November-December 2019
ISSN: 0193-4120 Page No. 2246 - 2254

A Study on Improved Outlier Detection and Prediction Based on Hybrid Machine Learning

Eun Soo Choi [1], Min Soo Kang [2*]
[1]Department of Medical IT Marketing, Eulji University, Korea, ces951030@gmail.com
[2]Department of Medical IT, Eulji University, Korea, mskang@eulji.ac.kr (*corresponding author)

Article Info
Volume 81
Page Number: 2246 - 2254
Publication Issue:
November-December 2019

Article History
Article Received: 5 March 2019
Revised: 18 May 2019
Accepted: 24 September 2019
Publication: 12 December 2019

Abstract
Data collection and preprocessing are important processes for training prediction models. If researcher cannot find outliers of wrongly collected, the noise data will be learned in the prediction model learning as well. Therefore, it can be considered that the removal of the outliers is essential for preprocessing process in order to learn prediction model more accurately.

In this paper, proposing hybrid machine learning for outlier detection and prediction by applying two algorithms. Two medical data were selected for the experiment, and the information gain was calculated before applying the DBSCAN, and two attributes with the highest relevance to the label value were extracted. DBSCAN parameters were selected based on extracted attributes. First, in the outlier detection process, in this paper, experiment was implemented with the proposed algorithm using the characteristics of DBSCAN, which is a density-based algorithm, and preprocessed the learning data three times. Second, the preprocessed data was evaluated by Neural Network and Boosted Decision Tree. Experimental results show that model accuracy of preprocessed data is similar or better than model accuracy of raw data. Applying the hybrid model proposed in this paper, it is expected higher accuracy and generalization of outliers and frequent medical data.

***Keywords**: Outlier Detection, Data Preprocessing, Machine Learning, Hybrid Model, DBSCAN*

1. INTRODUCTION

In recent years, the exponential growth of information technology has led to an increase in the number of hospitals that collect medical data, and accordingly the volume of medical data [1]. As the amount of medical data increases, there is a lot of information in the medical field, but the knowledge using information is still insufficient [2]. Therefore, more and more people are using artificial intelligence prediction models to gain knowledge by utilizing medical data. Data collection and preprocessing before training prediction model is an important process that controls the performance of the prediction model. In particular, if you have to manually enter data in the data collection process, the data may be entered incorrectly. Furthermore, the data may be glitch data. It is mainly caused by medical area, which collects and inputs data directly. Medical databases produce a lot of information, and the data types of information are various [3]. Medical database has many data types and attributes such as age, sex, x-ray, numerical, nominal and image, etc. [4]. The medical database has a large number of data types and attributes, causing errors in the data collection process. In particular, outlier detection is very important in medical research because medical data is large information in each case [5]. If it cannot be find outliers of wrongly collected data in the data collection process, the noise data will

be used in the prediction model training as well. Therefore, the removal of outliers can be regarded as an essential element of the data preprocessing process in order to learn the prediction model more accurately.

In this paper, proposed an improved outlier detection algorithm by utilizing the density-based characteristic of DBSCAN algorithm, and preprocessing was done using medical data. In addition, Two-Class Neural Network algorithm and Two-Class Boosted Decision Tree algorithm were applied to evaluate preprocessed data.

2. BACK GROUND

2.1 Information Gain

Feature Selection is a method for efficiently learning various machine learning and data mining. It is a method to select the feature that best represents the dataset among high-dimensional datasets [6]. Information Gain is one of the feature selection methods, Information Theoretical based method [7]. Information gain is also called Mutual Information Maximization (MIM) [8].

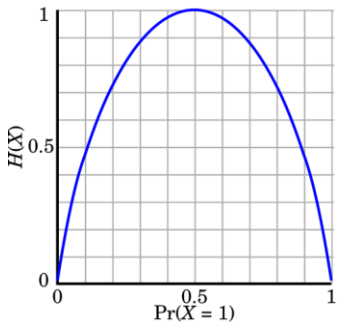


Figure 1 : Binary entropy function

Figure 1 shows the graph of the entropy function relative to binary classification. Entropy value is between 0 to 1. Entropy is impurity and purity in an arbitrary collection of examples. The larger the information gain value, the more likely the feature is describing the label [9].

