IPTEK The Journal of Technology and Science, 32(2), 2021 (e/pISSN:2088-2033, 0853-4098)
DOI: 10.12962/j20882033.v32i2.10483 | Received 22 Aug, 2021; Accepted 11 Oct, 2021

ORIGINAL RESEARCH

PERFORMANCE STUDY OF UNCERTAINTY BASED FEATURE SELECTION METHOD ON DETECTION OF CHRONIC KIDNEY DISEASE WITH SVM CLASSIFICATION

Lailly Syifa'ul Qolby | Joko Lianto Buliali* | Ahmad Saikhu

[1]Dept. of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Correspondence

*Joko Lianto Buliali, Dept of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. Email: joko@if.its.ac.id

Present Address

Gedung Teknik Informatika, Jl. Teknik Kimia, Surabaya 60111, Indonesia

Abstract

Chronic Kidney Disease (CKD) is a disorder that impairs kidney function. Early signs of CKD patients are very difficult until they lose 25% of their kidney function. Therefore, early detection and effective treatment are needed to reduce the mortality rate of CKD sufferers. In this study, the authors diagnose the CKD dataset using the Support Vector Machine (SVM) classification method to obtain accurate diagnostic results. The authors propose a comparison of the result on applying the feature selection method to get the best feature candidates in improving the classification result. The testing process compares the Symmetrical Uncertainty (SU) and Multivariate Symmetrical Uncertainty (MSU) feature selection method and the SVM method as a classification method. Several experimental scenarios were carried out using the SU and MSU feature selection methods using the CKD dataset. From the results of the tests carried out, it shows that using the MSU feature selection method with 80%: 20% data split produces nine important features with an accuracy value of 0.9, sensitivity 0.84, specification 1.0, and when viewed on the ROC graph, the MSU method graph shows the true positive value is higher than the false positive value. So the classification using the MSU feature selection method is better than the SU feature selection method by 90% accuracy.

KEYWORDS:

Chronic Kidney Disease, Feature Selection, Support Vector Machine, Uncertainty

1 | INTRODUCTION

At this time, the health care system is supported by advanced capabilities such as machine learning, data mining, and artificial intelligence to provide health services smarter and more trusted. Data mining is used in health service management, information health, patient care systems, and others in detecting and predicting disease. In addition, data mining also has a major role in analyzing the survivability of a disease. One of the harmful diseases that attack our organs is the kidney, namely Chronic Kidney Disease (CKD). CKD is a heterogeneous disorder that affects kidney structure and function. CKD is an important cause of death because early signs of CKD sufferers are very difficult to detect until the patient loses 25% of kidney function [1]. Therefore it is needed early detection and effective treatment to reduce patient mortality.

In the research to diagnose CKD, some types of feature selection reduce CKD dataset dimensions. The results show that the classification with the SVM method and the feature selection method Best First Search has a high level of accuracy in diagnosing CKD compared with other selected methods [2]. In Sosa-Cabrera et al. [1], to analyze the behavior of Multivariate Symmetrical Uncertainty (MSU) using statistical simulation techniques with a mix of randomly generated informative and non-informative features and MSU as part of the feature selection process. This research shows how the number of attributes, cardinality, and sample size affect MSU. The results obtained for conditions that maintain good quality at MSU under different combinations of the three factors provide useful new criteria to help the dimension reduction process. Based on research that has been done previously, in this study, the proposed process of literature study on the method Uncertainty-based feature selection for diagnosis CKD uses the Support Vector Machine. This research is expected to obtain the best feature selection method and the optimal feature set for CKD classification.

2 | PREVIOUS RESEARCHES

There is some previous research related to our research. Preventing CKD is one of the most attractive tasks for health workers. The main objective of this research is to analyze the comparison results of the Naïve Bayes Method, Multi-Layer Perceptron, and Support Vector Machine. Several pre-processing techniques are used, such as unsupervised discretization and normalization, to increase the accuracy value. The value of accuracy and the time required for classification are taken as the results of the study. This study states that the implementation results using SVM are superior to other classification methods [3].

There are two types of selection features to diagnose CKD; namely, wrappers and filters were chosen to reduce the dimensions of the CKD dataset. The results show that the SVM classification method and the Best First Search as the wrapper evaluation subset have a higher level of accuracy in the diagnosis of CKD than other methods selected [2]. Another research proposed a Symmetrical Uncertainty (SU) based feature subset generation and ensemble learning method for the electricity customer classification. The results show that the proposed method efficiently finds useful feature subsets and improves classification performance [4].

