Jurma : Jurnal Program Mahasiswa Kreatif
Universitas Ibn Khaldun Bogor, Indonesia
Vol. 9, No. 1, Juni 2025, pp. 37~44
ISSN: 2615-8019, DOI: 10.32832/pkm
http://pkm.uika-bogor.ac.id/index.php/pkm-p

Detection of Thyroid Disease Using a Comparison of Merchine Learning Methods: Support Vector Machine, Decision Tree and Random Forest

Adzkia Nur Nasution *, Ardilla Syahfitri Lubis, Fanny Ramadhani, Sri Dewi
Universitas Islam Negeri Sumatera Utara
Jl. William Iskandar Ps. V, Medan Estate, Kec. Percut Sei Tuan, Kabupaten Deli Serdang, Sumatera Utara 20371, Indonesia

Article Info

Article history:

Received October 18, 2024
Revised October 30, 2024
Accepted November 30, 2024

Keywords:

Decision Tree
Random Forest
Support Vector Machine (SVM)
Thyroid
Machine Learning

ABSTRACT

Thyroid disease occurs when the thyroid gland undergoes changes in function and shape, which can lead to disorders such as hypothyroidism or hyperthyroidism, depending on the amount of hormones produced. This disorder can be treated by a specialist, either through medication or surgery. Early detection is essential to prevent further complications, but conventional diagnostic methods are often time-consuming and costly. The purpose of this study is to develop a machine learning-based prediction model to improve the speed and accuracy in detecting thyroid disease. By using a comparison of machine learning algorithms, it is hoped that the diagnosis process can be carried out more efficiently than conventional methods. This study uses datasets from the UCI Machine Learning Repository, in this study the methods used are decision tree algorithms, random forests, and support vector machines (SVM). The data is divided into two sets, namely training data and test data, to evaluate the performance of each model. The results show that SVM has the highest accuracy rate with an F1-score of 0.9607, followed by Random Forest (0.9601), and Decision Tree which has the lowest performance (0.9177). This study proves that the use of machine learning can improve the speed and accuracy of diagnosing thyroid disease compared to conventional methods.

Corresponding Author:

Adzkia Nur Nasution
Universitas Islam Negeri Sumatera Utara
Email: adzkiaaanur41@gmail.com

INTRODUCTION

The thyroid gland is one of the largest endocrine glands in the human body and is located just below the larynx and in front of the trachea. These glands produce two main hormones that regulate the body's metabolism: thyroxine (T4) and triiodothyronine (T3). The production of thyroid hormones is regulated by thyroid-stimulating hormone (TSH). Thyroid disease is a disease that causes changes in the function and shape of the thyroid gland. Most thyroid diseases are enlarged thyroid glands, which can be classified into two forms, namely diffuse or nodular hypertrophy (enlargement of the thyroid in the form of nodules). Thyroid abnormalities can be benign or malignant gland enlargement. In the early 1800s, an enlarged thyroid gland was considered a swelling of the airways in the neck area. In Latin, this condition is called tumid guttur, which means "swelling of the throat". Later, the term was translated into French as goitre and English also became goitre, while in America it was known as goiter. In Indonesia, doctors also use the term goiter to refer to the enlargement of the adenoids. There is also another term, namely struma, which means swelling or tumefaction (Setiawati , 2022).

In the past, scientists thought that the thyroid originated from the laryngeal gland (the base of the throat), until finally in 1656, Wharton became the first person to name it the thyroid gland. At that time, there was no definite understanding of the function of the neck lump. Some argue that the lump is an enlarged vocal cord, while others think it is a lymph node, blood vessel, or food duct. In the late 19th century, when surgery or removal of the lump in the neck was performed, patients often experienced swelling, weakness, and even death. In 1896, Baumann discovered that the thyroid contained essential substances. Then, in 1915, Kendall successfully identified the hormone thyroxine (T4) from the thyroid gland, and in 1926, Harington and Barger discovered the chemical structure of the hormone T4.

Research continues, and various thyroid disorders are discovered. In 1825, Parry, then Graves in 1835, and Von Basedow in 1840, reported the same disease, which was hyperthyroidism. The condition is characterized by symptoms such as heart palpitations, weight loss, trembling, and eyes that appear protruding. This disease is known as Parry's disease, Graves' disease, or Von Basedow's disease. In addition, thyroid inflammation (thyroiditis) was discovered by Riedel in 1896, Hashimoto in 1912, and De Quervain in 1936. Various types of thyroid cancer also began to be known in the late 19th century. In 1909, Theodor Kocher, a surgeon from Switzerland, became the first doctor to successfully perform surgery on a goiter patient. He also managed to prove the important role of thyroid hormones in the growth and function of the body's organs. For this achievement, Dr. Kocher was awarded the Nobel Prize in medicine. (Tandra, 2013).