2.2 DBSCAN

DBSCAN (Density Based Spatial Clustering of Applications with Noised) is a density-based clustering algorithm that is suitable for handling spatial data including noise and can distinguish arbitrary shape and size clusters. The goal of the DBSCAN algorithm is to find high density regions and make them into the same cluster. At this time, the cluster and noise are intuitively distinguished based on the density of the points.

Algorithm 1 DBSCAN Algoritm
Input: dataset D, minimum point $minPts$, radius ϵ
Output: clustered data, outlier data
for *each point P in dataset D* **do**
  if *P is not visited* **then**
    mark P as visited
    $neigbrPts \leftarrow$ points in ϵ-*neigborhood* of P
    if $sizeof(neigbrPts) < minPts$ **then**
      mark P as outlier
    end
    else
      add P to new cluster C
      for *each ponit P' in $neigbrPts$* **do**
        if *P' is not visited* **then**
          mark P' as visited
          $neigbrPts' \leftarrow$ points in ϵ-*neigborhood* of P'
          if $sizeof(neigbrPts') \geq minPts$ **then**
            $neigbrPts \leftarrow neigbrPts + neigbrPts'$
          end
        end
        if *P' is not a member of any cluster* **then**
          add P' to cluster C
        end
      end
    end
  end
end

Figure 2 : DBSCAN Algorithm

Figure 2 shows DBSCAN Algorithm. The parameter eps defined the radius of neighborhood around a point x and parameter minPts is the minimum number of neighbors within the eps radius [10].

3. LITERATURE REVIEW

3.1 Hybrid Machine Learning

Table 1 summarized the hybrid machine learning research. Kim implemented the hybrid machine learning by applying Deep Neural Network algorithm after applying Cardiovascular Disease Diagnosis Prediction Data to SVM, Gaussian, Naive Bayes, DNN, Decision Tree and Cardio-DNN algorithm [11]. K-fold cross-validation was applied to verify the algorithm. As a result, F1_Score value is 0.981, and Hybrid Prediction Model is successfully implemented.

Table 1 : Hybrid machine learning research

Author	Application domain	Hybrid techniques	Evaluation method
Kim[11]	Cardiovascular Disease Diagnosis Prediction	Classification, Classification	F1 Score
Kim[12]	Ransomware Detection	Clustering, Classification	Accuracy Rate
Bala et al. [13]	Satellite data Pattern Recognition	Genetic Algorithm, Classification	Error rates
Tsai and Chen [14]	Credit rating Prediction	Clustering, Classification (Combination)	Accuracy rates
Ijaz et al. [15]	Hypertension Prediction	Clustering, Classification	Accuracy rates

Kim attempted to categorize Ransomware files and normal files in the first place through K-means algorithm, and applied logistic remediation using clustering information, Ransomware API information, and File System Activity's log monitoring information [12]. As a result of the application, the Accuracy Rate is 88.49%, which can be verified that the system successfully detected the Ransomware with the Hybrid Prediction Model.

Bala et al. proposed a hybrid machine learning algorithm using the Genetic Algorithm and ID3 algorithm [13]. Satellite and facial image data were preprocessed using Genetic Algorithm and classified through ID3 Algorithm. As a result of experiment, the proposed algorithm decreased error rates and decreased complexity.

Tsai and Chen studied which combination of classification and clustering algorithms was best for applying Hybrid Machine Learning [14]. The algorithms applied in the experiment are Linear Regression, Extension maximization, Neural Network. As a result, the results of hybridizing the Classification and Classification algorithms showed the best Accuracy with 83.44%.

Ijaz et al. was preprocessed through the DBSCAN clustering algorithm for the presentation of Hypertension and classified as the Random Forest algorithm [15]. As a result, it was confirmed that Hypertension was successfully predicted to 92.55%.

3.2 DBSCAN for Data Preprocessing

Table 2 : DBSCAN for data preprocessing research

Author	Preprocessing Target	Evaluation method
Ijaz et al. [15]	Hypertension Data	Accuracy rates
Abid et al. [16]	Wireless sensor networks Data	Accuracy rates
Alfian et al. [17]	Network communication glitch Data	3-Sigma Rule
ElBarawy et al. [18]	Social Network Data	Cluster statement
Tian et al. [19]	Time series Sensing Data	Error Rate

Ijaz et al. extracted two features by applying Feature Selection using Information Gain to preprocess Hypertension Data. The outlier was removed by using two features of DBSCAN, and it was confirmed that the result was successfully preprocessed to 92.55%. Abid et al. has applied the DBSCAN algorithm to detect normal sensing data in wireless sensor network data, and it can be confirmed that it was successfully preprocessed through the result of accuracy of 99% [16].

