Vol 6 No 1 (2025) E-ISSN: 2735-1009

International Journal of Management, Finance and Accounting

Decade-Long Analysis of Sustainable Development Goals Compliance and Financial Performance Tiers in All Banking Companies Listed on Bursa Malaysia

Kien Seng Lok[1], Khai Wah Khaw[1,*], XinYing Chew[2], Wei Chien Ng[1]

*Corresponding author: khaiwah@usm.my
[1]School of Management, Universiti Sains Malaysia, Penang, Malaysia
[2]School of Computer Sciences, Universiti Sains Malaysia, Penang, Malaysia

Abstract

This study examines the relationship between Sustainable Development Goals (SDGs) compliance and financial performance in Malaysian banks from 2013 to 2022. Using advanced machine learning techniques, including Support Vector Machines, Decision Trees, K-nearest Neighbours, Extra Trees, Gradient Boosting, and Random Forests, banks were classified into financial performance tiers. Gradient Boosting was the most effective, achieving 80% accuracy in categorising medium and low-performance tiers. Significant correlations were found between SDGs 10 and 15, as well as financial metrics like market capitalisation and asset turnover. These findings highlight the benefits of integrating specific SDGs into banking strategies and the need for supportive policy frameworks, contributing to a deeper understanding of sustainable banking practices.

Keywords: Sustainable Development Goals (SDGs), Financial Performance, Malaysian Banking Sector, Machine Learning, ESG Compliance, Sustainability in Banking, Financial Analysis, Sustainability Reporting.

Received on 15 April 2024; Accepted on 10 June 2024; Published on 28 February 2025.

Published by MMU PRESS. URL https://journals.mmupress.com/ijomfa

1.0 Introduction

The Sustainable Development Goals (SDGs) have driven global banks, including those in Malaysia, to integrate Environmental, Social, and Governance (ESG) practices (United Nations Global Compact Network Malaysia & Brunei, 2022). Malaysia, led by Bank Negara Malaysia (BNM) and Bursa Malaysia, has implemented regulatory frameworks and mandatory sustainability disclosures to promote sustainable banking (Bank Negara Malaysia, 2022, 2023; Bursa Malaysia Berhad, 2021; Bursa Malaysia Securities Berhad, 2015). Major banks have adopted SDG-compliant practices, such as green financing and ESG integration into governance (BusinessToday, 2022; Cheong, 2020; Chu & Donovan, 2020; The Star Online, 2023). Despite these efforts, challenges remain in fully embedding SDGs, transitioning from unsustainable practices, and developing sustainable solutions (Decker & Kingdom, 2020; Global Banking & Finance Review, Tapestry Networks, 2020; World Business Council for Sustainable Development, 2018).

Sustainable banking offers significant societal benefits, including environmental sustainability and community well-being (Beal et al., 2022). The relationship between SDG compliance and financial performance is complex, with varying correlations reported in research. This study aims to explore this relationship in Malaysian banks from 2013 to 2022 using machine learning (ML) models to classify financial performance tiers. By analysing the impact of SDG compliance on financial stability, the research contributes to policy development and promotes a more responsible financial system.

1.1 Problem Statement

In modern banking, the drive towards sustainability has emerged as a central concern, dovetailing with the global commitment to SDGs. However, a significant challenge persists in quantifying and analysing SDG compliance's impact on banking institutions' financial performance. This challenge is particularly pronounced in the context of Malaysian banks listed on Bursa Malaysia, where there is a conspicuous gap in comprehensive, long-term analyses that correlate SDG compliance with financial performance tiers. While there is an increasing recognition among Malaysian banks of the importance of integrating ESG principles into their operations, there remains a critical need to understand how these practices tangibly influence financial metrics. Traditional financial analysis methods often fail to capture the complexities and nuances of the interplay between sustainable practices and financial outcomes, highlighting the need for a more sophisticated approach to grasp these dynamics.

1.2 Research Objectives

This study aims to assess the relationship between SDG compliance and financial performance in Malaysian banks using ML models. The objectives are to: (1) classify Malaysian banks into financial performance tiers (‘Top’, ‘High’, ‘Medium’, and ‘Low’) based on SDG compliance; (2) evaluate the impact of SDG compliance on financial performance, focusing on metrics like Return on Assets (ROA), Return on Equity (ROE), and Market Capitalisation; (3) compare the effectiveness of various ML models, including Support Vector Machines (SVM), Decision Trees (DT), K-Nearest Neighbours (KNN), Extra Trees (ET), Gradient Bossting (GB), and Random Forests (RF), in analysing SDG compliance; and (4) perform a correlation analysis to explore connections between traditional financial metrics and SDG compliance. This research aims to enhance understanding and inform sustainable finance practices.

1.3 Significance of the Study

This study is crucial for integrating sustainable practices in Malaysia’s banking sector, which is a vital part of the national economy. It investigates the relationship between SDG compliance and financial performance, providing insights to embed sustainability more effectively in banking operations. Using ML algorithms to analyse financial data, the research offers a novel understanding of how SDG adherence can benefit financial performance, encouraging banks and regulators to prioritise sustainable practices. The findings support national goals like poverty reduction and climate action, emphasising the importance of transparency and accountability in the banking sector. This study also lays the groundwork for future research in sustainable finance, highlighting the potential for ML in financial analysis.

2.0 Literature Review for Theoretical Background

The nexus between sustainability practices and financial performance in banking is complex. Notable studies such as those by Brogi and Lagasio (2018) and Buallay et al. (2021) demonstrate a positive correlation between ESG disclosure and financial metrics like ROA. However, this relationship is not consistently positive. Forcadell and Aracil (2017), along with Buallay et al. (2021), report instances where ESG practices did not enhance or even negatively impacted financial performance, underscoring the context-dependent nature of sustainability efficacy in banking, influenced by factors like regional economic conditions and corporate governance.

Recent research has also explored the moderating role of non-performing loans (NPLs) in the relationship between SDGs and financial performance. Iqbal and Nosheen (2023) found that high NPL ratios can offset the financial benefits of SDG adoption, emphasising the importance of effective loan management. However, Al Lawati and Hussainey (2022) in Oman observed a positive relationship between comprehensive SDG disclosure and ROE, reinforcing the value of transparency and commitment to sustainability goals.

The impact of SDG adoption on financial performance also varies across industries and regions. Ramos et al. (2022) found no significant relationship across six industries, while Ighosewe (2021) reported mixed impacts in Nigeria. In Asia, some studies have shown positive effects (Burhan & Rahmanti, 2012; Laskar & Maji, 2018), suggesting that the specific dimensions of sustainability practices and the maturity of sustainability reporting within a country can significantly impact financial performance.

The evolving regulatory landscape of sustainable banking also plays a role. Kumar and Prakash (2019) advocate for tailored frameworks in different banking environments, while Ellili and Nobanee (2022) stress the importance of sustainability reporting in financial outcomes.

Recent research has also focused on the temporal dimension, with Lassala et al. (2021) highlighting the potential long-term benefits of sustainability-based business models and Alshehhi et al. (2018) showing a predominantly positive impact of sustainability practices on financial performance over time.

Data-driven approaches and technology adoption are crucial for sustainable finance. Adeoye et al. (2024) proposed a framework for data-driven sustainable finance in the green energy transition, emphasising advanced data analytics and ML. Bahl et al. (2023) highlighted the need for integrating digitalisation and training to enhance banking performance for specific SDGs in India. Chang et al. (2024) demonstrated the potential of ML in predicting corporate financial performance, finding that incorporating SDG-related information improved accuracy.

Despite the broad exploration of this field, several gaps remain. As noted by Khan et al. (2011), disparities in sustainability reporting practices and the lack of empirical studies in areas like the impact of corporate governance on sustainability initiatives call for further research. Other studies highlight the influence of financial performance and ownership structure on sustainability reporting quality. Nechita et al. (2020) found that chemical companies in Central and Eastern Europe with higher R&D investments and lower financial leverage tend to have better SDG reporting. Doğan and Kevser (2021) revealed a negative correlation between concentrated ownership and sustainability scores in Turkish banks, implying that concentrated ownership might prioritise short-term profits over long-term sustainability.

As observed by Khan et al. (2011), there is a notable disparity in sustainability reporting practices among banks in Bangladesh, highlighting gaps in adherence to established guidelines. Similarly, Adu (2022) emphasises the crucial role of corporate governance in enhancing sustainability initiatives in Sub-Saharan Africa’s banking sector, suggesting that both areas require further empirical investigation. Notably, the study by Jasni and Yusoff (2021) examines the relationship between ESG performance and financial performance across different risk-level sectors in Malaysia, highlighting the complexity of these relationships and suggesting the potential for varying impacts across sectors.

A systematic review by Muhmad and Muhamad (2020) further reinforces the positive link between sustainable business practices and financial performance, particularly after adopting SDGs. The review highlights the increasing global awareness and commitment to sustainable development, especially in developing economies. It emphasises the importance of incorporating SDGs into business strategies for long-term financial gains.

According to Pham et al. (2021), focusing on Swedish companies, there is a positive relationship between corporate sustainability practices and various financial performance metrics, such as earnings yield, ROA, ROE, and ROCE. The study suggests that while the general impact of sustainability on financial performance is favourable, the results for market-based financial measures, specifically Tobin's Q, are inconclusive, highlighting the complexity of these relationships. Moreover, Ameer and Othman (2011), who demonstrated a positive association between superior sustainability practices and financial performance among top global companies, further emphasise the need for context-specific analyses and consideration of multiple dimensions of sustainability.

Finally, while ML's role in banking, particularly in credit analysis and risk prediction, is highlighted by studies like Huang et al. (2007) and Doko et al. (2021), a notable gap exists in the application of ML to analyse the relationship between SDG compliance and financial performance, suggesting an unexplored avenue for future research.

2.1 Hypotheses and Conceptual Framework

Given the existing literature's mixed findings on the relationship between sustainability practices and financial performance, this study aims to clarify this relationship in the context of Malaysian banks listed on Bursa Malaysia. Drawing upon prior research and relevant theoretical frameworks, this paper hypothesises that the relationship between a bank's overall SDG compliance score and its financial performance may not be strictly linear. Instead, this study anticipates a potential U-shaped or inverted U-shaped curve.

This suggests that very low and very high SDG compliance could be associated with suboptimal financial performance compared to moderate levels. This hypothesis is grounded in stakeholder theory, which posits that firms must balance various stakeholders' demands (Agudo-Valiente et al., 2015; Freeman, 1984; Mitchell et al., 1997). Excessive focus on SDG compliance might divert resources from core business activities, leading to suboptimal financial performance, while too little focus could alienate ethically-minded investors and customers.

Additionally, this study hypothesises that the impact of individual SDGs on financial performance will vary. This study anticipates that certain SDGs, particularly those related to environmental sustainability (e.g., SDG 7; SDG 13) and social responsibility (e.g., SDG 8), will positively influence financial outcomes. This expectation is based on the literature’s findings that environmental and social practices can enhance a bank’s reputation, reduce risks, and attract customers, aligning with legitimacy theory, which suggests that firms aligning their actions with societal values are perceived as more legitimate and may experience better financial performance (Clementino, 2021; Dowling & Pfeffer, 1975; Suchman, 1995)

While SDG compliance is increasingly recognised as important, this study also hypothesises that traditional financial indicators (e.g., market capitalisation, valuation ratios like P/E and price-to-book ratios, profitability metrics like ROE and ROA) will remain significant predictors of a bank’s financial performance. This is consistent with signalling theory and information asymmetry theory, which suggest that companies disclose both financial and non-financial information to signal their quality and reduce information asymmetry with investors, influencing their financial performance (Akerlof, 1970; Dhaliwal et al., 2011; Spence, 1973).

To investigate these hypotheses, this study employs a conceptual framework that views SDG compliance as a complex, multi-dimensional construct with diverse pathways for influencing financial performance. This paper recognises that effective environmental performance may reduce costs and risks, while robust social practices could enhance customer loyalty and brand value. However, the framework also acknowledges the potential for trade-offs and the importance of considering the specific

SDG, bank characteristics, and broader economic context to fully understand the multi-faceted relationship between SDG compliance and financial performance in Malaysian banks.

