Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 4, 1-11, 2024
Publisher: Learning Gate
DOI: 10.55214/25768484.v8i4.1097

Prediction of CO_2 emissions using machine learning

Kanyarat Bussaban[1*], Kunyanuth Kularbphettong[1], Nareenart Raksuntorn[1], Chongrag Boonseng[2]
[1]Faculty of Science and Technology, Suan Sunandha Rajabhat University, Bangkok, Thailand; kanyarat.bu@ssru.ac.th (K.B.), kunyanuth.ku@ssru.ac.th (K.K.), nareenart.ra@ssru.ac.th (N.R.)
[2]Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand; chongrag.bo@kmitl.ac.th (C.B.)

Abstract: Carbon dioxide (CO_2) contributes significantly to climate change as a greenhouse gas. The Earth's atmosphere is naturally kept warm enough to support life by greenhouse gases which trap heat in the atmosphere. However, human activity has significantly increased the amount of CO2 in the atmosphere because of deforestation and the use of fossil fuels. One of the key concerns with human evolution that fuels global climate change is carbon dioxide (CO_2). It is released as fuels burn and as a result, people worldwide are gradually becoming more conscious of environmental issues. Effective policy formulation requires an investigation of the factors influencing CO_2 emissions, yet tiny datasets and traditional research methodologies have hampered prior investigations. This research uses three prediction models to estimate CO_2 trapping efficiency among CO_2 emissions, energy use and GDP: Multiple Linear Regression (MLR), Support Vector Machine (SVM) and Random Forest (RF). The machine learning (ML) techniques used in this work have demonstrated strong performance with multiple linear regressions, support vector machines and random forest models with mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE). The investigation has proposed a technique for approximating CO_2 emissions and the results indicate that Support Vector Machine (SVM) can attain the highest degree of precision. The outcome could be a useful model for the decision support system to enhance an appropriate course of action for reducing CO_2 emissions worldwide.

Keywords: *Carbon dioxide (CO_2), CO_2 emissions, Multiple linear regression, Random forest, Support vector machine.*

1. Introduction

The effect of energy entering and leaving the Earth is the balanced temperature available for any living organism particularly humans. The radiation of the sun reaching the surface of the planet is either absorbed by the earth or reflected into space. The former subsequently releases some of the energy back into the atmosphere as heat, aka. This process helps balance the temperature on Earth and plays a crucial role in supporting life. Infrared radiation warms the planet but the latter does not. The energy is emitted and then absorbed with some specific gases, aka. The greenhouse effect is where infrared radiation is trapped by certain gases in the Earth's atmosphere leading to warming. This process is essential for maintaining Earth's temperature but human activities have intensified the greenhouse effect leading to concerns about climate change. Greenhouse gases already exist in the atmosphere and inevitably delaying or preventing some residual heat from entering space. The main reason for the unexpected rise in global warming is this mechanism which unintentionally acts as an internal barrier preventing emission from going back into space. The greenhouse effect is generally natural and fairly essential to supporting life. Nevertheless, their recent accumulation of greenhouse

History: Received: 22 January 2024; Revised: 5 April 2024; Accepted: 26 April 2024; Published: 17 May 2024
* Correspondence: kanyarat.bu@ssru.ac.th

gases has already altered the earth's climate attributing dangerously to the well-being of humans and also to ecosystems [1].

According to the World Meteorological Organization (WMO), the primary greenhouse gases being observed, analyzed and reported are carbon oxide (CO_2), methane , nitrogen oxide , halocarbon and hydrogen [2]. These gases are either released naturally or emitted as part of human activities. The former involves decomposition, respiration and the latter includes but is not limited to the burning of fossil fuels, industrial production, crop plantations or raising livestock. CO_2 plays a significant and critical role in the discern of global warming particularly the marked change in its level within a short time concerning its so-called carbon cycle resulting in the impact as already witnessed, of rising annual global temperature as the major contributing factor to climate change [3]. This change does not happen in isolation but is related to others which collectively accelerate and exacerbate the planet's overall climate at present. The altered climate phenomenon from the past and present is considered climate change. According to Canada [4] climate change is due to a long-term shift in the average weather conditions of a region such as its typical temperature, rainfall and windiness meaning that the range of conditions expected in many regions will change over the coming decades implying that there will be changes in extreme conditions.