Alfian et al. has applied the DBSCAN algorithm to catch the network communication glitch and has been successfully preprocessed [17]. It is an example showing that preprocessing is possible through DBSCAN even for normal sensor data.

ElBarawy et al. removed the outlier by using DBSCAN, and then applied the clustering techniques to discover the community [18]. As a result of the application, the number of clusters

November-December 2019
ISSN: 0193-4120 Page No. 2246 - 2254

became smaller and the density became higher, and it was confirmed that the preprocessing was successful.

Tian et al. applied DBSCAN to preprocess Time Series Sensing Data and verified it according to 3-sigma rule [19].

4. EXPERIMENT

4.1 Proposed Outlier Detection Algorithm

Algorithm 2 Proposed outlier detection algoritm
Input: dataset D, number of preprocessing NP, minimum point $minPts$, radius ϵ
Output: clustered data, outlier data
for *each point* P *in dataset* D **do**
  if P *is not visited* **then**
    mark P as visited
    $neigbrPts \leftarrow$ points in ϵ-*neigborhood* of P
    if $sizeof(neigbrPts) < minPts$ **then**
      mark P as outlier
    end
    else
      add P to new cluster C
      for *each ponit* P' *in* $neigbrPts$ **do**
        if P' *is not visited* **then**
          mark P' as visited
          $n \leftarrow n + 1$
          $O_n \leftarrow$ Stack uncalssed points
          $neigbrPts' \leftarrow$ points in ϵ-*neigborhood* of P'
          if $sizeof(neigbrPts') \geq minPts$ **then**
            $neigbrPts \leftarrow neigbrPts + neigbrPts'$
          end
        end
        if P' *is not a member of any cluster* **then**
          add P' to cluster C
        end
      end
    end
  end
  if $NP \geq 2$ **then**
    mark $O.LIFO(NP-1)$ as outlier
  end
end

Figure 3 : Proposed Algorithm

Figure 3 is the pseudo code of proposed outlier detection algorithm. According to the change of the reachable point of P' unclassed points list is stacked and treated as outliers, and LIFO (Last In First Out) is applied according to the number of preprocessing.

4.2 Chronic Kidney Disease Data Set

Experiments were conducted using data published by the UCI Machine Learning Repository [20]. The data was used chronic kidney disease data. It was provided by Dr. P. Soundarapandian, M.D., D.M, from Apollo Hospitals, Tamilnadu, India [21]. The data set consists of a total of 24 feature and 400 patient data, which is diagnostic data for chronic kidney disease (CKD) or not.

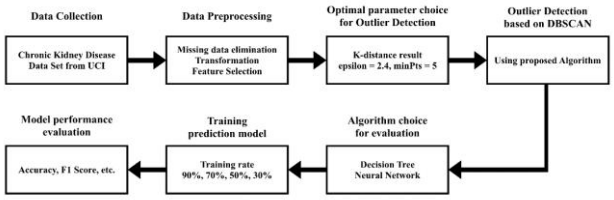


Figure 4 : Process of chronic kidney disease experiment

Figure 4 shows process of chronic kidney disesase experiment.

Table 3 : Result of feature selection

Feature	Explanation	Information Gain
hemo	Hemoglobin (numerical, hemo in gms)	0.5175
sc	Blood Urea (numerical, mgs/dl)	0.4979
sg	Specific Gravity (nominal, 1.005 - 1.025 (5 Cases))	0.4486
pcv	Packed Cell Volume (numerical)	0.4461
al	Albumin (nominal, 0~5)	0.3907
htn	Hypertension (nominal, Y/N)	0.3224
dm	Diabetes Mellitus (nominal, Y/N)	0.2921
rc	Red Blood Cell Count (numerical, millions/cmm)	0.2711
bu	Blood Urea (numerical, mgs/dl)	0.2617
bgr	Blood Glucose Random (numerical, mgs/dl)	0.2285
sod	Sodium (numerical, mEq/L)	0.1902
bp	Blood Pressure (numerical, mm/Hg)	0.1739
appet	Appetite (nominal, Good/Poor)	0.1566
pc	Pus Cell (nominal, Present/Not Present)	0.1471
pe	Pedal Edema (nominal, Y/N)	0.1434
pot	Potassium (numerical, mEq/L)	0.1308
rbc	Red Blood Cells (nominal, Normal/Abnormal)	0.1188
su	Sugar (nominal, 0 – 5 (6 Cases))	0.1170
ane	Anemia (nominal, Y/N)	0.1096
age	Age in year (numerical, Year)	0.1046

wc	White Blood Cell Count (numerical, cells/cumm)	0.0850
pcc	Pus Cell clumps (nominal, Present/Not Present)	0.0692
cad	Coronary Artery Disease (nominal, Y/N)	0.0577
ba	Bacteria (nominal, Present/Not Present)	0.0349

