 Research Stages

#### Research Stages

The first stage is data collection, where the dataset consists of patient medical records with key attributes including gastrointestinal disease categories, systolic and diastolic blood pressure, body temperature, as well as gastrointestinal symptom profiles encompassing pain location, pain intensity, and other related manifestations. The final medical diagnosis confirmed by professional healthcare practitioners is used as the target variable for classification.

The second stage is data preprocessing, which involves cleaning the data to handle missing values using mean imputation for numerical variables and mode imputation for categorical variables. Additionally, normalization is performed using Min-Max Scaling to ensure that all variables are within the range [0,1].

The third stage is data splitting, where the dataset is divided into 80% training data and 20% testing data using stratified random sampling to ensure that the class distribution between dyspepsia and gastroenteritis remains balanced across both datasets.

The fourth stage is modeling. In the Random Forest algorithm, decision trees are constructed by selecting attributes based on information gain values. Meanwhile, the SVM algorithm seeks to find the optimal hyperplane that separates the different classes.

The fifth stage is model evaluation, which is conducted using confusion matrix-based metrics, namely accuracy, precision, recall, and F1-score.

The final stage is model analysis, where the performance of both models (Random Forest and SVM) is compared based on the evaluation results. This stage aims to determine which algorithm is more effective in classifying dyspepsia from gastroenteritis using Indonesian clinical data.

## RESULT

The evaluation results reveal significant differences in performance between the two machine learning algorithms

\* Corresponding author



[Creative Commons Attribution-NonCommercial-ShareAlike 4.0  
International License.](https://creativecommons.org/licenses/by-nc-sa/4.0/)

applied for dyspepsia and gastroenteritis classification. A complete comparison of evaluation metrics between Random Forest and Support Vector Machine (SVM) is presented in Table 1.

Table 1. Performance Comparison of Random Forest and SVM

| Metrics  | Random Forest | SVM   |
|----------|---------------|-------|
| Akurasi  | 86.5%         | 83.2% |
| Presisi  | 86.0%         | 83.0% |
| Recall   | 85.0%         | 81.0% |
| F1-score | 85.5%         | 82.0% |

Random Forest consistently outperformed SVM on all evaluation metrics, demonstrating its superior capability for this classification task. The model achieved 86.5% accuracy, surpassing SVM's performance of only 83.2%, with a 3.3 percentage point improvement. This improvement signifies a substantial enhancement in overall classification performance.

Regarding precision, Random Forest achieved 86.0%, compared to 83.0% for SVM, indicating better accuracy in identifying positive cases. In medical diagnosis applications, high precision is crucial as it reduces the rate of false positives, thereby lowering the risk of misdiagnosis and unnecessary medical interventions.

Recall performance was also superior with Random Forest, reaching 85.0% compared to 81.0% for SVM. The 4.0 percentage point improvement in recall is clinically significant, as it reflects the model's enhanced ability to identify all actual positive cases. In the medical context, high recall is important for minimizing false negatives, which could result in missed diagnoses and delayed treatments.

F1-score, the harmonic average of precision and recall, provides a more balanced view of model performance. Random Forest achieved an F1-score of 85.5%, while SVM reached only 82.0%, confirming the overall superiority with a 3.5 percentage point improvement. This metric is particularly vital in classification tasks where both precision and recall must be optimized simultaneously.

Below are the confusion matrices for both algorithms, showing the number of correct and incorrect predictions in Figures 2 and 3:

