

# Implementation Model Linear Regression, Neural Network and Support Vector Machine in Simple Mathematical Pendulum Experiments to Determine the Value of Gravity

Irfan Syafar Farouk, Mustafa Mamat, W. S Mada Sanjaya, Wahid Abdurahman, Hani Rubiani, Mohamad Afendee Mohamed

*machine in mathematical pendulum experiments to find the value of gravity. There were 4 data obtained from mathematical pendulum experiments which were then interpolated to obtain more data (13 data), then the data was used as training data for each model. Each model is tested to get a gravity value of 26 including training data, then compared with reference gravity values [17,18,19]. The results of the model Neural network proved to be the most accurate with an error value of 2.53%. The support vector machine model is the most accurate model with a standard deviation value of 0.03 and the error deviation of 0.058 is the smallest value of the three models in this paper.*

*Index Terms: comparison, linear regression, neural network, support vector machine, mathematical pendulum experiment and value of gravity*

## I. INTRODUCTION

Mathematical pendulum experiment is a simple experiment that simulates a point with a pendulum that is hung at a fixed point with a rope [1]. With the oscillation approach that will be explained in advance we can get the acceleration value of gravity by measuring the length of the rope and the number of oscillations every second or period. This simple experiment is often studied at school or in college, although this experiment is simple but the results of this experiment are often good with high accuracy values [2]. In addition to using ordinary mathematical calculations, the results of this experimental data are often calculated using the linear regression method.

### Revised Manuscript Received on April 25, 2019.

**Irfan Syafar Farouk**, Department of Informatic and Computation, Universiti Sultan Zainal Abidin, Terengganu, Malaysia.

**Mustafa Mamat**, Faculty of Informatic and Computing, Universiti Sultan Zainal Abidin, Terengganu, Malaysia.

**Mada Sanjaya WS.**, Department of Physics, UIN Sunan Gunung Djati Bandung, Indonesia.

**Wahid Abdurahman**, Department of Physics, UIN Sunan Gunung Djati Bandung, Indonesia.

**Hani Rubiani**, Department of Electrical Engineering, Universitas Muhammadiyah Tasikmalaya, Indonesia.

**Mohamad Afendee Mohamed**, Faculty of Informatic and Computing, Universiti Sultan Zainal Abidin, Terengganu, Malaysia.

At present the use of learning machine methods is very rapidly developing both for classification and prediction, learning machines have many models including neural networks and support vector machines, in the last fifteen years neural networks and support vector machines are models that often compare their performance in predictions and estimates [3-12]. The author observes ten studies in the range of 2005-2018 that the prediction of wind speed and direction support vector machines is superior [3], for intrusion detection support vector machines are superior [4], for forecasting electrical energy consumption the support vector machine is superior [5], for image support vector machine is superior [6], for Estimating Function Based on Evolution Differential support vector machine is superior [7], to predict cooling load every hour in building support vector machine is superior [8], for microcalcification from mammograms Neural networks are superior [9], for predictions of performance of diesel engines in neural networks is superior [10], for diagnosis of students with neural network learning disabilities [11], for estimating the cost of building neural network schools is superior [12], but this cannot be used as a benchmark for which model is superior because the authors took samples randomly.

Various results obtained from comparisons of neural network models and support vector machines make the authors feel interested in writing similar comparisons using data from the results of mathematical pendulum experiments to determine the value of gravity, considering the previous calculation [2] using mathematical methods and linear regression to get results with accuracy above 90%, this can be called good.

Therefore, in this study, the accuracy of the three methods above (i.e. linear regression, neural network, and support vector machine) was compared to calculating gravitational values using data from mathematical pendulum experiments, so that the gravity value calculation model adapted two techniques (neural network and support vector machine) can be examined through linear regression.