Table 3 is the result of feature selection by Information Gain. As a result, attribute hemo and sc used for defining optimal epsilon value.

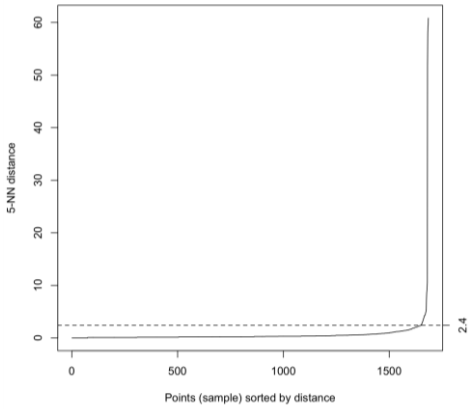


Figure 5 : K-distance graph for epsilon value

Figure 5 shows the optimal eps value of dataset. The K-distance graph was implemented in R Studio [22]. In the k-distance graph, the sharpest gradient is the optimal eps value. The sharpest gradient is 2.4, so the optimal eps value is 2.4 [23]. Thus, as parameter values for the experiment, MinPts and eps were defined as 5 and 2.4.

4.3 Breast Cancer Data Set

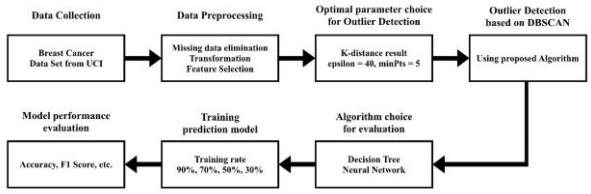


Figure 6 : Process of breast cancer data set

Experiments were conducted using data published by the UCI Machine Learning Repository. The data was used breast cancer data. It was provided by Dr. William H. Wolberg, from Clinical Sciences Center, Wisconsin, USA [24]. The data set consists of a total of 33 feature and 569 patient data, which is diagnostic data for malignant patient or benign patient. Figure 6 shows process of breast cancer data set.

Table 4 : Result of feature selection

Feature	IG Value
Perimeter_worst	0.6850
Area_worst	0.6686
Radius_worst	0.6665
concave points_worst	0.6478
concave points_mean	0.6347
perimeter_mean	0.5623
area_mean	0.5479
radius_mean	0.5410
concavity_mean	0.5171
area_se	0.5170
concavity_worst	0.4735
radius_se	0.3679
perimeter_se	0.3663
compactness_worst	0.3204
compactness_mean	0.3040
concavity_se	0.2225
concave points_se	0.1970
texture_worst	0.1881
texture_mean	0.1593
symmetry_worst	0.1492
compactness_se	0.1303
smoothness_worst	0.1235
symmetry_mean	0.0988
smoothness_mean	0.0971
fractal_dimension_worst	0.0747
fractal_dimension_se	0.0346
symmetry_se	0.0228

November-December 2019
ISSN: 0193-4120 Page No. 2246 - 2254

Table 4 is the result of feature selection by Information Gain. As a result, perimeter_worst and area_worst used for defining optimal epsilon value.

