Implementation of Sample Sample Bootstrapping for Resampling Pap Smear Single Cell Dataset

Anita Desiani[a1], Sugandi Yahdin[a2], Azhar Kholiq Affandi[b3], Shania Putri Andhini[a4], Yuli Andirani[a5], Muhammad Arhami[c6]

[a]Mathematics Department, Mathematics and Natural Science faculty, Universitas Sriwijaya
Jl. Palembang-Prabumulih Km 32, Inderalaya, Sumatera Selatan, Indonesia
[1]anita_desiani@unsri.ac.id (Corresponding author)
[2]sugandi@unsri.ac.id
[4]shania.andhini@gmail.com
[5]yuliandrianii@unsri.ac.id

[b]Physics Department, Mathematics and Natural Science faculty, Universitas Sriwijaya
Jl. Palembang-Prabumulih Km 32, Inderalaya, Sumatera Selatan, Indonesia
[3]azharka@unsri.ac.id

[c]Informatics Engineering Department, Politeknik Negeri Lhokseumawe
Jl. Medan - Banda Aceh, Kota Lhokseumawe, Aceh, Indonesia
[6]muhammad.arhami@pnl.ac.id

Abstract

The purpose of this study was to determine how the effect of using Bootstrapping Samples for resampling the Harlev dataset in improving the performance of single-cell pap smear classification by dealing with the data imbalance problem. The Harlev dataset used in this study consists of 917 data with 20 attributes. The number of classes on the label had data imbalance in the dataset that affected single-cell pap smear classification performance. The data imbalance in the classification causes machine learning algorithms to produce poor performance in the minority class because they were overwhelmed by the majority class. To overcome it, The resampling data could be used with Sample Bootstrapping. The results of the Sample Bootstrapping were evaluated using the Artificial Neural Network and K-Nearest Neighbors classification methods. The classification used was seven classes and two classes. The classification results using these two methods showed an increase in accuracy, precision, and recall values. The performance improvement reached 10.82% for the two classes classification and 35% for the seven classes classification. It was concluded that Sample Boostrapping was good and robust in improving the classification method.

Keywords: *Pap Smear, Imbalance Data, Sample Boostrapping, Artificial Neural Networks, K-Nearest Neighbor.*

1. Introduction

The imbalance in the classification data causes machine learning algorithms to produce poor performance in the minority class because they are overwhelmed by the majority class [1]. Several studies have addressed data imbalance in several ways. The first is to change the class distribution through various resampling, and the second is to set different priorities by modifying the algorithm structure [2]. The problem of unbalanced data like that often occurs in some machine learning application research [1], [3]–[7]. The problem of data imbalance is related to the accuracy of predictions because predictions are biased towards the majority class, while sometimes, prediction accuracy in the minority class is also required. One solution to overcome the data imbalance is to use the resampling technique [7]. Resampling techniques have become a concern, especially in big data [8]–[11]. There are many ways to increase the accuracy of the

minority class based on the resampling technique because resampling can balance the number of minority classes with the majority class [3], [12], [13].

The Harlev dataset contains single-cell pap smear data, which has seven diagnostic classes. However, the dataset has a problem with an unbalanced number of classes, where the majority class is very far from the minority class. The Superficial Epithelial class has 74 data, the Intermediate Epithelial class has 70 data, the Columnar Epithelial class has 98 data, the Mild Light Dysplasia class has 182 data, the Severe Dysplasia class has 146 data, Moderate Dysplasia class has 197 data, and the last Carcinoma In Situ class has 150 data. The seven classes of labels can be grouped into two normal and abnormal groups. The use of 7 classes shows that the majority class is Moderate Dysplasia and the minority class is Epithelial which has a far data range of 127 data. Unbalanced data affects classification performance, accuracy, precision, and recall because it is difficult to find information on minority classes [14]. Meanwhile, the use of 2 classes also has a disproportionate amount of 242 for the normal category and 675 for the abnormal category.

Several studies have shown that the use of resampling techniques can improve classification performance. One of the methods commonly used in resamples technique is Sample Bootstrapping. The Sample Bootstrapping method has several advantages. The method does not require any assumptions about the distribution of the data. It can resample the sample data up to thousands of times even though the number of samples was limited, and the method has simple calculation [15], [16]. Thanathamathee and Lursinap [12] used the Sample Bootstrapping method to resample data for classification on the Monk2 dataset. This study showed that the resample Sample Bootstrapping technique increased the accuracy value from 82.13% to 85.96%. Research from Al-Luhaybi et al. [17] also used resampling with the Sample Bootstrapping method to classify student datasets at Brunel University. The accuracy increased after the resampling technique was carried out from 75.59% to 93.1%. Several other studies also used sample bootstrapping as a resampling method for classification [18]–[21].