II. LITERATURE REVIEW

A. Mathematical Pendulum Experiment

Mathematical pendulum experiments are points that are simulated with a pendulum that is hung at a fixed point with a rope[1]. The pendulum is swung with the initial angle of deviation of  $\theta$  to the vertical line, the restoring force can be written[13,14]:

$$F = -mg\sin\theta \tag{1}$$

For a small  $\theta$  value (in radians)  $\sin\theta = \theta = \frac{s}{l}$ , where  $s$  is an arc in a circle and is the length of the rope, so the following equation is obtained[13,14]:

$$F = \frac{mgs}{l} \tag{2}$$

Assuming there is no rotation or friction, with  $g$  as gravity. harmonious vibration equation with the T period can be written[13,15,16]:

$$T = 2\pi\sqrt{\frac{l}{g}} \tag{3}$$

Values  $l$  and  $T$  can be measured in experiments with heavy metal balls suspended on very mild wires. Determining the  $g$  in this way the results are quite accurate if the rope is milder than the ball, the deviation must be small (angle  $\theta < 15^\circ$ ), friction with the air is very small so that it can be ignored and there is no torsional force[13,17,18].

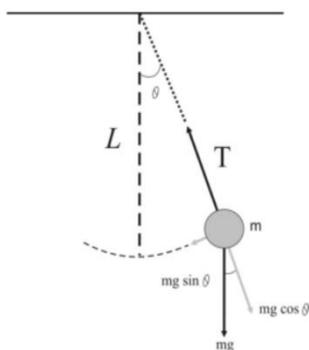


Fig. 1. Illustration of mathematical pendulum experiments[13]

B. Linear Regression

Linear regression (LR) is a statistical method that has 3 uses, including; 1) description of data phenomena, 2) control, and 3) predict. LR is able to describe the phenomenon of data through the formation of a numerical relationship model. LR is able to describe the phenomenon of data through the formation of a numerical relationship model. LR can also be used to control what is being observed through the use of the LR model obtained. In addition, the LR model can also be used to make predictions for the dependent variable. LR is used to connect two variables, one is called the dependent variable Y, the other is called the independent variable X[20].

With  $e$  called error. error is the difference between the observation and the line, LR expressed in mathematical functions as follows:

$$y = a_0 + a_1x + e \tag{4}$$

LR provides a technique for matching the most appropriate lines of data. However, this technique depends on the fact

that the link between the dependent and independent variables is linear. In the LR analysis, the first step should be a graph drawing to check whether the data applies to a linear relationship. Some data that is not linear can be synchronized with a data transformation[21].

C. Newton's Interpolation

Newton interpolation is one of the data matching methods that have the requirements of the data being analyzed must have very high accuracy, so this method will go through every point in the data. Generally, data point  $(x_0; y_0), (x_1; y_1), \dots, (x_n; y_n)$ , can be matched with a degree polynomial that has a general form as follows[21]:

$$P_{n-1}(X) = a_1 + (x - x_1)a_2 + (x - x_1)(x - x_2)a_3 + \dots + (x - x_1)(x - x_2)\dots(x - x_{n-1})a_n \tag{4}$$

So the solution for the polynomial coefficient is as follows[21]:

$$a_1 = y_1, a_2 = \nabla y_2, a_3 = \nabla^2 y_3, \dots, a_n = \nabla^n y_n \tag{5}$$

D. Neural Network (NN)

Neural network (NN) is a system that simulates the workings of the human brain based on a simplified neuron model and the relationship between these neurons [12]. NN are modeled in a mathematical way to mimic the intelligence of the human brain for use in various technologies. NN are processes in which free parameters of neural networks are adapted through a process of stimulation by the environment in which the network is embedded. The type of learning is determined by the way the parameter changes take place. NN consists of many processing elements that join together to form a network that matches the weighting function that can be adjusted for each input. These processing elements are usually arranged in three or more layers, there are: input layers where data is presented to the network through an input buffer, an output layer with a buffer that holds the output response to a given input, and one or more intermediate or "hidden layers" [22].

Multi-layer Perception (MLP) is one of the NN methods that uses a supervised learning algorithm that studies the  $f(\cdot): R^m \rightarrow R^o$  function by training on a data set, where  $m$  is the number of dimensions for input and  $o$  is the number of dimensions for output. With a set of  $X = x_1, x_2, \dots, x_m$  features and targets  $y$ , he can study the approximation of non-linear functions for classification or regression. This is different from logistic regression, between the input and the output layer, there can be one or more non-linear layers, called the hidden layer [23].

