

Currently the hospital is a place that is very vulnerable to the transmission of Covid-19, so giving birth in a hospital is very risky. In addition, the hospital currently only accepts cesarean deliveries, while mothers who can give birth vaginally are recommended to give birth in a midwife because the chances of being exposed to Covid-19 are much lower. In general, this study aims to examine the performance of the LDA-SVM method in predicting whether a prospective mother needs to undergo a C-section or simply give birth normally. The aims of this study are:

1) to determine the best parameters for building the detection model;

2) to determine the best accuracy from the model;

3) to compare the accuracies with the other methods.

The data used in this study is the dataset of caesarian section. This data consists of the results of 80 pregnant women following C-section with the most important characteristics of labor problems in the clinical field. Based on the results of the experiments that have been carried out, several parameter values that provide the best results for building the detection model are obtained, namely σ (sigma) -5.9 for 70 % training data, $\sigma=4$, -6.1 and 6.6 for 80 % training data and $\sigma=4$ and 16 for 90 % training data. Besides, the results obtained show that the LDA-SVM method is able to classify the C-section method properly with an accuracy of up to 100 %. This research is also able to surpass the methods in previous studies. The results show that LDA-SVM for this case study generates an accuracy of 100.00 %. This method has great potential to be used by doctors used as an early detection to determine whether a mother needs to go through a C-section or simply give birth vaginally. So that mothers can prevent the transmission of Covid-19 in the hospital

Keywords: caesarian section, cesarean deliveries, LDA, SVM, Covid-19, pregnant women

PERFORMANCE EVALUATION OF LINEAR DISCRIMINANT ANALYSIS AND SUPPORT VECTOR MACHINES TO CLASSIFY CESAREAN SECTION

Abdul Azis Abdillah

Corresponding author

Master of Mathematics, Assistance Professor*

E-mail: abdul.azis.a@mesin.pnj.ac.id

A z w a r d i

Master of Computer, Assistance Professor*

Sulaksana Permana

Doctor of Engineering in Metallurgy and Materials

Centre of Mineral Processing and Corrosion Research

Department of Metallurgy and Materials

Universitas Indonesia

Jl. Margonda Raya, Pondok Cina, Kecamatan Beji, Kota Depok,

Jawa Barat, Indonesia, 16424

Iwan Susanto

Doctor of Materials Science and Engineering,

Assistance Professor *

Fuad Zainuri

Doctoral of Mechanical Engineering, Assistance Professor*

Samsul Arifin

Doctoral of Mathematic, Assistance Professor

Statistics Department

School of Computer Science

Bina Nusantara University

Jl. K. H. Syahdan No. 9, Kemanggis, Palmerah,

Jakarta, Indonesia, 11480

*Department of Mechanical Engineering

Politeknik Negeri Jakarta

Jl. Prof. DR. G.A. Siwabessy, Kukusan, Kecamatan Beji,

Kota Depok, Jawa Barat, Indonesia, 16424

Received date 10.08.2021

How to Cite: Abdillah, A. A., Azwardi, A., Permana, S., Susanto, I., Zainuri, F., Arifin, S. (2021). Performance evaluation of linear discriminant analysis and support vector machines to classify caesarean section. *Eastern-European Journal of Enterprise Technologies*,

Accepted date 25.09.2021

5 (2 (113)), 37–43. doi: <https://doi.org/10.15587/1729-4061.2021.242798>

Published date 29.10.2021

1. Introduction

Giving birth in a hospital during the Covid-19 outbreak is very risky, this is because the hospital is a very vulnerable place for transmitting Covid-19. Currently, hospitals are prioritized for mothers who will give birth by cesarean section (C-section), because only hospitals have facilities for cesarean delivery. However, if mothers tend to be able

to give birth normally, it is recommended to give birth in a midwife because the chance of being exposed to Covid-19 is much lower. Cesarean section or C-section is the process of giving birth to a baby, which is done by cutting the abdomen to the uterus of the mother. The incision in the abdomen is a way out of the baby from the womb. Doctors usually make long incisions in a horizontal direction just above the pubic bone. The method of giving birth by C-section is usually

done around the 39th week, or when doctors recommend a mother to undergo this operation. Usually, the doctor will recommend a C-section if your pregnancy is at risk.