Research on the Herlev dataset was conducted by Kurniawati [22] that research applied SVM to cervical cancer classification with seven classes without using the resampling technique. This study resulted in a low accuracy value of 78.67%. The research from Kusy et al. [23] also classified cervical cancer data using an artificial neural network without resampling and resulted in an accuracy value that was still not good at 71.87%. Several studies on pap smears for detecting cervical cancer disorders used two classes, including Bora et al., study [24], which applied the KNN method. The results indicated that the accuracy, precision, and recall values were excellent above 80%. Likewise, research from Oka et al., [25] uses two classes with the Artificial Neural Network method gave excellent results of 88.8%. From this research, it could be seen that the performance of the classification method was influenced by an imbalance data problem.

The Herlev data set had an unbalanced number of classes. This study focused on applying the resample technique using the Sample Bootstrapping method to classify cervical cancer. The Sample Bootstrapping method was applied to the Herlev pap smear single-cell data to classify the types of cervical cancer disorders. The results of the Sample Bootstrapping application were evaluated by the Artificial Neural Network and K-Nearest Neighbors classification methods to determine the extent to which Sample Bootstrapping was able to improve the performance of the classification method.

2. Research Methods

For the training and testing process, the data was split by 10-fold cross-validation. Algorithm performance was measured based on the accuracy, precision, and recall of ANN and KNN methods.

2.1. Dataset

The dataset used was the Harlev dataset developed by the pathology department of Harlev University Hospital with the Danish Technical University Automation department. This dataset consisted of 917 single cell images, which have been classified into seven classes by cyto-

technicians and specialists [26]. The dataset consists of 20 attributes which were described in Table 1.

**Table 1.** Attribute Description On Harlev Dataset

Attribute	Attribute Description	Data Type
Kerne_A	The area of the nucleus	Real
Cyto_A	The area of the cytoplasm	Real
K/C	Comparison of the area of the nucleus and cytoplasm	Real
Kerne_Ycol	Nuclear light intensity degree	Real
Cyto_Ycol	Cytoplasmic light intensity	Real
KerneShort	The shortest diameter of the nucleus	Real
KerneLong	The longest diameter of the nucleus	Real
KerneElong	Nuclear stretching	Real
Kerne Rund	Nuclear roundabout	Real
CytoShort	The shortest diameter of the cytoplasm	Real
CytoLong	The longest diameter of the cytoplasm	Real
CytoElong	Cytoplasm stretching	Real
CytoRund	Cytoplasmic roundness	Real
KernePeri	Nucleus Limit	Real
KernePos	Nucleus position	Real
KerneMax	Maximum number of nuclear pixels	Integer
KerneMin	Minimum number of pixel nucleus	Integer
Cyto_Max	Maximum number of cytoplasmic pixels	Integer
CytoMin	Maximum number of cytoplasmic pixels	Integer
Class	Diagnosis (Cell type)	Polynomial

As previously mentioned, there were seven classes in this dataset, namely Superficial Epithelial, Intermediate Epithelial, Columnar Epithelial, Mild Light Dysplasia, Severe Dysplasia, Moderate Dysplasia, and Carcinoma In Situ. Cervical cells were grouped into two categories, namely normal and abnormal categories, the types of cells that fall into the normal and abnormal categories could be seen in Table 2 [27].

Table 2. Cell Type in The Harlev Dataset

Categories	Class	Cell Type
Normal	1	Superficial Epithelial
	2	Intermediate Epithelial
	3	Columnar Epithelial
Abnormal	4	Mild Light Dysplasia
	5	Severe Dysplasia
	6	Moderate Dysplasia
	7	Carcinoma In Situ

From Table 1, the normal category has three types of cell types, and the abnormal category has four types of cell types. Although some cells have the same category, each cell type has a different cell shape. Cell shapes for each category can be seen in Figures 1 and 2.

a   b   C
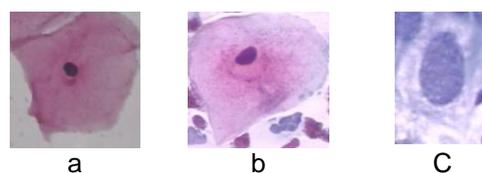
Figure 1. Images of Cervical Cells in Normal Category (a)Superficial Epithelial (b) Intermediate Epithelial (c) Columnar Epithelial

Figure 1 showed the shape of cells that fall into the normal category, which consisted of 3 classes. Then Figure 2 showed the shape of cells that fall into the abnormal category, which consisted of 4 classes.
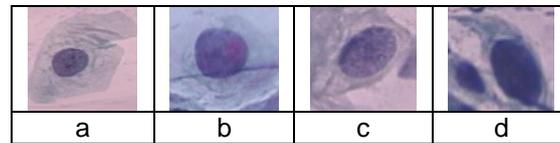
Figure 2. Images of Cervical Cells in Abnormal Category (a) Mild Light Dyplasia (b) Severe Dysplasia (c) Moderate Dysplasia (d) Carcinoma In Situ.