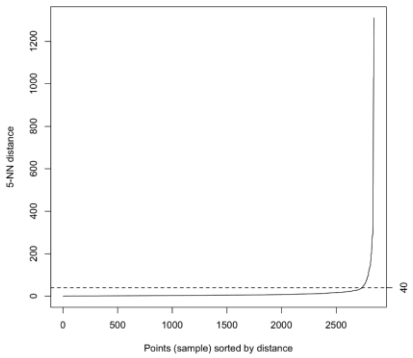
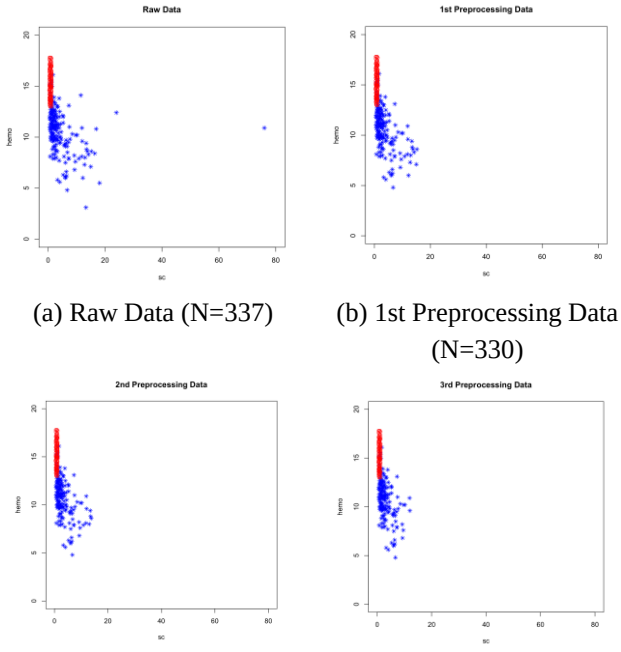


Figure 7 : K-distance graph for epsilon value

Figure 7 shows the optimal eps value of dataset. The figure 7 was implemented in R Studio. In the k-distance graph, the sharpest gradient is the optimal eps value. The sharpest gradient is 40, so the optimal eps value is 40. Thus, as parameter values for the experiment, MinPts and eps were defined as 5 and 40.

5. RESULTS

5.1 Chronic Kidney Disease Data Set



(a) Raw Data (N=337) (b) 1st Preprocessing Data (N=330)

(c) 2nd Preprocessing Data (N=325) (d) 3rd Preprocessing Data (N=319)

Figure 8 : Preprocessing result graph

Figure 8 shows the result of preprocessing graph. Whenever the number of preprocessing increases, data is preprocessed based on the cluster having a low density.

Table 5 : Result of preprocessing

Data	Number of Data	CKD Patient	Not CKD Patient
Raw Data	N = 337	N = 198	N=139
1st Preprocessing Data	N = 330	N = 191	N=139
2nd Preprocessing Data	N = 325	N = 186	N=139
3rd Preprocessing Data	N = 319	N = 180	N=139

Table 5 is the result of Preprocessing about Chronic Kidney Disease Data Set. Table 5 shows that label (CKD Patient) is preprocessed.

Table 6 : Evaluation result

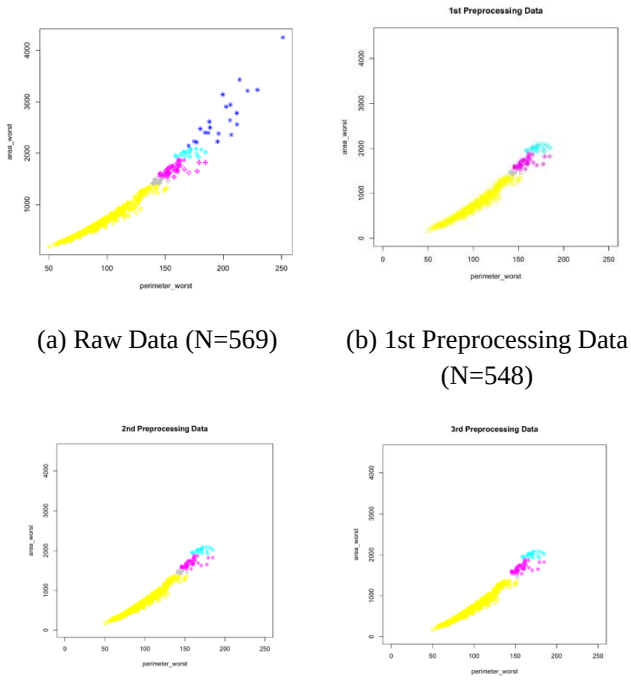
Data		**Training Data Rate**			
		90%	**70%**	**50%**	**30%**
Raw Data	*NN*	100%	100%	97.5%	96.5%
	DT	100%	97.9%	97.5%	97.4%
1st Preprocessing Data	*NN*	100%	100%	97.0%	95.1%
	DT	100%	100%	97.0%	98.1%
2nd Preprocessing Data	*NN*	100%	100%	97.2%	98.1%
	DT	100%	97.8%	97.2%	99.0%
3rd Preprocessing Data	*NN*	100%	100%	97.1%	98.1%
	DT	100%	100%	95.6%	96.2%

The results in Table 6 were implemented in Microsoft Azure Machine Learning Studio [25]. Table 6 shows the results after training and evaluating preprocessed data. The performance of Neural Network and Decision Tree algorithm is similar or improved. In addition, the number of data to be learned by preprocessing is reduced, but the accuracy is maintained or increased.

