IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 17 July 2023, accepted 11 October 2023, date of publication 16 October 2023, date of current version 23 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3324725

RESEARCH ARTICLE

Predictive Analytics Model for Optimizing Carbon Footprint From Students' Learning Activities in Computer Science-Related Majors

MICHAEL HANS, ERNA HIKMAWATI, AND KRIDANTO SURENDRO
School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung 40116, Indonesia
Corresponding author: Michael Hans (23522011@std.stei.itb.ac.id)

ABSTRACT Global warming poses a significant challenge to environmental sustainability due to the high greenhouse gas emissions originating from human activities. Among the various sectors contributing to these emissions, the education sector, particularly at the university level, plays a crucial role in generating carbon footprints. While several studies have examined carbon footprints within universities, few have focused on understanding the behavior and activities of individuals. Statistics show that students are the primary contributors to carbon emissions in higher education. Despite the availability of various methods for calculating carbon emissions, limited studies have utilized such data to predict future trends. Therefore, this study aimed to develop a predictive analytics model that leveraged students' learning activities as a significant factor to predict future trends in university carbon emissions. Institut Teknologi Bandung (ITB) was used as a case study, especially computer science-related majors. The carbon emission calculation utilized a formula that incorporated various emission sources, including electricity, transportation, and paper consumption. In the 2022/2023 academic year, ITB generated 612.8 tons of CO_2e. The prediction modeling employed the SVR algorithm and utilized historical data, such as carbon emissions from the last 30 days, and external information, such as weather, event-related data, and university data. The model's performance was evaluated using metrics, yielding values of 129.41 for MAE, 158.03 for RMSE, 0.98 for R2, and 15.83 for MAPE. The results provided insights for universities to assess their carbon footprint and raised awareness among the academic community, supporting decision-making to optimize carbon emissions at the students' level.

INDEX TERMS Carbon footprint, information technology, predictive analytics, students.

I. INTRODUCTION

Global warming has emerged as a critical issue demanding attention to ensure environmental and planetary sustainability. The increasing climate change worldwide was primarily driven by the rise in greenhouse gases [1]. Among these gases, carbon emissions from human activities are one of the largest contributors. These emissions originated from various sources, including transportation, electricity consumption, and the industrial sector [2]. Between 1970 and 2011, global carbon emissions resulting from fossil fuel combustion witnessed a significant increase of up to 90% [2]. Within various sectors, higher education, as part of the education sector, was accountable for greenhouse gas due to its operational activities.

Several universities have undertaken assessments of their carbon footprints, providing valuable insights into their emissions. For instance, a study conducted in Mexican universities revealed that electricity usage accounted for 42% of greenhouse gas emissions, while transportation contributed to 50% [3]. Similarly, at the University of Patras in Greece, students' commuting and electricity usage were found to contribute to 60.2% of the total emissions [4]. In-depth studies conducted at the Norwegian University of Technology and Science demonstrated that students in social departments emitted 50% less carbon compared to those in technology

The associate editor coordinating the review of this manuscript and approving it for publication was Michele Magno.

and natural sciences departments [5]. While numerous studies have estimated carbon emissions, only a few have delved into exploring human behavior as a causative factor.

Behavioral factors and habits were responsible for approximately 30 and 50% of overall heating and cooling energy consumption, respectively [6]. By implementing behavioral adjustments, organizations can achieve energy savings of 10-20% without incurring additional equipment costs [7]. In the context of higher education, students' behavior was particularly influential. Carbon footprint analysis indicated that the average emissions per student were 3.84 tons of CO_2e, with 65, 20, and 15% attributed to daily activities, transportation, and academic activities such as studying, respectively [8].

The case study university chosen for this research was the Bandung Institute of Technology (ITB), which was projected to have 23,000 students in 2022 [9]. The academic community at ITB was primarily composed of students, accounting for 94.28% of the total, emphasizing their substantial contribution to carbon emissions. These emissions result from their daily activities, electronic usage, and transportation to campus. Although various carbon accounting methods exist to calculate emissions from these activities, the results have not been effectively utilized for predicting future trends.

Predictive Analytics offered a concept for anticipating future trends and has discovered broad applications across different industries [10], [11]. In the banking sector, it was employed for fraud detection [12]. This analytics aids in disease diagnosis by analyzing health records from thousands of patients with similar conditions [13]. Furthermore, it was utilized by manufacturers for inventory management [14] and to develop effective pricing strategies based on market trends [10], such as in stock market analysis [15]. Sustainability Analysis, specifically Life Cycle Assessment (LCA), employed predictive analytics to identify stages of the product life cycle with the most significant environmental impact and calculate the Carbon Footprint [16]. Therefore, the utility spans diverse fields, making it a valuable tool in various industries.

Several studies applied machine learning to predict future carbon emissions [17], [18], [19], [20], [21]. The majority utilized global or country carbon emission data for predicting the model-building process [17], [18], [19]. From the historical carbon data, emission in the next few years was forecasted utilizing machine learning [17], [18], [19]. Some studies relate the data to other factors, such as a country's GDP per capita, human population, and the energy intensity of industries [20]. However, most of these studies use historical carbon emission data on a national scale. There are very few studies that utilize day-to-day data from human activities and take it further to trend prediction. It is a big challenge to obtain real-time activity data that can be processed into carbon emission data that can be used in the process of building a prediction model. Therefore, this research will try to raise the integration of activity data, especially at the university level, to determine the trend of carbon emissions in universities based on student learning activities.

This information is used to reduce carbon emissions at the institutional level. The measured emissions serve as inputs for prediction, as well as for identifying patterns or behaviors of students that significantly contribute to carbon production. By leveraging the concept of predictive analytics, it becomes possible to predict carbon emissions resulting from students' activities and behaviors. This leads to the identification of specific behaviors that can optimize energy consumption related to student learning activities. Therefore, this study aims to develop a predictive analytics model to predict the use of university carbon emissions based on the behavior and learning activities conducted by students. The model is expected to identify actionable steps regarding students' learning activities to reduce carbon emissions, utilizing the generated predictive analytics model. In the initial phase, this research will be limited to students within the computer science-related majors as an initial case study in developing a prediction model for student learning activities, with subsequent expansion to the entire university level.

II. RELATED WORKS

This section will discuss all related works which are used as the main reference and used in developing solutions.

A. CARBON FOOTPRINT CALCULATION

The carbon emissions of each source were calculated with the following formula as in (1), where E_s, AD_s, and EF_s denote the emission from source S, the quantitative measure to express a unit (such as kWh in electricity use), and the coefficient to convert to CO_2e, respectively [22].

$$E_S = AD_S \cdot EF_S \tag{1}$$

Based on the general formula above, it was possible to calculate carbon emissions from each source, given the availability of the emission factor. To achieve this, the obtained list of these factors was then used to derive the calculation formula for each defined source.

1) COMMUTING EMISSION CALCULATION

To calculate carbon emissions from commuting activities, the following formula in (2) will be used. It is important to note that the emission factors of each transportation mode may vary and the figures used will be specific to Indonesia [23].

$$E_{commuting} = Ef_{commuting}(v) \,.\, S \tag{2}$$

Description:

$E_{commuting}$ = Emissions generated from commuting activities (g CO_2e)

$Ef_{commuting}(v)$ = Emission factor from transportation mode v for each kilometer (g CO_2/km)

S = Travel distance from origin to destination (km)

2) ELECTRICITY EMISSION CALCULATION

To carbon emissions from electricity, the following formula in (3) will be applied. The electricity emission factor used was the value applicable in Indonesia [24].

$$E_{electricity} = Ef_{electricity}.W$$
$$W = \frac{P.t}{1000} \tag{3}$$

Description:
$E_{electricity}$ = Emissions from the use of electricity
$Ef_{electricity}$ = Emission factor produced for each kWh
W = Electrical energy consumed (kWh)
P = Power of an electronic device (Watt)
t = Consumption time of electricity (hour)

3) PAPER WASTE EMISSION CALCULATION

To calculate carbon emissions from paper consumption, the following formula in (4) will be applied. The emission factor used was obtained from the paper type in Indonesia [25].

$$E_{paper} = Ef_{paper}.N_{paper} \tag{4}$$

Description:
E_{paper} = Emissions from paper consumption
Ef_{paper} = Emission factor from each ream of paper
N_{paper} = The amount of paper used (reams)

4) CLASSROOM EMISSION CALCULATION

To calculate carbon emission from classroom usage, the following formula in (5) will be applied.

$$E_{classroom} = \frac{Ef_{classroom}.t}{N_{students}} \tag{5}$$

Description:
$E_{classroom}$ = Emissions from classroom usage (kg CO_2)
$Ef_{classroom}$ = Emission factor from the use of a classroom for each hour (kg CO_2/hour)
t = Total time of classroom usage (hours)
$N_{students}$ = Number of students using the classroom at a time.

A classroom or room consisted of a collection of electronic devices. The following steps can be used to determine the emission factor value.

a. Identify all electronic assets in the room.
b. Estimate the time of use of each electronic asset in the room.
c. Calculate the energy usage (kWh) of each electronic asset using the power consumption formula.
d. Normalize the energy usage of all equipment in the room by following the formula in (6).

$$Ef_{classroom} = \frac{\sum_{i=1}^{n} P_i t_i}{\sum_{i=1}^{n} t_i}.Ef_{electricity} \tag{6}$$

Description:
$Ef_{classroom}$ = Emission factor in emissions resulting from the use of a classroom or room for each hour (kg CO_2/hour)
$Ef_{electricity}$ = Emission factor in emissions generated for each kWh of electricity
P_i = Power of the i-th electronic device (Watt)
t_i = Consumption time of the i-th electronic device (hour)

B. CALCULATE CARBON EMISSIONS IN UNIVERSITY

This section discusses various related studies that are relevant to the process of calculating carbon emissions in universities from the perspective of students. These included the methodology employed and the results of comparing the emissions between universities within each study.