2.2. Implementation of Sample Sample Bootstrapping

The implementation of Sample Sample Boostrapping (SB) was in the pre-processing stage, which occurred before the data entered the classification process. Sample Bootstrapping was a method used to estimate the deviation from the standard error [28]. Sample Boostrapping used statistical procedures by changing the data from the existing sample and replicating the sample data (resampling) randomly to get new simulation data. The Sample Boostrapping took samples with the replacement method, which replaced the original data randomly with a specific label. The data in the process had an equal chance of being selected. The data could be re-selected in the following process [29]. Based on several studies, the advantages of Sample Boostrapping were the ability to study any statistic of interest and handle sampling error by creating a specific model [30]. The working steps in Sample Boostrapping could not reduce data errors but only estimated standard errors in the data [20]. The steps on the Sample Boostrapping method were [31]:

a. Construct a distribution of n Sample Bootstrapping sample ($\hat{F}$) by assigning a probability of 1/n to each data (Xi) for i=1,2,3,…,n.
b. Take a Sample Bootstrapping sample of size n at random with the return of the distribution of stage 1.
c. Choose the replication of each sample ($\hat{\theta}$) statistic from the Sample Bootstrapping sample was referred to as $\hat{\theta}_1^*$.
d. Repeat steps 2 and 3 until B times, so you got $\hat{\theta}_1^*, \hat{\theta}_2^*, …, \hat{\theta}_B^*$.
e. Estimate the standard error (se$_B$) using the standard deviation B times with Equation 1.

$$se_B = \left\{ \sum_{b=1}^{B} [\ \hat{\theta}*(b) - \ \hat{\theta}*(.)]^2 \right\}^{1/2} \qquad (1)$$

Where

$$\hat{\theta}P^{*p}(.) = \sum_{b=1}^{B} \hat{\theta}*(b)/B \qquad (2)$$

2.3. Evaluation of Sample Bootstrapping Using Classification Methods

To evaluate the performance of Sample Boostrapping samples, it would be implemented to classification methods and analyzed the performance result of the classification. Classification is the process of forming a model to predict an unknown class pattern [32]. In this study, the methods used in the classification for evaluation were the ANN and KNN method.

2.3.1. Artificial Neural Networks

Artificial Neural Networks (ANN) was a method that could process large data [33]. The ANN consisted of several layers, namely the input layer, hidden layer, and output layer. Hidden layers used were one hidden layer with learning rates of 0.01 and training cycles of 200. The form of the ANN architecture used could be seen in Figure 3.
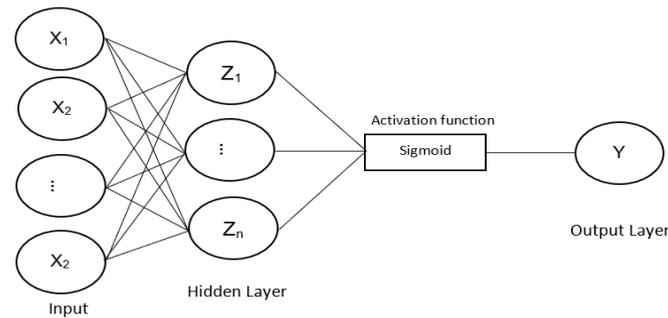
Figure 3. ANN Architecture on Pap smear Single Cell Classification

This study used ANN Backpropagation. This algorithm was one of the most frequently used algorithms [34]. The stages of the ANN Backpropagation method were [35]:

a. Initialize all weights with a small random number.
b. Enter to feedforward stage.
c. Forward Input Data (i = 1,2,…,n) from input layer to the hidden layer.
d. Calculate the forwarded data (z_j) on the hidden layer with Equation 3.

$$z\_in_j = v_{0j} + \sum_{i=1}^{n} v_{ij} x_i \tag{3}$$

Where z_in_j was the j-th data input (j=1,2,3,…,n) in the hidden layer, v_{0j} was the weight value for bias for unit z_j and v_{ij} is the weight value for unit x_i. Then the data that came out of the hidden layer to the output layer was calculated by Equation 4.

$$z_j = f(z\_in_j) \tag{4}$$

where s was the activation function used in the hidden layer. After all the data was calculated, then proceed to the next layer.

e. Calculate the forwarded data on the output layer with Equation 5.

$$y\_in_k = w_{0k} + \sum_{j=1}^{m} w_{jk} z_j \tag{5}$$

where y_in_k Was the k-th data input at the output layer d, and w_{0k} was the weight for bias to the output unit.

f. Calculate the data that comes out as output with Equation 6.

$$y_k = f(y\_in_k) \tag{6}$$

g. Prepare for the backpropagation stage.
h. Calculate all output data (y_k, k=1,2,3,…,m) for all target patterns. Calculate the factor error (δ) with Equation 6.

$$\delta_k = (t_k - y_k) f'(y\_in_k) \tag{7}$$

Where δ_k was the error used when the layer weight changes, t_k was the output target. Next, update the weight value of w_{jk} by calculating the change in weight using acceleration α using Equation 8.

$$\Delta w_{jk} = \alpha \delta_k z_j \tag{8}$$

Update the value of the bias b by calculating the value of the change in bias using Equation 9.

$$\Delta w_{k0} = \alpha \delta_k \tag{9}$$

Then the calculated value was sent to the previous layer.