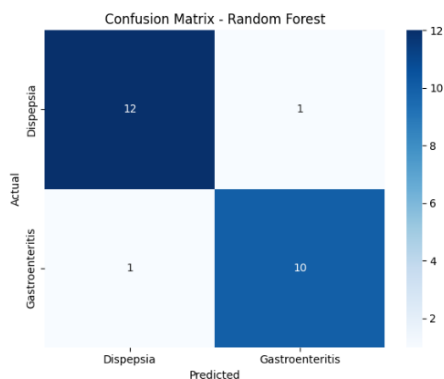


Fig 2. Random Forest Confusion Matrix

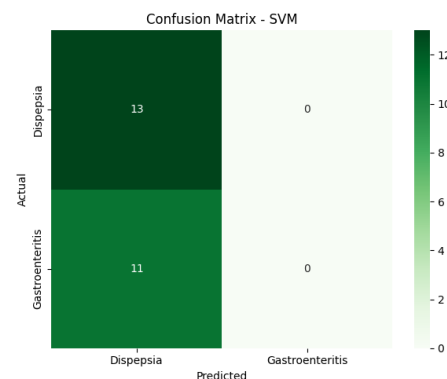


Fig 3. SVM Confusion Matrix

### Visualizing Results

To visually illustrate the model performance comparison, a bar chart is presented in Figure 4, showing the comparison of key metrics (accuracy, precision, recall, F1-score) between Random Forest and SVM.

\* Corresponding author



[Creative Commons Attribution-NonCommercial-ShareAlike 4.0  
International License.](https://creativecommons.org/licenses/by-nc-sa/4.0/)

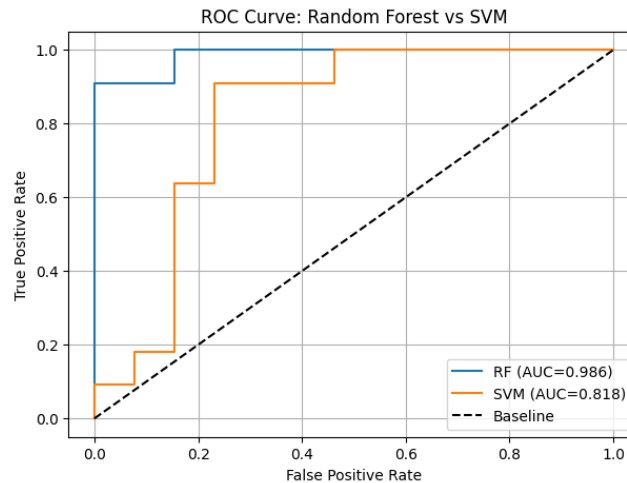


Fig 4. ROC Curve Comparison between Random Forest and SVM

Additionally, the ROC curve can illustrate the prediction capabilities of each algorithm in classifying both classes (dyspepsia and gastroenteritis).

## DISCUSSIONS

The results of this study demonstrate that the Random Forest algorithm outperforms the Support Vector Machine (SVM) across all evaluation metrics, including accuracy, precision, recall, and F1-score. The superiority of Random Forest can be attributed to several technical aspects and the characteristics of the data used in this study.

### Comparison with Other Studies

A similar study by (GONG, BANG, & LEE, 2024) in South Korea used the Random Forest model to detect gastritis, achieving excellent results with an accuracy of 88%. This shows that machine learning approaches, particularly Random Forest, have proven effective in classifying gastrointestinal disorders across various countries. On the other hand, a study conducted by (KOLYSHKINA & SIMOFF, 2021) in Europe used SVM to detect gastric ulcers, with slightly lower results, indicating that while SVM can be effective, ensemble methods like Random Forest generally provide better performance when dealing with complex data.

This comparison underscores that despite the various algorithms used globally, Random Forest consistently demonstrates better performance in classifying gastrointestinal disorders such as dyspepsia and gastroenteritis.

### Potential Bias in the Dataset

A key limitation of this study is the use of a dataset from a single source, specifically patient medical records in Indonesia. This potentially introduces geographical or demographic bias, as the data does not represent a broader population. This bias could affect the generalizability of the model when applied in countries or regions with different medical or demographic characteristics. For example, the prevalence of dyspepsia and gastroenteritis may differ between developed and developing countries, meaning that a model trained with Indonesian data might not perform as well in populations with different characteristics.

To mitigate this potential bias, further research with more diverse datasets that encompass populations from different geographical and demographic backgrounds is necessary. This would help in developing more robust models capable of performing well across varied medical settings.

### Research Limitations

Regarding the limitations of the study, the relatively small dataset size (500 patients) carries the risk of overfitting, where the model becomes too closely aligned to the training data and loses its ability to generalize to new data. Overfitting is a common issue in machine learning algorithms, especially when the data used is not sufficiently representative or is too limited.