1) CARBON FOOTPRINT OF THE UNIVERSITY OF PATRAS, GREECE

The University of Patras conducted comprehensive calculation on carbon emissions calculations using campus facilities and individual behaviors. Data were collected through questionnaires and relevant departments from over 30,000 students and various staff members. The Campus Carbon Calculator tool was used, and the results showed a total of 32,882.5 tons of CO_2e in 2015. Transportation by students and electricity usage are the primary contributors, accounting for 60.2% of emissions. On average, each student at the University produced about 0.931 tons of CO_2e per year [4].

2) CARBON FOOTPRINT OF STUDENTS' BEHAVIOR IN TONGJI UNIVERSITY, CHINA

This study focused on estimating carbon emissions from the behavior and activities of university students. The methodology was tested at Tongji University in Shanghai, where data were collected through a structured survey. The activities were categorized into daily, academic, and transportation activities, with the corresponding carbon emission measurements. The average emissions per student were observed to be 3.84 tons CO_2e, with 65, 20, and 18% from daily activities, transportation, and academic activities. Specifically, teaching and learning activities accounted for approximately 0.691 tons per student [8].

3) CARBON FOOTPRINT OF ENVIRONMENTAL SCIENCE STUDENTS AT SUAN SUNANDHA RAJABHAT UNIVERSITY, THAILAND

This study focused on calculating carbon emissions from students enrolled in the Environmental Science program at Suan Sunandha Rajabhat University, Thailand. Three main categories, namely transportation, food consumption, and electricity usage, were examined. The survey involved 35 students and the tool employed was a carbon footprint calculation tool provided by the Thailand Greenhouse Gas Management Organization (TGO). The results showed that each student generated 2.15 tons of CO_2e per year, with 1.05, 0.70, and 0.40 tons originating from electricity consumption, food consumption, and transportation, respectively [26].

4) CARBON FOOTPRINT OF ERASMUS UNIVERSITY ROTTERDAM

This study aimed to develop a model for calculating carbon emissions specifically at Erasmus University, which could subsequently be applied to other campuses. The model considered three scopes, namely on-campus sources, purchased electricity and heat, and various other factors such as commuting, travel, water usage, paper consumption, waste, and electricity losses. To collect data on transportation behavior, surveys were conducted on a sample of 1,548 students and 1,028. The results showed that Erasmus University produced 12,601 tons of CO_2e annually, with an average of 0.521 tons of CO_2e per student. However, by excluding commuting students, the emissions from teaching and learning activities on campus were 0.098 tons of CO_2e per student. This represented 18% of the students' total CO_2e emissions [27].

5) CARBON FOOTPRINT OF UNIVERSITI TEKNOLOGI MALAYSIA

This study focused on measuring carbon emissions at Universiti Teknologi Malaysia, with a specific emphasis on energy consumption from human activities. The two main sources of emissions considered were electricity usage and transportation activities. This study covered teaching and learning activities, residential accommodations, administrative activities, and the information and communication technology sector. For the teaching and learning activities, energy consumption was analyzed for students, administrative staff, and the information and communication technology sector [28].

C. APPLIED MACHINE LEARNING IN PREDICT CARBON EMISSIONS

This section provides an overview of relevant studies that utilize machine learning for building predictive models to predict carbon emissions, discussing their background, modeling techniques, experimental results, and key insights.

1) PREDICTING GLOBAL CO_2 EMISSION USING SARIMA

This study aimed to develop a CO_2 emission predicting model that covered past, present, and future emission patterns. SARIMA was used as the time series model, and the dataset spanned from 1751 to 2018. Furthermore, 4 prediction models were created based on different periods, including pre-COVID-19, start-COVID-19, trans-COVID-19, and post-COVID-19 periods. The models were trained and tested using an 8:2 ratio, with MAPE being applied as the evaluation parameter. The post-COVID-19 model had the best accuracy by predicting a decrease in CO_2 emissions in 2022. Additionally, it highlighted the impact of the pandemic in decreasing global emissions and raising awareness of the significant reductions [19].

2) MACHINE LEARNING-BASED TIME SERIES MODELS FOR EFFECTIVE CO2 EMISSION PREDICTION IN INDIA

This study was driven by the growing CO_2 emissions in India, which have reached 1.8 tons per capita. The aim was to develop a predictive model for predicting CO_2 emissions over the next 10 years using univariate time-series data from 1980 to 2019. Several models, including 3 statistical (ARIMA, SARIMAX, Holt-Winters), 2 machine learning (linear regression and random forest), and 1 deep learning model, were experimented with using LSTM. Based on the results and evaluation using 9 regression testing metrics, the LSTM, SARIMAX, and Holt-Winters exhibited the most accurate performance. Among them, the LSTM model was more predictive with MAPE, RMSE, and MedAE values of 3.101%, 60.635, and 28.898, respectively [18].

3) INFLUENCING FACTORS OF CARBON EMISSIONS AND THEIR TRENDS IN CHINA AND INDIA: A MACHINE LEARNING METHOD

This study aimed to identify factors influencing CO_2 emissions using machine learning, with a focus on China and India, the largest fossil fuel consumers, and most populous countries. Factors such as energy consumption in industry, financial development, gross domestic product, human population, and renewable energy, were analyzed using a dataset from 1990 to 2014 [17]. A carbon emission prediction model based on the LSTM algorithm showed that energy consumption in industry and renewable energy had the greatest and least impact, respectively. By adjusting the coefficients associated with renewable energy and industrial energy consumption, this study observed reductions in carbon emissions in China and India by 2022 and 2023, respectively.

4) CARBON EMISSION PREDICTION MODEL AND ANALYSIS IN THE YELLOW RIVER BASIN BASED ON A MACHINE LEARNING METHOD

This study focused on studying the factors influencing carbon emissions in the Yellow River Basin to provide insights into the reduction in China. The approaches employed were the quadratic assignment procedure regression and the machine learning approach using LSTM. The regression approach identified regional differences in factors that affected carbon emissions, while LSTM analyzed the contributing factors based on annual data including population, GDP per capita, industrial structure, urbanization, energy intensity, and carbon emissions. The results showed that GDP per capita primarily affected emissions from 2000 to 2010, but it shifted to the human population after 2010. Finally, the LSTM model achieved a MAPE of 44.38% [20].

Table 1 presents a comparative analysis of machine learning applications for carbon emission prediction across various studies. This comprehensive overview serves as a foundational reference for the solutions developed in this research.

TABLE 1. Comparative analysis of applied machine learning for predicting carbon emissions.

Research Title	Data Used	Method	Metrics	Advantages	Limitation
Predicting Global CO_2 Emission Using SARIMA [19]	Global carbon emissions	SARIMA	MAPE	Create four types of datasets for the model experiments	Single-variable, restricted to historical carbon emission data only; relatively small dataset (<50 rows); only conducted experiments with SARIMA.
Machine Learning-Based Time Series Models for Effective CO_2 Emission Prediction in India [18].	India's carbon emissions	ARIMA, SARIMAX, Holt-Winters, Linear Regression, Random Forest, LSTM	MAPE, RMSE, MedAE	Experiment with various methods (statistical learning, machine learning, deep learning) and various metrics	Single-variable, restricted to historical carbon emission data only; relatively small dataset (<50 rows).
Influencing Factors of Carbon Emissions and Their Trends in China and India: A Machine Learning Method [17]	Energy consumption in industry, financial development, gross domestic product, human population, renewable energy, carbon emissions	LSTM	MSE, MAPE	Multi-variable, enabling the observation of relationships between variables and carbon emissions	The data used are at the national scale; only conducted experiments with LSTM.
Carbon Emission Prediction Model and Analysis In The Yellow River Basin Based on a Machine Learning Method [20]	Population, GDP per capita, industrial structure, urbanization, energy intensity, carbon emissions	LSTM	MAPE	Multi-variable, enabling the observation of relationships between variables and carbon emissions	The data used are at the national scale; exhibits a relatively high MAPE value (44.38%); only conducted experiments with LSTM.

III. PROPOSED METHOD

This study continues to build upon the proposed solution in a previous research, which developed a predictive analytics model for optimizing carbon emissions from students' activities and behaviors [29]. However, after further review, several improvements and changes were made to the proposed model. The primary distinctions between the solution in this study and the previous research are delineated as follows:

1) Inclusion of external data sources considered to exert an influence on carbon emissions stemming from student learning activities.
2) Simplification of the predictive analytics model to a singular model, specifically, the predictive model for carbon emissions that employs a supervised classification with target variables expressed as continuous data.
3) The integration of the analysis section from the prior study into the output section.

Apart from that, some parts are retained such as the data used and the final output of the developed system. The system design for predictive analytics consisted of three parts, namely input, system, and output, with each being described further below as shown in Figure 1.

On a broad aspect, students, university, and external data were used to build predictive analytics models. The following is an explanation and details of each data group used in this study.

A. STUDENTS' DATA

Students' activity was a critical factor in calculating and predicting carbon emissions. This study focused on the learning activities of students as the main subject contributing to these emissions. Data on learning behavior and activities are collected from students in Computer Science-related majors from 2019 to 2021. The recorded data about students' activities that were deemed relevant to carbon emissions included:

a) Classroom schedule: This schedule outlined lectures received by students daily, to measure emissions during a single session of teaching-learning activities.
b) Practicum schedule: These are activities students engage in, which typically involve equipment either in the laboratory or outside.
c) Student questionnaire: This was used to collect data related to students' behavior in the teaching and learning process, both inside and outside lectures. Furthermore, this questionnaire covered various aspects such as traveling to campus, the mode of transportation, participating in learning during lectures, and students' habits regarding gadgets usage during lectures.

