Improving the Decision Support System in Sustainability Management with the Machine Learning

Jia Ying Ng[1] & Kok Weng Tan[2*]

[1]Universiti Tunku Abdul Rahman
Faculty of Engineering and Green Technology
31900 Kampar Perak Malaysia
Email: jiaying01911@gmail.com
[2]Universiti Tunku Abdul Rahman
Faculty of Engineering and Green Technology
31900 Kampar Perak Malaysia
Email: tankokweng@utar.edu.my

ABSTRACT

This study presents initial results on the integration of machine learning into a sustainability decision support system (MLDSS) designed to enhance organizational sustainability performance management. The proposed framework combines rule-based compliance assessment, ESG scoring, progress measurement, and strategic recommendations to support companies in aligning sustainability goals with operational priorities. A random forest regressor was employed to predict ESG scores from sustainability data extracted from corporate reports, focusing on greenhouse gas emission indicators. Experiments conducted with datasets from 5, 10, and 15 companies demonstrated that model performance improved with larger sample sizes, as reflected in rising R^2 (0.64 - 0.86) and explained variance scores, alongside reduced error margins. Meanwhile, the RMSE between the actual and model output for the GHGs Emission Score is reduced from 18.58 to 14.6 .The smaller datasets revealed limitations due to imbalance and variability, results from the 15-companies sample size indicated promising predictive capability and underscored the importance of continuous data enrichment. Beyond technical performance, the system's visualization and tracking tools facilitate transparent stakeholder communication, reinforcing consumer trust and investor confidence. These findings highlight the potential of machine learning–assisted DSS to automate compliance, optimize sustainability strategies, and strengthen ESG reporting, offering a scalable pathway for organizations to advance sustainable practices.
Keywords: Sustainability management, ESG, machine learning, decision making

ACCEPTED MANUSCRIPT

1. Introduction

Sustainability performance management has become increasingly critical for businesses worldwide, including in Malaysia. Aligning with sustainability agendas is vital for long-term business survival, influencing financial prospects, corporate image, and customer trust. By early 2020, global investments in sustainable projects had reached $35.3 trillion (Global Sustainable Investment Alliance, 2021), while the sustainable economic transition in Southeast Asia is expected to unlock $1 trillion in economic opportunities. Anthropogenic climate change has negatively impacted business operations through financial losses, supply chain disruptions, and physical risks (Kalogiannidis *et al*., 2024). Embracing sustainability performance management is essential for businesses to thrive in a climate-affected future.

Sustainability performance management facilitates corporate sustainability by integrating Environmental, Social, and Governance (ESG) principles into business strategies. Corporate sustainability, rooted in the triple bottom line and stakeholder theory, seeks to balance financial profitability with economic, social, and environmental responsibilities through leadership and management practices (Alhaddi, 2015). ESG aligns with this objective by guiding capital investments toward companies that demonstrate environmental stewardship, social responsibility, and strong governance. The relationship between ESG factors and company financial performance remains complex and debated in academic literature. Several studies indicate a positive correlation between ESG implementation and financial value creation (Castilla‐Polo & Guerrero‐Baena, 2023; Matsumura, Prakash and Vera-Muñoz, 2014; Oncioiu et al., 2020; Segura et al., 2024; Thayaraj & Karunarathne, 2021). However, Chen, Hung and Wang (2018) highlighted potential drawbacks, suggesting that ESG investments may reduce short-term profitability. Beyond profitability, ESG initiatives contribute to risk mitigation strategies, enhancing corporate financial stability (Kim, Lee and Kang, 2021). ESG-driven measures support stock price stability (Flammer, 2013) and strengthen corporate reputation (Koh et al., 2022), reinforcing long-term resilience. ESG initiatives extend beyond financial and risk considerations, serving as key drivers within sustainability performance management. As businesses integrate ESG principles into their strategies, sustainability performance management provides the essential framework for measuring, refining, and optimizing these efforts.

Sustainability performance management typically follows an iterative process involving goal setting, action planning, progress tracking, reporting achievements, and refining strategies to achieve better outcomes. A critical aspect of this process is sustainability or Environment, Social, and Governance (ESG) reporting, which shares a company’s initiatives and progress toward sustainable practices. ESG reporting has gained traction as a key driver for attracting investors, improving transparency, and fostering consumer trust. Despite the global emphasis on sustainability, Malaysia’s industries have shown limited progress. A report from Malaysia’s Ministry of Investment, Trade, and Industry highlights significant challenges, including a lack of expertise and resources in private companies, leaving many unaware of their sustainability standing. The unavailability of accessible assessment tools further hinders the adoption of sustainable practices (Ministry of Investment, 2023). The i-ESG framework offers a foundational approach for integrating sustainability principles into private companies. However, challenges such as compliance with reporting standards and the effective interpretation of results remain obstacles to widespread adoption.

There is a wide usage of machine learning (ML) models in research on sustainability performance management. Natural language processing was used by Fischbach *et al.* (2023) and Kang and Kim (2022) in text mining task on data of social media sourced and corporate sustainability report. Modapothala *et al.* (2010) conducted the analysis on the reporting variables (organizational, environmental, social and economic performance) selected by different industries using One-Way ANOVA and Multivariate Discriminant Analysis. The Classification and

ACCEPTED MANUSCRIPT

Regression Tree (CART) and Adaptive Neuro-Fuzzy Inference System (ANFIS) were applied to assess country sustainability through large number of indicators set (Nilashi *et al.,* 2019). Shahi et al. (2012) used Naïve Bayes, Neural Network, C4.5 and Decision Table to conduct the automated scoring of Corporate Sustainability Reports following Global Reporting Initiative (GRI) standards. Vivas *et al.* (2019) utilized Multiple Linear Regression (MLR), a generalized linear regression algorithm, in the development of a hybrid multi-criteria decision analysis (MCDA) model that assesses sustainability performance to carry out prediction task. Laskar (2018) used Logistic Regression to determine the effects of Corporate Social Reporting on Asian firm Value. However, the researchers highlighted the overconfidence of logistic regression model as one of the limitations of the study. Least-squares regression, panel data regression, and logistic regression were used by Wang (2017) to study the relationship between sustainability disclosure and firm characteristics of the Taiwan 50 Index-listed companies. Logistic regression was used again by Akbulut and Kaya (2019) to study the relationship between sustainability reporting, firm value and financial leverage in the automobile industry.

Chang *et al.* (2019) used multiple regression model to investigate the factors affecting sustainability reporting quality in the financial sector and the impact of equator principle on moderation. Sariyer and Taşkın (2022) used Kmeans++ clustering algorithm to conduct cluster analysis on ESG score of companies included on the Borsa Istanbul (BIST) Sustainability Index for the identification of the relationship between ESG score and firm financial and ESG performance. Kanmani et al. (2020) constructed a framework for the assessment of environmental sustainability of countries utilizing Self-Organized Maps, a clustering technique, to form clusters of countries with similar environmental performance and to compare the countries in each cluster in different timeframe. Galindo *et al.* (2015) applied hierarchical clustering Ward's methods to form clusters of companies of alike sustainability profile to understand their contribution on their country's sustainability performance. Kmeans algorithm was used by Li and Rockinger (2024) to investigate the changes in bank sustainability reporting focus over the years.

The rapid advancements in machine learning (ML) technology have transformed data processing capabilities, enabling innovative solutions to pressing societal challenges, particularly those related to sustainability. The integration of machine learning with decision support systems (DSS) offers a transformative opportunity to facilitate sustainability transitions across industries by producing actionable insights for strategic decision-making. There is a critical need for a decision making system capable of evaluating a company's sustainability performance, analyzing comprehensive sustainability reports, and deriving actionable recommendations to optimize sustainability practices. A hybrid methodology combining supervised learning models with Integrated Rule Engines (IREs) provides a robust framework for sustainability-related decision-making. Supervised learning algorithms contribute by generating predictive outputs such as classifications or probability scores, while IREs apply deterministic, rule-based logic to ensure decisions remain transparent, consistent, and interpretable. Integrating an IRE into a machine learning system enhances its capacity for logical inference and enables automated decision execution through the application of predefined rules to real-time data streams. Within sustainability systems, this synergy facilitates reliable verification against established reporting standards, including the Global Reporting Initiative (GRI) and the Sustainability Accounting Standards Board (SASB), thereby reinforcing accountability and regulatory compliance.

