

Comparative Analysis of SVM and Perceptron Algorithms in Classification of Work Programs

Jaka Tirta Samudra , Rika Rosnelly , Zakarias Situmorang
Universitas Potensi Utama, Medan, Indonesia

Article Info

Article history:

Received November 01, 2022
Revised February 01, 2023
Accepted March 14, 2023

Keywords:

Classification
Government
SVM
Perceptron
Sigmoid
Work Program

ABSTRACT

Government agencies are required to mobilize every aspect of publication, which is carried out every year and must be accounted for and also carried out for each device that receives it, such as assisted villages by utilizing available APBD funds in maximizing work programs designed so that they can be implemented optimally and effectively. Getting the best from all aspects of the work program implementation, of course, there are important points in designing an annual work program without exception. Data mining itself can help the department of population, family planning, women's empowerment, and child protection in analyzing each work program design from before it is implemented onwards to look at various aspects of past data whose grouping is in the form of classification. This study aimed to build a classification model by adding a sigmoid activation function that used SVM and perceptron to obtain a comparison value for the accuracy of the algorithm used to obtain the best working program design. The classification results were used to get the best value for classifying the best P2KBP3A work program dataset, where it can be seen that the average accuracy value was 87.5%, the f1 value was 82.2%, the precision value was 80.2%, and the recall value is 87.5% so that the final result of the research results obtained a good accuracy value.

Copyright ©2022 The Authors.
This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Jaka Tirta Samudra, +62822 7474 8215,
Computer Science Study Program,
Universitas Potensi Utama, Medan, Indonesia,
Email: jakatirta135@gmail.com

How to Cite: J. Samudra, R. Rosnelly, and Z. Situmorang, "Comparative Analysis of SVM and Perceptron Algorithms in Classification Of Work Programs", *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 22, no. 2, pp. 285-298, Mar. 2023.

This is an open access article under the CC BY-SA license (<https://creativecommons.org/licenses/by-sa/4.0/>)

1. INTRODUCTION

Support vector machine, commonly abbreviated as SVM, is a method that solves problems such as predictions, whether in the form of classification or regression [1]. This algorithm has the advantage of high accuracy and does not require a lot of data samples to avoid overfitting [2]. This algorithm can also solve problems by using datasets that have a large feature space [3]. In addition to SVM, there is also another algorithm that is also good for use in prediction problems in the form of classification, namely the perceptron, a simple supervised learning neural network algorithm that can be used to recognize patterns in data [4]. As a method to enter into machine learning and focus on the type of supervised learning, it can be used to form behavioral patterns from data based on a collection of data samples that have been labeled [5]. By using the desired input and output values from a data set, supervised learning can be used to solve classification or regression problems depending on the data being processed [6]. As one part of supervised learning, the classification itself can be used to build a learning model, where the computer learns the input data and generates a classification function to categorize the data used as test material [7].

This study aims to compare the classification between the SVM algorithm and the perceptron, both of which have the addition of using the sigmoid function. The sigmoid function is a kernel function often used in non-linear classification problems using the SVM algorithm [8]. It can also be used for the perceptron algorithm with the sigmoid activation function in performing comparative analysis. In reference to previous research, the accuracy of SVM was obtained with a value of 98.28% from [9], with a value of 88.2% [10], with a value of 92.3% from [11], then reference to previous research obtained a perceptron accuracy value with a value of 81.1% from [12], with a value of 80.5% from [13], with a value of 85.3% from [14]. Data mining is one of the sciences in computer science, which can be in the form of grouping to be used in solving problems such as predictions, classification, or regression [15, 16]. Machine learning itself can also be defined as a type of application to computers and mathematical learning algorithms that are adopted by following learning concepts that come from data and also to produce predictions in the future. Machine learning itself is also associated as a branch of data mining learning that has been widely used to solve a problem with classification by utilizing a neural network model [17, 18]. Based on the number of hidden layers, the ANN model is divided into two types: single-layer ANN, which has one hidden layer, and multi-layer ANN, which has more than one hidden layer [19, 20].

Support Vector Machine (SVM) is a method that follows self-classification, which Vapnik first introduced in 1998. Based on this concept is the modeling of the method so that it can work well in defining the boundary between two classes that are different from the maximum distance from the nearest available data [21]. To be able to execute obtaining by getting the maximum limit between two different classes, it must be formed into a hyperplane or the best-dividing line on the input space or space obtained by measuring at the hyperplane margin and also to finding the maximum point. Single layer perceptron is a network that has one layer with weighted layers connected. The advantages also, among others, are that they can obtain necessary knowledge from the uncertainty in the data used to perform tests on the data, which can also be made from the results generalized to extraction [22]. The single-layer perceptron method is the most basic method for machine learning cases and the simplest at the case level. Single-layer perceptron is a feedforward type, a type of NN where neurons in one layer can only connect with neurons in different layers. Therefore, in this case, the Single Layer perceptron network method includes supervised learning because the learning method is carried out by studying examples of known input and output [23].

This study takes a case study of population control, family planning, women's empowerment, and child protection, or P2KBP3A Deli Serdang. Every year, several different work programs are arranged, such as family planning team training activities, provision of correspondence services, child rights convention training, coordination meetings, training in the use of tools, and consultations outside and within the region. Other work programs totaling up to 62 work programs and even more than 62 work programs [24] by analyzing performance programs. To get a feasible work program to be implemented and continues to exist every year, the work program that has been evaluated and, of course, makes improvements. Feasibility and performance of work programs that are ready to be implemented and accounted for to be able to provide significant results for progress [25].

