<u>How accurately can a machine learning model predict the frequency of a simple pendulum compared to the theoretical formula, across a range of release angles?</u>

PHYSICS HL EXTENDED ESSAY

Word Count: 3520

May 2026

1.0) Introduction

The simple pendulum is one of the most studied and one of the most basic systems in the area of physics. A system as simple as connecting a mass to a piece of rope and letting it swing has been the study of many scientists for centuries. Pendulums are also a part of our daily lives they are used in: Clocks, Seismometers and also on some architectural buildings. The general formula to produce the meaning of the pendulum has some limitations. It only produces correct results in small angles. When it comes to great angles it fails to give us a correct measurement. This means that the formula does not provide an exact solution in some real world problems. However this problem could be solved by a machine learning system.

In this study a simple pendulum experiment measuring frequency using different angles was done and the results were used to train a fully functioning neural system(Artificial Intelligence). This model predicts the frequency of a pendulum when the theoretical formula becomes ineffective at great angles. The advantage and the interesting part of using a machine learning model that learns from real data is that it could be used in situations where the formula fails.

The reason for me to pick this topic is that we have been introduced to the pendulum and I have been able to attend a real pendulum experiment in one of our classes. Our teacher put real emphasis on the limitations of the theoretical as it is limited to some angles. I am also interested in coding and creating new things using technology, and after our lesson I wanted to combine these two areas: coding and physics. As a result I produced this Extended Essay research in the area of Physics.

1.1) Research Question

The research question of this exploration is: How accurately can a machine learning model predict the frequency of a simple pendulum compared to the theoretical formula, across a range of release angles.

2.1) The Definition of Period and Frequency

A simple pendulum consists of a mass swinging around while connected to a rope. And there are two general terms to define this movement: frequency and period. Period is the time taken for the pendulum to complete one full oscillation and it is measured in the unit seconds. Frequency is the number of oscillations per second and Hertz(Hz) is the unit of it. These two terms are connected to each other. Specifically, frequency is the inverse of Period.

Therefore the relation is formulated as:

$$F = \frac{1}{T}$$

2.2) The Theoretical Formula

The formula for simple pendulum is shown as:

$$T = 2\Pi\sqrt{\frac{L}{g}}$$

The value g represents Gravitational Acceleration(g = 9.81 meters per second squared) and the value L represents the length of the pendulum in meters.[1]

2.3) The Limitation of the Formula: Small Angle Approximation

However this formula relies on an estimation, it does not take account great angles appropriately. The full equation of the pendulum movement includes $sin\theta$, and the solution of this equation analytically is pretty complicated. To overcome this problem physicists use the $sin(\theta) \approx \theta$(in radians) approach in small angles.[2] By using this they found the simple pendulum we mentioned earlier in this exploration. However as we have said this formula is limited only to small angles.[3] When the angle increases a gap occurs between the formula value and the real experimental value. In reality the pendulum oscillates more slowly when released from larger angles therefore the frequency changes. This situation creates a huge problem in the calculation of frequency.[4]

2.4) Neural Networks

To overcome the problem we have mentioned in 2.3 Artificial intelligence or machine learning models could be used. These models get their base from neural networks. Neural networks are models that are created by getting inspiration from the human brain. It includes segments that are connected to each by layers. These layers include the input layer, secret layers and the output layer. They get the information from the input layer, process it in the secret layer and give the result from the output layer.

These networks learn from the data you give them. In the training of these networks so much input and output data are given to the model. And the model tries to learn the pattern between input and output. The more you train and give new data to the model the more it will be accurate and realistic. The reason why a machine learning model is suitable for this problem is that the model is given the release angle as an input and the frequency as an output. We trained the model using the data from our simple pendulum experiment, which we will mention in the further stages of this Extended Essay.[5]

3.1) Material List

1. Pendulum bob(50g)
2. String(20cm)
3. Protractor(±0.5°)
4. Smartphone chronometer(±0.1 seconds)
5. Retort stand and clamp
6. Tape Measure(0.5 cm)

3.2) Hypothesis

It is hypothesized that as the release angle of the pendulum increases, the measured frequency will decrease, deviating from the constant value predicted by the theoretical formula.

