Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control
Journal homepage: http://kinetik.umm.ac.id
ISSN: 2503-2267
Vol. 4, No. 4, November 2019, Pp. 315-324

Performance improvement of non invasive blood glucose measuring system with near infra red using artificial neural networks

Rizaldi Ramdlani Pamungkas*[1], Aji Gautama Putrada[2], Maman Abdurohman[3]
Telkom University, Indonesia[1,2,3]

Article Info

Keywords:
Diabetes, ANN, IoT, Glucose, Blood

Article history:
Received 29 May 2019
Revised 17 June 2019
Accepted 5 October 2019
Published 30 Ocotber 2019

Cite:
Pamungkas, R., Putrada, A., & Abdurohman, M. (2019). Performance Improvement of Non Invasive Blood Glucose Measuring System With Near Infra Red Using Artificial Neural Networks. Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control, 4(4). doi:http://dx.doi.org/10.22219/kinetik.v4i4.844

*Corresponding author.
Rizaldi Ramdlani Pamungkas
E-mail address:
rizaldipamungkas@student.telkomuniversity.ac.id

Abstract

Measurement of body blood sugar levels is one of the important things to do to reduce the number of people with diabetes mellitus. Non-invasive measurement techniques become a blood sugar measurement technique that is more practical when compared to invasive techniques, but this technique has not shown too high levels of accuracy, specificity and sensitivity. For this reason, the non-invasive measurement model using NIR and ANN is proposed to improve the performance of non-invasive gauges. Non-invasive blood sugar measuring devices will be built using a nodemcu board with photodiaodes and NIR transmitters whose data is then processed using ANN models compared to invasive blood sugar data obtained from 40 data. 40 data obtained then used as raw data to build ANN models which 75% percent of it use as training data and 25% od it will be use as testing data to validate accuration of the model been built, the split of data doing randomly without any interference from programmer or model designer. All the data gathered are data collected from all volunteers which willingly to test their blood glucose using invasive glucose meter and non invasive glucose meter which been built. The invasive glucose meter used to gather raw data of blood glucose is SafeAccu-2 with 95% level of accuracy so the accuracy and error parameter calculated in this research are based on that 95% level accurcy of the invasive device.

1. Introduction

In Indonesia according to the International Diabetes Federation as of 2017 there are as many as 10,276,000 people who are indicated to have diabetes. Of all the countries listed in IDF Indonesia was in No. 6 as the country with the highest number of Diabetes sufferers [1].

One of the best efforts to suppress the increase in diabetics is to prevent it. Preventive measures that can be taken include knowing blood sugar levels in the body, so far the blood sugar checking technique that is carried out is by invasive or injuring the body to get the blood to be checked. When viewed from the side of practicality, invasive techniques are not too easy to do, while checking blood sugar is a process that should be done regularly.

Due to the ineffectiveness of invasive techniques, non-invasive techniques for detecting blood sugar have emerged, this non-invasive technique is already quite widely used by implementing various types of calculation techniques.[2][3][4] and the most widely used technique is linear regression techniques [5][6]. Detection results produced through non-invasive devices do not have a high degree of accuracy and the level of error margins of non-invasive devices is still quite large.

To increase the accuracy of the non-invasive blood sugar detection techniques the author will implement the technique of detecting non-invasive blood sugar through the NIR (Near Infra-Red) specstropy [4][7][8][9] using ANN(Artificial Neural Network) [10][11]. This technique will be implemented on the nodeMCU microcontroller [12] which is equipped with a photodioda sensor and an NIR transmitter that will be analyzed and compared with the level of accuracy and error with a technique that has been widely used, namely non-invasive techniques with linear regression.

To build a model of ANN and compared with the linear regression model then takes the raw data, the raw data used to build the model is obtained by taking the data of sugar obtained from the detection instrument blood sugar invasive brand SafeAccu-2 as a feature output and the value derived from detection devices non-invasive blood sugar that uses NIR with a wavelength of 850nm as an input feature. The input features obtained from noninvasive devices will then be extracted and converted into 4 features which include the average value, the highest value, the lowest value, and the standard deviation of the electric wave received from the infrared photodiode that penetrates the ring finger meat.