ACCEPTED MANUSCRIPT

This study presents initial results on integrating machine learning into a sustainability decision support system (DSS). It proposes a framework that enables efficient evaluation and management of sustainability performance within organizations.

2. Methodology

2.1 Conceptual framework for the decision support system

The conceptual framework for the decision support system is developed through case studies of several research papers of decision support system development and applications of decision support systems in different industries. The process for the decision support system is adapted from Di Matteo *et al.* (2021) and shown below in **Figure 1**.

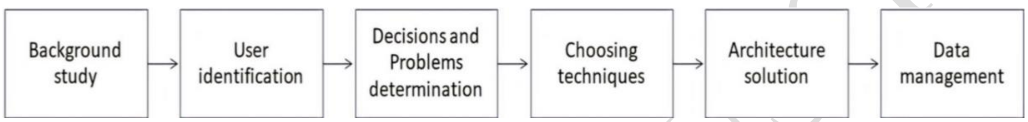


Figure 1: Development of Decision Support System

The desktop study was conducted to have an in-depth understanding of each component. There are mainly three components in a decision support system which are (i) data management system for the storage and handling of data necessary for the decision making process, (ii) the model management system which holds the machine learning model that enables the functioning of the system, (iii) knowledge management system which comprise of information collected from all sources including input from experts and excerpts of sustainability strategy from company sustainability reports with excellent sustainability score and the user interface that allows interactions between users and the systems.

2.2 Integration of decision support system with machine learning

This stage involved 3 main activities: data collection, model training and model performance assessment. The ESG scores for 15 electrical and electronic companies in Malaysia were extracted from LSEG ESG Score website (https://www.lseg.com/en) and recorded in an Excel database. The LSEG ESG Score was chosen because it is an easily accessible and its scoring categories can be traced back to sustainability indicator. The sampled companies were selected randomly and consisted of Malaysian and international companies of different sizes. The information extracted includes ESG scores and the corresponding year of data collection. The scoring categories in the LSEG Scoring System are shown in **Figure 2**.

The company sustainability reports of those companies were collected from the companies' website according to the year of data collection. The environmental data were extracted from the sustainability reports according to the indicators listed in the Global Reporting Initiative (GRI) Standards. The indicators considered were GRI 302, GRI 303, GRI 305, GRI 306. Three environmental indicators were excluded, namely GRI 301 Material, 304 Biodiversity and 308

ACCEPTED MANUSCRIPT

Supplier Environmental Assessment, as these data points were absent in the sustainability reports of many companies. The exclusion primarily aimed to minimize missing data, as missing values could negatively impact the prediction model's performance. Since the extracted data had varying units, unit conversion was performed to standardize them and prevent inconsistencies in the prediction results.

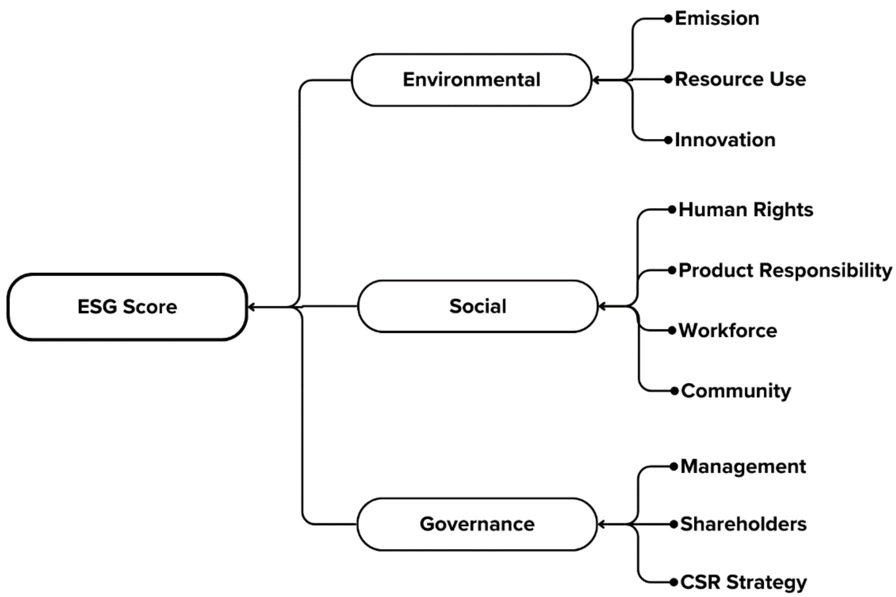


Figure 2: Structure of LSEG ESG Score

For this experiment, the datasets were segmented into three sample sets of sizes 5, 10, and 15 (ESG reporting companies). Due to the limited dataset sizes, the regression model's ability to produce accurate ESG score predictions is constrained. Nonetheless, the use of varying sample sizes allows for an evaluation on how data volume influences the Random Forest Regressor's performance in predicting ESG scores from sustainability report data. The dataset was partitioned into 60% training and 40% testing subsets using the *train_test_split* function from the *Scikit Learn* library to facilitate model evaluation.

i. Model Training

The training of the machine learning was conducted on Jupyter Notebook (https://jupyter.org/) using the Python programming language. The Python libraries such *as Pandas, Numpy, Matplotlib, Sickit-learn* were used for different purposes. The Pandas library provided data structures and functions to support data manipulation and analysis operations including data cleaning and exploration. The *Numpy* is a Python package for scientific computing that provided arrays, matrices and mathematical functions for multi-dimensional arrays like the datasets used in this study. Matplotlib can generate various types of charts such as line and scatter plots and bar charts for the visualization of data and results. All machine learning related tools and functions including the machine learning model were called from *Sickit-learn*, a machine learning library in Python.

Random Forest Regressor, a supervised learning algorithm, was used for the scoring model in this study. It is an ensemble learning model that produces prediction through combining the output of multiple decision trees in the regressor. The aggregation of multiple output from different decision trees allows random forest regressor to have the benefit of accurate and stable prediction. Random forest regressor also has the capability to quantify feature importance which is useful in determining indicators influential to sustainability performance.

The training of a random forest regressor consists of three steps: bootstrap sampling, decision tree building and prediction generation. Bootstrap sampling is the random selection of datapoints to create multiple subsets of the training data. A sample data can coexist in multiple sample subsets as it is randomly selected. This method is called bootstrapping. A decision tree is constructed for each sample subset, randomly selected some features to create split node in the tree as this can reduce the correlation in the decision trees. Therefore, each tree uses different sets of features to form the split node. The prediction made by each decision tree is aggregated to produce the final prediction of the random forest regressor.

The performance of the regression model is influenced by parameters such as *n_estimators* and *random_state*. The *n_estimators* parameter determines the number of trees in the random forest regressor, where a higher value typically improves performance by reducing variance. The random state parameter is related to the randomness of data sampling, and assigning it a fixed integer ensures reproducibility of the training results. This reproducibility is critical for conducting consistent and comparable experiments. For this study, the *n_estimators* parameter was set to 30, and the *random_state* value was fixed at 30 to maintain consistency across experiments

ii. Model Performance Assessment

The task of the regression model is to successfully map out a formula that can explain the relationship between the independent variables and dependent variables, or the input and output data. The performance of the regression model is evaluated with various parameters. This research uses the Coefficient of Determination R^2, Root Mean Squared Error (RMSE) and the Explained Variance Score (EVS).

The model testing was conducted through the prediction of ESG scores in the testing datasets of the sample. The testing data were selected randomly by the *training_testing_split* function. The comparison of predicted and the actual values of the ESG scores and the analysis of the evaluation metrics were performed to observe the performance of the regression model.

3. Results and Discussions

3.1 Proposed Integrated Framework of Machine Learning and Decision Support System

The proposed framework of the machine learning and decision support system (MLDSS) is displayed in **Figure 3**. The details of each component are discussed in subsequent sections.