3.0 Methodology

3.1 Data Sources

The study's dataset is compiled from several reputable sources to ensure a comprehensive analysis of Malaysian banks' financial performance and sustainability practices (see Table 1). Bloomberg Terminal provided market capitalisation data, while Morningstar contributed additional financial metrics. Sustainability metrics were sourced from banks' annual and sustainability reports, and macroeconomic indicators were obtained from the World Development Database. This diverse dataset allows an in-depth exploration of the relationship between banks' financial performance, sustainability efforts, and the broader economic landscape.

3.2 Tools

This research uses Jupyter Notebook (Anaconda 3) with Python for data analysis and exploration. Jupyter Notebook, acclaimed for its adaptability and interactivity, allows iterative composition, execution, and refinement of Python code. Its integration with libraries like Pandas, NumPy, and Scikit-learn supports efficient data handling, statistical modelling, and ML algorithm implementation. The platform's visualisation capabilities and Markdown documentation enhance clarity and reproducibility, while its text-based nature facilitates version control, promoting transparency and collaboration.

Table 1: Summary of Dataset Variables, Sources and Rationale

Variable(s)	Source(s)	Rationale
'Company Name'	N/A – Identifier	Identifies the banking institutions within the dataset.
'Fiscal Year'	N/A - Time reference	Indicates the period for which the data is relevant.
'Market_Capitalisation_Fiscal_Year' 'Market_Capitalisation_Calendar_Year'	Bloomberg Terminal	Reflects the market value of the banks' equity at specific time points for trend analysis and year-over-year comparison.
Financial Performance Metrics ('Return_on_Asset' through 'Total_Return')	Morningstar	Captures a comprehensive array of financial indicators to assess the banks' economic value, operational efficiency, and financial stability.
SDG Compliance Metrics ('SDG1_Compliance' through 'SDG17_Compliance')	Company Annual and Sustainability Reports.	The banks' adherence to SDGs reflects their commitment to sustainable banking practices.
Sustainability Report Metrics ('Sustainability_Report' through 'External_Assurance_for_Sustainability')	Company Annual and Sustainability Reports.	Details the banks' sustainability efforts, workforce, and the verification of sustainability reporting. 'Approximate_Employee_Count_(Over)' is derived from 'Reported_Employee_Count'.
Economic Indicators ['GDP_Growth_(Annual_Percentage)' through 'Real_Interest_Rate']	World Development Database	Provides critical economic context, including growth, inflation, and interest rates, to frame the financial and sustainability data within the broader macroeconomic environment.

4.0 Data Analysis

4.1 Data Description and Initial Assessment

This research utilises a dataset comprising 100 entries with 53 distinct variables, including financial metrics, sustainability compliance indicators, and macroeconomic factors, detailed in Table 1. Key financial metrics like ‘Market Capitalisation’, ‘Return on Asset’, and ‘Debt/Equity Ratio’ are included, offering insights into the financial health of the banks across a varied landscape.

Sustainability efforts are quantified through SDG-related variables, which are crucial for examining the impact of sustainable practices on financial performance. Additionally, macroeconomic indicators such as GDP growth, inflation rates, and real interest rates provide a contextual backdrop, enhancing the analysis of the banks’ operational environments.

The dataset also includes data on employee counts, with ‘Reported_Employee_Count’ available for 71 entries and ‘Approximate_Employee_Count_(Over)’ for 17 entries. This data is instrumental in understanding the banks' operational scale and its correlation with financial and sustainability metrics.

4.2 Data Exploration

Data exploration is a crucial initial phase in data analysis, involving meticulous examination to understand dataset attributes, discern patterns, and extract insights. This process includes summarising key characteristics, visualising features, and identifying data anomalies or gaps. The primary goals are to gain a deeper understanding of the dataset, uncover variable interconnections and lay the groundwork for subsequent data processing and ML applications. This will enhance people's grasp of the data’s nuances and set a solid foundation for advanced analytical techniques (Magdum, 2022).

4.2.1 Descriptive Statistics and Data Quality

4.2.1 (a) Descriptive Statistics and Data Quality

	count	mean	std	min	25%	50%	75%	max
Fiscal Year	100	2017.5	2.886751	2013	2015	2017.5	2020	2022
Market_Capitalisation_Fiscal_Year	1.00E+02	3.29E+10	3.08E+10	3.67E+09	7.10E+09	2.03E+10	5.20E+10	1.06E+11
Market_Capitalisation_Calendar_Year	1.00E+02	3.27E+10	3.08E+10	2.94E+09	6.78E+09	2.03E+10	5.20E+10	1.06E+11
Return_on_Asset	100	0.00937	0.004106	-0.0226	0.008125	0.0096	0.0113	0.0144
Fixed_Assets_Turnover	100	7.7048	4.970837	1.68	4.43	6.415	8.3125	26.02
Debt_/_Equity_Ratio	100	0.468	0.256621	0.09	0.2675	0.385	0.6625	1.3
Earnings_Yield	100	0.084675	0.052437	-0.3966	0.0766	0.0872	0.103625	0.1622
EPS_Percentage_Year_Over_Year	100	1.616751	10.328066	-50.41	-0.14315	0.0526	2.07245	37.93
Price_/_Earnings_Ratio	100	11.5985	2.886161	0	9.6275	11.435	13	23.33
Enterprise_Value	1.00E+02	4.44E+10	4.27E+10	5.45E+09	8.69E+09	2.82E+10	7.17E+10	1.57E+11
Net_Margin_Percentage	100	0.332481	0.151871	-0.7248	0.2655	0.32435	0.389975	0.5944
Tax_Rate	100	0.235159	0.064382	0	0.212475	0.2357	0.254	0.674
Return_on_Equity	100	0.104155	0.048327	-0.2304	0.0864	0.10325	0.132775	0.2119
Price_/_Sales	100	4.0248	1.835073	1.61	2.74	3.425	4.285	8.77
Price_/_Book	100	1.2332	0.547778	0.38	0.8675	1.17	1.525	3.52
Asset_Turnover	100	0.0293	0.006237	0.02	0.03	0.03	0.03	0.05
Financial_Leverage	100	10.9619	2.296072	7.27	9.6325	10.565	11.8325	17.98
Book Value_/_Share	100	6.1084	4.036903	1.02	3.56	5.12	6.7375	20.82
Capital_Expenditure_as_a_Percentage_of_Sales	100	0.124523	0.822165	0.0106	0.0268	0.03675	0.046675	8.26
Free_Cash_Flow_/_Sales	100	7.367813	24.388447	-3.0895	-0.05445	0.404	1.2123	126.44
Free Cash Flow_/_Net Income	100	1.3204	4.856662	-12.25	-0.18	1.145	2.685	35.99
Free Cash Flow_/_Share	100	0.7738	1.973882	-3.81	-0.41	0.42	1.62	8.64
Dividend_Per_Share	100	0.247	0.162627	0	0.12	0.2	0.38	0.67
Total_Yield	100	0.036423	0.015246	0	0.0267	0.0337	0.043825	0.0839
Payout_Ratio	100	0.413615	0.199424	0	0.29915	0.39765	0.494475	1.211
Price_/_Fair_Value	100	0.8974	0.184723	0	0.84	0.91	1.01	1.15
Total_Return	100	0.628163	3.167061	-3.05	-0.04905	0.0455	0.1368	25
Reported_Employee_Count	71	18718.02817	14865.64055	36	6312.5	14139	33724	47771
Approximate_Employee_Count_(Over)	17	5923.529412	1760.869605	3700	4800	5600	6400	9500
GDP_Growth_(Annual_Percentage)	100	0.041563	0.035333	-0.055344	0.044132	0.047684	0.058127	0.086943
Inflation_Consumer_Prices_(Annual_Percentage)	100	0.019579	0.014153	-0.011387	0.008847	0.021047	0.03143	0.038712
Real_Interest_Rate	100	0.022758	0.025874	-0.023939	0.007989	0.030665	0.044675	0.048016

Figure 1: Description of Dataset

Descriptive statistics play a pivotal role in illuminating the dataset's characteristics, guiding towards a nuanced understanding of the Malaysian banking sector and its engagement with sustainable practices. This study delves into central tendency, dispersion, variability, and data quality measures to uncover patterns, anomalies, and potential limitations. Key indicators such as mean, median, mode, and quartiles are used to describe the central tendency of the data. In contrast, measures like variance, standard deviation, and quartile deviation illuminate the extent of data spread. These descriptive statistics are presented in Figure 1, including counts, mean, standard deviation, minimum, 25th percentile (lower quartile), median (50th percentile), 75th percentile (upper quartile), and maximum for each variable, thereby offering a comprehensive overview of the dataset's characteristics in the context of present research focus.

4.2.1 (b) Missing Value Analysis

Visualising missing values is crucial for understanding their extent and pattern, influencing data treatment decisions. While most of the dataset is complete, ‘Reported_Employee_Count’ and ‘Approximate_Employee_Count_(Over)’ have significant missing entries, with only 71 and 17 non-null entries, respectively. Contrasting these variables against those with high completeness highlights the data gaps. This approach provides immediate insight into data integrity. It informs methods like data imputation to address these gaps, ensuring a robust foundation for analysis and enhancing the research’s reliability and validity.

4.2.2 Distribution and Relationship Analysis

4.2.2 (a) Distribution and Relationship Analysis

Financial and sustainability data analysis for banks listed on Bursa Malaysia reveals vital trends and variations across financial metrics and sustainability efforts. Histograms and boxplots indicate diverse values in metrics such as market capitalisation, ROA, price-to-earnings (P/E) ratio, free cash flows, and dividends per share, predominantly displaying right-skewed distributions. This pattern suggests a concentration of financial strength among a few banks with exceptionally high profitability or valuation.

Operational efficiency metrics like fixed asset turnover, net profit margin, asset turnover, and capital expenditure as a percentage of sales also vary significantly, indicating different management strategies and capabilities in asset utilisation and cost control. Similarly, leverage levels and cash flow generation efficiency vary across banks, reflecting diverse capital structures and financial management approaches.

Macroeconomic factors, including GDP growth, inflation, and interest rates, further illustrate banks’ fluctuating business conditions, necessitating strategic choices in growth, leading and cost optimisation amidst external uncertainties and internal challenges.

From a sustainability perspective, commitment levels to different SDGs vary, with higher compliance focused on decent work and economic growth (SDG 8) and responsible consumption and production (SDG 12). Following that, SDG 9 (Industry, Innovation, and Infrastructure) showed a high compliance rate, indicating a strong focus on fostering innovation and infrastructure development. However, environmental and climate-centric goals, such as SDG 6 (Clean Water and Sanitation), SDG 13 (Climate Action), and SDG 14 (Life below Water), see lower engagement, highlighting a focus on social and economic sustainability over environmental issues.

Despite the prevalence of sustainability reporting among banks, external assurance and strategic integration of SDGs remain limited, indicating potential areas for improvement in sustainability governance. Last but not least, employee counts also showed a right-skewed distribution, with a few banks employing a large workforce, affecting talent management practices and workplace culture across the sector.

4.2.2 (b) Correlation Analysis

A detailed correlation analysis in the banking sector reveals critical financial and sustainability metrics insights. Table 2 displays the correlation interpretation ranges. High correlations among various financial metrics reflect common banking practices and principles, enhancing understanding of banks' financial aspects.