B. UNIVERSITY DATA

Learning activities undertaken by students involve the utilization of shared classrooms and other university-owned facilities. Data related to equipment and facilities at the university were needed to help calculate carbon emissions in universities resulting from the use of shared equipment during teaching and learning activities. Carbon conversion data from electricity usage differed for each piece of equipment. Therefore, it was crucial to collect data specific to ICT equipment assets are needed. The data from the university consisted of two groups, namely:

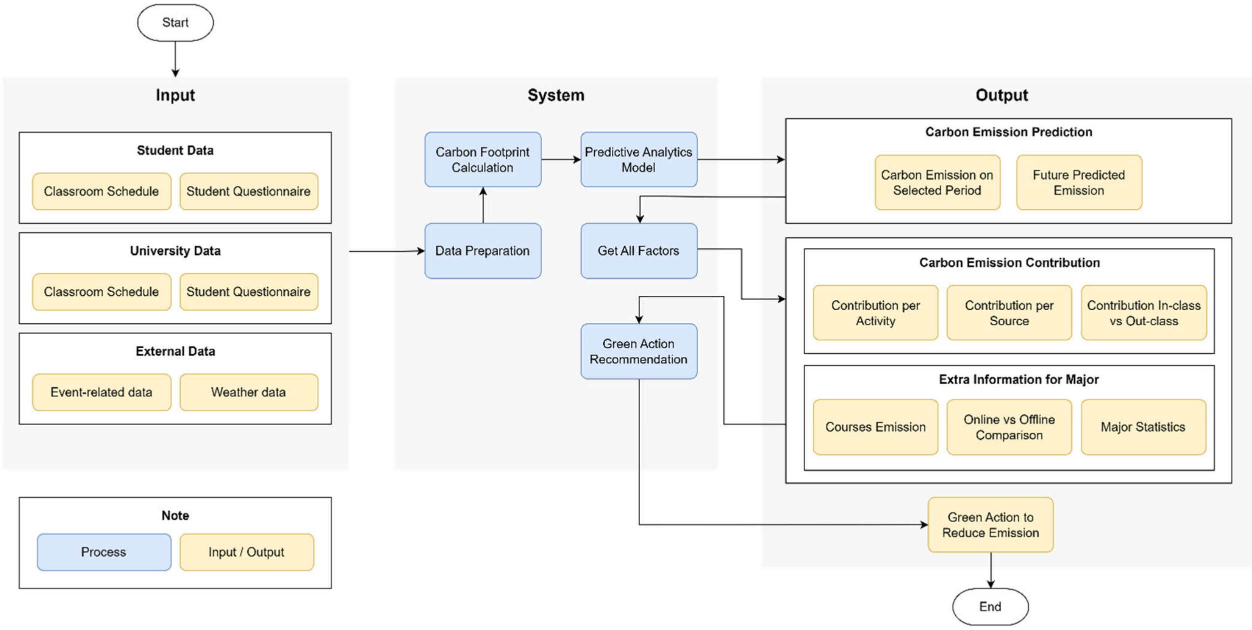


FIGURE 1. System design.

a) Room assets: This comprised a list of electronic equipment in each classroom or laboratory at ITB university. It included all the necessary facilities utilized to support lectures, practical sessions, and examination activities.

b) Electricity bills: The equipment in each room will certainly consume electrical energy to function. Consequently, the college maintained a breakdown of costs related to electricity bills, which need to be paid monthly, considering the amount of energy consumed, typically measured in (kWh). This data will serve as an initial benchmark for examining the relationship between electrical energy consumption and carbon emissions generated by students or study programs.

C. EXTERNAL DATA

External data were sourced from outside the students and university datasets as well as utilized as input for the predictive analytics model. Furthermore, it encompassed event-related information, such as whether a certain date falls within a pandemic, holiday, or exam period. Weather data such as temperature, wind speed, and humidity were also incorporated into the model. These external data sources provided valuable contextual information that helped identify patterns and correlations, aiding in the prediction of future outcomes. However, it is important to note that the variables of this data were not directly utilized in the calculation of the carbon footprint but instead served as inputs for the predictive analytics model, aiming to enhance its accuracy and effectiveness.

a) Event-related data: This type of data includes specific information such as the possibility of a certain date falling within a pandemic, holiday, or corresponding to an exam period. These indicators helped to contextualize and understand patterns in the data.

b) Weather data: This encompasses variables such as temperature, wind speed, and humidity. These measurements provided insights into the prevailing weather conditions during a given period and were valuable in predicting various outcomes or identifying correlations between weather patterns and other data points.

The obtained data was prepared before being used in the calculation and modeling stages. After collecting all the data, it was examined separately to understand its characteristics and identify interesting insights. Modeling was conducted using ERD, as shown in Figure 2, to determine the relationship and connections between datasets.

Following the understanding of the relationship, processing, and cleaning were performed. This involved ensuring consistent data format, data type, and filling in empty values in each data table. Feature engineering processes were applied to incorporate additional relevant attributes obtained from a deep understanding. Subsequently, all the data was integrated to create a complete dataset that could be utilized in the carbon emission calculation and predictive modeling processes. The flow of data integration was presented in Figure 3.

The first step involved deriving data related to the assets, which were organized into a table containing a list of emission factors for each class. This was accomplished by applying the Classroom Emission Calculation Formula. Subsequently, the data regarding the course participants were divided into 2 tables, with the first table describing the distribution of students for each class, and the second table comprising a list of unique courses. This list of courses was then combined with information obtained from the student questionnaire, including whether the exams were paper-based or electronic. The distribution of students for each class was merged with

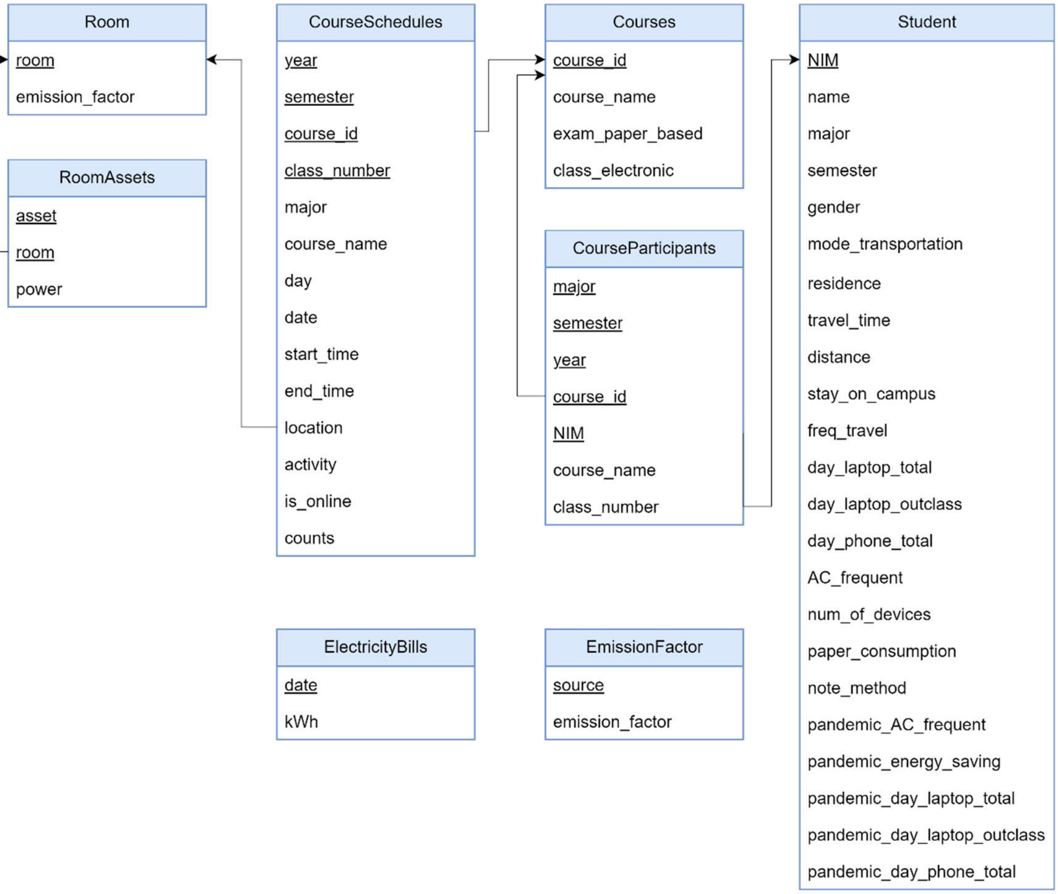


FIGURE 2. Entity relationship diagram.

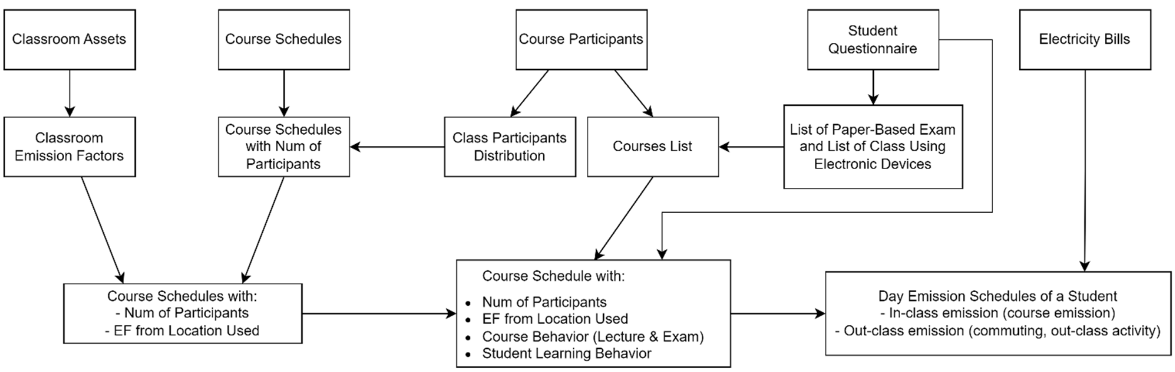


FIGURE 3. Data integration flow.

the course schedules data, which lacked information on the number of participants per class. This combined data was then integrated with classroom emission factors to calculate emissions for each course schedule. To expand the emission calculation and include activities outside of lectures but still relevant to the learning process undertaken by students, the calculation was transformed to a daily level for each lecture schedule. Additionally, information on electricity bills was included as supporting data to calculate learning emissions on a daily basis.