ACCEPTED MANUSCRIPT

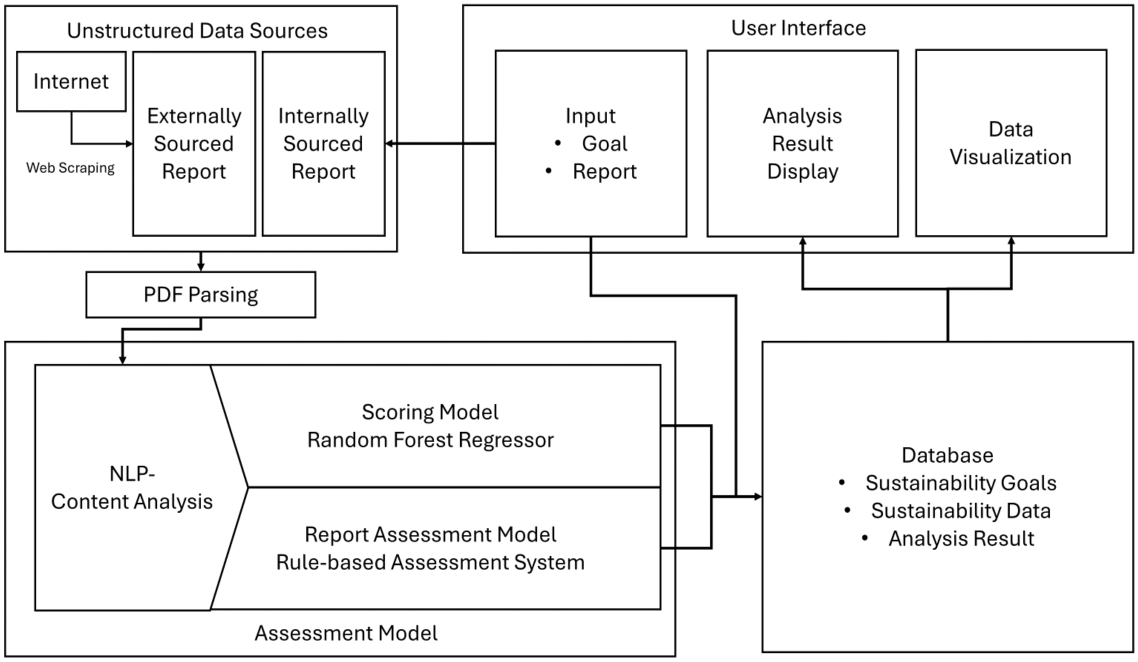


Figure 3: The Proposed Integration Framework of Machine Learning and Decision Support System (DSS)

The first component of the MLDSS is a data management system consisting of three subsystems: data collection, extraction, storage and processing systems. The data collection system utilizes search query command and web scraping tools to enable the automation of company sustainability report collection. The search query command searches the internet using keywords such as "sustainability report", "Electrical and Electronics Company", "ESG Data", etc. to collect the URL of target data into a list. The web scraping tool is used to extract the reports in Portable Document Format (PDF) and ESG data from the URL.

The portable document format (PDF) parsing is used to extract information from the company sustainability reports. As of current, most sustainability reports are only available in PDF documents that are not machine readable. PDF parsing allows computers to access the information in the sustainability reports through the extraction of text, tables and images from the PDF files. Besides, the PDF parsing techniques such as image extraction detect image objects like charts in the PDF file and save them in standard image formats like the Joint Photographic Experts Group (JPEG) and the Portable Network Graphics (PNG). The data in tables which are used extensively in sustainability reports can also be extracted using PDF parsing techniques. The PDF parser can identify the table structure in the PDF file and convert them into comma separated values (*.csv) files and Excel formats that are accessible to computers. There are several python libraries that can perform PDF parsing, including *PyPDF2* and *textract*. The extracted data are stored in a structured database for accessibility.

ACCEPTED MANUSCRIPT

Data processing is conducted using Natural Language Processing (NLP) techniques to understand the content of the extracted data and categorize them under the three pillars of ESG. NLP will also identify the reporting standards from the content of the report. Charts are widely used in the reports to display company ESG data, hence, computer vision is used to interpret and extract data points from the images. The system can be adapted to any reporting standards used by the companies. For the assessment of sustainability reports following the GRI standards, the data extracted will be categorized according to the indicators in the universal and topic standards.

The structure of the database is demonstrated in **Figure 4**. Each box represents a table in the database. The data included in the database are information of the company, their sustainability goals, information of the sustainability reports and the ESG data extracted from the sustainability reports and recommendations on sustainability strategy improvement.

CRIPT

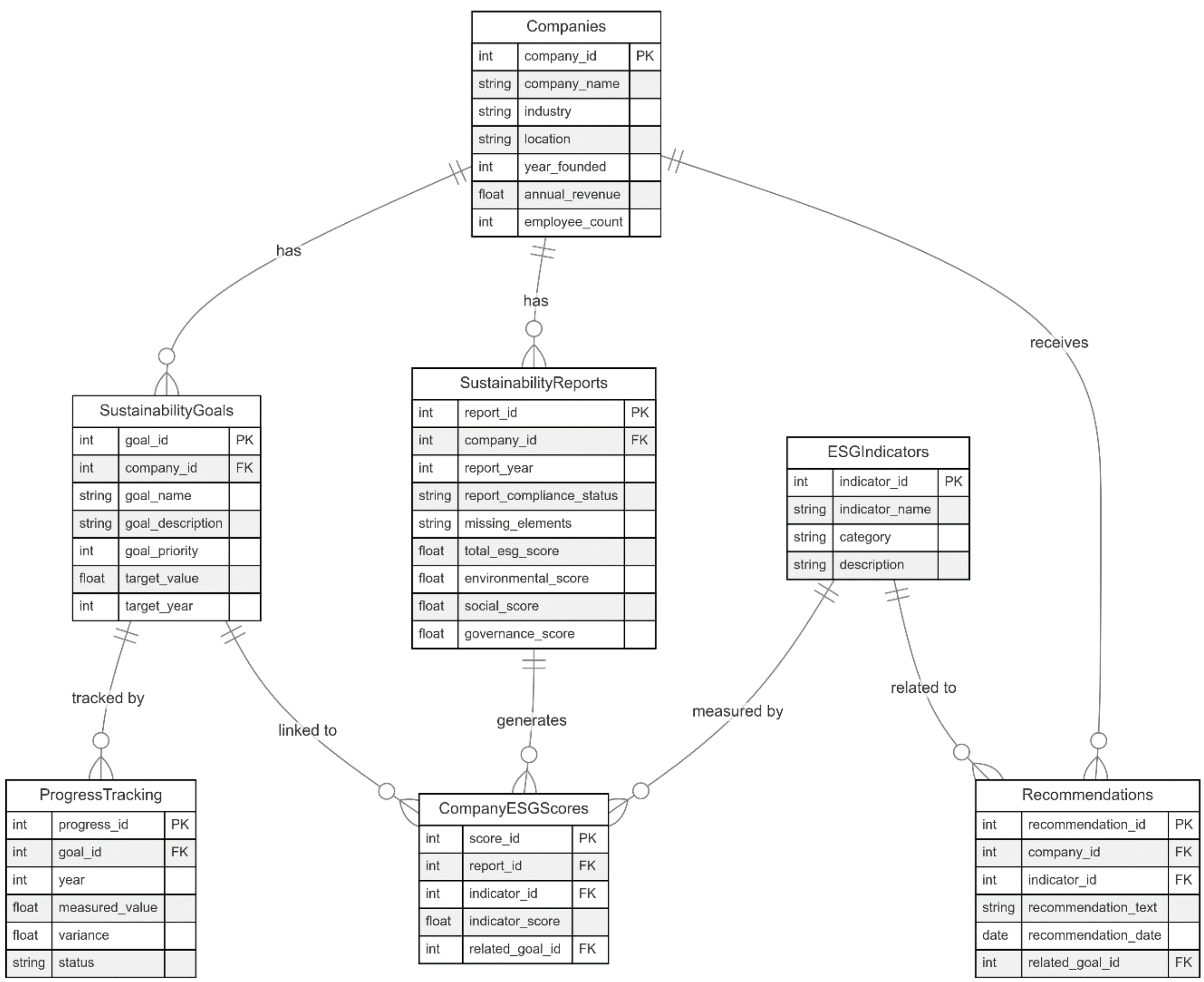


Figure 4: The Structure of ESG Database

The information of the company and sustainability goals are input by the users through the user interface. The data for the sustainability report and the ESG indicators are extracted through the data collection system. The data for progress tracking, company ESG score and recommendations are generated by the decision support system.