i. Calculate each input data from the output layer with Equation (10).

$$\delta\_in_j = \sum_{k=1}^{m} \delta_k w_{jk} \tag{10}$$

The input data that has been obtained would be multiplied by the inverse function of the activation function using equation 11.

$$\delta_j = \delta_{in_j} f' \left(z_{in_j} \right) \tag{11}$$

j. Calculate the weight changes with equation 12 and the value of the change in bias to update the weight and bias values in the hidden layer.

$$\Delta v_{ji} = \alpha \delta_j x_i \tag{12}$$
$$\Delta v_{j0} = \alpha \delta_j \tag{13}$$

k. Update the bias and weight values on the output layer with equation 14 and the hidden layer with equation 15.

$$w_{jk}(new) = w_{jk}(previous) + \Delta w_{kj} \tag{14}$$
$$v_{ji}(new) = v_{ji}(previous) + \Delta v_{ji} \tag{15}$$

l. Repeat Steps a-i for each training data.
m. Perform steps a-j for each iteration.

2.3.2. K-Nearest Neighbors

The second method used for classification was K-Nearest Neighbors (KNN). KNN was an algorithm that worked on the shortest distance from the query instance to the training sample. The goal was to classify an object based on attributes and training samples. For this research, the value of k used was k=5. The following are the steps in the KNN algorithm [36]:

a. Determine the parameter k to be used.
b. Calculate the distance between the new data and all training data using Euclidean Distance with Equation 16 [32].

$$d_i = \sqrt{\sum_{i=1}^{n}(x_{2i} - x_{1i})^2} \tag{16}$$

Where d_i was defined as the distance between x_1 and x_2, x_1 was the sample data and x_2 was the test data, i was the data variable.

c. Sort the distance calculation results from the smallest to the largest and determine the nearest neighbor based on the kth minimum distance.
d. Claim the class wher it was taken based on the highest number of class.

2.4. Algorithm Performance Assessment

Algorithm performance assessment was based on the confusion matrix that appeared after the training and testing process. There were two classes in the confusion matrix, namely positive class and negative class. The true positive (TP) was a positive class that was guessed correctly. The False positive (FP) was a positive class that was guessed wrong. As well as true negative (TN) was a negative class that was guessed right, and The False Negative (FN)was a negative class that was guessed wrong. If the case had more than two classes, one class became a positive class, and the rest became a negative one.

The confusion matrix calculated the accuracy value used to measure the accuracy of the classification results. In addition, from the confusion matrix, the precision value used to calculate the accuracy of the prediction results against the requested information could also be calculated and the recall value used to calculate the ratio of the selected relevant items to the actual value[37].

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \text{ x } 100\% \tag{17}$$

$$\text{Precision} = \frac{TP}{(FP+TP)} \text{ x } 100\% \tag{18}$$

$$\text{Recall} = \frac{TP}{(TP+FN)} \text{ x } 100\% \tag{19}$$

3. Result and Discussion

In this study, sample Sample Bootstrapping was used for resampling the single-cell pap smear dataset before classification using ANN and KNN. The steps taken were the research method described above, preprocessing using the SB method for sampling. Then, the classification method used was ANN and KNN. in the sampling process, the parameters used include relative or ration (0-1). Furthermore, the resampled data was validated using n-fold cross-validation. The value of n used was 10-fold. The data set was divided into ten partitions, nine partitions as training data, and one partition for test data. This process was repeated ten times for each section so that every part of the ten sections had become testing data. These stages could be seen in the form of a flowchart which could be seen in Figure 4.

A comparative test of results was carried out in this study to find out the result of applying the Sample Bootstrapping method. The results were compared by comparing the test results on the original data (which does not use Sample Bootstrapping) with the test results obtained in the Sample Bootstrapping implementation. This test was carried out on a dataset group with two classes and a group of 7 classes. In the classification of 7 classes, the data were grouped into seven classes based on all types of pap smear cell images in table 2. in the classification of 2 classes, it grouped from 7 classes into two categories normal and abnormal in table 2.
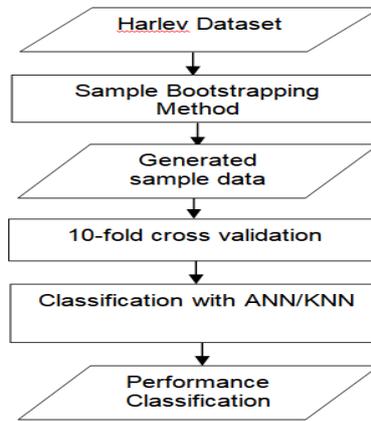