Based on the background of the classification problem for the prediction above, the authors collect data directly, and also, in this case, there has been no related research, and the data used has not been processed [26, 27] in analyzing to classify the best work program from the P2KBP3A case study using the sigmoid function, which is applied to the classification problem using the SVM and perceptron algorithms. The comparison analysis used is to find the values of accuracy, precision, recall, and f-measure obtained from the evaluation of k-fold cross-validation [13, 14] to then assess which algorithm has the best performance. The novelty of this research is basically to find out how the process of the results of the classification of values that have been tested against the accuracy values obtained is to make comparisons between each work program that has been implemented in 2018, 2019, 2020, and 2021. so that each work program is feasible and implemented returns and work programs that are not feasible are not re-implemented. This study aims to build a classification model with the addition of a sigmoid activation function that uses svm and perceptron to obtain a comparison value for the accuracy of the algorithm used to obtain the best working program design.

2. RESEARCH METHOD

The classification system for the architectural level of research is designed to look for the best work programs of case studies P2KBP3A, which is where the course of this system goes according to the procedures addressed in Figure 1.

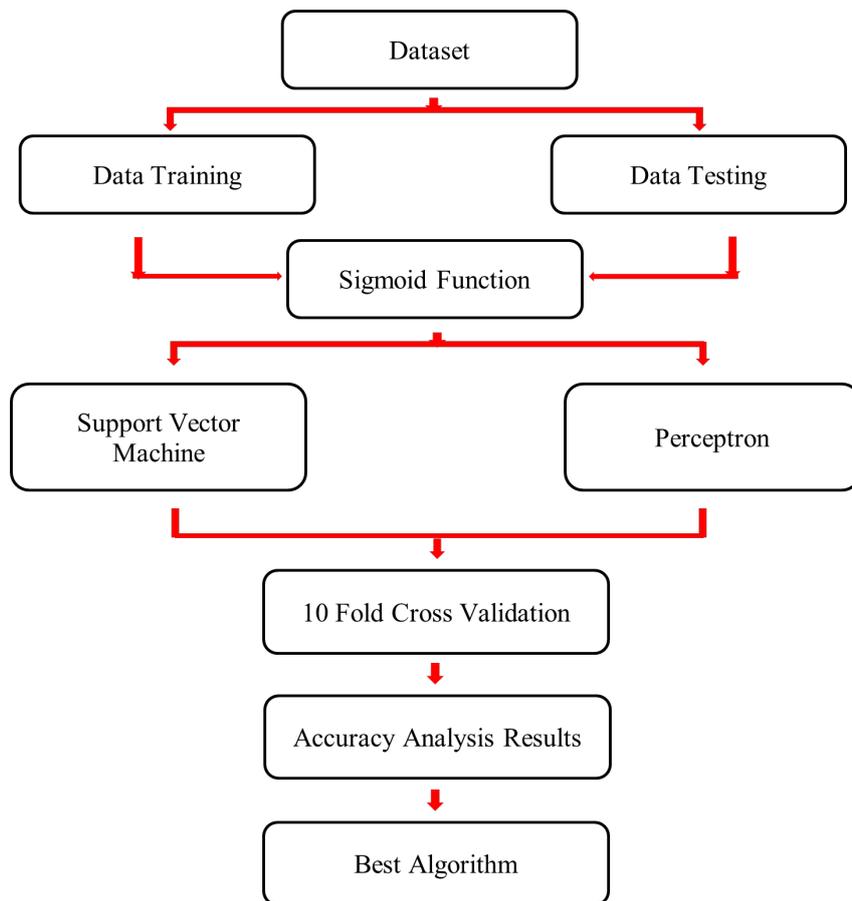


Figure 1. Research Design Architecture

1. Model Dataset

In conducting research, datasets are used in the form of a range from 2018 to 2021 in the sense of using data that has been grouped for three years, thus testing validation data from every three years to find out from the value of the determination of accuracy every year after testing based on the year of testing. From the dataset itself, it has more than 600 data on annual performance work program plans where the average annual performance plan range is 40% to 28%, which means that from one year, the work plan from P2KBP3A is around 180 to 260 work plan data made and must be carried out annually. This study applied the sigmoid activation function to the SVM and perceptron algorithms to classify the best work program at P2KBP3A Deli Serdang based on pre-determined categories. The results of the classification with the output of target achievement and realization of work programs with low, medium, and high categories are then analyzed using cross-validation with a value of $K = 10$ to obtain the accuracy, precision, recall, and f-measure values for each algorithm. These values are then compared to determine which algorithm, between SVM and perceptron, the sigmoid activation function results in a better classification of the best working program problem.

This study compares the support vector machine and perceptron, which consists of the sigmoid function, to the best work program dataset in the P2KBP3A case study using the Orange 3.30 application, which uses different annual calculations from 2018 to 2021, which is carried out annually in the form of a model (as shown in Figure 2).