Multivariate Symmetrical Uncertainty (MSU) is measured as an extension of the SU to the multivariate case. It is applied to feature selection on synthetic and real-world data; it can be used as a new feature subset evaluation method to capture linear and non-linear correlation and interactions [1]. Given the previous studies, we compared the result applying the feature selection method to get the best feature candidates to improve the classification result. The testing process compares the SU and MSU feature selection method and the SVM method as a classification method.

3 | MATERIAL AND METHOD

This research presents a literature study on the method of Uncertainty-based feature selection for diagnosis of CKD using the Support Vector Machine. Symmetrical Uncertainty (SU) and Multivariate Symmetrical Uncertainty (MSU) are feature selection methods that give a different classification result using SVM.

3.1 | Dataset

The dataset used in this study is Chronic Kidney Disease (CKD) data obtained from the UCI Machine Learning Website with a total of 400 data with 250 data labeled “ckd” and 150 labeled as “notckd." The dataset has 24 attributes and 1 class, 11 of which are numeric, and the other 13 are nominal attributes shown in Table 1 . In the dataset, 1012 missing values can affect the level of accuracy.

TABLE 1 The dataset description.

Attributes	Definition	Value	Attributes	Definition	Value
age	Age	Years	pot	Potassium	mEq/L
bp	Blood Pressure	Mm/Hg	hemo	Hemoglobin	gms
sg	Specific Gravity	1.005,1.01,1.015,1.020,1.025	pcv	Packed Cell Volume	0,1,2,...
al	Albumin	0,1,2,3,4,5	wbcc	White Blood Cell Count	cells/cumm
su	Sugar	0,1,2,3,4,5	rbcc	Red Blood Cell Count	millions/cumm
rbc	Red Blood Cells	Normal, Abnormal	htn	Hypertension	Yes,No
pc	Pus Cell	Normal, Abnormal	dm	Diabetes Mellitus	Yes,No
pcc	Pus Cell Clumps	Present, Not Present	cad	Coronary Artery Disease	Yes,No
ba	Bacteria	Present, Not Present	appet	Appetite	Good, Poor
bgr	Blood Glucose Random	mgs/dl	pe	Pedal Edema	Yes,No
bu	Blood Urea	mgs/dl	ane	Anemia	Yes,No
sc	Serum Creatinine	mgs/dl	class	CKD, notCKD	CKD, Not CKD
sod	Sodium	mEq/L			

3.2 | Missing Data

Missing value or missing data is a condition when some value that was originally wanted to be obtained during data collection cannot be obtained for several reasons [5]. Missing data can cause various problems. First, the absence of data reduces statistical power, which refers to the probability that the test will reject the null hypothesis when it is false. Second, missing data can lead to bias in parameter estimation. Third, it can reduce the representativeness of the sample. Fourth, it complicates the research analysis. At the same time, the machine learning algorithms for classification require a complete dataset [6].

To overcome missing values in the dataset, data imputation is applied. The selection of the imputation method is usually determined by how the values are missing [7]. There are some missing data handling methods such as missing data ignoring technique, missing data imputation methods, and missing data model base technique [8].

In this study missing data imputation method is used. Because there are two types of attributes, nominal and numeric, mean and mode substitution are used. This method replaces the missing value with the middle or median value of the variable. This method maintains the sample size and is easy to use. Still, the variability in the data is reduced so that the standard deviation and variance estimates tend to be underestimated. The magnitude of covariance and correlation is also reduced by limiting variability, and this method often leads to biased estimates. In addition, this approach does not add new information but only increases the sample size and leads to underestimating errors.

3.3 | Discretization

Discretization is finding a set of cut-points for some continuous features by partitioning the range into a small number of intervals with good class coherence [9]. The interval label can then be used to replace the actual data values. In this study, numerical discretization is performed on a dataset with nominal type, namely the value of an unordered set. This is because some classification algorithms can only accept categorical attributes.

3.4 | Feature Selection

Many input features are a fundamental problem in many fields, especially forecasting, classification, bioinformatics, and object recognition. A typical solution is to use specific techniques to reduce the dimensionality of the original problem, eliminating redundant, irrelevant, or noise data. Feature selection, which builds a subset of the original features, is advantageous when interpretability and knowledge extraction is crucial, as in medicine. However, sometimes this comes at the cost of losing some accuracy [10]. From the feature selection stage, optimal features are obtained, accelerating the data mining process, improving the quality and performance of data mining, and improving the completeness of mining results. The main challenge of feature reduction is identifying the best feature subset to achieve the best classification results.

