IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Vol.x, No.x, July xxxx, pp. 1~5
ISSN (print): 1978-1520, ISSN (online): 2460-7258
DOI: 10.22146/ijccs.xxxx

Optimal Feature Selection in Diabetes Classification Using the MLP Algorithm

Maulana Muhammad Jogo Samodro *[1], Muhammad Kunta Biddinika [2], Abdul Fadlil [3]
[1,2]Master of Informatic Universitas Ahmad Dahlan, Yogyakarta
[3]Department of Electrical Universitas Ahmad Dahlan, Yogyakarta
e-mail: *[1]2108048022@webmail.uad.ac.id, [2]muhammad.kunta@mti.uad.ac.id
[2]fadlil@mti.uad.ac.id

Abstrak

Tahun 2021, sebanyak 531 juta penduduk dunia terkena penyakit diabetes, dengan 90% di antaranya merupakan tipe 2. Diabetes dapat menjadi komorbiditas bagi penyakit lain seperti ginjal dan jantung. Penelitian ini bertujuan untuk mengklasifikasi penyakit diabetes menggunakan machine learning. Algoritma MLP (Multilayer Perceptron) digunakan dalam eksperimen ini untuk mendeteksi diabetes pada tahap awal sehingga dapat memberikan peringatan kepada pasien untuk mengontrol kondisi tubuh mereka. Data yang digunakan sebagai penelitian berasal dari UCI database yang terdiri dari 16 atribut dan 2 kelas yang menunjukkan positif dan negatif diabetes. Tahap selanjutnya dilakukan preprocessing untuk cleaning dan preparing data. Model yang dikembangkan untuk klasifikasi menggunakkan algoritma MLP. Arsitektur yang dirancang dalam penelitian ini neuron input 16 berasal dari atribut, neuron hidden dilakukan 4 variasi neuron output 2 yaitu positif dan negatif. Variasi neuron hidden layer terdiri: tipe 'a', tipe 't', tipe 'i' dan tipe 'o'. Hasil pengujian yang dilakukan menggunakan data training sebanyak 520 data dan dilakukan pengujian (testing) dengan data yang sama. Hasil akurasi yang didapatkan sebesar 98.85%, 98.85%, 99.42% dan 98.46% untuk tipe 'a', 't', 'i', dan tipe 'o' secara beurutan.

Kata kunci*—Diabetes, Multilayer Perceptron, Machine Learning*

Abstract

In 2021, a total of 531 million people worldwide were affected by diabetes, with 90% of them being type 2. Diabetes can become a comorbidity for other diseases such as kidney and heart diseases. This research aims to classify diabetes disease using machine learning. The MLP (Multilayer Perceptron) algorithm is used in this experiment to detect diabetes at an early stage, thus providing warnings to patients to control their body conditions. The data used for the research is sourced from the UCI database, consisting of 16 attributes and 2 classes indicating positive and negative diabetes. The next stage involves preprocessing for data cleaning and preparation. The model developed for classification uses the MLP algorithm. The architecture designed in this research consists of 16 input neurons derived from attributes, 4 variations of hidden neurons, and 2 output neurons indicating positive and negative classes. The variations of the hidden layer neurons are: 'a' type, 't' type, 'i' type, and 'o' type. Testing was performed using 520 training data and then tested with the same data. The accuracy results obtained were 98.85%, 98.85%, 99.42%, and 98.46% for the 'a', 't', 'i', and 'o' types respectively.
Keywords—*Diabetes, Multilayer Perceptron, Machine Learning*

1. INTRODUCTION

In 2021, 531 million people worldwide were affected by diabetes. Type 1 diabetes is caused by the pancreas breaking down cells for the production of the hormone insulin. Thus, insulin cannot be produced, and requires external intake such as insulin injections. While type 2