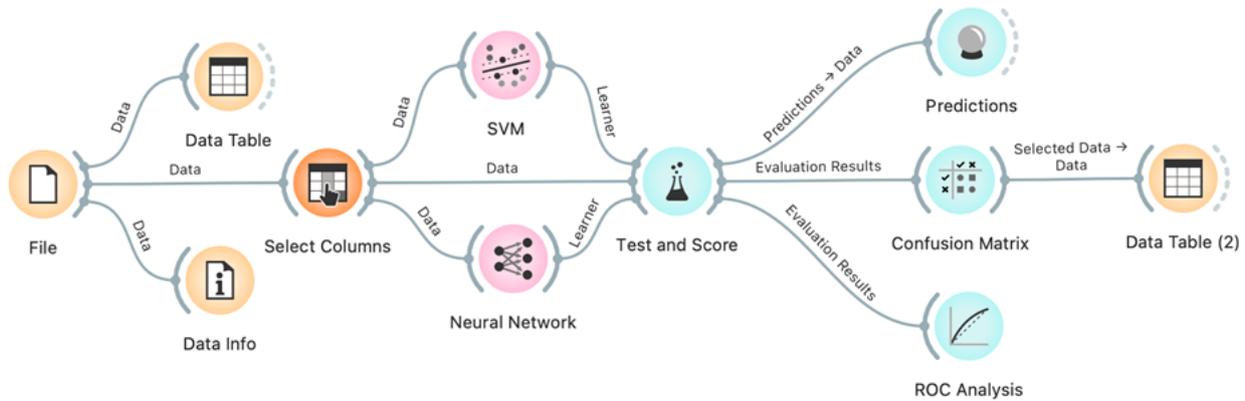


Figure 2. Classification Model

This study also has details to identify the classification of the models used from SVM and perceptron to activate the sigmoid function with the caption Figure 3 SVM Model and Perceptron model.

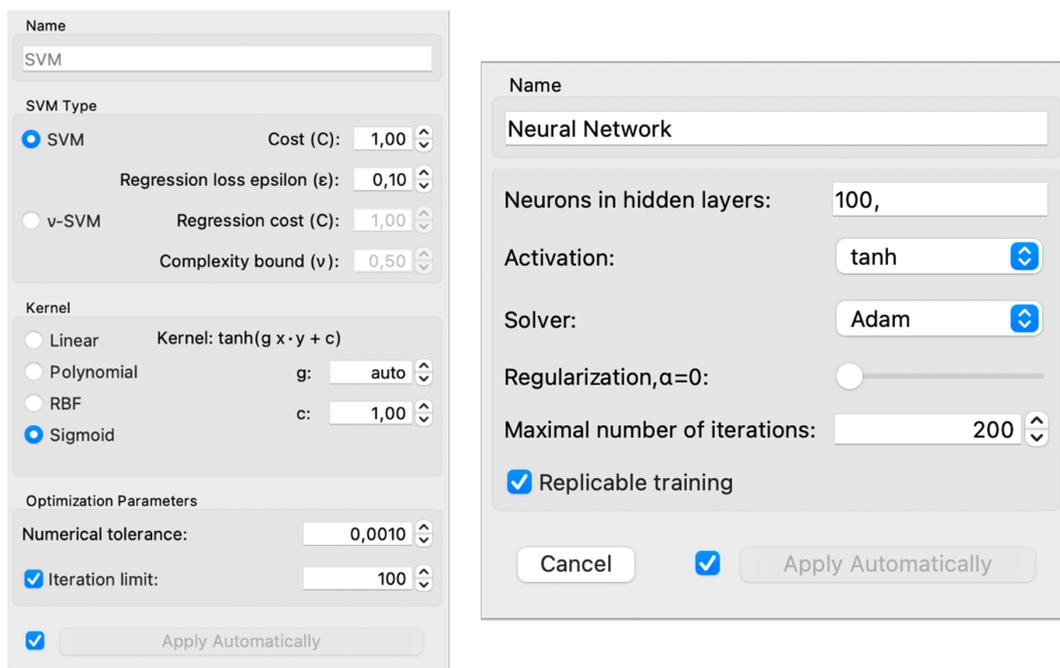


Figure 3. Sigmoid Function Model of SVM and Perceptron

2. Work Program Plan Data

Data from the work program plan from P2KBP3A, which is used as reference material to determine the feasibility of the work program, from the work program plan that is used as data taken with a range of years starting from 2018 to 2021 as test material. From each work program data, each year has a different range of work programs with the following specifications:

1. Data for 2018 with a total number of work program plans of as many as 262 implementers.
2. Data for 2019 with a total number of work program plans of as many as 189 implementers.
3. Data for 2020 with a total number of work program plans of as many as 174 implementers.
4. Data for 2021 with a total number of work program plans of as many as 161 implementers.

The result is that each year has a reduced work program from the initial year, which has filtered the feasibility of the work program but does not demand the possibility of increasing each work program. Therefore, the features used in determining the criteria for the work program plan are 1 and 1 as the target. The following features are used, and the explanations (are shown in Table 1).

Table 1. Classification Features and Targets

| Number | Feature |
|--------|---|
| 1 | Participation Number of Assisted Villages |
| 2 | Activity Duration |
| 3 | Budget Source |
| 4 | Budget |
| 5 | Work Program Infrastructure |
| 6 | Number of Employee Participation |
| 7 | Number of Participants |
| 8 | Sisa Anggaran Biaya |
| 9 | Number of District Activities 1 Month |
| 10 | Acceptance of Implementation Results |
| 11 | Activity Level |
| 12 | Indicative Fund Remaining Initial Fund Budget |
| 13 | Work Plan Year |
| 14 | Achievement Targe |

The description of the features and targets in Table 1 used is as follows:

1. The number of assisted villages participating explains how many villages have participated in the implementation of the work program; the more assisted villages that participate in these activities, the better.
2. Activity duration explains how long it takes to carry out activities from the implementation of the work program, where the longer the implementation, the better.
3. Sources of budget funds, explaining the sources of funds available for the initial design of the work program to be implemented.
4. Budget explains the budget issued by the government for the implementation of work programs where the more budget spent, the better.
5. The infrastructure work program describes the location of the infrastructure available from the implementation of the work program.
6. The number of participants describes the participants who took part in activities from each village; the more participants, the better.
7. The number of sub-district activities for one month explains the number of work program activities that have been implemented for one month, and the more work programs implemented, the better.
8. Remaining budget, explaining the remaining budget after carrying out a work program where the less remaining budget, the better.
9. Performance achievement targets for one month, explaining the achievements of the work programs that have been implemented where the more they are implemented, the better.
10. Acceptance of implementation results explains the achievement value of what has been implemented in the work program where the value range is between 0 and 100.
11. Activity level describes the activity level of the city, district, sub-district, and village.
12. The remaining funds are indicative of the initial budget, explaining the difference between budgeted costs and the remaining budgeted funds; the smaller the difference, the better.
13. The year of work plan describes the range of years the work program has been implemented.
14. Target achievement describes the achievement target of the work program obtained, namely feasible or not feasible.

The following is a sample of the data for the work program plan carried out starting in 2018 (as seen in Table 2).

Table 2. 2018 Work Program Plan Data

| Number | Work Program Plan Indicator |
|--------|---|
| 1 | Coordinating meetings and consultations outside and within the region |
| 2 | Provision of office administration services |
| 3 | Provision of Water Installation Equipment |
| 4 | Provision of Equipment and Work Equipment |
| 5 | Provision of office publication services |
| 6 | Development of Women's Organizations |
| 7 | Education and Training Activities to Increase Participation and Gender Equality |
| 8 | Counseling Activities for Housewives in Building a Prosperous Family |
| 9 | Business Management Guidance Activities for Women in Business Management |
| 10 | Exhibition of Women's Work in the Development Sector |

The following is a sample of the data for the work program plan carried out starting in 2019 (as seen in Table 3).

Table 3. 2019 Work Program Plan Data

| Number | Work Program Plan Indicator |
|--------|---|
| 1 | Provision of Reading Materials and Legislation |
| 2 | Provision of Food and Beverages |
| 3 | Coordination and Consultation Meetings Outside the Region |
| 4 | Provision of Office Administration Services |
| 5 | Provision of work tools and equipment |
| 6 | Construction of the Office House |
| 7 | Office Building Construction |
| 8 | Procurement of Service/Operational Vehicles |
| 9 | Procurement of furniture |
| 10 | Routine / Periodic Maintenance of Office Buildings |

The following is a sample of the data for the work program plan carried out starting from 2020 in Table 4.

Table 4. 2020 Work Program Plan Data

| Number | Work Program Plan Indicator |
|--------|--|
| 1 | Strengthening the integration of BKB Posyandu PADU |
| 2 | Formulation of Policies for Improving the Quality of Life of Women in the Field of Science and Technology |
| 3 | Formulation of Policy to Increase the Role and Position of Women in the Field of Politics and Public Positions |
| 4 | Implementation of Socialization related to Gender equality, |
| 5 | Women's Empowerment and Child Protection |
| 6 | Monitoring, evaluation, and reporting |
| 7 | Convention on the Rights of the Child (CRC) Training |
| 8 | Socialization of the Formation of Elderly Women's Empowerment Institutions (LPPLU) |
| 9 | Socialization and Advocacy of Women's Representation in Parliament |
| 10 | Facilitating the Development of Integrated Service Centers for Women's Empowerment (P2TP2) |

The following is a sample of the data for the work program plan carried out starting from 2021 in Table 5.

Table 5. 2021 Work Program Plan Data

| Number | Work Program Plan Indicator |
|--------|--|
| 1 | Strengthening the integration of BKB Posyandu PADU |
| 2 | Formulation of Policies for Improving the Quality of Life of Women in the Field of Science and Technology |
| 3 | Formulation of Policy to Increase the Role and Position of Women in the Field of Politics and Public Positions |
| 4 | Implementation of Socialization related to Gender equality, |
| 5 | Women's Empowerment and Child Protection |
| 6 | Monitoring, evaluation, and reporting |
| 7 | Convention on the Rights of the Child (CRC) Training |
| 8 | Socialization of the Formation of Elderly Women's Empowerment Institutions (LPPLU) |
| 9 | Socialization and Advocacy of Women's Representation in Parliament |
| 10 | Facilitating the Development of Integrated Service Centers for Women's Empowerment (P2TP2) |

Explanations from Tables 2, 3, 4, and 5 are sample work program data that have been implemented in 2018, 2019, 2020, and 2021.

As for the split of the scale of each feature into a range of numbers starting from 1-5 on each type of feature for the extraction of the initial data value and also for the target used in the target name, the achievement is called the feasible and unfeasible extraction result where each target category. This achievement is given a score for worthy of being given a value of 1 while it is not worthy of being given a value of 0, the following is an explanation of the results of the features and extraction targets (Show Table 6, Table 7, Table 8, Table 9, Table 10, Table 11, Table 12, Table 13, Table 14, Table 15, Table 16, Table 17, and Table 18).

