ISSN: 2637-6067 | DOI: 10.6918/IJOSSER.202208_5(8).0034

A Research on the Relationship between Indoor Environment and Learning Efficiency based on Symbolic Regression

Xuebing Ba[1, a], Fengnan Wang[2, b], Yonghui Wang[2, c] and Xueying Zhang[3, d]

[1] School of Energy and Power Engineering, Dalian University of Technology, Dalian 116024, China

[2] School of Economics and management, Dalian University of Technology, Dalian 116024, China

[3] School of Economics and management, Dalian University of Technology, Dalian 116024, China

[4] School of Marxism, Dalian University of Technology, Dalian 116024, China

[a]baxb@dlut.edu.cn, [b] wangfn@mail.dlut.edu.cn, [c] wangyonghui0302@mail.dlut.edu.cn,

[d] 15660963238@163.com

Abstract

The research of the learning environment has become an important part of education, because indicating how the learning environment affects learning efficiency has great significance. This research focuses on the indoor environment in summer, and aims to figure out which environmental factors have substantial influences on learning efficiency. The indoor air data and questionnaire data were collected in summer, and the nonlinear relationship between environmental factors and learning efficiency was identified by applying the automatic data mining function of symbolic regression. A model of this relationship was also established. In this paper, three major conclusions were made: (1) In the summer indoor environment, the temperature and carbon dioxide concentration are primary influential factors, especially the temperature; (2) The optimal temperature of the indoor environment is about 25°C, and the lighting condition also has a certain effect on learning efficiency; (3) The higher the carbon dioxide concentration, the lower the learning efficiency.

Keywords

Symbolic regression; Learning efficiency; Learning environment.

1. Introduction

The advancement of technology has accelerated the development of learning methods. Mobile learning, virtual learning, formal and informal learning and else have become increasingly popular new ways of learning. As an important research direction, the learning environment has gained growing attention from a lot of researchers, implying that it will attach more importance to the field of education in the future. Indeed, dating back to the early 19th century, researchers had just started to pay attention to the influence of the environment on working and learning efficiency. The studies then focused on how the thermal environment and air quality affected respondents' subjective performance, cognitive performance, and physiological performance. In more detail, subjective performance includes subjective perception, emotional fluctuation, satisfaction with physical components, fatigue symptoms, and other factors influenced by the external environment; Cognitive performance comprises problem-solving ability, attention, memory, percipience, creativity, etc. Physiological

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202208_5(8).0034

performance generally covers the heart rate, finger-skin temperature and blink rate of the respondents, etc.

J Jiang et al. (2019) found that in a relatively cold environment, students' learning efficiency was higher, through subjective questionnaires and tests on learning cognitive ability [1]. Kim H et al. (2020) adopted the Electroencephalogram (EEG) measuring method and found that learning efficiency was the highest at around 25.7 °C. However, when the temperature turned too high or too low, the learning efficiency in these two cases both decreased [2]. Ming-Xue et al. (2012) applied on-site measurements and the questionnaire-survey method to research a naturally ventilated classroom environment at a college during the winter. The results revealed the individual influences of indoor light, PMV, and CO_2 concentration on learning efficiency, with CO_2 being the primary influential factor [3]. Chang K F et al. quantified satisfaction with the quality of the indoor environment by utilizing the laboratory simulation method. The factors like temperature, humidity, lighting intensity, ventilation, and so on were considered to design the indoor environment, so that they investigated the relationship between the quality of the indoor environment and learning results as well as working efficiency [4]. C Jung et al. (2021) improved the air quality by adding the green plants in the classroom, and found that the improvement of air quality has a certain positive effect on learning concentration [5]. Mao P et al. (2009) implemented experiments in a classroom at university and noticed higher satisfaction with a comfortable environment but a more obvious Sick Building Syndrome, demonstrating that students' subjective moods and learning efficiency could be significantly affected by indoor temperature variations [6]. Norazman N et al. (2018) focused on how lighting conditions of classrooms have a certain effect on learning efficiency. The results revealed that learning performance could be directly or indirectly influenced by the lighting quality of the classroom. Good lighting conditions played a significant role in encouraging learning effects and preventing vision or headache-related issues [7]. As compared to the usual influences of thermal environment and air quality, Wyon D P et al. (2018) argued that working efficiency was comprehensively affected by all indoor environmental factors in the short term [8].