Compared to giving birth normally through the vagina, giving birth by C-section does require a longer healing time. That is why it is important to consult a doctor first before deciding to undergo this birthing procedure. C-section should generally be done if you experience certain complications in pregnancy. These complications can usually complicate the process of giving birth normally or through the vagina. Even if forced to carry out normal childbirth, the risk is endangering the health and safety of the mother and baby. This is where the doctor will suggest the option of having a C-section delivery. C-section can be planned from the beginning or the middle of pregnancy, or when complications arise before the time of giving birth.

The development of artificial intelligence in the medical field is very rapid. Some of them are diabetes detection [1–3], brain detection [4–6], cancer [7, 8], heart disease [9–11] and others. Disease detection using artificial intelligence has also been used by the medical team to be an early diagnosis in detecting an abnormal condition.

2. Literature review and problem statement

Currently, around the world, the number of babies born using cesarean section is increasing. By doing C-section, the woman or baby does not get any health benefits. Because of the importance of limiting the increased use of C-section in childbirth, this year, new guidelines on non-clinical interventions specifically designed to reduce unnecessary C-section were published by WHO. In line with the directives provided by WHO, several researchers have also conducted research to reduce the potential use of the cesarean section method for delivery by developing an intelligent system to predict whether a mother needs to undergo a cesarean section or only give birth vaginally.

Research related to birth detection was conducted by [12] using machine learning methods such as Support Vector Machines (SVM), Random Forests (RF), Naive Bayes (NB), K-Nearest Neighbors (KNN) and Logistics Regression (LR). The results of the study [12] show that the average accuracy of machine learning detection results is relatively low. Meanwhile, research [13] used Neural Network and only produced an average accuracy of 80 %. This is probably due to the large data dimensions so that the classification method is not able to produce a good model. The same thing shown by research [14] is that the more features used, the greater the possibility of redundancy or irrelevance so that it can cause a decrease in classifier performance. One way to solve this problem is to do pre-processing in the form of dataset feature selection. That approach was also used to detect C-section births using numerical data and signals [15]. The results of the study [15] showed the positive impact of using selection features in the form of increasing detection accuracy. While the study [16] related to cardiotocography (CTG) with 17 features, after reducing the dimensions, the accuracy obtained was more than 94 %. In addition, the same constraint was also found in the study [17] in the case of diabetes detection. Feature selection was used to identify important features in the data set. Feature selection is the key in increasing the accuracy of early diabetes detection using SVM. Meanwhile, research [18] shows that in addition to increasing detection accuracy, feature selection also short-

ens computational time. Furthermore, in research [19], feature selection was applied to the voice-based data used in brain detection. The results showed that the SVM method combined with Kernel RBF on pre-processed data could produce high sensitivity and specificity values for detecting Parkinson's disease. By obtaining the parameters needed to build an optimal model, this system is expected to help the medical field to make an early diagnosis of the birth method before taking action for cesarean section or normal delivery.

3. The aim and objectives of the study

The aim of the study is to examine the performance of the LDA-SVM method in predicting whether a prospective mother needs to undergo a C-section or simply give birth normally.

To achieve this aim, the following objectives are accomplished:

- to determine the best parameters for building the detection model;
- to determine the best accuracy from the model;
- to compare the accuracies with the other methods.

4. Materials and methods

In this section, we will discuss the machine learning methods used in this study such as Linear Discriminant Analysis as a reduction dimension model and Support Vector Machines as a classifier, research framework, dataset and evaluation model.

4.1. Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) [20–22] is a statistical method that is quite popular in solving cases related to pattern recognition. Initially, LDA was used to perform facial recognition. This method is known as Fisher's Linear Discriminant. However, at this time, LDA is more widely used to extract important information from data and reduce data dimensions.

This method is basically intended to find linear combinations of features that are the main characteristics, so that the data can be separated into two or more object classes or groups. Once the data has been processed using LDA, new data will have a more dispersed distribution and can finally increase recognition success. After that, the new data combination obtained can be reprocessed using both linear and non-linear classification methods.