After all data were integrated, the carbon emission calculation process began at the students' level, as the lowest level of granularity. Carbon Footprint Calculation was limited to the students' learning activities only which was the emission from scope 2 (use of faculty facilities) and scope 3 (students' behavior outside the university) [30], [31]. The activity scope encompassed lectures, exams, practicum, coursework (outside class), and commuting activity. The carbon emission calculation process followed the flow chart presented in Figures 4, 5, and 6. To obtain the emission from

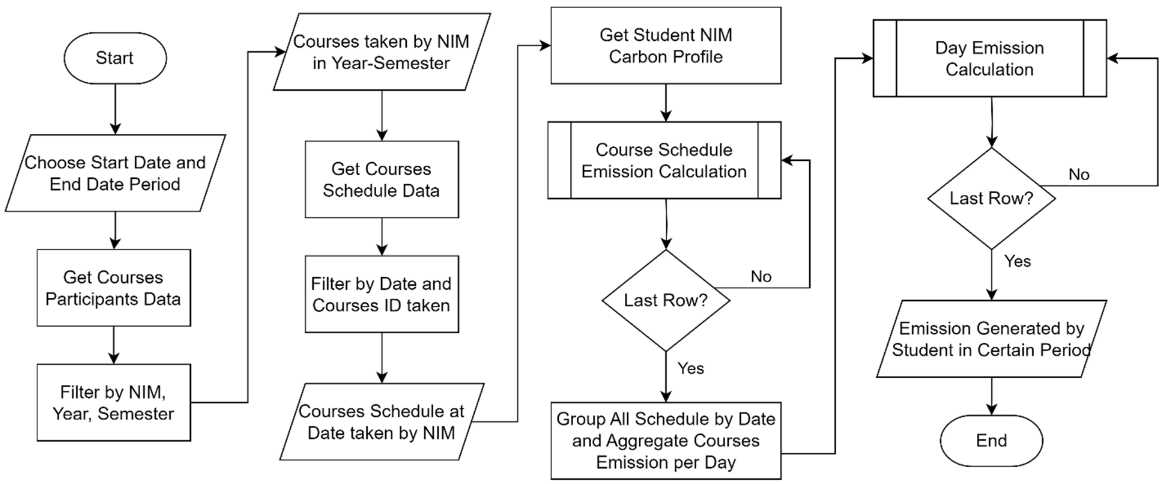


FIGURE 4. Carbon footprint calculation flow chart.

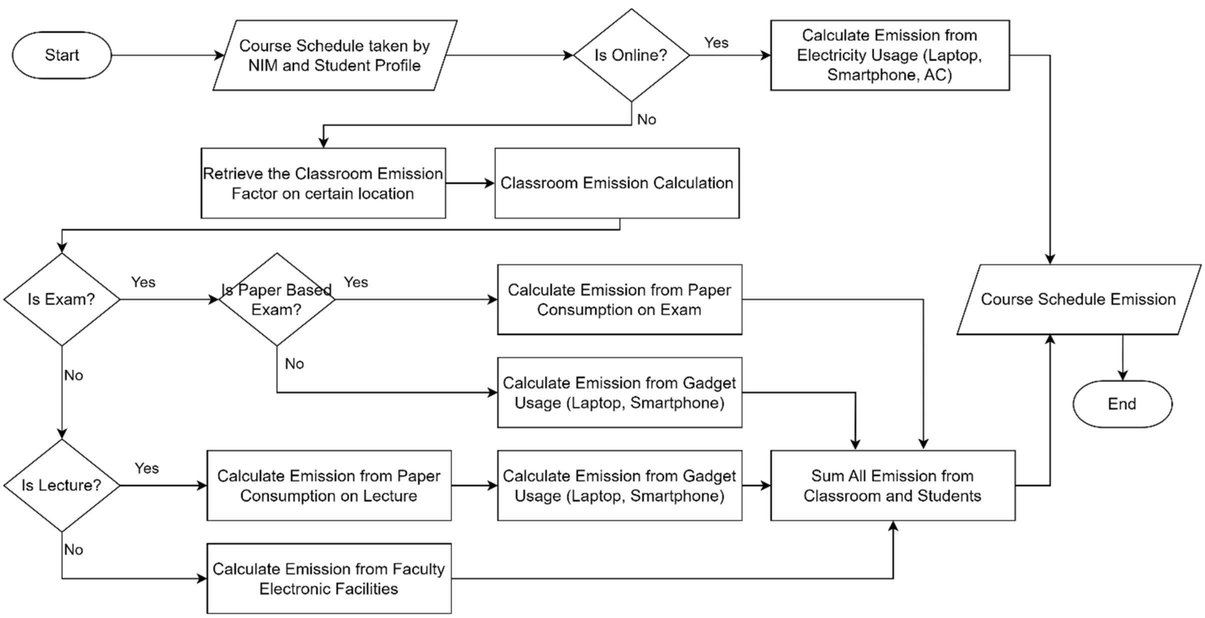


FIGURE 5. Course schedule emission calculation flow chart.

the university, the emission of all students was calculated using the same process and then summed up to derive the total value. The outcome of the calculation represented the total carbon emissions generated on a particular day, which served as the target variable of the final dataset.

After obtaining total carbon emission data from all dates, a new dataset was created, which involved external data as a dataset in the construction of predictive models. A model is stated to be predictive when it applies historical and external data. Therefore, the previous carbon emission dataset was combined with external data per day to obtain the final predictive dataset with the following attributes in Table 2.

The selected task for predictive analytics was a regression, a suitable algorithm for predicting continuous values from data. In this case, the goal was to predict the value of carbon emissions generated by a certain level of granularity over a specific period. The following are the Machine Learning algorithms used in the modeling:

1) Long Short Time Memory (LSTM)
2) XGBoost Regressor
3) Linear Regression
4) Support Vector Regression (SVR)
5) Decision Tree Regressor

The predictive model used the dataset from August 2021 to May 2023. This dataset was divided into training and test data, using the train test split technique with division ratios of 70% and 30%, respectively. With this scheme, the training dataset contained carbon emission data and its supporting attributes from August 2021 to December 2022. Meanwhile, the testing dataset consisted of data from January 2023 to May 2023. In general, three models were developed based on the entity, and they included predictive models

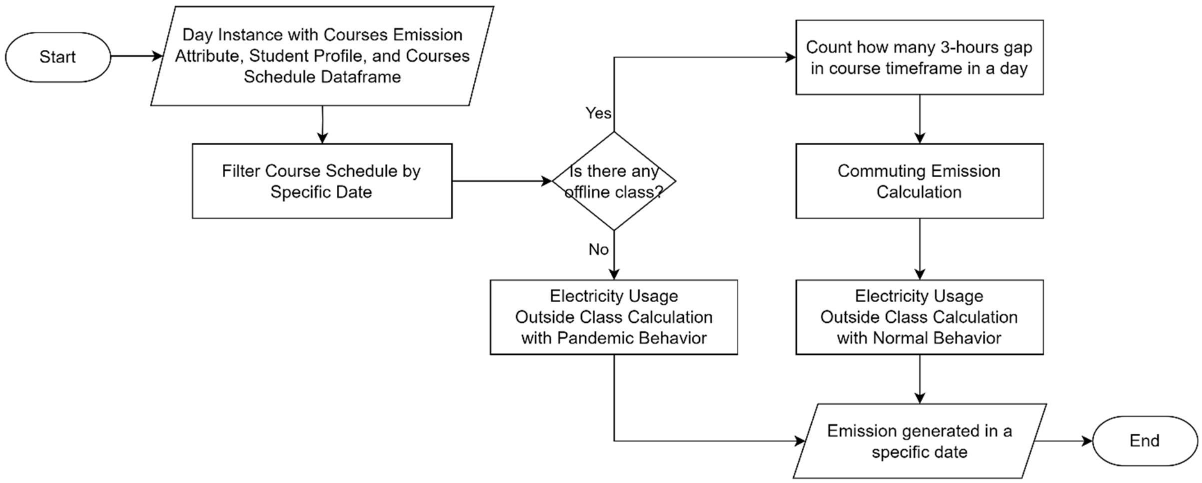


FIGURE 6. Day emission calculation flow chart.

for students, majors, and ITB. The model will predict future carbon emissions based on these attributes. The two outputs from the prediction were carbon footprint calculation results and future predicted emissions. The system will then utilize these outputs to provide comprehensive details regarding the factors influencing carbon emissions.

The factor details were broken down into carbon emission contribution and extra information for major. Carbon emission contribution provided details regarding the contribution of each factor, which were categorized into three groups, namely contributions by activity, contribution by emission source, and a comparison between contributions inside and outside lectures. Extra information for major contained relevant insights for the study program. It aimed to provide a more detailed understanding of carbon emissions from each course, including the comparison between online and offline periods, as well as paper-based and electronic-based exam methods. All these factors were considered in determining green actions that can be recommended to the study program and students. These actions were sorted based on their impact on emission reduction, starting from the most impactful measures.

IV. RESULTS AND DISCUSSION

This section will discuss all findings of the study and provides an in-depth analysis and interpretation of the results.

A. ANALYZE THE PERFORMANCE OF THE PREDICTIVE MODEL

The carbon footprint prediction model was evaluated using four testing metrics, namely the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R2 score, and Mean Average Percentage Error (MAPE) metrics. MSE and RMSE help gauge predictive accuracy and error magnitude, with MSE emphasizing substantial deviations, while RMSE offers an interpretable measurement in the target variable's units. R2 assesses the model's capacity to explain variance in the dependent variable, representing its explanatory power. MAPE, which expresses accuracy in percentages, is particularly useful for models dealing with data of varying scales. Employing this combination of metrics provides a comprehensive assessment of the model's capabilities and aids in informed decision-making for its application within a specific context. The evaluation results are presented in Table 3.