In a study conducted by Beynon & Pinneri (2016), it was found that there were 22,368 cases of death due to thyroid cancer in the United States. This figure is lower than deaths caused by non-cancerous factors in people with thyroid cancer. However, the death rate needs to continue to be suppressed so that the number of deaths caused by thyroid disorders can be reduced. One way to reduce this number is to take advantage of technological developments. In conventional methods such as recording the prediction of recurrence of thyroid patients through clinical assessment and determination of treatment steps, there are limitations in terms of accuracy, which often leads to prediction failure. As more and more medical data becomes available, researchers now have the opportunity to utilize machine learning techniques with specific models to achieve better levels of accuracy (Joseph & Hidayatulloh , 2024).

Machine Learning can improve the decision-making process by providing automated, predictive insights. As part of artificial intelligence, this technology continues to evolve in pattern recognition and computational learning, as well as being able to predict outcomes and support automated decisions based on existing data. In the context of thyroid disease detection, Machine Learning can analyze patients' medical data quickly and accurately, assisting doctors in diagnosing thyroid disease early and with high precision (Wardhana et al., 2023)

One of the commonly used algorithms is the Support Vector Machine (SVM), which works for regression and classification, including in the Support Vector Classification and Support Vector Regression applications. SVM utilizes the hypothetical space in the form of linear functions on high-dimensional features and is trained with optimization theory-based algorithms. The algorithm operates as a linear classifier, but can be adapted to handle non-linear problems by using kernel tricks, allowing operation in higher dimensional spaces (Severian & Supatman , 2024) (Yoga et al., 2021)

Decision trees are also a popular and effective method for classification and prediction. This method turns large data sets into decision trees that represent various rules. These rules are easy to understand in natural language and can be converted into database languages such as Structured Query Language (SQL) for searching data in specific categories. In addition, decision trees are useful in data exploration and uncover hidden relationships between input variables and (Partogi et al., 2021)

The Random Forest algorithm is an ensemble method that uses multiple decision trees as the basic classifier that is built and combined. This method involves random sampling to form a prediction tree, with each tree using a random predictor. Predictions are generated by combining the results of all trees, either through majority votes for classification or calculating averages for regression. (Afdhal et al., 2022)

Therefore, this study aims to explore the use of machine learning technology in detecting thyroid disease, with the hope that this approach can help in providing faster and more accurate diagnoses, as well as make it easier for doctors to make clinical decisions.

RESEARCH METHODS

This research focuses on the Algortima aspect of machine learning by applying the Decision Tree, Random Forest and Support Vector Machine methods to predict the accuracy of thyroid disease. In this study, in the experimental process we sequentially carried out various processing steps starting from the stages of data collection, data preprocessing, feature selection, modeling, and evaluation, as seen in the figure.