In this study, LDA is used to find the main characteristics of the Caesar dataset and if possible, from the existing data, the researcher also reduces the dimensions so that it is expected that the results of the dimensional reduction can be classified properly by SVM. The following is the LDA algorithm. First of all, we will define the covariance matrix in class S_W and the inter-class covariance matrix S_B , each of which is exemplified as follows:

$$S_W = \sum_{i=1}^C \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T; \quad (1)$$

$$S_B = \sum_{i=1}^C N_i (x_k - \mu_i)(x_k - \mu_i)^T, \quad (2)$$

where X_k is data in k-position, C is total class, N_i is the amount of data in class, μ is the average from all data and μ_i is the average data in class i .

So that the covariance matrix in class (S_W) can be minimized then we maximize the covariance matrix between classes (S_B). Next, we can determine the eigenvector (V) so that we can maximize the ratio of (3):

$$\frac{\det(VS_B V^T)}{\det(VS_W V^T)} \tag{3}$$

Resulting in a solution

$$S_B V = \lambda S_W V \tag{4}$$

Then the eigenvalue (λ) and eigenvector (V) values of the covariance matrix (4) are found, namely:

$$\text{cov} = S_B (S_W)^{-1} \tag{5}$$

Because the LDA method discussed in this paper is the LDA method with 2 classes, the eigenvector used is c^{-1} . The eigenvector value can be chosen based on the largest eigenvalue of the covariance matrix. After the eigenvector is known, the LDA feature can be determined by calculating:

$$F_x = \sum_{i=1}^k (x_i - \mu)^T \times V \tag{6}$$

Furthermore, the data that has been reduced is processed using SVM.

4.2. Support Vector Machines

The method used in this study is the Support Vector Machine (SVM). In addition to use in the health sector, SVM had also been widely applied in other areas, i. e. pattern recognition [23], information retrieval [24], military [25],

forest detection [26], etc. SVM is a classification method in data mining. SVM can also make predictions in both classification and regression. Basically, SVM has a linear principle, but at this time SVM has evolved and then it can be utilized to solve non-linear cases. The way SVM works in non-linearly separable issues is to include the kernel concept. In this dimensionless field, we will seek a separator or often called a hyperplane. The hyperplane which SVM want to find is the one that can maximize the margin between two classes. We can find the best hyperplane between two data by estimating the edge and afterward finding the maximum point from the separator. The effort to find the best hyperplane as a class separator is the core of the process in the SVM method. The kernel function used in this study is the RBF kernel:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \tag{7}$$

RBF kernel can be recognized as the squared Euclidean distance between the two feature vectors. Meanwhile, σ (sigma) is an independent parameter that needs to be set to generate the SVM model.

4.3. Research Framework

The framework of this study can be seen in Fig. 1. In this section, we will explain the dataset and evaluation methods used by this research.

In this study, first the data was pre-processed using LDA by reducing the dimensions. Then the data is divided into training and testing data. Furthermore, the training data is used to form a detection model that was developed based on the SVM method. In the next stage, the model that has been built is validated by testing it on training data and testing data.