Machine learning plays an important role in reducing CO_2 emissions by identifying areas where emissions can be reduced or avoided, optimizing energy systems and improving the accuracy of emissions measurements. For example, machine learning algorithms can be used to analyse satellite data to track land use changes and identify areas where deforestation is occurring which can be a significant source of CO_2 emissions. By analyzing this data, policymakers and companies can make more informed decisions about where to focus their efforts to reduce emissions.

In this paper, we propose a carbon dioxide emissions prediction model based on machine learning techniques. This paper is structured and organized as follows: a background introduction is presented in this section. Section 1 is followed by section 2 on the related works. Sections 3 and 4 describe the system framework and methods employed along with the corresponding results. Section 5 is the last section that concludes the paper.

2. Literature Review

Several investigations have been carried out to track the problems with CO_2 emissions Yuan, et al. [5], Liu and Hao [6], Munir, et al. [7] and Luccioni, et al. [8]. Masini, et al. [9] state that high-dimensional models and supervised machine learning (ML) are taken into consideration for time series in predicting and mixing ensemble and hybrid model elements from several options. The SVM model was proposed [10] and the model's evaluation was conducted using the RMSE to estimate carbon dioxide (CO_2) emissions from energy consumption. CO_2 emissions were calculated using a linear model (LM) and support vector machines in the framework of the cost-effectiveness analysis [11]. Decision tree machine learning models are used recursively to split data that minimizes the residual sum of squares and random forest techniques have been used in the past to investigate carbon sinks in soils and forests [12, 13]. The goal of energy resources and economic development is to create and implement a machine learning approach that will increase the predictive accuracy of GDP predictions [14].

These assessments of the literature point to the application of machine learning algorithms as a more precise algorithm for CO_2 emission prediction. Nonetheless, the best approach depends on the particular dataset and problem.

3. Experimental Setup

This section mainly discusses the data processing, selected models, experiment framework and evaluation metrics used in the study.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 4: 1-11, 2024
DOI: 10.55214/25768484.v8i4.1097

3.1. Materials and Data Processing

The research used available annual data for CO_2 emissions from 1997 to 2021. The primary data were retrieved from this repository [15-17]. The features and labels were extracted from the raw datasets. The dataset comprises GDP, per capita income, power industry, transportation and construction. These factors are all crucial elements that can impact the release of CO_2 into the atmosphere. GDP statistics and CO_2 emissions serve as vital metrics that offer valuable insights into both the economic and environmental performance of a nation. There exists a well-documented positive relationship between GDP and CO_2 emissions as countries with higher GDP tend to exhibit higher levels of CO_2 emissions. The power industry stands out as a major contributor to CO_2 emissions, responsible for a significant share of global emissions. Similarly, the transport sector represents another substantial source of CO2 emissions contributing roughly one-fifth of global emissions. This phenomenon can be attributed to the heavy reliance of transportation on fossil fuels like gasoline and diesel. One-third of all CO2 emissions come from buildings which is a significant contribution. This is primarily due to the energy consumption in buildings for heating, cooling and lighting often derived from fossil fuel reservoirs.

Data preprocessing is a significant step in the data analysis process that cleans and transforms raw data into an understandable format. Data preprocessing can help improve the accuracy and robustness of the model. The data set is necessary to check for errors like duplicate data, incorrect data, missing value or outliers. It is the first and crucial step in creating a machine learning model. Table 1 summarizes the descriptive statistics of the raw data.

Table 1.
Descriptive statistics of the raw data set.