3.5 | Symmetrical Uncertainty (SU)

SU is shown to be effective for large-scale datasets [11]. SU is a simple and efficient feature subset selection method to evaluate the goodness of classification features. Features that have a higher SU value get a higher weight. The SU correlation measure is an information-based measure that uses entropy values and conditional conditions to determine correlations between pairs of features[1].

The measure of SU correlation is a measure of the uncertainty of the random variable. The entropy H of the discrete random variable X, with $x_1, ..., x_n$ as the possible values and the probability mass function $P(X)$, is a measure of the uncertainty in predicting the value of X defined as Equation 1.

$$H(X) := -\sum_{j} P(x_i) log_2(P(x_i)) \tag{1}$$

Where $H(X)$ can be interpreted as a measure of an endless variation of X, or the amount of information needed to predict or describe the outcome of X. With the discrete random variable Y, the conditional entropy $H(X|Y)$ quantifies the amount of information needed to describe the result X given that the value of Y is known and defined as Equation 1.

$$H(X|Y) := -\sum_{j} \left[P(x_i|y_j) log_2(P(x_i|y_j))\right] \tag{2}$$

Where P is the prior probability of the value of Y and is the posterior probability of the value for the variable X given that the value of the variable Y is . The Information Gain (IG(X|Y)) of variable X for a given variable Y measures the reduction in uncertainty about the value of X when the value of Y is known, defined as Equation 3.

$$IG(X|Y) := H(X) - H(X|Y) \tag{3}$$

IG measures how much knowledge about Y makes the value of X easier to predict so that it can be used as a correlation measure. It can be shown that $IG(X_j|Y)$ is a symmetric measure which is a convenient property for pairwise sizes. The IG value can be normalized using the two entropies derived from the SU size expressed as Equation 4.

$$SU(X, Y) := 2\left[\frac{IG(X|Y)}{H(X)H(Y)}\right] \tag{4}$$

The main limitation of SU is that it only considers pairwise interactions to fail to detect redundancy when dealing with more than two features.

3.6 | Multivariate Symmetrical Uncertainty (MSU)

MSU is a method developed from SU which aims to improve SU's shortcomings to measure redundancy between more than two features[12]. To overcome the lack of SU, MSU must be determined by defining the total correlation for n variables according to the following Equation 5:

$$C(X_{1:n}) := 2\sum_{i=1}^{n} H(X_i) - H(X_{1:n}) \tag{5}$$

Where $H(X_{1:n})$ is obtained from

$$H(X_{1:n}) := H(X_1, ..., X_n) := \sum_{x_1} ... \sum_{x_n} P(x_1, ..., x_n) log_2\left[P(x_1, ..., x_n)\right] \tag{6}$$

It is the entropy combination of the random variables $X_1, ..., X_n$. So that the definition of MSU [0,1] is obtained as Equation 7.

$$MSU(X_{1:n}) := \frac{n}{n-1}\left[\frac{C(x_{1:n})}{\sum_{i1}^{n} H(X_i)}\right] \tag{7}$$

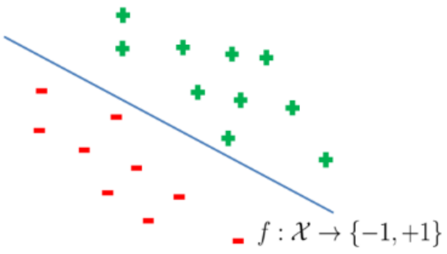

FIGURE 1 The binary classification.

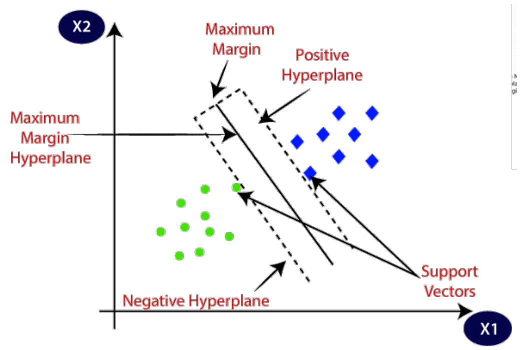

FIGURE 2 The hyperplane diagram with two feature.