Table 3 shows that the LSTM model performed the worst. Additionally, it exhibited the highest error values across all test metrics, with MAE and RMSE reaching 210.47 and 296.65, respectively. The MAPE value of the LSTM model reached 66.76%, indicating a relatively large error. Moreover, the R2-score value stood at 0.37, indicating a low correlation between the independent variables influencing the target variable.

Models other than LSTM showed similar performance, with relatively minimal differences in error values. Figure 7 provides a comprehensive comparison of these models based on their algorithms. These models were quite effective in predicting carbon emissions at the university level. The order of evaluation indicated that SVR was the best model, followed by Linear Regression, XGBoost Regressor, Decision Tree Regressor, and LSTM. The MAE, RMSE, R2, and MAPE values of this model were 129.41, 158.03, 0.98, and 15.82 respectively. These parameters have errors near zero, indicating that the SVR model has the best performance in predicting carbon emissions. Based on these results, the regression model with the Support Vector Machine algorithm was selected as the best model.

The regression model built was classified as predictive analytics because it involves historical and external data in the development process. The influence of each factor or attribute in the dataset on the prediction of carbon emissions was explained. The influence of each factor was in the form of a score that stated how useful and significant a feature was in predicting the target variable. Given that the predictive model built was a regression model, the representation of the influence of each feature was observed from their coefficients. The Support Vector Regression model already had a coefficient

TABLE 2. Predictive dataset attributes.

Data Group	Attribute	Data Type	Definition
Historical data	total_emission _1	Float	Total carbon emission data for the previous 1 day
	total_emission _2	Float	Total carbon emission data on the previous 2 days
	…	…	...
	total_emission _30	Float	Total carbon emission data for the previous 30 days
University and Students Data	total_laptop_u sages	Float	Total laptop usage in one day of all students
	total_ distances	Float	Total distance traveled by all students from residence to campus
	total_classes	Integer	Number of classes held
	online_classes	Integer	Number of offline classes held
	offline_classes	Integer	Number of offline classes held
	students_atten ds	Integer	Number of students who attended the class
	online_attends	Integer	Number of students attending online classes
	itb_electricity _usage	Float	Total electricity usage used in one day
Event data	is_exam_perio d	Boolean	States if the day is an exam period
	is_pandemic	Boolean	State if the day is still in the pandemic period
	is_holiday	Boolean	States if the day is a public holiday
Weather data	temperature	Float	Average temperature from 7 am to 5 pm in one day
	wind	Float	Wind speed during the day
	humidity	Float	Water humidity during the day

TABLE 3. Predictive model evaluation results.

MODEL	MAE	RMSE	R2	MAPE
LSTM	540.16	791.65	0.51	211.25
XGBoost Regressor	141.45	188.49	0.97	7.45
Linear Regression	149.61	179.29	0.97	15.37
Decision Tree Regressor	213.48	301.55	0.93	11.61
Support Vector Regression	129.41	158.03	0.98	15.82

value for each feature and an intercept-like linear regression. Table 4 presents the coefficient value for each feature.

Based on the coefficient of each feature on the predictive dataset, 5 features including total_laptop_usage, total_distances, online_classes, students_attends, and online_attends had a significant and large value. The total_laptop_usage and distance had coefficient values of 0.42896 and 0.37561, respectively. This indicated that the total laptop usage by students and the transportation distance from residence to campus had a major influence on predicting carbon emissions generated in a day. Similarly, the variable "online_classes," which indicated the number of classes held online, had a significant impact. The negative value indicated that the more classes held online, the fewer carbon emissions were generated. In addition to the number of classes, the number of students attending lectures simultaneously also played a crucial role, showing a direct proportionality. However, this factor also considered students attending online lectures, which explained the negative coefficient value. Consequently, an increase in the number of students opting for online lectures resulted in a reduction in carbon emissions.

TABLE 4. Coefficient of each attributes.

Attribute	Coefficient	Attribute	Coefficient
total_laptop_usages	0.42896	total_emission_lag9	0.01112
total_distances	0.37561	total_emission_lag10	0.00841
total_classes	0.01911	total_emission_lag11	-0.0011
online_classes	-0.0725	total_emission_lag12	-0.02155
offline_classes	0.00751	total_emission_lag13	-0.00476
students_attends	0.11044	total_emission_lag14	0.02834
online_attends	-0.11724	total_emission_lag15	0.00363
kWh_day	-0.0032	total_emission_lag16	0.00854
is_exam_period	-0.01864	total_emission_lag17	-0.00454
is_pandemic	-0.05109	total_emission_lag18	0.00077
is_holiday	-0.00467	total_emission_lag19	-0.02
temperature	-0.01694	total_emission_lag20	-0.00975
wind	-0.00064	total_emission_lag21	0.00495
humidity	-0.01175	total_emission_lag22	-0.01656
total_emission_lag1	0.01695	total_emission_lag23	0.00778
total_emission_lag2	0.00134	total_emission_lag24	0.01803
total_emission_lag3	0.00304	total_emission_lag25	-0.00578
total_emission_lag4	0.01298	total_emission_lag26	-0.01197
total_emission_lag5	0.00123	total_emission_lag27	-0.00735
total_emission_lag6	-0.00262	total_emission_lag28	0.0253
total_emission_lag7	0.01267	total_emission_lag29	-0.00409
total_emission_lag8	0.00257	total_emission_lag30	0.00982

In terms of time, lectures held during exams, the pandemic, and vacation periods yielded relatively lower carbon emission values compared to regular offline lecture periods in normal conditions. This was indicated by the negative coefficient associated with each of these time-related attributes. Weather data, including factors such as temperature, wind speed, and humidity, also exerted influence on the predicted emissions. This influence could be observed in real-world scenarios where certain lectures were conducted online due to rainy days characterized by high water humidity values close to 100% and lower temperatures than usual. Based on the attributes of historical data, the model utilized past carbon emission data to make predictions for a given day. This was supported by the autocorrelation plot, which revealed a repeating pattern in the data every 30 days. A seasonal-like pattern was observed in the carbon emission behavior of students. This historical data served as the foundation for the

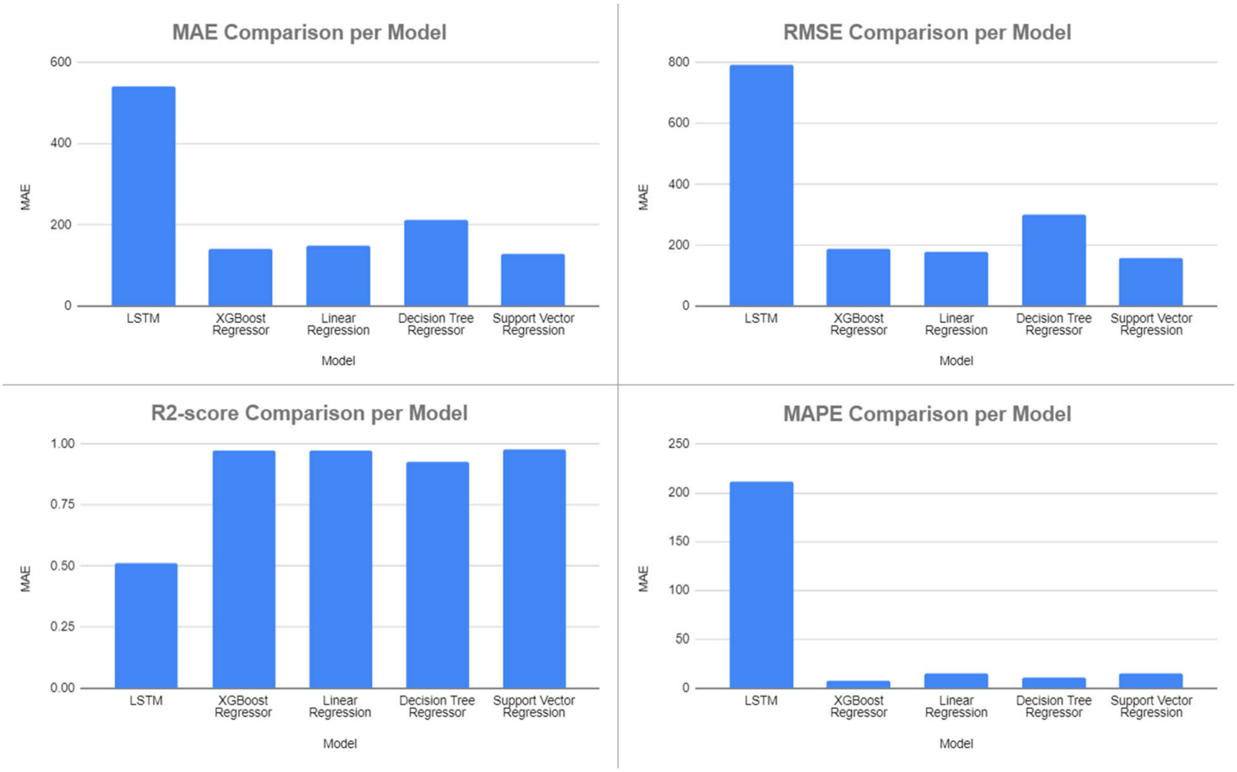


FIGURE 7. Predictive model evaluation results.

model to analyze trends in carbon emission values over time and make future predictions. However, when incorporating external data, future predictions may deviate, especially in cases where the data from the last 30 days were relatively similar but certain events had a significant impact on emissions. The following examples illustrate cases where external data affected the emission values:

1) The value of carbon emissions during the pandemic was certainly different from the hybrid and normal periods. This can be seen from the absence of transportation activities undertaken by students to campus, resulting in a reduction in carbon emissions related to motor vehicle use.
2) The value of carbon emissions that occurred during the exam was always lower than the normal lecture period. This was primarily attributed to the fact that each course holds an exam only once within a specific period, while lecture activities took place multiple times based on the credit hours (SKS).
3) Weather data such as temperature, wind speed, and humidity affected the lecture schedule on a certain day. These changes in the lecture schedule are often predicted based on the weather conditions observed in the previous week, which follow a periodic pattern. Several courses changed to an online system or cancel lectures due to extreme weather.