Table 6. Data Extraction Feature 1 – Participation Number of Assisted Villages

| Number | Extraction | Value Range |
|--------|------------|-------------|
| 1 | >20 | 5 |
| 2 | >15 | 4 |
| 3 | >10 | 3 |
| 4 | >5 | 2 |
| 5 | >0 | 1 |

Table 7. Data Extraction Feature 2 – Activity Duration

| Number | Extraction | Value Range |
|--------|------------|-------------|
| 1 | >15 | 5 |
| 2 | >10 | 4 |
| 3 | >7 | 3 |
| 4 | >4 | 2 |
| 5 | >0 | 1 |

Table 8. Data Extraction Feature 3 – Budget Source

| Number | Extraction | Value Range |
|--------|------------|-------------|
| 1 | APBD | 5 |
| 2 | APBD/BAK | 1 |

Table 9. Data Extraction Feature 4 – Budget

| Number | Extraction | Value Range |
|--------|--------------|-------------|
| 1 | >250 Million | 5 |
| 2 | >150 Million | 4 |
| 3 | >75 Million | 3 |
| 4 | >25 Million | 2 |
| 5 | >0 Rupiah | 1 |

Table 10. Data Extraction Feature 5 – Work Program Infrastructure

| Number | Extraction | Value Range |
|--------|-------------|-------------|
| 1 | Very good | 5 |
| 2 | Good | 4 |
| 3 | Very Enough | 3 |
| 4 | Enough | 2 |

Table 11. Data Extraction Feature 6 – Number of Employee Participation

| Number | Extraction | Value Range |
|--------|------------|-------------|
| 1 | >0 | 5 |
| 2 | >5 | 4 |
| 3 | >10 | 3 |
| 4 | >17 | 2 |
| 5 | >25 | 1 |

Table 12. Data Extraction Feature 7 – Number of Participants

| Number | Extraction | Value Range |
|--------|------------|-------------|
| 1 | >500 | 5 |
| 2 | >300 | 4 |
| 3 | >150 | 3 |
| 4 | >75 | 2 |
| 5 | >0 | 1 |

Table 13. Feature Extraction Data 8 – Number of District Activities 1 Month

| Number | Extraction | Value Range |
|--------|------------|-------------|
| 1 | >10 | 5 |
| 2 | >8 | 4 |
| 3 | >5 | 3 |
| 4 | >2 | 2 |
| 5 | >0 | 1 |

Table 14. Data Extraction Feature 9 – Remaining Budget

| Number | Extraction | Value Range |
|--------|-------------|-------------|
| 1 | >0 Rupiah | 5 |
| 2 | >5 Million | 4 |
| 3 | >10 Million | 3 |
| 4 | >25 Million | 2 |
| 5 | >50 Million | 1 |

Table 15. Data Extraction Feature 10 – Performance Achievement Target 1 Month

| Number | Extraction | Value Range |
|--------|------------|-------------|
| 1 | >10 | 5 |
| 2 | >7 | 4 |
| 3 | >4 | 3 |
| 4 | >2 | 2 |
| 5 | >0 | 1 |

Table 16. Feature Extraction Data 11 – Acceptance of Implementing Results

| Number | Extraction | Value Range |
|--------|------------|-------------|
| 1 | >85 | 5 |
| 2 | >70 | 4 |
| 3 | >60 | 3 |
| 4 | >50 | 2 |
| 5 | >0 | 1 |

Table 17. Feature Extraction Data 12 – Activity Level

| Number | Extraction | Value Range |
|--------|-------------|-------------|
| 1 | Village | 5 |
| 2 | Subdistrict | 4 |
| 3 | Regency | 3 |
| 4 | City | 2 |

Table 18. Feature Extraction Data 13 – Indicative Fund Remaining Initial Fund Design

| Number | Extraction | Value Range |
|--------|-------------|-------------|
| 1 | >0 Rupiah | 5 |
| 2 | >5 Million | 4 |
| 3 | >10 Million | 3 |
| 4 | >15 Million | 2 |
| 5 | >25 Million | 1 |

Table 19. Target Extraction Data – Achievements

| Number | Extraction | Value Range |
|--------|--------------|-------------|
| 1 | Worthy | 1 |
| 2 | Not feasible | 0 |

Explanations from Tables 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, and 18 are data extraction values from descriptions of features and targets from Table 1 by using work program data in Table 2, 3, 4 and 5 with work program data consisting of 2018, 2019, 2020 and 2021 that have been implemented.

The classification process with the SVM and perceptron algorithms uses the number of folds for cross-validation of 10 and the sigmoid activation function.

3. Evaluation

Based on the evaluation of the classification, it is scored to the size of the results of classification accuracy, F1, precision, and recall which are calculated using the equations (1), (2), (3), and (4) [28] :

$$Precision = \frac{tp}{tp + fp} \quad (1)$$

$$Recall = \frac{tp}{tp + fn} \quad (2)$$

$$Fscore = (1 + \beta^2) \frac{Precision + Recall}{\beta^2 + Precision + Recall} \quad (3)$$

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn} \quad (4)$$

Regarding the evaluation of the results used, one of which is a confusion matrix which is obtained from the results of accuracy, precision, and recall as well as from the ROC curve to measure the AUC value. That way, the larger the area under the curve (AUC), the better the prediction results. The following is Table 19 of the confusion matrix [29].

Table 20. Confusion Matrix

| Actually | Prediction | |
|----------|------------|-------|
| | True | False |
| True | TP | FN |
| False | FP | TN |

3. RESULT AND ANALYSIS

The results obtained in testing the data from the total dataset used are 754 data from 2018, 2019, 2020, and 2021, with detailed data as much as the work program in 2018 there are 262 data, in 2019 there are 189 data, in 2020 there are 174 data. In 2021 there are 161 there, and overall the data for very good results by carrying out tests and a similar level on the results of work program contributions where the existing results for carrying out the results of the classification experiment between SVM and perceptron are very good and more dominant in SVM as at 89, 2% from [30] then there is 80.5% from [31] as well as from the perceptron with a value of 96.2%, with a value of 85.0% from [14].