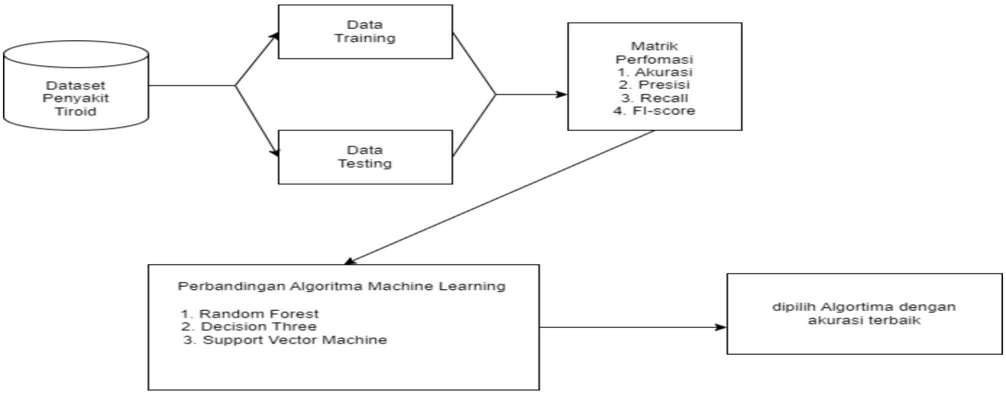


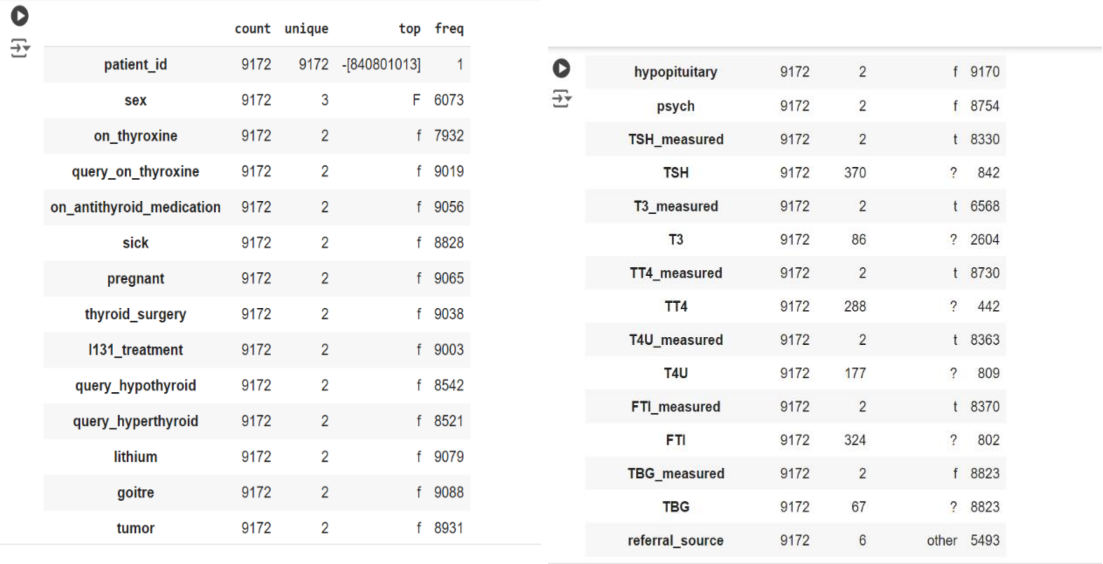
Figure 1. Research Methods

In Figure 1, the research flow begins by searching for heart disease datasets through UCI machine learning reports, followed by the stages of dividing the dataset into two parts, namely training data and data testing with a ratio of 80:20. Data training is used to train several types of machine learning algorithms, namely decision three, random forest, support vector machine while data testing is used to test the accuracy of the model. Once these algorithms are trained, their performance is tested on test data, Performance is measured using metrics such as accuracy, precision, acquisition, and F1 score. Then, based on the results of the performance test, the most accurate algorithm will be selected as the most effective model to be used to detect thyroid disease.

Each stage of processing this research is described in detail in each subchapter.

Data Collection

The dataset consists of 9172 rows and 30 columns that can be viewed from



	count	unique	top	freq
patient_id	9172	9172	-[840801013]	1
sex	9172	3	F	6073
on_thyroxine	9172	2	f	7932
query_on_thyroxine	9172	2	f	9019
on_antithyroid_medication	9172	2	f	9056
sick	9172	2	f	8828
pregnant	9172	2	f	9065
thyroid_surgery	9172	2	f	9038
I131_treatment	9172	2	f	9003
query_hypothyroid	9172	2	f	8542
query_hyperthyroid	9172	2	f	8521
lithium	9172	2	f	9079
goitre	9172	2	f	9088
tumor	9172	2	f	8931
hypopituitary	9172	2	f	9170
psych	9172	2	f	8754
TSH_measured	9172	2	t	8330
TSH	9172	370	?	842
T3_measured	9172	2	t	6568
T3	9172	86	?	2604
TT4_measured	9172	2	t	8730
TT4	9172	288	?	442
T4U_measured	9172	2	t	8363
T4U	9172	177	?	809
FTI_measured	9172	2	t	8370
FTI	9172	324	?	802
TBG_measured	9172	2	f	8823
TBG	9172	67	?	8823
referral_source	9172	6	other	5493

Figure 2. Data Collection

Pre Processing Data

The pre-processing stage of data begins with preparing raw data, then continues with handling, feature cleaning, outliers, and relevant missing values before entering the next process.