2. Research Metodology

In general, seen in Figure 1, the tool that will be made will be built using a nodemcu board which is accompanied by an ESP8266 wireless communication module in it then combined with an NIR LED [8] with different wavelengths of light ranging from 850nm, 940nm and 1550nm. The light emitted by the NIR LED will be emitted to the hand ring finger and the light that successfully penetrates the skin will be detected by the potodiode sensor which is also connected to the nodemcu board which later the data will be transferred to the server and processed and the calculation results will be displayed on the LCD screen.

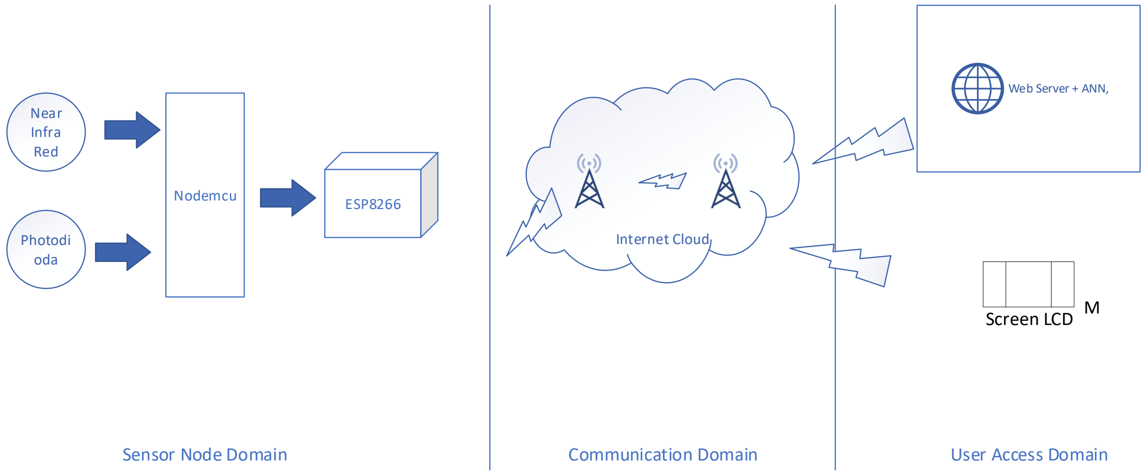

Figure 1. Block Diagram of System

The shape of the device that will be used to scatter and receive the beam of light that has been through the flesh of the body will be shaped like a hole where the body part will be used as a measurement object in which case the ring finger is inserted into it.

Different types of LEDs with different non wavelengths are intended to be used as analysis to compare the results of data reception because with different NIR wavelengths the sensitivity of data reception will also vary. In Figure 2 shows the general design of the tool that will be built later in which there are several components.

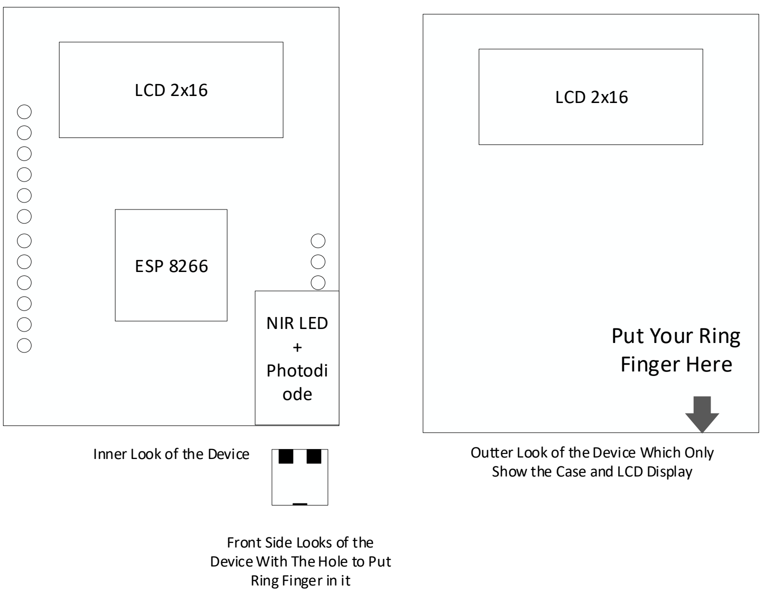

Figure 2. Device's Design

The components that can be seen in Figure 2 include the main components, namely the motherboard that has been integrated with the ESP8266 wireless communication module, then there is a component to display information to the user, namely the 2x16 LCD module, and the last component which is the most fundamental component of this tool namely photodiode and NIR sensors which can be seen in the front side cross section of the tool.