3.2 The MLDSS Model Management System

The model management system includes three functions, each serving a distinct role in assessing and managing sustainability performance namely i. conduct compliance assessment of sustainability report to its reporting standards, ii. perform analysis on corporate sustainability performance and iii. recommend sustainability strategy improvement. Compliances with sustainability reporting standards is rewarding on company financial performance through the increase of firm value (Moses, 2022; Sreepriya *et al.*, 2023). However, assessing the level of standard compliance requires some levels of expertise and familiarity with the sustainability reporting standards which is not easily accessible to most companies in Malaysia as sustainability reporting is still in its infancy in the country. The proposed solution to the first problem of standard compliance assessment is a system that utilizes both machine learning and rules to identify the required reporting matters in the report. This combination was proposed by Hamdani et al. (2021) to automate the assessment of General Data Protection Regulation (GDPR) compliance of company data privacy policies in Europe.

Although the classification task was not targeted towards sustainability reporting standards, they share some common attributes in term of requirement of mandatory information disclosure and having text-based data source that are not directly machine readable. Natural Language Processing (NLP) model were used as text classifier that assigns categories to the text segments to assist the rule-based approach on the checking of compliance requirement of GDPR. Transformer-based language models pretrained on database of the niche languages used in sustainability reporting can be applied for text classification. Transformers differentiate themselves from other language models through their ability to understand words in the context of their usage (Brugger *et al.*, 2023). Text-to-Text Transfer Transformer (T5 Transformer) demonstrated high performance on text classification task (Hamdani *et al.*, 2021). Webersinke *et al*., (2021) had pretrained the climateBERT on a large database of climate related excerpts which improved its performance on tasks such as text classification and sentiment analysis. ClimateBERT is a transformer-based language model that can conduct text classification tasks of climate related texts. Research conducted by Brugger *et al*. (2023) focused on the classification of text from social pillar in the reports. Their sentence transformer text classifier demonstrated promising result in text classification of text related to human rights in the constraint of limited database.

Following the categorization of text fragments, they can be processed through an automated compliance-checking mechanism based on the specified standard requirements. In this study, the reporting criteria outlined by the Global Reporting Initiative (GRI) standards serve as the basis for implementing the integrated rule engine within the sustainability reporting framework.The rule of reporting according to GRI standards are listed here:

i) Disclose all disclosures in GRI 2: General Disclosure 2012
ii) Disclose Materiality Assessment Process Using GRI 3: Material Topics 2021

ACCEPTED MANUSCRIPT

iii) Disclose Material Topics
iv) Disclose Non-Reporting Disclosures Under Material Topics and Reasons for Omissions of Disclosures Items
v) Publish GRI Index
vi) Produce Statement of Use

GRI 2: General Disclosure 2012 falls under universal standards which are applicable to all industry. The first two sections require companies to provide a general overview of company structure, operation and details concerning company sustainability practice. The third section focuses on the governance body and policy of the company. The last two sections should report on the company sustainability strategy development process and stakeholder engagement approach adopted by the company. Companies are allowed to exclude disclosure items from GRI 2 with permitted reasons justifying the exclusion except for Disclosure 2-1: Organizational Details, Disclosure 2-2: Entities included in the organization's sustainability reporting, Disclosure 2-3: Reporting Period, frequency and contact point, Disclosure 2-4: Restatements of information, Disclosure 2-5: External assurance. These five disclosures are the mandatory reporting items that should be included in the report. GRI allows the four reasons with explanations for the exclusion of disclosures (**Table 1**).

Table 1: Permitted Reasons for Disclosure Omissions

Reasons for Omission	**Required Explanation**
Not Applicable	Explain why the disclosure or the requirement is considered not applicable.
Legal Prohibitions	Describe the specific legal prohibitions.
Confidentiality Constraints	Describe the specific confidentiality constraints.
Information Unavailable/ Incomplete	Specify which information is unavailable or incomplete, specify which part is missing (e.g., specify the entities for which the information is missing). Explain why the required information is unavailable or incomplete. Describe the steps being taken and the expected time frame to obtain the information.

(Source: Global Reporting Initiative, 2023)

Companies shall include the material topics and their materiality assessment process in their report using GRI 3: Material Topics 2021. There are three sections in GRI 3. The first two sections must be reported while the third section on material topics management can be omitted with reasons and explanation included in **Table1**.

The rules are incorporated into the report standard compliance assessment system. The proposed decision-making processes for the system are shown in **Figure 5** for GRI standard

ACCEPTED MANUSCRIPT

compliance assessment. The categorized text segments go through the compliance verification process and conclude whether the report comply to the requirements of the reporting standards. The system checks the inclusion of reporting on GRI 2, materiality assessment, reporting of material topics and the GRI index. At the end of the checking process, if the report complies with all requirements, the result of "In Compliance with GRI Standards" is concluded, recorded and send to user interface for display. If the report is incompliant with the standards, the list of the missing elements and the conclusion will be recorded and sent to user interface.

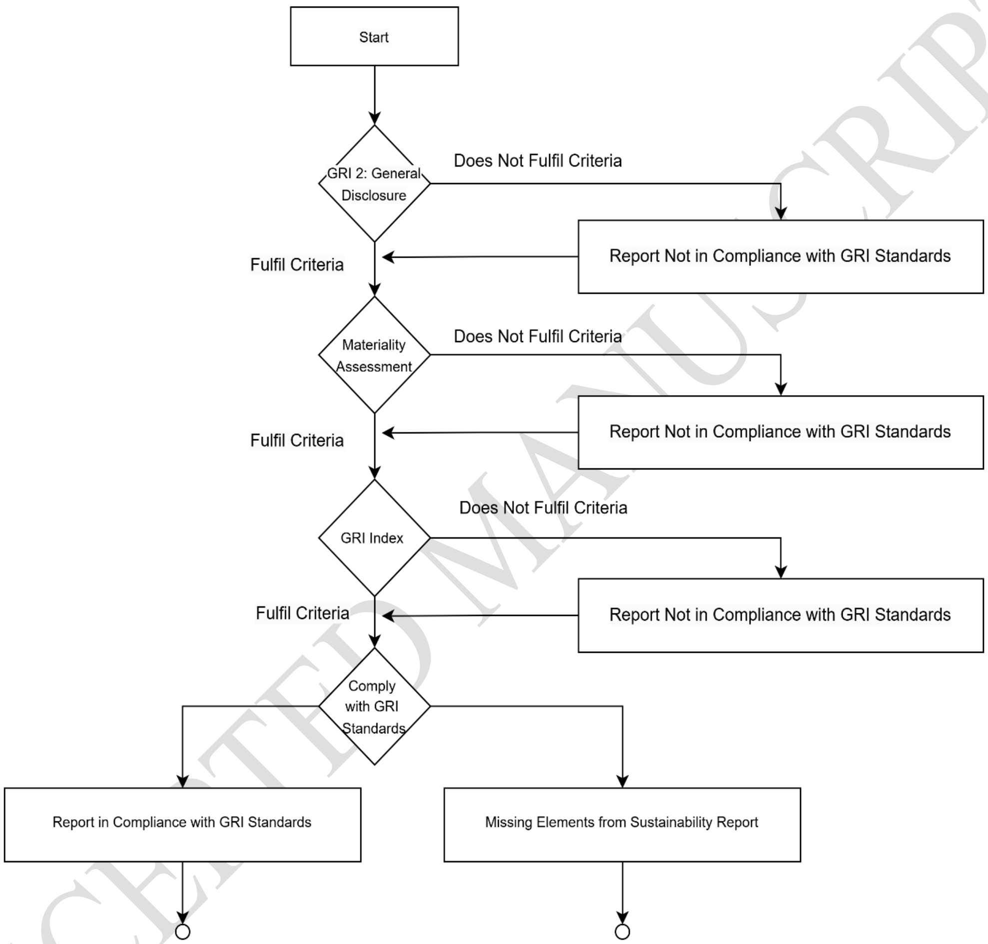


Figure 5: Proposed Decision-making Process for Rule-based Report Standard Compliance Assessment System

Sustainability scoring is used next to assess the sustainability performance of the company. Machine learning random forest regression model is applied in the system to predict the sustainability score of the company based on ESG data extracted from the sustainability report. ().