Figure 4. The Flowchart on The Proposed Method

3.1. Classification without Sample Sample Bootstrapping

The classification method was used to analyze the result of sample Sample Bootstrapping(SB) was implemented. To be fair, all parameters in both methods used the same parameters when using sample Sample Boostrapping and without it. The number of data before using SB could be seen in Figure 5.
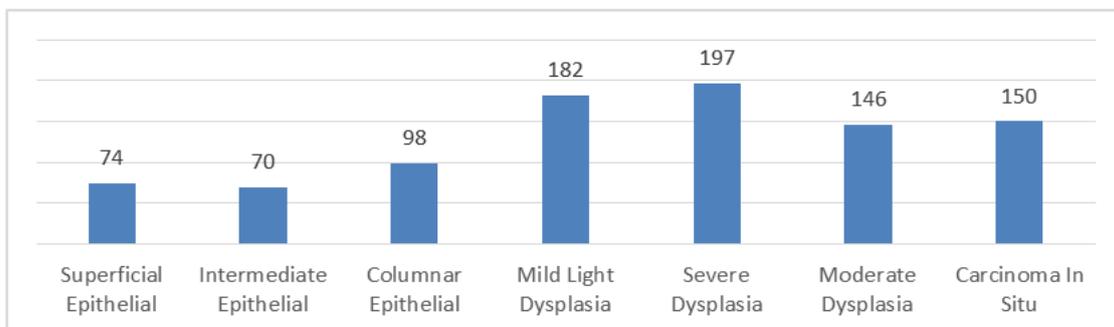

Figure 5. The Number of Data by Cell Type on Original Single Cell Pap Smear Dataset

The number of data by categories on pap smear data set on the normal class had 242 data, and the abnormal class had 675 data. By 10-fold cross-validation, the classification results using the original dataset or without using the SB could be seen in Table 4.

Table 4. Classification Result Without Sample Sample Boostrapping (%)

Classification	ANN			KNN		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall
2 classes	93,90	92,83	92,16	91,28	90,76	86,72
7 classes	62,70	67,74	67,43	50,59	53,90	52,87

From Table 5, it could be seen that the 7 class classifications had smaller accuracy, precision, and recall values. It could have happened because the data spread in 7 classes was smaller than two classes.

3.2. Classification with Sample Bootstrapping

The resampled dataset had the same size as the original dataset. In this study, the number of samples was determined, namely relative to the ratio 1. Resampling with SB was done five times. The number of each class after passing the SB process could be seen in Figure 6.
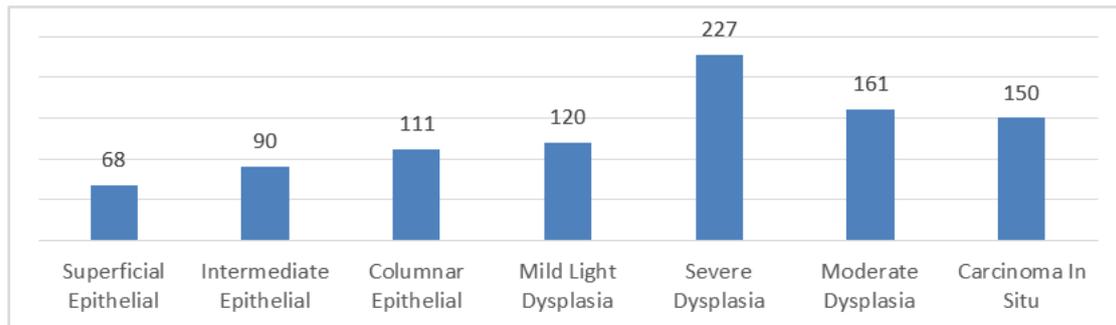
Figure 6. Amount of Data by Cell Type After Five Times Resampling with Sample Sample Bootstrapping

The results of the Sample Boostrapping sample performance on the ANN and KNN methods with two types of classification could be seen in Table 5. Based on Figure 6, it could be seen that the minority class experienced an increase in the amount of data so that the number between classes was not too far apart even though the majority class still had a wide range of values. The number of data by categories became 269 for normal class and 648 for abnormal class.

Table 5. Classification Result Using Sample Sample Boostrapping (%)

Classification	ANN			KNN		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall
2 classes	96,61	96,52	95,52	97,49	96,63	97,54
7 classes	81,13	84,87	82,42	86,49	87,77	87,54

Based on Table 5, classification in 7 classes still produced lower performance scores than two classes, but the results were much better than before. A comparison of results on the use of SB was discussed further in the next section.

3.3. Comparison of Results

To see more clearly, the comparison on the use of SB was be divided into 2 Tables. There was a Table for comparison of 2 classes and the other for a comparison of 7 classes. For comparison in the two classes could be seen in Table 6 below.

Table 6. Comparison of the results of the Single Cell Pap Smear classification in 2 classes (%)

Classification Methods	ANN			KNN		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall
Without SB	93,90	92,83	92,16	91,28	90,76	86,72
With SB	96,61	96,52	95,52	97,49	96,63	97,54
Difference	2,71	3,69	3,36	6,21	5,87	10,82

The SB method of classification could increase the accuracy value, especially the recall value, because the higher the recall value, the better the machine in finding information about a class. The recall value increased due to the machine recognizing the minority class, which was previously biased towards the majority class. Based on the classification method used, the SB

method worked well on KNN compared to ANN. It could be seen from the value of the increase that occurred in KNN reaching more than 5% in all performance values. In Table 6, it could be seen in the classification of 2 classes. Furthermore, to see the comparison of the seven classes could be seen in Table 7.