Table 2: Correlation Interpretation Ranges

Correlation Coefficient	Interpretation
0.5 and above	Strong Positive Correlation
0.3 to 0.499	Moderate Positive Correlation
0.1 to 0.299	Weak Positive Correlation
0 to 0.099	Negligible/Very Weak Correlation
-0.099 to 0	Negligible/Very Weak Correlation
-0.1 to -0.299	Weak Negative Correlation
-0.3 to -0.499	Moderate Negative Correlation
-0.5 and below	Strong Negative Correlation

Table 3: Comparative Analysis of Variable Correlations with Employee Count Metrics Before and After Imputation (Part 1)

Variable	'Reported_Employee_Count'		'Approximate_Employee_Count_(Over)'	
	Before Imputation	**After Imputation**	**Before Imputation**	**After Imputation**
'Fiscal Year'	Weak negative	Weak negative	Strong positive	Weak negative
'Market_Capitalisation_Fiscal_Year'	Strong positive	Strong positive	Strong positive	Strong positive
'Market_Capitalisation_Calendar_Year'	Strong positive	Strong positive	Strong positive	Strong positive
'Fixed_Assets_Turnover'	Weak negative	Weak negative	Moderate positive	Weak negative
'Debt_/_Equity_Ratio'	Strong positive	Strong positive	Moderate positive	Strong positive
'Earnings_Yield'	Weak negative	Weak negative	Strong positive	Weak negative
'Price_/_Earnings_Ratio'	Moderate positive	Moderate positive	Weak negative	Moderate positive
'Enterprise_Value'	Strong positive	Strong positive	Strong positive	Strong positive
'Tax_Rate'	Weak negative	Weak positive	Moderate negative	Weak positive
'Price_/_Sales'	Weak negative	Weak positive	Moderate negative	Weak positive
'Price_/_Book'	Weak negative	Weak positive	Moderate negative	Weak positive
'Book Value_/_Share'	Weak negative	Weak negative	Strong positive	Weak negative
'Free_Cash_Flow_/_Sales'	Weak positive	Weak positive	Moderate positive	Weak positive
'Free Cash Flow_/_Net Income'	Weak positive	Weak positive	Moderate positive	Weak positive
'Free Cash Flow_/_Share'	Weak negative	Weak negative	Moderate positive	Weak negative
'Dividend_Per_Share'	Moderate positive	Moderate positive	Weak negative	Moderate positive
'Total_Yield'	Moderate positive	Moderate positive	Weak negative	Moderate positive
'Payout_Ratio'	Strong positive	Strong positive	Moderate negative	Moderate positive
'Total_Return'	Moderate positive	Moderate positive	Weak positive	Moderate positive

Note. This table includes only those observations where at least one of the two variables demonstrates a correlation strength of moderate or higher. For instance, 'Fiscal Year' is included in the table because it shows a strong correlation with 'Approximate_Employee_Count_(Over)' before imputation. This inclusion is made despite the correlations being weak for both 'Reported_Employee_Count' (before and after imputation) and 'Approximate_Employee_Count_(Over)' after imputation.

Table 4: Comparative Analysis of Variable Correlations with Employee Count Metrics Before and After Imputation (Part 2)

Variable	'Reported_Employee_Count'		'Approximate_Employee_Count_(Over)'	
	Before Imputation	**After Imputation**	**Before Imputation**	**After Imputation**
SDG2_Compliance	Weak positive	Weak negative	Strong positive	Weak negative
SDG3_Compliance	Weak positive	Weak positive	Strong positive	Weak positive
SDG4_Compliance	Weak positive	Weak positive	Strong positive	Weak positive
SDG5_Compliance	Weak positive	Weak positive	Strong positive	Weak positive
SDG8_Compliance	Weak positive	Weak positive	Strong positive	Weak positive
SDG9_Compliance	Weak positive	Weak positive	Strong positive	Weak positive
SDG10_Compliance	Moderate positive	Moderate positive	Weak negative	Moderate positive
SDG11_Compliance	Weak positive	Weak positive	Strong positive	Weak positive
SDG12_Compliance	Weak positive	Weak positive	Strong positive	Weak positive
SDG15_Compliance	Moderate positive	Moderate positive	Weak negative	Moderate positive
SDG16_Compliance	Weak negative	Weak positive	Moderate negative	Weak positive
'Sustainability_Report'	Weak positive	Weak positive	Strong positive	Weak positive
'SDGs_Guided'	Weak positive	Weak positive	Strong positive	Weak positive
'Reported_Employee_Count'	Perfect positive	Perfect positive	Strong positive	Strong positive
'Approximate_Employee_Count_(Over)'	Strong positive	Strong positive	Perfect positive	Perfect positive
'External_Assurance_for_Sustainability'	Moderate positive	Strong positive	Strong positive	Strong positive
'Inflation_Consumer_Prices_(Annual_Percentage)'	Weak negative	Weak positive	Moderate negative	Weak positive

Note. The table selectively presents variables that exhibit at least a moderate correlation in one of the observed relationships. For example, the variable 'SDG2_Compliance' is included because it displays a strong positive correlation with 'Approximate_Employee_Count_(Over)' prior to imputation. This is in spite of the fact that 'SDG2_Compliance shows only weak positive and weak negative correlations with 'Reported_Employee_Count' before and after imputation, respectively, as well as a weak negative correlation with 'Approximate_Employee_Count_(Over)' following imputation.

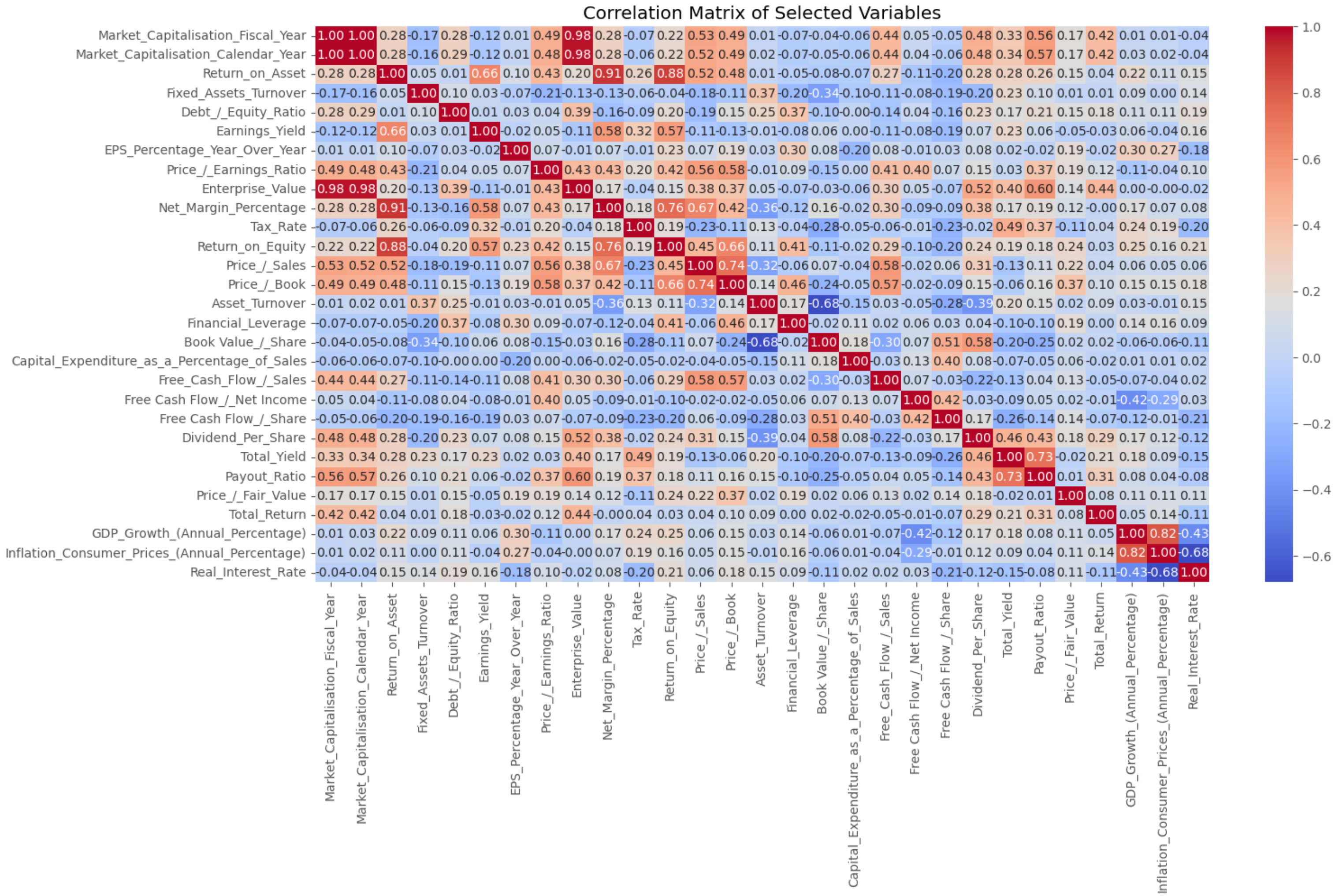

Figure 2: Correlation Matrix of the Financial Performance Metrics and Macroeconomic Indicators

A noteworthy finding is the near-perfect correlation between reported and approximate employee counts. Despite this, some variables demonstrated discrepancies in their correlation before and after imputation, as detailed in Tables 3 and 4. This necessitated imputation, explained in Section 4.3.1, to address missing value entries discussed in Section 4.2.1(b).

Figure 2 presents a correlation matrix between financial performance indicators and macroeconomic indicators. It highlights a nearly perfect correlation between fiscal and calendar year market capitalisations, suggesting stability in short-term bank valuations. Furthermore, a high correlation between market capitalisation and enterprise value indicates that market capitalisation is a robust indicator of a bank's financial structure. Significant correlations are also observed between profitability metrics, such as ROA and net margin percentage.

The analysis reveals several key trends. Larger banks, measured by market value, tend to employ more staff. Financial ratios, such as price-to-book and price-to-sales, often predict higher dividend payouts. Unexpectedly, this study found a link between employee count and a bank's debt-to-equity ratio, suggesting that workforce size may influence a bank's financial structure.

The results are mixed when looking at how companies' efforts to meet sustainability goals affect their bottom line. Overall, there does not seem to be a direct link between meeting these goals and immediate financial gains. In other words, the analysis generally reveals low correlations with SDGs, indicating limited direct financial benefits from SDG compliance. However, there are some exceptions. For example, companies prioritising reducing inequalities (SDG 10) and protecting life on land (SDG 15) tend to perform better in several financial areas, including their overall market value and the profits they share with shareholders.

Efforts to reduce poverty (SDG 1), address climate change (SDG 13), and foster partnerships (SDG 17) show a more consistent link with financial performance. These goals correlate with companies sharing more profits with shareholders (higher 'Payout Ratio'). In addition, prioritising goals 1 and 17 is linked with higher overall investment

Published by MMU PRESS. URL https://journals.mmupress.com/ijomfa

returns (‘Total Yield’), while focusing on SDG 17 is also associated with higher dividend payouts per share. Companies that prioritise SDG 13 tend to have a higher overall market value.

While the link between financial performance and efforts to achieve goals like Zero Hunger, Good Health and Well-being, and Quality Education is less pronounced, these goals are increasingly factoring into companies’ long-term financial strategies. This suggests that while these efforts may not yield immediate financial gains, companies recognise their importance for sustained success.

While most sustainability goals have a limited direct impact on financial performance, efforts towards clean water and sanitation (SDG 6) are moderately linked to how efficiently a company uses its assets to generate sales. On the other hand, efforts to conserve marine life (SDG 14) show little to no connection with financial measures. This suggests that investors and the market value different sustainability goals differently and that the impact of these efforts on financial results can vary widely.

When looking at how these goals have changed over time, the results indicate a growing trend for companies to consider a wider range of sustainability goals in their financial planning. Goals related to education (SDG 4), gender equality (SDG 5), clean energy (SDG 7), and others are showing increasingly strong links with companies' long-term financial strategies. This suggests a shift in the banking sector towards a more holistic approach to sustainability, where a broader spectrum of social and environmental issues is factored into financial decision-making.