Different researchers had leveraged various machine learning models in regression tasks to predict ESG scores with financial and non-financial data disclosed by the company. Del Vitto, Marazzina and Stocco (2023) used a variety of white box and black box regression models to reproduce Refinitiv ESG scores for companies from different backgrounds.

ACCEPTED MANUSCRIPT

Lin and Hsu (2023) had deployed a series of machine learning models on ESG scoring calculations for Taiwanese companies using indicators extracted from Taiwan Economic Journal (TEJ). Extreme Learning Machine (ELM), Support Vector Machines (SVM), eXtreme Gradient Boosting (XGBoost) and Random Forest (RF) were tasked with uncovering the relationship between the indicators and the final ESG score. All the models were able to understand the relationships between the indicators and the ESG scores and demonstrated the ability to generate close predictions to the actual ESG scores. The conclusions of this research pointed out that supervised machine learning is faster at solving complex prediction problems when compared with mathematical models. However, the result of the experiment had led to the conclusion of Random Forest not performing as well as other models. The potential reason for this phenomenon could be related to the calibration of the Random Forest models as the study had only applied 20 trees in the model which could lead to the low effectiveness of the model as the number of trees is an important factor in improving the model's performance. . Contreras *et al*., (2021) studied the optimization of Random Forest Regressor application on the modeling of rainfall runoff and forecasting in the Andean Mountains. They concluded that the increase of *n_estimator* value, which is the parameter that decides the number of trees in a random forest model, has a significant positive effect on model performance in the range of 0 to 100. This finding is consistent with the conclusion of Nadi and Moradi (2019) which stated that the model with large number of smaller trees demonstrate better performance. Hence, this study increases in the number of trees from 20 to 30 that results in improvement of model performance

The ESG score has multiple functions in sustainability performance management. The ESG score is indicative of the sustainability performance of the company. The sustainability scoring model uses data of the sustainability indicators to generate sustainability scores, quantifying sustainability performance of the company without human intervention. The data in this decision support system are sustainability reports from various companies that have been scored and categorized into different categories that differ by the sustainability performance of the companies. This study uses the LSEG scoring system as an example of explanation. The LSEG scoring system separates companies into four categories seen in **Table 2** according to their ESG scores.

Table 2: Categories in LSEG ESG Scoring System

Range of ESG Score	**Category**	**Descriptions**
0 to 25	First Quartile	• Poor Relative ESG Performance • Insufficient Degree of Transparency in Reporting Material ESG Data Publicly.
>25 to 50	Second Quartile	• Satisfactory Relative ESG Performance • Moderate Degree of Transparency in Reporting Material ESG Data Publicly.

ACCEPTED MANUSCRIPT

>50 to 75	Third Quartile	• Good Relative ESG Performance • Above Average Degree of Transparency in Reporting Material ESG Data Publicly.
>75 to 100	Fourth Quartile	• Excellent Relative ESG Performance • High Degree of Transparency in Reporting Material ESG Data Publicly

From the categories awarded to the companies, they can understand their standings in terms of sustainability performance. The predicted ESG Score can point out the weaknesses in the company sustainability strategy as the ESG score reflects the sustainability performance of the company. The score of each subcategory under ESG serves as measurements for the sustainability performance of the company. The scores can be traced back to the data of the related indicator set which is the benchmarks that allow the system to track the progress of the company on the sustainability goal achievement. The scores can also be related to a set of indicators that are indicative of the area of improvement. Using the environment score in the LSEG scoring system as examples, the score of each subcategory reveals information on the company's environmental performance on emissions, resource use and innovation. The system will suggest some improvements on the correlated indicators to improve their sustainability in the underperforming sectors. The sustainability indicators related to each category are displayed in **Table 3**.

Table 3: Sustainability Indicators Related to the Categories in LSEG ESG Scoring System

Pillar	**Category**	**Theme**
Environment	Emission	Emission Waste Biodiversity Environmental Management System
	Innovation	Product Innovation Green Revenues R&D and Capital Expenditure
	Resource Use	Water Energy Sustainable Packaging Environmental Supply Chain
Social	Community	Equally Important to All Industries, hence, a median weight of five is assigned to all
	Human Rights	Human Rights
	Product Responsibility	Responsible Marketing Product Quality Data Privacy
	Workforce	Diversity and Inclusion Carrer development and Training Working Conditions

ACCEPTED MANUSCRIPT

		Health and Safety
Governance	CSR Strategy	CSR Strategy ESG Reporting and Transparency
	Management	Structure (Independence, Diversity, Committees) Compensation
	Shareholders	Shareholder Rights Takeover Defences

(Source: Twinamatsiko and Kumar, 2022)

3.3 The functions of MLDSS in Sustainability Management

The proposed MLDSS has mainly three functions, standard compliance assessment of sustainability report, generation of sustainability score, analysis on sustainability data, sustainability progress measurement and recommendations for company sustainability strategy improvement. These functions allow the decision support system to contribute to company sustainability performance management through enhancement of reporting standard compliance, measurement of sustainability progress, achievement of company sustainability goal and support for stakeholder communication.

The utilization of machine learning assisted rule-based model in assessing sustainability reports automates and democratizes the checking for reporting standard compliance. The model checks the criteria for report compliance with the standards. In the situation where the report does not adhere to the standards, the model points out the missing elements of the report to help the company improve their reporting. This function assists companies to ensure their report adheres with the standards which enables them to receive the benefits of sustainability reporting like increased favours in green financing opportunities and strengthening of consumer trust from company transparency (Deloitte & Touche LLP, 2022). Compliance with sustainability reporting standards prevents financial penalties in situations of compulsory reporting for the Publicly Listed Companies (PLC) listed on the markets that requires sustainability reporting as a listing prerequisite and company operating in countries with mandatory reporting requirements.

The decision support system optimizes company sustainability efforts through strategic alignment with their sustainability goals. As all companies have their own sustainability priorities and financial limitations, a standardized sustainability management strategy might not be applicable. Hence, the company has the freedom to introduce their own set of sustainability goals that are in line with the company values and priorities to receive assistance from the DSS for goal achievement. As the recommendations are goal oriented, adoption of these measures ensures the company resources are used in areas with the most significant impact. The annual ESG scoring and benchmarking with the data of sustainability indicators quantify the impact of the current sustainability management strategy, allowing the company to monitor the progress and make timely amendments if necessary.

ACCEPTED MANUSCRIPT

The incorporation of ESG (Environmental, Social, and Governance) scoring visualization and tracking tools constitutes a strategic approach to enhancing stakeholder engagement. These tools facilitate clear and accessible communication of a company's sustainability performance to diverse stakeholder groups, including consumers, employees, investors, and shareholders. As noted by Boufounou et al. (2023), sustainability performance has increasingly influenced consumer purchasing behavior, underscoring the importance of transparent and reliable disclosure mechanisms. The visibility of ESG scores serves not only to reinforce consumer trust but also to provide a robust rationale for sustainability-related capital investments presented to shareholders. Moreover, the use of visualization tools for sustainability metrics is critical to effective and straightforward performance communication. Empirical evidence from Kim, Setlur, and Agrawala (2021) suggests that visual charts significantly outperform raw data in engaging and informing audiences when presented concurrently. Consequently, the strategic deployment of visual representations in ESG reporting can substantially improve comprehension and support informed decision-making among stakeholders, thereby reinforcing the organization's commitment to sustainable practices.