3.7 | Support Vector Machine (SVM)

SVM is a supervised machine learning technique used for both classification and regression problems[13]. SVM has a good performance on high dimensional data to avoid the curse of dimensionality[14]. It can efficiently perform non-linear classifications, implicitly mapping their inputs into high-dimensional feature spaces. The main goal of SVM is to find the decision rule with the maximum margin. A function f is determined by a decision boundary that separates positive and negative samples. With the decision boundary, it can determine the decision rule. The SVM decision rule for the binary classification problem, as shown in Figure 1 , is defined as Equation 8.

$$f(X) = \begin{cases} +1, & \textit{if } g(x) \geq 0 \\ -1, & \textit{if } g(x) < 0 \end{cases} \tag{8}$$

Where $g(x)$ is the boundary.

In Figure 1 , the area under the blue line is negative class because $f(x) = -1, g(x) \leq 0$, and the area upper the blue line is positive class because $f(x) = +1, g(x) \geq 0$. The blue line in Figure 1 is the linear function in between ensures the best possible separation between the two categories, called a hyperplane. SVM searches for the best hyperplane[15].

The hyperplane is a function that can be used to separate between classes. The dimensions of the hyperplane depend on the features in the dataset, which means if there are two features, the hyperplane will be a straight line, as shown in Figure 2 . A separator hyperplane is a hyperplane that maximizes the distance between two parallel hyperplanes or what is known as the maximum margin[16]. The larger the margin, the better the generalization error of the classifier.

The data point or vector that is closest to the hyperplane and affects the hyperplane's position is called a Support Vector because this vector supports the hyperplane. SVM has a basic principle of linear classifier, which is a case that can be linearly separated. Still, currently, SVM is also being developed so that it can also work on non-linear problems by adding the kernel concept to a high-dimensional workspace. Linear SVM is used for data separated linearly or datasets classified into two classes using a single straight line. Non-linear SVM is used for non-linearly separated data or datasets that cannot be classified using straight lines. In this study, the type of kernel used is a non-linear kernel. Because the dataset used is non-linear data and in several studies[2, 17] explain that SVM can improve generalization performance by mapping inputs to high dimensional areas and solving quadratic programming on optimization so this study will use the RBF kernel equation as an Equation 9.

$$K(x_i, x_j) = \gamma x_i^T x_j + r)^d, \gamma > 0 \tag{9}$$

```
     X   age    bp    sg    al    su   rbc    pc   pcc    ba   bgr    bu    sc   sod   pot  hemo   pcv  wbcc
  0.00  2.25  3.00 11.75 11.50 12.25 38.00 16.25  1.00  1.00 11.00  4.75  4.25 21.75 22.00 13.00 18.00 27.00
  rbcc   htn    dm   cad appet    pe   ane class
 32.75  0.50  2.00  1.00  0.50  0.50  0.25  0.75
```

FIGURE 3 Percentage of missing data from each attribute.

```
          Reference
Prediction ckd notckd
    ckd     83     12
    notckd  17     48

               Accuracy : 0.8188
                 95% CI : (0.7502, 0.8751)
    No Information Rate : 0.625
    P-Value [Acc > NIR] : 7.918e-08

                  Kappa : 0.6197

 Mcnemar's Test P-Value : 0.4576

            Sensitivity : 0.8300
            Specificity : 0.8000
         Pos Pred Value : 0.8737
         Neg Pred Value : 0.7385
             Prevalence : 0.6250
         Detection Rate : 0.5188
   Detection Prevalence : 0.5938
      Balanced Accuracy : 0.8150

       'Positive' Class : ckd
```

FIGURE 4 The first scenario results of SU method.

4 | RESULTS AND DISCUSSION

In this section, we experiment and analyze the results obtained by comparing the effect of feature selection using a dataset. The results were based on some research steps such as pre-processing, feature selection, and classification. The results of the test scenarios carried out are as follows :

4.1 | Pre-Processing

UCI begins pre-processing the data using the Chronic Kidney Disease (CKD) dataset obtained from the research. The purpose of pre-processing is to eliminate missing data to obtain more accurate results. It was done using the substitution mode method. In addition to missing data, the pre-processing also discretized numerical data.

Start by exploring the initial data, namely uploading the dataset used. It is known that the amount of each missing data on each attribute is as in Figure 3 . Given Figure 3 , the percentage can be obtained by filling empty data using two different ways because of two different data types. For numeric data types, fill in the empty value with the mean value of each attribute. The nominal or categorical data-type were filled using the mode value.

After making sure all the data is filled in, the next step is changing the categorical variables into dummy data. This process is called the discretization process numeric. The numerical discretization process is caused by a combination or mix of numeric types and categories—data with the category type converted to data dummy except for the target variable (class attribute).