Interestingly, some trade-offs are observed when companies heavily prioritise specific sustainability goals. For example, a strong focus on goals like Zero Hunger (SDG 2), SDG 8, Sustainable Cities (SDG 11), and Responsible Consumption (SDG 12) may be associated with slightly lower returns for investors in the short term. Similarly, prioritising SDG 4 or SDG 13 sometimes correlates with higher company debt levels, which can be a financial risk. These findings suggest that while focusing on important social and environmental goals is valuable, it can sometimes create short-term financial challenges.

Companies increasingly publish sustainability reports and align their activities with SDGs, demonstrating a growing commitment to social and environmental responsibility. However, this increased focus on sustainability sometimes coincides with slightly lower short-term returns for investors and a higher reliance on debt. This suggests that prioritising sustainability might come with some short-term financial trade-offs.

Additionally, the study found variations in how companies report their SDG compliance. Some companies appear to align their actions more closely with the SDGs than their formal reporting suggests, possibly indicating a gap between actual practices and how they are officially documented. This highlights the need for transparency and accuracy in sustainability reporting to give investors and stakeholders a clear picture of a company's commitment to sustainability.

Regarding broader economic factors, the relationship between a country's economic growth and companies' available cash is somewhat negative, suggesting a potential trade-off between growth and immediate liquidity. The link between inflation and employment levels is complex, with initial negative correlations becoming weakly positive when accounting for missing data (see Table 4). Real interest rates show weak connections with various economic factors, company financial strategies, and specific sustainability goals, particularly those related to clean energy and climate action.

4.2.3 Temporal Trends

The fluctuating market capitalisation trends in the Malaysian banking sector reflect responsiveness to internal and external factors like market conditions and corporate strategies. It measures the total value of a company's shares. The stability of ROA, which indicates how efficiently a bank generates income from its assets, suggests consistent operational efficiency. Variations in Earnings Yield, the ratio of earnings per share to the market price, reflect shifts in profitability and market valuation. Fluctuations in the Debt/Equity Ratio, which assesses financial risk by comparing total debt to shareholder equity, indicate strategic financial decisions. Changes in the P/E Ratio, the stock price divided by earnings per share, reveal evolving investor perceptions and market confidence. These metrics collectively offer a comprehensive view of the sector's financial health and market positioning.

A significant trend is the increasing integration of sustainability practices, marked by a rise in reporting aligned with the SDGs and increased external assurance for these reports. This shift reflects regulatory pressures and evolving societal expectations, positioning sustainability as a core element of modern banking.

Fluctuations in employee counts suggest a dynamic workforce landscape, potentially indicative of a shift towards more agile operational models. Additionally, the banking sector's sensitivity to macroeconomic factors like inflation and real interest rates underscores its interconnectedness with the broader economy.

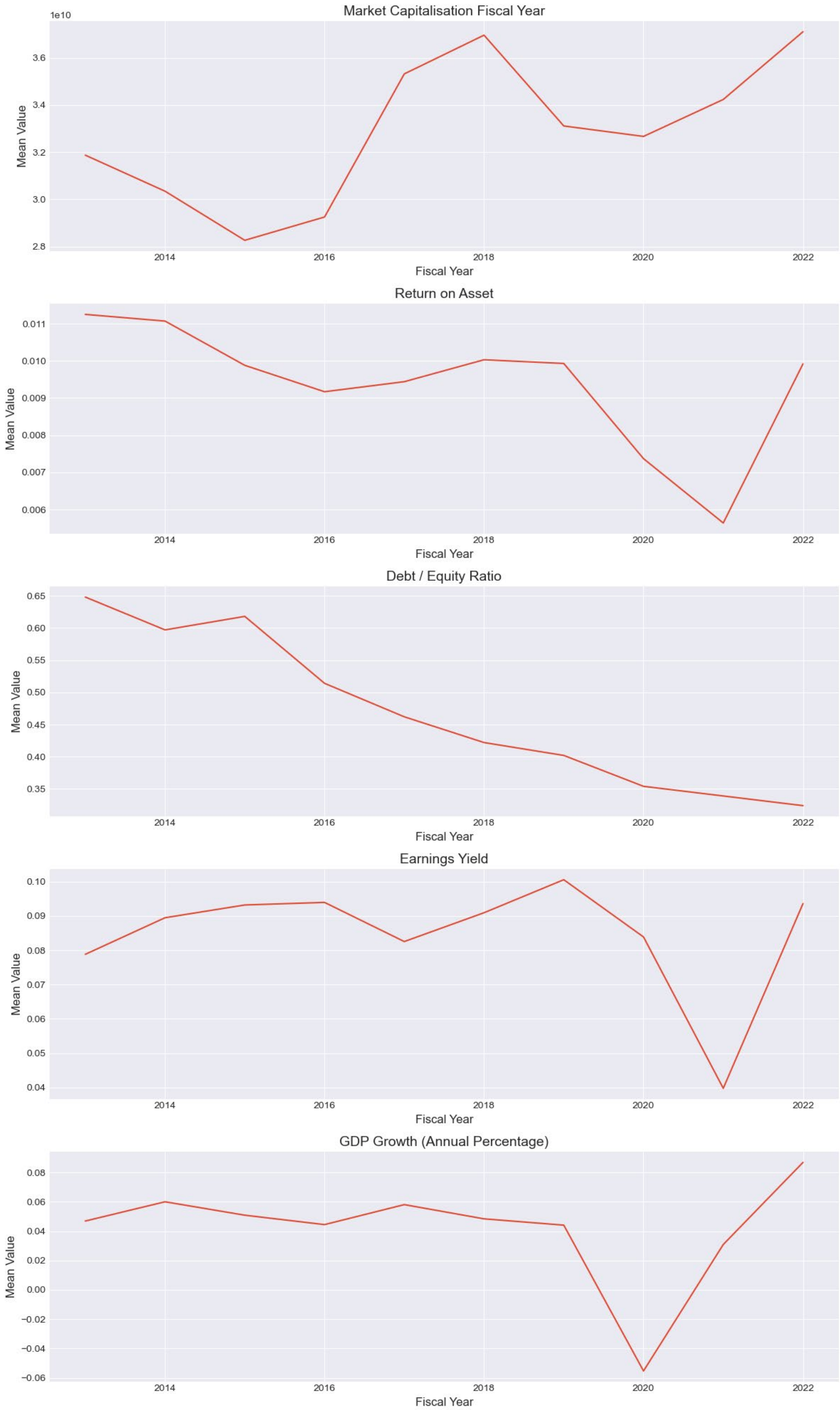

Figure 3: Time-Series Analysis of Key Financial and Economic Metrics

The Malaysian banking sector is evolving, balancing financial performance with increasing sustainability and social responsibility commitments. These insights highlight the sector's complex and multi-faceted nature, providing valuable information for stakeholders and emphasising the sector's critical role in fostering a sustainable and equitable future.

4.3 Data Preprocessing

Data preprocessing is a vital stage in this research, serving as the foundation for accurate and insightful analysis. This section outlines the systematic steps to refine the dataset, ensuring its analytical robustness and alignment with this research objectives.

4.3.1 Handling Missing Values

The dataset's missing values in 'Reported_Employee_Count' and 'Approximate_Employee_Count_(Over)' were addressed using Multiple Imputation by Chained Equations (MICE) with RandomForestRegressor as the estimator. This approach ensures data integrity and showcases proficiency in advanced data processing techniques. MICE effectively handles complex datasets, while RandomForestRegressor accommodates diverse data distributions, ensuring a robust imputation. While these variables may be excluded from later analyses, their comprehensive treatment reflects a commitment to understanding and preparing the dataset for informed decision-making. The meticulous handling of missing data through MICE aligns with the broader research objectives, enabling a thorough exploration of workforce dynamics in the banking sector and demonstrating a rigorous approach to data analysis.

4.3.2 Binary Encoding of Categorical Variables

The dataset, containing many 'Yes/No' responses, especially for SDG compliance, requires binary encoding for ML algorithms. This process converts 'Yes' to 1 and 'No' to 0, retaining the original data's essence while making it suitable for quantitative analysis. This is crucial for accurately assessing SDG compliance and translating qualitative assessments into quantifiable metrics.

4.3.3 Removing Unnecessary Variables

During data preprocessing, this study strategically removed non-essential variables to focus this research analysis on the impact of SDG compliance on the financial performance of Malaysian banks. External Assurance for Sustainability, while informative about reporting practices, does not directly influence this relationship. Similarly, 'Reported Employee Count' and 'Approximate Employee Count (Over),' although relevant to operational scale, were less impactful in analysing the relationship between SDG compliance and financial performance. Removing these variables addressed the missing values and streamlined the dataset, ensuring alignment with this research objectives and eliminating potential distractions.

4.3.4 Classification of Financial Performance Metrics

A comprehensive assessment of each bank's financial health is conducted using multiple financial indicators. Initially, these metrics are standardised through Robust Scaling to mitigate the impact of outliers. A composite score, calculated as the median of the scaled indicators, reflects each bank's overall financial standing. Banks are ranked based on this score and classified into four tiers ('Top,' 'High,' 'Medium,' and 'Low') using a quantile-based approach for balanced distribution. This enables nuanced comparisons across the banking sector, with each tier representing a segment with varying financial robustness and operational sustainability.

4.3.5 Indicating SDG Compliance

Firstly, relevant SDG compliance columns within the dataset are identified, encompassing a spectrum of SDG indicators and the overarching presence of SDG-guided practices. This comprehensive selection ensures that all facets of SDG adherence are captured. The cornerstone of the SDG compliance assessment is the binary classification method. A bank is considered SDG compliant if it explicitly demonstrates affirmative compliance in any of the designated SDG-specific columns or through general SDG-guided practices, provided that the bank has an existing sustainability report (indicated by a non-zero value in the 'Sustainability_Report' column). This conditional approach distinguishes between genuine SDG compliance and mere tokenism. Following this, the individual SDG-related columns used for the compliance score calculation are removed to avoid multicollinearity.

4.3.6 Data Scaling and Normalisation

Scaling and normalisation are essential preprocessing steps in data science, particularly for financial datasets with outliers and skewed distributions. They ensure ML model input variables are comparably scaled without distorting value ranges. Variables unsuitable for scaling, like those resulting in uniform zero values with the PowerTransformer, were excluded. The Yeo-Johnson method of the PowerTransformer was applied to the remaining metrics for its ability to handle both positive and negative values and stabilise variance, which is beneficial for algorithms favouring normally distributed features. RobustScaler, resilient to outliers, was used after splitting data into training and testing sets.

4.3.7 Data Preparation

This section outlines the data-splitting process, which is crucial for ML analysis. The dataset is divided into 80% training and 20% testing sets. The training set is used to train ML models to learn patterns, while the testing set, not used in training, assesses the model's ability to generalise to new data, preventing overfitting (Aslam et al., 2022 & Gillis, 2022).

4.4 Building Predictive Model

4.4.1 Support Vector Machine (SVM)

SVM is a supervised learning algorithm primarily used for binary classification. Known for its accuracy and flexibility, it operates as a 'black box,' making its decision-making process less interpretable (Barbella et al., 2009). This research uses the linear kernel for its superior performance in this context (Aslam et al., 2022). SVM excels in high-dimensional spaces, often outperforming other algorithms, and effectively generalises new data to avoid overfitting. It is efficient, scalable, and capable of non-linear classification through kernel functions (Bansal, 2022). SVM establishes a decision boundary that separates data into distinct classes, generating a hyperplane that maximises the margin for linearly separable data (Zuo & Carranza, 2011).