Based on existing research, it can be seen that designing experiments has mostly been the primary approach. By designing with multi-factors and multi-variables to form a single or cross-combined experimental environment, the patterns of environmental variables influencing working and learning efficiencies could be obtained. However, with the advancement of computer technology, machine learning has ushered in two developmental booms. The more notable one is the symbolic regression method based on genetic algorithms, which has steadily attracted academics' attention. Symbolic regression, first proposed in the 1990s, now has been used in numerous fields because of its outstanding capacity to automatically extract the underlying patterns in data.

MG Pizon et al. (2021) adopted symbolic regression to forecast and model Disease Burden in the Philippines [9]. H Wang et al. applied symbolic regression to build a model related to tool wear, and this regression model was further used to identify the state of tool wear [10]. Yang G et al. (2013) utilized symbolic regression to predict the relationship between the economy and environmental pollution in 283 Chinese cities, from which the results of linear regression, nonlinear regression, and symbolic regression were compared and analyzed respectively [11]. Besides, the future trend was also predicted by modelling oil production, which was similar to those of other approaches [12]. By identifying influential elements of energy intensity in China, the total population was regarded as the most influential factor [13]. Furthermore, the relationship between CO2 emissions and economic growth had also been modelled [14]. Li W et al. (2019) employed symbolic regression to conduct a cluster analysis of also the relationship between carbon dioxide emissions and economic development [15]. Wand Y et al. (2019) studied the application of symbolic regression in materials science [16]. Yang X et al. (2021)

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202208_5(8).0034

implemented symbolic regression to model relationships between retail prices and consumer reviews, aiming to assist experts in building relationship models efficiently and thus making corresponding decisions reasonably [17].

One contribution of this paper is proposing a novel approach to study the influential factors of students' learning efficiency, and to find the most important ones in the summer indoor environment. Different from traditional research methods, the advantages of symbolic regression are the capability to automatically mine potential relationships hidden in data, which were applied to predict the influence trend of each factor and to identify primary influential factors from within.

Another contribution is implementing a detailed analysis of the primary influential factors from the perspective of symbolic regression. Generally speaking, statistical perspectives such as linear regression have mostly been used in relevant studies, yet this research established a model between the primary influential factors and students' learning efficiency from the perspective of symbolic regression. Through the analysis of their influence trends, this research found the best range of indoor temperature in summer, which agreed with some previous studies. In addition, the conclusions about the effect of indoor air quality on learning efficiency were also consistent with some existing studies. Thus, the model established by symbolic regression is proven to be more diverse in form, and after more influential factors are considered, more comprehensive results will be achieved.

2. Methods

Symbolic regression, an evolutionary method of function discovery based on genetic programming, was first proposed by Koza J R in 1992 [18]. Unlike traditional regression methods that predefine the formula, symbolic regression can discover the hidden mathematical formulas inspired by the process of biological genetic evolution, and apply characteristic variables to forecast the target variables. Genetic algorithms, such as replication and mutation of individuals, could be adopted to achieve automatic evolution within a randomly generated population. When a gene benefits from the current environment, the promising individuals will have more chances to survive during the evolutionary process, eventually with only the best genes remaining. As a result, the optimal solutions at present will be efficient to acquire benefits from the symbolic regression method, simultaneously the importance and influence of each factor are also displayed.

There are numerous factors affecting students' learning efficiency, including temperature, humidity, carbon dioxide and so on, and the interactions among these factors are complex. But the commonly applied regression methods have to set a predefined structure among those factors, which is difficult to develop from the above nonlinear relations. On the contrary, the symbolic regression method could effectively solve this problem by automatically discovering linear or even nonlinear relations without a predefined regression structure. Furthermore, significantly influential factors will automatically emerge from these discovered relations.

The advantage of symbolic regression is that no prior knowledge or models are required to identify the relationship among those factors. However, such an advantage still requires a large number of models to be introduced into the calculation, demanding huge computing procedures during the process. Therefore, this research only selected some commonly-examined factors to save computing time and resources. But there will still be numerous models to be processed, posing a great challenge to identifying the primary influential factors and their relationships. To simplify the screening and optimization process, this research established a Pareto front focusing on the solutions with minimum fitting errors and minimum model complexity. As shown in Figure 1, the dots represent the Pareto optimal solutions, and the line constituted by those dots is the Pareto front. The infeasible solution is on the left side

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202208_5(8).0034

of the Pareto front, while the solution on the right side is the feasible one. According to Occam's razor, when there are two models with the same accuracy, the one with less complexity will be preferred. As a result, a limited number of Pareto optimal solutions will be selected for further consideration. Therefore, these are four steps to determine the primary influential factors:

Select target variables and possible influential variables;

Apply the symbolic regression on the collected data to discover plenty of candidate models;

Establish the Pareto front to select the limited set of Pareto optimal solutions;

Investigate all the factors to determine the most frequently emerging ones in the Pareto optimal models.

