MATEMATIKA, MJIAM, Volume 40, Number 3, 143–167

Sustainable Energy Management: Artificial Intelligence-Based Electricity Consumption Prediction in Limited Dataset Environment for Industry Applications

[1]Zun Liang Chuan*, [2]Lit Ken Tan, [3]Angel Wee Chi Chyin, [3]Yim Hin Tham, [3]Shao Jie Ong, [3]Jia Yi Low and [4]Chong Yeh Sai*

[1] Centre for Mathematical Sciences, Universiti Malaysia Pahang Al-Sultan Abdullah, Lebuh Persiaran Tun, Khalil Yaakob, 26300, Kuantan Pahang, Malaysia

[2]Department of Mechanical Precision Engineering, Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia.

[3]Faculty of Chemical and Process Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Lebuh Persiaran Tun Khalil Yaakob, 26300, Kuantan Pahang, Malaysia.

[4] Ever AI Holdings Sdn Bhd, 12, Jalan Anggerik Aranda 31/170C, 40460 Kota Kemuning, Shah Alam, Selangor, Malaysia

*Corresponding author: chuanzl@umpsa.edu.my, cysai@ever-technologies.com

Article history
Received: 1 August 2024
Received in revised form: 24 November 2024
Accepted: 10 December 2024
Published on line: 31 December 2024

Abstract Electricity has been a key driver of global socioeconomic development and sustainability for both developed and developing nations. In Malaysia, electricity is primarily generated by burning fossil fuels, emitting greenhouse gases (GHG) that adversely impact the environment and public health. Therefore, accurately predicting electricity consumption is crucial for economic management, security analysis, facility scheduling for generation and distribution, and maintenance planning. This study aimed to develop a modified stacked ensemble multivariable Artificial Intelligence (AI)-based predictive algorithm, specifically Stacked Simple Linear Regression and Multiple Linear Regression (SLR-MLR), and Stacked Simple Linear Regression and Multiple Non-Linear Regression (SLR-MNLR) utilizing the Cross Industry Standard Process for Data Mining (CRISP-DM) data science methodology. The proposed AI-based predictive algorithm aimed to provide predictive insights and interpret the impact of significant economic, environmental, and social clustered determinants on electricity consumption in Malaysia. The analysis revealed that the SLR-MLR predictive algorithm better fits Malaysias limited electricity consumption dataset compared to the existing Stacked SLR and ϵ-Support Vector Regression (SLR-ϵ-SVR) and SLR-MNLR predictive algorithms. It identified key economic and environmental clustered determinants that significantly impact electricity consumption in Malaysia. In academia, this study proposed an innovative SLR-MLR predictive algorithm and utilized a novel statistical approach to evaluate and select the superior predictive algorithm. Practically, it offered valuable insights for policymakers to craft efficient regulations, manage the energy sector proactively, and anticipate electricity generation and consumption trends. These contributions align with Malaysias economic and environmental sustainability goals outlined in the Twelfth Malaysia Plan, the Madani

Economy Framework, the National Energy Policy 2022-2040, and the National Energy Transition Roadmap (NETR) agenda.

Keywords multivariable AI-based predictive algorithm; electricity consumption; modified stacked ensemble; CRISP-DM data science methodology; sustainability goals.

Mathematics Subject Classification 62P20; 62P25; 65Z05; 68T99.

1 Introduction

Electricity is the primary driver of sustainable socioeconomic development globally, impacting both developed and developing nations, and is crucial for individuals and national economic activities. At the individual level, electricity enhances daily life by powering household appliances utilized for food preparation, personal care, and, leisure and by enabling access to information and communication technologies essential for education, healthcare, and social connectivity. On a national scale, electricity powers manufacturing industries facilitates the production of goods and services, drives local economies, and creates jobs, particularly for small and medium-sized enterprises (SMEs). Additionally, accurate prediction of electricity consumption is paramount for economic management, electricity security analysis, facility scheduling for electricity generation and distribution, and maintenance planning. Reliable prediction algorithms optimize electricity generation, distribution, and pricing, ensuring a stable supply, and meeting demand efficiently. Effective electricity management is essential for achieving Sustainable Development Goals (SDGs) such as No Poverty (SDG1), Zero Hunger (SDG2), Good Health and Well-Being (SDG3), Affordable Clean Energy (SDG7), Decent Work and Economic Growth (SDG8), Industry, Innovation and Infrastructure (SDG9), Sustainable Cities and Communities (SDG11), and Climate Action (SDG13).

In Malaysia, the primary method of electricity generation heavily relies on the combustion of fossil fuels [1, 2], particularly coal and natural gas. This reliance leads to significant emissions of greenhouse gases (GHGs), notably carbon dioxide (CO2), contributing significantly to global warming and climate change. In response to these environmental concerns, national and international initiatives, such as the Paris Agreement, have been implemented to promote a low-carbon economy, aiming for cleaner air, energy security, job creation for economic sustainability, and achieving zero CO2 emissions by 2050. Malaysia has recently introduced a revised National Energy Policy (DTN) 2022-2040 and National Energy Transition Roadmap (NETR) agenda [3, 4], designed to drive socioeconomic development and environmental sustainability by enhancing macroeconomic resilience, ensuring energy security, promoting social equity and affordability, and fostering environmental sustainability. Despite commitments to the Low Carbon Nation Aspiration and the Long-Term Low Emission Development Strategy (LT-LEDS) targeting net-zero GHG emissions by 2050, fossil fuels continue to dominate electricity generation in Malaysia. Figure 1 illustrates the limited progress in transitioning from traditional to renewable energy sources. While renewable energy significantly mitigates GHG emissions, it also poses challenges across the four principal pillars of the SDGs: economic, environmental, human, and social. Renewable energy projects can directly and indirectly impact ecosystem change and biodiversity loss, influenced by factors such as habitat alteration, climate change, invasive species introduction, overexploitation, and pollution [5,6]. Moreover, renewable energy sources such as biofuels, hydropower, solar photovoltaic, and waste face economic, financial, and technological barriers hindering their widespread adoption. These barriers include high initial investment costs, limited market competitiveness without subsidiaries, economic barriers, and variability in electricity generation due to meteorological conditions [2, 5, 6].

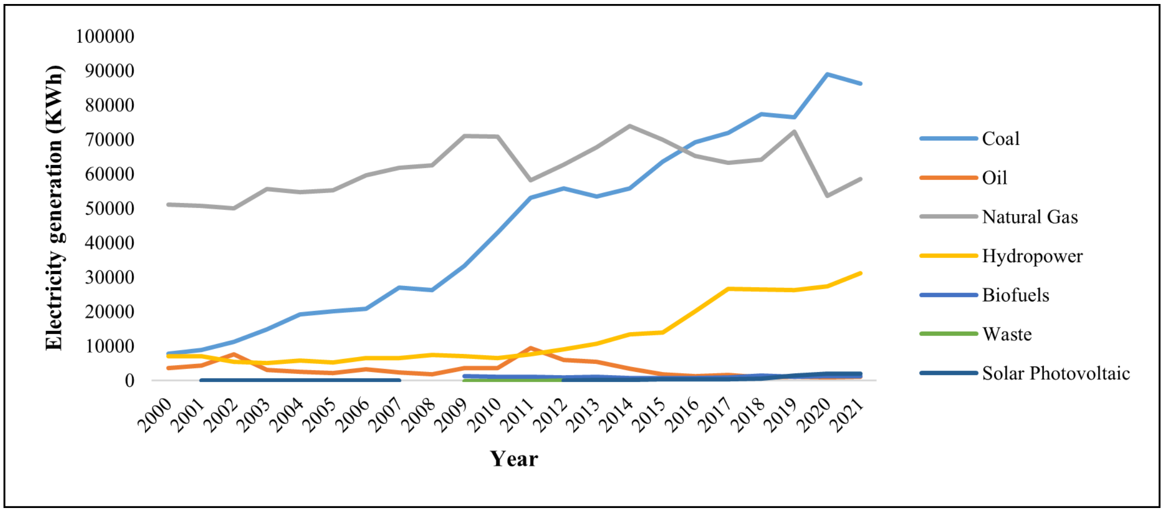

Figure 1: Electric Evolution: Charting Malaysias Power Sources from 2000 to 2021 Source: International Energy Agency [7].

Therefore, the principal objective of this study is to propose innovative modified stacked ensemble multivariable AI-based predictive algorithms, including Stacked Simple Linear Regression and Multiple Linear Regression (SLR-MLR) and Stacked Simple Linear Regression and Multiple Non-Linear Regression (SLR-MNLR) based on the Cross Industry Standard Process for Data Mining (CRISP-DM) data science methodology, and assess their efficacy compared to existing modified stacked ensemble univariate AI-based predictive algorithms (Stacked SLR and ϵ-Support Vector Regression (SLR-ϵ-SVR)) for modeling and forecasting electricity consumption in Malaysia. These predictive algorithms are designed to identify statistically significant determinants impacting electricity consumption, aligning with the core pillars of SDGs encompassing economic, environmental, and social. However, the human pillar was not considered in this study due to insufficient relevant time-series datasets. This study is crucial for economic management, electricity security analysis, facility scheduling for electricity generation and distribution, and maintenance planning. Specifically, an accurate prediction of electricity consumption is essential to avoid overestimation leading to inefficient energy resource utilization and distribution issues, or underestimation causing inadequate reserves and high costs during peak periods. In summary, the proposed innovative AI-based predictive algorithms provide valuable insights for policymakers, enabling them to craft more efficient regulations, proactively manage the energy sector, and gain early insights into electricity generation and consumption trends. This contributes to environmental sustainability goals outlined in the Twelfth Malaysia Plan, the Madani Economy Framework, the National Energy Policy (DTN) 2022-2040, and the National Energy Transition Roadmap (NETR) agenda. To achieve the principal objective of this article, the remainder of this article is structured as follows: Section 2 provided an overview of the literature review and identified research gaps, while Section 3 detailed the research methodology based on the CRISP-DM data science methodology. Additionally, Section 4 presented the findings from Exploratory Data Analysis (EDA), predictive performance evaluation, and deployment of the proposed AI-based predictive algorithms. Finally, Section 5 concluded the study with a summary of key insights and recommendations.