Received June 1st,2012; Revised June 25th, 2012; Accepted July 10th, 2012

diabetes is caused by the body's insulin that is not fulfilled as a result of the pancreas gland that works not optimally. This resulting in insulin not functioning optimally. Type 1 sufferers are mostly children and adolescents. While type 2 is mostly suffered by adults. The current condition is 90% of diabetes suffered from type 2. People with diabetes have characteristics such as easy thirst, drastic weight loss, blurred vision, lack of energy, and fatigue. In addition, obesity, age, ethnicity and family history can also be factors that trigger diabetes. This disease is usually not detected quickly, the average patient knows they have diabetes for a long time. [6]. Diabetes can be a comorbid disease for other diseases. Kidney, heart, brain diseases are among the comorbid diseases that result from diabetes [7][8]. Rapid development of science, disease classification can be done with machine learning [9][10]. Classification is needed to assist clinical treatment decisions, stimulate research into aetiopathology, provide a basis for epidemiological studies [11].

Machine learning (ML) is a branch of science in the field of statistics that has algorithms to classification an object. ML is divided into 3 types, supervised learning, unsupervised learning and reinforcement learning [12]. The development of reinforcement learning in ML was then developed into deep learning. Algorithms used in classification include Naïve Bayes (NB), Support Vector Machine, K-Nearest Neighbor, MLP, and MLP [8][13] [14]. These algorithms have often been used for classification, prediction and clustering. These algorithms can be used because they are able to manage large amounts of data and integrate in various systems [15]. The MLP algorithm works like a human neural network. MLP has 3 layers, namely input layer, hidden layer, output layer.

Research on diabetes classification has been conducted by several previous researchers. Research [1] utilized the Naïve Bayes algorithm, SVM, and Decision Tree for diabetes classification. The study data were sourced from the Pima Indian Diabetes Database (PIDD), comprising 768 individuals. The accuracy levels achieved by the Naïve Bayes algorithm were 76.30%, by SVM were 65.10%, and by Decision Tree were 73.82%. Additionally, Researcher [2] conducted research to classify diabetes using logistic regression algorithms, SVM, KNN, Naïve Bayes, Decision Tree, and Random Forest. The research dataset included 952 individuals aged 18 years and above, with 580 male participants and 372 female participants. The accuracy results for the logistic regression algorithm were 85.70%, for KNN were 77.30%, for SVM were 86.50%, for Naïve Bayes were 80.60%, for Decision Tree were 84.00%, and for Random Forest was 94.10%. Furthermore, research [3] classified diabetes mellitus using a dataset from Sylhet Hospital, employing various algorithms such as NB, LR, DT, RF, SVM, KNN, ANN, and XGBoost. The average result obtained from this research was 95.16%. Research conducted by [4] on diabetes classification utilized decision trees (C4.5). The research dataset, comprising 520 data points, was obtained from Sylhet Hospital. There are a total of 16 attributes, with 13 used for classification. The results obtained with the Decision Tree algorithm were 90.38%.

Based on the research that has been carried out, experimental innovations will be carried out with the MLP algorithm model. The classification of diabetes will be processed using Weka software [16]. The hidden layer in the MLP algorithm is the main parameter in determining system performance. The performance of the hidden layer is determined by the number of neurons in the hidden layer [17]. Changes in these parameters will affect the working results of the system [18]. The best results will be obtained by using the appropriate parameters. By conducting experiments in this study, it is hoped that it will be able to find the right model in classifying diabetics.

2. METHODS

2.1 Model Diagram

Research on 'Optimal Feature Selection in Diabetes Classification' is depicted in Figure 1. The initial stage of the research involves collecting datasets. The dataset format is .*csv. The next stage is preprocessing, which is carried out to clean and prepare the data. Data cleaning is used to remove data that does not comply with the specifications. Data preparation involves data that has undergone the cleaning process and is ready to be used for system testing.