3.2) Variables

Variable Type	**Variable**	**Description**
Independent Variable	Release Angle	Tested from 5° to 90° at the increase of each 5 degrees. It was measured using a protractor.
Dependent Variable	Frequency	Calculated from the measured period at each angle.
Controlled Variable	Pendulum Length	Kept constant at 20 cm throughout all trials. It was measured using a tape measure.

Controlled Variable	Pendulum Weight	The same weight was used throughout all trials. A 50 gram pendulum bob was used as the constant weight.
Controlled Variable	Experiment Environment and Temperature	All trials were conducted in the same location and the temperature was also therefore constant.

3.3) Measurement Process

1. The Pendulum is set to the chosen angle using the protractor.
2. The pendulum is released and at the same time we start the chronometer.
3. When the pendulum completes 10 full oscillations the chronometer is stopped at the time it is recorded.
4. Then the frequency is calculated using $f = \frac{10}{t}$
5. This process is completed 5 times at the same angle
6. This whole is process is repeated for all angles from 5 to 90 degrees(in angles that are the multiple of 5)

3.4) The Creation of the Machine Learning Model

After we completed the pendulum experiment, the acquired 90 trial data were used to train the machine learning model. The model is a feedforward neural network and it has the structure: 1 input layer(release angle), 2 hidden layers(each 64 neurons) and an output layer(frequency). This model is trained in 300 epochs. At each epoch the whole data set is processed under groups of 16(batch size 16) and compared its guesses to real data. During this process the validity of the estimates of the model were checked and it was made sure that the model was not overfitting.

Only the release angle was given to the model as input and the frequency value that matches to that angle was given out as output. By this way the model learned the correlation between angle and frequency and has become able to guess the frequency of release angles that have not been tested or that have decimals.

The following are photos taken from the coding environment of the Machine Learning Model:

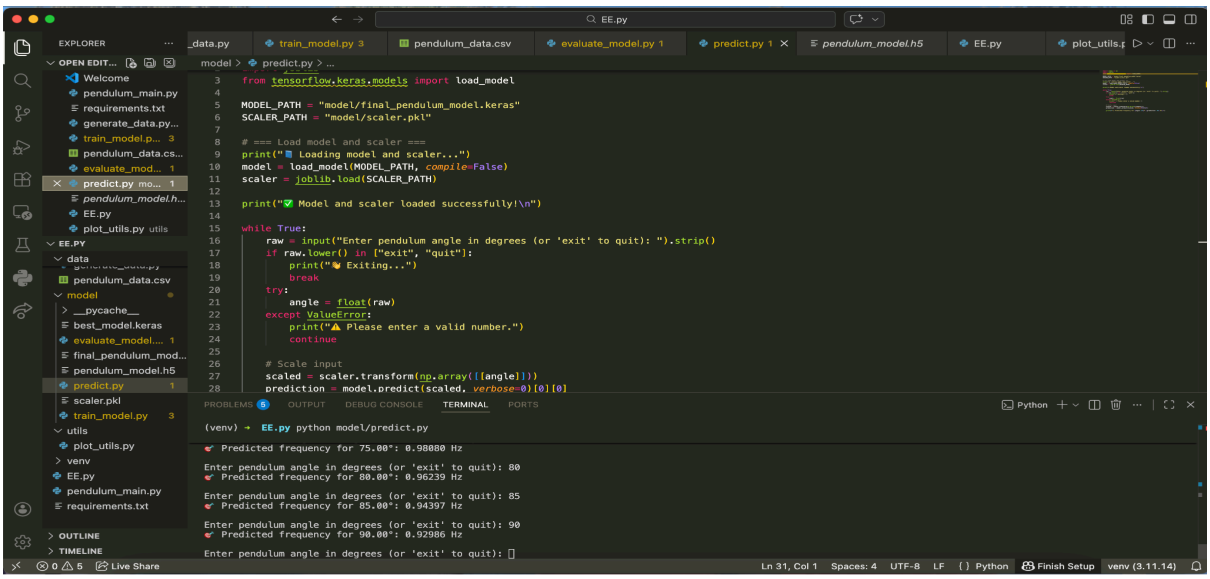


Image 3.4: Image taken from the working environment of this project, working on the predict.py mechanism.