The results obtained in testing the data results are very good by carrying out tests and similar levels on the results of work program contributions where the existing results for carrying out the results of classification experiments between SVM and perceptron are very good and more dominant in SVM as in the value of 89.2% of Application Design for Signature Pattern Recognition Using the Support Vector Machine (SVM) Method, then there is 80.5% of the SVM Method for Classification of Primary School Teacher Education Journal Articles as well as from perceptron with a value of 96.2% of Student Knowledge Assessment with Perceptron Algorithmic Neural Networks, with a value of 85.0% of Cataract Detection Using Single Layer Perceptron Based on Smartphone.

Each classification model from Figure 3 is inputted into the SVM and perceptron models in Figure 2 with a maximum number of epochs of 100. Accuracy, F1, precision, and recall values are obtained from the classification results using variations of the sigmoid function using 10-Fold Cross Validation (shown in Table 21).

Table 21. Cross Validation Results

| Model | Activation Function | Fold | Year | Cross Validation | | | |
|------------|---------------------|------|---------|------------------|------|-----------|--------|
| | | | | Accuracy | F1 | Precision | Recall |
| SVM | Sigmoid | 10 | 2018 | 86.3 | 80.2 | 75.0 | 86.3 |
| | | | 2019 | 88.4 | 82.9 | 78.1 | 88.4 |
| | | | 2020 | 87.4 | 81.5 | 76.3 | 87.4 |
| | | | 2021 | 88.8 | 83.6 | 78.9 | 88.8 |
| | | | Average | 87.5 | 82.2 | 77.2 | 87.5 |
| Perceptron | Sigmoid | 10 | 2018 | 87.4 | 82.8 | 85.9 | 87.4 |
| | | | 2019 | 86.8 | 82.1 | 77.9 | 86.8 |
| | | | 2020 | 85.6 | 80.6 | 76.1 | 85.6 |
| | | | 2021 | 85.7 | 82.8 | 80.6 | 85.7 |
| | | | Average | 86.5 | 82.2 | 80.2 | 86.5 |

Table 21 can be seen from the results that used the sigmoid activation function in the evaluation with 10-fold cross-validation resulting in the highest accuracy value of 88.8% in 2021 for the SVM model, the highest F1 value of 83.6% in 2021 for SVM the SVM model, the highest precision value is 85.9% in 2018 for the perceptron model, and the highest recall value is 88.8% in 2021 for the SVM model.

For the results of using the algorithm with the SVM model and perceptron, which uses the sigmoid function with the use of work program data that has been implemented, the best on average is obtained every year from 2018, 2019, 2020 up to 2021 for an accuracy value of 87.5%, for the F1 value of 82.2%, for the precision value of 78.7%, for the recall value of 87.5%. Thus, the results in Table 21 get good results for 2021 which are already in the work program plan and are also higher than the previous year's average results in 2018, 2019, and 2020.

The results that occur for classification can be seen in the comparison of research results from the differences in accuracy values (shown in Table 22).

Table 22. Comparison Of Research Results

| Model | Result | | Model |
|-------|-----------|--------|------------|
| SVM | Accuracy | 87.5 % | 86.5 % |
| | F1 | 82.2 % | 82.2 % |
| | Precision | 77.2 % | 80.2 % |
| | Recall | 87.5 % | 86.5 % |
| | | | Perceptron |

The following are the results of the classification that has been done, which can be seen in the following confusion matrix display (Shown in Figure 4, Figure 5, Figure 6, and Figure 7).