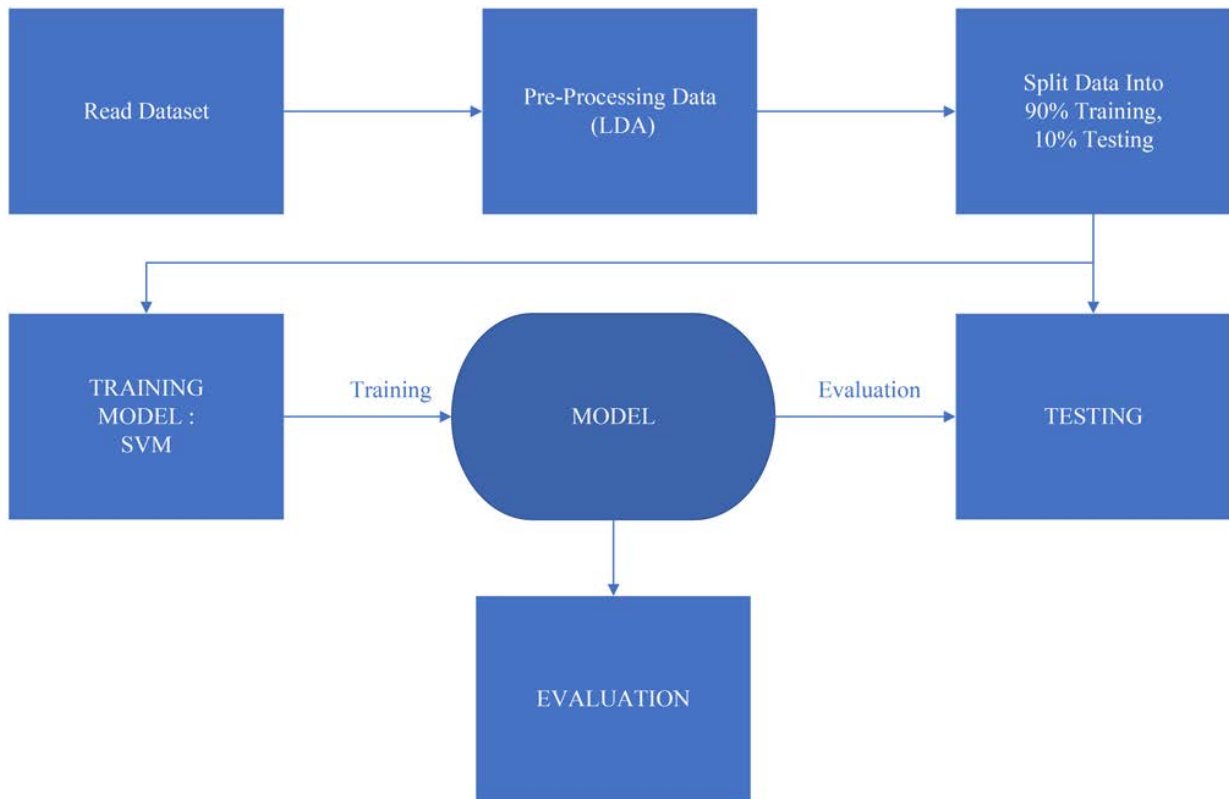


Fig. 1. Research framework for LDA-SVM model

4. 4. Cesarean Dataset

The dataset utilized in this investigation is the dataset of C-section [27]. This dataset contains data about cesarean section consequences of 80 pregnant mothers with the most significant attributes of giving birth issues in the clinical field. Table 1 shows the features used in the dataset.

Table 1

Explanation of dataset features

Attribute	Value
'Age'	{22,26,28,27,32,36,33,23,20,29,25,37,24,18,30,40,31,19,21,35,17,38}
'Delivery number'	{1,2,3,4}
'Delivery time'	{0,1,2}->{0=timely, 1=premature, 2=latecomer}
'Blood pressure'	{2,1,0}->{0=low, 1=normal, 2=high}
'Heart problems'	{1,0}->{0=apt, 1=inept}
Cesarean class	{0,1}->{0=No, 1=Yes}

The dataset comprises 80 cases, made out of five ascribes, which are the most significant attributes of childbirth issues, specifically age, delivery number, delivery time, blood pressure and heart status. In this study, all features were extracted using LDA so that only important features were obtained that would be used in the classification stage.

4. 5. Confusion Matrix

To evaluate the performance of the classifier, in this paper the author uses a confusion matrix. The confusion matrix is a method commonly used to calculate accuracy in the concept of data mining or decision support systems. To measure the performance of the LDA-SVM method, researchers used accuracy as an evaluator. Accuracy can be described as the appropriate number of predictions divided by the total of all data. Prediction of the accuracy of the formation of classification models can be obtained as follows [28]:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{8}$$

where TP or True Positive is the number of Caesar patients, which are correctly classified as C-section treatment. Then, TN or True Negative is the number of non-C-section patients, which are correctly classified as non-C-section treatment. Furthermore, FP or False Positive is the number of C-section patients who received non-C-section treatment. The last, FN or False Negative is the number of non-C-section patients who received C-section treatment.