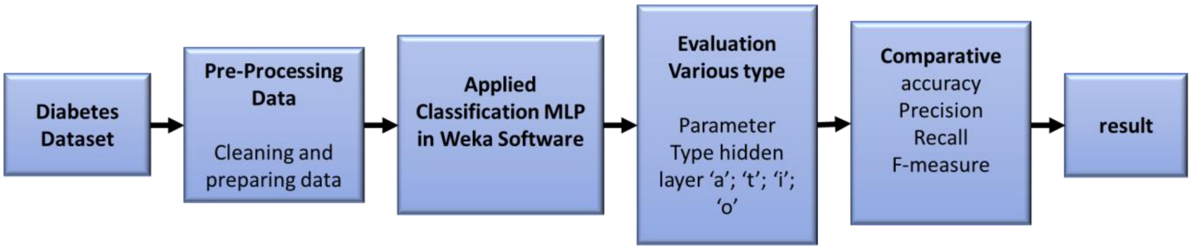

Figure 1. Classification Process

Data that has undergone the data preprocessing stage will proceed to the classification process. The classification process utilizes the Multilayer Perceptron (MLP) algorithm, which falls under the neural network algorithm type. This research employs the Weka application for the classification process. Weka is software that incorporates various machine learning algorithms. It offers several features, including classification, regression, and clustering. Various algorithms can be applied using this application, such as MLP, Naïve Bayes, SVM, and K-NN. Weka provides a variety of settings with diverse functions for each parameter [5].

The results of the classification will be evaluated for performance using a confusion matrix. At this stage, the results will be compared based on accuracy, precision, recall, and F-measure derived from the confusion matrix table. Through this comparison, the optimal outcomes from the classification utilizing the MLP algorithm will be determined.

2. 2 Multilayer Perceptron

Neural Network is a method in artificial intelligence (AI) that teaches computers to process data. The inspiration for the neural network method is like the human brain. This system uses neurons that are interconnected in a layered structure that resembles the human brain.[19]. A neural network is an adaptive system used by computers to learn from their mistakes and improve continuously [20].

The development of neural networks is currently undergoing rapid changes. Multilayer perceptron (MLP) is modeling in neural networks with better weight value characteristics compared to other models. MLP is a development of the perceptron algorithm. The perceptron algorithm has one hidden layer, while the MLP has several hidden layers between the input and output [21][22].

MLP is characterized by having 3 layers, namely the input layer, hidden layer and output layer. The input layer is a layer that functions as vector data input. Hidden layer is a layer that is between input and output. The output layer is a representation of the vector results [23][24].

MLP implementation in Weka uses backpropagation which is used as a classification. Basically, the backpropagation algorithm is a partial derivative of the cost function with weights on each neuron $\frac{\partial C}{\partial w}$. The quadratic form of the cost function is as follows:

$$C = \frac{1}{2n}\sum_{x} ||y(x) - a^{L}(x)||^{2} \tag{1}$$

Based on the equation, the variable n is the total number of trainings; the x variable is the sum of the training; y = y(x) is the desired output; L is the number of layers in the neuron; $a^{L} = a^{L}(x)$ is the vector of the network activation output when x is the input. After the algorithm runs the first iteration, the error in the output layer δ^{L} is calculated based on the equation:

$$\delta_{j}^{L} = \frac{\partial C}{\partial a_{j}^{L}}\sigma'(z_{j}^{L}) \tag{2}$$

After the error on the output layer, then calculate the error on the hidden layer. Error equation δ^{l} on the next layer δ^{l+1} That is:

$$\delta^{l} = ((w^{l+1})^{T}\delta^{l+1}).\sigma'(z') \tag{3}$$

After calculating all errors, gradient descent is used to minimize the cost function and get the optimal solution to the system [25][26].