3.4 The Performance of Scoring Model

The ESG scoring model was executed using random forest regressor, with supervised learning style. The task of the model was to predict ESG scores of companies from sustainability data extracted from company sustainability reports according to environmental indicators listed in the GRI standards. The number of trees in a random forest model was set to be 30 and remained for the entire experiment. The experiments were run three times respectively with three samples consisting of sustainability data from a total of 5 companies (Sample 1), 10 companies (Sample 2) and 15 companies (Sample 3). The dataset was split into 60% training and 40% testing data. The results are demonstrated in **Table 4.**

Table 4: The Performance Assessment for 3 Sets of samples

N_estimator	30		30		30	
Sample Label	1		2		3	
No. of ESG disclosure Companies	5		10		15	
Dataset	Training	Testing	Training	Testing	Training	Testing
R^2 Score	0.64	-1,379.64	0.81	-23.19	0.86	0.25
EVS	0.66	-83.64	0.84	0.18	0.86	0.58
RMSE	13.75	18.58	8.96	37.2	8.35	14.62

The performance of the scoring model was assessed through three parameters, the Coefficient of Determination (R^2 score), Explained Variance Score (EVS), and Root Mean

ACCEPTED MANUSCRIPT

Squared Error (RMSE). Overall, the model demonstrated improvement in the prediction of emission score with the increase of sample size. The value of Coefficient of Determination or the R^2 score for training dataset consistently increased from 0.64 to 0.86 when the data size grew from 5 to 15 sets of data. The R^2 score for testing data set started from a negative score of -1,379.64 for sample 1 climbed to -23.19 for sample 2 and reached 0.25 for sample 3 that contains 15 sets of data. The negative R^2 score for sample 1 and 2 meant that the model did not fit the data well. The model could not capture the variability in the relationship between the ESG data and the emission score with the data provided. Despite the relatively low R^2 score observed in testing model output for sample 3, its positive value indicates a gradual improvement in the model's predictive capability. The progression of R^2 scores across both datasets further suggests that the incorporation of additional data has the potential to enhance model performance. This trend underscores the importance of continuous data enrichment in refining predictive accuracy and optimizing model robustness.

The results of explained variance score (EVS) were applied to determine the variance in the actual data is explained by the model's predictions. The EVS for both training and testing datasets increased with the data size. For training data, the EVS for sample 1 was 0.66. After the addition of 5 sets of data, the EVS rose to 0.81. The EVS for sample 3 was 0.86. Similar to R^2 score, the testing EVS for sample 1 was of negative value as well, standing at -83.64. The additional 5 datasets lead to substantial improvement in EVS, causing it to leap to 0.18. The improvement sustained, the EVS for sample 3 was 0.58. From the result of the EVS, the performance of the regression model appeals to benefiting from the increase of data.

The root means squared error (RMSE) is taken by square rooting the MSE. This metric reveals information on the error margin of the prediction made by the model. The training RMSE stood at 13.75 for the smallest data size. The increase of data brought the RMSE down to 8.96 (Sample 2) and 8.35 (Sample 3). The RMSE for the testing data started at 18.58 and grew to 37.2 after doubling the data size and decreased to 14.62 for the data size of 15 companies.

Based on the performance metrics, the performance of the regression model improves with the growth of data size. The R^2 score and EVS demonstrate consistent improvement with the introduction of additional data. Their behaviours communicate that, with large amount of training data, the regression model can capture the variance in the input and output data. The access to new data causes irregular changes in the error evaluation metrics. For the training data, new data can generally effectively bring down the value of the metrics but the metrics for the testing data show a rise-and-fall pattern when new data are introduced. To understand the reasons for such behaviour of the error metrics, attentions are shifted to the composition of the dataset used in this experiment. **Table 5** shows the statistical information of the three samples used. The range of mean for three sample is 72.8 – 75.60, . The 50 percentiles of all three sets of samples had reached the value of 80 while the lowest value in all three sets of samples is only 44.

ACCEPTED MANUSCRIPT

Table 5: Statistical Information of the Samples

Sample Label	**1**	**2**	**3**
No. of ESG Disclosure Companies	**5**	**10**	**15**
Mean	75.40	72.8	75.60
Standard Deviation	23.15	21.57	20.89
Minimum	44.00	44.00	44.00
25 Percentile	58.00	52.00	55.50
50 Percentile	88.00	81.00	85.00
75 Percentile	89.00	88.75	93.00
Maximum	98.00	98.00	99.00
Score > 80	3	6	9
Score < 60	2	4	5

The value of the dependent variable in the database is skewed towards a higher end. During the database development stage, the companies selected to be included in the database were done randomly without consideration to the balance of greenhouse gases (GHGs) emission score in the database. This leads to an imbalanced database consisting of mostly companies with high emission scores. Due to the small sample size and the randomness in the split of training and testing datasets, the distribution of the data points in the training and testing dataset became asymmetric. The training datasets for sample 1 and 2 were mostly populated by low data points, while having most high data points in the testing dataset. The training and testing datasets of Sample 3 had a better mix of low and high data points. However, it must be noted that the database of this experiment does lack diversity.

In addition, this study selects the single component i.e. GHGs emission component for pilot run. The minimum emission score in the database consisting of 15 sets of data was 44 while the median stood at 85, a result of the 8 emission scores having values higher than 80 out of the 15 scores included. A comparison of the predicted and the actual value of the testing datasets for all samples is shown in **Table 6**. The comparison of the prediction and the actual values for the testing data of sample 1 shows that the model has yet to understand the relationship between the data of the sustainability indicators and the emission scores with the training data size of 3. There is a wide gap between the actual and predicted values which, indicating the model is not fitting to the data well.

Table 6: The Actual and Model Output for the GHGs Emission Score in the Testing Dataset

Sample	**1**		**2**		**3**	
No. of ESG disclosure Companies	**5**		**10**		**15**	
No.	Actual	Predicted	Actual	Predicted	Actual	Predicted
1	89	75.6	77	50.70	66	53.30
2	88	65.4	98	52.43	88	86.83
3			85	48.47	95	78.63

ACCEPTED MANUSCRIPT

4			89	51.17	44	53.7
5					77	63.17
6					85	61.20
RMSE		18.58		37.2		14.6

The sample 2 has 10 sets of data with a comparatively balanced data composition. As shown in **Figure 6**, the gap between the actual values and model outputs remains significant. The model begins to show signs of understanding the pattern of scoring, seen through the behaviour of the trend lines between actual values and model output with less intense variability. However, all the actual emission scores are still much higher than the predictions of the model.

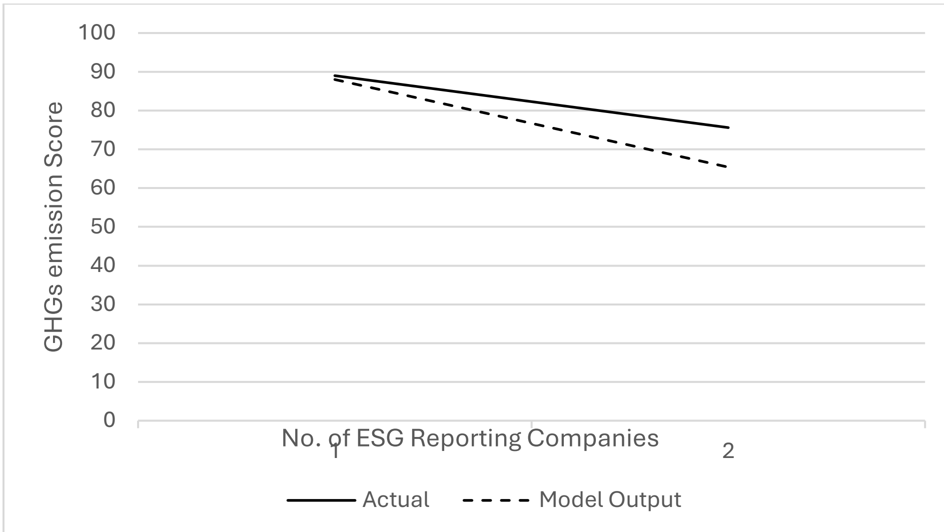
Sample 3 has 15 data points in the testing dataset. The gap between the between actual values and model output has reduced significantly, meaning that the predictions are getting closer to the actual emission scores. Although the model output's trend begins to align with the actual value, there is still some discrepancy between the actual and model output values of several data points. This suggests that the model still requires further training to produce accurate predictions consistently.

Close examination of the predicted emission score provides explanation for the behaviours of the performance metrics. The predicted emission scores for the testing dataset of sample 1 and 2 are lower than the actual emission scores by at least 20 points. The large difference in the actual and the predicted values lead to the surge of all the error related metrics for sample 2. The model also cannot capture the variance of the relationship of the sample 1 and 2 datasets which explains the negative R^2 Score for both sample and the EVS of sample 1.