Table 7. The Comparison of the results of the Single Cell Pap Smear classification in 7 classes (%)

Classification Methods	ANN			KNN		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall
Without SB	62,70	67,74	67,43	50,59	53,90	52,87
With SB	81,13	84,87	82,42	86,49	87,77	87,54
Difference	18,43	17,44	14,99	35,9	33,87	34,67

In Table 7, the SB method was much better in the 7 class classification because it could be seen from the difference, which was quite far up to 35%. Significant accuracy, precision, and recall values indicate that resampling using SB greatly improved classification results on unbalanced data. Although the amount of data generated (Figure 6) was still not very balanced, it has made an excellent classification. Based on the method used, KNN also had a better performance value than ANN. It could be seen that the highest difference was in the KNN method.

Although the performance in 7 classes has increased, the numbers produced were not as good as the 2 class classification. However, the SB method worked very well on the KNN method on both classifications because it has increased the accuracy value considerably. This showed that SB was very good at improving the performance of the KNN method for seven classes. To analyze the results of this study further, a comparison of the results with previous studies was carried out. The Comparison of research results for single-cell pap smear classification could be seen in Table 8.

Table 8. The Comparison of Results on Classification of Single Cell Pap Smear

Author / Dataset / Methods	The Increasing Accuracy	The Increasing Presicion	The Increasing Recall
Zughrat et. Al (2014)/Rail Dataset/SB-SVM [1]	18,2%	-	-
Thanathamathee and Lursinsap(2013)/ Monk2/SB-ANN [12]	8,98%	-	-
Thanathamathee and Lursinsap(2013)/ Abalone/SB-ANN [12]	3,14%	-	-
Saez et. Al (2015)/Abalone/Smote-C.45 [38]	3%	-	-
Arifin and Rachman(2020)/Harlev Dataset 2 Classes/DecisionTree-PSO [39]	5,37%	-	-
Proposed Method/Harlev Dataset 2 Classes/SB-ANN	2,71%	3,69%	3,36%
Proposed Method/Harlev Dataset 2 Classes/SB-KNN	6,21%	5,87%	10,82%
Proposed Method/Harlev Dataset 7 Classes/SB-ANN	18,43%	17,44%	14,99%
Proposed Method/Harlev Dataset 7 Classes/SB-KNN	35,9%	33,87%	34,67%

In Table 8, it could be seen that several studies used resampling techniques for unbalanced data. In other studies, it only showed differences in accuracy values as it was known that the accuracy value was not enough to determine an algorithm works well. If there is so much the number of majority classes, the machine could only predict the majority class. In contrast, a good accuracy value could occur because the number of minorities that could not be predicted is few. From Table 8, it could also be seen that only the proposed method could display other performance values

and was the advantage of this research. In addition, this study had the highest improvement value compared to the others. Although the classification of 2 classes using the ANN method was still lower than Arifin and Rachman's [39] research, this study had an increase in other performance values not shown in that study. From this comparison, it could be concluded that SB was very good and robust in improving the classification method.

4. Conclusion

The sample Sample Bootstrapping method was very good and robust for resampling on an imbalanced data problem. That is indicated by the improved classification performance of the study. The highest increase occurred using the KNN method in the classification of 2 classes and seven classes. The highest difference value is the KNN method on seven classes classification, with an increasing value is 35.9% for accuracy, 33.87% for precision, and 34.67% for recall. With a significant increase, it can be concluded that sample Sample Bootstrapping can improve the classification of labels that have many classes.

References

[1]   A. Zughrat, M. Mahfouf, Y. Y. Yang, and S. Thornton, "Support Vector Machines for Class Imbalance Rail Data Classification with Sample Boostrapping- based Over-Sampling and Under-Sampling, " *IFAC Proceedings Volumes*, vol. 47, no. 3. IFAC, 2014.

[2]   A. Tharwat, T. Gabel, and T. Gabel, "Parameters optimization of support vector machines for imbalanced data using social ski driver algorithm," *Neural Computing and Applications*, vol. 0123456789, 2019, doi: 10.1007/s00521-019-04159-z.

[3]   R. Ghorbani and R. Ghousi, "Comparing Different Resampling Methods in Predicting Students ' Performance Using Machine Learning Techniques," *IEEE Access*, vol. 8, pp. 67899–67911, 2020, doi: 10.1109/ACCESS.2020.2986809.

[4]   J. . Sanz, D. Bernardo, F. Herrera, H. Bustince, and H. Hagras, "A compact evolutionary interval-valued fuzzy rule-based classification system for the modeling and prediction of real-world financial applications with imbalanced data," *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 4, pp. 973–990, 2015.

[5]   W. Wei, J. Li, L. Cao, Y. Ou, and J. Chen, "Effective detection of sophisticated online banking fraud on extremely imbalanced data," *World Wide Web*, vol. 16, no. 4, pp. 449–475, 2013.