2. 3 Evaluating System

Accuracy calculations need to be done to ensure the system runs well. The results of the accuracy calculation are influenced by the training data and the algorithm used. The following criteria are used to assess accuracy[27][28][29]. Calculation of accuracy in this system is obtained based on the equation:

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \tag{4}$$

to find out the percentage of the actual class with all classes:

$$Precision = \frac{TP}{FP + TP} \tag{5}$$

to find out the percentage of a class to the whole class:

$$Recall = \frac{TP}{FN + TP} \tag{6}$$

to find out the average of precision and recall using the equation:

$$F - Measure = 2 * \frac{precision * recall}{precision + recall} \tag{7}$$

TP = True Positive: positive predictive value with positive result
TN = True Negative: positive predictive value with negative result

FP = False Positive: negative predictive value with negative result
FN = False Negative: negative predictive value with positive result

3. RESULTS AND DISCUSSION

The initial stage of the research is to collect data. The dataset was obtained from Sylhet Hospital, Bangladesh. The amount of data is 520 data. Data pre-processing is done to select data, separate data so that the data is ready for system testing. The results of preprocessing obtained 520 data to be tested. The number of attributes in the data is 16 attributes with a total of 2 classes. The age range of people with diabetes ranges from 20 years to 65 years consisting of men and women.

Table 1. Research Database

No	1	2	3	…	…	…	518	519	520
Age	40	58	41	…	…	…	58	32	42
Gender	Male	Male	Male				Female	Female	Male
Polyuria	No	No	Yes	…	…	…	Yes	No	No
Polydipsia	Yes	No	No	…	…	…	Yes	No	No
Sudden Weight Loss	No	No	No	…	…	…	Yes	No	No
Weakness	Yes	Yes	Yes	…	…	…	Yes	Yes	No
Polyphagia	No	No	Yes	…	…	…	Yes	No	No
Genital Thrush	No	No	No	…	…	…	No	No	No
Visual Blurring	No	Yes	No	…	…	…	Yes	Yes	No
Itching	Yes	No	Yes	…	…	…	No	Yes	No
Irritability	No	No	No	…	…	…	No	No	No
Delayed Healing	Yes	No	Yes	…	…	…	No	Yes	No
Partial Paresis	No	Yes	No	…	…	…	Yes	No	No
Muscle Stiffness	Yes	No	Yes	…	…	…	Yes	No	No
Alopecia	Yes	Yes	Yes	…	…	…	No	Yes	No
Obesity	Yes	No	No	…	…	…	Yes	No	No
Class	Positive	Positive	Positive	…	…	…	Positive	Negative	Negative

The attributes in table 1 are symptoms for people with diabetes. The data that will be tested on the system later is 100% of the total data available. The attributes in the data are symptoms for people with diabetes. The description of the data attributes can be seen in Table 2.

Table 2. Description Dataset

NO	Attribute	Value
1.	Age	20-65
2.	Gender	Male, Female
3.	Polyuria	Yes, No
4.	Polydipsia	Yes, No
5.	sudden weight loss	Yes, No
6.	weakness	Yes, No

NO	Attribute	Value
7.	Polyphagia	Yes, No
8.	Genital thrush	Yes, No
9.	visual blurring	Yes, No
10.	Itching	Yes, No
11.	Irritability	Yes, No
12.	delayed healing	Yes, No
13.	partial paresis	Yes, No
14.	muscle stiffness	Yes, No
15.	Alopecia	Yes, No
16.	Obesity	Yes, No
17.	class	Positive, Negative

Data that has been pre-processed will then be tested using the MLP algorithm. Testing is done on the Weka application. Seen in Figure 1 Testing is done by changing the number of hidden layers. The results of changing the number of hidden layer parameters will be compared. In this study, the test uses 4 hidden layers that are available in the Weka software.

