



**T.C.
YALOVA ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

**İŞLETME ANABİLİM DALI
İŞLETME BİLİM DALI**

ESG PERFORMANCE PREDICTION BY MACHINE LEARNING

YÜKSEK LİSANS TEZİ

ÇİSEM TURĞAY

DANIŞMAN: DOÇ. DR. SİNEM ATEŞ

**YALOVA
TEMMUZ 2025**



T.C.
YALOVA ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

İŞLETME ANABİLİM DALI
İŞLETME BİLİM DALI

ESG PERFORMANCE PREDICTION BY MACHINE LEARNING

YÜKSEK LİSANS TEZİ

ÇİSEM TURĞAY
238127003

DANIŞMAN: DOÇ. DR. SİNEM ATEŞ

YALOVA
TEMMUZ 2025

ETİK BEYAN

Yalova Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Yazım Kuralları'na uygun olarak hazırladığım “ESG PERFORMANCE PREDICTION BY MACHINE LEARNING” başlıklı bu tez çalışmada; tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi, tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu, tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi, kullanılan verilerde herhangi bir değişiklik yapmadığımı, bu tezde sunduğum çalışmanın özgün olduğunu bildirir, aksinin tespiti halinde doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi taahhüt ve beyan ederim.

Çisem TURĞAY

ETHICAL STATEMENT

I hereby declare that the thesis titled “ESG PERFORMANCE PREDICTION BY MACHINE LEARNING”, which I have prepared in accordance with the Thesis Writing Guidelines of Yalova University Institute of Graduate Studies, complies with the principles of academic and ethical conduct. I affirm that all data, information, and documents presented in this thesis have been obtained and used in line with academic and ethical standards, and that all sources I benefited from have been properly cited and referenced. I confirm that I have made no alterations to the original data, and that the work presented in this thesis is original and my own. I accept full legal responsibility for any contrary situation that may arise.

Çisem TURĞAY

PREFACE

I would like to express my sincere gratitude to my esteemed advisor, Assoc. Prof. Dr. Sinem ATEŞ, who has always guided me academically throughout my master's studies and generously shared her knowledge and experience. I am deeply thankful to my family for always being by my side and for their unwavering emotional support especially to my dear brother, İlyas Turğay, to whom I am truly grateful.

Temmuz – 2025

Çisem TURĞAY



TABLE OF CONTENTS

ETİK BEYAN	i
PREFACE	ii
TABLE OF CONTENTS	iii
LIST OF SYMBOLS AND ABBREVIATIONS	v
LIST OF TABLES	vi
LIST OF FIGURES	vii
ÖZET	viii
ABSTRACT	ix
1. INTRODUCTION	1
2. CONCEPTUAL AND THEORETICAL FRAMEWORK	6
2.1. Conceptual Foundations of ESG Performance	6
2.2. Theoretical Foundations of The Impact of Financial Structure on ESG Scores	8
2.2.1. Stakeholder theory	9
2.2.2. Legitimacy theory	10
2.2.3. Theory of slack resources	11
3. LITERATURE REVIEW	13
4. RESEARCH FRAMEWORK	20
4.1. Sample and Variables	20
4.2. Research Hypotheses	21
4.3. Data Cleaning and Preprocessing	21
4.4. Feature Selection and Modeling	23
4.4.1. Feature selection process	23
4.4.2. Modeling approach	24
4.5. Model Performance Evaluation Methods	27
4.5.1. RMSE (Root Mean Squared Error)	27
4.5.2. R ² (Determination Coefficient):	28
4.5.3. MAE (Mean Absolute Error):	29
4.5.4. MSE (Mean Squared Error):	30
5. FINDINGS	32
5.1. Model Performance Results	33

5.1.1. Baseline modeling	33
5.1.2. Modeling with 5-fold cross-validation	34
5.1.3. Comparison of training-test and 5-fold cross-validation performance	36
5.2. Key Factors Determining ESG Scores	37
5.3. Model Reliability and Consistency	39
6. DISCUSSION AND CONCLUSION	40
6.1. Evaluation of Findings	40
6.2. Recommendations for Policies and Practices.....	42
6.3. Limitations of the Study	42
6.4. Recommendations for Future Research	43
REFERENCES.....	45
APPENDICES.....	55
ÖZGEÇMİŞ	64