Features	Mean	Std.	Min	25%	50%	75%	Max
CO2 emissions per GDP	26.988	1.836	23.500	26.350	27.040	28.100	30.750
CO2 emission capita	3.453	0.434	2.640	3.080	3.540	3.840	3.970
Power industry	76.221	13.232	54.271	62.967	79.296	87.837	92.380
Building	15.832	2.550	11.497	13.902	15.666	17.474	21.230
Transport	57.178	8.245	44.846	51.976	56.128	61.741	73.624
CO2 emission	2.062	0.089	1.930	2.000	2.030	2.130	2.270

The data set was scaled according to the equation to make the data equal in size.

$$x'_i = \frac{x_i - \bar{x}}{S.D} \quad (1)$$

Let x be the data at that point. $\bar{x}$ is the mean of the data and $S.D.$ is the standard deviation of the data.

Figure 1 shows the proposed framework for CO_2 emissions. The steps involved in using machine learning algorithms for CO_2 emission prediction are as follows:

1. Collect a dataset of CO_2 emissions and corresponding input factors. This dataset should include a range of input factors that are relevant to CO_2 emissions.
2. Preprocess the dataset by cleaning the data removing any missing or invalid values and normalizing the input factors to ensure that they are on the same scale.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 4: 1-11, 2024
DOI: 10.55214/25768484.v8i4.1097

3. Split the dataset into two sets: a training set and a testing set. The training set is used to train the machine learning model while the testing set is used to evaluate the performance of the model. Cross-validation was applied at this stage to divide the data set.
4. Train the machine-learning model using the training set. This involves selecting an appropriate kernel function, tuning the hyper parameters of the model and optimizing the weight coefficients and the bias term.
5. Test the machine learning model using the testing set. This involves using the trained model to predict the CO_2 emissions for each input combination in the testing set and then comparing the predicted values to the actual values.
6. Evaluate the performance of the machine learning model using the MAPE metric. The metric can help determine how well the machine learning model can predict CO_2 emissions based on the input factors.