Table 3. Hidden Layer Type

Hidden Layer	**Description**	**Hidden Layer Total**
a	(attributes + classes) / 2	9
t	Attributes + classes	18
i	attributes	16
o	classes	2

Table 3 is the types of hidden layers in Weka that will be used for system testing. The code in the hidden layer comes from the attribute or class in the research data. The results of these changes will be compared to the accuracy value. Hidden layer is very influential in the MLP algorithm method. Hidden layer will affect the success rate of the system. Figure 2 shows the architecture of the MLP algorithm. The input of the MLP architecture in the system test has 16 input neurons. The number of hidden layer neurons in this study consists of 9, 18, 16, hidden layers.

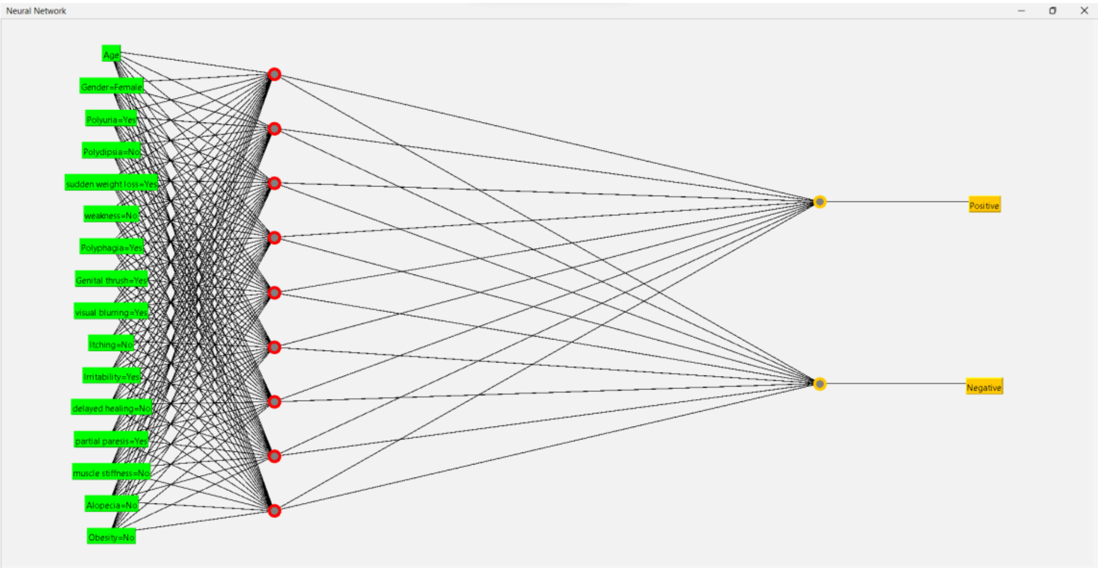

Figure 2. Architecture MLP Algorithm with Weka Software

In Figure 2 the MLP algorithm architecture in the input layer consists of 16 attributes. The attributes are as in table 2. Neurons in the hidden layer will be made several changes. The first trial with a hidden layer of 9. The second trial uses a hidden layer of 18. The third trial uses a hidden layer of 16, and the fourth trial uses a hidden layer of 2. The output layer consists of 2 neurons, namely yes and no.

Table 4. Comparison results with hidden layer variations

Hidden Layer Type	Positive	Negative	Status
a	317	3	Positive Prediction
	3	197	Negative Prediction
t	317	3	Positive Prediction
	3	197	Negative Prediction
i	320	0	Positive Prediction
	3	197	Negative Prediction
o	317	3	Positive Prediction
	5	195	Negative Prediction

Table 4 is the confusion matrix of the research results. There are positive and negative columns and rows of positive predictions and negative predictions. Based on the results of the confusion matrix in table 4, the accuracy value of each type of hidden layer is known.

Table 5. MLP Classification Results with Hidden Layer Type 'a'

Class	Hidden Layer	TP Rate	FP Rate	Precision	Recall	F-Measure	Accuracy (%)
Positive	a	0.991	0.015	0.991	0.991	0.991	98.85
Negative		0.985	0.009	0.985	0.985	0.985	
Mean		0.988	0.013	0.988	0.988	0.988	

Table 5 shows the results of the MLP algorithm on hidden layer type 'a'. Hidden layer in this type has an average TP rate of 0.991 and an average FP rate of 0.013. The precision, recall and f-measure values are 0.988. Hidden layer type 'a' has an accuracy of 98.85%.