Based on the aforementioned external attributes, the model predicted future carbon emission values by taking into account these conditions, despite having historical data that exhibited seasonality. These attributes indicated the presence of external factors that influenced the fluctuations in carbon emissions resulting from the learning activities attended by students.

B. USE OF PREDICTIVE MODEL IN APPLICATION

The best predictive model was integrated into an application called Student Carbon Footprint Tools, which presented a range of information and insights regarding carbon emissions generated from students' learning activities. This application allowed users to choose a specific date range for calculation, and it generated predictions for the number of days within the selected period. Figure 8 showed the interface of the application and the results of carbon emission predictions. The application also provided comparative results of carbon emissions between each major that can be seen in Figure 9.

Based on the results of carbon emission calculations, ITB university produced 612.8 tons of CO_2e during the 2022/2023 academic year. This emission was attributed to 1,071 students enrolled in the Computer Science related majors, namely Undergraduate of Informatics (IF), Undergraduate of Information System and Technology (STI), and Master of Informatics (MIF). Therefore, the average value produced by students in a year and daily were 571.19 kg CO_2e and 3.08 kg CO_2e, respectively. ITB students have an average laptop usage and travel distance from their residence of 8.9 hours per day and 3.06 km.

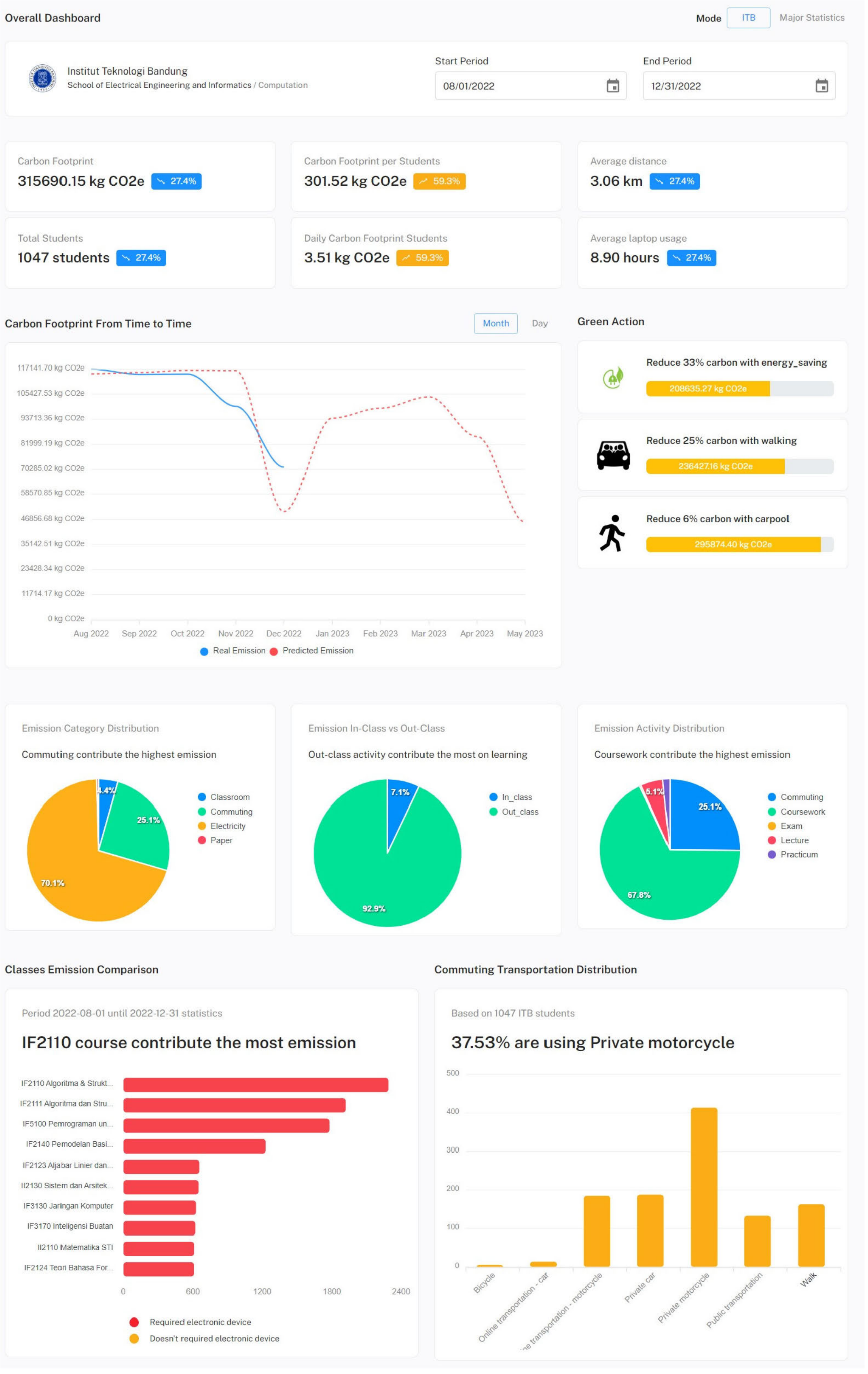


FIGURE 8. Student carbon footprint tools user interface.

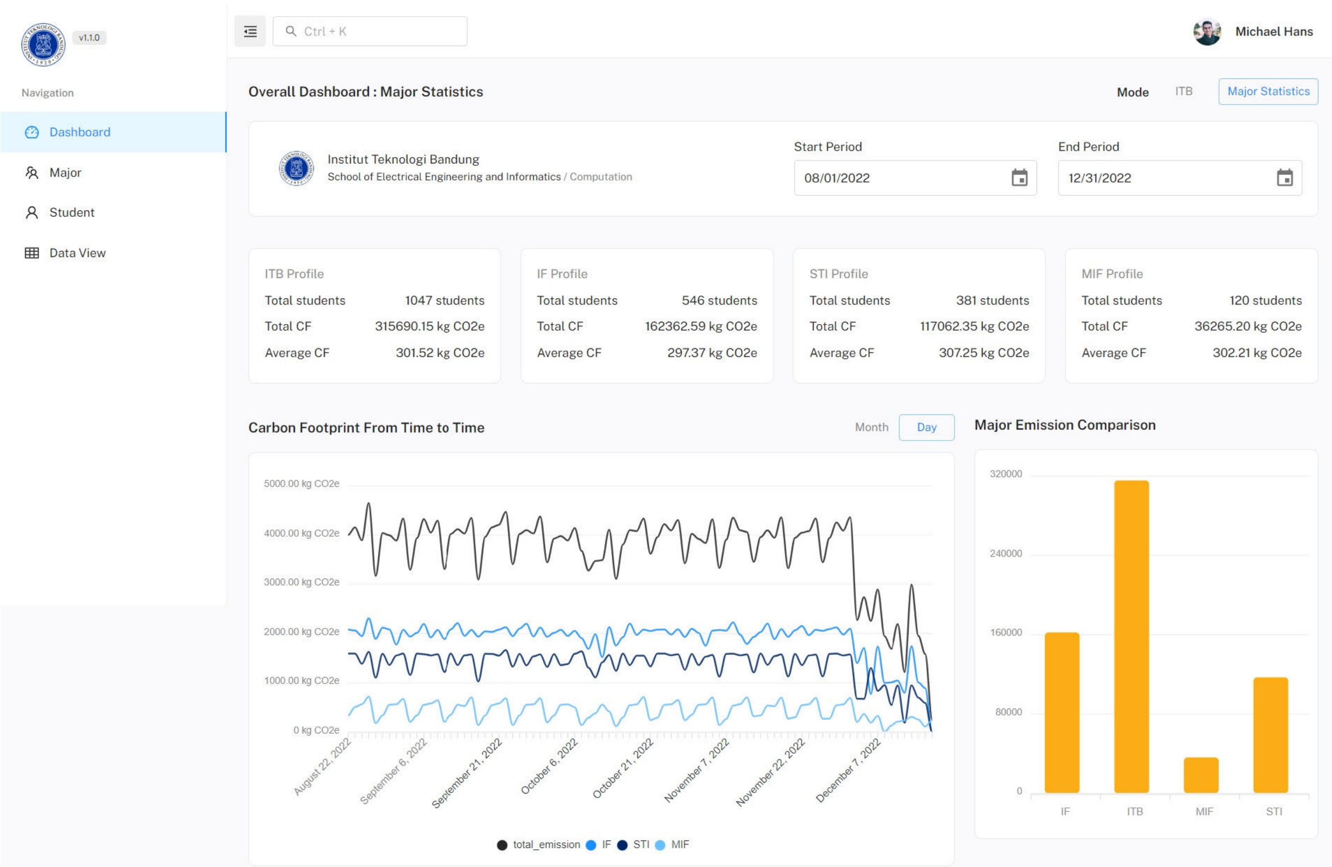


FIGURE 9. Major statistics page.

According to Figure 9, the IF major produced the largest carbon emissions of 323.6 tons CO_2e. This was influenced by the number of students enrolled in this program. There is a positive correlation between the amount of emissions produced and the total number of students enrolled. Furthermore, the number of credits affected the intensity of lectures attended. From Figure 9, it can be observed that IF and STI majors have average carbon emissions per student that are relatively similar to each other, namely 591.6 kg CO_2e and 589.98 kg CO_2e, while that of the MIF major was at 469.02 kg CO_2e. The number of credits usually taken by IF and STI students is in the range of 18 to 24, while for MIF, it was between 12 to 15. The number of credits can express the number of lectures attended such that the value of MIF carbon emissions was relatively smaller than undergraduate IF and STI students. This is because the number of credits that MIF students can take was also less than the undergraduate. However, it is worth noting that MIF students have a relatively large average laptop usage time. One of the influencing aspects was that they spend more time outside of lectures and allocate a significant portion of their time to self-study activities.