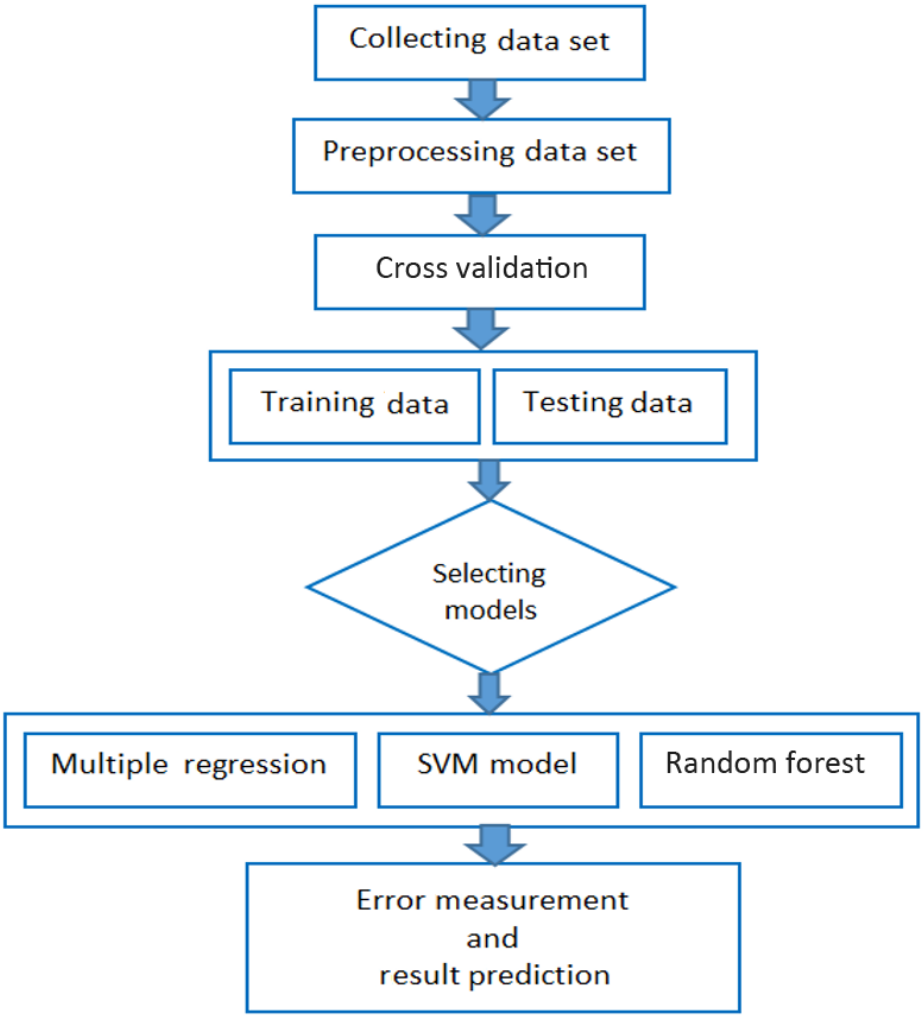

Figure 1.
The proposed framework for CO_2 emissions.

SVM, random forest and MLR are all machine learning algorithms that can be used for CO_2 emission prediction or other related tasks. Each algorithm has its strengths and weaknesses.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 4: 1-11, 2024
DOI: 10.55214/25768484.v8i4.1097

3.2. Methods

3.2.1. Cross-Validation

In machine learning, cross-validation is a method used to assess a model's performance on a dataset. A training set and a testing set are often created from the dataset. After training on the training set, the model is assessed on the testing set. K-fold cross-validation is the most popular type of cross-validation. In this method, the dataset is divided into k folds of equal size. The model is trained k times using one fold as the training set and the remaining folds as the testing set.

The formula for k-fold cross-validation can be shown as follows:

1. Divide the data into k equal-sized subsets.
2. For each subset
   a. Train the model on the remaining k-1 subsets.
   b. Evaluate the model on the current subset.
3. Compute the average evaluation metric across all k subsets to obtain the final model performance.

Cross-validation is a crucial technique in machine learning for ensuring that models are robust and generalizable to new data.

3.2.2. Multiple Linear Regression (MLR)

A proficient approach for simulating the linear relationship between the input and target variables is multiple linear regression (MLR). MLR can be beneficial in determining the main variables influencing CO_2 emissions and quantifying the extent of those variables' effects. The link between energy use and CO_2 emissions in different nations has been modelled using MLR. The independent and dependent variables' causal linkages are determined through the use of multiple linear regression analysis. The MLR model in the equation is as follows:

$$\hat{y} = w_0 + w_1x_1 + w_2x_2 + \cdots + w_Dx_D + \epsilon \quad (2)$$

$$W = (Xb^TXb)^{-1}Xb^TY \quad (3)$$

Where w_0, w_1, w_2,…, w_d causes the least error

$$Error = \sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (4)$$

In this equation, w represents the coefficients of independent variables. ϵ is referred to as error terms that are not detected from the data.

3.2.3. Support Vector Machines (SVM)

Support Vector Machines (SVM) are one of the powerful algorithms of machine learning used for both classification and regression analysis. SVMs operate by determining the best hyperplane for predicting a continuous target variable or by dividing the data into different categories. They are also robust to overfitting and can handle noisy or incomplete data. SVM has been used to predict CO_2 emissions in power plants, forecast CO_2 emissions in many studies and perform other related tasks.

$$h_\theta(x) = w_1x_1 + w_2x_2 + \cdots + w_nx_n + b \quad (5)$$

$$= w^Tx + b$$

Where $x_i \in R$, R is the domain of real numbers.
w_i is the n-dimensional weight vector.
b is the bias value.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 4: 1-11, 2024
DOI: 10.55214/25768484.v8i4.1097

Finding the optimal hyperplane is to find the optimal w and b and it can be expressed as follows:

$$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i \tag{6}$$

$$\xi_i \geq 0$$

$$y_i(\vec{w}\vec{x}_i + b) \geq 1 - \xi_i \tag{7}$$

$$-\xi_i \leq 0$$

$$1 - \xi_i - y_i(\vec{w}\vec{x}_i + b) \leq 0 \tag{8}$$

According to the LaGrange function [18], it can be minimized as follows:
Constraint: $g_i(x) = 0$; i = 1, 2 . . . , n

$$L(x, \lambda) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i - \sum_{i=1}^{n}\lambda_i\,[y_i(w^T x_i + b) - 1 - \xi_i)] \tag{9}$$

Where
x_i is the ith data point in the training set.
y_i is the corresponding target value.
w is the weight vector that defines the separating hyperplane.
b is the bias term that shifts the separating hyperplane.
ξ_i is the slack variable that allows for misclassification errors.
λ_i is the Lagrange multiplier that enforces the constraints.
C is the penalty parameter that controls the trade-off.