- Missing Value
  Based on the data in the Hepatitis dataset, there are 4 records that contain missing values. To overcome this, the author decided to delete the record that has the missing value.
- Feature Cleaning
  Feature cleaning is used to remove irrelevant columns from the dataset, the features that are removed from the dataset because they are useless are 'TSH_measured', 'T3_measured', 'TT4_measured', 'T4U_measured', 'FTI_measured', 'TBG_measured', 'referral_source', and 'patient_id'

Feature Cleaning (drop fitur yang tidak perlu)

```
[ ] data.drop(['TSH_measured', 'T3_measured', 'TT4_measured', 'T4U_measured', 'FTI_measured', 'TBG_measured', 'referral_source', 'patient_id'],axis=1 ,inplace=True) #1 kolom, 0 baris jadi
```

```
data.shape #Kolom berhasil didrop
```

```
(9172, 22)
```

Figure 3. Pre Processing Data

Performance Matrix Calculation

Performance matrices are used to evaluate how well a machine learning model performs in classification tasks by comparing the model's predictions to the actual value. The author uses the 'classification_report' function to calculate metrics such as precision, recall, F1-score, and support. Precision measures the accuracy of predictions of positive classes, recall measures the model's ability to identify all positive cases, while F1-score is the harmonic average of precision and recall. These metrics are calculated for each class and help compare model performance, such as Random Forest, Decision Tree, and SVM. Click or tap here to enter text.

Matiks Performa

```
from sklearn.metrics import classification_report
import pandas as pd

report_rf = classification_report(y_test, y_pred_rf, output_dict=True)
report_dt = classification_report(y_test, y_pred_dt, output_dict=True)
report_svm = classification_report(y_test, y_pred_svm, output_dict=True)

df_rf = pd.DataFrame(report_rf).transpose()
df_dt = pd.DataFrame(report_dt).transpose()
df_svm = pd.DataFrame(report_svm).transpose()

metrics_rf = df_rf[['precision', 'recall', 'f1-score']].iloc[:-1]
metrics_dt = df_dt[['precision', 'recall', 'f1-score']].iloc[:-1]
metrics_svm = df_svm[['precision', 'recall', 'f1-score']].iloc[:-1]

print("Random Forest Metrics:\n", metrics_rf)
print("\nDecision Tree Metrics:\n", metrics_dt)
print("\nSVM Metrics:\n", metrics_svm)
```

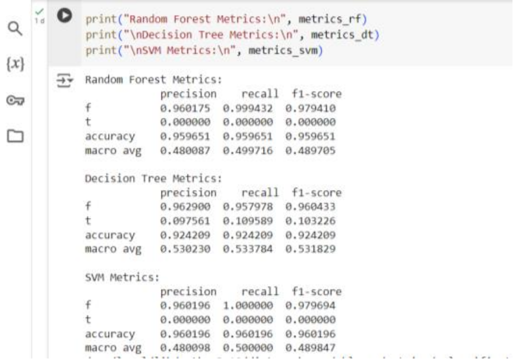
```
Random Forest Metrics:
              precision    recall  f1-score
f             0.959039  0.998294  0.978273
t             0.000000  0.000000  0.000000
accuracy      0.957470  0.957470  0.957470
macro avg     0.479519  0.499147  0.489136

Decision Tree Metrics:
              precision    recall  f1-score
```

Figure 4. Performance Matrix Calculation

ISSN: 2615-8019

RESULTS AND DISCUSSION

This study uses SVM, Random Forest and Decision Three algorithms. The data was analyzed using google collab and divided into training and testing subsets.



```
print("Random Forest Metrics:\n", metrics_rf)
print("\nDecision Tree Metrics:\n", metrics_dt)
print("\nSVM Metrics:\n", metrics_svm)
```

```
Random Forest Metrics:
            precision    recall  f1-score
f           0.960175  0.999432  0.979410
t           0.000000  0.000000  0.000000
accuracy    0.959651  0.959651  0.959651
macro avg   0.480087  0.499716  0.489705

Decision Tree Metrics:
            precision    recall  f1-score
f           0.962900  0.957978  0.960433
t           0.097561  0.109589  0.103226
accuracy    0.924209  0.924209  0.924209
macro avg   0.530230  0.533784  0.531829

SVM Metrics:
            precision    recall  f1-score
f           0.960196  1.000000  0.979694
t           0.000000  0.000000  0.000000
accuracy    0.960196  0.960196  0.960196
macro avg   0.480098  0.500000  0.489847
```

Figure 5. Results and Discussion

Here is a visual matrix of how each model performs.

SVM Matrix (Support Vector Machine)

This figure shows the performance metrics of the support vector machine (SVM) model used for classification. Precision, recall, and F1-score have values of around 0.968, 1,000, and 0.984 for false, respectively, while precision, recall, and F1-score for overall accuracy are at 0.968.