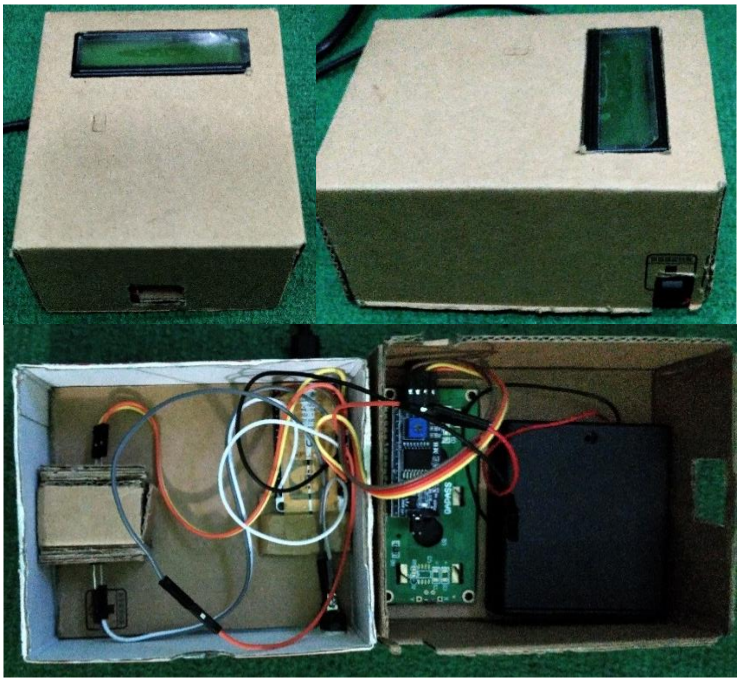

Figure 3. Picture of Non-Invasive Blood Sugar Measuring Instrument

Figure 3 displays an image of the glucose meter tool that has been assembled with the image at the top showing the outside display of the device that shows the 2x16 LCD screen, the hole for inserting the ring finger and the power button. At the bottom of the picture shows the part in the image which consists of microelectric components such as photodioda, near infrared, nodemcu, lcd board with adapter, battery holder, power supply and connecting cables.

Overall the tools built will not be too complicated and seem simple with only the LCD screen that will be displayed and instructions for entering the pinkie. It is expected that a simple design can facilitate its use later. In the description of the tool design shown in Figure 2, the user can only use the ring finger to be scanned on the tool.

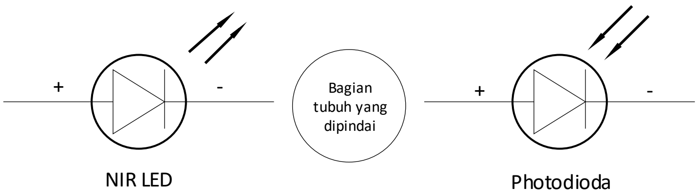

Figure 4. Photodioda and NIR Sensor Diagram

The sensor used in this device is a photodiode sensor which is combined with an NIR transmitter. In general, the design of sensors and NIRs used can be seen in Figure 4.

For testing of the equipment built, it will be carried out on approximately 40 people who will undergo a non-invasive blood sugar test right before finally testing invasive blood sugar, so that the body's chemical state does not change too much so that the test parameters can be compared to get more accurate comparison results. Of the 40 data obtained from the test shown in Table 1, 25% of the data will be used as test data and the remaining 75% will be used as training data.

Table 1. Test and Training Data Split Schemes

Data Type	Number of Data	Percentage
Train Data	30	75%
Test Data	10	25%
Total	40	100%

Using a 1: 3 ratio, in the case of 40 data, the first 30 data received will be used as training data and the last 10 data entered will be used as test data to measure the accuracy and mean squared error of the calculation algorithm.