This image is taken from VisualStudio Code environment where we coded this project. At this particular time I was testing the prediction abilities of the model and was developing the predicting part. I was particularly trying to improve the layers output so that it would become easier for me to understand the values coming out of the model.

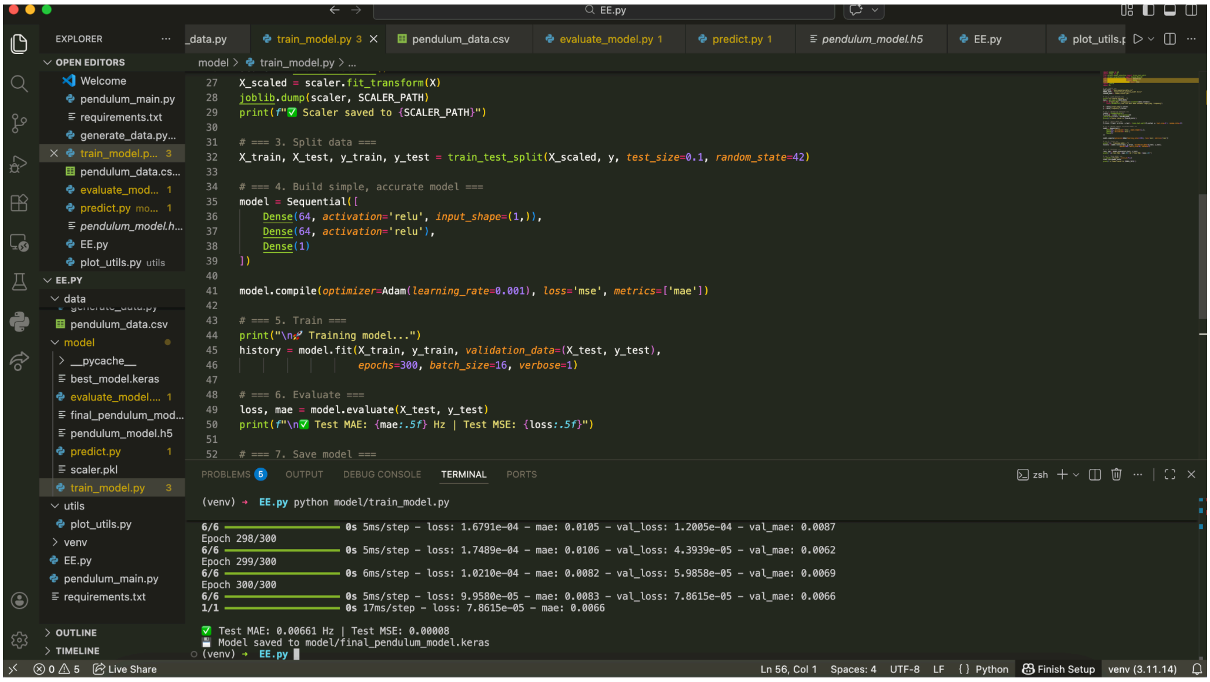


Image 3.4.2: This image is taken again from VisualStudio Code but this time it is a picture of the train_model.py mechanism I developed. In the terminal the model is currently training itself in 300 epochs as it is shown.[6]

4.1) Data Table and Results

The sample Calculation of Uncertainty for Raw frequency is shown as:

Given t = 9.1s, $\Delta t = \pm 0.1\ s,\ f = 1.10\ Hz$

$\Delta f = f \times (\frac{\Delta t}{t})\ \Delta f = 1.10 \times (\frac{0.1}{9.1})\ \Delta f = \pm 0.0121\ Hz \approx 0.01\ Hz$

Angle(±0.5)	**Trials**	**Time taken for 10 Oscillations(±0.1 seconds)**	**Frequency(±0.01 Hz)**
5	1	9.1	1.10
5	2	9.0	1.11
5	3	9.1	1.10
5	4	9.0	1.11
5	5	8.9	1.12
10	1	9.1	1.1
10	2	9.1	1.1
10	3	9.1	1.1
10	4	8.9	1.12
10	5	9.0	1.11
15	1	9.0	1.11