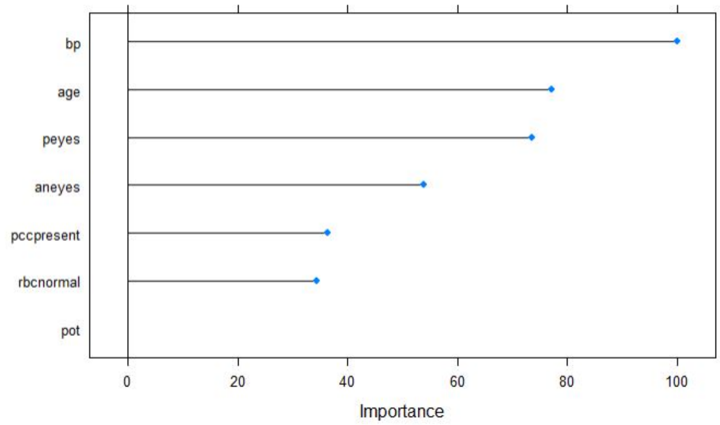

FIGURE 5 The important attributes of first scenario SU method plot graph.

```
Confusion Matrix and Statistics

          Reference
Prediction ckd notckd
    ckd     39      4
    notckd  11     26

               Accuracy : 0.8125
                 95% CI : (0.7097, 0.8911)
    No Information Rate : 0.625
    P-Value [Acc > NIR] : 0.0002292

                  Kappa : 0.6178

 Mcnemar's Test P-Value : 0.1213353

            Sensitivity : 0.7800
            Specificity : 0.8667
         Pos Pred Value : 0.9070
         Neg Pred Value : 0.7027
             Prevalence : 0.6250
         Detection Rate : 0.4875
   Detection Prevalence : 0.5375
      Balanced Accuracy : 0.8233

       'Positive' Class : ckd
```

FIGURE 6 The second scenario results of SU method.

4.2 | First Scenario of SU Method

In the first trial, SVM classification used the Symmetrical Uncertainty (SU) feature selection method. The data divided into 60% training data and 40% testing data is then entered into the SVM classification method in a linear model and is repeated five times. From the results of the classification obtained results as shown in Figure 4 .

Figure 4 shows that the accuracy is 0.8188, sensitivity is 0.83, specificity is 0.8. In Figure 5 , seven features are considered the most important: bp, age, peyes, aneyes, pccpresent, rbcnormal, and pot.

4.3 | Scenario of SU Method

In the second trial, SVM classification uses the Symmetrical Uncertainty (SU) feature selection method. The data divided into 80% training data and 20% testing data is then entered into the SVM classification method in a linear model and is repeated five times. From the classification results obtained results as shown in Figure 6 .

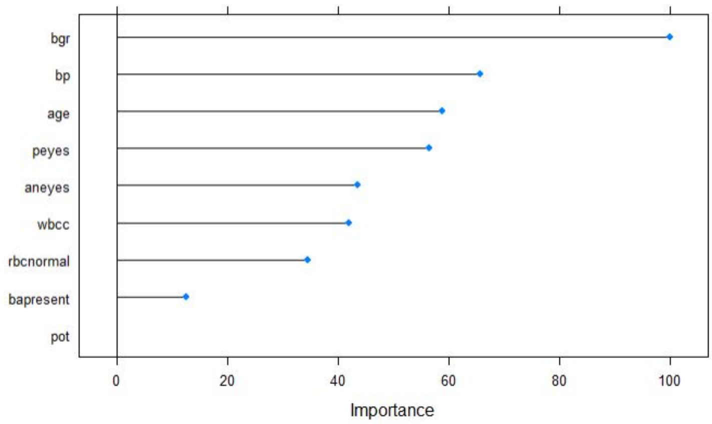

FIGURE 7 The important attributes of second scenario SU method olot graph.

```
Confusion Matrix and Statistics

          Reference
Prediction ckd notckd
    ckd     86     10
    notckd  14     50

               Accuracy : 0.85
                 95% CI : (0.7851, 0.9015)
    No Information Rate : 0.625
    P-Value [Acc > NIR] : 3.023e-10

                  Kappa : 0.6842

 Mcnemar's Test P-Value : 0.5403

            Sensitivity : 0.8600
            Specificity : 0.8333
         Pos Pred Value : 0.8958
         Neg Pred Value : 0.7812
             Prevalence : 0.6250
         Detection Rate : 0.5375
   Detection Prevalence : 0.6000
      Balanced Accuracy : 0.8467

       'Positive' Class : ckd
```

FIGURE 8 The First Scenario Results of MSU Method.