The mathematical underpinnings of SVM include an observation feature vector $x_i = (x_i, x_2, \dots \dots, x_n)^T$ in the training $\overline{set}$, with y_i denoting the corresponding label to x_i and w as the weight vector $= (w1, w2, \dots \dots wp)^T$ where $\left((w)\right)^2 = 1$

$$f(x_i) = w^t x_i + b \tag{1}$$

The model outcome is then specified as follows:

$$\hat{Y}_t = \begin{Bmatrix} 1 & for & f(x_i) \geq 0 \\ 0 & for & f(x_i) < 0 \end{Bmatrix} \tag{2}$$

An optimisation constraint is defined as:

$$w^T x_i + b > 0 \quad for \quad y_i = +1 w^T x_i + b < 0 \quad for \quad y_i = -1 y_i \quad (w^T \quad x_i + b) \geq 1 \tag{3}$$

Minimisation objective function:

$$\left((w)\right)^2 + \quad C \quad \sum_i \grave{\varepsilon}_i \tag{4}$$

where $\grave{\varepsilon}_i$ indicates the slack variable, while the constraints on the optimisation problem are:

$$y_i \quad (w x_i + b) \geq 1 - \quad \grave{\varepsilon}_i \quad where \quad \grave{\varepsilon}_i \quad \geq 0 \tag{5}$$

The kernel function $k\left(x_i, x_j\right)$ for non-linear SVM is defined as:

$$R_{SVM}^2 + \quad C \quad \sum_i \grave{\varepsilon}_i\left((x_i \quad - \quad a)\right)2 \quad \leq \quad R_{SVM}^2 + \quad \grave{\varepsilon}_i L = \quad \sum_i \alpha \quad (x_i \quad x_i) - \quad \sum_{i \quad j} \alpha_i \alpha_j \tag{6}$$

where $0 \leq \alpha_i \leq C$ for all i, the hypersphere centre is calculated by:

$$a = \sum_{i} \alpha_i \emptyset(xi) \tag{7}$$

4.4.2 Decision Tree (DT)

DT, a versatile supervised ML tool, excels in classification and regression tasks. They consist of internal nodes (feature tests), branches (test outcomes), and leaf nodes (classifications/predictions). As data increases, a DT visually grows in complexity, resembling a tree's branching structure (Bansal, 2022). DTs, like the Classification and Regression Tree (CART) algorithm, are adaptable, effectively managing numerical and categorical data (Sharma, 2021). They identify key factors, discard irrelevant features, and handle outliers/missing values. Their interpretability, simplicity, and proficiency in high-dimensional scenarios make them suitable for various learning tasks (Singh & Giri, 2014).

DTs operate by recursively splitting the dataset into subsets based on the most informative attribute, guided by entropy and information gain criteria. When a subset's observations belong to a single class, the node is labelled accordingly. This process continues until all nodes are labelled, creating a comprehensive model for classification or regression (Paul et al., 2018).

The core mechanism of SVM involves establishing a decision boundary that separates data into distinct classes. For linearly separable data, SVM generates a hyperplane that maximises the margin, defined as the shortest distance to the nearest data points on either side of the hyperplane (Zuo & Carranza, 2011).

$$Entropy(S) = -\sum_{x \in A} P(x) \, log_2 \, P(x) \tag{8}$$

$$Gain(S, A) = Entropy(S) - \sum_{x \in A} \frac{|S_v|}{|s|} \times Entropy(S_v) \tag{9}$$

Calculating information gain, which reflects entropy changes in relation to an independent feature, is a critical step in this process. By repeating this computation at each node, the DT model develops a nuanced understanding of the data, enabling accurate predictions of new observations based on their characteristics.

4.4.3 K-Nearest Neighbours (KNN)

KNN is a supervised, non-parametric ML algorithm used in instance-based learning. The algorithm retains training data and predicts new data points based on similarity to the nearest examples. KNN is adept at image classification, object identification, and anomaly detection but can be computationally demanding with large or high-dimensional datasets (Zhang, 2016). Its simplicity and intuitive classification based on proximity to neighbours make them easy to implement (Liu & Zhang, 2012).

KNN, without assuming data distribution, handles complex and non-linear data (Li & Zhang, 2014). Its resilience to noisy or irrelevant data and adaptability to classification and regression tasks make it a versatile tool (Liu & Zhang, 2012; Alam et al., 2019).

The workings of KNN involve computing the distance between an unlabelled data point and its neighbours in the dataset. The most common distance measure is the Euclidean distance, defined as:

$$D(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2} \tag{10}$$

where p and q represent the subjects being compared across n attributes (Cost & Salzberg, 1993). Alternative distance measures, such as the Manhattan distance, are also viable options:

$$d(x,y) = \left(\sum_{i=1}^{m} |x_i - y_i|\right) \tag{11}$$

The choice of k, the number of nearest neighbours is crucial for the algorithm's performance. While a higher k value can diminish the effects of random noise; it may also obscure significant patterns. Adjusting the weighting of each neighbour's influence on the prediction is another approach to enhancing the algorithm's effectiveness.

4.4.4 Extra Trees (ET)

ET, also known as Extremely Randomised Trees, is an ML algorithm derived from random forests designed to reduce the risk of overfitting. Like RF, ET employs a random subset of features to train each base estimator. However, ET distinguishes itself in its approach to constructing DT. During training, DT within the Extra Trees Forest is generated by segmenting data at each node based on pre-specified mathematical criteria, followed by a random selection from the remaining features, as detailed by Sharma et al. (2022).

ET's feature selection process, guided by the forest structure's unique characteristics, is referred to as "The Featured Gini Importance" (Dsouza & Ansari, 2022). This technique simplifies the mathematical computations for determining split features, allowing users to prioritise features based on their Gini values (Dsouza & Ansari, 2022).

The Gini impurity method is the default criterion in the classification tasks using ETs. Conversely, the algorithm typically employs mean absolute and squared error metrics for regression tasks. The relevant formulas are shown in Equations 12 and 13.

$$Gini\ Impurity = \sum_{j=1}^{O} f_j \left(1 - f_j\right) \tag{12}$$

$$Entropy = \sum_{j=1}^{O} -f_j \quad log\left(f_j\right) \tag{13}$$

Here, f_j represents the frequency of label j at a node, and O denotes the number of unique labels.

4.4.5 Gradient Boosting (GB)

GB is a method formulated to tackle convex optimisation problems sequentially, facilitating the development of models that linearly represent basic predictors (Biau & Cadre, 2021). This approach aims to construct a generalisable model capable of optimising a range of arbitrary loss functions. Key to this method is the application of the gradient-descent algorithm, which is instrumental in reducing the loss in new tree additions. GB proves particularly effective in both classification and regression contexts.

The essence of GB lies in creating a cohesive ensemble of models, as follows:

$$f(x) = \sum_{m=0}^{M} f_m(x) \tag{14}$$

This can be conceptualised as an adaptive function model:

$$f(x) = \theta_0 + \sum_{m=0}^{M} \theta_m \emptyset_m(x) \tag{15}$$

Here, the function $f(x) = \theta_0, f_m(x) = \theta_m \emptyset_m(x)$ for $m = 1, \ldots, M$, where $\emptyset_m$ represents a basis function incrementally added to improve the current model's fit.

In the GB process, the goal is to minimise prediction errors by integrating the best new model with existing ones. This technique adjusts target outcomes from previous models to reduce errors further. Customising the GB function across different loss functions allows it to suit various hyperparameters, as Hassan et al. (2023) detailed.

4.4.6 Random Forest (RF)

The RF algorithm has established itself as a powerful tool in geomorphology and remote sensing, widely applied due to its ability to enhance geomorphic and susceptibility mapping techniques. Its speed and ease of parameter optimisation have garnered significant interest from the scientific community (Akinci et al., 2020). RF's versatility allows it to perform feature selection, classification, and regression simultaneously.

In RF, the training dataset is divided into in-bag (training) and out-of-bag (validation) samples. Each tree within the forest is formed using bootstrap sampling (Kavzoglu & Teke, 2022). The algorithm then aggregates individual tree predictions based on a majority vote to determine the final predictive outcome. This process is encapsulated as follows:

$$H(x) = arg_z^{max} \sum_{i=1}^{k} I(h_i(x) = Z) \tag{16}$$

In the model, $H(x)$, each decision tree is denoted as h_i. The output variable is Z, and $I(.)$ represents the indicative function, which integrates the outputs of the individual trees (h_i) to determine the final output (Z) for a given input (x).

RF construction involves two key steps. Firstly, using the bagging sampling technique, N training subsets are created from the original training set, each containing roughly a third of the total samples. Secondly, N decision trees are formed to build the random forest. Each tree is based on a separate subset, developed using the CART algorithm, which employs the Gini coefficient principle. This principle minimises misclassification by assigning objects to classes based on estimated probabilities. The Gini coefficient, which is used to measure the probability of misclassification, is calculated using Equation 17, where $p(i|t)$ is the probability of an object being assigned to class i and $p(i|t)$ is the actual probability of it belonging to class j at a given node t.

$$Gini = i \neq jJp(i|t)p(j|t) \tag{17}$$

4.5 Evaluating Predictive Models

Table 5: Model's Evaluation Metrics Formulas

Evaluation metric	**Formula**
Precision	$P = \frac{TP}{TP + FP}$
Recall	$R = \frac{TP}{TP + FN}$
F-measure	$F = \frac{(1 + \beta^2) \times Precision \times Recall}{(\beta^2 \times Precision) + Recall}$
Accuracy	$ACC = \frac{TP + TN}{TP + FP + TN + FN}$

In assessing the performance of various predictive models, this study employs a set of established evaluation metrics (Hanczar & Nadif, 2013; Manning et al., 2008; Powers, 2011). The specific formulas for these evaluation metrics are outlined in Table 5.

4.6 Analysing Feature Importance and SDG Compliance Impact

4.6.1 Methodology for Feature Importance Analysis

For each model, except SVM and KNN, the 'feature_importances_' attribute was used to represent the importance of each feature in the model's decision-making process. Higher values indicate more significant features. SVM and KNN were excluded from feature importance analysis due to their limitations: SVM, especially with non-linear kernels, lacks a direct method for interpreting feature importance, and KNN, based on feature space distances, does not assess feature significance. This approach ensures a robust and interpretable analysis using the most insightful models.

4.6.2 Methodology for Partial Dependence Analysis

Partial Dependence Plots (PDPs) were used to elucidate the relationship between SDG compliance and financial performance tiers. Using all six ML models, PDPs for the

'SDG_Compliance' feature were generated across all performance tiers with the 'PartialDependenceDisplay. from_estimator' method in 'sklearn.inspection.' This provided visual insights into the effect of SDG compliance on the probability of classification into each tier while holding other features constant. The analysis was replicated for each model to assess the consistency and robustness of the 'SDG_Compliance' effect across different predictive frameworks.

4.7 Experiment of Machine Learning (ML) Algorithms

4.7.1 Support Vector Machine (SVM)

The SVM model utilised in this study to predict Malaysian bank financial performance tiers demonstrates a nuanced level of effectiveness. While a training accuracy of 51.25% suggests a moderate ability to classify the training data correctly, the 50% test accuracy indicates limited generalisability, as it performs no better than random guessing on unseen data.

Tables 6 and 7 indicate that the model excels in identifying 'Top' tier banks, correctly predicting all instances, but fails to recognise any 'High' tier banks. Performance for 'Medium' and 'Low' tiers is mixed, showcasing some capability but lacking high precision.

Table 6: Confusion Matrix

ML Models	Test Dataset Confusion Matrix				
		Actual Top	Actual High	Actual Medium	Actual Low
SVM	Predicted Top	4	0	0	0
	Predicted High	1	0	2	1
	Predicted Medium	1	0	1	4
	Predicted Low	0	0	1	5
DT	Predicted Top	3	1	0	0
	Predicted High	1	1	1	1
	Predicted Medium	1	2	2	1
	Predicted Low	0	0	0	6
KNN	Predicted Top	4	0	0	0
	Predicted High	1	0	2	1
	Predicted Medium	1	0	1	4
	Predicted Low	0	0	2	4
ET	Predicted Top	3	1	0	0
	Predicted High	0	2	1	1
	Predicted Medium	0	1	3	2
	Predicted Low	0	0	1	5
GB	Predicted Top	3	1	0	0
	Predicted High	0	2	1	1
	Predicted Medium	0	0	5	1
	Predicted Low	0	0	0	6
RF	Predicted Top	3	1	0	0
	Predicted High	1	1	1	1
	Predicted Medium	0	1	4	1
	Predicted Low	0	0	1	5

Table 7: Metric of Support Vector Machine

	Top	High	Medium	Low
Precision	0.67	0.00	0.25	0.50
Recall	1.00	0.00	0.17	0.83
F1-measure	0.80	0.00	0.20	0.62

Precision, recall, and F1 scores further illuminate the model's performance. The 'Top' tier boasts high precision (0.67) and perfect recall (1.00), resulting in a strong F1-score (0.80). The 'Low' tier shows decent precision (0.50) and high recall (0.83), yielding a respectable F1-score (0.62). However, performance in the ‘High’ and ‘Medium’ tiers is weak, with low scores across all three metrics.