3. Results and Discussion

In carrying out the testing, researchers conducted invasive blood sugar data collection using a Safe-Accu 2 blood sugar detection device with 95% accuracy in the device. In addition to taking invasive blood sugar data, raw blood sugar data is also obtained through NIR with a wavelength of 850nm because the wavelength is the largest wave of light absorbed by blood and water.

3.1 Data

The raw data obtained from the results of non-invasive measurements consists of 4 features obtained from the voltage level received by the potodiode. The four features include the average voltage (mean), standard voltage deviation (deviation), the lowest value of the voltage (low), and the highest value of the voltage (peak). The four data features will be used as input features to be trained in data with the output feature is the actual value obtained from the search results of data through the detection of invasive blood sugar.

Features such as mean and deviation are formed based on the basic formula of statistics. The equation to get the mean feature which is the average [13] can be seen in Equation 1.

$$\bar{x} = \frac{\sum_i^n x_i}{n} \tag{1}$$

Where:
$\bar{x}$ = Mean
x_i = Data x order of i
n = Sum of data

Next to get the deviation feature is formed using a standard deviation formula [14] which can be seen in the Equation 2.

$$s = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}} \tag{2}$$

Where:
s = Standard deviation
$\bar{x}$ = Mean
x_i = Data x order of i
n = Sum of data

The model formation process will take a combination of several features from the four input features, there will be a model that only uses one of the input features, until there is a model that uses all input features and all models that have been built will be compared and searched for the most optimum performance .

The feature selection process is based also on the correlation between features calculated using the Pearson correlation coefficient [15]. Formula to calculate the amount of Pearson correlation coefficient between x and y shown in the Equation 3.

$$\mathrm{r} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{3}$$

Where:
r = Pearson correlation coefficient
x_i = Data x order of i
y_i = Data y order of i
$\bar{x} = \frac{\sum x}{n}$
$\bar{y} = \frac{\sum y}{n}$
n = Sum of data

The results of the calculation of Pearson correlation show the relationship between variables in the data set. The results of the calculation of Pearson correlation from raw blood sugar data can be seen in Figure 5.