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOLS

R^2 : Coefficient of Determination

ABBREVIATIONS

ANN	: Artificial Neural Network
BIST	: Borsa İstanbul
BGC	: Bagging Classifier
BRR	: Bayesian Ridge Regression
CSR	: Corporate Social Responsibility
CatBoost	: Categorical Gradient Boosting
EBIT	: Earnings Before Interest and Taxes
ELM	: Extreme Learning Machines
EPS	: Earnings Per Share
ESG	: Environmental, Social, and Governance
GLM	: Generalized Linear Model
GBDT	: Gradient Boosting Decision Tree
KNN	: K-Nearest Neighbors
LGBM/ LightGBM	: Light Gradient Boosting Machine
MAE	: Mean Absolute Error
MAPE	: Mean Absolute Percentage Error
ML	: Machine Learning
MLR	: Multiple Linear Regression
MSE	: Mean Squared Error
NBC	: Naive Bayes Classifier
RF	: Random Forest
RMSE	: Root Mean Squared Error
ROA	: Return on Assets
ROE	: Return on Equity
SVM	: Support Vector Machines
SVR	: Support Vector Regression
SHAP	: SHapley Additive exPlanations
XGBoost	: eXtreme Gradient Boosting

LIST OF TABLES

Table 5.1. Comparison of Baseline Model Performance	33
Table 5.1. (Continued) Comparison of Baseline Model Performance.....	34
Table 5.2. 5-Fold Cross-Validation Results	35
Table 5.3. Train-Test and 5-Fold Cross-Validation Performance Metrics.....	36



LIST OF FIGURES

Figure 5.1. Yearly Evolution of ESG Scores	32
Figure 5.2. SHAP Value Ranking of Variables Determining ESG Scores	38



MAKİNE ÖĞRENMESİ KULLANILARAK ESG PUAN TAHMİNLEMESİ

ÖZET

Son yıllarda çevresel, sosyal ve yönetim (ESG) performansı, şirketlerin sürdürülebilirlik ve etik sorumluluklar açısından değerlendirilmesinde önemli bir ölçüt haline gelmiştir. ESG puanlarının finansal performansla olan ilişkisini anlamaya yönelik artan ilgi, bu alanda kullanılan ölçümleme yöntemlerinin geliştirilmesini zorunlu kılmaktadır. Ancak mevcut ESG derecelendirme sistemlerinde karşılaşılan metodolojik farklılıklar ve veri tutarsızlıkları, değerlendirmelerin güvenilirliğini azaltmakta ve yatırım kararlarını güçleştirmektedir. ESG performansına yönelik artan toplumsal, çevresel ve yatırımcı beklentileri ise, sürdürülebilirlik temelli yatırım kararlarında veri odaklı analiz ve tahmin modellerine olan ihtiyacı artırmaktadır.

Bu tez çalışmasının temel amacı, şirketlerin ESG skorlarının yalnızca kurumsal itibar açısından değil, aynı zamanda yatırımcı kararlarını etkileyen stratejik bir unsur olarak ele alınması gerektiği varsayımından hareketle, ESG skorlarının finansal göstergeler aracılığıyla tahmin edilebilirliğini analiz etmektir. Ayrıca, ESG skorları üzerinde etkili olan finansal değişkenlerin belirlenmesi ve bu ilişkinin kuramsal temellerle açıklanması da çalışmanın hedefleri arasındadır.