Factors affecting carbon emissions can be grouped into three parts, namely emission sources, scope of learning, and activities. The details and percentage of each factor are shown in Table 5 and Figure 8. Based on the distribution results, it was observed that electricity consumption was the largest contributor to carbon emissions at 70%, followed by commuting at 25.3%. The use of paper does not seem to have a significant impact on computer science majors. When looking at the scope of learning, most emissions come from out-of-class activities, accounting for 93% of the total. These include independent activities and large assignments outside the classroom. According to Figure 8, top five courses with the biggest emission contributors were those that involved practicum, such as Algorithm & Data Structure and Object-Oriented Programming.

Insights from the calculation and prediction of carbon emissions in 2022/2023, were compared with the 2021/2022 academic year when lectures were mostly held online. In 2021/2022, ITB university produced 408.3 tons kg CO_2e, 204.8 tons less than in 2022/2023. However, when considering the average carbon emissions per student, the value was 567.86 kg CO_{2e}, which was relatively lower but not significantly different from those produced during the offline lecture period. Several factors contribute to this similarity in average carbon emissions between the pandemic and normal times.

During the online lecture period, all lectures are conducted from students' respective homes, resulting in higher

TABLE 5. Carbon emission distribution per factors.

Group	Factor	Portion (%)
Emission Sources	Electricity	70.0%
	Commuting	25.3%
	Classroom	4.3%
	Paper	0.4%
Scope of Learning	In-class	93.0%
	Out-class	7.0%
Activities	Commuting	25.3%
	Coursework	67.7%
	Exam	0.3%
	Lecture	5.2%
	Practicum	1.5%

electricity usage. Electricity contributions during the online and offline lecture periods were 93.5% and 60%, respectively. Lectures held at home caused the emissions generated by students to be significantly different from one another. This was because there was no shared use of facilities or electronic devices, such as using the classroom together. Despite the smaller contribution of electricity during the offline period, there was an additional emission source in the form of commuting, which contributed 25% of the overall value.

To validate the accuracy of the prediction results, a comparison is conducted between the predicted and calculated carbon emissions for the 2023 period, using it as test data. The comparison graph of the prediction and calculation results of carbon emissions in 2023 is shown in Figure 10. The calculation outcome obtained at a certain period was used as historical data to predict future carbon emissions with a predictive model developed based on machine learning. Furthermore, the model was developed using a dataset of emissions from August 2021 to December 2022 and will be used to predict the period of January to May 2023. The results obtained are carbon emissions generated in the second semester of the 2022/2023 academic year, and it decreased when compared to the first semester of 2022/2023. The predictions were then compared with real data in the form of calculations with data available in 2023.

When solely considering the prediction results, carbon emissions generated in the 2022/2023 academic year were relatively higher than in 2021/2022, for both odd and even semesters. This was evidenced by the emissions in the even semester of the 2022/2023 academic year, which were predicted to decrease but remain higher than in the even semester of 2021/2022. From the comparison, it was observed that the pandemic factor was quite influential in the emissions produced by students. Emissions during the pandemic were smaller than in the non-pandemic period. However, the real carbon emission data in the even semester of the 2022/2023 academic year was relatively smaller than that of the previous year. This indicated that students produced smaller carbon emissions evidenced by a more significant decrease compared to the predicted results.

C. OPTIMIZE THE CARBON EMISSIONS

The calculation and prediction of carbon emissions through the application offered valuable insights for stakeholders to make informed decisions regarding the reduction. The application and predictive model built were used as consideration for decision-making both at students and the study program stakeholder levels. The Student Carbon Footprint Tools provided a set of green action recommendations that can be performed to reduce carbon emissions. Figure 8 presents the recommended green action. The recommendations were sorted based on the actions with the most reduction rate. Green actions were simple activities that can be performed at the individual level. Based on the calculation and prediction, ITB promoted students to implement this measure in the form of energy-saving utilization, as it could reduce up to 33% of the original emissions. Furthermore, the decrease can extend to 25 and 6% when all students walk to campus and engage in vehicle sharing per 4 persons, respectively.

From a students' perspective, the results and information presented by this application provided significant benefits in reducing carbon emissions. By utilizing the results of the calculations over specific periods, students can gain a clear understanding of their carbon footprint and develop awareness regarding its environmental impact. Furthermore, the inclusion of variables such as laptop usage, travel distance, and other activities, are to be optimized were identified to reduce carbon emissions. For instance, there should be a focus on reducing energy consumption when using a laptop or choosing environmentally friendly transportation alternatives. Carbon emission graphs presented with historical data and future predictions also provided a comprehensive view of the impact of actions taken. The green action information presented based on significant reduction rates will provide practical guidance for students to adopt effective measures in their efforts to decrease carbon emissions. With concrete numbers and quantified prediction results, students can become more aware of the emissions they produced and be greatly motivated to take simple steps to ensure an effective decrease.

From the perspective of study program stakeholders, the information and results provided by this application held immense value in guiding efforts to mitigate carbon emissions among students within the study program. By accessing the calculated carbon emission data of students over specific periods, as well as comprehensive carbon emission profiles encompassing various contributing factors, the head of the study program could discern patterns and identify areas characterized by high levels of carbon emissions. This insight empowered study program stakeholders to focus their endeavours and initiatives on reducing carbon footprints. For example, awareness could have been raised about the importance of energy efficiency, promotion of walking among students, and advocacy for the use of eco-friendly technologies. Additionally, the availability of carbon emission graphs, which incorporated historical data and future projections,

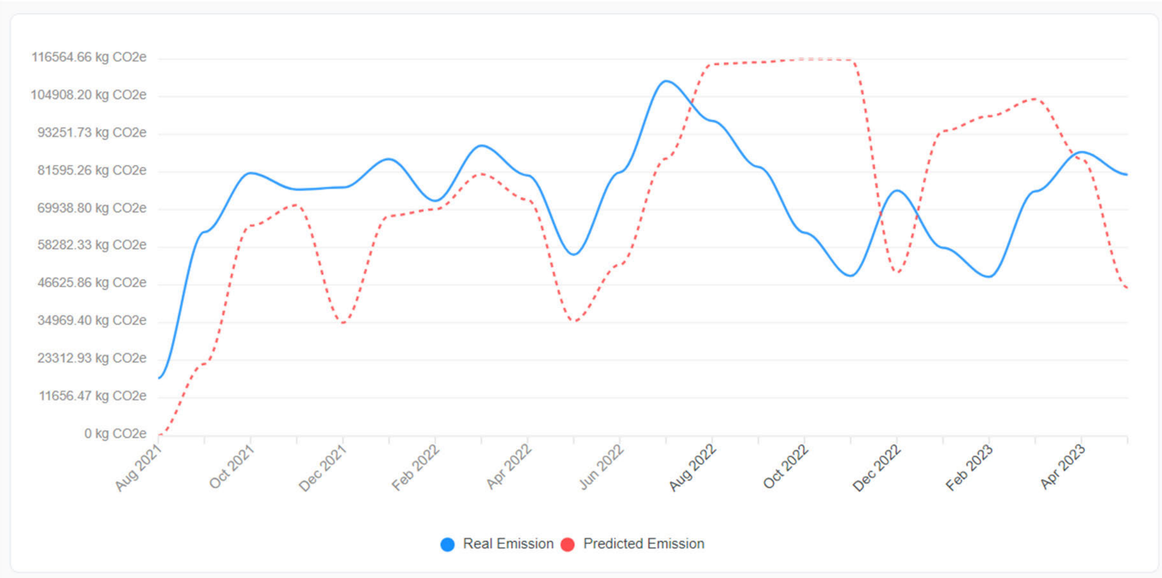


FIGURE 10. Carbon footprint prediction comparison with real values.

would have provided a clear understanding of emission trends and facilitated the formulation of long-term strategies for carbon emission reduction. Furthermore, the data regarding the distribution of carbon emissions based on emission sources, areas of study, and activities undertaken would have assisted the head of the study program in designing more targeted and efficient emission reduction programs tailored to the unique requirements of the study program.

The Student Carbon Footprint Tools application can be developed into a comprehensive information system related to students' carbon emissions. The Information System of this tool provided services for ITB management to determine the pattern of emissions generated or students' behavior in the context of learning. The Student Carbon Footprint Tools system was built as a means for ITB to understand the quantification of emissions generated through teaching and learning activities, to simultaneously provide awareness. With green actions sorted by impact, the first step to reduce the emission rate was talking, starting from the most impactful.

V. CONCLUSION

In conclusion, the predictive analytics model was developed in two stages, namely CF calculation and prediction. These stages required data from various sources, including questionnaire data, lecture schedules, and room assets. The integration of the best predictive model into an application allowed users to measure and predict their carbon emissions within a specific timeframe. Implementing the predictive analytics model quantified ITB's carbon emissions, revealing an average carbon emission per student in a year of 572.19 kg CO_2e. This revelation enabled students and stakeholders to gain awareness of their carbon emissions by providing emission profiles, tracking trends over time, displaying distribution, and recommending prioritized green actions for a significant reduction.

For future studies, several potential developments should be considered. Firstly, it would be beneficial to expand the scope of carbon emissions studied to include students' behavior and activities outside of the campus. This expansion would involve incorporating emissions from students' travel behavior, food consumption, waste generation, and clothing patterns. Additionally, developing applications to assist users in measuring the progress of recommended green actions would be valuable. The objective is to empower students not only with knowledge about emissions and available initiatives but also with the ability to actively implement and monitor the green actions they undertake. This approach would allow students to evaluate their current behavior and measure its sustainability in the future, along with predictive results if they continue to maintain environmentally friendly practices. Furthermore, there are opportunities to enhance predictive models by incorporating longer historical data spanning several decades, as shown in previous studies. Expanding the range of external data involved in the development of predictive models could improve their accuracy and effectiveness. Some experiments may also be needed to measure the scalability of the developed model if the data structure and volume of data received by the system include carbon emission data of all students with a wider scope of learning activities.