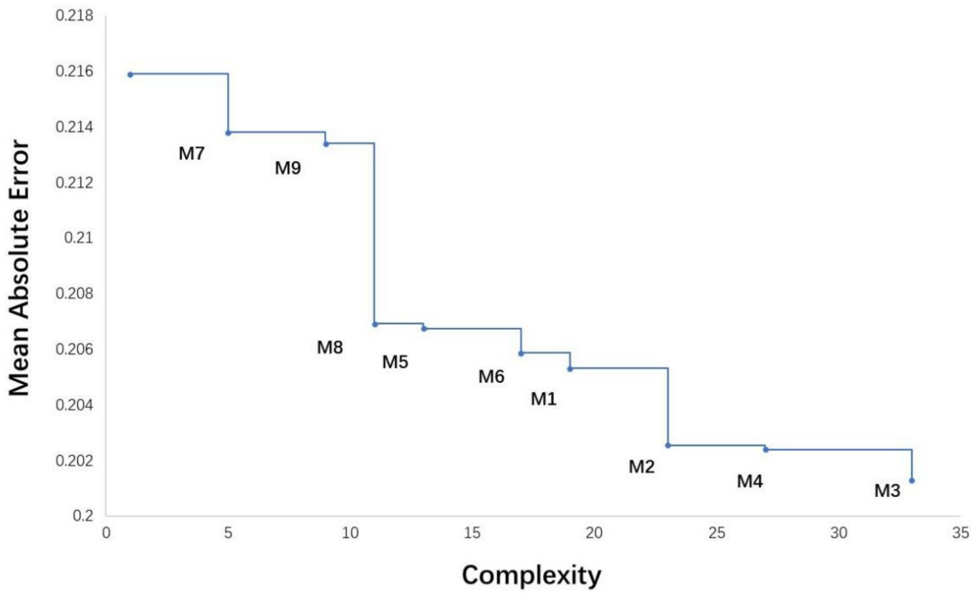

Figure 1. The Pareto front of this study.

3. Data

The original data used in this paper consists of the respondents' questionnaires an-swers and air detection. To guarantee these data's consistency and availability, this re-search considers the influences of both time and space. Therefore, this research collected questionnaire data three times a day and air detection data once an hour. Under the in-door environment in summer, the final calculation results are influenced by the selection of diverse variables. As a result, multiple air indexes are taken into account by the calculation, and the following variables are defined as:

Y – Learning Efficiency

CH2O - Average Formaldehyde Concentration, Unit: $\mu g \cdot m^{-3} \cdot h^{-1}$

CO2 – Average Carbon Dioxide Concentration, Unit: ppm

HUM – Average Humidity, Unit: $none$

PM10 – Average PM10 Concentration, Unit: $\mu g \cdot m^{-3} \cdot h^{-1}$

PM25 – Average PM2.5 Concentration, Unit: $\mu g \cdot m^{-3} \cdot h^{-1}$

TEMP – Average Temperature, Unit: °C

TVOC – Average Concentration of Total Volatile Organic Compounds, Unit: $\mu g \cdot m^{-3} \cdot h^{-1}$

Where Y is the dependent variable, CH_2O, CO_2, HUM, PM_{10}, $PM_{2.5}$, TEMP, TVOC are all independent variables.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202208_5(8).0034

In this study, the ratio of actual learning time to planned learning time is taken as the reference value, and the relevant data will be regarded as invalid and eliminated when its value of planned learning time is zero. Additionally, the values of all variables are normalized and deviated in the calculation, i.e., all the variables minus their mean value and then divided by the standard deviation, so that they all have the same scale and deviation, which enhances the model's calculation efficiency and accuracy.

4. Results and Discussions

4.1. Results Based on Symbolic Regression

For symbolic regression problems, this research needs to further select the best solutions among millions of possible candidate models discovered by the evolutionary process. Mean Absolute Error (MAE) and node count in the tree are two objectives newly introduced to determine models' superiority, with the former representing models' accuracy and the latter being a measurement of complexity. Therefore, another Pareto front could be built based on MAE and node count to select the optimal solutions among the numerous candidates.