Table 6. MLP Classification Results with Hidden Layer Type 't'

Class	Hidden Layer	TP Rate	FP Rate	Precision	Recall	F-Measure	Accuracy (%)
Positive	t	0.991	0.015	0.991	0.991	0.991	98.85
Negative		0.985	0.009	0.985	0.985	0.985	
Mean		0.988	0.013	0.988	0.988	0.988	

Table 6 is the result of the hidden layer in type 't'. This type has the same average TP rate as type 'a' which is 0.988, average FP rate '0.013'. The precision, recall and f-measure values are 0.988. The accuracy of type 't' is 98.85%. The accuracy value has not changed from before.

Table 7. MLP Classification Results with Hidden Layer Type 'i'

Class	Hidden Layer	TP Rate	FP Rate	Precision	Recall	F-Measure	Accuracy (%)
Positive	i	1.000	0.015	0.991	1.000	0.995	99.42
Negative		0.985	0.000	1.000	0.985	0.992	
Mean		0.994	0.009	0.994	0.994	0.994	

Table 7 has a different accuracy value from table 4. In table 5, the accuracy value has changed, so the accuracy rate for this type of hidden layer is 99.42%. The average TP rate in type 'i' is 0.994, the average FP rate is 0.013. The precision, recall and f-measure values are 0.994.

Table 8. MLP Classification Results with Hidden Layer Type 'o'

Class	Hidden Layer	TP Rate	FP Rate	Precision	Recall	F-Measure	Accuracy (%)
Positive	o	0.991	0.025	0.984	0.991	0.988	98.46
Negative		0.975	0.009	0.985	0.975	0.980	
Mean		0.985	0.019	0.985	0.985	0.985	

The results of table 7 have decreased in accuracy when compared to the results in tables 4, 5, and 6. The average TP rate on type 'o' is 0.985, the average FP rate is 0.019. The precision, recall and f-measure values are 0.985. The accuracy rate for this type has a value of 98.46%.

Based on these results, we can see that the hidden layer in the MLP algorithm influences the accuracy level of the research results. The 'i' type hidden layer has the highest accuracy level compared to other types. Testing the system using the 'i' type hidden layer in this study was able to achieve the best accuracy, at 99.42%. The number of neurons in this variation is 16. The 't' type has more neurons compared to the 'i' type, but the output of this type is not better than the 'i' type. Neurons in the 'a' type are fewer than in the 't' type, and the results in both have the same accuracy of 98.85%. The 'o' type has the lowest output results compared to other types, which is 98.46%. There are 2 neurons in this type. The hidden layer neurons greatly affect the results of the MLP algorithm. Accuracy in determining the number of neurons can produce the best output. Overall, the average accuracy obtained from the MLP algorithm is 99.04%. The results are better compared to previous research. This algorithm can be used for early-stage classification of diabetes disease, thus enabling the early prevention of diabetes.

5. CONCLUSIONS

The research used 520 data, with system testing using 100% of the data. Testing is carried out by changing the number of hidden layers in the MLP. Changes in the number of hidden layers in the MLP are based on the type in the Weka software. There are 4 types of Weka software, namely, types a, i, o and t. Changes to these 4 types produce different results. Type a produces an accuracy of 98.85%. Type i produces an accuracy of 99.42%. In type o it produces 98.46%. The t type produces an accuracy of 98.85%. Hidden layer type 'i' has the greatest accuracy compared to other hidden layers.

REFERENCES

[1] D. Sisodia and D. S. Sisodia, "Prediction of Diabetes using Classification Algorithms," *Procedia Comput. Sci.*, vol. 132, no. Iccids, pp. 1578–1585, 2018, doi: 10.1016/j.procs.2018.05.122.