3.2.4. Random Forest

Random forest is a supervised learning technique that combines several decision trees to improve prediction accuracy based on the ensemble learning method. It can handle high-dimensional input data and is resilient to noisy data [19]. Ensemble learning is a strategy that aggregates predictions from multiple machine learning algorithms to address a specific problem. Boosting, bootstrap aggregating and stacking are some examples of ensemble learning types [20, 21].

Conducting predictions using a random forest generally requires the aggregation of predictions made by each individual tree to derive a conclusive prediction as shown in Figure 2. Each decision tree within the random forest algorithm autonomously generates a prediction by analyzing the input features. Once all of the trees have generated their forecasts, the ultimate forecast for a specific input is derived by combining the forecasts from all of the trees. Random forests have the capability to generate predictions that are not only more precise and consistent than those produced by singular decision trees but also to alleviate the issue of overfitting by combining the forecasts generated by numerous trees. This attribute stands out as a fundamental advantage of the random forest algorithm.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 4: 1-11, 2024
DOI: 10.55214/25768484.v8i4.1097

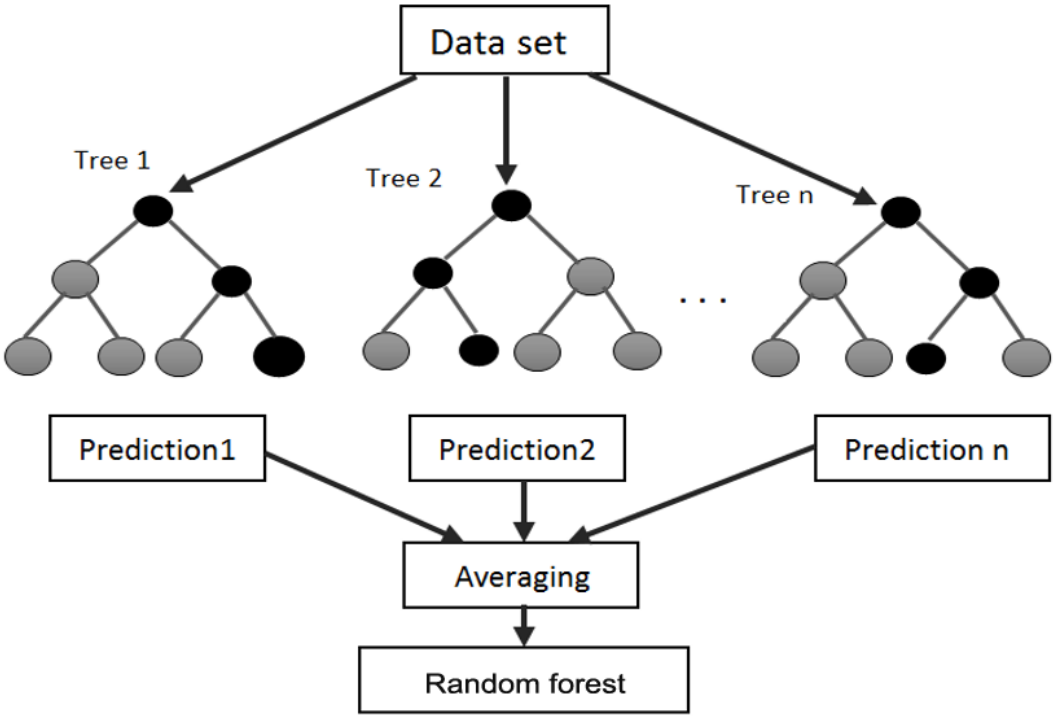

Figure 2.
The structure of a random forest.