The elements involved in the evolutionary process are necessitated during symbolic regression calculation. This study chooses the most common symbols that appear in the regression models as follows: constant, the input variable, + (addition), - (subtraction) and * (multiplication). According to the Pareto front (Figure 1), there are a limited number of models from evolutionary programming, and these Pareto optimal solutions will be further investigated to find out which factor presented major impacts on learning efficiency. Based on the symbolic regression method, nine Pareto optimal models are screened out by using model fitting. The calculated models are sorted by Applicability Ratio, and the results are shown in Table 1.

Table 1. Nine valid models after screening

ID	Model	C	MAE	Ratio
M1	$Y = 0.95 * TEMP + 0.012 * CO_2 - 16.61 - 0.00037 * CO_2 * TEMP - 1.05 * 10^{-6} * {CO_2}^2 - 0.012 * TEMP^2$	19	0.205	0.279
M2	$Y = 491.05 + 0.0093 * CO_2 + 3.54 * TEMP^2 + 0.00068$	23	0.203	0.261
M3	$Y = 613.65 + 0.013 * CO_2 + 4.41 * TEMP^2 + 0.00084 * TEMP^4 - 85.45 * TEMP - 0.00038 * CO_2 * TEMP - 9.88 * 10^{-7} * {CO_2}^2 - 0.010 * TEMP^3$	33	0.201	0.261
M4	$Y = 618.67 + 0.0099 * CO_2 + 4.43 * TEMP^2 + 0.00085 * TEMP^4 - 85.97 * TEMP - 0.00034 * CO_2 * TEMP - 0.10 * TEMP^3$	27	0.202	0.253
M5	$Y = 0.90 * TEMP + 0.011 * CO_2 - 15.47 - 0.00035 * CO_2 * TEMP - 0.012 * TEMP^2$	13	0.207	0.240
M6	$Y = 0.82 * TEMP + 0.010 * CO_2 - 14.13 - 0.00035 * CO_2 * TEMP - 0.010 * TEMP^2$	17	0.206	0.195
M7	$Y = 1.17 - 0.00042 * CO_2$	5	0.214	0.177
M8	$Y = 0.14 * TEMP + 0.0068 * CO_2 - 3.04 - 0.00025 * CO_2 * TEMP$	9	0.213	0.176
M9	$Y = 1.32 - 0.00047 * CO_2 - 0.0042 * TEMP$	11	0.207	0.165

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202208_5(8).0034

Table 1 demonstrates that the models with higher complexity are ranking relatively near the top, because models' accuracy generally improves with the increasing complexity. Model M1 shows the best applicability among all, as it can be applied to 27.9% of the respondents, or 183 out of 655. Beyond that, Model M5 also exhibits its great applicability in the less complex models, matching roughly 24% of the respondents.

4.2. Discussions

4.2.1. Analysis of Primary Factors Affecting Learning Efficiency

It can be observed that Pareto optimal models are the ones with relatively high accuracy and low complexity, while further research on them will benefit the study on the relationship between variables' characteristics and interactive effects. Figure 2 depicts the statistical graph of the variables' occurrence frequency, in which the left one shows the number of models each with a certain variable, and the right one describes the total number of occurrences of each variable. It has been seen that the primary factors that influence learning efficiency were TEMP and CO2. However, other factors like CH2O, HUM, PM2.5, PM10, TVOC did not appear at all, implying their weak influence on learning efficiency in this study.