4.7.2 Decision Tree (DT)

The DT model performs strongly in classifying bank financial performance tiers, with a training accuracy of 78.75%. However, test accuracy is lower at 60%, indicating a need to improve the generalisability of new data.

Tables 6 and 8 reveal varied success across tiers. The model excels in the ‘Low’ category, predicting all instances correctly, and performs well in the ‘Top’ tier. However, its performance is less consistent for the ‘High’ and ‘Medium’ tiers.

Table 8: Metric of Decision Tree

	Top	High	Medium	Low
Precision	0.60	0.25	0.67	0.75
Recall	0.75	0.25	0.33	1.00
F1-measure	0.67	0.25	0.44	0.86

Precision, recall, and F1 scores further detail the model’s effectiveness. The ‘Low’ category achieves high precision (0.75) and perfect recall (1.00), leading to an impressive F1-score (0.86). The ‘Top’ category also performs well with decent precision and recall, resulting in a good F1 score. However, the ‘High’ and ‘Medium’ categories show lower scores, particularly in precision and recall for the ‘High’ tier.

4.7.3 K-Nearest Neighbours (KNN)

The KNN model, while showing moderate accuracy (56.25%) on training data, demonstrated reduced effectiveness on unseen test data with an accuracy of 45%. This suggests a potential overfitting issue and limited generalisability to new data.

Notably, the model excels in identifying ‘Top’ tier banks, correctly predicting all instances, resulting in high precision (0.67) and perfect recall (1.00). However, its performance is inconsistent across other categories.

Table 9: The metric of K-Nearest Neighbours

	Top	High	Medium	Low
Precision	1.00	0.00	0.40	0.86
Recall	1.00	0.00	0.50	0.86
F1-measure	1.00	0.00	0.44	0.86

The model achieves moderate accuracy in the ‘Low’ category (precision of 0.44, recall of 0.67). The ‘Medium’ tier shows lower performance (precision of 0.20, recall of 0.17), and the model fails to predict any ‘High’ tier banks correctly.

4.7.4 Extra Trees (ET)

The ET classifier exhibited perfect accuracy (100%) on training data but demonstrated a decrease in performance on unseen test data with an accuracy of 65%, highlighting the need for improved generalisation.

The model excelled in predicting ‘Top’ and ‘Low’ tier banks, achieving a precision of 100% and a recall of 75% from the ‘Top’ tier, and a precision of 0.62 and a recall of 0.83 for the ‘Low’ tier. However, the model showed moderate performance in the ‘High’ category, with a precision and recall of 0.50 each, and in the ‘Medium’ category, it achieved a precision of 0.60 and a recall of 0.50.

Table 10: Metric of Extra Trees

	Top	High	Medium	Low
Precision	1.00	0.50	0.60	0.62
Recall	0.75	0.50	0.50	0.83
F1-measure	0.86	0.50	0.55	0.71

4.7.5 Gradient Boosting (GB)

The GB classifier performs exceptionally in the training and testing phases, achieving 100% and 80% accuracy, respectively. This demonstrates the model's ability to generalise effectively to unseen data, indicating minimal overfitting.

The model excels in the 'Medium' and 'Low' categories, with high precision and recall, accurately identifying most banks within these tiers. While demonstrating robust performance in the 'Top' category (precision of 100%, recall of 75%), the model shows moderate accuracy in the "High" category (precision of 67%, recall of 50%).

Table 11: Metric of Gradient Boosting

	Top	High	Medium	Low
Precision	1.00	0.67	0.83	0.75
Recall	0.75	0.50	0.83	1.00
F1-measure	0.86	0.57	0.83	0.86

4.7.6 Random Forest (RF)

The RF classifier achieved perfect training accuracy but a lower test accuracy of 65%, suggesting potential overfitting and reduced generalisability to new data.

The model performs well in the 'Top' and 'Medium' categories, demonstrating a balanced precision and recall of 75% for 'Top' and 67% for 'Medium.' However, it faces challenges with the 'High' category, with a precision of 33% and recall of 25%. Conversely, it excels in the 'Low' category, with high precision (71%) and recall (83%).

Table 12: Metric of Random Forest

	Top	High	Medium	Low
Precision	0.75	0.33	0.67	0.71
Recall	0.75	0.25	0.67	0.83
F1-measure	0.75	0.29	0.67	0.77

4.8 Model Comparison

Table 13 summarises the performance of various models in classifying bank financial performance tiers. The GB model is the top performer with 80% test accuracy, demonstrating robust classification ability across tiers, particularly in the 'Medium' and 'Low' categories.

Table 13: Summary of Accuracy Rate for Each Machine Learning Model

Model	Test Accuracy
Gradient Boosting (GB)	0.80
Random Forest (RF)	0.65
Extra Trees (ET)	0.65
Decision Tree (DT)	0.60
Support Vector Machine (SVM)	0.50
K-Nearest Neighbours (KNN)	0.45

The DT model shows moderate effectiveness with 60% test accuracy, indicating balanced performance across tiers, especially in the 'Low' and 'Top' categories, but with room for improved predictive accuracy.

The ET and RF classifiers both achieve 65% test accuracy. ET shows promise in distinguishing banks across tiers, particularly in the 'Low' and 'Medium' categories, while RF, despite excellent training accuracy, may suffer from overfitting. Both models indicate the potential for further refinement.

The SVM model, with 50% test accuracy, and the KNN model, with 45% accuracy, highlight challenges in generalising to unseen data. The SVM model struggles particularly in the ‘High’ and ‘Medium’ tiers, while the KNN model shows limitations across all tiers, especially in the ‘High’ category. Both models require significant optimisation to enhance their predictive capabilities.

4.9 Analysis of Feature Importance and SDG Compliance Impact

4.9.1 Feature Importance Analysis

A nuanced analysis of feature importance unveils a multi-faceted landscape of factors influencing the financial performance tiers of Malaysian banks. Market capitalisation emerges as a consistent and dominant force, reflecting the significant influence of market valuation on financial standing. Valuation ratios, remarkably P/E and price-to-book ratios, stand alongside profitability metrics like ROE as primary drivers, showcasing the industry’s focus on financial health and potential growth. While models recognise ‘SDG_Compliance’, it sits on the periphery of prediction, suggesting a secondary role compared to traditional financial indicators within the current context. This nuanced understanding compels a critical reassessment of how sustainability metrics are prioritised within the banking sector’s strategic framework.

4.9.2 Partial Dependence Analysis (PDA) on SDG Compliance

The PDA across various ML models demonstrates diverse interactions between SDG compliance and financial performance tiers. It indicates how changes in the feature of interest affect the predicted outcome while other features are held constant. For GB, an upward trend for the ‘Top’ and ‘Low’ tiers suggests a positive relationship with SDG compliance, whereas the ‘High’ and ‘Medium’ tiers exhibit a downward trend, indicating a potential inverse relationship. RF and ET models show a similar pattern, with SDG compliance potentially bolstering the ‘High’ and ‘Low’ tiers but not ‘Top’ and ‘Medium’.

The DT, SVM, and KNN models’ horizontal lines across all tiers suggest no discernible influence of SDG compliance on performance tier classification. This heterogeneity in trends underscores the complexity of the relationship between SDG adherence and financial outcomes, warranting a deeper investigation into the underlying mechanisms.

5.0 Discussion

5.1 Impact on Poverty Reduction, Climate Action, and Social Equality

The correlation analysis revealed moderate positive associations between SDG 1 and key financial metrics such as payout ratio and total yield. Notably, the GB model, which boasts the highest accuracy in classifying banks according to financial performance, indicates that SDG 1 compliance might positively influence banks in the 'Medium' and 'Low' tiers. This suggests that banks engaging in poverty alleviation could see improved financial outcomes, creating a virtuous cycle where better financial health fuels further social initiatives. However, for 'High' tier banks, factors like market capitalisation and valuation ratios appear more influential in determining financial success.

Regarding climate action, a strong positive correlation between SDG 13 compliance and Enterprise Value suggests that investors highly value proactive environmental stewardship. This bolsters a bank's financial standing and minimises vulnerability to environmental risks, making climate-focused strategies both an ethical and economically prudent choice.

The PDA underscores the varied impact of SDG compliance across different bank tiers and models. For instance, the GB model suggests a financial uplift for banks in the 'Top' and 'Low' tiers with increased SDG adherence, whereas the effects are less definitive for banks in the 'High' and 'Medium' tiers. This variability could stem from how banks implement diverse SDGs (e.g., SDGs 5, 10, 11) to promote social equality, with effectiveness likely varying based on each bank's specific strategies and execution. These findings underline the complexity of the relationships between SDG compliance and financial performance, emphasising the need for more detailed research to decipher the underlying mechanisms.

5.2 Comparison with Past Studies and Theoretical Implications

The research findings reveal a complex interplay between SDG compliance, financial performance, and macroeconomic factors in Malaysian banks, echoing the nuanced relationships observed in the broader literature. Consistent with concerns raised in prior research (Buallay, 2019), the varying impact of ESG disclosures on financial performance indicators like ROA, ROE, and Tobin's Q underscores the critical need for rigorous data management in sustainability studies to ensure accurate and reliable results.

Building on this foundation, this study correlation matrix emphasises the stability of short-term bank valuations in Malaysia. The strong correlation between market capitalisation and enterprise value further substantiates market capitalisation as a robust financial indicator, echoing findings from global studies (Ameer & Othman, 2011). This is consistent with signalling theory (Spence, 1973), where market capitalisation serves as a signal of firm value to investors. Additionally, the significant correlations between profitability metrics (e.g., ROA) are in line with the broader literature linking financial performance to sustainability (Brogi & Lagasio, 2018).

This study uses ML analysis to delve deeper into these relationships, utilising diverse models. The superior accuracy of the GB model identifies market capitalisation, valuation ratios (P/E, price-to-book), and ROE as critical predictors of financial performance tiers, thus aligning with information asymmetry theory (Akerlof, 1970) and underscoring how investors rely on traditional financial signals amidst information disparities.

However, the marginal importance of the aggregated 'SDG_Compliance' feature in most models contrasts with this study's initial expectations. This aligns with the mixed findings in the broader literature on the SDG-financial performance nexus, particularly in the Asian context (Burhan & Rahmanti, 2012; Laskar & Maji, 2018). This suggests that the relationship between SDG compliance and financial performance is likely context-dependent and not always straightforward.

PDA reveals intriguing nuances that shed light on this complexity. The upward trend for the 'Top' and 'Low' tiers in GB, RF, and ET models suggests a potential positive association between SDG compliance and these tiers, aligning with stakeholder theory (Agudo-Valiente et al., 2015), where addressing societal concerns can enhance long-term financial performance. Conversely, downward trends in the ‘High’ and ‘Medium’ tiers for these models hint at potential short-term trade-offs between SDG investments and profitability, echoing Forcadell and Aracil (2017). Notably, the absence of clear trends in DT, SVM, and KNN models further underscores the complexity of this relationship, similar to the context-dependent findings of Pham et al. (2021).

This heterogeneity in PDA trends aligns with this study's conceptual framework, suggesting multiple pathways through which SDG compliance can influence financial performance. The varying impact on different tiers could be attributed to the specific SDGs pursued, the banks’ characteristics (Doğan & Kevser, 2021), or the dynamics of the banking industry. Future research should investigate these nuances to understand the underlying mechanisms.