Figure 6 shows that the accuracy is 0.8125, sensitivity is 0.78, specificity is 0.8667. In Figure 7 , there were nine features which were considered the most important, namely bgr, bp, age, peyes, aneyes, wbcc, rbcnormal, bapresent, pot.

4.4 | First Scenario of Multivariate Symmetrical Uncertainty (MSU) Method

In the third trial, SVM classification uses the Multivariate Symmetrical Uncertainty (MSU) feature selection method. The data divided into 60% training data and 40% testing data is then entered into the SVM classification method in a linear model and is repeated five times. From the results of the classification obtained results as shown in Figure 8 .

Figure 8 shows that the accuracy is 0.85, sensitivity is 0.8600, specificity is 0.8333. In figure 9, seven features are considered the most important: bp, age, peyes, aneyes, pcopresent, rbcnormal, and pot.

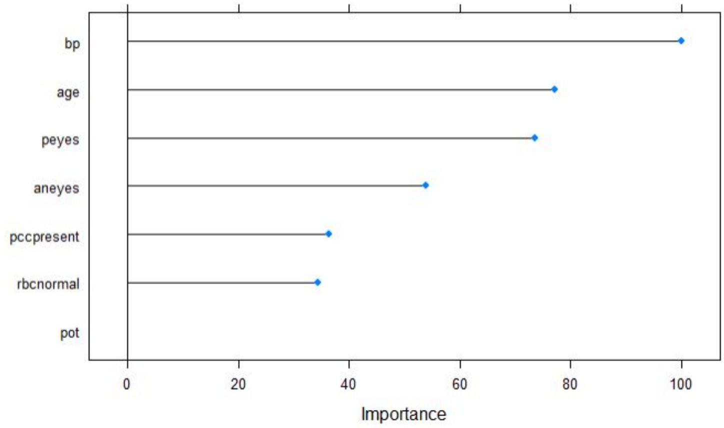

FIGURE 9 The important attributes of first scenario MSU method plot graph.

```
Confusion Matrix and Statistics

          Reference
Prediction ckd notckd
    ckd     42      0
    notckd   8     30

               Accuracy : 0.9
                 95% CI : (0.8124, 0.9558)
    No Information Rate : 0.625
    P-Value [Acc > NIR] : 2.771e-08

                  Kappa : 0.7975

 Mcnemar's Test P-Value : 0.01333

            Sensitivity : 0.8400
            Specificity : 1.0000
         Pos Pred Value : 1.0000
         Neg Pred Value : 0.7895
             Prevalence : 0.6250
         Detection Rate : 0.5250
   Detection Prevalence : 0.5250
      Balanced Accuracy : 0.9200

       'Positive' Class : ckd
```

FIGURE 10 The second scenario results of MSU method.

4.5 | Second Scenario of Multivariate Symmetrical Uncertainty (MSU) Method

In the fourth trial, SVM classification uses the Multivariate Symmetrical Uncertainty (MSU) feature selection method. The data divided into 80% training data and 20% testing data is then entered into the SVM classification method in a linear model and is repeated five times. From the classification results obtained results as shown in Figure 10 .

Figure 10 shows that the accuracy is 0.9, sensitivity is 0.8400, specificity is 1.0000. The graph of the features that are considered important can be seen in Figure 11 . Nine features are considered the most important: bgr, bp, age, peyes, aneyes, wbcc, rbcnormal, bapresent, and pot.

4.6 | Evaluation of Trial Comparison Results

To evaluate the results of the classification process, a confusion matrix can be calculated to measure the performance value of the method used. From the four trials that have been carried out, the results of the two feature selection methods used can be

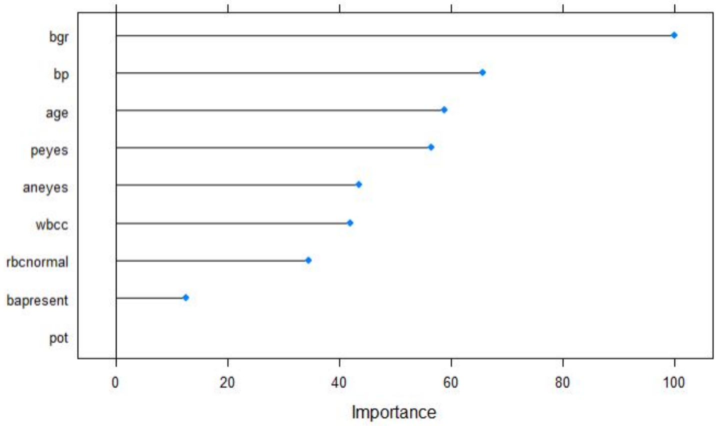

FIGURE 11 The important attributes of second scenario MSU method plot graph.