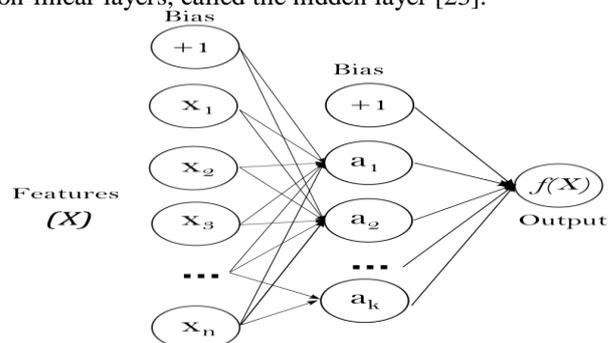


Figure 2. Illustration of MLP one hidden layer[23]

### III. RESEARCH METHOD

NN has the following advantages: 1) they can be used to build high-level nonlinear function estimation models; and 2) its use does not impose a limit on the number of features. The disadvantage of NN mentioned in previous research is that the black box technique and knowledge acquisition process is very time consuming [12].

#### E. Support Vector Machine (SVM)

Support Vector Machines (SVM) is a set of supervised learning methods used for classification, regression, and detection of outliers [23]. In regression  $\epsilon - SV$  the purpose of SVM is to find the function  $f(x)$  which has the most deviation  $\epsilon$  from the actual  $y$  target obtained for all training data, and at the same time as uniform as possible [25,26]. The input pattern (support vector) is mapped into the feature space by the map  $\Phi$ . Then, point products are calculated by drawing training patterns below the map  $\Phi$ . This matches the evaluation of kernel functions  $k_{xi}, x$ . Point products are collected using weights  $V_i = \alpha_i - \alpha_i$ . Finally, the final prediction results are calculated by adding a constant value (b) [12].

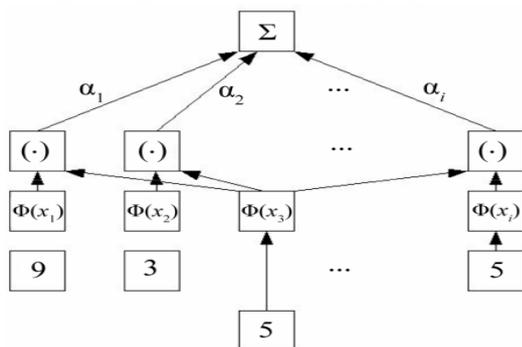


Figure 3. Support Vector Regression Structure [12]

The advantages of supporting vector machines are: 1) Effective in high dimensional space. 2) Still effective in cases where the number of dimensions is greater than the number of samples. 3) Using a subset of training points in the decision function (called a support vector), so that it also saves memory. 4) Versatility: Different Kernel functions can be specified for decision functions. Common kernels are provided, but it is also possible to specify special kernels. The disadvantages of supporting vector machines include: 1) If the number of features is far greater than the number of samples, avoid excessive installation in choosing Kernel functions and the term regularization is very important. 2) SVM does not directly provide a probability estimate, this is calculated using an expensive five-fold cross validation [23].

#### F. Evaluation of Accuracy

An experiment is said to have high accuracy if the experimental error rate is small, and an experiment is said to have high accuracy if the trial error rate is small [13].

$$\text{error} = \left| \frac{x_{xx} - x_{rxf}}{x_{rxf}} \right| \times 100\% \quad (6)$$

$$\text{Accuracy} = \left( 1 - \left| \frac{x_{xx} - x_{rxf}}{x_{rxf}} \right| \right) \times 100\% \quad (7)$$

Mathematical pendulum experiments are carried out by collecting period data from pendulum swings obtained from how long the pendulum is for one swing. By using the measured string length value, the  $g$  gravity value will be obtained using equation (3). Then interpolated the experimental data using Newton's interpolation to obtain data that was not obtained in the experiment to get more data from the results of the experiment.