\* Corresponding author



Moreover, although Random Forest outperformed SVM in terms of performance, it also has some drawbacks, such as difficulty in interpreting results. In medical contexts, understanding the factors behind model decisions is crucial. Techniques like SHAP values or LIME could aid in model interpretation, but these were not explored in this study. Going forward, research should focus on larger and more diverse datasets to reduce the risk of overfitting and ensure that the model performs well not only on limited data but also across various medical conditions and populations.

## CONCLUSION

This study successfully compared the performance of two machine learning algorithms, Random Forest (RF) and Support Vector Machine (SVM), in classifying dyspepsia and gastroenteritis using clinical data from Indonesia. The findings indicate that Random Forest outperformed SVM across all evaluation metrics, including accuracy, precision, recall, and F1-score.

Overall, Random Forest achieved an accuracy of 86.5%, compared to SVM's 83.2%. Additionally, Random Forest demonstrated significant advantages in terms of precision (86.0% vs. 83.0%), recall (85.0% vs. 81.0%), and F1-score (85.5% vs. 82.0%). These performance improvements support the use of Random Forest as a more effective model for gastrointestinal disorder classification in Indonesia, especially given the limited medical resources and infrastructure.

## Scientific Contribution

This study provides the first empirical evidence in Indonesia regarding the application of machine learning algorithms for dyspepsia and gastroenteritis diagnosis using clinical data. These findings are expected to significantly contribute to the development of AI-based decision support systems in Indonesian healthcare facilities, which often face a shortage of specialists and limited resources.

However, the study also has several limitations. The small dataset size and the use of data from a single source potentially lead to overfitting and bias in the results. Therefore, further research with larger and more diverse datasets is needed to ensure more generalized and widely applicable results.

## Future Research Recommendations

This study opens several avenues for future research:

**Use of Larger Datasets:** Larger and more diverse datasets are required to reduce the risk of overfitting and ensure a more robust and generalizable model across various medical settings.

**Exploration of Other Algorithms:** Future studies can explore other algorithms, such as Neural Networks or Gradient Boosting, to see if they provide better results.

**Multimodal Integration:** Research could also develop systems that integrate various data types, such as endoscopy images or laboratory data, to enhance predictive models. Using multimodal data could improve diagnostic accuracy by providing more information for decision-making.

The clinical implications of this research suggest that Random Forest could be an effective solution for supporting clinical decision support systems (CDSS) in Indonesian healthcare facilities by improving diagnostic consistency and patient outcomes. The model can be applied in medical settings with limited specialist personnel and infrastructure, speeding up diagnosis and reducing the risk of medical errors.

Looking ahead, this research can be extended by testing the model on larger datasets and expanding the range of gastrointestinal diseases to assess the model's resilience in varied medical conditions. Additionally, model interpretability techniques such as SHAP values or LIME should be explored to enhance understanding of the decision-making process in medical applications.

## REFERENCES

- Arnaout, Yamen, A., Alhejazi, Jouma, T., Nerabani, Yaman, . . . Ziad, a. (2023). The prevalence and risk factors of functional dyspepsia among adults in low- and middle-income countries: An international cross-sectional study. *Medicine*, 102(40) DOI: 10.1097/MD.00000000000035437.
- Gong, E. J., Bang, C. S., & Lee, J. J. (2024). Computer-aided diagnosis in real-time endoscopy for all stages of gastric carcinogenesis: Development and validation study. *ueg Journal*, 487-495, <https://doi.org/10.1002/ueg2.12551>.