[6]   H. Yu, J. Ni, and J. Zhao, "ACOSampling: An ant Colony Optimizaiton-Based Undersampling Method for Classifying Imbalanced DNA Microarray Data," *Neurocomputing*, vol. 101, pp. 309–318, 2013.

[7]   T. Sasada, Z. Liu, T. Baba, K. Hatano, and Y. Kimura, "A resampling method for imbalanced datasets considering noise and overlap," *Procedia Computer Science*, vol. 176, pp. 420–429, 2020, doi: 10.1016/j.procs.2020.08.043.

[8]   I. Triguero, S. del Rio, V. L´opez, J. Bacardit, J. . Ben´ıtez, and F. Herrera, "ROSEFW-RF: The winner algorithm for the ECBDL14 big data competition. An extremely imbalanced big data bioinformatics problem," *Knowledge-Based Systems*, vol. 87, pp. 69–79, 2015.

[9]   M. Koziarski and M. Wozniak, "CCR: A combined cleaning and resampling algorithm for imbalanced data classification," *International Journal of Applied Mathematics and Computer Science*, vol. 27, no. 4, pp. 727–736, 2017, doi: 10.1515/amcs-2017-0050.

[10] T. . Hoens, R. Polikar, and N. . Chawla, "Learning from streaming data with concept drift and imbalance: An overview," *Progress in Artificial Intelligence*, vol. 1, no. 1, pp. 89–101, 2012.

[11] F. Fernandez-Navarro, C. Hervas-Martinez, and P. . Gutierrez, "A dynamic over-sampling procedure based on sensitivity for multi-class problems," *Pattern Recognition*, vol. 44, no. 8, pp. 1821–1833, 2011.

[12] P. Thanathamathee and C. Lursinsap, "Handling Imbalanced Data Sets with Synthetic Boundary Data Generation Using Sample Bootstrapping Re-sampling and AdaBoost Techniques," *Pattern Recognition Letters*, vol. 34, no. 12, pp. 1339–1347, 2013, doi:

10.1016/j.patrec.2013.04.019.

[13] A. Elhassan, M. Aljourf, F. Al-Mohanna, and M. Shoukri, "Classification of Imbalance Data using Tomek Link (T-Link) Combined with Random Under-sampling (RUS) as a Data Reduction Method Technology & Optimization," *Global Journal of Technology & Optimization*, vol. 1, no. 111, pp. 1–11, 2017, doi: 10.4172/2229-8711.S1.

[14] A. Desiani, S. Yahdin, and A. Kartikasari, "Handling the imbalanced data with missing value elimination SMOTE in the classification of the relevance education background with graduates employment," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 2, pp. 346–354, 2021, doi: 10.11591/ijai.v10.i2.pp346-354.

[15] I. Ivanov, "Tenfold Boostrap Procedure for Support Vector Machine," *Computer Science*, vol. 21, no. 2, pp. 253–268, 2020.

[16] I. Rodliyah, "Perbandingan Metode Sample Bootstrapping Dan Jackknife (Comparison of Sample Bootstrapping and Jackknife Methods To," *Jurnal Matematika dan Pendidikan Matematika*, vol. 1, no. 1, pp. 76–86, 2016.

[17] M. Al-Luhaybi, L. Yousefi, S. Swift, S. Counsell, and A. Tucker, *Predicting academic performance: A Sample Boostrapping approach for learning dynamic bayesian networks*, vol. 11625 LNAI. Springer International Publishing, 2019.

[18] T. Agus, S. M. Adib, and A. Karomi, "Penerapan Metode Sample Sample Boostrapping untuk Meningkatkan Performa kNearest Neighbor pada Dataset Berdimensi Tinggi," *IC-Tech,* vol. XII, no. 1, pp. 9–14, 2017.

[19] T. A. Setiawan, R. Satria, and A. Syukur, "Integrasi Metode Sample Sample Boostrapping dan Weighted Principal Component Analysis untuk Meningkatkan Performa k Nearest Neighbor pada Dataset Besar," *Journal of Intelligent Systems*, vol. 1, no. 2, pp. 76–81, 2015.

[20] E. Siswanto, Suprapedi, and Purwanto, "Metode Sample Boostraping Pada K-Nearest Neighbor Untuk Klasifikasi Status Desa," *Jurnal Teknologi Informasi*, vol. 14, no. 1, pp. 13–23, 2018.

[21] E. Jumiati and M. R. Kamal, "Integrasi Sample Sample Boostrapping Pada K-Nearest Neighbor untuk Klasifikasi Herregistrasi Calon Mahasiswa Baru," *IC-Tech*, vol. 12, no. 1, pp. 23–32, 2017.

[22] Y. E. Kurniawati, A. E. Permanasari, and S. Fauziati, "Comparative study on data mining classification methods for cervical cancer prediction using pap smear results," *Proceeding 2016 1st International Conference on Biomedical Engineering (IBIOMED) 2016*, 2017, doi: 10.1109/IBIOMED.2016.7869827.