2 Related Works

This section reviews the literature on time-series predictive algorithms for predicting electricity consumption across various timescales, with a focus on studies conducted in Malaysia, aligning with the primary objective of this study. Malaysian case studies generally utilize both univariate and multivariable classical statistical predictive algorithms, as well as AI-based predictive algorithms. A key limitation across these approaches is inadequate feature engineering, which, along with constraints in predictive performance, impacts the accuracy and practicality of these predictive algorithms for forecasting future electricity consumption. Additionally, while AI-based predictive algorithms are widely utilized in developed and developing nations [8–13], they have received limited attention in Malaysian studies. Furthermore, some previous studies [14, 15] reveal misconceptions from a statistical perspective implementing AI-based predictive algorithms, highlighting a gap that this study aims to address.

To illustrate the emphasis on classical statistical predictive algorithms in Malaysia research, Razak et al. [16] evaluated several classical univariate predictive algorithms, including Nave, Exponential Smoothing (ES), Seasonal Holt-Winters (SHW), Autoregressive Moving Average (ARMA), AutoRegressive AutoRegressive (ARAR), and Regression with ARMA Errors. Their study forecasted monthly mean maximum electricity load demand in Malaysia spanning from September 2000 to December 2004, concluding that the second-order Autoregressive (AR(2)) predictive algorithm outperformed other predictive algorithms. Similarly, Karim and Alwi [17] compared predictive algorithms for forecasting electricity load consumption at Universiti Teknologi Petronas (UTP), finding that the ES predictive algorithm outperformed the Moving Average (MA) predictive algorithm in both semester-on and semester-off seasons.

In another study, Ozoh *et al.* [18] introduced the modified Newton (MN) predictive algorithm to predict monthly electricity consumption at Universiti Malaysia Sarawak (UNIMAS) from 2009 to 2012. Their comparison with the univariate Forecasting Time-Series Technique (FTST) and univariate Artificial Neural Network (ANN) predictive algorithm showed that the MN predictive algorithm outperformed both benchmark FTST and ANN predictive algorithms. Meanwhile, Lee *et al.* [19] investigated several classical statistical univariate predictive algorithms, including Simple Moving Average (SMA), Weighted Moving Average (WMA), Simple Exponential Smoothing (SES), Holt Linear Trend (HLT), Holt-Winters (HW), and Centered Moving Average (CMA). Analyzing the monthly electricity consumption dataset from January 2011 to December 2017 at Universiti Tun Hussein Onn (UTHM), they found that the HW predictive algorithm had the lowest Mean Absolute Percentage Error (MAPE) among the tested predictive algorithms, although their assessment may be limited by a lack of comprehensive consideration of other Goodness-of-Fit (GoF) measures presented in their study.

Sim *et al.* [20] proposed forecasting monthly electricity consumption from January 2009 to December 2018 at Universiti Tun Hussein Onn Malaysia (UTHM) utilizing the Seasonal Autoregressive Integrated Moving Average (SARIMA) predictive algorithm, due to the dataset exhibiting seasonal-periodic fluctuations. In a different approach, Abdullah and Leong [21] investigated the linear association between the annual final energy consumption, and three economic determinants (Gross Domestic Product (GDP), population, and tourism) spanning from 2001 to 2012 utilizing a classical statistical multivariable multiple linear regression (MLR) predictive algorithm. Their analysis indicated that the sample data of final energy consumption was well-fitted for the MLR predictive algorithm, considering all three economic determinants, revealing a coefficient of determination higher than 90%. Recently, Ahmad *et al.* [22] applied the Autoregressive Integrated Moving Average (ARIMA) predictive algorithm to forecast electricity load from solar rooftop datasets in a public university. However, their study exhibits several methodological issues. They erroneously treated the univariate ARIMA predictive algorithm as a multivariable predictive algorithm, a fundamental statistical error. Additionally, their time-series analysis methodology was flawed; they clustered time-series datasets from April, October, and November 2019 into usage categories (low, medium, and high), neglecting the sequential nature of time-series datasets. This approach introduces bias and undermines the reliability of their findings. Moreover, this study lacks a robust statistical approach necessary for accurate time-series analysis and fails to evaluate predictive performance comprehensively. In summary, Ahmad *et al.* [22]s study on forecasting electricity load based on solar rooftop datasets suffers from significant methodological shortcomings, casting doubt on the validity of its conclusions. In contrast to the classical statistical predictive algorithms, AI-based predictive algorithms have gained prominence in recent years for modeling and forecasting electricity consumption on various scales, from individual premises to international levels. For instance, Shapi *et al.* [14] conducted a study comparing three univariate AI-based predictive algorithms-ANN, k-nearest Neighbor (k-NN), and SVR-for predicting the electricity consumption of smart commercial buildings in the Klang Valley from June 2018 to December 2018 utilizing Microsoft Azure's cloud-based machine learning platform. However, this study contains inaccuracies regarding fundamental machine learning concepts, particularly in confusing regression with classification paradigms, and misinterpreting predictive performance evaluation methods. Consequently, the findings of this study are not discussed further in this section. Additionally, the conclusion presented in this article is biased due to solely identifying the superior predictive algorithm based on Root Mean Square Error (RMSE), whereas this study has evaluated the predictive performance utilizing other GoF measures such as Normalized Root Mean Square Error (NRMSE) and MAPE.

Similarly, Lee *et al.* [15] proposed comparing the effectiveness of various AI-based predictive algorithms in predicting monthly electricity consumption across ten nations, including Malaysia. These AI-based predictive algorithms considered in this study include ANN, Adaptive Neuro-Fuzzy Inference System (ANFIS), Least Squares Support Vector Machines (LSSVMs), and Fuzzy Time-Series (FTS). The findings revealed that no universal AI-based predictive algorithm is well-fitted for all nations due to the distinct electricity availability and the development levels of the region. However, this article presented a misconception about the evaluation and forecasting of predictive performance. Meanwhile, Chuan *et al.* [23] conducted a comprehensive comparison of 72 modified stacked ensemble univariate SVR-based predictive algorithms to predict annual electricity consumption in Malaysia. They investigated various insensitive loss functions (ϵ- and ν-insensitive) and kernel functions (linear, polynomial, radial basis function, and sigmoid). Their analysis revealed that the -insensitive loss function associated with a third-degree polynomial kernel function is the most effective for short-term annual electricity consumption forecasting, particularly in a limited dataset environment. Furthermore, this study validated the superior AI-based predictive algorithm by deploying it to forecast electricity consumption for the next five years, demonstrating reasonable and practical results. In contrast, Mustapa *et al.* [24] proposed a multivariable Non-linear AutoRegressive with Exogenous Input-Artificial Neural Network (NARX-ANN) predictive algorithm for predicting daily electricity consumption for the Faculty of Electrical Engineering, Universiti Teknologi MARA in Pasir Gudang Johor during semester. They employed a multivariable MLR predictive algorithm for benchmark comparisons. This study concluded that the NARX-ANN predictive algorithm outperformed the MLR predictive algorithm, although the determinants involved were selected instinctively. However, the results may be biased and unreliable due to the analysis skipping some daily time points, treating them as missing values for their study.

Furthermore, Ayodele *et al.* [25] compared the effectiveness of two multivariable Multiple Non-Linear Regression (MNLR) and three multivariable NARX predictive algorithms in predicting annual final energy consumption per capita from 1980 to 2018, focusing solely on socioeconomic determinants. Their findings demonstrated that the MNLR predictive algorithm based on an exponential function outperformed the MNLR based on a polynomial function, and NARX predictive algorithms, However, the methodology for concluding the superiority of the proposed AI-based predictive algorithm could be improved due to conflicts among the GoF measures.

In a similar vein, Sena *et al.* [26] proposed predicting household monthly electricity consumption based on techno-socioeconomic determinants (socio-demographic, building characteristics, occupant behavior, and appliance characteristics) utilizing a multivariable ANN predictive algorithm. This study collected a primary dataset utilizing a simple random sampling method from 214 Universiti Teknologi Malaysia International Kuala Lumpur Campus students via survey questions from November 2017 to January 2018. Their findings revealed that the multivariable ANN predictive algorithm, incorporating socio-demographic, occupant behavior, and appliance characteristics is superior to other predictive algorithms. However, this studys approach of identifying determinants based on all possible combinations may lead to suboptimal analysis results due to ignoring statistical evidence, being computationally expensive, time-inefficient, and resulting in unreliable predictive algorithms.

Previous studies have shown that no universal AI-based and non-AI-based predictive algorithm can effectively manage the complexity of grid consumption across different timescales and levels. Classical and univariate AI-based predictive algorithms frequently struggle with predictive performance for short-term, mid-term, and long-term periods due to limitations in incorporating impacting economic, environmental, and social determinants. Additionally, while previous studies emphasize predictive performance, they frequently lack the in-depth analysis needed to make these predictive algorithms practical for future electricity consumption predictions. A key limitation in these classical and AI-based predictive algorithms is inadequate feature engineering, especially in feature selection, which frequently respectively leads to a lack of determinants and redundant determinants that degrade predictive accuracy. To address these gaps, this article proposes a modified stacked ensemble multivariable AI-based predictive algorithm with enhanced feature engineering, designed to accurately predict annual electricity consumption in Malaysia.