5. Research results of the birth detection method

5. 1. Finding the Best Parameters

The results of all the experiments can be seen in Table 2. In the early stages of implementation, we preprocessed the cesarean dataset using LDA. The results obtained are that there are two features with the largest eigenvalue or greater than 0. The two features are Delivery Number and Delivery Time with the eigenvalues results being 0.214713766357188

and 1.70605365610297e-17, respectively. In the next stage, the data was reduced into two dimensions using only two corresponding features, then the data was divided into training data and testing data to then enter the classification stage. The researcher divides the data into training data and testing data randomly ten times so that each same experiment will produce different data groups. The training data and testing data consisted of three groups, namely 70 % training data (30 % testing data), 80 % training data (20 % testing data) and 90 % training data (10 % testing data).

The results of the performance evaluation for the optimal parameter values for both training and testing data can be seen in Table 2. We used two parameters in this classification stage using SVM, namely C and Sigma. The C parameter value used is default while the sigma parameter used consists of three sets, namely:

- 1) sigma value between 2² and 2⁵ with an exponential increment of 1;
- 2) sigma values between -10 and -3 in 0.1 increments;
- 3) sigma values between 2.4 and 10 in 0.1 increments.

Table 2

Accuracy for All Experiments

Exp	Sig- ma	70 % Training	30 % Testing	Sig- ma	80 % Training	20 % Testing	Sig- ma	90 % Training	10 % Testing
1	5.9	67.86	70.83	-9.6	98.44	93.75	-9.6	97.22	100.00
2	3.6	62.50	58.33	-9.4	96.88	81.25	-9.4	98.61	100.00
3	2.4	62.50	70.83	-6.6	100.00	100.00	-6.6	97.22	100.00
4	2.4	64.29	70.83	-6.1	100.00	100.00	-6.1	97.22	100.00
5	2.4	64.29	62.50	-4.4	98.44	100.00	-4.4	98.61	100.00
6	2.4	62.50	62.50	4.0	100.00	100.00	4.0	98.61	100.00
7	4.0	66.07	70.83	4.0	100.00	100.00	4.0	97.22	100.00
8	4.0	55.36	58.33	4.0	98.44	100.00	4.0	100.00	100.00
9	4.0	66.07	58.33	16.0	93.75	81.25	16.0	98.61	100.00
10	4.0	66.07	83.33	16.0	95.31	100.00	16.0	100.00	100.00
Average	-	63.75	66.67	Average	98.13	95.63	Average	98.33	100.00

5. 2. Finding the best accuracy from the model

Fig. 2 shows the results of the average accuracy of the training and testing data from all experiments. It can be seen that there is a trend of increasing the accuracy value, which is directly proportional to the amount of training data used from each experiment carried out.

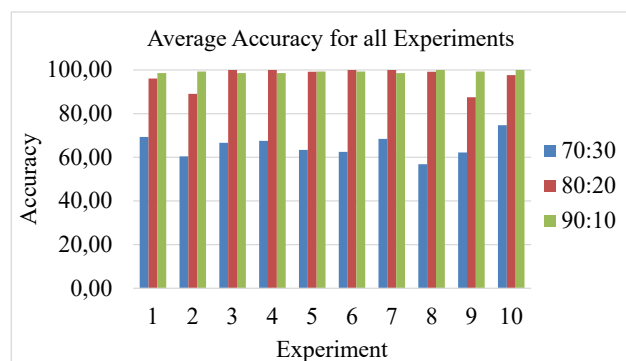


Fig. 2. Average Accuracy for All Experiments

The lowest average accuracy was obtained in the eighth experiment using the data composition of 70:30 with an accuracy

value of 58. On the other hand, the highest average accuracy was also obtained in the eighth experiment with an average accuracy=99 % for the data composition of 80:20 and 100 % for the data composition of 90:10. Meanwhile, the gap between each experiment is also relatively the same, namely the proportion for 70:30 and 80:20 is around 30 % and between 80:20 and 90:10 is only less than 3 %. This is quite different for experiments 2 and 9 where the distance between each proportion of data can reach 1020 %.