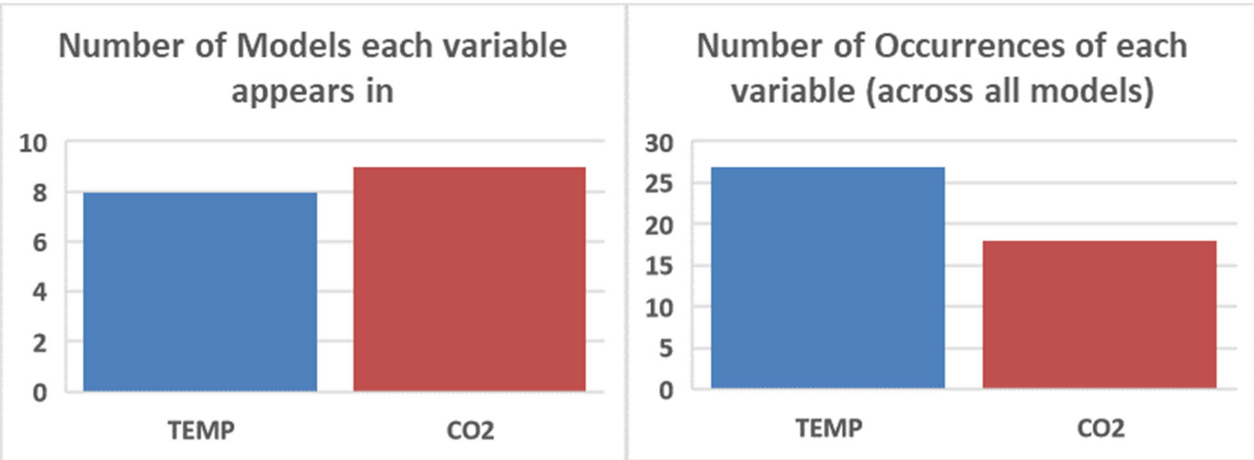

Figure 2. Statistical graphs of frequency of each variable's occurrence.

In this study, there are strict standards on air indexes (CH_2O, HUM, $PM_{2.5}$, PM_{10}, TVOC) in college classrooms. Because the variations of the above indexes are always within human bodies' permissible ranges, their impacts are fairly minor. However, the values of CO_2 and TEMP, on the other hand, change significantly in the indoor environment, which makes them easily perceived and hence exerted a greater influence.

From Figure 1, this research could observe that CO_2 appears in more models than TEMP does, while its total frequency is lower than that of TEMP, making it difficult to pinpoint the key factor affecting learning efficiency. Therefore, to obtain a more accurate result, this research calculated both the positive and negative correlations, as well as the respective sensitivities of learning efficiency to TEMP and CO_2, the results of calculations above are displayed in Table 2:

It can be seen from Table 2 that learning efficiency is more sensitive to the variable TEMP, indicating that temperature presents a greater influence on learning efficiency than carbon dioxide does. In other models, carbon dioxide mostly exerts a negative effect on learning efficiency, while those of temperature are equally positive and negative.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202208_5(8).0034

Table 2. Sensitivities and positive & negative correlations of indoor learning efficiency.

ID	Variable	Sensitivity	% Positive	Positive Magnitude	% Negative	Negative Magnitude
M1	TEMP	1.964	48%	2.582	52%	1.388
	CO2	0.191	43%	0.189	57%	0.193
M2	TEMP	1.445	36%	0.985	64%	1.703
	CO2	0.019	0%	0	100%	0.019
M3	TEMP	1.435	37%	1.026	63%	1.678
	CO2	0.134	45%	0.145	55%	0.126
M4	TEMP	1.453	36%	1.009	64%	1.708
	CO2	0.010	0%	0	100%	0.010
M5	TEMP	1.891	48%	2.459	52%	1.361
	CO2	0.017	100%	0.017	0%	0
M6	TEMP	1.669	49%	2.187	51%	1.177
	CO2	0.191	0%	0	100%	0.191
M7	CO2	1.000	0%	0	100%	1.000
M8	CO2	0.866	0%	0	100%	0.866
	TEMP	0.066	100%	0.066	0%	0
M9	CO2	1.095	0%	0	100%	1.095
	TEMP	0.276	0%	0	100%	0.276