November-December 2019
ISSN: 0193-4120 Page No. 2246 - 2254

5.2 Breast Cancer Data Set

A.



(a) Raw Data (N=569) (b) 1st Preprocessing Data (N=548)

(c) 2nd Preprocessing Data (N=543) (d) 3rd Preprocessing Data (N=535)

Figure 9 : Preprocessing result graph

Figure 9 shows the result of preprocessing graph. Whenever the number of preprocessing increases, data is preprocessed based on the cluster having a low density.

Table 7 : Result of preprocessing

Data	Number of Data	Benign Patient	Malignant Patient
Raw Data	N = 569	N = 357	N=212
1st Preprocessing Data	N = 548	N = 357	N=191
2nd Preprocessing Data	N = 543	N = 357	N=186
3rd Preprocessing Data	N = 535	N = 357	N=178

Table 7 is the result of preprocessing about Breast Cancer Data Set. Table 7 shows that label (Malignant Patient) is preprocessed.

Table 8 : Evaluation result

Data		Training Data Rate			
		90%	**70%**	**50%**	**30%**
Raw Data	***NN***	96.5 %	96.5 %	97.5 %	96.2%
	DT	100%	97.1 %	93.7 %	94.2%
1st Preprocessing Data	***NN***	100%	97.6 %	97.8 %	92.4%
	DT	98.2 %	97.6 %	96.7 %	95.3%
2nd Preprocessing Data	***NN***	96.3 %	98.8 %	96.3 %	94.2%
	DT	92.6 %	96.3 %	96.3 %	94.5%
3rd Preprocessing Data	***NN***	98.1 %	96.9 %	97.0 %	95.7%
	DT	96.2 %	96.3 %	95.1 %	93.6%

The result in Table 8 was implemented in Microsoft Azure Machine Learning Studio. Table 8 shows the results after training and evaluating preprocessed data. The performance of Neural Network and Decision Tree algorithm is similar or improved. In addition, the number of data to be learned by preprocessing is reduced, but the accuracy is maintained or increased [27-29].

6. CONCLUSION

This study proposes an improved algorithm using the characteristics of DBSCAN. The neural network and decision tree algorithms were applied to evaluate the prediction and the results based on Hybrid Machine Learning. The medical data to verify the performance were the chronic kidney disease data and the breast cancer data in the UCI repository.

As a result, it can be confirmed that preprocessing of only the data of the other class in the binary class and the preprocessing through

November-December 2019
ISSN: 0193-4120 Page No. 2246 - 2254

the evaluation results are successful. If we use the preprocessing algorithm proposed in this paper for medical data preprocessing, high performance can be expected.

ACKNOWLEDGEMENT

"This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2017R1D1A1B03034411)."

REFERENCES

[1] Gaspar, J., Catumbela, E., Marques, B., & Freitas, A., **A Systematic Review of Outliers Detection Techniques in Medical Data-Preliminary Study**, In HEALTHINF, 2011, pp. 575-582

[2] Lincoln, T. L., & Builder, C., **Global healthcare and the flux of technology. International journal of medical informatics**, Vol.53 No.2-3, 1999, pp. 213-224

[3] Kumar, V., Kumar, D., & Singh, R. K., **Outlier mining in medical databases: an application of data mining in health care management to detect abnormal values presented in medical databases**, International Journal of Computer Science and Network Security (IJCSNS), Vol.8, No.8, 2008, pp.272-277

[4] Deneshkumar, V., Senthamaraikannan, K., & Manikandan, M., **Identification of outliers in medical diagnostic system using data mining techniques**, International Journal of Statistics and Applications, Vol.4, No.6, 2014, pp.241-248

[5] Chandola, V., Banerjee, A., & Kumar, V., **Anomaly Detection: A Survey**, ACM Computing Surveys (CSUR), Vol. 41 No.3, 2009

[6] Marsland, S., **Machine learning: an algorithmic perspective**. Chapman and Hall/CRC., USA., 2011

[7] Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., & Liu, H., **Feature selection: A data perspective**, ACM Computing Surveys (CSUR), Vol.50, No.6, 2018

[8] Lewis, D. D., **Feature selection and feature extraction for text categorization**, In Proceedings of the workshop on Speech and Natural Language. Association for Computational Linguistics, 1992, pp. 212-217

[9] Tom, M., **Machine Learning**, Hill McGraw., USA, 1997

[10] Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X., **DBSCAN revisited, revisited: why and how you should (still) use DBSCAN**, ACM Transactions on Database Systems (TODS), Vol.42, No.3, 2017, pp 19.