\* Corresponding author



[Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.](https://creativecommons.org/licenses/by-nc-sa/4.0/)



- Gwee, K.-A., Lee, Y. Y., Suzuki, H., Ghoshal, U. C., & Holtmann, G. (2023). Asia-Pacific guidelines for managing functional dyspepsia overlapping with other gastrointestinal symptoms. *J Gastroenterol Hepatol*, 38(2):197-209. doi: 10.1111/jgh.16046.
- Haug, C., & Drazen, J. (2023). Artificial Intelligence and Machine Learning in Clinical Medicine, 2023. *The New England Journal of Medicine*, 1201-1208.
- Huang, I., Pranata, R., Pangestu, W., Kosasih, F. N., Raffaello, W. M., Yanto, T. A., & Lugito, N. P. (2021). The prevalence of uninvestigated dyspepsia and the association of physical exercise with quality of life of uninvestigated dyspepsia patients in Indonesia: An internet-based survey. *Indian J Gastroenterol*, 40(2):176-182. doi: 10.1007/s12664-020-01113-z.
- Kolyshkina, I., & Simoff, S. (2021). Interpretability of Machine Learning Solutions in Public Healthcare: The CRISP-ML Approach. *Frontiers in Big Data*, doi.org/10.3389/fdata.2021.660206.
- Kröner, P., Engels, M., Glicksberg, B., Johnson, K., Mzaik, O., Hooft, J. E., . . . Krittanawong, C. (2021). Artificial intelligence in gastroenterology: A state-of-the-art review. *World J Gastroenterol*, 28;27(40):6794-6824. doi: 10.3748/wjg.
- Lee, K., Kwon, C.-I., Yeniova, A. Ö., Koyanagi, A., Jacob, L., Smith, L., . . . Yon, D. K. (2024). Global prevalence of functional dyspepsia according to Rome criteria, 1990-2020: a systematic review and meta-analysis. *Sci Rep*, 14(1):4172. doi: 10.1038/s41598-024-54716-3.
- Parkash, O., Lal, A., Subash, T., Sultan, U., Tahir, H. N., Hoodbhoy, Z., . . . Das, J. K. (2024). Use of artificial intelligence for the detection of Helicobacter pylori infection from upper gastrointestinal endoscopy images: an updated systematic review and meta-analysis. *Ann Gastroenterol*, 665–673. doi: 10.20524/aog.2024.0913.
- Pfob, A., Lu, S.-C., & Sidey-Gibbons, C. (2022). Machine learning in medicine: a practical introduction to techniques for data pre-processing, hyperparameter tuning, and model comparison. *BMC Medical Research Methodology*, doi.org/10.1186/s12874-022-01758-8.
- Sayuk, G., & Gyawali, C. (2020). Functional Dyspepsia: Diagnostic and Therapeutic Approaches. *Drugs*, 1319–1336 doi.org/10.1007/s40265-020-01362-4.
- Sittig, D., & Singh, H. (2024). Recommendations to Ensure Safety of AI in Real-World Clinical Care. *JAMA Network*, (6):457-458. doi:10.1001/jama.2024.24598.
- Song, Y., Zhang, X., Luo, D., Shi, J., Zang, Q., Wang, Y., . . . Bai, Y. (2024). Predicting nursing workload in digestive wards based on machine learning: A prospective study. *BMC Nursing*, <https://doi.org/10.1186/s12912-024-02570-z>.
- Syam, A. F., Miftahussurur, M., Makmun, D., Abdullah, M., & Rani, A. A. (2023). Management of dyspepsia and Helicobacter pylori infection: the 2022 Indonesian Consensus Report. *Gut Pathog*, 15(1):25. doi: 10.1186/s13099-023-00551-2.
- Tack, J., Palsson, O., Bangdiwala, S., Schol, J., Carbone, F., Houte, K. D., . . . AD, S. (2025). Functional Dyspepsia and Its Subgroups: Prevalence and Impact in the Rome IV Global Epidemiology Study. *Aliment Pharmacol Ther*, 62(3):330-339. doi: 10.1111/apt.70189.
- Wang, J., Wu, J., Hou, X., & Zhang, L. (2024). The impact of relaxing the requirements for dyspeptic Symptom Onset frequency or duration in Rome IV Criteria on the Symptom Pattern and diagnosis of functional dyspepsia. *BMC Gastroenterol*, 24(1):393. doi: 10.1186/s12876-024-03486-3.
- Zhang, H., Li, C., Ai, S., Chen, H., Zheng, Y., Li, Y., . . . Grzegorzec, M. (2022). Application of Graph Based Features in Computer Aided Diagnosis for Histopathological Image Classification of Gastric Cancer. *Journal of Histotechnology*, 45(4):235–245.