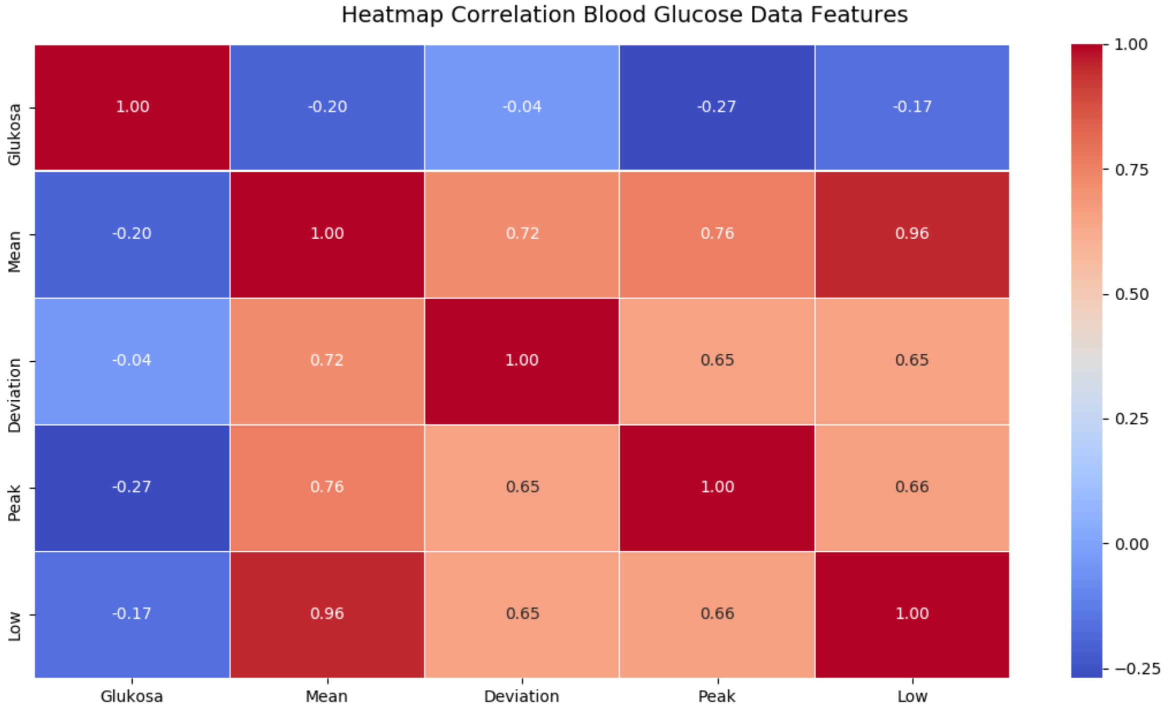

Figure 5. Heatmap Correlation

The results of the Pearson correlation calculation show the bond relationship between variables in the data set. The results of the calculation of Pearson correlation from raw blood sugar data can be seen in Figure 5. Judging from the results of the correlation calculation, almost all the correlation values feature input to the output (glucose) feature which does not exceed 0.5 or not too strong and the deviation feature value is a feature with the smallest correlation coefficient so that the combination of feature selection for building a model only includes the Mean, Peak and Low features, and for deviations it will enter the part of the model that is built using the whole feature.

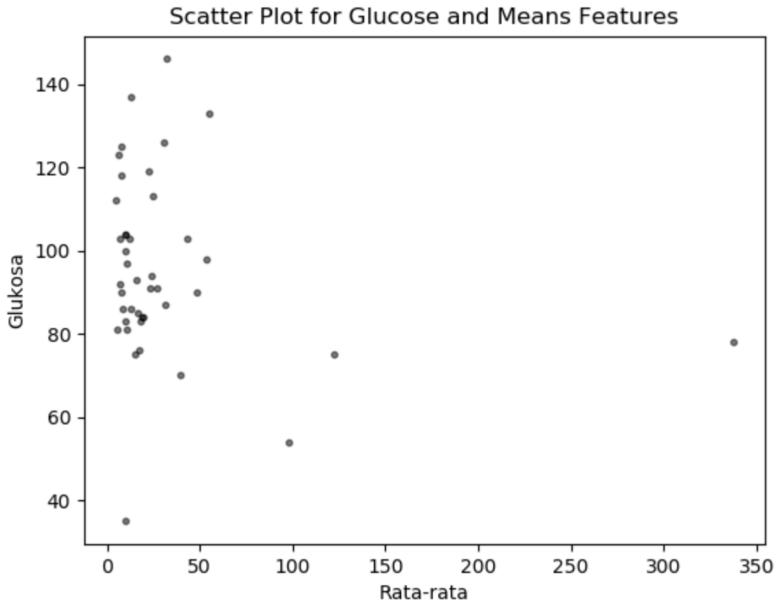

Figure 6. Plots of Glucose and Mean Features

In Figure 6 and Figure 7 show a pattern of data distribution that is more or less polarized in the upper left corner of the Cartesian coordinates. In Figure 6 the average data distribution pattern for glucose does not have a correlation that is too strong but by looking at a polarized and not very diffuse distribution pattern this feature has a fairly strong relation compared to other input features of the glucose output feature.