Random forest has been used to predict CO_2 emissions in power plants, forecast CO_2 emissions and perform other related tasks.

3.2.5. Mean Absolute Error (MAE)

The Mean Absolute Error (MAE) is a commonly used metric for evaluating the performance of machine learning models including those used to predict CO_2 emissions.

The MAE is calculated as follows:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|\hat{y}_i - y_i| \quad (10)$$

Where n is the number of observations in the dataset.
y_i is the actual value of the target variable for the i^{th} observation.
$\hat{y}_i$ is the predicted value of the target variable for the i^{th} observation.

The MAE is computed by adding up all of the individual observations' absolute differences between the predicted and actual values dividing the result by the total number of observations and then calculating the mean absolute error. The MAE would show how effectively the model can forecast the actual CO_2 emissions based on these predictors.

3.2.6. Mean Absolute Percentage Error (MAPE)

The mean absolute percentage error (MAPE) was used to assess the model. A forecasting model's prediction accuracy in terms of estimating the percentage error between the actual and predicted outcomes is measured by the mean absolute percentage error or MAPE. Following is the calculation of the MAPE:

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 4: 1-11, 2024
DOI: 10.55214/25768484.v8i4.1097

$$MAPE = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \quad (11)$$

Where n is the number of observations in the dataset.
y_i is the actual value of the target variable for the i^{th} observation.
$\hat{y}_i$ is the predicted value of the target variable for the i^{th} observation.

MAPE is commonly used because it's easy to interpret and understand and MAPE is expressed as a percentage. A lower MAPE indicates better forecasting accuracy.

3.2.7. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is a generally used metric for evaluating the performance of machine learning models including those used to predict CO_2 emissions. Like MAE, the RMSE should be applied in conjunction with other metrics and qualitative analysis to gain a more complete understanding of the model's strengths and weaknesses. The formula for RMSE is as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{y}_i - y_i)^2} \quad (12)$$

Where n is the number of observations in the dataset.
y_i is the actual value of the target variable for the i^{th} observation.
$\hat{y}_i$ is the predicted value of the target variable for the i^{th} observation.

The RMSE is calculated by taking the squared difference between the predicted and actual values of the target variable for each observation summing these squared differences dividing by the total number of observations and taking the square root of the result.

4. Results

The datasets in the aforementioned section were partitioned into training and testing sets post-data preparation using the K-fold cross-validation methodology. The implementation of the suggested framework was carried out using the Sci-Kit library in Python. Performance evaluations were conducted using MAE, MAPE and RMSE metrics within this study. The outcomes of multiple linear regression, SVM and random forest regression are presented in Table 2 and in Figure 3. A proposed method for estimating CO_2 emissions has been introduced in the research revealing that SVM exhibits the highest level of accuracy. Following SVM, random forest emerges as the subsequent best algorithm due to its superior MAE value and the second-highest rankings for MAPE and RMSE. In contrast, the support vector machine approach performs commendably with the lowest values for MAE, MAPE and RMSE.

Table 2.
Model comparison.

Methods	MAE	MAPE	RMSE
Multiple linear regression	0.22	3.15	0.27
Support vector machine	0.17	2.75	0.21
Random forest	0.18	2.91	0.23