15	2	9.1	1.1
15	3	9.0	1.11
15	4	9.1	1.1
15	5	9.0	1.11
20	1	8.9	1.12
20	2	9.1	1.1
20	3	9.0	1.11
20	4	9.1	1.1
20	5	9.0	1.11
25	1	9.2	1.09
25	2	9.0	1.11
25	3	9.2	1.09
25	4	9.2	1.09
25	5	9.2	1.09
30	1	9.0	1.11
30	2	9.1	1.1
30	3	9.0	1.11
30	4	9.1	1.1
30	5	9.1	1.1
35	1	9.2	1.09

35	2	9.4	1.06
35	3	9.2	1.09
35	4	9.2	1.09
35	5	9.2	1.09
40	1	9.1	1.1
40	2	9.2	1.09
40	3	9.3	1.08
40	4	9.1	1.1
40	5	9.3	1.08
45	1	9.3	1.08
45	2	9.2	1.09
45	3	9.4	1.06
45	4	9.3	1.08
45	5	9.3	1.08
50	1	9.4	1.06
50	2	9.3	1.08
50	3	9.4	1.06
50	4	9.5	1.05
50	5	9.3	1.08
55	1	9.6	1.04

55	2	9.5	1.05
55	3	9.5	1.05
55	4	9.7	1.03
55	5	9.4	1.06
60	1	9.7	1.03
60	2	9.7	1.03
60	3	9.6	1.04
60	4	9.7	1.03
60	5	9.7	1.03
65	1	9.7	1.03
65	2	9.6	1.04
65	3	9.6	1.04
65	4	9.8	1.02
65	5	9.9	1.01
70	1	9.9	1.01
70	2	10.0	1.0
70	3	9.8	1.02
70	4	10.0	1.0
70	5	10.0	1.0
75	1	10.2	0.98

75	2	10.2	0.98
75	3	10.1	0.99
75	4	10.1	0.99
75	5	10.0	1.0
80	1	10.2	0.98
80	2	10.1	0.99
80	3	10.4	0.96
80	4	10.3	0.97
80	5	10.3	0.97
85	1	10.5	0.95
85	2	10.5	0.95
85	3	10.6	0.94
85	4	10.4	0.96
85	5	10.4	0.96
90	1	10.6	0.94
90	2	10.5	0.95
90	3	10.7	0.93
90	4	10.6	0.94
90	5	10.6	0.94

Table 4.1: Data Table for raw experimental measurements.

4.2) Calculating Theoretical Frequency and Comparing with the Experimental Results

The theoretical frequency was calculated using $T = 2\Pi\sqrt{\frac{L}{g}}$

L = 0.20m

g = 9.81 m/s^2

The calculations are therefore:

$T = 2\Pi\sqrt{\frac{0.20}{9.81}}$

$T = 2\Pi \times 0.1428$

$T = 0.8974$

And from $f = \frac{10}{t}$

$f = 1.1147\ Hz$

The Average time uncertainty was calculated as:

Trials at 5 degrees: 9.1, 9.0, 9.1, 9.0, 8.9s

Mean = 9.02 s

Deviations squared:

(9.1 − 9.02)^2 = 0.0064

$(9.0 - 9.02)^2 = 0.0004$

$(9.1 - 9.02)^2 = 0.0064$

$(9.0 - 9.02)^2 = 0.0004$

$(8.9 - 9.02)^2 = 0.0144$

Sum = 0.0280

Deviation = $\sqrt{\frac{0.0280}{4}} = 0.0837$

Uncertainty = $\frac{0.0837}{\sqrt{5}} = \pm\ 0.0374$

The Average Experimental Frequency Uncertainty was calculated as:

Trials at 5 degrees: 1.10, 1.11, 1.10, 1.11, 1.12 Hz

Mean = 1.108 Hz

Deviations squared:

$(1.10 - 1.108)^2 = 0.000064$

$(1.11 - 1.108)^2 = 0.000004$

$(1.10 - 1.108)^2 = 0.000064$

$(1.11 - 1.108)^2 = 0.000004$

$(1.12 - 1.108)^2 = 0.000144$

Sum = 0.000280

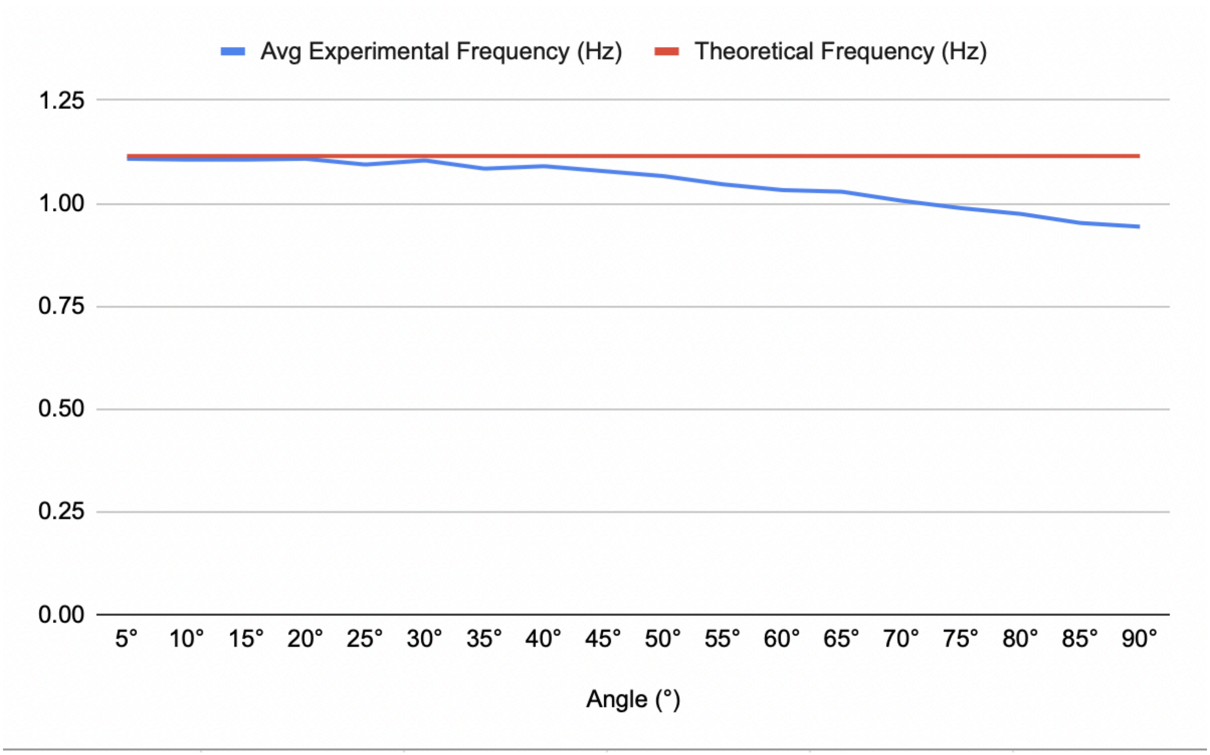
Deviation= $\sqrt{\frac{0.000280}{4}}$ = 0.0084 Hz

Uncertainty = $\frac{0.0084}{\sqrt{5}}$ = ± 0.0037 Hz

Angle (±0.5)	Average Time for 10 Oscillations. (±0.06s)	Average Experimental Frequency (Hz)
5°	9.02	1.108±0.004
10°	9.04	1.106±0.004
15°	9.04	1.106±0.002
20°	9.02	1.108±0.004
25°	9.16	1.094±0.004
30°	9.06	1.104±0.002
35°	9.24	1.084±0.006
40°	9.20	1.090±0.005
45°	9.30	1.078±0.005
50°	9.38	1.066±0.006
55°	9.54	1.046±0.005
60°	9.68	1.032±0.002
65°	9.72	1.028±0.006
70°	9.94	1.006±0.004

75°	10.12	0.988±0.004
80°	10.26	0.974±0.005
85°	10.48	0.952±0.004
90°	10.60	0.940±0.003

Table 4.2: Table for comparison between theoretical frequency and the experimental average frequency for each angle tested.



Graph 4.2: Graph comparing theoretical frequency and the experimental frequency.