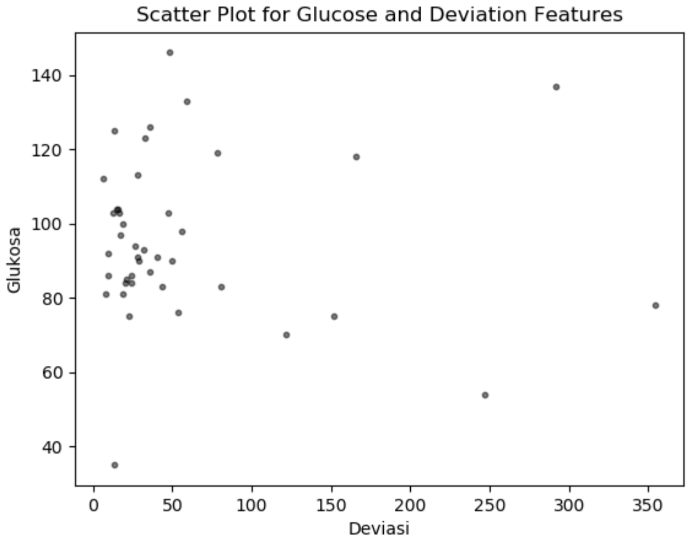

Figure 7. Plot of Glucose and Deviation Features

In the scatter plot the average features of the data are more collected and patterned which causes the correlation between the average features to have a higher relation with the glucose feature than the standard deviation.

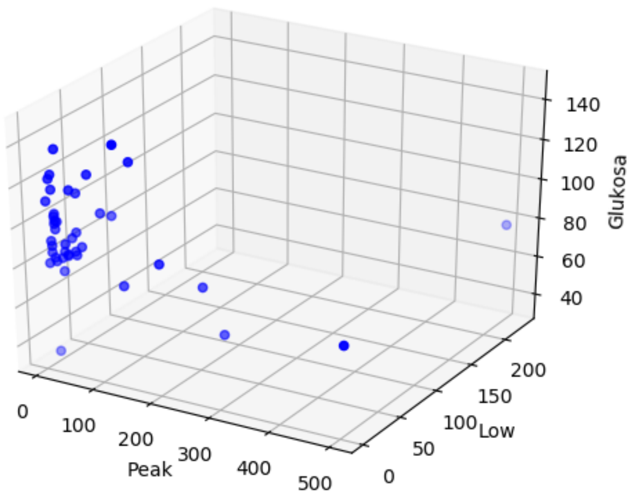

Figure 8. Feature Plots of Glucose, Peak, and Low

Next in Figure 8 shows the level of correlation between peak and low data with glucose which is generally polarized and has a fairly strong pattern that makes the low and peak features strong enough to be used as input features for model training.

3.2 Test Results Using Artificial Neural Networks and Linear Regression

ANN model is built using 3 layers of nerves with the first layer consisting of 12 input neurons then in the second layer consists of 8 hidden layer neurons with ReLu activation function (Rectified Linear Unit) [16] and in the release layer there is a neuron output with linear activation function [17]. The model was trained to use 32 training data and 9 validation data with 150 epochs with a batch size of 50 and the loss function and optimizer used were MSE and Adam [18]. The model is built with a number of diverse features, the model with one feature uses the mean feature, the 2 feature model is built from peak and low features as input and the last model is built using all the features in the data set of 4 can be seen in Table 2.

Table 2. ANN Model Test Results

Volunteer Number	Age	Sex	Glucose(mg/dl)	Prediction (mg/dl)		
				1 Feature	2 Features	4 Features
1	19	M	119	96,63	101,81	101
2	19	M	103	98,53	102,58	100
3	22	F	133	92,69	95,93	92,7
4	42	M	118	98,49	102,68	112
5	22	M	97	98,12	101,32	97,9
6	23	M	91	96,13	96,55	96,4
7	23	M	104	98,21	101,86	98,4
8	20	M	84	97,07	99,69	96
9	26	M	86	98,39	102,2	98,3
10	21	M	84	97,03	100,23	96,7
11	19	M	87	95,62	100,13	95,4
12	23	M	103	94,13	94,95	91
13	20	M	94	96,48	97,96	94,5
14	22	M	97	98,12	101,32	97,9

In Table 3 show in to construct a linear regression model the training data structure and the test data applied are the same as in the ANN model as well as many and the types of features used to build it.

Table 3. Results of Testing the Linear Regression Model

Volunteer Number	Sex	Glucose(mg/dl)	Prediction(mg/dl)		
			1 Feature	2 Features	4 Features
1	M	119	97,47	96,63	100,96
2	M	103	98,6	97,56	95,39
3	F	133	95,14	93,26	93,31
4	M	118	98,58	97,75	108,3
5	M	97	98,36	97,7	95,38
6	M	91	97,17	90,95	86,88
7	M	104	98,41	97,71	95,66
8	M	84	97,73	97,04	94,48
9	M	86	98,52	97,85	95,39
10	M	84	97,71	96,98	95,2
11	M	87	96,87	96,48	97,06
12	M	103	95,98	95,34	93,31
13	M	94	97,38	96,66	93,55
14	M	97	98,36	97,7	95,38

ANN model testing was conducted on 14 test data from a total of 41 data obtained. 14 data used as tests were obtained randomly. The results of the tests show 3 types of prediction results from 3 types of ANN models.