[11] Nayyar, A., & Puri, V. (2017). **Comprehensive Analysis & Performance Comparison of Clustering Algorithms for Big Data**. Review of Computer Engineering Research, 4(2), 54-80.

[12] Jin-Won Kim., **Cardiovascular Disease Diagnosis System Construction using Hybrid Learning Algorithm (Thesis)**, Jeonju, Jeollabuk-do, Chonbuk National University. 2019

[13] Ji Won Kim. **A study on Machine Learning-based Ransomware Detection Model using Hybrid Model (Thesis)**, Seoul, Konkuk University. 2016

[14] Bala, J., Huang, J., Vafaie, H., DeJong, K., & Wechsler, H., **Hybrid learning using genetic algorithms and decision trees for pattern classification**, In IJCAI, 1995, pp. 719-724.

[15] Tsai, C. F., & Chen, M. L., **Credit rating by hybrid machine learning techniques**, Applied soft computing, Vol.10, No.2, 2010, pp 374-380.

[16] Ijaz, M., Alfian, G., Syafrudin, M., & Rhee, J., **Hybrid Prediction Model for Type 2 Diabetes and Hypertension Using DBSCAN-Based Outlier Detection, Synthetic Minority Over Sampling Technique (SMOTE), and Random Forest**, Applied Sciences, Vol.8, No.8, 2018, pp 1325.

November-December 2019
ISSN: 0193-4120 Page No. 2246 - 2254

[17] Abid, A., Kachouri, A., & Mahfoudhi, A., **Outlier detection for wireless sensor networks using density-based clustering approach**, IET Wireless Sensor Systems, Vol.7, No.4, 2017, pp 83-90.

[18] Alfian, G., Syafrudin, M., & Rhee, J., **Real-time monitoring system using smartphone-based sensors and NoSQL database for perishable supply chain. Sustainability**, Vol.9, No.11, 2017, pp 2073.

[19] ElBarawy, Y. M., Mohamed, R. F., & Ghali, N. I., **Improving social network community detection using DBSCAN algorithm**, In 2014 World Symposium on Computer Applications & Research (WSCAR) IEEE, 2014, pp 1-6

[20] Tian, H. X., Liu, X. J., & Han, M, **An outliers detection method of time series data for soft sensor modeling**, In 2016 Chinese Control and Decision Conference (CCDC) IEEE, 2016, pp 3918-3922

[21] UCI Repository, Available Website : http://archive.ics.uci.edu/ml/index.php

[22] Chronic Kidney Disease Data Set, Available Website : https://archive.ics.uci.edu/ml/datasets/chronic_kidney_disease

[23] Microsoft Azure Machine Learning, Available Website : https://studio.azureml.net/

[24] Zhou, A., Zhou, S., Cao, J., Fan, Y., & Hu, Y., **Approaches for scaling DBSCAN algorithm to large spatial databases**, Journal of computer science and technology, 15(6), 2000, pp 509-526.

[25] Breast Cancer Data Set, Available Website : https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)

[26] R Studio, Website : https://www.r-project.org/

[27] Fabus, M., Dubrovina, N., Guryanova, L., Chernova, N., Zyma, O., 2019. Strengthening financial decentralization: driver or risk factor for sustainable socio-economic development of territories. Entrepreneurship and Sustainability Issues, 7(2), 875-890. http://doi.org/10.9770/jesi.2019.7.2(6)

[28] Sasongko, G.; Huruta, A.D.; Wardani, A. 2019. Does the Wagner's Law exist in a strategic national area? An evidence from Kedungsepur - Indonesia, Insights into Regional Development 1(2): 99-117. https://doi.org/10.9770/ird.2019.1.2(2)

[29] Jabarullah, N.H. (2019) Production of olefins from syngas over Al2O3 supported Ni and Cu nano-catalysts, Petroleum Science and Technology, 37 (4), 382 – 385.