3 Research Methodology

The CRISP-DM data science methodology was employed as the research framework for this study due to its dynamic nature and flexibility. Firstly, it effectively guided the research process from defining business objectives to deploying research outcomes, ensuring and delivering commercial value when utilized properly. Secondly, this methodology was adaptable and has been successfully applied in various fields, including agriculture economics [27], education [28, 29], energy economics [23], healthcare [30], and management [31]. Theoretically, the CRISP-DM data science methodology comprised six primary phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. These phases are detailed in subsequent sections and were tailored to this study. Although CRISP-DM data science methodology supported iterative transitions between phases, this study described a linear progression through this methodology.

3.1 Business Understanding

The primary business objective of this study was to deliver valuable short-term predictive insights into annual electricity consumption and identify statistically significant determinants impacting electricity consumption in Malaysia. Consequently, the core aim of data mining in this research was to propose innovative modified stacked ensemble multivariable AI-based predictive algorithms. These insights were crucial for optimizing economic strategies, ensuring electricity security analysis, efficiently scheduling electricity generation and distribution facilities, and planning maintenance operations.

Despite its importance, this research encountered various challenges, including misconceptions, issues with the reliability of research outcomes, and financial limitations. To address these challenges, this study ensured that all analyses adhered to key statistical concepts, including EDA, modeling, and evaluation. The pre-defined assumptions of the multivariable AI-based predictive algorithms were rigorously tested through diagnostic checks on the trained predictive algorithms. Additionally, the dataset was split into various training-to-test ratios (TTR) to investigate potential overfitting issues. These analyses were essential for ensuring the reliability of research outcomes before deploying the proposed AI-based predictive algorithm for forecasting.

Moreover, this study utilized open-source datasets and employed free tools such as Microsoft Excel and R for statistical analyses in a mid-range computing environment (Intel(R) CoreTM i5-10210U CPU@1.60GHz), due to limited financial resources. Despite these efforts to incorporate a broad range of relevant determinants, specific potential environmental determinants such as humidity and solar radiation, economic determinants such as the share of Gross Domestic Product (GDP) in commercial, residential, transport, and agriculture, and coal and natural gas prices, as well as social determinants including household size, income distribution, educational levels, and employment rates across economic sectors, were unavailable in the open-source databases or had high rates of missing values, and therefore were not considered in this study. Additionally, human determinants such as public health determinants primarily related to GHGs, which significantly impact electricity consumption were also excluded. To achieve the principal objective, this study followed a strategized plan based on the linear progression of the CRISP-DM data science methodology. Detailed descriptions of the employed statistical analysis tools are provided in subsequent sections corresponding to each phase.

3.2 Data Understanding

The primary objective of the data understanding phase was to establish the foundational aspects of business understanding. This involved comprehensive data collection, thorough description of data, ensuring data quality through verification, and conducting detailed data exploration. Specifically, for this study, the environmental datasets ($\boldsymbol{X_1}$: average minimum surface air temperature (^oC); $\boldsymbol{X_2}$: average maximum surface air temperature (^oC); $\boldsymbol{X_3}$: annual precipitation (mm)) were sourced from the Climate Change Knowledge Portal (CCKP), while the electricity consumption dataset ($\boldsymbol{Y}$: final electricity consumption (ktoe)), and economic ($\boldsymbol{X_4}$: coal and coke supply (ktoe); $\boldsymbol{X_5}$: crude oil supply (ktoe); $\boldsymbol{X_6}$: natural gas supply (ktoe); $\boldsymbol{X_7}$: coal import (ktoe); $\boldsymbol{X_8}$: crude oil import (ktoe); $\boldsymbol{X_9}$: diesel import (ktoe); $\boldsymbol{X_{10}}$: natural gas import (ktoe); $\boldsymbol{X_{11}}$: coal export (ktoe); $\boldsymbol{X_{12}}$: crude oil export (ktoe); $\boldsymbol{X_{13}}$: diesel export (ktoe); $\boldsymbol{X_{14}}$: natural gas export (ktoe); $\boldsymbol{X_{15}}$: Gross GDP per capita at current prices (RM)) and social ($\boldsymbol{X_{16}}$: population (000)) datasets were sourced from the Malaysia Energy Information Hub (MEIH). Additionally, the economic datasets ($\boldsymbol{X_{17}}$: exports goods and services (constant 2015 US$); $\boldsymbol{X_{18}}$: imports goods and services (constant 2015 US$); $\boldsymbol{X_{19}}$: Foreign Direct Investment (FDI) net inflows (% in Gross Domestic Product (GDP)); $\boldsymbol{X_{20}}$: FDI net outflows (% in Gross Domestic Product (GDP)); $\boldsymbol{X_{21}}$: share of manufacturing in GDP (%); $\boldsymbol{X_{22}}$: world crude oil (US$)), environmental ($\boldsymbol{X_{23}}$: carbon dioxide ($CO_2$) by coal (tonnes); $\boldsymbol{X_{24}}$: CO_2 by gas (tonnes); $\boldsymbol{X_{25}}$: CO_2 by oil (tonnes)), and social ($\boldsymbol{X_{26}}$: percentage urban (%)) were sourced from Our World In Data (OWID).

These datasets are sourced from reputable open-access databases, widely utilized in academic publications by leading publishers such as Elsevier and Springer, the selection of these sources was based on their credibility, open-access nature, and comprehensive coverage of relevant determinants that align with this study's objectives. Additionally, the datasets were structured in Comma-Separated Value (CSV) format. In this study, the annual final electricity consumption ($\boldsymbol{Y}$) served as the endogenous variable, while the remaining variables ($\boldsymbol{X_i}; i = 1, 2, , 26$) represented potential determinants directly and indirectly impacting $\boldsymbol{Y}$. It is noteworthy that all variables involved in this study were a mix of discrete and continuous determinants which were selected based on literature and the core pillars of SDGs encompassing economic ($\boldsymbol{X_4}, \boldsymbol{X_5}, \boldsymbol{X_6}, \boldsymbol{X_7}, \boldsymbol{X_8}, \boldsymbol{X_9}, \boldsymbol{X_{10}}, \boldsymbol{X_{11}}, \boldsymbol{X_{12}}, \boldsymbol{X_{13}}, \boldsymbol{X_{14}}, \boldsymbol{X_{15}}, \boldsymbol{X_{17}}, \boldsymbol{X_{18}}, \boldsymbol{X_{19}}, \boldsymbol{X_{20}}, \boldsymbol{X_{21}}$, and $\boldsymbol{X_{22}}$), environmental ($\boldsymbol{X_1}, \boldsymbol{X_2}, \boldsymbol{X_3}, \boldsymbol{X_{23}}, \boldsymbol{X_{24}}$, and $\boldsymbol{X_{25}}$) and social ($\boldsymbol{X_{16}}$ and $\boldsymbol{X_{26}}$), highlighting the quantitative research nature of this study. However, due to the large number of potential determinants, this article does not delve into the specific impact of each determinant but focuses on the core SDG pillars discussed in the previous section.

In verifying the quality of the datasets acquired from these databases and conducting data exploration, this study utilized frequency distributions to tabulate the occurrence of missing values, confirming that no missing values were present in the acquired datasets. This approach was essential to avoid the need for imputation, thereby preserving the integrity of the time-series datasets. In simpler terms, imputing missing values, such as row omissions, or utilizing techniques like Probabilistic Principal Component Analysis-based (PPCA-based) algorithms, was unnecessary since no missing values were identified in the acquired datasets. Additionally, graphical representations such as line graphs were employed for each acquired variable to verify data quality, ensuring that the datasets exhibited equally successive time points, thereby overcoming misconceptions from previous studies detailed in Section 2. The correlogram and

Shapiro-Wilks normality test were also employed to explore the linear association among the variables and the normality of each variable, respectively. Furthermore, the first four statistical L-moments were utilized for summarizing the core characteristics of all variables. Mathematically, L-mean (LM_1), L-Coefficient of variation (LM_2), L-Skewness (LM_3), and L-Kurtosis (LM_4) can be expressed in Equations (1)-(4).

$$LM_1 = \binom{n}{1}^{-1} \sum_{j=1}^{n} (\lambda_j)_{(j:n)} \tag{1}$$

$$LM_2 = \frac{1}{2}\binom{n}{2}^{-1} \sum_{j=1}^{n} \left\{ \binom{j-1}{1} - \binom{n-j}{1} \right\} (\lambda_j)_{(j:n)} \tag{2}$$

$$LM_3 = \frac{1}{3}\binom{n}{3}^{-1} \sum_{j=1}^{n} \left\{ \binom{j-1}{2} - 2\binom{j-1}{1}\binom{n-j}{1} + \binom{n-j}{2} \right\} (\lambda_j)_{(j:n)} \tag{3}$$

$$LM_4 = \frac{1}{4}\binom{n}{4}^{-1} \sum_{j=1}^{n} \left\{ \binom{j-1}{3} - 3\binom{j-1}{2}\binom{n-j}{1} + 3\binom{j-1}{1}\binom{n-j}{2} - \binom{n-j}{3} \right\} (\lambda_j)_{(j:n)} \tag{4}$$

where n represented the sample size of the acquired time-series datasets, $\binom{\cdot}{\cdot}$, represented the combination function,$\lambda_j; j = 1, 2, ..., n$ represented the j th measurements of time-series dataset of size n corresponding to each $\boldsymbol{Y} = \left[y_j\right]_{n\times 1}$ and $\boldsymbol{X}_i; i = 1, 2, ..., 26$ and $\cdot_{j;n}$ represented the order statistics.