5. 3. Comparison with the other methods

Table 3 shows the results of accuracy comparison using other methods that has have been carried out by other researchers [12, 13] on Caesar’s data. When compared with other methods that have been done before, namely SVM, Random Forest, Naive Bayes, Logistic Regression, K-Nearest Neighbors and Neural Network, the performance of LDA-SVM is the best among the five other classifiers.

Table 3

Performance comparison with nother classifiers

Number	Classifier	Accuracy
1	SVM [12]	76.3 %
2	Random Forest [12]	95.0 %
3	Naïve Bayes [12]	76.3 %
4	Logistic Regression [12]	77.5 %
5	K-Nearest Neighbors [12]	95.0 %
6	Neural Network [13]	77.5 %
7	LDA-SVM*	100.0 %

Note: (*) conducted in this study

We can say that because of the right combination of methods and use of sigma parameters. We hope this research can help the medical field in determining the birth method appropriately both in Indonesia and throughout the world.

6. Discussion of the research results of the birth detection method

In the preprocessing stage, there are only two features that have eigenvalues greater than zero, namely Delivery Number and Delivery Time. The results obtained indicate that only two features play an important role in forming a good model. While other features such as age, blood pressure and heart problems can be ignored to avoid a decrease in classification performance. Based on the results of feature extraction using LDA, now we only need to use the new data with two dimensions. The results of Caesar’s data classification using two dimensions showed mixed results. The highest accuracy of the LDA-SVM method was obtained when the data used for training were 80 % and 90 %, respectively. The classification results obtained are 100 % of the data can be classified correctly. Sigma parameters that produce perfect scores are 4, -6.1 and -6.6 on 80 % training data and 4 and 16 on 90 % training data.

Based on Fig. 2, it can be seen that the highest average accuracy in the experiment with a composition

of 70 % training data was 74.70 %, which was obtained in the 10th experiment with the sigma parameter used of 4.0. Meanwhile, the composition of 80 % training data is 100 %, which is obtained in experiments 3–7 with the sigma parameters used of -6.6, -6.1, 4.0 and 4.0, respectively. While the composition of 90 % training data is 100 % obtained in the 8th and 10th experiments with the sigma parameters of 4 and 16. Based on Table 2 and Fig. 2, it can also be seen that the resulting increase in accuracy is not directly proportional to the greater sigma parameters used. In converse, based on the results of the average accuracy of each group of 70 %, 80 % and 90 %, the greater the data used as training data, the greater the accuracy obtained.

Eventually, when compared with the SVM method [12] with no reduction in dimensions through LDA, which only reached 76 %, in the case of Caesar’s data, it can be concluded that the accuracy of classification using SVM is affected by dimensional reduction using LDA. So, this model has great potential to be used by doctors as an early detection to determine whether a mother needs to go through a C-section or simply give birth vaginally.

This study has several limitations, one of which is that the data used must have the features as described in Table 1. Furthermore, because the determination of the delivery method is very decisive in determining the safety and health of mothers and babies, this research needs to get scientific validation from experts especially in the medical field. In addition, currently this study only evaluated one dimension reduction method and one kernel type for the classification method. For further development, researchers can try various other detection methods combined with other feature extraction models to compare and evaluate the performance.

7. Conclusions

1. Based on the results of dimension reduction using LDA, it is known that the most important features for making a classification model on cesarean data are Delivery Number and Delivery Time. Furthermore, the proportion of dataset used as training and testing data in this study was 70:30, 80:20 and 90:10. The best parameters from each experiment were obtained when the parameters (sigma) used were -5.9 for 70:30 data, sigma=4, -6.1 and -6.6 for 80:20 training data and sigma=4 and 16 for 90:10 training data.

2. The best accuracy obtained during the experiments for training and testing data respectively is 67.86 % and 70.83 for 70:30 proportion, 100 % and 100 % for 80:20 and 90:10.

3. Based on the results of experiments and comparisons with previous methods that have been done, it can be seen that LDA-SVM produces better accuracy than previous experiments.

Acknowledgments

The authors would like to thank Politeknik Negeri Jakarta, for the support of research funding, through the Research Grant in the Field of Science and Institutional Development.