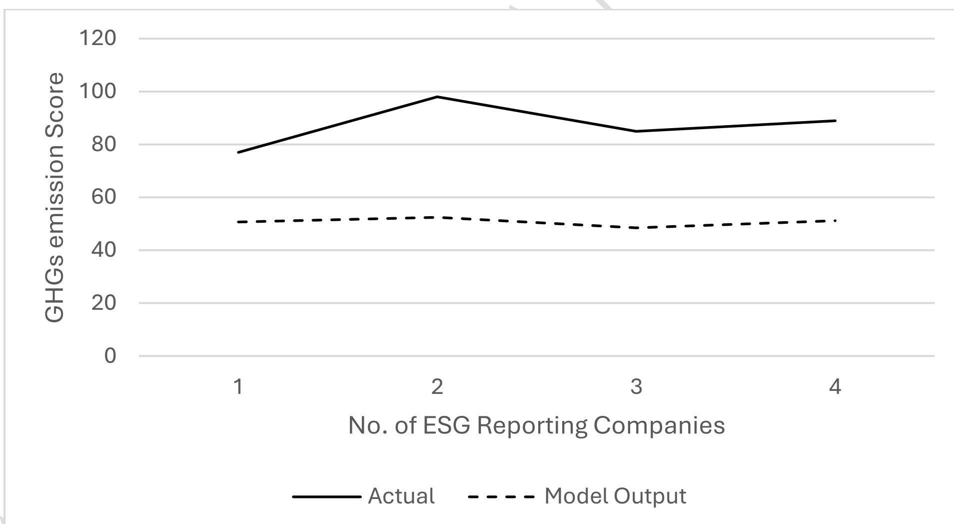
The random forest regressor has shown promise in ESG score prediction, despite limitations arising from a small sample size and asymmetric data composition. The findings highlight a positive relationship between data size and model performance in emission score prediction, corroborating results from earlier research. Cui and Gong (2018) examined datasets ranging from 20 to 700 samples and concluded that regression models stabilize and improve in performance as sample size increases. Similarly, Bouasria et al. (2023) observed that sample size significantly enhances R^2 scores up to 300 samples, after which the impact diminishes. This aligns with Bailly et al. (2022), who reported negligible performance improvements when expanding sample sizes from 1,000 to 100,000. These studies collectively indicate that the effect of data size on regression model performance diminishes after a threshold. Although the dataset in this experiment was insufficient for the random forest regressor to produce consistently stable

ACCEPTED MANUSCRIPT

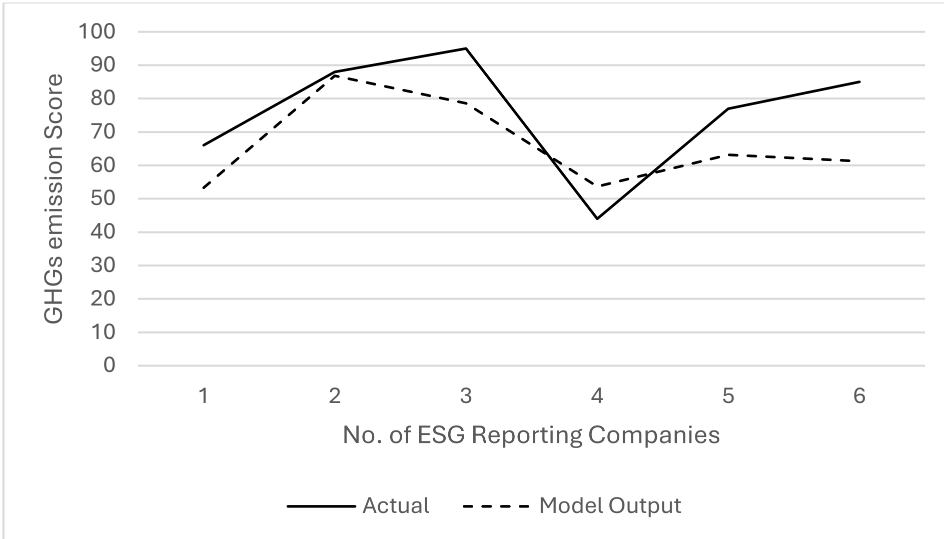
predictions, the promising results from sample 3 suggest that supervised machine learning models hold significant potential for ESG scoring applications.



(a)



(b)



(c)

Figure 6: Model Output and Actual Value of Testing Dataset for (a) Sample 1 (b) Sample 2 and (c) Sample 3

It has been previously emphasized that the database is inherently imbalanced, a result of oversight during its creation. Consequently, the database fails to represent the actual distribution of ESG scores, introducing bias in the emission score prediction model (Gu and Oelke, 2019). This bias causes the model to favour lower emission scores when lower scores dominate the training dataset, as reflected in the prediction outcomes for samples 1 and 2. Variability between the distributions of the training and testing datasets further hinders the model's ability to generalize effectively, resulting in significant increases in MAE, MSE, and RMSE for sample 2 (Ben-David et al., 2010). In comparison, sample 3 exhibits a relatively balanced distribution across datasets, which contributes to enhanced learning performance in the random forest regressor. This observation is supported by its elevated performance metrics, indicating superior predictive accuracy relative to the other samples evaluated.

4. Conclusion

This study develops a machine learning decision support system (MLDSS) framework for managing company sustainability performance using machine learning techniques. The MLDSS analyses sustainability reports for compliance with reporting standards and extracts relevant data to evaluate corporate sustainability performance. A rule-based approach, enhanced by Natural Language Processing (NLP), is employed for the compliance evaluation process. Additionally, a machine learning regression model is utilized to score company performance on sustainability metrics. By integrating machine learning into the DSS, the framework facilitates data-driven decision-making and evolves as new data becomes available. Results comparing predicted and actual sustainability scores highlight the potential of the machine learning model, which effectively captures scoring patterns despite small and imbalanced datasets. This is evident in the predictions for sample 3, where the model closely aligns with actual database values, demonstrating its ability to identify underlying scoring patterns. Given the limitations of the database used in this study,

ACCEPTED MANUSCRIPT

future research could benefit from utilizing a larger and more comprehensive dataset to enhance the reliability of findings.

Acknowledgement

Special thanks to Universiti Tunku Abdul Rahman (UTAR) to provide Research Fund for this study (Project No.: IPSR/RMC/UTARRF/2022-C2/T03)

References

Akbulut D.H. and Kaya I. (2019). Sustainability reporting and firm performance. *Pressacademia*, 9(**9**), 81–84.

Alhaddi H. (2015). Triple Bottom Line and Sustainability: A Literature Review. *Business and Management Studies*, 1(**2**), 6–10.

Bailly A., Blanc C., Francis E., Guillotin T., Jamal F., Wakim B., Roy P. (2022). Effects of dataset size and interactions on the prediction performance of logistic regression and deep learning models. *Computer Methods and Programs in Biomedicine*, 213.

Ben-David S. Luu T., Lu T., Pal D. (2010). Impossibility Theorems for Domain Adaptation. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, PMLR **9**:129-136.

Bouasria A., Bouslihim G., Taghizadeh-Mehrjardi R., Hengl T. (2023). Predictive performance of machine learning model with varying sampling designs, sample sizes, and spatial extents. *Ecological Informatics*, 78.

Boufounou P. Moustairas I., Toudas K., Malesios C. (2023),. ESGs and Customer Choice: Some Empirical Evidence. *Circular Economy and Sustainability*, 3(**4**), 1841–1874.

Brugger F., Etienne D., Hool J., Küpper R., Weiss D. (2023). *Analysing Sustainability Reports Using Machine Learning*. ETH Zurich.

Castilla-Polo, F. and Guerrero-Baena, M.D. (2023). The business case for sustainability reporting in SMEs: consultants' and academics' perceptions. *Sustainable Development*, 31(**5**), 3224–3238. https://doi.org/10.1002/sd.2576.

Chang W.F. Amran A., Iranmanesh M., Foroughi B. (2019). Drivers of sustainability reporting quality: financial institution perspective. *International Journal of Ethics and Systems*, 35(**4**), 632–650.

Chen Y.-C., Hung M. and Wang Y. (2018). The effect of mandatory CSR disclosure on firm profitability and social externalities: Evidence from China. *Journal of Accounting and Economics*, 65(**1**), 169–190. https://doi.org/10.1016/j.jacceco.2017.11.009.