3.3 Data Preparation

Data exploration during the data understanding phase provided insights into hidden patterns within the acquired time-series dataset. Consequently, the primary aim of the data preparation phase, also known as data wrangling, was to prepare the final dataset for modeling. This included data selection, cleaning, transformation, construction, integration, formatting, splitting, and feature engineering. However, data normalization and data construction were not applied in this study. Data normalization, a technique frequently employed to scale or transform determinants into a specific range, was deemed unnecessary for the proposed predictive algorithms, as the acquired datasets were in a suitable quantitative form for modeling. Additionally, since the acquired time-series datasets were in quantitative form, they could be straightforwardly utilized for modeling, making the creation of dummy variables for qualitative variables unnecessary.

For data selection, all aforementioned acquired variables were included in this study. Regarding data cleaning, outlier detection was performed utilizing the 1.5 interquartile range (IQR) and 3IQR rules to identify mild and extreme outliers, respectively. Mild outliers were identified as observations falling below the Lower Inner Fence (LIF) (Equation 5) or exceeding the Upper Inner Fence (UIF) (Equation 6) of the 1.5IQR rule. Extreme outliers were defined as observations below the Lower Outer Fence (LOF) (Equation 7) or exceeding the Upper Outer Fence (UOF) (Equation 8) of the 3IQR rule. To address extreme outliers, a capping technique was applied by replacing them with the LOF and UOF limits, thereby retaining the datasets structure while minimizing the impact of extreme values. No adjustments were required for mild outliers, as they did not significantly impact the overall data distribution.

$$LIF_i = Q_1(\lambda_i) - 1.5(Q_3(\lambda_i) - Q_1(\lambda_i)) \tag{5}$$

$$UIF_i = Q_3(\lambda_i) + 1.5(Q_3(\lambda_i) - Q_1(\lambda_i)) \tag{6}$$

$$LOF_i = Q_1(\lambda_i) - 3(Q_3(\lambda_i) - Q_1(\lambda_i)) \tag{7}$$

$$UOF_i = Q_3(\lambda_i) + 3(Q_3(\lambda_i) - Q_1(\lambda_i)) \tag{8}$$

where $Q_1(\lambda_i)$ and $Q_3(\lambda_i)$ represented the 25th and 75th percentile observation of the acquired time-series of $\boldsymbol{\lambda} = \left[\lambda_j\right]_{n\times 1}$ corresponding to each $\boldsymbol{Y}$ and $\boldsymbol{X}_i$ respectively.

Given that the time-series datasets in this study were sourced from three primary open-source databases, they were integrated to prepare for modeling, and compiled into a medium-dimensional dataset stored in a CSV file. The compiled dataset was divided into training and test sets utilizing various pre-defined TTR of 60:40, 70:30, 80:20, and 90:10, following the holdout cross-validation method for evaluating predictive performance. The rationale behind utilizing different TTRs was to determine the appropriate sample size for training a robust modified stacked ensemble AI-based predictive algorithm while evaluating overfitting. The ideal TTR depends on factors such as data quality and dimensionality, with smaller data dimensions requiring larger training ratios to improve algorithm performance. While feature engineering is an essential part of data preparation, the detailed feature engineering process, such as feature selection is discussed in a separate section. This is because the feature selection performed in this study was specifically tailored to the proposed modified stacked ensemble AI-based predictive algorithms.

3.4 Modeling

The divided training set based on various pre-defined TTRs was utilized for training a parsimonious modified stacked ensemble multivariable AI-based predictive algorithm, which is the primary objective of data mining in this article. To pursue this objective, a well-established automated wrapper stepwise features selection (AWSFS) method was utilized, combining forward selection and backward selection methods associated with SLR-MLR and SLR-MNLR predictive algorithms in resulting a parsimonious predictive algorithm, where MLR and MNLR predictive algorithms as expressed respectively in Equations (9)-(10) were employed. It is noted that the SLR predictive algorithm, while not specified in this article, represents a special case of the MLR predictive algorithm, sharing closely similar mathematical equations.

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{25} \hat{\beta}_i X_i \tag{9}$$

$$\hat{y} = exp\left(\hat{\beta}_0 + \sum_{i=1}^{25} \hat{\beta}_i X_i\right) \tag{10}$$

This study emphasized that both MLR and MNLR predictive algorithms employed a machine learning approach distinct from the classical statistical modeling approach. Furthermore, the proposed modified stacked ensemble multivariable AI-based predictive algorithms (SLR-MLR and SLR-MNLR) were adapted from the traditional stacked ensemble concept in machine learning. In traditional stacked ensembles, multiple machine learning algorithms are utilized at the base layer, with a meta-predictive algorithm assigning specific weights to the outcomes. However, this classical approach frequently suffers from challenges related to interpretability and complexity in weight assignment, making it impractical in certain contexts such as energy prediction.

To overcome these challenges, this study utilized a modified stacked ensemble approach that employed a single predictive algorithm at the base layer, which enhanced interpretability and reduced complexity. Specifically, the SLR predictive algorithm was selected for the base layer to forecast future observations for each statistically significant determinant in the resulting parsimonious SLR-MLR and SLR-MNLR predictive algorithms. This approach enabled the MLR and MNLR predictive algorithms to serve dual roles. They were not merely a meta-algorithm for combining forecasts but also crucial components in the feature selection process via the AWSFS method. The inclusion of the SLR-MNLR predictive algorithm as an effectiveness comparison was particularly relevant, given that Malaysias annual electricity consumption exhibits a compound growth pattern, which aligns with Equation (10) and captures these complex relationships more effectively.

In pursuit of the parsimony principle, this study initiated the feature selection process utilizing the AWSFS method to refine the feature set. AWSFS combines forward selection and backward selection elimination, iteratively adding and removing variables to identify a subset of determinants that minimizes the Akaike Information Criterion (AIC). AIC serves as a statistical measure that balances algorithm complexity and GoF, guiding the selection of a parsimonious predictive algorithm. While AWSFS provided an initial set of selected determinants, further refinement was carried out through rigorous statistical hypothesis testing for each $\boldsymbol{X}_i$. Merely those determinants with a statistically significant impact on $\boldsymbol{Y}$ were retained for inclusion in the proposed modified stacked ensemble multivariable AI-based predictive algorithm. Specifically, this study employed the parametric t-test to assess the significance of each determinant, ensuring that non-significant determinants were excluded. This process of determinant refinement is essential to ensure the final predictive algorithms include merely the most relevant determinants, thus enhancing both the performance and interpretability of the proposed modified stacked ensemble multivariable AI-based predictive algorithm.

This dual-layer approach ensured the final predictive algorithm was both parsimonious and statistically robust, preserving its predictive accuracy while avoiding overfitting. Notably, Equation (10) employed in this study differs slightly from the predictive algorithm proposed by [25], which utilized $\hat{y} = \hat{\beta}_0 exp\left(\sum_{i=1}^{25} \hat{\beta}_i log(X_i)\right)$, where $\log(\cdot)$ denoted the natural logarithm utilized for dataset stabilization. By integrating AWSFS and statistical hypothesis testing, this study ensured the proposed modified stacked ensemble multivariable AI-based predictive algorithm effectively balances simplicity and predictive performance.

Moreover, Ayodele *et al.* [25] advocate for the Non-Linear Least-Squares (NLLS) method for estimating regression parameters. In contrast, this study advocated for the Ordinary Least Squares (OLS) parameters estimation method post-linearization of Equation (10). The rationale behind this selection is twofold: first, OLS offered computational efficiency suitable for medium-dimensional time-series datasets, avoiding the prolonged convergence times and resource-intensive computations associated with NLLS. Second, OLS provided stable parameter estimates, less susceptible to divergence caused by inappropriate initial guess values frequently encountered with NLLS.

Regarding the impact of post-linearization, the transformation was introduced by applying the logarithmic function to both sides of Equation (10) to stabilize the data and make the relationship between variables more linear, which is a necessary condition for the OLS parameters estimation method to perform effectively. Importantly, this transformation does not compromise the integrity of the original dataset. While it changes the form of the equation, it preserves underlying relationships among the variables, ensuring that the modeling process remains consistent with the data's original structure. This leads that despite the linearization, the datasets integrity is retained, and the transformation aligns with the mathematical principles that underpin the predictive algorithm. Consequently, OLS remains a robust and efficient method for parameter estimation in this context, providing a practical and efficient alternative to NLLS in complex time-series modeling scenarios.

Furthermore, traditional predictive algorithms such as ARIMA and MLR were not employed as benchmarks in this study. ARIMA is a univariate predictive algorithm that does not account for economic, environmental, and social determinants, representing a research gap this study aims to address. Similarly, the MLR predictive algorithm, when not associated with a stacked ensemble, is limited for short-term forecasting, as discussed in Section 4.2. These limitations highlight the principal reason ARIMA and MLR predictive algorithms are inappropriate as benchmark comparisons for the proposed modified stacked ensemble AI-based predictive algorithm. This further emphasizes the advantages of the proposed predictive algorithm in addressing the complexities of the datasets.

Meanwhile, the ANN-based predictive algorithm was also not utilized as a benchmark in this study due to several limitations. While ANN-based predictive algorithms can automatically perform feature selection, they do not provide statistical evidence, such as p-value, which can lead to sub-optimal decision-making and interpretability. Moreover, the complexity of determining the optimal number of epochs, hidden layers, and neurons, along with constraints related to forecasting future electricity consumption, limits their practical applicability for this study.