In conclusion, this study of ML analysis contributes a novel perspective to the ongoing discourse, confirming the importance of traditional financial indicators while revealing the non-linear and context-dependent nature of SDG compliance’s impact on financial outcomes. This aligns with Lassala et al.’s (2021) findings on the long-term benefits of sustainability, underscoring the need for policymakers, investors, and banking professionals to reassess the role of sustainability in shaping financial success. It also highlights the need for further context-specific research in the Malaysian banking sector.

6.0 Conclusion

6.1 Summary

The analysis reveals that financial performance in Malaysian banks is influenced by multiple factors, with market capitalisation and traditional financial health indicators playing key roles. While SDG compliance is significant, its impact varies across performance tiers.

The GB model, with 80% accuracy, uncovered the complex relationship between financial performance and SDG compliance, especially in non-linear data. Ensemble models like RF and ET showed a moderate impact, while DT, SVM, and KNN indicated that SDG compliance influence is not always straightforward.

Correlation analysis highlighted that low-compliance SDGs (6, 13, 14) have positive correlations with financial metrics, suggesting potential financial improvements through strategic focus. Conversely, highly-complied SDGs (8, 12) show potential negative correlations, necessitating a strategic reassessment to balance ethical and financial sustainability.

The study underscores the importance of aligning SDG initiatives with financial goals, particularly in areas where compliance has a significant impact. This alignment can help banks contribute to global sustainability while ensuring financial sustainability.

6.2 Limitations

This study faces limitations, particularly in obtaining specific SDG compliance data, which is challenging and costly. Relying on official reports from listed banks may miss insights from non-listed institutions and bias towards larger banks. The focus on publicly listed Malaysian banks limits their broader applicability to private banks or different regions. Aggregating SDG compliance metrics into a composite score may not capture the nuanced impacts of each SDG. Future research should use more disaggregated analyses and diverse SDG metrics to assess distinct impacts. Exploring advanced models like time series or hybrid approaches could better capture the dynamic nature of SDG compliance and financial performance. Potential biases from imbalanced datasets necessitate cautious interpretation, highlighting the need for diverse, high-quality data in future research.

6.3 Recommendations

To better understand the link between banks' sustainability efforts and financial performance, partnerships with ESG data providers or in-house frameworks are recommended. This approach would provide detailed insights into banks' SDG engagement, revealing how specific initiatives impact financial outcomes. Encouraging third-party data providers to integrate comprehensive SDG data with financial metrics would enhance investment decisions and sustainability reporting.

To improve data quality, banks should minimise missing values and inaccuracies in sustainability reports by including detailed metrics for specific targets. This granularity allows for deeper analysis and targeted strategies. Exploring a combination of established ML models like RF, with advanced techniques such as time series analysis and Artificial Intelligence architectures could uncover hidden patterns in SDG compliance data, further clarifying its relationship with financial performance.

Establishing standardised SDG reporting frameworks within the banking sector is crucial for consistency and comparability, fostering robust analysis and benchmarking. Incentivising SDG compliance through regulations, financial rewards, or preferential treatment can accelerate sustainable banking practices. Policy interventions like regulatory changes, tax incentives, or sustainability-linked financial products can significantly influence banks' focus on SDGs.

Future research should explore the impact of diverse SDG performance metrics on financial outcomes, as different measurement methodologies offer varied insights. Continuous monitoring and refinement of ML models with new data and insights are essential to maintain their relevance and accuracy. Collaboration with academia, industry experts, and policymakers is crucial for developing a long-term research roadmap to understand the evolving relationship between SDG compliance and financial performance.

Author Contributions Statement: The authors worked together for this paper. Conceptualization: K.S.L, K.W.K; Methodology: K.S.L, K.W.K, X.Y.C, W.C.N; Formal analysis and investigation: K.S.L, K.W.K, X.Y.C, W.C.N; Writing - original draft preparation: K.S.L, K.W.K; Writing - review and editing: K.W.K, X.Y.C, W.C.N; Resources: K.W.K, X.Y.C; Supervision: K.W.K, X.Y.C. All authors have read and agreed to the published version of the manuscript.

Funding Statement: No funding was received to assist with the preparation of this manuscript.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgement: Special thanks were extended to the School of Management, Universiti Sains Malaysia.

Conflict of Interest Statement: The authors have no competing interests to declare that are relevant to the content of this study.

References

Adeoye, O. B., Ani, E. C., Ninduwesuor-Ehiobu, N., Montero, D. J., Usman, F. O., & Olu-lawal, K. A. (2024). A conceptual framework for data-driven sustainable finance in green energy transition. *World Journal of Advanced Research and Reviews, 21*(02), 1791–1801. https://doi.org/10.30574/wjarr.2024.21.2.0620

Adu, D. A. (2022). Sustainable banking initiatives, environmental disclosure and financial performance: The moderating impact of corporate governance mechanisms. *Business Strategy and the Environment, 31*(5), 2365–2399. https://doi.org/10.1002/bse.3033

Agudo-Valiente, J. M., Garcés-Ayerbe, C., & Salvador-Figueras, M. (2015). Corporate social performance and stakeholder dialogue management. *Corporate Social Responsibility and Environmental Management, 22*(1), 285-304. https://doi.org/10.1002/csr.1324

Akerlof, G. A. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics, 84*(3), 488–500. https://doi.org/10.2307/1879431

Akinci, H., Kilicoglu, C., & Dogan, S. (2020). Random forest-based landslide susceptibility mapping in coastal regions of Artvin, Turkey. *International Journal of Geo-Information, 9*(9), 553. https://doi.org/10.3390/ijgi9090553

Al Lawati, H., & Hussainey, K. (2022). Does Sustainable Development Goals disclosure affect corporate financial performance? *Sustainability, 14*(13), Article 7815. https://doi.org/10.3390/su14137815

Alam, T., Ahmed, C. F., Zahin, S. A., Khan, M. A. H., & Islam, M. T. (2019). An effective recursive technique for multi-class classification and regression for imbalanced data. *IEEE Access, 7*, 127615–127630. https://doi.org/10.1109/ACCESS.2019.2939755

Alshehhi, A., Nobanee, H., & Khare, N. (2018). The impact of sustainability practices on corporate financial performance: Literature trends and future research potential. *Sustainability, 10*(2), Article 494. https://doi.org/10.3390/su10020494

Ameer, R., & Othman, R. (2011). Sustainability practices and corporate financial performance: A study based on the top global corporations. *Journal of Business Ethics, 108*, 61–79. https://doi.org/10.1007/s10551-011-1063-y

Aslam, F., Hunjra, A. I., Ftiti, Z., Louhichi, W., & Shams, T. (2022). Insurance fraud detection: Evidence from artificial intelligence and machine learning. *Research in International Business and Finance, 62*, Article 101744. https://doi.org/10.1016/j.ribaf.2022.101744

Bahl, K., Kiran, R., & Sharma, A. (2023). Scaling up banking performance for the realisation of specific Sustainable Development Goals: The interplay of digitalisation and training in the transformation journey. *Sustainability, 15*(18), Article 13798. https://doi.org/10.3390/su151813798

Bank Negara Malaysia. (2022). *Climate risk management and scenario analysis.* https://www.bnm.gov.my/documents/20124/938039/PD_Climate-Risk-Mgmt-Scenario-Analysis-Nov2022.pdf

Bank Negara Malaysia. (2023). *Annual Report 2022.* https://www.bnm.gov.my/documents/20124/10150308/ar2022_en_book.pdf

Bank Negara Malaysia. (n.d.). *Climate change.* Retrieved November 21, 2023, from https://www.bnm.gov.my/climatechange

Bansal, M., Goyal, A., & Choudhary, A. (2022). A comparative analysis of K-nearest neighbor, genetic, support vector machine, decision tree, and long short term memory algorithms in machine learning. *Decision Analytics Journal, 3*, Article 100071. https://doi.org/10.1016/j.dajour.2022.100071

Barbella, D., Benzaid, S., Christensen, J., Jackson, B., Qin, X. V., & Musicant, D. (2009, July 13–16). *Understanding support vector machine classifications via a* *recommender system-like approach* [Paper presentation]. Proceedings of the 2009 International Conference on Data Mining, Las Vegas, NV, United States.

Beal, D., Reeves, K. N., Husain, Z., & John, K. (2022, December 14). *Banks can bet big on social impact.* Boston Consulting Group. https://www.bcg.com/publications/2022/role-of-social-impact-banking-in-society

Biau, G., & Cadre, B. (2021). Optimisation by gradient boosting. In A. Daouia & A. Ruiz-Gazen (Eds.), *Advances in Contemporary Statistics and Econometrics* (pp. 23–44). Springer, Cham. https://doi.org/10.1007/978-3-030-73249-3_2

Brogi, M., & Lagasio, V. (2018). Environmental, social, and governance and company profitability: Are financial intermediaries different? *Corporate Social Responsibility and Environmental Management, 26*(3), 576–587. https://doi.org/10.1002/csr.1704

Buallay, A. (2019). Is sustainability reporting (ESG) associated with performance? Evidence from the European banking sector. *Management of Environmental Quality, 30*(1), 98–115. https://doi.org/10.1108/meq-12-2017-0149

Buallay, A., Fadel, S. M., Alajmi, J., & Saudagaran, S. (2021). Sustainability reporting and bank performance after financial crisis: Evidence from developed and developing countries. *Competitiveness Review, 31*(4), 747–770. https://doi.org/10.1108/cr-04-2019-0040

Burhan, A. H. N., & Rahmanti, W. (2012). The impact of sustainability reporting on company performance. *Journal of Economics, Business, and Accountancy Ventura, 15*(2), 257–272. https://doi.org/10.14414/jebav.v15i2.79

Bursa Malaysia Berhad. (2021, April 27). *Integrating SDGs into sustainability reporting.* https://bursasustain.bursamalaysia.com/droplet-details/sustainability/integrating-sdgs-into-sustainability-reporting

Bursa Malaysia Securities Berhad. (2015). *Sustainability Reporting Guide.* https://www.bursamalaysia.com/sites/5bb54be15f36ca0af339077a/content_entry

5ce3b5005b711a1764454c1a/5ce3c83239fba2627b286508/files/bursa_malaysia_sustainability_reporting_guide-final.pdf?1570701456

BusinessToday. (2022, February 4). A*re Malaysian banks ESG compliant?* https://www.businesstoday.com.my/2022/02/04/are-malaysian-banks-esg-compliant/

Chang, A., Lee, T.-S., & Lee, H.-M. (2024). Applying sustainable development goals in financial forecasting using machine learning techniques. *Corporate Social Responsibility and Environmental Management, 31*(3), 2277–2289. https://doi.org/10.1002/csr.2694

Cheong, A. (2020, November 13). *RHB Banking Group pledges RM5 billion in support of green financing by 2025.* RinggitPlus. https://ringgitplus.com/en/blog/bank-news/rhb-banking-group-pledges-rm5-billion-in-support-of-green-financing-by-2025.html

Chu, M. M., & Donovan, K. (2020, December 8). *Malaysia's CIMB commits to phase out coal financing by 2040.* Reuters. https://www.reuters.com/article/uk-malaysia-cimb-idUKKBN28I0X0/

Clementino, E., & Perkins, R. (2021). How do companies respond to Environmental, Social and Governance (ESG) ratings? Evidence from Italy. *Journal of Business Ethics, 171*, 379–397. https://doi.org/10.1007/s10551-020-04441-4

Cost, S., & Salzberg, S. (1993). A weighted nearest neighbor algorithm for learning with symbolic features. *Machine Learning, 10*, 57–78. https://doi.org/10.1023/A:1022664626993

Decker, O. S., & Kingdom, O. J. (2020). Banks and the Sustainable Development Goals. In S. Idowu, R. Schmidpeter, N. Capaldi, L. Zu, M. D. Baldo, & R. Abreu (Eds.), *Encyclopedia of Sustainable Management* (pp. 1–9). Springer, Cham. https://doi.org/10.1007/978-3-030-02006-4_526-1

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review, 86* (1): 59–100. https://doi.org/10.2308/accr.00000005

Doğan, M., & Kevser, M. (2021). Relationship between sustainability report, financial performance, and ownership structure: Research on the Turkish banking sector. *Istanbul Business Research, 50*(1), 77–102. https://doi.org/10.26650/ibr.2021.51.0115

Doko, F., Kalajdziski, S., & Mishkovski, I. (2021). Credit risk model based on central bank credit registry data. *Journal of Risk and Financial Management, 14*(3), Article 138. https://doi.org/10.3390/jrfm14030138

Dowling, J., & Pfeffer, J. (1975). Organisational legitimacy: Social values and organisational behavior. *The Pacific Sociological Review, 18*(1), 122–136. https://doi.org/10.2307/1388226

Dsouza, K. J., & Ansari, Z. A. (2022). Histopathology image classification using hybrid parallel structured Deep-CNN models. *Applied Computer Science, 18*(1), 20–36. https://doi.org/10.35784/acs-2022-2

Ellili, N. O. D., & Nobanee, H. (2022). Impact of economic, environmental, and corporate social responsibility reporting on financial performance of UAE banks. *Environment, Development and Sustainability, 25*, 3967–3983. https://doi.org/10.1007/s10668-022-02225-6

Forcadell, F. J., & Aracil, E. (2017). European banks' reputation for corporate social responsibility. *Corporate Social Responsibility and Environmental Management, 24*(1), 1–14. https://doi.org/10.1002/csr.1402

Freeman, R. E. (2010). *Strategic management: A stakeholder approach*. Cambridge University Press.