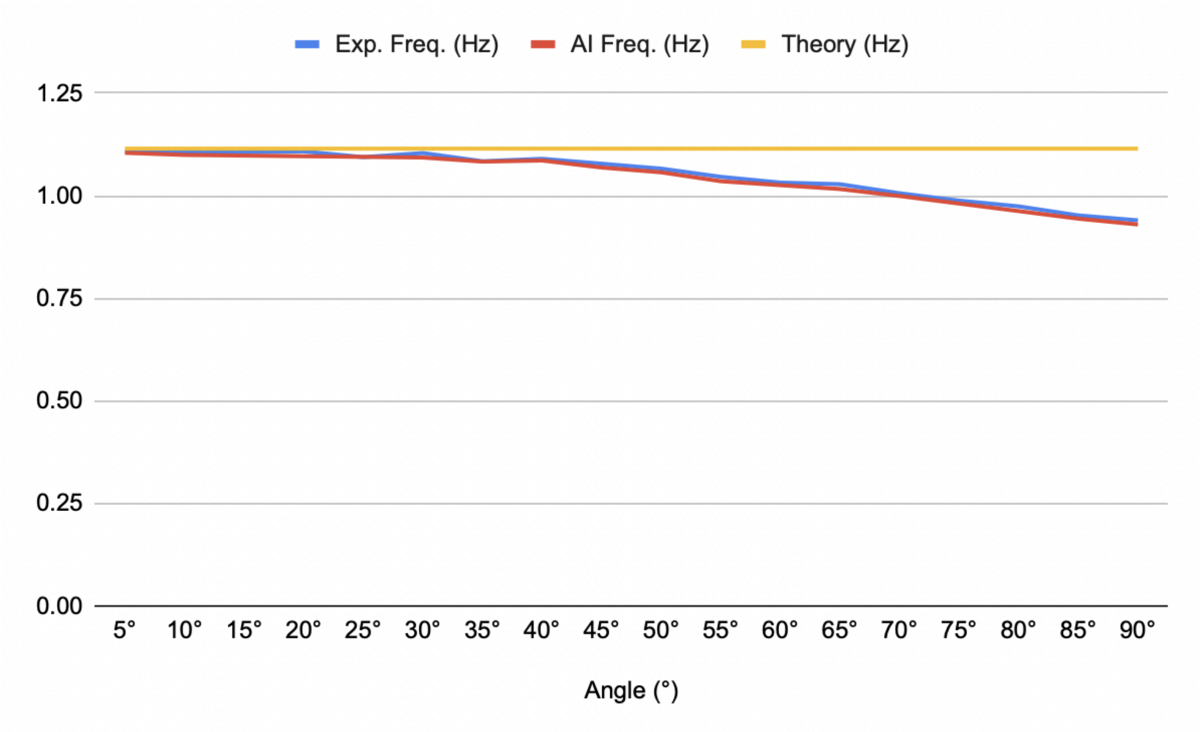
If we analyze table 4.1, we can see that 5 trials were done for each release angle and between the trial values there is only a slight difference meaning the experiment's reproducibility rate and accuracy is high. When we look at the average frequency table we can see that the experimental frequency of angles from 5 to 25 degrees are almost identical to the theoretical frequency value. However after 25 degrees there is a clear difference occurring between theoretical and experimental values. This is a clear example of the limitation of the theoretical formula we discussed in section 2.0. At 90 degrees average frequency was measured to be 0.940 Hz and this value is approximately 19% smaller than the theoretical frequency. This finding gives us the knowledge that the pendulum formula does not provide results that are similar to real experimental data. In the graph the red line represents the theoretical frequency and the blue line represents the frequency results from our experiment. We can see that after 25 degrees the frequency rate starts to drop. And we can see the huge difference between the theory and the experiment as the gap between the lines they represent starts to increase.

5.1) Analysis and comparison of the AI Model

Angle (°)	Experiment Freqency (Hz)	Model Frequency. (Hz)	Model % Error	Formula % Error
5°	1.108	1.10450	0.32%	0.60%
10°	1.106	1.09937	0.60%	0.79%
15°	1.106	1.09779	0.74%	0.79%

20°	1.108	1.09621	1.06%	0.60%
25°	1.094	1.09462	0.06%	1.89%
30°	1.104	1.09303	0.99%	0.97%
35°	1.084	1.08311	0.08%	2.83%
40°	1.090	1.08580	0.39%	2.27%
45°	1.078	1.06868	0.86%	3.40%
50°	1.066	1.05697	0.85%	4.57%
55°	1.046	1.03524	1.03%	6.57%
60°	1.032	1.02565	0.62%	8.01%
65°	1.028	1.01622	1.15%	8.43%
70°	1.006	0.99922	0.67%	10.81%
75°	0.988	0.98080	0.73%	12.82%
80°	0.974	0.96239	1.19%	14.45%
85°	0.952	0.94397	0.84%	17.09%
90°	0.940	0.92986	1.08%	18.59%

Table 5.1: This table provides the data from the Machine learning model in contrast with the experiment and with the theoretical value.