In contrast, the univariate SLR-ϵ-SVR predictive algorithm was employed for benchmark comparison. Unlike the ARIMA predictive algorithm, it does not suffer from similar limitations and allows effective incorporation of identified significant determinants. Moreover, its methodology closely resembles that of the proposed multivariable AI-based predictive algorithm, making it a more suitable selection for comparison. Additionally, while Chuan *et al.*s predictive algorithm employed a univariate approach that focused on a narrower set of variables, this study proposed a predictive algorithm that adopts a multivariable framework. This approach integrates multiple significant determinants identified through robust statistical testing on the feature selection stage, reflecting a broader scope and higher flexibility in modeling complex relationships, particularly in the context of electricity consumption impacted by diverse economic, environmental, and social factors.

3.5 Evaluation

The evaluation phase aimed to technically assess the trained AI-based predictive algorithms, focusing on results evaluation, reviewing processes, and determining the next steps, whether to iterate previous phases, initiate new projects, or proceed to deployment. For this study, the evaluation of the results and the reviewing process were conducted by diagnosing overfitting in the trained predictive algorithms and implementing a newly proposed systematic approach based on nonparametric tests to verify its absence. This approach compared the GoF measures evaluated across training and test sets utilizing pre-defined ratios. Specifically, parameters for both SLR-MLR and SLR-MNLR predictive algorithms were estimated utilizing the OLS parameters estimation method. The adequacy of the predictive algorithms was assessed based on principal assumptions reliant on the OLS parameters estimation method, including normality, homoscedasticity, independence of residuals, and multicollinearity, verified through utilizing the nonparametric, including Shapiro-Wilks normality test, Breusch-Pagan test, Run test, and Variance Inflation Factor (VIF), with a threshold value of 10. The diagnostic check for these OLS assumptions is critical for ensuring the reliability and interpretability of the resulting parsimonious predictive algorithm. Importantly, since these statistical tests are nonparametric, they are not tied to the normality assumption typically required by traditional parametric statistical tests, enhancing their applicability in this context.

Furthermore, this study proposed a systematic approach to monitor overfitting, aiming to improve upon conventional approaches reliant solely on the adjusted-R2. While high adjusted-R2 values typically indicate superior predictive algorithm fit, however, it may conflict with other GoF measures. To provide a rigorous evaluation, this study introduced the nonparametric Wilcoxon Signed Ranked (WSR) test to verify for overfitting. GoF measures considered included Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) as defined in Equations (11)-(13).

$$RMSE_k = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_{jk} - y_{jk}^{pred}\right)^2} \tag{11}$$

$$MAE_k = \frac{1}{n}\sum_{j=1}^{n}\left|y_{jk} - y_{jk}^{pred}\right| \tag{12}$$

$$MAPE_k = \frac{100}{n}\sum_{j=1}^{n}\left|\frac{y_{jk} - y_{jk}^{pred}}{y_{jk}}\right| \tag{13}$$

where $y_{jk}^{pred}; k = 1, 2, ..., 10$ represented the predicted annual electricity consumption based on the k th trained AI-based predictive algorithm.

Mathematically, these GoF measures quantify the discrepancy between y_{jk} and y_{jk}^{pred} utilizing various similarity distance metrics. These measures approach zero when the predictive algorithms achieve a superior fit to the data. Each metric has distinct statistical properties that complement one another in monitoring predictive performance. RMSE penalizes larger errors more heavily due to the squaring of errors, making it particularly suitable for applications where significant deviations are highly undesirable. This property aligns with minimizing squared errors in regression, where the least-squares estimation method is widely utilized due to its optimality under normal error distribution assumptions. MAE computes the average of absolute errors, treating all discrepancies equally. While this makes MAE less sensitive to the magnitude of errors compared to RMSE, it is not inherently robust to outliers. Its effectiveness in a limited data environment depends on the data preparation phase, such as mitigating the impact of anomalies. MAPE provides the average percentage of relative error, making it easy to interpret and compares across difference scale.

In summary, these metrics offer complementary insights into algorithms predictive performance. RMSE emphasizes large deviations, MAE provides a straightforward average of absolute errors, and MAPE enables percentages-based error interpretation, making them collectively useful across a range of data scenarios, including limited data environment when supported by appropriate data preparation. Additionally, adjusted-R^2 was excluded from consideration in this study due to its tendency to indicate predictive algorithm adequacy based on sample data fit, which may conflict with other GoF measures. Following a linear progression through the CRISP-DM data science methodology, therefore the next step of this study was the deployment of the proposed modified stacked ensemble multivariable AI-based predictive algorithm rather than iterating previous phases and initiating new projects.

3.6 Deployment

The deployment phase represented the final phase in executing pre-defined tangible business and data mining objectives, aimed at improving decision-making and operational efficiency. This phase encompassed planning deployment, monitoring, and maintenance, producing a final report, and reviewing the research. Deployment efforts ranged from presenting a final report to stakeholders to potentially developing a technological prototype for commercialization. In this study, the deployment phases focused on achieving Technology Readiness Level 3 (TRL3) by planning deployment and producing a final report. The superior predictive algorithm was deployed to forecast the future 5-year annual electricity consumption for business and data mining purposes, and the proof-of-concept, as presented in Section 4, was documented in a final report.

To gather constructive feedback from experts in related fields, the predictive results were also submitted as a research manuscript to a relevant journal for peer review. Although this study remains at TRL3, with no immediate plans for monitoring or maintenance, it acknowledges the current reliance on historical datasets as a limitation. This approach reflects the focus on establishing foundational insights and validating the algorithms predictive capabilities in a controlled environment.

While the integration of real-time data streams from environmental and economic clustered determinants is beyond the scope of this study, it is recognized as a significant area for future research. Similarly, the findings lay the groundwork for potential real-world applications, such as embedding the predictive algorithm into operational energy management systems utilized by utility managers and energy policymakers. Future efforts could focus on building user-friendly dashboards, software platforms, or Internal of Things (IoT)-integrated tools to make predictions actionable for decision-makers.

Additionally, future efforts could involve testing the proposed predictive algorithm with key stakeholders, such as utility managers, energy regulators, and industrial users to gather invaluable feedback on its usability and functionality in practical decision-making contexts. These steps will ensure a gradual, cost-effective progression toward aligning the predictive algorithms refinement, ensuring its alignment with industry needs and standards. Together, these efforts will aid this study's progress to higher TRL levels and ultimately exhibit commercialized value.

4 Analysis Results

The section presented a proof-of-concept based on analysis results utilizing real-world datasets, which were fully analyzed utilizing Microsoft Excel and R for statistical analysis. This study noted that analysis results did not correspond to every single phase of CRISP-DM data science methodology due to their inappropriateness. However, this study presented them by combining the first (business understanding, data understanding, and data preparation) and the last (modeling, evaluation, and deployment) three phases as detailed in Sections 4.1-4.2, respectively.

4.1 Business Understanding, Data Understanding, and Data Preparation

Table 1 provided a numerical summary utilizing the first four L-moments to analyze electricity consumption across economic activity from 1980 to 2021, corresponding to the period for which datasets were available from reputable open-source databases. The analysis revealed that industrial economic activity exhibited the highest average electricity consumption, while transportation economic activity had the lowest. This variation reflected the dependency on usage demand in various economic activities and socioeconomic growth in Malaysia. Meanwhile, Table 2 focused on significant determinants selected utilizing AWSFS and parametric t-tests, clustered based on the principal pillars of SDGs, including economic, environmental, and social factors. Additionally, Table 2 included Shapiro-Wilk analysis to objectively assess the normal distribution of each variable, avoiding subjective evaluations such as LM_3 and LM_4. Due to the extensive number of variables considered, graphical or tabular representations of the correlogram and outlier analysis were omitted. However, correlations between these determinants and $\boldsymbol{Y}$ were investigated.

Determinants $\boldsymbol{X_{12}}, \boldsymbol{X_{18}}$, and $\boldsymbol{X_{21}}$ showed weak negative (-0.0196), weak negative ($r - 0.2611$), and weak positive ($r = 0.0014$) correlations with $\boldsymbol{Y}$, respectively. In contrast, determinants $\boldsymbol{X}_3(0.3321)$, and $\boldsymbol{X}_{11}(0.4407)$ displayed moderate positive correlations with $\boldsymbol{Y}$. The rest of the determinants exhibited strong correlations ($|r| > 0.5$) with $\boldsymbol{Y}$, as classified based on Chuan *et al.* [32]. Weakly and moderately correlated determinants with Y were retained during this stage, even if they did prove significant in the proposed AI-based predictive algorithm after adjusting for other determinants. Notably, the modeling analysis results indicated that determinants such as $\boldsymbol{X}_3, \boldsymbol{X}_{11}$, and $\boldsymbol{X}_{21}$ were significant for the superior AI-based predictive algorithm, as detailed in Section 4.2.

This study also noted that highly correlated pairs of determinants were not eliminated at this stage due to potential sub-optimal decisions. These pairs provided early alerts regarding multicollinearity. However, findings could alter with feature engineering and effect adjustments in the modeling phase. Therefore, the elimination of highly correlated pairs of determinants was not mandatory. Conversely, this study explored multicollinearity corresponding to the top-ranked AI-based predictive algorithms through diagnostics checks when necessary. For mild and extreme outliers detection, the analysis identified determinants with mild outliers, including

$\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_9, \boldsymbol{X}_{11}$, and $\boldsymbol{X}_{19}$ while $\boldsymbol{X}_{11}$ exhibited extreme outliers. However, no corrections were made for extreme outliers utilizing methods such as capping, as extreme outliers may not necessarily represent influential observations, potentially impacting parameter estimation accuracy. Future work proposed investigating superior methods beyond Cooks Distance to identify influential observations fitting the acquired datasets.

Table 1: Numerical Summary Utilizing the First Four L-moments Across Economic Activity (1980-2021).