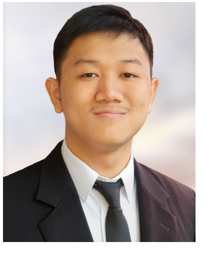
ACKNOWLEDGMENT

The authors would like to thank to Institut Teknologi Bandung for supporting this research.

REFERENCES

[1] *Kyoto Protocol*, United Nations, Kyoto, Japan, 1997.

[2] EPA. (2022). *Global Greenhouse Gas Emissions Data*. Accessed: Dec. 5, 2022. [Online]. Available: https://www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data
[3] L. P. Güereca, N. Torres, and A. Noyola, ''Carbon footprint as a basis for a cleaner research institute in Mexico,'' *J. Cleaner Prod.*, vol. 47, pp. 396–403, May 2013, doi: 10.1016/j.jclepro.2013.01.030.
[4] A. Vrachni, A. Christogerou, G. A. Thomopoulos, C. Marazioti, and G. N. Angelopoulos, ''Carbon footprint of the university of patras in greece: Evaluating environmental culture and campus' energy management towards 2030,'' *Pollutants*, vol. 2, no. 3, pp. 347–362, Aug. 2022, doi: 10.3390/pollutants2030024.
[5] H. N. Larsen, J. Pettersen, C. Solli, and E. G. Hertwich, ''Investigating the carbon footprint of a university—The case of NTNU,'' *J. Cleaner Prod.*, vol. 48, pp. 39–47, Jun. 2013, doi: 10.1016/j.jclepro.2011.10.007.
[6] K. Steemers and G. Y. Yun, ''Household energy consumption: A study of the role of occupants,'' *Building Res. Inf.*, vol. 37, nos. 5–6, pp. 625–637, Nov. 2009, doi: 10.1080/09613210903186661.
[7] J. Langevin, P. L. Gurian, and J. Wen, ''Reducing energy consumption in low income public housing: Interviewing residents about energy behaviors,'' *Appl. Energy*, vol. 102, pp. 1358–1370, Feb. 2013, doi: 10.1016/j.apenergy.2012.07.003.
[8] X. Li, H. Tan, and A. Rackes, ''Carbon footprint analysis of Student behavior for a sustainable university campus in China,'' *J. Cleaner Prod.*, vol. 106, pp. 97–108, Nov. 2015, doi: 10.1016/j.jclepro.2014.11.084.
[9] Institut Teknologi Bandung. (2022). *Tentang ITB*. Accessed: Dec. 5, 2022. [Online]. Available: https://www.itb.ac.id/tentang-itb
[10] IBM. (2022). *What is Predictive Analytics*. Accessed: Dec. 5, 2022. [Online]. Available: https://www.ibm.com/topics/predictive-analytics
[11] D. Abbott, *Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst*. Indianapolis, IN, USA: Wiley, 2014.
[12] D. Huang, D. Mu, L. Yang, and X. Cai, ''CoDetect: Financial fraud detection with anomaly feature detection,'' *IEEE Access*, vol. 6, pp. 19161–19174, 2018, doi: 10.1109/ACCESS.2018.2816564.
[13] P. Khan, M. F. Kader, S. M. R. Islam, A. B. Rahman, M. S. Kamal, M. U. Toha, and K.-S. Kwak, ''Machine learning and deep learning approaches for brain disease diagnosis: Principles and recent advances,'' *IEEE Access*, vol. 9, pp. 37622–37655, 2021, doi: 10.1109/ACCESS.2021.3062484.
[14] M. E. Hoque, A. Thavaneswaran, S. S. Appadoo, R. K. Thulasiram, and B. Banitalebi, ''A novel dynamic demand forecasting model for resilient supply chains using machine learning,'' in *Proc. IEEE 45th Annu. Comput., Softw., Appl. Conf. (COMPSAC)*, Madrid, Spain, Jul. 2021, pp. 218–227, doi: 10.1109/COMPSAC51774.2021.00040.
[15] K. Pahwa and N. Agarwal, ''Stock market analysis using supervised machine learning,'' in *Proc. Int. Conf. Mach. Learn., Big Data, Cloud Parallel Comput. (COMITCon)*, Faridabad, India, Feb. 2019, pp. 197–200, doi: 10.1109/COMITCon.2019.8862225.
[16] J. B. Dunn and P. Balaprakash, *Data Science Applied to Sustainability Analysis*. Amsterdam, The Netherlands: Elsevier, 2021, doi: 10.1016/C2018-0-02415-9.
[17] M. Ahmed, C. Shuai, and M. Ahmed, ''Influencing factors of carbon emissions and their trends in China and India: A machine learning method,'' *Environ. Sci. Pollut. Res.*, vol. 29, no. 32, pp. 48424–48437, Jul. 2022, doi: 10.1007/s11356-022-18711-3.
[18] S. Kumari and S. K. Singh, ''Machine learning-based time series models for effective CO_2 emission prediction in India,'' *Environ. Sci. Pollut. Res.*, Jul. 2022, doi: 10.1007/s11356-022-21723-8.
[19] Y. Meng and H. Noman, ''Predicting CO_2 emission footprint using AI through machine learning,'' *Atmosphere*, vol. 13, no. 11, p. 1871, Nov. 2022, doi: 10.3390/atmos13111871.
[20] J. Zhao, L. Kou, H. Wang, X. He, Z. Xiong, C. Liu, and H. Cui, ''Carbon emission prediction model and analysis in the yellow river basin based on a machine learning method,'' *Sustainability*, vol. 14, no. 10, p. 6153, May 2022, doi: 10.3390/su14106153.
[21] N. Zheng, S. Li, Y. Wang, Y. Huang, P. Bartoccid, F. Fantozzid, J. Huang, L. Xing, H. Yang, H. Chen, Q. Yang, and J. Li, ''Research on low-carbon campus based on ecological footprint evaluation and machine learning: A case study in China,'' *J. Cleaner Prod.*, vol. 323, Nov. 2021, Art. no. 129181, doi: 10.1016/j.jclepro.2021.129181.
[22] K. Valls-Val and M. D. Bovea, ''Carbon footprint in higher education institutions: A literature review and prospects for future research,'' *Clean Technol. Environ. Policy*, vol. 23, no. 9, pp. 2523–2542, Nov. 2021, doi: 10.1007/s10098-021-02180-2.
[23] *Peraturan Kementerian Lingkungan Hidup, No. 12 Tahun 2010*, Kementerian Lingkungan Hidup, Kementerian Lingkungan Hidup, Jakarta, Indonesia, 2010.
[24] *Climate Transparency Report Comparing G20 Climate Action and Responses to the COVID-19 Crisis*, Climate Transparency, Jakarta, Indonesia, 2020.
[25] D. Negoro, D. R. Jati, and K. P. Utomo, ''Penentuan Nilai Jejak Karbon (CO_2, CH_4, N_2O) dari Aktivitas Kampus Fakultas Teknik Universitas Tanjungpura,'' *Jurnal Rekayasa Lingkungan Tropis*, vol. 2, no. 2, 2021. [Online]. Available: https://jurnal.untan.ac.id/index.php/jurlis/article/view/48257
[26] T. Utaraskul, ''Carbon footprint of environmental science students in suan sunandha rajabhat university, Thailand,'' *Proc.-Social Behav. Sci.*, vol. 197, pp. 1156–1160, Jul. 2015, doi: 10.1016/j.sbspro.2015.07.371.
[27] S. Sprangers, ''Calculating the Carbon Footprint of Universities,'' Erasmus Univ. Rotterdam, Rotterdam, The Netherlands, Tech. Rep., 2011. [Online]. Available: https://thesis.eur.nl/pub/10488/10488-Sprangers.pdf
[28] I. A. Abdul-Azeez and C. S. Ho, ''Realizing low carbon emission in the university campus towards energy sustainability,'' *Open J. Energy Efficiency*, vol. 04, no. 02, pp. 15–27, 2015, doi: 10.4236/ojee.2015.42002.
[29] M. Hans, E. Hikmawati, and K. Surendro, ''Predictive analytics model for optimizing carbon emission from Student activity and behavior with a design thinking approach: Case study: Bandung institute of technology,'' in *Proc. 12th Int. Conf. Softw. Comput. Appl.*, Kuantan, Malaysia, Feb. 2023, pp. 77–82, doi: 10.1145/3587828.3587841.
[30] J. Ranganathan, L. Corbier, P. Bhatia, S. Schmitz, P. Gage, and K. Oren, ''GHG protocol standard,'' WRI/WBCSD, Geneva, Switzerland, Tech. Rep., 2004.
[31] *Stanford University CY19 Scope 3 Emissions Inventory Public*, Stanford Univ., Stanford, CA, USA, 2022.



MICHAEL HANS received the B.S. degree in computer science from Institut Teknologi Bandung, in 2022, where he is currently pursuing the master's degree in information systems with the thesis topic related with Green IT. His research interests include software engineering, information systems, data science, and machine learning.

ERNA HIKMAWATI received the B.S. degree in computer science from Pasim National University, in 2013, the master's degree in information systems from STMIK LIKMI, in 2017, and the Ph.D. degree in electrical engineering and informatics from Institut Teknologi Bandung, in 2023. She is currently a Lecturer in information systems with Institut Teknologi Bandung, Pasim National University. Her research interests include data science, machine learning, and information systems.

KRIDANTO SURENDRO received the B.Eng. and master's degrees in industrial engineering from Institut Teknologi Bandung (ITB), Indonesia, in 1987 and 1991, respectively, and the Ph.D. degree in computer science from Keio University, Japan, in 1999. He is currently an Associate Professor in computer science with ITB. His research interests include data science, soft computing, and IT governance.