4.2.2. Study on the Influence of Thermal Environment on Learning Efficiency

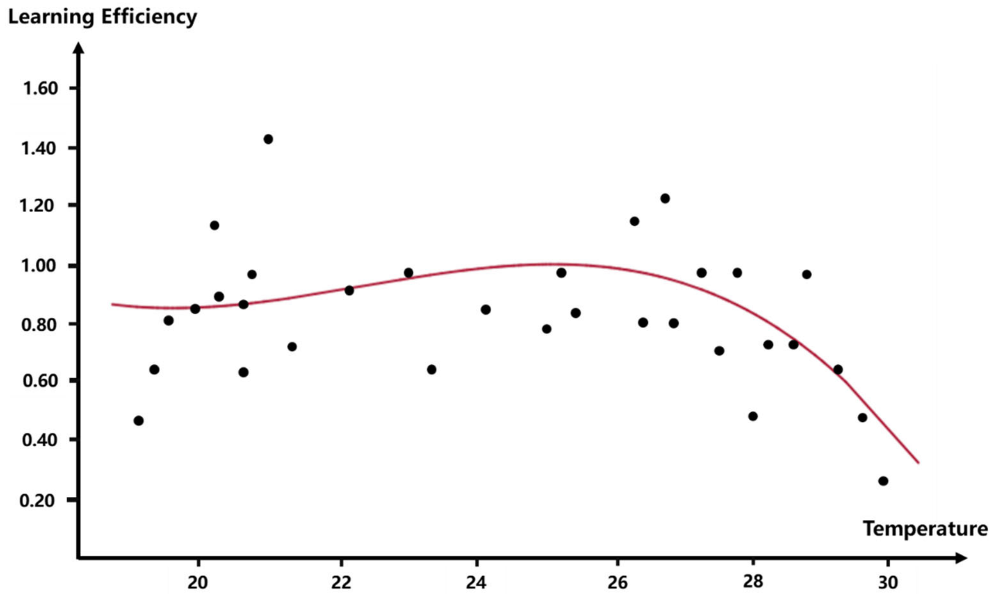

Figure 3. The relationship between learning efficiency&satisfaction and indoor average temperature

The relationship between learning efficiency and average indoor temperature(TEMP) is shown in Figure 3. It can be expressed as follows after curve-fitting:

$$Y = 20.05 - 2.64 * TEMP + 0.120 * TEMP^2 - 0.00179 * TEMP^3 \quad (1)$$

Figure 3 indicates that students' learning efficiency and satisfaction with the thermal environment follow a declining trend as indoor air temperature rises. In addition, it could be

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202208_5(8).0034

found that the temperature around 25°C resulted in both the highest learning efficiency and the highest satisfaction.

It can be seen that students' learning efficiency will be higher if the indoor temperature is below 19°C, such temperature generally occurs when the air conditioner is turned on. Compared to a higher temperature, people's cold sensation induces excitement at an appropriately low temperature.

The secretion of excitability hormone is encouraged correspondingly and no strain is put on the body's adaption mechanism, which benefits students' learning. Therefore, students will perform higher efficiency around 19 °C, while presenting a much lower efficiency at the temperature between 19 °C and 23 °C. During the experiment, people's physiological and psychological states could both be affected by some other objective factors associated with temperature. In this study, the low indoor temperature in summer is more common in the evenings, and less light causes both physiological and psychological discomforts, affecting learning efficiency negatively. Similarly, when the indoor air temperature at daytime is between 23°C and 27°C, sufficient sunlight will encourage a better mood and more sufficient energy, leading to the higher efficiency of learning. However, in a comfortable environment, which is commonly regarded as a warm one for learning, students will be more prone to fatigue, and their psychology will become more relaxed, which conflicts with the demand for strong concentration on learning. So, students' learning efficiency drops slightly with the rise of temperature; Besides, at a temperature exceeding 27°C, students' learning efficiency will decrease drastically. Studies suggested that higher temperature caused students to sweat more, increased a higher metabolism, which made them feel sleepy and tired. Furthermore, higher temperatures also made students more irritable, causing a significant decrease in learning efficiency.