Economic Activity	Numerical Summary			
	$\mathbf{LM}_1$	$\mathbf{LM}_2$	$\mathbf{LM}_3$	$\mathbf{LM}_4$
Industrial*	2884.9048	0.4056	0.1219	0.0006
Commercial*	1857.6667	0.3893	0.1426	-0.0718
Residential*	1182.5476	0.4882	0.1304	-0.0257
Agriculture*	13.6905	0.7359	0.4841	0.0714
Transport*	10.1191	0.6704	0.3968	0.0454
Total*	5948.9286	0.4168	0.1300	-0.0377

*Economy activity is not statistically normal distributed

Table 1 provided comprehensive insights into the dataset, indicating that industrial and transport economic activities had the highest and lowest average electricity consumption (LM_1) from 1980 to 2021. Since the 1970s, the government shifted focus to promoting export-oriented industries, leading to significant changes in economic policies from inward-oriented trade to outward-oriented industrialization. This transformation profoundly impacted the labor market structure, with the commercial (wholesale and retail trade, accommodation, food, and beverage) and industry (manufacturing) sectors experiencing rapid expansion in the 1980s and becoming the largest contributors to national GDP. Statistical analysis of GDP at current prices (RM million) from 1987 to 2020 across economic activities such as agriculture, construction, industrial, livestock, forestry and fishing, mining & querying, and transport supported these findings [33].

In contrast, transportation economic activity indicated the lowest average electricity consumption from 1980 to 2021. Policymakers launched the New Industrial Master Plan 2030 (NIMP 2030), complementing existing initiatives for Electric Vehicle (EV) development under the National Automotive Policy (NAP), and budget incentives in 2022 and 2023 to boost demand and charging infrastructure. However, the adoption of EVs in Malaysia remained low due to market price disparities with Petrol Vehicles (PV), battery capacity and cost, charging behavior, and compatibility, charging station availability, charging times, grid capability, limited traveling distance, and barriers related to renewable energy, and climate mitigation efforts [34–37]. This was further supported by EV sales in Malaysia, where EVs accounted for merely 1.27% (10159 out of 799731 units) [38] of total vehicle sales in 2023, despite increasing demand in 2022. Electrified vehicles (EV and hybrid) accounted for 4.78% (38214 out of 799731 units) in 2023, highlighting low electricity consumption in transportation compared to other economic activities in Table 1.

Moreover, Table 1 revealed variation (LM_2) across economic activity due to dependency on usage demand and socioeconomic growth in Malaysia [39]. While LM_3 and LM_4 could reflect the data distribution shape, their subjective justification could provide sub-optimal insights. To address this limitation, this study utilized the Shapiro-Wilk test to assess normal distribution, confirming non-normal distribution across electricity consumption datasets, including $\boldsymbol{Y}$ (total).

Table 2: Key Determinants of MLR and MNLR Predictive Algorithms: Insights into SDG Pillars Across TTR

TTR	Economic																		Environmental						Social	
	X_4^*	X_5^*	X_6^*	X_7^*	X_8	X_9^*	X_{10}^*	X_{11}^*	X_{12}	X_{13}^*	X_{14}^*	X_{15}^*	X_{17}^*	X_{18}^*	X_{19}	X_{20}^*	X_{21}^*	X_{22}^*	X_1	X_2	X_3	X_{23}^*	X_{24}^*	X_{25}^*	X_{16}^*	X_{26}^*
60:40																										
70:30		√							√	√		√	√	√		√		√	√	√	√					
80:20				√								√	√			√	√	√	√	√	√	√				√
90:10	√	√	√		√	√		√				√					√	√	√	√	√					
60:40																										
70:30	‡	‡			‡				‡						‡							‡			‡	
80:20	‡			‡					‡			‡						‡								
90:10	‡		‡											‡					‡							‡

Table 2 extended the analysis to all determinants, revealing normal distribution exceptions for $\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3, \boldsymbol{X}_8, \boldsymbol{X}_{12}$, and $\boldsymbol{X}_{19}$. Despite non-normal distributions, the proposed AI-based predictive algorithms adhered to OLS normality assumptions. Thorough diagnostic checks of residuals ensured predictive algorithm reliability and robustness. Additionally, Table 2 detailed significant determinants for SLR-MLR and SLR-MNLR predictive algorithms across different TTR, selected via AWSFS and parametric t-tests. Significant determinants varied across TTR, highlighting their impact on AI-based predictive algorithms. Section 4.2 discussed the selected determinants significance, while superior modified stacked ensemble AI-based predictive algorithms were evaluated beyond this stage.

4.2 Modeling, Evaluation, and Deployment

In Section 3.3, this article discussed that feature engineering such as determinant selection was associated with training the proposed modified stacked ensemble multivariable AI-based predictive algorithms. To authenticate the effectiveness and select the superior proposed AI-based predictive algorithm, this article also employed the superior univariate SLR-ϵ-SVR predictive algorithm proposed by Chuan *et al.* [23]. This predictive algorithm was selected as a benchmark after comparing its effectiveness with other SVR-based predictive algorithms proposed in the literature. The principal reason for employing the SLR-ϵ-SVR predictive algorithm as the benchmark was its close resemblance to the proposed AI-based predictive algorithm. In contrast, classical statistical-based time-series predictive algorithms exhibited different approaches compared to machine learning, and ANN-based predictive algorithms posed limitations as highlighted in Section 3.4 including barriers to forecasting and interpretability. The lack of forecasting and interpretability did not fulfill the data science scope of this study, making classical statistical-based time-series and ANN-based predictive algorithms inappropriate as benchmarks.

Table 3: Internal and Predictive Performance Evaluation Utilizing GoF Measure Across TTR

Predictive Algorithm	TTR	GoF Measures (Training)			GoF Measures (Test)			MAD	Ranked
		RMSE	MAE	MAPE	RMSE	MAE	MAPE		
SLR-ε-SVR	60:40	172.1998	138.3402	8.9769	2257.0960	1714.2549	13.9717	1221.9352	5
SLR-MLR									
SLR-MNLR									
SLR-ε-SVR	70:30	184.0239	157.1767	8.8890	3137.5328	2683.4854	21.5546	1830.8277	7
SLR-MLR		86.9543	67.6214	3.7951	2276.8461	2152.4342	19.0706	1429.9934	6
SLR-MNLR		146.3616	101.4885	3.8625	5914.6354	5434.5076	45.2467	3714.2257	9
SLR-ε-SVR	**80:20**	**235.0790**	**197.4922**	**10.5379**	**610.1665**	**515.4531**	**4.0686**	**228.8597**	**2**
SLR-MLR		92.4446	80.8861	3.8341	5302.7250	4949.5399	38.5567	3371.2189	8
SLR-MNLR		339.1590	229.7977	5.5289	107058.9846	93864.5914	726.0674	67025.0526	10
SLR-ε-SVR	**90:10**	**257.2080**	**212.2115**	**10.4429**	**787.9359**	**609.0078**	**4.6016**	**307.2276**	**3**
SLR-MLR		**114.2531**	**87.1801**	**4.0915**	**494.1536**	**360.3756**	**2.7213**	**217.2419**	**1**
SLR-MNLR		195.6275	146.5494	4.1267	1341.6621	1170.2705	8.8410	724.8233	4

The prediction evaluation results among SLR-ϵ-SVR, SLR-MLR, and SLR-MNLR across different TTR were presented in Table 3. These results included both internal and hold-out cross-validation analyses, utilizing selected GoF measures. Its important to note that the analysis results for the SLR-ϵ-SVR predictive algorithm were slightly different from those presented in Chuan *et al.* [23], primarily due to differences in timescale coverage. Their analysis covered the period from 1978 to 2020, whereas this study maintained consistency across variables with a timescale spanning from 1980 to 2021. Additionally, GoF measures of SLR-MLR and SLR-MLNR at 60:40 TTR were not available in Table 3. This was due to these predictive algorithms being incompetent to train in the scenario with an n-to-number of determinants ratio (26:26) approaching one, a considerably high-dimensional dataset. To address this challenge, one potential solution was applying dimensional reduction or a Bayesian regression-based predictive algorithm, which was beyond the scope of this article and could lead to biased decision-making when compared with other predictive algorithms in Table 3 under different environments.

Based on these GoF measures between the split training and test sets presented in Table 3, statistical hypothesis testing based on WSR verified that there was no statistical significance of the GoF measures across all predictive algorithms corresponding to each TTR. This result was further strengthened by utilizing a parametric bootstrap paired t-test with 10000 bootstrap samples, which yielded similar analysis results. In simple terms, these inferential results indicated the absence of overfitting for these predictive algorithms. Due to the difficulty in identifying the superior predictive algorithm, this study proposed a novel ranking of the superiority of these predictive algorithms by taking the absolute difference average of the GoF measures between training and test sets (MAD) as depicted in Table 3. Particularly, ranking based on MAD was implemented by ordering the predictive algorithms in ascending order of their MAD values, with the smallest MAD indicating superior performance due to minimal deviation between the training and test sets. This approach was rational as there was no conflict among the GoF measures, and better predictive performance was indicated when these GoF measures approached zero, addressing the limitations of decision-making approaches in literature as highlighted in Section 2.

Consequently, the three top-ranked predictive algorithms (SLR-MLR with 90:10 TTR, SLR-ϵ-SVR with 80:20 TTR, and SLR-ϵ-SVR with 90:10 TTR) were employed in forecasting 5-year future electricity consumption as presented in Figure 2. This analysis revealed that the top-ranked predictive algorithms frequently required a high percentage of the training set, which was consistent with machine learning theory as this study involved a limited sample size of 42. To ensure the validity, diagnostic analyses were conducted to verify the key pre-defined assumptions of the SLR-MLR predictive algorithm. These include the Shapiro-Wilks normality test for normality assumption (p-value=0.4760), the Breusch-Pagan test (p-value=0.9461) for homoscedasticity assumption, and the Run test (p-value=0.0704) for independence assumption. These analyses verified the adequacy of the SLR-MLR predictive algorithm.