With more data, it is possible to use it as good training data. The results of interpolation data including experimental data are used as training data to be used in the LR model and learning machines to predict other gravity values outside the range of experimental results.

Then evaluate the prediction results of each model by looking for error values, standard deviation values, and percentage accuracy values for reference values.

### IV. RESULT AND DISCUSSION

An experiment has been carried out to determine the value of gravity using the pendulum swing method. Where the method is carried out by collecting period  $T$  data from pendulum swings obtained from how long the pendulum takes for one swing. If we know the length of the rope  $l$  used by the pendulum, the gravitational  $g$  value will be obtained using equation (3). Gravitational values are taken which are closest to the reference gravity value ( $9.8 \text{ kg.m/s}^2$ ) [17,18,19] for each different rope length. So that the data is obtained as follows:

Table 1. Test results data [2]

No.	$l(m)$	$T(s)$	$g(kg.m/s^2)$	Acc(%)
1	0.22	0.967	9.28	94.68
2	0.24	1.010	9.28	94.68
3	0.26	1.041	9.46	96.55
4	0.28	1.076	9.54	97.33

Then the data in the Table 1. was interpolated using Newton's interpolation to obtain data that was not obtained in the experiment, so that 13 data were obtained.

Table 2. Data interpolated

No.	$l(m)$	$T(s)$	$g(kg.m/s^2)$	Acc(%)
1	0.22	0.967	9.28	94.68
2	0.225	0.981	9.22	94.09
3	0.23	0.992	9.22	94.06
4	0.235	1.002	9.23	94.19
5	0.24	1.010	9.28	94.68
6	0.245	1.018	9.32	95.14
7	0.25	1.025	9.38	95.76
8	0.255	1.033	9.42	96.17
9	0.26	1.041	9.46	96.55
10	0.265	1.049	9.50	96.91

# Implementation Model Linear Regression, Neural Network and Support Vector Machine in Simple Mathematical Pendulum Experiments to Determine the Value of Gravity

11	0.27	1.058	9.51	97.07
12	0.275	1.067	9.53	97.21
13	0.28	1.076	9.54	97.33

Table 2. has more data than in Table 1. This allows it to be used as training data that is good to be applied to machine learning which is then used to predict the gravity  $g$  values for rope length  $l$  and pendulum swing periods  $T$  that are greater than values and experiments. So that the data is obtained as follows:

**Table 3. Data on overall results of interpolation and prediction**

No.	$l(m)$	$T(s)$	$g$ in units of $(kg. m/s^2)$			
			Theory	RL	NN	SVM
1	0.22	0.967	9.28	9.28	9.19	9.35
2	0.225	0.981	9.22	9.24	9.23	9.34
3	0.23	0.992	9.22	9.24	9.27	9.33
4	0.235	1.002	9.23	9.25	9.30	9.34
5	0.24	1.010	9.28	9.29	9.33	9.35
6	0.245	1.018	9.32	9.33	9.36	9.37
7	0.25	1.025	9.38	9.38	9.39	9.38
8	0.255	1.033	9.42	9.42	9.42	9.39
9	0.26	1.041	9.46	9.46	9.45	9.40
10	0.265	1.049	9.50	9.49	9.47	9.42
11	0.27	1.058	9.51	9.52	9.51	9.43
12	0.275	1.067	9.53	9.54	9.54	9.44
13	0.28	1.076	9.54	9.57	9.57	9.43
14	0.285	1.085	9.55	9.59	9.60	9.42
15	0.29	1.095	9.54	9.61	9.63	9.40
16	0.295	1.104	9.55	9.63	9.66	9.39
17	0.3	1.113	9.55	9.66	9.68	9.39
18	0.305	1.115	9.68	9.68	9.70	9.39
19	0.31	1.130	9.57	9.72	9.73	9.39
20	0.315	1.138	9.59	9.76	9.75	9.39
21	0.32	1.145	9.63	9.81	9.77	9.39
22	0.325	1.152	9.66	9.86	9.79	9.39
23	0.33	1.159	9.69	9.91	9.81	9.39
24	0.335	1.167	9.70	9.95	9.83	9.39
25	0.34	1.175	9.71	9.99	9.85	9.39
26	0.345	1.184	9.71	10.01	9.87	9.39
Average of error rate (%)			3.06	2.77	2.53	4.22