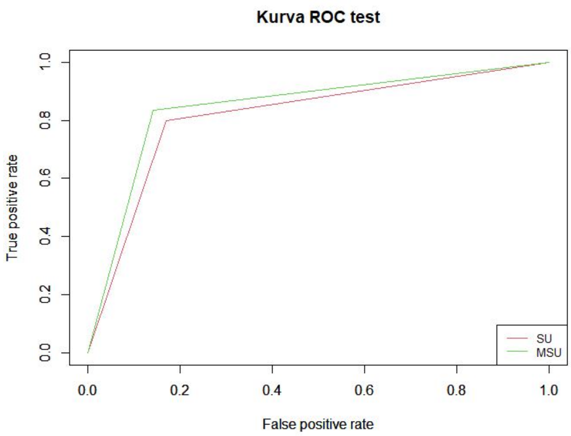

FIGURE 12 The SU and MSU ROC curves with a data ratio of 60%: 40%.

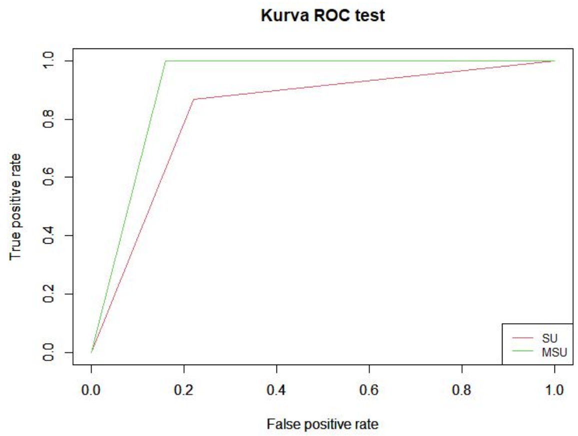

FIGURE 13 The SU and MSU ROC curves with a data ratio of 80%: 20%.

compared. In addition to using the results of the confusion matrix calculation, it can also be seen the results of comparing the two methods from the two comparisons of the amount of data in Figures 12 and Figure 13 .

Figures 12 and 13 are graphs or curves of Receiver Operating Characteristics (ROC) or Precision-Recall curves. The ROC curve is made based on the values that have been obtained in the calculation with the confusion matrix, namely between the False Positive Rate and the True Positive Rate. The curve is considered bad if the resulting curve is close to the baseline line or a line crosses the 0.0 point. In contrast, the curve will be considered good if the curve is close to the point 0.1.

It can be concluded that the MSU method or the green line is better than the SU method or the red line because the red line is closer to the baseline than the green line. Some errors that can affect the test results in this study are influenced by several things, namely the number of datasets used accompanied by some missing data and test scenarios carried out based on the number of comparisons of training data and test data.

5 | CONCLUSION

In classification feature selection is one of the important stages that can impact the result of classification. To get the best feature candidates in improving the classification result, two feature selection methods, i.e., Symmetrical Uncertainty (SU) and Multivariate Symmetrical Uncertainty (MSU), were used. This paper compares both feature selections to get the best result. The results show that SU has a lower value in the confusion matrix than MSU. Also, in the ROC graph, the MSU shows true positive value is higher than the false positive value compared with the SU. This paper concludes that the classification using the MSU feature selection method is better than the SU feature selection method. However, other factors can also influence these results, such as percentages of missing values, feature selection sizes, and datasets size.

CREDIT

Lailly Syifa'ul Qolby: Conceptualization, Methodology, Writing-original draft preparation and supervision; Formal analysis and investigation; **Joko Lianto Buliali:** Supervision, Writing-review and editing.**Ahmad Saikhu:** Supervision, Writing-review and editing.

References

1. Sosa-Cabrera G, García-Torres M, Gomez-Guerrero S, Schaerer CE, Divina F. A Multivariate Approach to The Symmetrical Uncertainty Measure: Application to Feature Selection Problem. Information Sciences 2019 August;494:1–20. https://doi.org/10.1016/j.ins.2019.04.046.

2. Polat H, Mehr HD, Cetin A. Diagnosis of Chronic Kidney Disease Based on Support Vector Machine by Feature Selection Methods. Journal of Medical Systems 2017 February;41(4):1–11. https://doi.org/10.1007/s10916-017-0703-x.