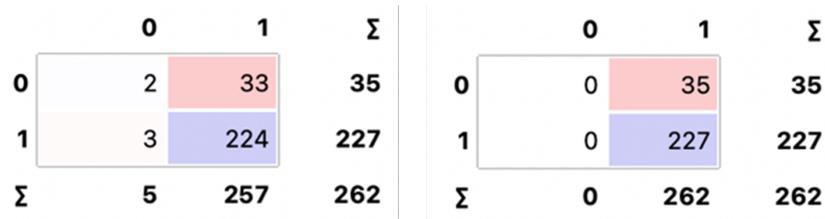


Figure 4. Confusion Matrix SVM Perceptron The year 2018

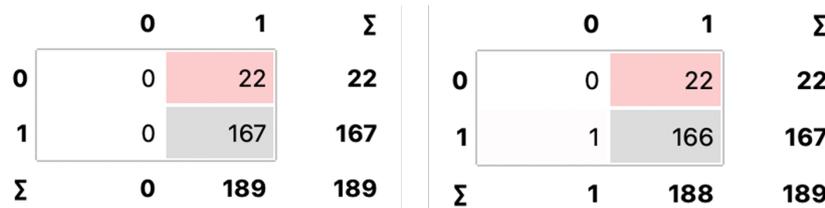


Figure 5. Confusion Matrix SVM Perceptron The year 2019

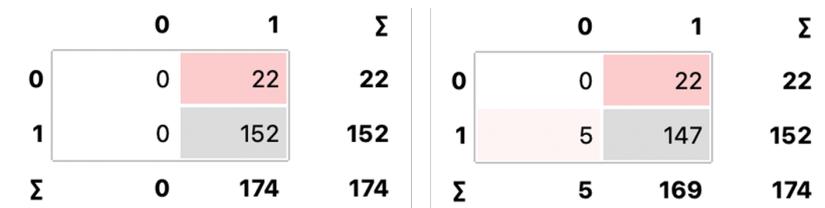


Figure 6. Confusion Matrix SVM Perceptron The year 2019

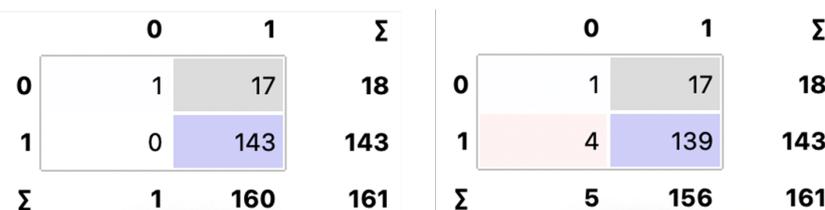


Figure 7. Confusion Matrix SVM Perceptron The year 2021

4. CONCLUSION

From the evaluation results obtained in finding the best implementation of the work program from government agencies in P2KBP3A itself by using a comparison of the SVM and perceptron algorithms that use the sigmoid function and also the results of the first test for the dataset tested, it can be concluded from the results of the research stated that the SVM algorithm model is more dominant in the new P2KBP3A data collection from 2018 to 2021 which has the highest level of accuracy in the field of the model. On this occasion, in conducting research, the data used is original data and only tested for new cases. It is also hoped that further research from this study can test using other than k-fold 10 and also adopting other algorithm models or can also use function models other than sigmoid. For further research, it is recommended to use the model with other models or make comparisons with other model algorithms and also use cross-validation using different numbers such as K-5 or K-20. Then, the level of distribution of training data and test data can be used with a comparison of 70% and 30% or other because, in this study, 80% and 20% were used. With backtesting, you can also use the model by combining the existing activation functions or also using several models that are more than the existing ones.

5. ACKNOWLEDGEMENTS

Thank you to Mrs. Dr. Rika Rosnelly, S.Kom., M.Kom. and Mr. Dr. Zakarias Situmorang, MT., for finishing this scientific work to the best of his ability.

6. DECLARATIONS

AUTHOR CONTRIBUTION

The first author is me, Jaka Tirta Samudra, S.Kom. as the author for each data collection and data extraction on the dataset; the second author is Mrs. Dr. Rika Rosnelly, S.Kom., M.Kom. as analytical research on every aspect of existing data and also conducting research in research methodology. The third author is Dr. Zakarias Situmorang, MT. as a helper and correction of the results of each result of data mining object and also checks the results of the classification.

FUNDING STATEMENT

Statements regarding financing in conducting research are the personal property of authors 1, 2, and 3.

COMPETING INTEREST

I have no declaration under financial, general, and institutional competing interests.

REFERENCES

- [1] S. Tamaela, Y. Andry Lesnussa, V. Y.I. Ilwaru, and A. M. Balami, "Analysis of Support Vector Machine (SVM) Method and Simple Additive Weighting (SAW) Method in Making Decisions to Choose Specialization," *Jurnal Matematika MANTIK*, vol. 6, no. 2, pp. 104–113, 2020.
- [2] S. Hartati, A. Harjoko, R. Rosnelly, I. Chandradewi, and Faizah, *Performance of SVM and ANFIS for Classification of Malaria Parasite and Its Life-Cycle-Stages in Blood Smear*. Springer Singapore, 2019, vol. 937.
- [3] N. A. Sinaga, Ramadani, K. Dalimunthe, M. S. A. A. Lubis, and R. Rosnelly, "Komparasi Metode Decision Tree, KNN, dan SVM untuk Menentukan Jurusan Di SMK," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 3, no. 2, pp. 94–100, 2021.
- [4] I. Permadi and A. K. Nugroho, "Klasifikasi Citra Menggunakan Kombinasi Jaringan Syaraf Tiruan Model Perceptron dan Algoritma One vs Rest," *INOVTEK Polbeng - Seri Informatika*, vol. 4, no. 2, p. 193, 2019.
- [5] A. Roihan, P. A. Sunarya, and A. S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review Paper," *IJCIT (Indonesian Journal on Computer and Information Technology)*, vol. 5, no. 1, pp. 75–82, 2020.