References

1. Abdilllah, A. A., Suwarno, S. (2016). Diagnosis of Diabetes using Support Vector Machines with Radial Basis Function Kernels. *International Journal of Technology*, 7 (5), 849. doi: <https://doi.org/10.14716/ijtech.v7i5.1370>
2. Abdilllah, A. A., Prianto, B. (2019). Pembelajaran Mesin Menggunakan Principal Component Analysis dan Support Vector Machines untuk Mendeteksi Diabetes. *Jurnal Matematika Dan Sains*, 24 (1), 10–14. doi: <https://doi.org/10.5614/jms.2019.24.1.2>
3. Bernardini, M., Romeo, L., Misericordia, P., Frontoni, E. (2020). Discovering the Type 2 Diabetes in Electronic Health Records Using the Sparse Balanced Support Vector Machine. *IEEE Journal of Biomedical and Health Informatics*, 24 (1), 235–246. doi: <https://doi.org/10.1109/jbhi.2019.2899218>
4. Zhang, Y.-D., Jiang, Y., Zhu, W., Lu, S., Zhao, G. (2017). Exploring a smart pathological brain detection method on pseudo Zernike moment. *Multimedia Tools and Applications*, 77 (17), 22589–22604. doi: <https://doi.org/10.1007/s11042-017-4703-0>
5. Ahmmed, R., Swakshar, A. S., Hossain, M. E., Rafiq, M. A. (2017). Classification of tumors and it stages in brain MRI using support vector machine and artificial neural network. 2017 International Conference on Electrical, Computer and Communication Engineering (ECCE). doi: <https://doi.org/10.1109/ecace.2017.7912909>
6. Birare, G., Chakkarwar, V. A. (2018). Automated Detection of Brain Tumor Cells Using Support Vector Machine. 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT). doi: <https://doi.org/10.1109/iccant.2018.8494133>
7. Menaka, K., Karpagavalli, S. (2013). Breast Cancer Classification using Support Vector Machine and Genetic Programming. *International Journal of Innovative Research in Computer and Communication Engineering*, 1 (7), 1410–1417.
8. Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics & Proteomics*, 15 (1), 41–51. doi: <https://doi.org/10.21873/cgp.20063>
9. Tomar, D., Agarwal, S. (2014). Feature Selection based Least Square Twin Support Vector Machine for Diagnosis of Heart Disease. *International Journal of Bio-Science and Bio-Technology*, 6 (2), 69–82. doi: <https://doi.org/10.14257/ijbsbt.2014.6.2.07>
10. Yang, C., An, B., Yin, S. (2018). Heart-Disease Diagnosis via Support Vector Machine-Based Approaches. 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC). doi: <https://doi.org/10.1109/smc.2018.00534>
11. Nilashi, M., Ahmadi, H., Manaf, A. A., Rashid, T. A., Samad, S., Shahmoradi, L. et. al. (2020). Coronary Heart Disease Diagnosis Through Self-Organizing Map and Fuzzy Support Vector Machine with Incremental Updates. *International Journal of Fuzzy Systems*, 22 (4), 1376–1388. doi: <https://doi.org/10.1007/s40815-020-00828-7>
12. Amin, M. Z., Ali, A. (2017). Performance Evaluation of Supervised Machine Learning Classifiers for Predicting Healthcare Operational Decisions. C-Section Classification Database Report, UCI Machine Learning Repository. doi: <http://dx.doi.org/10.13140/RG.2.2.26371.25127>
13. Suwarno, Santo, P. (2019). Performance Evaluation Of Artificial Neural Network Classifiers For Predicting Cesarean Sections. *International Journal Of Scientific & Technology Research*, 8 (9), 1843–1846. Available at: <http://www.ijstr.org/final-print/sep2019/Performance-Evaluation-Of-Artificial-Neural-Network-Classifiers-For-Predicting-Cesarean-Sections.pdf>
14. Desyani, T., Saifudin, A., Yulianti, Y. (2020). Feature Selection Based on Naive Bayes for Caesarean Section Prediction. *IOP Conference Series: Materials Science and Engineering*, 879, 012091. doi: <https://doi.org/10.1088/1757-899x/879/1/012091>
15. Fergus, P., Selvaraj, M., Chalmers, C. (2018). Machine learning ensemble modelling to classify caesarean section and vaginal delivery types using Cardiocography traces. *Computers in Biology and Medicine*, 93, 7–16. doi: <https://doi.org/10.1016/j.combiomed.2017.12.002>
16. Ricciardi, C., Improta, G., Amato, F., Cesarelli, G., Romano, M. (2020). Classifying the type of delivery from cardiocographic signals: A machine learning approach. *Computer Methods and Programs in Biomedicine*, 196, 105712. doi: <https://doi.org/10.1016/j.cmpb.2020.105712>
17. Lukmanto, R. B., Suharjo, Nugroho, A., Akbar, H. (2019). Early Detection of Diabetes Mellitus using Feature Selection and Fuzzy Support Vector Machine. *Procedia Computer Science*, 157, 46–54. doi: <https://doi.org/10.1016/j.procs.2019.08.140>
18. Howsalya Devi, R. D., Bai, A., Nagarajan, N. (2020). A novel hybrid approach for diagnosing diabetes mellitus using farthest first and support vector machine algorithms. *Obesity Medicine*, 17, 100152. doi: <https://doi.org/10.1016/j.obmed.2019.100152>
19. Lahmiri, S., Shmuel, A. (2019). Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine. *Biomedical Signal Processing and Control*, 49, 427–433. doi: <https://doi.org/10.1016/j.bspc.2018.08.029>
20. Brusco, M. J., Voorhees, C. M., Calantone, R. J., Brady, M. K., Steinley, D. (2018). Integrating linear discriminant analysis, polynomial basis expansion, and genetic search for two-group classification. *Communications in Statistics - Simulation and Computation*, 48 (6), 1623–1636. doi: <https://doi.org/10.1080/03610918.2017.1419262>
21. Ren, R., Han, K., Zhao, P., Shi, J., Zhao, L., Gao, D. et. al. (2019). Identification of asphalt fingerprints based on ATR-FTIR spectroscopy and principal component-linear discriminant analysis. *Construction and Building Materials*, 198, 662–668. doi: <https://doi.org/10.1016/j.conbuildmat.2018.12.009>
22. Ali, L., Zhu, C., Zhang, Z., Liu, Y. (2019). Automated Detection of Parkinson's Disease Based on Multiple Types of Sustained Phonations Using Linear Discriminant Analysis and Genetically Optimized Neural Network. *IEEE Journal of Translational Engineering in Health and Medicine*, 7, 1–10. doi: <https://doi.org/10.1109/jtehm.2019.2940900>