Cui Z. and Gong G. (2018). The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *NeuroImage,* **178**, 622-637.

Contreras P., Orellana-Alvear J., Muñoz P., Bendix J., Célleri R. (2021). Influence of Random Forest Hyperparameterization on Short-Term Runoff Forecasting in an Andean Mountain Catchment. *Atmosphere,* **12**, 238.

Del Vitto A., Marazzina D. and Stocco D. (2023). ESG ratings explainability through machine learning techniques. *Annals of Operations Research* [Preprint].

Deloitte & Touche LLP (2022). *Purpose-driven ESG in the consumer industry: Sustainability as a value-creator*. Available at:

ACCEPTED MANUSCRIPT

https://www2.deloitte.com/content/dam/Deloitte/us/Documents/risk/us-risk-purpose-driven-esg-in-the-consumer-industry-infograhic.pdf (Accessed: 15 August 2024).

Di Matteo E., Roma, P., Zafonte S. Panniello, U., Abbate L. (2021). Development of a Decision Support System Framework for Cultural Heritage Management. *Sustainability*, 13(**13**), 7070.

Flammer C. (2013).Corporate Social Responsibility and Shareholder Reaction: The Environmental Awareness of Investors. *Academy of Management Journal*, 56(**3**), 758–781. https://doi.org/10.5465/amj.2011.0744.

Fischbach J., Adam M., Dzhagatspanyan V., Mendez D., Frattini J., Kosenkov O., Elahidoost P. (2023). Automatic ESG Assessment of Companies by Mining and Evaluating Media Coverage Data: NLP Approach and Tool. *Proceedings - 2023 IEEE International Conference on Big Data, BigData 2023*. Institute of Electrical and Electronics Engineers Inc., pp. 2823–2830. https://doi.org/10.1109/BigData59044.2023.10386488.

Galindo P.V., Vaz E. and de Noronha T. (2015). How corporations deal with reporting sustainability: Assessment using the multicriteria logistic biplot approach. *Systems*, 3(**1**), 6–26.

Global Sustainable Investment Alliance (2021). *Global Sustainable Investment Review 2020*. Available at: www.gsi-alliance.com.

Gu J. and Oelke D. (2019).*Understanding Bias in Machine Learning*. University of Munich.

Hamdani R., Mustapha M., Amariles R.D., Troussel A., Meeùs S., Krasnashchok K. (2021). A combined rule-based and machine learning approach for automated GDPR compliance checking, *Proceedings of the 18th International Conference on Artificial Intelligence and Law, ICAIL 2021*. Association for Computing Machinery, Inc, 40–49. https://doi.org/10.1145/3462757.3466081.

Kalogiannidis S., Kalfas D., Patitsa C. (2024). Relationship between Climate Change and Business Risk: Strategies for Adaptation and Mitigation: Evidence from a Mediterranean Country. *WSEAS Transactions on Environment and Development*, **20**,276–294.

Kanmani A.P., Obringer R., Rachunok B., & Nateghi R. (2020). Assessing Global Environmental Sustainability Via an Unsupervised Clustering Framework. *Sustainability*, *12*(**2**), 563.

Kang H. and Kim J. (2022). Analyzing and Visualizing Text Information in Corporate Sustainability Reports Using Natural Language Processing Methods. *Applied Sciences (Switzerland),* 12(**11**). https://doi.org/10.3390/app12115614.

Kim D.H., Setlur V. and Agrawala M. (2021). Towards understanding how readers integrate charts and captions: A case study with line charts, in *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery.

Kim S., Lee G. and Kang H. (2021). Risk management and corporate social responsibility. *Strategic Management Journal*, 42(**1**), 202–230.

Koh H.-K., Burnasheva, R., & Suh, Y. G. (2022). Perceived ESG (Environmental, Social, Governance) and Consumers' Responses: The Mediating Role of Brand Credibility, Brand Image, and Perceived Quality. *Sustainability*, *14*(8), 4515. https://doi.org/10.3390/su14084515

Laskar N. (2018). Impact of corporate sustainability reporting on firm performance: an empirical examination in Asia. *Journal of Asia Business Studies*, 12(**4**), 571–593.

Li Y. and Rockinger M. (2024). Unfolding the Transitions in Sustainability Reporting, *Sustainability*. 16(**2**), 809.

Lin H.Y. and Hsu B.W. (2023). Empirical Study of ESG Score Prediction through Machine Learning—A Case of Non-Financial Companies in Taiwan. *Sustainability*, 15(**19**),14106.

ACCEPTED MANUSCRIPT

Matsumura, E.M., Prakash, R. and Vera-Muñoz, S.C. (2014). Firm-Value Effects of Carbon Emissions and Carbon Disclosures.*The Accounting Review*, 89(**2**), 695– 724. https://doi.org/10.2308/accr-50629.

Ministry of Investment, T. and I. (2023). *i-ESG National Industry Environmental, Social & Governance Framework Phase 1.0:"Just Transition"*. Malaysia.

Modapothala J.R., Issac B. and Jayamani E. (2010). Appraising the corporate sustainability reports - Text mining and multi-discriminatory analysis, in *Innovations in Computing Sciences and Software Engineering*. Kluwer Academic Publishers, 489–494.

Moses L.T. (2022). Sustainability Reporting Compliance and Financial Performance of Companies Listed on the Nigeria Stock Exchange. *European Journal of Accounting, Auditing and Finance Research*, 10(**5**), 25-73.

Nadi A. and Moradi H. (2019). Increasing the views and reducing the depth in random forest.*Expert Systems with Applications*, 138, 112801. https://doi.org/10.1016/j.eswa.2019.07.018

Nilashi M., Rupani PF., Rupani M.M., Kamyab H., Shao W., Ahmadi H., Rashid T.A., Aljojo N. (2019). Measuring sustainability through ecological sustainability and human sustainability: A machine learning approach. *Journal of Cleaner Production*, 240. https://doi.org/10.1016/j.jclepro.2019.118162

Oncioiu I., Petrescu A.G., Bîlcan F.R., Petrescu M, Popescu D.M., Anghel E. (2020). Corporate Sustainability Reporting and Financial Performance., *Sustainability*, 12(**10**), 4297. https://doi.org/10.3390/su12104297.

Sariyer G. and Taşkın D. (2022). Clustering of firms based on environmental, social, and governance ratings: Evidence from BIST sustainability index, *Borsa Istanbul Review*. Borsa Istanbul Anonim Sirketi, S180–S188.

Segura L.C., Naser A., Abreu R., Perez-Lopez J.A. (2024). ESG Dimensions and Corporate Value: Insights for Sustainable Investments. *Sustainability*, 16(**17**), 7376. https://doi.org/10.3390/su16177376.

Shahi A.M., Issac B. and Modapothala J.R. (2012). Intelligent Corporate Sustainability Report Scoring Solution Using Machine Learning Approach to Text Categorization. *IEEE Conference on Sustainable Utilization and Development in Engineering and Technology (STUDENT)*, Kuala Lumpur, Malaysia.

Sreepriya J., Suprabha K.R. and Prasad K. (2023). Does GRI compliance moderate the impact of sustainability disclosure on firm value?. *Society and Business Review*, 18(**1**), 152–174.

Thayaraj M.S. and Karunarathne W.V.A.D. (2021). The Impact of Sustainability Reporting on Firms' Financial Performance. *Journal of Business and Technology*, 5(**2**), 51–73. https://doi.org/10.4038/jbt.v5i2.33.

Twinamatsiko E. and Kumar D. (2022). Incorporating ESG in Decision Making for Responsible and Sustainable Investments using Machine Learning, in *2022 International Conference on Electronics and Renewable Systems (ICEARS)*. IEEE, 1328–1334.

Vivas R., Sant'anna Â., Esquerre K. and Freires F. (2019). Measuring Sustainability Performance with Multi Criteria Model: A Case Study. *Sustainability*, 11(**21**), 6113.

Wang M.C. (2017). The relationship between firm characteristics and the disclosure of sustainability reporting. *Sustainability*, 9(**4**) 624. https://doi.org/10.3390/su9040624

Webersinke N. *et al.* (2021). *ClimateBert: A Pretrained Language Model for Climate-Related Text*. Technical Report, Cornell University.

ACCEPTED MANUSCRIPT