Graph 5.1: This graph shows the comparison between the model, the experiment and the theoretical formula.

We can see from this line graph that the model results and the experimental results are almost identical and that it is hard in some parts to understand which one is which. However the theory sits above and does not express a difference throughout different angles.

When we analyze table 5.1, it is understandable that ,for all the angles we tested in the experiment, the machine learning model predicted frequency values that are pretty close to real experimental values. The percentage error for the model compared to the experiment was seen to be in the range of 0.06% to 1.19%. Compared to this accuracy, the theoretical formula's percentage error begins to extremely increase after 25 degrees. At 90 degrees the theoretical formula's error was measured to be 18.59% while the model's error was 1.08%.

This finding suggests that at great angles, especially above 25 degrees, the model puts a more realistic and accurate estimate compared to the formula.

Angle (°)	**Model Predicted Frequency (Hz)**	**Model against Theory % Difference**
7°	1.10032	1.29%
12°	1.09874	1.43%
18°	1.09684	1.60%
33°	1.08709	2.48%
47°	1.06459	4.50%
63°	1.01929	8.56%
78°	0.96975	13.00%

Table 5.1.2: This table shows the predicted frequency values for some of the angles that have not been tested in the experiment.

When we analyze table 5.1.2, we can see the frequency values estimated by the model on the angles that have not been tested. For example, for the angles 7 and 78 degrees the model estimated values based on the correlation it created from the experimental values. However the theoretical formula is not sufficient to provide these frequency values accurately. We can again see the difference between the model and the theoretical formula when the angle increases as the percentage error increases to 13.00% at 78 degrees. This situation lets us know that the model does not only learn the data from the experiment but it learns the correlation between angle and frequency.

Angle (°)	Model Frequency. (Hz)	Model against Theory % Difference.
7.5°	1.10016	1.30%
23.5°	1.09510	1.76%
37.5°	1.08305	2.84%
52.3°	1.04546	6.21%
67.2°	1.00967	9.42%
78.5°	0.96791	13.17%
84.6°	0.94508	15.22%

Table 5.1.3: Table showing the results when the frequency is estimated in decimal angles in the model.

When we look at table 5.1.3, we can see one of the greatest advantages of this model. The model as seen in the examples of 7.5 and 84.6 degrees can predict the frequency in decimal placed angles. This finding creates an ease to calculate the frequency of the decimal placed angles. These results suggest that the model is useful at all angles and provide a practical alternative to the theoretical formula.

In this Extended essay 3 different frequency values were compared, the one from formula, the one from experiment and the one from the machine learning mode we created. The theoretical formula gave the constant frequency 1.1147 Hz. However this value can not match the real experimental value when the angle is above 25 degrees. The experimental values show that the frequency drops as angle increases. However the experimental results are only applicable to the tested angles. The experiment does not give us a chance to calculate other angles. The machine learning model however overcomes both the limitations of Theoretical formula and the experiment. The model showed a small percentage error of 1.19% allowing us to interpret that the model is able to match real experimental data. The model also provides accurate data after 25 degrees of release angle, meaning it is able to provide data that is correlated with the real experimental data. In addition to these the model also provides angles that have not been tested before. These angles could be anything between the range of 0-90 degrees and it could even be decimal placed angles. This feature of the model allows us to learn the frequency values of angles that have not been tested in the experiment, and therefore the model provides a meaningful result of these angles as well. To sum up, the model showed similarity with both the formula and the experiment under 25 degrees of release angle. However after that limit, where the theoretical formula becomes insufficient, the model provides estimates that are coherent with the real experimental data. This finding gives a meaningful answer to the research question, a thing we will discuss in the conclusion part of this extended essay research.