[2] N. P. Tigga and S. Garg, "Prediction of Type 2 Diabetes using Machine Learning Classification Methods," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 706–716, 2020, doi: 10.1016/j.procs.2020.03.336.

[3] A. Yilmaz, "Prediction of Type 2 Diabetes Mellitus Using Feature Selection-Based Machine Learning Algorithms," *Heal. Probl. Civiliz.*, vol. 16, no. 2, pp. 128–139, 2022, doi: 10.5114/hpc.2022.114541.

[4] B. A. C. Permana, R. Ahmad, H. Bahtiar, A. Sudianto, and I. Gunawan, "Classification of diabetes disease using decision tree algorithm (C4.5)," *J. Phys. Conf. Ser.*, vol. 1869, no. 1, 2021, doi: 10.1088/1742-6596/1869/1/012082.

[5] M. A. Karan Manchandia*, Navdeep Khare, "Weka As a Data Mining Tool To Analyze Students' Academic PERFORMANCES USING NAÏVE BAYES CLASSIFIER- A SURVEY," *Int. J. Eng. Sci. Res. Technol.*, vol. 6, no. 3, pp. 431–434, 2017, doi: 10.5281/zenodo.438104.

[6] M. Diana, E. Boyko, and I. Diabetes, *International Diabetes Federation*, vol. 10, no. 2. 2021.

[7] K. Kannadasan, D. R. Edla, and V. Kuppili, "Type 2 diabetes data classification using stacked autoencoders in deep neural networks," *Clin. Epidemiol. Glob. Heal.*, vol. 7, no. 4, pp. 530–535, 2019, doi: https://doi.org/10.1016/j.cegh.2018.12.004.

[8] S. A. Paschou, G. I. Sydney, K. J. Ioakim, K. Kotsa, and D. G. Goulis, "Comment on the systematic review and meta-analysis titled 'Gestational diabetes and the risk of cardiovascular disease in women,'" *Hormones*, vol. 19, no. 3, pp. 447–448, 2020, doi: https://doi.org/10.1007/s42000-019-00158-w.

[9] A. Pradhan, S. Prabhu, K. Chadaga, S. Sengupta, and G. Nath, "Supervised Learning Models for the Preliminary Detection of COVID-19 in Patients Using Demographic and Epidemiological Parameters," *Inf.*, vol. 13, no. 7, pp. 1–28, 2022, doi: 10.3390/info13070330.

[10] R. Cheheltani *dkk.*, "Predicting misdiagnosed adult-onset type 1 diabetes using machine learning," *Diabetes Res. Clin. Pract.*, vol. 191, p. 110029, 2022, doi: 10.1016/j.diabres.2022.110029.

[11] World Healt Organization, *Classification of diabetes mellitus*. 2019.

[12] M. Batta, "Machine Learning Algorithms - A Review," *Int. J. Sci. Res. (IJ*, vol. 9, no. 1, pp. 381–386, 2020, doi: 10.21275/ART20203995.

[13] A. Iyer, J. S, and R. Sumbaly, "Diagnosis of Diabetes Using Classification Mining Techniques," *Int. J. Data Min. Knowl. Manag. Process*, vol. 5, no. 1, pp. 01–14, 2015, doi: 10.5121/ijdkp.2015.5101.

[14] A. V. Srinivas, A. Ramya, G. T. Chandralekha, B. Vaagdevi, and K. Anand Goud, "Prediction of Diabetes Using Machine Learning," *Ymer*, vol. 21, no. 5, pp. 485–492, 2022, doi: 10.37896/YMER21.05/54.

[15] I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, "Machine Learning and Data Mining Methods in Diabetes Research," *Comput. Struct. Biotechnol. J.*, vol. 15, pp. 104–116, 2017, doi: 10.1016/j.csbj.2016.12.005.