4.2.3. Study on the Influence of Indoor Air Quality on Learning Efficiency

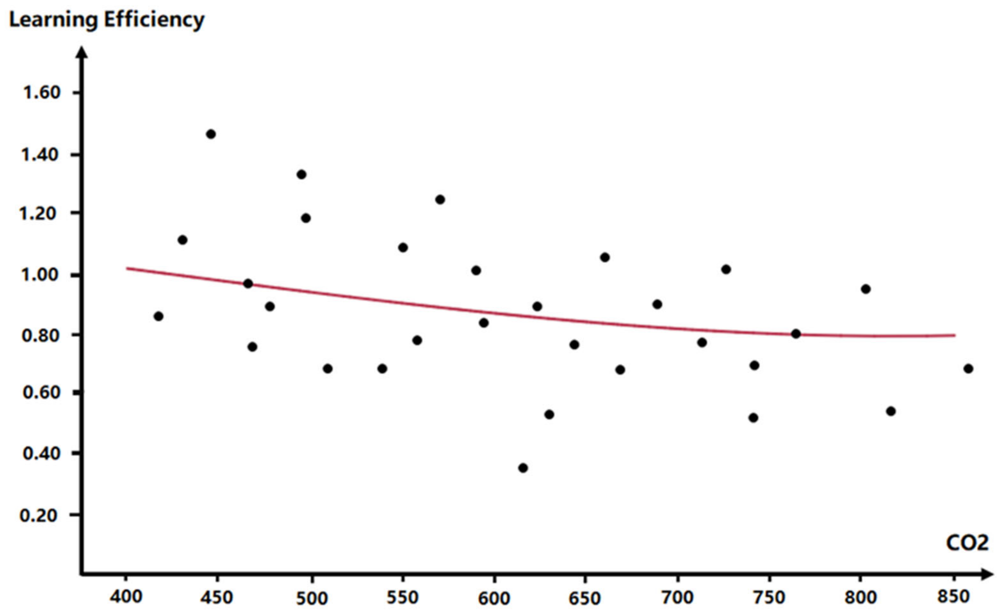

Figure 4. The relationship between learning efficiency and indoor carbon dioxide.

The relationship between learning efficiency and indoor carbon dioxide is shown in Figure 4. It can be expressed as follows after curve-fitting:

$$Y = 1.23 - 1.95 * 10^{-6} * {CO_2}^2 + 1.59 * 10^{-9} * {CO_2}^3 \tag{2}$$

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202208_5(8).0034

The results indicate that the higher the average CO_2 concentration of the indoor environment, the lower the satisfaction with indoor air quality and the learning efficiency. In particular, when CO_2 concentration is below 800ppm, students' learning efficiency decreases rapidly after the indoor average CO_2 concentration increases. However, students' learning efficiency remains stable when CO_2 concentration grows beyond 800ppm.

In the summer environment, the air conditioner in the classroom is mostly turned on while the windows are shut. Due to this lack of ventilation, carbon dioxide in the classroom keeps accumulating. Long-term exposure to high levels of carbon dioxide causes negative effects on a diversity of organs and systems, with symptoms of deepened respiration and delayed response. And in some cases, the indoor air environment is polluted by perfume, food and other odorous substances, substantially reducing the air freshness. As a consequence, the harmful influence of carbon dioxide on learning efficiency will be ignored when people are distracted by those peculiar smells. Moreover, the rising intensity of individuals further enhances CO_2's influence, resulting in more people suffering from decreasing learning efficiency.

5. Conclusion

This study uses a questionnaire and on-site measurements to collect the original data. In addition, the impacts of multiple environmental indexes on students' learning efficiency are studied by using symbolic regression, and two primary factors are discovered to affect the learning efficiency through Pareto optimal solutions. After the detailed analyses in the above sections, the conclusions are listed as follows:

(1) In this paper, the reference environmental indexes are $TEMP$, CO_2, CH_2O, HUM, $PM_{2.5}$, PM_{10}, $TVOC$, in which $TEMP$ and CO_2 occured most frequently in the models, whereas CH_2O, HUM, $PM_{2.5}$, PM_{10} and $TVOC$ did not appear at all. As a result, temperature and carbon dioxide are regarded as the primary factors of students' learning efficiency within the indoor environment in summer. The reason why other factors didn't appear is mainly related to the strict requirements for indoor air indexes. In addition, it can be observed that temperature exhibited a greater impact than CO_2 by comparing their frequencies and sensitivities.

(2) As the relationship between temperature and learning efficiency is investigated to further fit the formula being visualized afterwards, this research discovers that: students will show both the highest learning efficiency and the highest satisfaction with the thermal environment when the indoor temperature is around 25°C in summer, which also agrees with conclusions of the previous studies[4] in references. When the temperature turns too high, however, students' learning efficiency will drop drastically. Besides, learning efficiency could also be adversely impacted by insufficient indoor lighting.

(3) Similarly, the relationship between carbon dioxide levels and learning efficiency is also explored to further fit the formula being analyzed next with the graph, this research finds that: students' learning efficiency continuously declines as carbon dioxide concentration successively rises. Further investigation reveals that physiological discomfort could be derived from increased carbon dioxide levels, which interfered with the learning state. Moreover, negative impacts of carbon dioxide on learning state can be easily neglected due to other peculiar smells or perceived things. This conclusion is also supported by the prior studies [7] in references.