To calculate the performance of the built model used relative accuracy parameters [19] and root mean squared error (RMSE) [20]. The formula for calculating the relative accuracy of each result is used Equation 4.

$$Accuracy_{relative} = \frac{Actual\ Value - |Predicted\ Value - \ Actual\ Value|}{Actual\ Value} \times 100 \tag{4}$$

Where:
$Accuracy_{relative}$ = Accuracy value in percent
$Actual\ Value$ = Actual value
$Predicted\ Value$ = Predicted value

While the formula for calculating RMSE is shown in Equation 5.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e^2} \tag{5}$$

Where:
RMSE = Rooted Mean Squared Error
e = Difference between actual value and predicted value
n = Sum of data

Using the RMSE and Relative Accuracy parameters are obtained shown in Figure 9 and Figure 10.

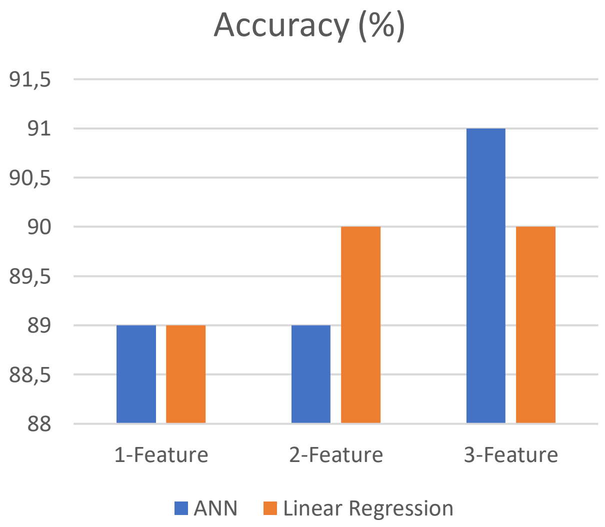

Figure 9. Inter Model Features Accuracy Chart

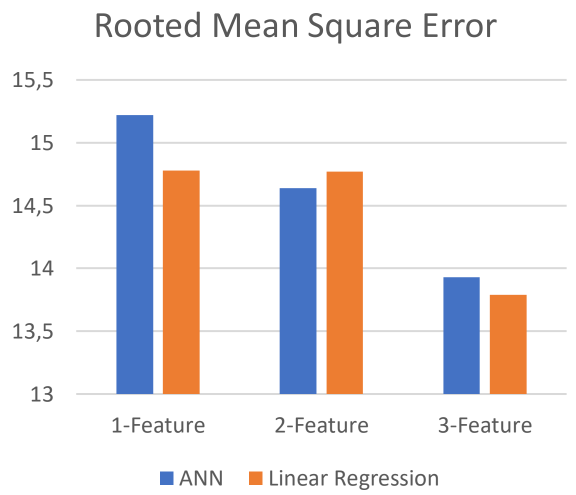

Figure 10. Inter Model Features RMSE Chart

In general, the best performance of the ANN model is obtained from the ANN model with 4 features as well as the Linear Regression model, the best level of performance of the two types of models is not too far but the ANN model is slightly superior with 91% accuracy, but not too good when seen from the RMSE level which is slightly larger than the Linear Regression model. But when viewed from the diversity of output values using the standard deviation formula the best ANN model has a higher level of data diversity of 4.93 compared to the level of diversity of the output data from the best model Linear Regression is 4.68.

4. Conclusion

After doing the research, it was found that invasive blood sugar level calculation techniques using the ANN model could have a much better performance as well as the scalability of machine learning models ANN is more flexible to adjust to the output data using various methods (loss, mse, mae, etc.). While linear regression is less able to adjust to the output data because the predicted output generated is fixed on the linear function of the model. It can be seen from the results of standard data deviation. All linear regression models have a low value compared to the results of the data output standard deviation of all ANN models.