23. A Abdillah, A. A. (2019). Support vector machines untuk menyelesaikan masalah klasifikasi pada pengenalan pola. *Jurnal Poli-Teknologi*, 18 (2). doi: <https://doi.org/10.32722/pt.v18i2.1432>
24. Abdillah, A. A., Murfi, H., Satria, Y. (2015). Uji Kinerja Learning to Rank dengan Metode Support Vector Regression. *IndoMS Journal on Industrial and Applied Mathematics*, 2 (1), 15–25. Available at: <https://adoc.pub/uji-kinerja-learning-to-rank-dengan-metode-support-vector-re.html>
25. Ratna S, D., Setyono, B., Herdha, T. (2016). Bullet Image Classification using Support Vector Machine (SVM). *Journal of Physics: Conference Series*, 693, 012009. doi: <https://doi.org/10.1088/1742-6596/693/1/012009>
26. Rokhmatuloh, Murfi, H., Ardiansyah (2013). A method to derive optimal decision boundary in SVM method for forest and non-forest classification in Indonesia. 34th Asian Conference on Remote Sensing 2013, ACRS (2013). Bali, 2431–2442.
27. Amin, M. Z., Ali, A. (2018). Caesarian Section Classification Dataset Data Set. UCI Machine Learning Repository. Available at: <https://archive.ics.uci.edu/ml/datasets/Caesarian+Section+Classification+Dataset>
28. Gorunescu, F. (2011). *Data Mining: Concepts, Models and Techniques*. Springer, 360. doi: <https://doi.org/10.1007/978-3-642-19721-5>