Gillis, A. S. (2022). *Data splitting*. TechTarget. Retrieved November 21, 2023, from https://www.techtarget.com/searchenterpriseai/definition/data-splitting

Girón, A., Kazemikhasragh, A., Cicchiello, A. F., & Panetti, E. (2020). Sustainability reporting and firms' economic performance: Evidence from Asia and Africa. *Journal of the Knowledge Economy, 12*, 1741–1759. https://doi.org/10.1007/s13132-020-00693-7

Global Banking & Finance Review. (n.d.). *UN's Sustainable Development Goals: The critical role of banks in creating a sustainable future*. https://www.globalbankingandfinance.com/uns-sustainable-development-goals-the-critical-role-of-banks-in-creating-a-sustainable-future/

Hanczar, B., & Nadif, M. (2013, June 16–21). *Precision-recall space to correct external indices for biclustering* [Conference session]. Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, United States. https://proceedings.mlr.press/v28/hanczar13.pdf

Hassan, M. M., Hassan, M. M., Yasmin, F., Khan, M. A. R., Zaman, S., Galibuzzaman, Islam, K. K., & Bairagi, A. K. (2023). A comparative assessment of machine learning algorithms with the Least Absolute Shrinkage and Selection Operator for breast cancer detection and prediction. *Decision Analytics Journal, 7*, Article 100245. https://doi.org/10.1016/j.dajour.2023.100245

Huang, C., Chen, M., & Wang, C. (2007). Credit scoring with a data mining approach based on support vector machines. *Expert Systems with Applications, 33*(4), 847–856. https://doi.org/10.1016/j.eswa.2006.07.007

Ighosewe, E. F. (2021). Corporate sustainability disclosure and the Nigerian industrial/consumer goods sector's performance: A panel data approach. *International Journal of Research and Innovation in Social Science, 5*(3), 622–627.

Iqbal, S., & Nosheen, S. (2023). Moderating impact of non-performing loans on the relationship between sustainable development goals and the financial performance of bank. *Future Business Journal, 9,* Article 46. https://doi.org/10.1186/s43093-023-00224-1

Jasni, N. S., & Yusoff, H. (2021). Corporate environmental social governance and financial performance: A cross-sector risk analysis in Malaysia. *International Journal of Academic Research in Business and Social Sciences, 11*(12), 956–975. http://dx.doi.org/10.6007/IJARBSS/v11-i12/11834

Kavzoglu, T., & Teke, A. (2022). Predictive performances of ensemble machine learning algorithms in landslide susceptibility mapping using random forest, extreme gradient boosting (XGBoost) and natural gradient boosting (NGBoost). *Arabian Journal for Science and Engineering, 47*, 7367–7385. https://doi.org/10.1007/s13369-022-06560-8

Khan, M. H.-U.-Z., Islam, M. A., Fatima, J. K., & Ahmed, K. (2011). Corporate sustainability reporting of major commercial banks in line with GRI: Bangladesh evidence. *Social Responsibility Journal, 7*(3), 347–362. https://doi.org/10.1108/17471111111154509

Khan, P. A., Johl, S. K., & Akhtar, S. (2022). Vinculum of Sustainable Development Goal practices and firms' financial performance: A moderation role of green innovation. *Journal of Risk and Financial Management, 15*(3), 96. https://doi.org/10.3390/jrfm15030096

Kumar, K., & Prakash, A. (2019). Developing a framework for assessing sustainable banking performance of the Indian banking sector. *Social Responsibility Journal, 15*(5), 689–709. https://doi.org/10.1108/srj-07-2018-0162

Laskar, N., & Maji, S. G. (2018). Disclosure of corporate sustainability performance and firm performance in Asia. *Asian Review of Accounting, 26*(4), 414–443. https://doi.org/10.1108/ARA-02-2017-0029

Lassala, C., Orero-Blat, M., & Ribeiro-Navarrete, S. (2021). The financial performance of listed companies in pursuit of the Sustainable Development Goals (SDG). *Economic Research-Ekonomska Istraživanja, 34*(1), 427–449. https://doi.org/10.1080/1331677X.2021.1877167

Li, Y., & Zhang, X.-M. (2014). Diffusion maps based k-nearest-neighbor rule technique for semiconductor manufacturing process fault detection. *Chemometrics and Intelligent Laboratory Systems, 136*, 47–57. https://doi.org/10.1016/j.chemolab.2014.05.003

Liu, H.-W., & Zhang, S.-C. (2012). Noisy data elimination using mutual k-nearest neighbor for classification mining. *Journal of Systems and Software, 85*(5), 1067–1074. https://doi.org/10.1016/j.jss.2011.12.019

Magdum, R. (2022). What is data exploration? and its importance in data analytics. *International Research Journal of Engineering and Technology, 9*(1), 1482–1485. https://www.irjet.net/archives/V9/i1/IRJET-V9I1265.pdf

Malayan Banking Berhad. (2022, September 1). *Maybank launches sustainable product framework to drive sustainable finance.* https://www.maybank.com/en/news/2022/09/01.page

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval.* Cambridge University Press.

Mitchell, R. K., Agle, B. R., & Wood, D. J. (1997). Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *The Academy of Management Review, 22*(4), 853–886. https://doi.org/10.2307/259247

Muhmad, S. N., & Muhamad, R. (2020). Sustainable business practices and financial performance during pre- and post-SDG adoption periods: a systematic review. *Journal of Sustainable Finance & Investment, 11*(4), 291–309. https://doi.org/10.1080/20430795.2020.1727724

Nechita, E., Manea, C. L., Nichita, E.-M., Irimescu, A.-M., & Manea, D. (2020). Is financial information influencing the reporting on SDGs? Empirical evidence from central and eastern European chemical companies. *Sustainability, 12*(21), Article 9251. https://doi.org/10.3390/su12219251

Paul, A., Mukherjee, D. P., Das, P., Gangopadhyay, A., Chintha, A. R., & Kundu, S. (2018). Improved random forest for classification. *IEEE Transactions on Image Processing, 27*(8), 4012–4024. https://doi.org/10.1109/TIP.2018.2834830

Pham, D. C., Do, T. N. A., Doan, T. N., Nguyen, T. X. H., & Pham, T. K. Y. (2021). The impact of sustainability practices on financial performance: empirical evidence from Sweden. *Cogent Business & Management, 8*(1), Article 1912526. https://doi.org/10.1080/23311975.2021.1912526

Powers, D. M. W. (2011). Evaluation: From precision, recall and f-measure to ROC, informedness, markedness & correlation, *Journal of Machine Learning Technologies, 2*(1), 37–63.

Ramos, D. L., Chen, S., Rabeeu, A., & Abdul Rahim, A. B. (2022). Does SDG coverage influence firm performance? *Sustainability, 14*(9), Article 4870. https://doi.org/10.3390/su14094870

Sehgal, V., Garg, N., & Singh, J. (2022). Impact of sustainability performance & reporting on a firm’s reputation. *International Journal of System Assurance Engineering and Management, 14*, 228–240. https://doi.org/10.1007/s13198-022-01782-3

Shakil, M. H., Mahmood, N., Tasnia, M., & Munim, Z. H. (2019). Do environmental, social and governance performance affect the financial performance of banks? *Management of Environmental Quality, 30*(6), 1331–1344. https://doi.org/10.1108/meq-08-2018-0155

Sharma, D., Kumar, R., & Jain, A. (2022). Breast cancer prediction based on neural networks and extra tree classifier using feature ensemble learning. *Measurement: Sensors, 24*, Article 100560. https://doi.org/10.1016/j.measen.2022.100560

Sharma, S. (2021). Classification and regression trees: The use and significance of trees in analytics, *Journal on Recent Innovation in Cloud Computing, Virtualization & Web Applications, 5*(1), 1–24.

Singh, S. & Giri, M. (2014). Comparative study ID3, CART and C4.5 decision tree algorithm: A survey. *International Journal of Advanced Information Science and Technology, 3*(7), 47–52. https://www.ijaist.com/wp-content/uploads/2018/08/ComparativeStudyId3CartAndC4.5DecisionTreeAlgorithmASurvey.pdf

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics, 87*(3), 355–374. https://doi.org/10.2307/1882010

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *The Academy of Management Review, 20*(3), 571–610. https://doi.org/10.2307/258788

Tapestry Networks. (2020, May 7). *How banks are defining and internalising sustainability goals.* Ernst & Young Global Limited. https://www.ey.com/en_us/banking-capital-markets/how-banks-are-defining-and-internalizing-sustainability-goals

The Star Online. (2023, April 24). *Maybank mobilises RM34.37bil in sustainable finance.* https://www.thestar.com.my/business/business-news/2023/04/24/maybank-mobilises-rm3437bil-in-sustainable-finance

The World Bank. (2019, April 18). *Connecting Malaysia's Islamic and sustainable finance to the world.* https://www.worldbank.org/en/news/feature/2019/04/18/islamic-finance-initiatives-from-the-hub

United Nations Global Compact Network Malaysia & Brunei. (2022). *Malaysia Businesses Sustainability Pulse Report 2022.*

United Nations. (2018). *The sustainable development agenda.* https://www.un.org/sustainabledevelopment/development-agenda-retired/

Whelan, T., Atz, U., Holt, T. V., & Clark, C. (2021). ESY and financial performance*: Uncovering the relationship by aggregating evidence from 1,000 plus studies published between 2015 – 2020.* NYU Stern.

https://www.stern.nyu.edu/sites/default/files/assets/documents/ESG%20Paper%20Aug%202021.pdf

World Business Council for Sustainable Development. (2018, July 10). *Latest trends and challenges regarding business integration of the Sustainable Development Goals.* https://www.wbcsd.org/Programs/People-and-Society/Sustainable-Development-Goals/News/Latest-trends-and-challenges-regarding-business-integration-of-the-SDGs

Zhang, Z.-H. (2016). Introduction to machine learning: K-nearest neighbors. *Annals of Translational Medicine, 4*(11), 218. https://doi.org/10.21037/atm.2016.03.37

Zuo, R.-G., & Carranza, E. J. M. (2011). Support vector machine: A tool for mapping mineral prospectivity. *Computers & Geosciences, 37*(12), 1967–1975. https://doi.org/10.1016/j.cageo.2010.09.014