Table 3 shows that the NN model has the lowest error rate compared to the other models in this paper, this shows that the accuracy of the NN model is best when compared with the RL model and SVM model for this case. In contrast, the SVM model has the lowest error rate when compared to other models in this paper. this shows that the accuracy of the SVM model is the worst when compared to the RL model and SVM model for this case.

Next we observe the analysis of the calculation of the error rate. As shown in Table 4.

**Table 4. Description of analysis from error rate (%) calculations**

Model	Mean	STDev	STDer	Lower	Upper
-------	------	-------	-------	-------	-------

Theory	3.06	1.63	0.32	0.89	5.94
RL	2.77	1.78	0.35	0.09	5.70
NN	2.53	1.97	0.3864	0.05	6.23
SVM	4.22	0.30	0.058	3.67	4.0

Table 4 shows that the SVM model has the lowest STDev and STDer values compared to other models. In contrast, the NN model has the lowest error rate when compared to other models in this paper.

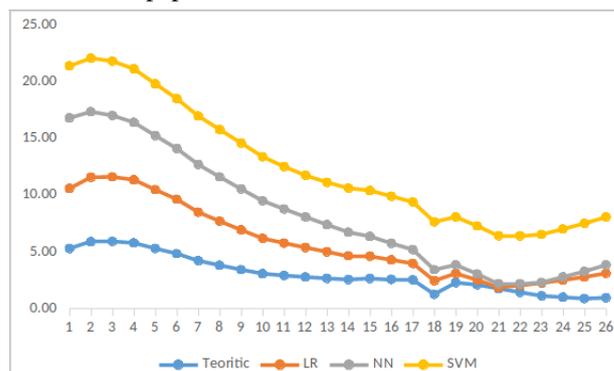


Figure 4. Comparison of error rates for each model

It has been discussed previously, in this case the NN model error rate is smaller than LR and SVM, the NN model produces a more accurate calculation value than LR and SVM, even more accurate than conventional mathematical theory calculations. NN has proven useful and suitable for dealing with complex problems and developing predictable models that are easy to use in the case of this experiment. They are able to detect patterns found in the data and provide greater opportunities to investigate various problems with value predictions. However, SVM models have advantages with the smallest standard deviations and deviation errors, they are very suitable for processing data that really requires standard deviations and small deviations from errors as in this experiment.

## V. CONCLUSION

The LR, NN and SVM models work well to be applied in mathematical pendulum experiments, the NN model is the most accurate compared to the three models above. However, the SVM model has advantages with standard deviation values and the smallest deviation of errors. In this experiment NN is suitable for research that requires better accuracy, SVM is suitable for research requiring standard deviation and smaller deviation of errors.

## ACKNOWLEDGMENT

The authors thank the Government of Malaysia for funding this research under the Fundamental Research Grant Scheme (FRGS/1/2018/ICT03/UNISZA/02/2) and also Universiti Sultan Zainal Abidin, Terengganu, Malaysia