These results support the adequacy of the SLR-MLR predictive algorithm. However, multicollinearity, assessed via VIF, was present and could inflate coefficient standard errors. While this could impact coefficient interpretation, the robustness of predictions remained unimpacted, as this study was identifying significant determinants rather than interpreting the estimated regression coefficients of the SLR-MLR predictive algorithm. Furthermore, the small sample size may obscure the practical effects of multicollinearity [40]. For both SLR-ϵ-SVR predictive algorithms, which are semi-parametric predictive algorithms, the pre-defined assumptions were not explicitly verified. This selection reflects the flexibility of these predictive algorithms, which do not rely on strict parametric assumptions. To mitigate multicollinearity challenges, future studies could employ methods such as Ridge regression, Principal Component Analysis (PCA), or the Bayesian methods.

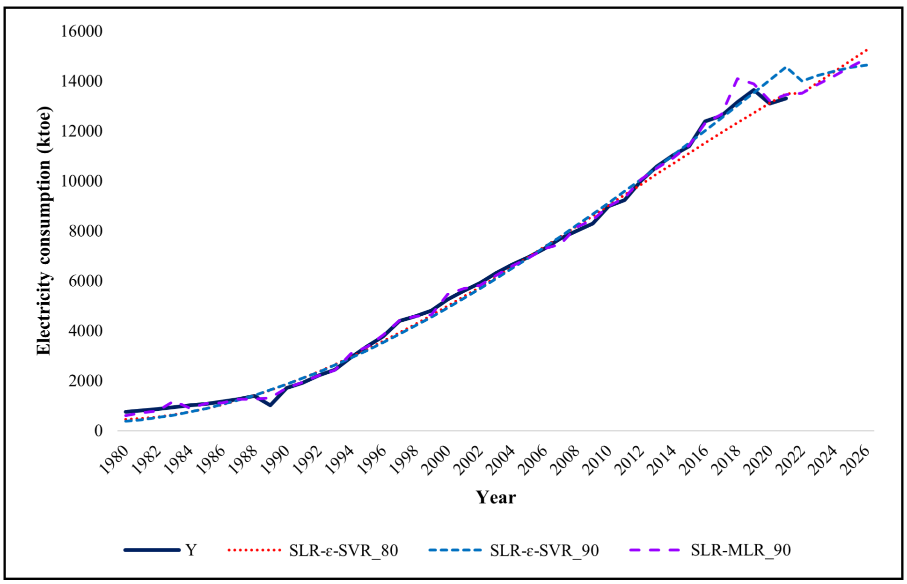

Figure 2: Vibration Signal of Square-shaped Stainless Steel Material at the Impact Force 459N.

Furthermore, the short-term forecasting of electricity consumption for the 5-year future was reasonable as Figure 2 revealed a continual increase in electricity consumption in Malaysia. This could be explained by global climate change, the expected continual growth of the Malaysian population, and accumulated technological advancements due to the arrival of the Industrial Revolution 4.0 (IR4.0), and several collaborative initiatives by policymakers in attracting potential Domestic Direct Investment (DDI), and Foreign Direct Investment (FDI), especially in the manufacturing, services and primary economic sectors [23,41,42]. This fact was further supported by the significant determinants included in the top-ranked AI-based predictive algorithm resulting from this study. These determinants included $\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3, \boldsymbol{X}_4, \boldsymbol{X}_5, \boldsymbol{X}_6, \boldsymbol{X}_8, \boldsymbol{X}_9, \boldsymbol{X}_{11}$, $\boldsymbol{X}_{15}, \boldsymbol{X}_{21}$, and $\boldsymbol{X}_{22}$, belonging to the economic and environmental clusters. In summary, the superior proposed modified stacked ensemble multivariable AI-based predictive algorithm in this study proved invaluable for short-term forecasting and identifying significant determinants impacting annual electricity consumption in Malaysia. Its applications spanned economic management, electricity security analysis, facility scheduling for electricity generation and distribution, and maintenance planning.

5 Conclusions and Recommendations

This study aimed to propose two innovative modified stacked ensemble multivariable AI-based predictive algorithms (SLR-MLR and SLR-MNLR) for annual electricity consumption in Malaysia, following the CRISP-DM data science methodology. To authenticate the effectiveness of these proposed AI-based predictive algorithms, the previously proposed modified stacked ensemble univariate AI-based (SLR-SVR) predictive algorithm was employed as a benchmark comparison. It is noteworthy that classical statistical predictive algorithms were not considered benchmarks due to their univariate nature, which limits interpretability. Similarly, ANN-based predictive algorithms for electricity-relevant predictions in the limited Malaysian literature were not utilized as benchmarks, as they were merely limited to predictive performance evaluation and were inadequate for forecasting.

To pursue this principal objective, open-source datasets were acquired from three well-known databases, including CCKP, MEIH, and OWID in CSV format. These time-series datasets underwent the necessary preprocessing before being integrated and split into training and test sets utilizing pre-defined TTRs. Specifically, the TTRs considered in this study included 60:40, 70:30, 80:20, and 90:10. Consequently, a total of ten AI-based predictive algorithms, comprising SLR-ϵ-SVR, SLR-MLR, and SLR-MNLR were trained across these TTRs. The analysis results revealed that SLR-MLR with a TTR of 90:10 outperformed SLR-ϵ-SVR, and SLR-MNLR across all TTRs, as ranked by the lowest MAD between the average of GoF measures of training and test sets. Additionally, the analysis identified that economic (coal and coke supply ($\boldsymbol{X}_4$), crude oil supply ($\boldsymbol{X}_5$), natural gas supply ($\boldsymbol{X}_6$), crude oil import ($\boldsymbol{X}_8$), diesel import ($\boldsymbol{X}_9$), coal export ($\boldsymbol{X}_{11}$), Gross GDP per capita at current prices ($\boldsymbol{X}_{15}$), share of manufacturing in GDP ($\boldsymbol{X}_{21}$), world crude oil ($\boldsymbol{X}_{22}$)), and environmental (average minimum surface air temperature ($\boldsymbol{X}_1$), average maximum surface air temperature ($\boldsymbol{X}_2$), and annual precipitation ($\boldsymbol{X}_3$)) clustered determinants significantly impacted electricity consumption in Malaysia, unlike the social clustered determinants.

In summary, this article presented dual contributions aimed at both academic and industrial applications. Academically, this study proposed the SLR-MLR predictive algorithm, enhancing predictive efficiency and interpretability from a practical perspective. This study also proposed a systematic approach to evaluate and mitigate overfitting in supervised learning scenarios utilizing time-series datasets. Industrially, this study introduced a practical multivariable AI-based predictive algorithm designed for computational efficiency. This predictive algorithm proved invaluable for short-term forecasting and identifying significant determinants impacting annual electricity consumption in Malaysia. Its applications include economic management, electricity security analysis, facility scheduling for electricity generation and distribution, and maintenance planning. These practical benefits align with Malaysias economic and environmental sustainability goals as outlined in the Twelfth Malaysia Plan, the Madani Economy Framework, the National Energy Policy 2022-2040, and the National Energy Transition Roadmap (NETR) agenda.

Future research directions include refining the predictive algorithm by investigating extreme outliers that may influence parameter estimation accuracy and integrating the three top-ranked modified stacked ensemble AI-based predictive algorithms for improved robustness beyond pre-defined OLS assumptions. The inclusion of carbon tax determinants, expected to be implemented in 2026, is also recommended to enhance the predictive algorithms relevance for sustainable energy planning. A key limitation of this study lies in its reliance on historical datasets, which underscores the need to transition to real-time data streams from environmental and economic clustered determinants to improve forecasting accuracy and applicability.

To advance real-world applications, this study recommends embedding the predictive algorithm into operational energy management systems utilized by utility managers and energy policymakers. Additionally, user-friendly tools such as dashboards, software platforms, or Internal of Things (IoT)-integrated systems should also involve testing the proposed predictive algorithm with key stakeholders, including utility managers, energy regulators, and industrial users, to refine its functionality and usability in practical contexts. By addressing these areas, this study aims to progress from its current TRL3 to higher TRLs, ensuring alignment with industry needs and fostering commercial value.

Acknowledgments

The authors sincerely appreciate the Ministry of Higher Education for financial support under the Fundamental Research Grant Scheme (Grant No. FRGS/1/2024/SS01/UTEM/02/2). They also thank the Climate Change Knowledge Portal (CCKP), Malaysia Energy Information Hub (MEIH), and Our World in Data (OWID) for providing the open-source datasets utilized in this study.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the authors used ChatGPT to improve the readability and language of this work. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the publication.

References

[1] Noh, A. M., Mat, S. and Ruslan, M. H. Development and performance analysis of new solar dryer with continuous and intermittent ventilation. *Jurnal Kejuruteraan*. 2018. 1(3): 1-8.

[2] Azni, M. A., Khalid, R. M., Hasran, U. A. and Kamarudin, S. K. Review of the effects of fossil fuels and the need for a hydrogen fuel cell policy in Malaysia. *Sustainability*. 2023. 15(5): 4033.

[3] Economic Planning Unit. *National Energy Policy*. Putrajaya: Economic Planning Unit Prime Ministers Department. 2022.

[4] Ministry of Economy. *National Energy Transition Roadmap: Energising the Nation, Powering Our Future*. Putrajaya: Ministry of Economy. 2023.

[5] Gasparatos, A., Doll, C. N. H., Esteban, M., Ahmed, A. and Olang, T. A. Renewable energy and biodiversity: implications for transitioning to a Green Economy. *Renewable and Sustainable Energy Reviews*. 2017. 70: 161-184.