The thing to note in building ANN models is to avoid overfiting of models that might cause prediction errors if the model is given the task of predicting data output from truly new data. More amount of data will also improve the performance of ANN models.

Acknowledgment

To all 40 volunteers who participate in this research by willingly agree to tested their blood glucose.

References

[1] T. I. D. Federation, "IDF Diabetes Atlas Eighth edition," 2017.
[2] X. A. Guo, D. X.; Shang, Y. Z.; Peng, R.; Yong, S. S.; Wang, "Noninvasive Blood Glucose Measurement," *J. Biosci. Med.*, Vol. 3, 2015. http://dx.doi.org/10.4236/jbm.2015.36007
[3] S. Dai, Juan; Ji, Zhong; Du, Yubao; Chen, "In vivo noninvasive blood glucose detection," *IOS Press*, Pp. 229–239, 2018. https://dx.doi.org/10.3233%2FTHC-174592
[4] M. M. Buda, R. A.; Addi, "A Portable Non-Invasive Blood Glucose Monitoring Device," in *Conference on Biomedical Engineering and Sciences*, 2014. https://doi.org/10.1109/IECBES.2014.7047655
[5] V. A. Barbur, D. C. Montgomery, and E. A. Peck, "Introduction to Linear Regression Analysis.," in *The Statistician*, 2006, Vol. 43, No. 2, Pp. 339.
[6] M. A. Boatemaa and S. Doss, "Non-Invasive Glucose Estimation Based on Near Infrared Laser Diode Spectroscopy," *Asian J. Biomed. Pharm. Sci.*, Vol. 7, No. 60, 2017.
[7] H. S. Kamasahayam, Swathi; K, "Non Invasive Estimation of Blood Glucose using Near Infra red Spectroscopy and Double Regression Analysis," 2013. https://doi.org/10.1109/ICSensT.2013.6727729
[8] J. M. Derrick, Michele; Stulik, Dusan; Landry, "Infrared Spectroscopy in Conservation Science (Tools for Conservation)," *1st ed. Los Angeles,* 1999.
[9] S. Menon, K. A. Unnikrishna; Gayathri, B.; K, "Non-invasive Blood Glucose Monitoring Using Near Infrared Spectroscopy," 2017. https://doi.org/10.1109/ICCSP.2017.8286555
[10] K. Gurney, "An Introduction to Neural Networks," *1st ed. Boca Raton: CRC Press,* 1997.
[11] G. I. Webb, "Encyclopedia of Machine Learning and Data Mining Second Edition," *2nd ed. New York: Springer,* 2017.
[12] *ESP8266 Technical Reference*. 2019.
[13] K. R. Tripalupi, Lulup Endah; Suwena, "Statistika Dasar," *1st ed. Makassar: Graha Ilmu.*
[14] S. YL, "Statistika Dasar," *1st ed. Yogyakarta: Andi Publisher,* 2014.
[15] M. Parsian, "Data Algorithms," *1st ed. Sebastopol: O'Reilly Media, Inc,* 2015.
[16] A. F. M. Agarap, "Deep Learning using Rectified Linear Units (ReLU)," *Manila,* 2018.
[17] F. Agostinelli, M. Hoffman, P. Sadowski, and P. Baldi, "Learning Activation Functions To Improvedeep Neural Networks," in *International Conference on Learning Representations*, 2015.
[18] S. Bock, J. ; Goppold, and M. Weiß, "An improvement of the Convergence Proof of the ADAM-Optimizer," in *Clusterkonferenz*, 2018.
[19] J. M. Chen, Chao; Twycross, Jamie; Garibaldi, "A New Accuracy Measure Based On Bounded Relative Error for Time Series Forecasting," *PLoS One*, 2017.
[20] R. R. Chai, T ; Draxler, "Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? – Arguments Against Avoiding RMSE in the Literature," *Geosci. Model Dev.*, Vol. 7, Pp. 1247–1250, 2014. https://doi.org/10.5194/gmd-7-1247-2014