Figure 6. Matrix Support Vector Machine

Random Forest Matrix

This image shows the performance metrics of the Random Forest model used for classification. Precision, recall, and F1-score have values of around 0.968, 0.999, and 0.983 for false, respectively, while precision, recall, and F1-score for overall accuracy are at 0.967.

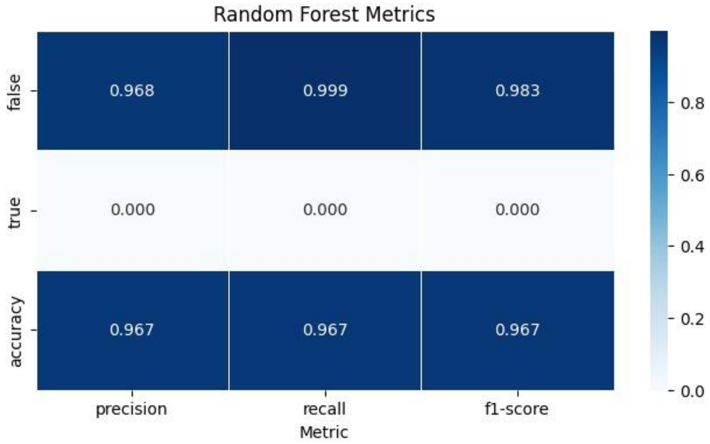


Figure 7. Random Forest Matrix

Decision Tree Matrix

This image shows the performance metrics of the Decision Tree model used for classification. Precision, recall, and F1-score have values of around 0.969, 0.963, and 0.966 for false, respectively, while precision, recall, and F1-score for overall accuracy are at 0.934.

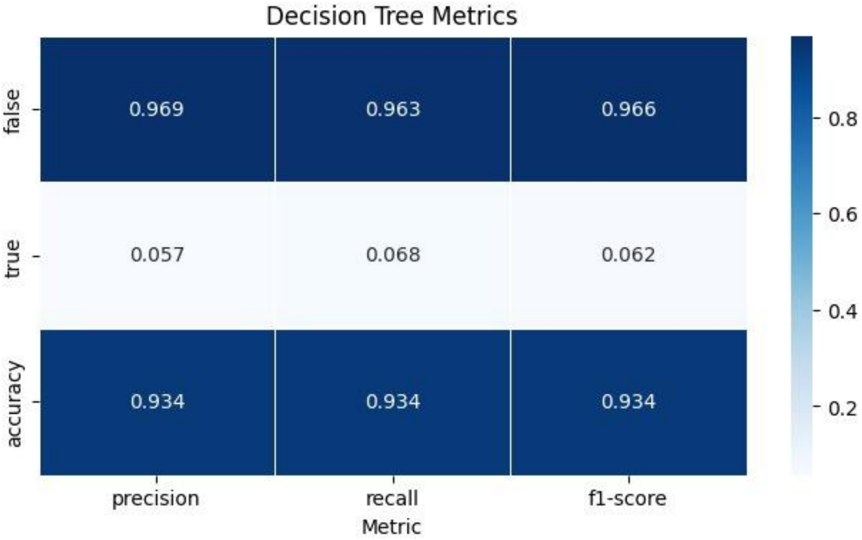


Figure 8. Decision Tree Matrix

F1-Score Comparison Chart of Multiple Models

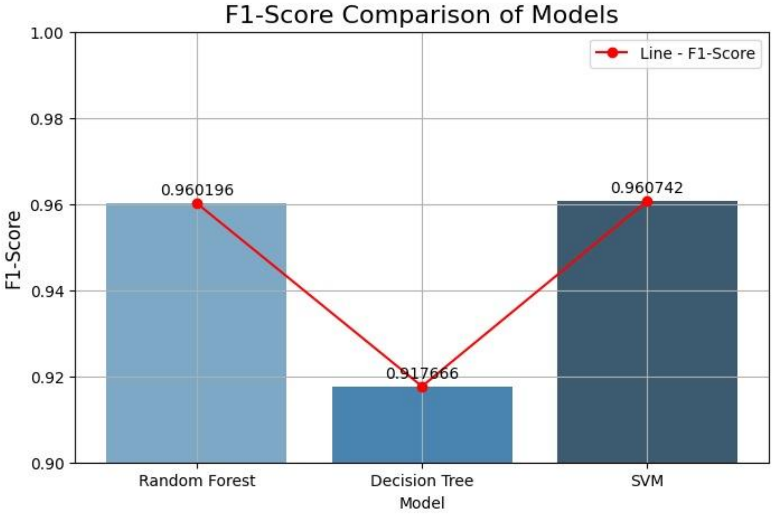


Figure 9. Comparison Diagram of F1-Score of multiple models

From the figure above, the F1-score value of each model shows that the SVM model has the highest accuracy for detecting thyroid disease with an F1-score of 0.960742. The F1-score of Random Forest stands at 0.960196, which indicates that both models have excellent predictive abilities. In contrast, Decision Tree shows the lowest performance with an F1-score of 0.917666 which indicates its performance is lower than the other two models in the data classification.