[23] M. Kusy, B. Obrzut, and J. Kluska, "Application of gene expression programming and neural networks to predict adverse events of radical hysterectomy in cervical cancer patients," *Medical & Biological Engineering & Computing*, vol. 51, no. 12, pp. 1357–1365, 2013, doi: 10.1007/s11517-013-1108-8.

[24] K. Bora, M. Chowdhury, and L. B. Mahanta, "Automated classification of Pap smear images to detect cervical dysplasia," *Comput Methods Programs Biomed*, vol. 138, pp. 31–47, 2017, doi: 10.1016/j.cmpb.2016.10.001.

[25] N. P. A. Wiastini Oka, I. K. G. Darma Putra, and K. S. Wibawa, "Klasifikasi Sel Nukleus Pap Smear Menggunakan Metode Backpropagation Neural Network," *Jurnal Ilmiah Merpati*, vol. 7, no. 3, pp. 182–192, 2019.

[26] D. Riana, D. H. Widyantoro, T. Latifah, and R. Mengko, "Ekstraksi dan Klasifikasi Tekstur Citra Sel Nukleus Pap Smear," *Jurnal TICOM*, vol. 1, no. 3, pp. 62–70, 2013.

[27] Y. Ramdhani and D. Riana, "Hierarchical Decision Approach Based on Neural Network and Genetic Algorithm Method for Single Image Classification of Pap Smear," *Second International Conference on Informatics and Computing (ICIC)*, pp. 1–6, 2017, [Online]. Available: doi: 10.1109/IAC.2017.8280587.

[28] R. E. McRoberts, S. Magnussen, E. O. Tomppo, and G. Chirici, "Parametric, Sample Bootstrapping, and jackknife variance estimators for the k-Nearest Neighbors technique with

illustrations using forest inventory and satellite image data," *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3165–3174, 2011, doi: 10.1016/j.rse.2011.07.002.

[29] T. Siswanto, "Optimalisasi Sosial Media Sebagai Media Pemasaran Usaha Kecil Menengah," *Liquidity*, vol. 2, no. 1, pp. 80–86, 2018, doi: 10.32546/lq.v2i1.134.

[30] L. R. Zientek and B. Thompson, "Applying the Sample Bootstrapping to the multivariate case : Sample Bootstrapping component / factor analysis," *Behavior Research Methods*, vol. 39, no. 2, pp. 318–325, 2007.

[31] H. lin Shang, "Resampling Techniques for Estimating the Distribution of Descriptive Statistics of Functional Data," *Communication in Statistics-Simulation and Computation*, vol. 44, no. 3, pp. 614–635, 2015, [Online]. Available: doi: 10.1080/03610918.2013.788703.

[32] N. L. W. S. R. Ginantra, "Deteksi Batik Parang Menggunakan Fitur Co-Occurence Matrix Dan Geometric Moment Invariant Dengan Klasifikasi KNN," *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, vol. 7, no. 1, p. 40, 2016, doi: 10.24843/lkjiti.2016.v07.i01.p05.

[33] M. Hasanipanah, M. Noorian-Bidgoli, D. Jahed Armaghani, and H. Khamesi, "Feasibility of PSO-ANN model for predicting surface settlement caused by tunneling," *Engineering with Computers*, vol. 32, no. 4, pp. 705–715, 2016, doi: 10.1007/s00366-016-0447-0.

[34] D. Kristianto, C. Fatichah, B. Amaliah, and K. Sambodho, "Prediction of Wave-induced Liquefaction using Artificial Neural Network and Wide Genetic Algorithm," *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, vol. 8, no. 1, p. 1, 2017, doi: 10.24843/lkjiti.2017.v08.i01.p01.

[35] D. Graupe, *Principles of Artificial Neural Networks (2nd Edition)*, vol. 53, no. 9. University of Illinois, Chicago,USA, 2007.

[36] R. Apurb, S. Milan, A. Avi, and R. Dundigalla, "Heart disease prediction using machine learning classifiers," *International Journal of Advanced Science and Technology*, vol. 29, no. 6, pp. 1700–1707, 2020, doi: 10.37200/IJPR/V24I6/PR260661.

[37] S. Yahdin, A. Desiani, N. Gofar, K. Agustin, and D. Rodiah, "Application of the Relief-f Algorithm for Feature Selection in the Prediction of the Relevance Education Background with the Graduate Employment of the Universitas Sriwijaya," *Computer Engineering and Applications (ComEngApp)*, vol. 10, no. 2, pp. 71–80, 2021.

[38] J. A. Saez, J. Luengo, J. Stefanowski, and F. Herrera, "SMOTE–IPF: Addressing the noisy and borderline examples problem in imbalanced classification by a re-sampling method with filtering," *Information Sciences*, vol. 291, pp. 184–203, 2015.

[39] T. Arifin and R. Rachman, "Optimasi Decision Tree Menggunakan Particle Swarm Optimization Untuk Klasifikasi Sel Pap Smear," *(JATISI) Jurnal Teknik Informatika dan Sistem Informasi*, vol. 7, no. 3, pp. 572–579, 2020.