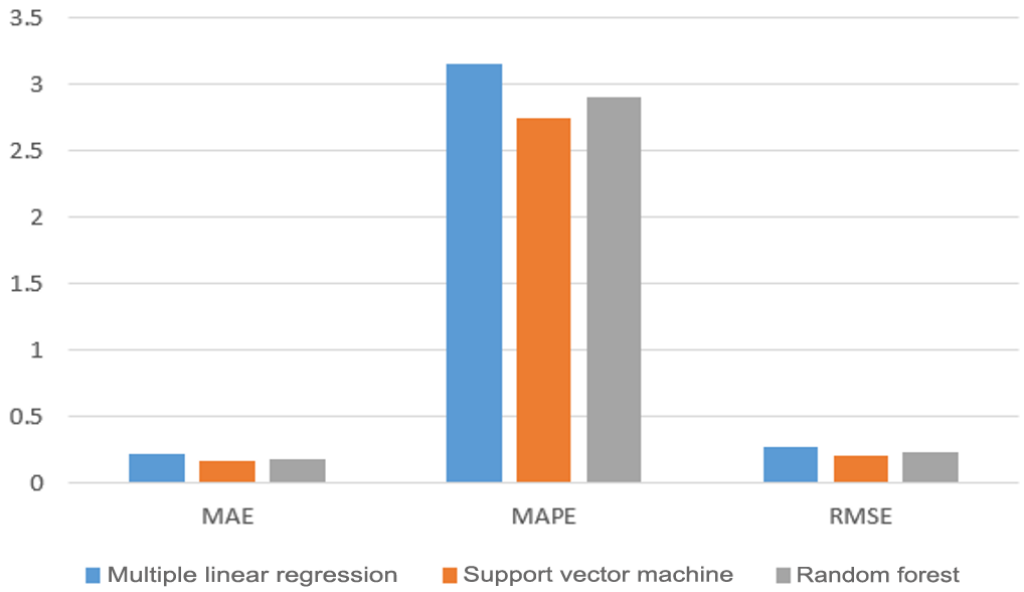

Figure 3.
The performance comparison graph of three machine learning algorithms.

5. Conclusion

CO_2 emissions are a significant contributor to climate change which has numerous effects on the environment and human societies like global warming, ocean acidification, air pollution, ecological changes and economic impacts. Machine learning has great potential to contribute to the reduction of CO_2 emissions by providing insights and tools for better decision-making, monitoring and optimization of energy consumption and emission sources. Moreover, multiple linear regression, linear regression, support vector machine and random forest are all sophisticated machine learning techniques used to develop models for CO_2 emissions and determining the factors influencing them. The application of SVM, MLR and random forest in the context of CO_2 emission modeling within this research necessitates meticulous data preprocessing, feature selection, model training, assessment and implementation to predict future CO_2 emissions based on historical data and external variables. Assessment of the machine learning algorithms' effectiveness in forecasting CO_2 emissions involves the utilization of performance metrics such as MAE, MAPE and RMSE. Model prediction values derived from this machine learning provided empirical data for trends in CO_2 emissions and supported policies for CO_2 reduction.

In a nutshell, reducing CO_2 emissions is critical to mitigating the effects of climate change. Machine learning can be a valuable tool to help us achieve this goal. Machine learning is a valuable approach to combating climate change helping us to better understand the causes and effects of CO_2 emissions and to develop effective strategies for reducing them. Future research in this area can be directed to analyze the other factors that are expected to influence the rate of power consumption, the energy efficiency index and carbon dioxide emissions. Furthermore, the development of web and mobile applications is predicted by models derived from machine learning.

Funding:

This research is supported by Suan Sunandha Rajabhat University, Thailand (Grant number: PD0 3-6602009).

Institutional Review Board Statement:

Not applicable.

Transparency:
The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Competing Interests:
The authors declare that they have no competing interests.

Authors' Contributions:
All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

Copyright:

References
[1] US EPA, "Climate change indicators: Heavy precipitation [Reports and Assessments]," Retrieved: https://www.epa.gov/climate-indicators/climate-change-indicators-heavy-precipitation. 2016.
[2] Greenhouse Gases, "World meteorological organization," Retrieved: https://wmo.int/topics/greenhouse-gases#:~:text=Greenhouse%20gases%20(GHGs)%20in%20the,warming%20of%201.1%20%C2%B0C.&text=Carbon%20dioxide%20(CO2)%20in%202021,higher%20than%20pre%2Dindustrial%20levels. 2022.
[3] L. Amarpuri, N. Yadav, G. Kumar, and S. Agrawal, "Prediction of CO2 emissions using deep learning hybrid approach: A case study in Indian context," presented at the 2019 Twelfth International Conference on Contemporary Computing (IC3). https://doi.org/10.1109/IC3.2019.8844902, 2019.
[4] E. C. C. Canada, "Climate change concepts," Retrieved: https://www.canada.ca/en/environment-climate-change/services/climate-change/canadian-centre-climate-services/basics/concepts.html. [Accessed 2018.
[5] J. Yuan, Y. Xu, and X. Zhang, "Income growth, energy consumption, and carbon emissions: The case of China," *Emerging Markets Finance and Trade*, vol. 50, no. 5, pp. 169-181, 2014.
[6] Y. Liu and Y. Hao, "The dynamic links between CO2 emissions, energy consumption and economic development in the countries along "the Belt and Road"," *Science of the Total Environment*, vol. 645, pp. 674-683, 2018. https://doi.org/10.1016/j.scitotenv.2018.07.062
[7] Q. Munir, H. H. Lean, and R. Smyth, "CO2 emissions, energy consumption and economic growth in the ASEAN-5 countries: A cross-sectional dependence approach," *Energy Economics*, vol. 85, p. 104571, 2020. https://doi.org/10.1016/j.eneco.2019.104571
[8] A. Luccioni, A. Lacoste, and V. Schmidt, "Estimating carbon emissions of artificial intelligence [opinion]," *IEEE Technology and Society Magazine*, vol. 39, no. 2, pp. 48-51, 2020. https://doi.org/10.1109/mts.2020.2991496
[9] R. P. Masini, M. C. Medeiros, and E. F. Mendes, "Machine learning advances for time series forecasting," *Journal of Economic Surveys*, vol. 37, no. 1, pp. 76-111, 2023.
[10] C. Saleh, N. R. Dzakiyullah, and J. B. Nugroho, "Carbon dioxide emission prediction using support vector machine," presented at the IOP Conference Series: Materials Science and Engineering, 2016.
[11] M. Rao, "Machine learning in estimating CO 2 emissions from electricity generation. In Engineering Problems-Uncertainties, Constraints and Optimization Techniques," 2021. https://doi.org/10.5772/intechopen.97452
[12] R. L. M. Tavares, S. R. d. M. Oliveira, F. M. M. D. Barros, C. V. V. Farhate, Z. M. d. Souza, and N. L. Scala Junior, "Prediction of soil CO 2 flux in sugarcane management systems using the random forest approach," *Scientia Agricola*, vol. 75, pp. 281-287, 2018. https://doi.org/10.1590/1678-992x-2017-0095
[13] J. Li, L. Pan, M. Suvarna, Y. W. Tong, and X. Wang, "Fuel properties of hydrochar and pyrochar: Prediction and exploration with machine learning," *Applied Energy*, vol. 269, p. 115166, 2020. https://doi.org/10.1016/j.apenergy.2020.115166
[14] D. Cogoljević, M. Alizamir, I. Piljan, T. Piljan, K. Prljić, and S. Zimonjić, "A machine learning approach for predicting the relationship between energy resources and economic development," *Physica A: Statistical Mechanics and its Applications*, vol. 495, pp. 211-214, 2018. https://doi.org/10.1016/j.physa.2017.12.082
[15] Energy Policy and Planning office (EPPO), "Energy policy and planning office," Retrieved: https://www.eppo.go.th/index.php/en/en-energystatistics/co2-statistic?fbclid=IwAR3lkNVWEVMxs_ZlSMAnHy9i501OOwy2NcGgy3fWyz5gyMW-g-rmenJGpGY. 2023.
[16] CO2 Emissions, "CO2 emissions," Retrieved: https://www.worldometers.info/co2-emissions/. 2023.

[17] EDGAR - Emissions Database for Global Atmospheric Research, "Emissions database for global atmospheric research," Retrieved: https://edgar.jrc.ec.europa.eu/. 2023.
[18] H. Li, *Lagrange multipliers and their applications; department of electrical engineering and computer science*. Knoxville, TN, USA: University of Tennessee, 2008.
[19] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 832-844, 1998. https://doi.org/10.1109/34.709601
[20] L. Breiman, "Random forests," *Journal Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
[21] S. Hartshorn, "Machine learning with random forests and decision trees: A Visual guide for beginners," *Kindle Edition*, 2016.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 4: 1-11, 2024
DOI: 10.55214/25768484.v8i4.1097