CONCLUSION

The use of machine learning models such as SVM, Random Forest, and Decision Three can significantly improve the speed and accuracy of thyroid disease diagnosis compared to conventional methods. Based on the evaluation results, SVM is the best performing model with an accuracy of 0.960742, followed by Random Forest with an accuracy of 0.960196 and Decision Tree as the model with the lowest performance of 0.917666. The SVM model gives the highest F1-score. This means that this model can better handle data classification with precision and balanced acquisition. Random Forest also performs well, albeit slightly lower than decision three and decision three has the lowest performance. This study proves that the use of machine learning can be an effective solution to detect diseases, enable faster and more accurate diagnosis, and provide significant benefits for clinical decision-making by doctors.

REFERENCES

[1] Afdhal, I., Kurniawan, R., Iskandar, I., Salambue, R., Budianita, E., Syafria, F. (2022). Application of the Random Forest algorithm to sentiment analysis of comments on YouTube about Islamophobia. National Journal of Computing and Information Technology, 5(1).

[2] Amelia, Y. (2023). COMPARISON OF MACHINE LEARNING METHODS FOR DETECTING HEART DISEASE. Indonesia Journal Information System (Vol. 6, Issue 2). http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

[3] Bria, N. A. R., & Witanti, A. (2023). ANALYSIS OF INDONESIAN PUBLIC SENTIMENT USING THE SUPPORT VECTOR MACHINE ALGORITHM REGARDING THE 2024 PRESIDENTIAL ELECTION. Informatics Engineering Student Journal (Vol. 7, Issue 6).

[4] Dwi Putra, A., Nurani, D., Dewi, M. M., & Nur Rahmi, A. (2024). Supervised Machine Learning Model for Hepatitis Disease Prediction. Indonesian Journal of Computer Science. https://archive.ics.uci.edu/ml/datasets/Hepatitis

[5] Pamuji, F. Y., & Ramdhan, V. P. (2021). Journal of Informatics Technology and Management: Comparing Random Forest and Decision Tree Algorithms to Predict the Success of Immunotherapy Info ABSTRACT Article. Journal of Informatics Technology and Management, 7, 46–50. http://http://jurnal.unmer.ac.id/index.php/jtmi

[6] Partogi, Y., Pasaribu, A., & Sutrisno. (2021). Designing the DecisionTree Method for the STMIK Kuwera Library System. Journal of Information Systems and Technology (SINTEK), 1(2), 20. https://sintek.stmikku.ac.id/index.php/SINTEK

[7] Setiawati, R. (2022). POCKET BOOK FOR THYROID DISEASE MANAGEMENT AND EARLY DETECTION. Airlangga University Press.

[8] Severianus, Y. K., & Supatman. (2024). SENTIMENT ANALYSIS OF PUBLIC OPINION RELATED TO WEATHER CHANGES IN INDONESIA USING THE SUPPORT VECTOR MACHINE ALGORITHM. In Informatics Engineering Student Journal (Vol. 8, Issue 2).

[9] Tandra, H. (2013). Prevent and overcome thyroid disease. Gramedia Pustaka Utama.

[10] Wardhana, R. G., Wang, G., & Sibuea, F. (2023). THE APPLICATION OF MACHINE LEARNING IN PREDICTING THE RATE OF DISEASE CASES IN INDONESIA. In Journal of Information System Management (JOISM) e-ISSN (Vol. 5, Issue 1).

[11] Yoga, V. W., Erfina, A., & Warman, C. (2021). Analysis of Sentiment Around the ITE Law using the Support Vector Machine Algorithm. Journal of Computer Science, 17. https://databoks.katadata.co.id/

[12] Yusuf, L., & Hidayatulloh, T. (2024). Implementation of Artificial Neural Network Algorithm with ReLu Activation: Triroid Classification. JOURNAL SWABUMI, 12(1), 2024.