## REFERENCES

1. PHWYE. "Physics Laboratory Experiments". *PHYWE SYSTEM GMBH*. 37070 Gottingen, Germany.
2. A. Wahid. "Object Tracking Menggunakan Metode Deteksi Tepi Canny Berbasis Python Open Cv". *Skripsi Jurusan Fisika, UIN Sunan Gunung Djati Bandung*. 2018.
3. K. Ali, Koury, R.N.N., L. Machado, Garcia, and J. Jose. "Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro-fuzzy inference system". *Sustainable Energy Technologies and Assessments*. 25. 146-160. 10.1016/j.seta.2018.01.001. 2018.
4. W. Chen *et al.* "Application of SVM and ANN for intrusion detection". *Computer and Operation Research*, volume 32. 2005.
5. A.S. Ahmad *et al.* "A review on applications of ANN and SVM for building electrical energy consumption forecasting". *Renewable and Sustainable Energy Reviews* 102-109. 2014.
6. W. Wong *et al.* "Application of SVM and ANN for image retrieval". *European Journal of Operational Research*, volume 173. 2006.
7. S. Zhou *et al.* "Parameters selection of SVM for function approximation based on Differential Evolution". *International Conference on Intelligent Systems and Knowledge Engineering*. Atlantis Press, ISSN 1951-6851, ISBN 978-90-78677-04-8. 2007.
8. L. Qiong *et al.* "Predicting hourly cooling load in the building: A comparison of support vectormachine and different artificial neural networks". *Energy Conversion and Management* 50 90-96. 2009.
9. R. Jinchang. "ANN vs. SVM: Which one performs better in classification of MCCs in mammogram imaging". *Knowledge-Based Systems; Volume* 26. 2012.
10. M. Noor, C. Wan, Mamat, Rizalman, N. Ahmed and Ali. "Comparative Study of Artificial Neural Network and Mathematical Model on Marine Diesel Engine Performance Prediction". *International journal of innovative computing, information & control: IICIC*. 14. 2018.
11. Tk. Wu, *et al.* "Evaluation of ANN and SVM classifiers as predictors to the diagnosis of students with learning disabilities". *Expert Systems and Application*, volume 34. 2008.
12. G. Kim, J. Shin, S. Kim, and Y. Shin. "Comparison of School Building Construction Costs Estimation Methods Using Regression Analysis, Neural Network, and Support Vector Machine". *Journal of Building Construction and Planning Research*, 1, 1-7. 2013.
13. T. Penyusun. "Modul Praktikum Fisika Dasar 1". *Laboratorium Fisika Fakultas Sains Dan Teknologi Uin Sunan Gunung Djati Bandung*. 2017
14. L. Prasetyo, T. K. Hien and S. Setiawan. "Mengerti Fisika". *Pertama penyunt. Andi Offset*. ISBN 979-533-088-8. Yogyakarta. 1992
15. A. Renreng. "Asas-asas ilmu alam universitas jilid I". *Badan Kerjasama Perguruan Tinggi Negeri Indonesia Bagian Timur*. Ujung Pandang. 1985
16. M. Shofwan. "Peranan Persamaan Diferensial Linier Orde Kedua pada Ayunan Bandul". *Universitas Negeri Malang*. Malang. 2003
17. P. Tipler. "Fisika Untuk Sains dan Teknik". *Erlangga*. Jakarta. 2001
18. Halliday and Resnik; P. Silaban and E. Sucipto, (1984). "Fisika Untuk Universitas, Edisi ketiga jilid 1". *Penerbit Erlangga*. Jakarta, p. 460. 1984
19. R. Serway. "Physics for Scientists & Engineers with Modern Physics". *James Madison University Harrison Burg*, Virginia. 1989
20. D. Kurniawan. "Regresi Linear ( Linear Regression)". *R. Development Core Team*. Vienna, Austria. 2008
21. S. W. S. Mada. "Metode Numerik Berbasis Python". *Penerbit Gava Media* ISBN 978-602-7869-91-2. Yogyakarta. 2015
22. R. E. Uhrig. "Introduction to artificial neural networks". 33 - 37 vol.1. 10.1109/IECON.1995.483329. 1995
23. Pedregosa *et al.* "Scikit-learn: Machine Learning in Python". *Journal of Machine Learning Research* 12, pp. 2825-2830. 2011.
24. S. W. S. Mada *et al.* "Colored Object Sorting using 5 DoF Robot Arm based Artificial Neural Network (ANN) Method". *Journal of Physics: Conference Series*. 1090. 012070. 10.1088/1742-6596/1090/1/012070. 2018
25. V. Vapnik. "The Nature of Statistical Learning Theory". *Springer*. New York. 1995
26. A. Smola and B. Schölkopf. "A Tutorial on Support Vector Regression". *Statistics and Computing*. 14.10.1023/B:STCO.0000035301.49549.88. 1998