3. Kumar CS, Thangaraju P. Improving Classifier Accuracy for diagnosing Chronic Kidney Disease Using Support Vector Machines. International Journal of Engineering and Advanced Technology 2019 August;8(6):3697–3706. https://doi.org/10.35940/ijeat.f9377.088619.

4. Piao M, Piao Y, Lee J. Symmetrical Uncertainty-Based Feature Subset Generation and Ensemble Learning for Electricity Customer Classification. Symmetry 2019 April;11(4):498–508. https://doi.org/10.3390/sym11040498.

5. Akmam EF, Siswantining T, Soemartojo SM, Sarwinda D. Multiple Imputation with Predictive Mean Matching Method for Numerical Missing Data. In: Proceedings of The 3rd International Conference on Informatics and Computational Sciences (ICICoS) IEEE; 2019. p. 1–6. https://doi.org/10.1109/icicos48119.2019.8982510.

6. Bertsimas D, Orfanoudaki A, Pawlowski C. Imputation of clinical covariates in time series. Journal of Machine Learning Research 2020 November;110(1):185–248. https://doi.org/10.1007/s10994-020-05923-2.

7. Abidin NZ, Ritahani A, A N. Performance analysis of Machine Learning Algorithms for Missing Value Imputation. International Journal of Advanced Computer Science Applications 2018;9(6):442–447. https://doi.org/10.14569/ijacsa.2018.090660.

8. Hartini E. Classification of Missing Values Handling Method During Data Mining: A Review. Sigma Epsilon 2017 February;19(1):11–18. https://doi.org/10.17146/tdm.2017.19.1.3159.

9. Tsai CF, Chen YC. The Optimal Combination of Feature Selection and Data Discretization: An Empirical Study. Information Sciences 2019 dec;505:282–293. https://doi.org/10.1016/j.ins.2019.07.091.

10. Remeseiro B, Bolon-Canedo V. A Review of Feature Selection Methods in Medical Applications. Computers in Biology and Medicine 2019 September;112:103375. https://doi.org/10.1016/j.compbiomed.2019.103375.

11. Moayedikia A, Ong KL, Boo YL, Yeoh WG, Jensen R. Feature Selection for High Dimensional Imbalanced Class Data Using Harmony Search. Engineering Applications of Artificial Intelligence 2017 January;57:38–49. https://doi.org/10.1016/j.engappai.2016.10.008.

12. Arias-Michel R, Garcia-Torres M, Schaerer C, Divina F. Feature Selection Using Approximate Multivariate Markov Blankets. In: Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics Springer International Publishing; 2016.p. 114–125. https://doi.org/10.1007/978-3-319-32034-2_10.

13. Almansour NA, Syed HF, Khayat NR, Altheeb RK, Juri RE, Alhiyafi J, et al. Neural Network and Support Vector Machine for the Prediction of Chronic Kidney Disease: A Comparative Study. Computers in Biology Medicine 2019 jun;109:101–111. https://doi.org/10.1016/j.compbiomed.2019.04.017.

14. Piao Y, Ryu KH. A Hybrid Feature Selection Method Based on Symmetrical Uncertainty and Support Vector Machine for High-Dimensional Data Classification. In: Artificial Intelligence and Lecture Notes in Bioinformatics Springer International Publishing; 2017.p. 721–727. https://doi.org/10.1007/978-3-319-54472-4_67.

15. Amirgaliyev Y, Shamiluulu S, Serek A. Analysis of Chronic Kidney Disease Dataset by Applying Machine Learning Methods. In: Proceedings of The 12th International Conference on Application of Information and Communication Technologies (AICT) IEEE; 2018. p. 120–123. https://doi.org/10.1109/icaict.2018.8747140.

16. Srivastava DK, Bhambhu L. Data Classification Using Support Vector Machine. Journal of Theoretical and Applied Information Technology 2010;12(1):1–7.

17. V RB, Sriraam N, Geetha M. Classification of Non-Chronic and Chronic Kidney Disease Using SVM Neural Networks. International Journal of Engineering Technology 2017 dec;7(1.3):191–194. https://doi.org/10.14419/ijet.v7i1.3.10669.

How to cite this article: Qolby L.S., Buliali J.L., Saikhu A. (2021), Performance Study Of Uncertainty Based Feature Selection Method On Detection Of Chronic Kidney Disease With SVM Classification, *IPTEK The Journal of Technology and Science*, *32(2):103-114*.