[6] Osman, A. I., Chen, L., Yang, M., Msigwa, G., Farghali, M., Fawzy, S., Rooney, D. W. and Yap, P. -S. Cost, environmental impact, and resilience of renewable energy under a changing climate: a review. *Environmental Chemistry Letters*. 2023. 21: 741-764.

[7] International Energy Agency. (n.d.) *Source of Electricity Generation (Malaysia)*. Retrieved from `https://www.iea.org/countries/malaysia/electricity`.

[8] Saglam, M., Spataru, C. and Karaman, O. A. Electricity demand forecasting with use of artificial intelligence: the case of Gokceada Island. *Energies*. 2022. 15(16): 5950.

[9] Jnr, E. O. and Ziggah, Y. Y. Electricity demand forecasting based on feature extraction and optimized backpropagation neural network. *E-Prime-Advances in Electrical Engineering, Electronics and Energy*. 2023. 6: 100293.

[10] Liu, Y. and Li, J. Annual electricity and energy consumption forecasting for the UK based on back propagation neural network, multiple linear regression, and least square support vector machine. *Processes*. 2023. 11(1): 44.

[11] Saglam, M., Spataru, C. and Karaman, O. A. 2023. Forecasting electricity demand in Turkey using optimization and machine learning algorithms. *Energies*. 2023. 16(11): 4499.

[12] Varelas, G., Tzimas, G. and Alefragis, P. A new approach in forecasting Greek electricity demand: from high dimensional hourly series to univariate series transformation. *The Electricity Journal*. 2023. 36(6): 107305.

[13] Wang, Z., Chen, Z., Yang, Y., Liu, C., Li, X. and Wu, J. A hybrid Autoformer framework for electricity demand forecasting. *Energy Reports*. 2023. 9: 3800-3812.

[14] Shapi, M. K. M., Ramli, N. A. and Awalin, L. J. Energy consumption prediction by using machine learning for smart building: case study in Malaysia. *Development in the Built Environment*. 2021. 5: 100037.

[15] Lee, M. H. L., Ser, Y. C., Selvachandran, G., Thong, P. H., Cuong, L., Son, L. H., Tuan, N. T. and Gerogiannis, V. C. A comparative study of forecasting electricity consumption using machine learning models. *Mathematics*. 2022. 10(8): 1329.

[16] Razak, F. A., Shitan, M., Hashim, A. H. and Abidin, I. Z. Load forecasting using time series models. *Jurnal Kejuruteraan*. 2009. 21(1): 53-62.

[17] Karim, S. A. A. and Alwi, S. A. Electricity load forecasting in UTP using moving averages and exponential smoothing techniques. *Applied Mathematical Sciences*. 2013. 7(80): 4003-4014.

[18] Ozoh, P., Abd-Rahman, S., Labadin, J. and Apperley, M. A comparative analysis of techniques for forecasting electricity consumption. *International Journal of Computer Applications*. 2014. 88(15): 8-12.

[19] Lee, Y. W., Tay, K. G. and Choy, Y. Y. Forecasting electricity consumption using time series model. *International Journal of Engineering and Technology*. 2018. 7(4.30): 218-223.

[20] Sim, S. E., Tay, K. G., Huong, A. and Tiong, W. K. Forecasting electricity consumption using SARIMA method in IBM SPSS software. *Universal Journal of Electrical and Electronic Engineering*. 2019. 6(5B): 103-114.

[21] Abdullah, L. and Leong, W. H. The relationship of economic variables and final energy consumption: multiple linear regression evidence. In *MATEC Web of Conferences, May 25-27*. Beijing, China: EDP Sciences: 2018. 189: 10025.

[22] Ahmad, N., Othman, J., Aziz, M. A. S. A., Sivaraju, S. S., Borhan, N., Mohtar, W. A. A. I. W. and Abdullah, M. F. Load profile forecasting using a time series model for solar rooftop and integrated carpark of a public university in Malaysia. *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences*. 2023. 111(2): 86-98.

[23] Chuan, Z. L., Jie, O. S., Hin, T. Y., Zain, S. N. S. B. M., Rashid, Y. A. B. A. and Kamarudin, A. N. B. AI-driven electricity consumption forecasting for sustainable energy policy: an SVR time-series algorithm in sparse data environments. *SSRN*: 2023.

[24] Mustapa, R. F., Dahlan, N., Yassin, I. M., Nordin, A. H. M. and Zabidi, A. Energy consumption prediction through linear and non-linear baseline energy model. *Indonesian Journal of Electrical Engineering and Computer Science*. 2020. 17(1): 102-109.

[25] Ayodele, B. V., Mustapa, S. I., Mohammad, N. and Shakeri, M. Long-term energy demand in Malaysia as a function of energy supply: a comparative analysis of Non-Linear Autoregressive Exogenous Neural Networks and Multiple Non-Linear Regression models. *Energy Strategy Reviews*. 2021. 38: 100750.

[26] Sena, B., Zaki, S. A., Rijal, H. B., Ardila-Rey, J. A., Yusoff, N. M., Yakub, F., Liana, F. and Hassan, M. Z. Development of an electrical energy consumption model for Malaysian households, based on techno-socioeconomic determinant factors. *Sustainability*. 2021. 13(23): 13258.

[27] Chuan, Z. L., Wei, D. C. T., Aminuddin, A. S. B. A., Fam, S. -F. and Ken, T. L. Comparison of multiple linear regression and multiple nonlinear regression models for predicting rice production. In *AIP Conference Proceedings, May 16-17*. Bangi, Malaysia: AIP Publishing. in press.

[28] Solano, J. A., Cuesta, D. J. L., Ibez, S. F. U. and Coronado-Hernndez, J. R. Predictive models assessment based on CRISP-DM methodology for students performance in Colombia-Saber 11 test. In *Procedia Computer Science, November 1-4*. Leuven, Belgium: Elsevier. 2022. 198: 512-517.

[29] Chuan, Z. L., Japashov, N., Yuan, S. K., Qing, T. W. and Ismail, N. Analyzing enrolment patterns: modified stacked ensemble statistical learning-based approach to educational decision-making. *Akademika*. 2024. 94(2):232-251.

[30] Lohaj, O., Parali, J., Bednr, P., Paraliov, Z. and Huba, M. Unraveling COVID-19 dynamics via machine learning and XAI: investigating variant influence and prognostic classification. *Machine Learning & Knowledge Extraction*. 2023. 5(4):1266-1281.

[31] Brzozowska, J., Pizo, J., Baytikenova, G., Gola, A., Zakimova, A. and Piotrowska, K. Data engineering in CRISP-DM process production data-case study. *Applied Computer Science*. 2023. 19(3): 83-95.

[32] Chuan, Z. L., Wei, D. C. T., Yan, C. L. W., Nasser, M. F. A., Ghani, N. A. M., Jemain, A. A. and Liong, C. -Y. A comparative of two-dimensional statistical moment invariants features in formulating an automated probabilistic machine learning identification algorithm for forensic application. *Malaysian Journal of Fundamental and Applied Sciences*. 2023. 19(4): 525-538.

[33] Department of Statistics Malaysia. (2022) Malaysia Economic Statistics-Time Series 2019 (National Accounts). Retrieved from https://www.dosm.gov.my/portal-main/time-series.

[34] Glandorf, J. (2020). On the move: unpacking the challenges and opportunities of electric vehicles. Retrieved from https://www.eesi.org/articles/view/on-the-move-unpacking-the-challenges-and-opportunities-of-electric-vehicles

[35] Muzir, N. A. Q., Mojumder, M. R. H., Hasanuzzaman, M. and Selvaraj, J. Challenges of electric vehicles and their prospects in Malaysia: a comprehensive review. *Sustainability*. 2022. 14(14), 8320.

[36] Alanazi, F. Electric vehicles: benefits, challenges, and potential solutions for widespread adaptation. *Applied Sciences*. 2023. 13(10), 6016.

[37] Yean, T. S. and Rebecca, N. H. Y. (2024) A study of the emerging electric vehicle (EV) supply chain in Malaysia. Retrieved from https://www.iseas.edu.sg/articles-commentaries/iseas-perspective/2024-33-a-study-of-the-emerging-electric-vehicle-ev-supply-chain-in-malaysia-by-tham-siew-yean-and-neo-hui-yun-rebecca/

[38] Chan, A. (2024) MAA: 10,159 EVs were sold in Malaysia in 2023, almost quadruple from 2022. Retrieved from https://autobuzz.my/2024/01/16/maa-10159-evs-sold-in-malaysia-in-2023-almost-quadruple-from-2022/#:~:text=The%20number%20of%20EVs%20sold,the%20ones%20recorded%20by%20MAA.

[39] Pei, T. L., Shaari, M. S. and Ahmad T. S. T. The effects of electricity consumption on agriculture, service and manufacturing sectors in Malaysia. *International Journal of Energy Economics and Policy*. 2016. 6(3), 401-407.

[40] Kroll, C. N. and Song, P. Impact of multicollinearity on small hydrologic regression models. *Water Resources Research*. 2013. 49(6), 3756-3769.

[41] Ahmad, F. (n.d.) Malaysia creates almost 90,000 jobs from RM225.0 billion approved investments for 9M2023, exceeding full-year annual target. Retrieved from https://www.mida.gov.my/media-release/malaysia-creates-almost-90000-jobs-from-rm225-0-billion-approved-investments-for-9m2023-exceeding-full-year-annual-target/

[42] Malaysia attracts RM76.1bil potential foreign investments as of March 2024-PM. (2024, March 21). *The Stars*. Retrieved from https://www.thestar.com.my/business/business-news/2024/03/21/malaysia-attracts-rm761bil-potential-foreign-investments-as-of-march-2024---pm