[16] E. G. Kulkarni and R. B. Kulkarni, "WEKA Powerful Tool in Data Mining General Terms," *International Journal of Computer Applications*, vol. 5, no. Rtdm, pp. 975–8887, 2016.

[17] I. N. A. Suprana, "Pengaruh Perbedaan Jumlah Hidden Layer dan Node pada Hidden Layer terhadap Performa Model Klasifikasi Diabetes," *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 2, no. 2, pp. 411–419, 2022, doi:

10.24002/konstelasi.v2i2.5330.

[18] M. K. Hasan, M. A. Alam, D. Das, E. Hossain, and M. Hasan, “Diabetes prediction using ensembling of different machine learning classifiers,” *IEEE Access*, vol. 8, pp. 76516–76531, 2020, doi: 10.1109/ACCESS.2020.2989857.

[19] A. S. Musliman, A. Fadlil, and A. Yudhana, “Identification of White Blood Cells Using Machine Learning Classification Based on Feature Extraction,” *J. Online Inform.*, vol. 6, no. 1, pp. 63–72, 2021, doi: 10.15575/join.v6i1.704.

[20] M. Fatima and M. Pasha, “Survey of Machine Learning Algorithms for Disease Diagnostic,” *J. Intell. Learn. Syst. Appl.*, vol. 09, no. 01, pp. 1–16, 2017, doi: 10.4236/jilsa.2017.91001.

[21] Y. S. Park and S. Lek, “Artificial Neural Networks: Multilayer Perceptron for Ecological Modeling,” in *Developments in Environmental Modelling*, vol. 28, Elsevier, 2016, pp. 123–140.

[22] Z. Car, S. B. Šegota, N. Anđelić, I. Lorencin, and V. Mrzljak, “Modeling the Spread of COVID-19 Infection Using a Multilayer Perceptron,” *Comput. Math. Methods Med.*, vol. 2020, p. 5714714, 2020, doi: 10.1155/2020/5714714.

[23] I. Lorencin, N. Andelić, V. Mrzljak, and Z. Car, “Genetic algorithm approach to design of multi-layer perceptron for combined cycle power plant electrical power output estimation,” *Energies*, vol. 12, p. 4352, 2019, doi: 10.3390/en12224352.

[24] Y. Ghanou and G. Bencheikh, “Architecture optimization and training for the multilayer perceptron using ant system,” *IAENG Int. J. Comput. Sci.*, vol. 43, no. 1, pp. 20–26, 2016.

[25] A. Music and S. Gagula-Palalic, “Classification of Leaf Type Using Multilayer Perceptron, Naive Bayes and Support Vector Machine Classifiers,” *Southeast Eur. J. Soft Comput.*, vol. 5, no. 2, pp. 16–20, 2016, doi: 10.21533/scjournal.v5i2.119.

[26] M. Nielsen, “How the backpropagation algorithm works,” 2019. http://neuralnetworksanddeeplearning.com/chap2.html (accessed Nov. 11, 2023).

[27] A. H. Meftah, Y. A. Alotaibi, and S.-A. Selouani, “Evaluation of an Arabic Speech Corpus of Emotions: A Perceptual and Statistical Analysis,” *IEEE Access*, vol. 6, pp. 72845–72861, 2018, doi: 10.1109/ACCESS.2018.2881096.

[28] V. R. Balaji, S. T. Suganthi, R. Rajadevi, V. Krishna Kumar, B. Saravana Balaji, and S. Pandiyan, “Skin disease detection and segmentation using dynamic graph cut algorithm and classification through Naive Bayes classifier,” *Meas. J. Int. Meas. Confed.*, vol. 163, pp. 1–14, 2020, doi: 10.1016/j.measurement.2020.107922.

[29] V. Doma and M. Pirouz, “A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals,” *J. Big Data*, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00289-7.