- [6] J. P. Pinto, S. Kelur, and J. Shetty, "Iris Flower Species Identification using Machine Learning Approach," *2018 4th International Conference for Convergence in Technology, I2CT 2018*, pp. 1–4, 2018.
- [7] Asmita Shukla, Ankita Agarwal, Hemlata Pant, and Priyanka Mishra, "Flower Classification using Supervised Learning," *International Journal of Engineering Research and*, vol. V9, no. 05, pp. 757–762, 2020.
- [8] C. Savas and F. DAVIS, "The Impact of Different Kernel Functions on the Performance of Scintillation Detection Based on Support Vector Machines," *Sensors (Switzerland)*, vol. 19, no. 23, pp. 1–16, 2019.
- [9] M. W. Huang, C. W. Chen, W. C. Lin, S. W. Ke, and C. F. Tsai, "SVM and SVM Ensembles in Breast Cancer Prediction," *PLoS ONE*, vol. 12, no. 1, pp. 1–14, 2017.
- [10] K. N and R. V, "Ensemble of Multiple Kernel SVM Classifiers for Detection of Online Spam Reviews," *SSRN Electronic Journal*, no. Iccinis, pp. 479–488, 2021.
- [11] A. Bekkanti, V. P. Gunde, S. Itnal, G. Parasa, and C. Z. Basha, "Computer-Based Classification of Diseased Fruit using K-means and Support Vector Machine," *Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology, ICSSIT 2020*, no. Iccsit, pp. 1227–1232, 2020.
- [12] I. M. Maysanjaya, "Comparative Study of Classification Method on Diagnosis of Plasmodium Phase," *Journal of Physics: Conference Series*, vol. 1516, no. 1, 2020.
- [13] M. W. Huang *et al.*, "Optimizing Stochastic Gradient Descent in Text Classification Based on Fine-Tuning Hyper-Parameters Approach. A Case Study on Automatic Classification of Global Terrorist Attacks," *Journal of Physics: Conference Series*, vol. 16, no. 1, pp. 1–6, 2019.
- [14] R. Sigit, E. Triyana, and M. Rochmad, "Cataract Detection using Single Layer Perceptron-Based on Smartphone," *ICICOS 2019 - 3rd International Conference on Informatics and Computational Sciences: Accelerating Informatics and Computational Research for Smarter Society in The Era of Industry 4.0, Proceedings*, 2019.
- [15] K. Hulliyah, "Predicting Airline Passenger Satisfaction with Classification Algorithms," *IJIS: International Journal of Informatics and Information Systems*, vol. 4, pp. 82–94, mar 2021.
- [16] B. Yanto, J. Jufri, A. Lubis, B. Hayadi, and E. Armita, NST, "Klarifikasi Kematangan Buah Nanas Dengan Ruang Warna Hue Saturation Intensity (Hsi)," *INOVTEK Polbeng - Seri Informatika*, vol. 6, no. 1, p. 135, 2021.
- [17] A. Wanto, B. Herawan Hayadi, P. Subekti, D. Sudrajat, R. Wikansari, G. W. Bhawika, E. Sumartono, and S. Surya, "Forecasting the Export and Import Volume of Crude Oil, Oil Products and Gas Using ANN," *Journal of Physics: Conference Series*, vol. 1255, no. 1, 2019.
- [18] B. H. Hayadi, I. G. I. Sudipa, and A. P. Windarto, "Model Peramalan Artificial Neural Network pada Peserta KB Aktif Jalur Pemerintahan menggunakan Artificial Neural Network Back-Propagation," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 1, pp. 11–20, 2021.
- [19] H. A. G. Al-kaf, K. S. Chia, and N. A. M. Alduais, "A Comparison between Single Layer and Multilayer Artificial Neural Networks in Predicting Diesel Fuel Properties using Near-Infrared Spectrum," *Petroleum Science and Technology*, vol. 36, no. 6, pp. 411–418, 2018.
- [20] B. Yanto, B. J. , and B. Hayadi, "Identifikasi Pola Aksara Arab Melayu dengan Jaringan Syaraf Tiruan Convolutional Neural Network (CNN)," *JSAI (Journal Scientific and Applied Informatics)*, vol. 3, no. 3, pp. 106–114, 2020.
- [21] I. C. R. Drajana, "Metode Support Vector Machine dan Forward Selection Prediksi Pembayaran Pembelian Bahan Baku Kopra," *ILKOM Jurnal Ilmiah*, vol. 9, no. 2, pp. 116–123, 2017.
- [22] Y. U. Budiman, P. Studi, T. Informasi, U. Bina, S. Informatika, M. Learning, S. Layer, and P. Network, "Identifikasi Kata Benda dan Bukan Kata Benda Menggunakan," vol. 01, no. 5, pp. 759–768, 2022.

- [23] E. Yuliani, A. N. Aini, and C. U. Khasanah, "Perbandingan Jumlah Epoch dan Steps Per Epoch pada Convolutional Neural Network untuk Meningkatkan Akurasi dalam Klasifikasi Gambar," *Jurnal INFORMA Politeknik Indonusa Surakarta*, vol. 5, no. 2442-7942, pp. 2–6, 2019.
- [24] J. T. Samudra and P. S. Ramadhan, "Sistem Pendukung Keputusan Mencari Pelaksana Program Kerja Terbaik Menggunakan Metode MOORA," *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika dan Komputer)*, vol. 21, no. 1, p. 10, 2022.
- [25] I. Romli and B. M. Putra, "Evaluasi Penilaian Kinerja Dalam Klasifikasi Data Mining dengan metode Naïve Bayes," vol. 1, no. 1, pp. 36–45, 2020.
- [26] N. Ketut, M. Astuti, N. W. Utami, I. G. Putu, and K. Juliharta, "Classification of Blood Donor Data Using C4 . 5 and K-Nearest Neighbor Method (Case Study : Utd Pmi Bali Province)," vol. 18, no. 1, 2022.
- [27] H. A. G. Al-kaf *et al.*, "Predicting Airline Passenger Satisfaction with Classification Algorithms," *IJIS: International Journal of Informatics and Information Systems*, vol. 4, no. 1, pp. 82–94, mar 2021.
- [28] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, R. Sun, and I. Garcíá-Magarinõ, "A Hybrid Intelligent System Framework for the Prediction of Heart Disease using Machine Learning Algorithms," *Mobile Information Systems*, vol. 2018, 2018.
- [29] A. W. Putri, "Implementasi Artificial Neural Network (ANN) Backpropagation untuk Klasifikasi Jenis Penyakit pada Daun Tanaman Tomat," *MATHunesa: Jurnal Ilmiah Matematika*, vol. 9, no. 2, pp. 344–350, 2021.
- [30] G. Gunawan and Y. Reswan, "Desain Aplikasi Pengenalan Pola Tanda Tangan Menggunakan Metode Support Vector Machine (Svm)," *Jurnal Media Infotama*, vol. 17, no. 1, pp. 8–12, 2021.
- [31] U. Pujianto, I. A. E. Zaeni, and N. O. Irawan, "SVM Method for Classification of Primary School Teacher Education Journal Articles," *ICEEIE 2019 - International Conference on Electrical, Electronics and Information Engineering: Emerging Innovative Technology for Sustainable Future*, no. 1, pp. 324–329, 2019.