According to the patterns concluded in this study, the following suggestions are given:

The local climate conditions are to be considered for the colleges and universities education in different regions, and classrooms should be designed more reasonably. Besides, temperature-control equipment is supposed to be installed appropriately, such as air conditioners and heating systems, to rationally adjust the indoor temperature of classrooms. Not only that, but also greater attention is required for the impact of indoor lighting on students' learning efficiency, so that the room's lighting standards can be improved in favor of providing students with a more efficient and comfortable environment for working and learning.

The ventilation of rooms at colleges and universities is also supposed to be strengthened to guarantee the air quality and increase the maximum intake of fresh air for reducing the concentration of carbon dioxide. Simultaneously, the maximum capacity of people in a room must be sensibly rearranged, which will benefit students' learning efficiency as well.

Acknowledgments

This research received no external funding.

References

[1] J. Jiang, D. Wang, Y. Liu, et al. Experimental study on relationship between indoor temperature and learning efficiency of primary and middle school students in winter. Heating Ventilating & Air Conditioning, Vol.49 (2019) No. 05, p.99-105+85.

[2] H. Kim, T. Hong, J. Kim, et al. A psychophysiological effect of indoor thermal condition on college students' learning performance through EEG measurement. Building and Environment, Vol.184(2020), 107223.

[3] M.X. Li, X.Y. Wu, D. Zhang, et al. Investigation of Relationship between Indoor Environment and Learning Efficiency of Classrooms in Winter. Building Energy & Environment, Vol.31 (2012) No. 06, p. 30-33+47.

[4] K.F. Chang, P.C. Chou, Y.T. Lui: A Study on the Psychology Satisfaction Influenced by the Indoor Physical Environment Factors-Taking the Thermal Comfort Factor. (2007) Information on: https://www.pws.stu.edu.tw/paul/B_Conference/B144.pdf

[5] C. Jung, J. Awad.: Improving the IAQ for Learning Efficiency with Indoor Plants in University Classrooms in Ajman, United Arab Emirates. Buildings, Vol.11 (2021) No. 7, p. 289.

[6] P. Mao, S. Bei, J. Li, et al. Effects of indoor temperature on learning efficiency in university classroom. Journal of South-east University (Medical Science Edition), Vol.38 (2019) No.1, p. 168-173.

[7] N. Norazman, A. Ani, NH Ja€Afar, et al. Indoor Lighting in Classroom Environment Influences on Students€ Learning Per-formance. The Journal of Social Sciences Research, (2018) No.6 , p. 986-990.

[8] D.P. Wyon, P. Wargocki. How indoor environment affects performance. ashrae journal, Vol.55 (2013)No.3, p. 46-52.

[9] M.G. Pizon, E.F. Sagrado: Forecasting Disease Burden In Philippines: A Symbolic Regression Analysis. arXiv e-prints, (2021). Information on: https://doi.org/10.48550/arXiv.2105.04813

[10] H. Wang, G. Dong, J. Chen. Application of genetic programming in the identification of tool wear. Engineering Computations, (2021) Information on:. https://doi.org/10.1108/EC-08-2020-0470

[11] G.F. Yang, T. Sun, L. Lian, et al. Modeling Relationship between Economy and Industrial Pollution in China by Symbolic Regression. (Ningbo, China, October 25-27,2013)

[12] G.F. Yang, X.N. Li, J.L. Wang, et al. Modeling oil production based on symbolic regression. Energy Policy, Vol.82 (2015), p. 48-61.

[13] G.F. Yang, W.L. Li, J.L. Wang, et al. A comparative study on the influential factors of China's provincial energy intensity. Energy Policy, Vol.88 (2016), p. 74-85.

[14] G.F. Yang, T. Sun, J.L. Wang, et al. Modeling the nexus between carbon dioxide emissions and economic growth. Energy Policy, Vol.86 (2015), p. 104-117.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202208_5(8).0034

[15] W.L. Li, G.F. Yang, X.N. Li, et al. Cluster analysis of the relationship between carbon dioxide emissions and economic growth. Journal of Cleaner Production, Vol. 225(2019), p. 459-471.

[16] Y.Q. Wang, N. Wagner, J.M. Rondinelli: Symbolic regression in materials science. MRS Communications, Vol.9(2019) No.3, p. 793-805.

[17] X. Yang, G.F. Yang, J. Wu, et al. Modeling relationships between retail prices and consumer reviews: A machine discovery approach and comprehensive evaluations. Decision Support Systems, Vol.145(2021), 113536.

[18] J.R. Koza: Genetic programming: on the programming of computers by means of natural selection(MIT Press, U.S.A 1992).