6.1) Evaluation and Limitations

In this experiment several factors may have influenced the accuracy and overall results of the data we gathered. Firstly, the smartphone chronometer we used to measure the period of the pendulum had an uncertainty of ±0.1. This value may have influenced the frequency values we have measured throughout the experiment and it may have been a significant source of error. The start-stop processes were done by hand on the chronometer, therefore human reaction time may have influenced the time taken for 1 period. Even though 10 oscillations were measured to minimise this source of error, it could still be considered as a limitation in this experiment. For the measurements of angles a protractor with ±0.5 uncertainty was used. This amount is significant and could have been a huge source of error when the experiment is taken as a whole. Addition to this air friction was not taken under control and it has influenced the movement of the pendulum. Especially in the greater angles, the pendulum may have slowed down due to friction. Therefore this is also a huge source of error.

The artificial intelligence model was only trained with a data set size of 90 trials, meaning it is a relatively small size dataset. This could have caused errors and is a limitation that could be solved by increasing the trial amount the the type of angle tested. The model learns the correlation of angle and frequency from the data set of trials we acquired from the pendulum expergment. If we test more angles with more trials we could get more accurate results from the model. And we may even make the percentage error or 1.08% lower than 1% and potentially reach below 0.5%. Therefore this is a limitation that may have caused errors and is a limitation that could be solved if further experiments are done. Beyond all these limitations, the theoretical formula gives a frequency value that is based on perfect conditions. The formula gives the value for example for the perfect pendulum, no friction and no other errors. However our experiment had multiple sources of potential errors. Therefore when comparing these two results an accuracy problem could arise. This is another limitation that may have affected the results of this experiment.

6.2) Conclusion

This investigation proves that a machine learning model could predict the outcomes of a simple pendulum much more accurately compared to a theoretical pendulum formula. The findings show us that the model was able to estimate frequency values with the percentage error as low as 1.19%. Compared to this the theoretical formula had a percentage error of 18% when the release angle reached 90 degrees. The model is able to give results more realistic and more precise when we compare it with the theoretical formula. The model was also able to give reliable estimates about the angle values that have not been tested in the experiment. Meaning the model is ready to be used as a tool that we can rely on to find both decimal placed values and release angle values that have not been tested yet. The model has been trained using the angles from 5 to 90 degrees, only the values that are a multiple of 5 were used. The reason the model is able to estimate any angle value is that it was trained to learn the correlation between angle and frequency. The small angle approximation the theoretical formula is limited on, blocks us to analyze any angle we want or we may need in the future. The model however gives accurate estimates of the desired angle. As a result we can answer the research question and say that the model has been more accurate and more advantageous than the formula and is able to provide more realistic estimates. We can say that machine learning models could provide realistic, accurate and precise estimates when producing knowledge related to the outcomes of a physical experiment.

References

1. Tsokos, K. A. (2023). *Physics for the IB Diploma* (7th ed.). Cambridge University Press.
2. Tolasa, D. G. (2025). Theoretical analysis of a simple pendulum experiment. *International Journal of Current Research in Science, Engineering and Technology, 8*(1), 214–218. https://doi.org/10.x
3. Aggarwal, N., Verma, N., & Arun, P. (2005). Simple pendulum revisited. *European Journal of Physics, 26*(3), 517–523. https://doi.org/10.1088/0143-0807/26/3/016
4. Beléndez, A., Rodes, J. J., Beléndez, T., & Hernández, A. (2009). Approximation for a large-angle simple pendulum period. *European Journal of Physics, 30*(2), L25–L28. https://doi.org/10.1088/0143-0807/30/2/L03
5. Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., & Piccialli, F. (2022). Scientific machine learning through physics-informed neural networks: Where we are and what's next. *Journal of Scientific Computing, 92*(3), 88. https://doi.org/10.1007/s10915-022-01939-z
6. Kashinath, K., Mustafa, M., Albert, A., Wu, J. L., Jiang, C., Esmaeilzadeh, S., & Hassanzadeh, P. (2021). Physics-informed machine learning: Case studies for weather and climate modelling. *Philosophical Transactions of the Royal Society A, 379*(2194), 20200093. https://doi.org/10.1098/rsta.2020.0093