İlgili hedefe ulaşmak amacıyla, Borsa İstanbul'da (BIST) faaliyet gösteren şirketlerin 2008–2023 dönemine ait ESG skorları ile kârlılık, likidite, borçluluk ve faaliyet verimliliği gibi çeşitli finansal göstergelerden oluşan bir veri seti kullanılmıştır. Random Forest (RF), Gradient Boosting, Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), Ridge ve K-Nearest Neighbors (KNN) algoritmaları ile regresyon modelleri oluşturulmuş; modellerin performansı Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Squared Error (MSE) ve Determination coefficient (R^2) metrikleriyle değerlendirilmiştir. Model yorumlanabilirliğini artırmak amacıyla SHAP (SHapley Additive exPlanations) yöntemi uygulanmış, ESG skorlarını en yüksek doğrulukla tahmin eden finansal değişkenler belirlenmiştir. Ayrıca, çapraz doğrulama ve hiperparametre optimizasyonu süreçleriyle model güvenilirliği sağlanmıştır. Kuramsal çerçevede ise paydaş teorisi, meşruiyet teorisi ve atıl kaynaklar teorisi temel alınarak ESG ile finansal yapı arasındaki ilişki açıklanmıştır.

Araştırmanın temel bulguları, ESG skorlarının yalnızca birkaç temel finansal göstergeden değil, şirketlerin genel mali yapısı, stratejik harcama tercihleri ve bilanço bileşenlerinden de önemli ölçüde etkilendiğini göstermektedir. Analiz sonuçları, özellikle maddi olmayan varlıklar, dönemsel etkiler, hisse başına kâr oranları, operasyonel gider kalemleri ve uzun vadeli borçluluğun ESG skorlarının öngörülmesinde belirleyici rol oynadığını ortaya koymuştur. Bu durum, ESG performansının sadece çevresel ya da sosyal politikalara bağlı olmadığını; aynı zamanda şirketlerin kaynak kullanımı, yatırım stratejisi ve finansal sürdürülebilirlik düzeyleriyle yakından ilişkili olduğunu göstermektedir.

Bu tez çalışması, ESG skorlarının tahmin edilmesine yönelik literatüre hem teorik hem de metodolojik katkılar sunmaktadır. Gelişmekte olan bir piyasa olan Türkiye özelinde, uzun dönemli ve sektör çeşitliliği içeren bir örneklem kullanılarak yürütülen bu araştırma; yatırımcılar, politika yapımcılar ve şirket yöneticileri açısından karar destek sistemlerinin geliştirilmesine katkı sağlamaktadır.

Anahtar Kelimeler: ESG, ESG Skoru Tahmini, Finansal Göstergeler, Makine Öğrenmesi, SHAP Analizi

ESG PERFORMANCE PREDICTION BY MACHINE LEARNING

ABSTRACT

In recent years, Environmental, Social, and Governance (ESG) performance has emerged as a significant criterion for assessing companies' commitment to sustainability and ethical practices. The growing interest in understanding the relationship of ESG scores with financial performance necessitates the improvement of measurement methodologies used in this field. However, methodological discrepancies and data inconsistencies in current ESG rating systems undermine the reliability of assessments and complicate investment decisions. Meanwhile, rising societal, environmental, and investor expectations regarding ESG performance have amplified the need for data-driven analysis and predictive models in sustainability-oriented investment decision-making.

The main purpose of this thesis study is to analyze the predictability of companies' ESG scores through financial indicators, based on the assumption that ESG performance should be considered not only as a component of corporate reputation but also as a strategic factor influencing investor decisions. In addition, identifying the financial variables that significantly impact ESG scores and explaining this relationship with theoretical framework are also among the objectives of the study.

In pursuit of this aim, a dataset comprising ESG scores and various financial indicators, like profitability, liquidity, leverage, and operational efficiency, of companies listed on Borsa Istanbul (BIST) for the period 2008–2023 was utilized. Regression models were developed using Random Forest (RF), Gradient Boosting, Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), Ridge, and K-Nearest Neighbors (KNN) algorithms, and their performance was assessed using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Determination coefficient (R^2) metrics. To enhance model interpretability, the SHAP (SHapley Additive exPlanations) method was applied, and the financial variables with the greatest influence on ESG score predictions were identified. In addition, model reliability was ensured through cross-validation and hyperparameter optimization processes. From a theoretical perspective, the relationship between ESG and financial structure was explained based on stakeholder theory, legitimacy theory, and theory of slack resources.

The main findings of the study show that ESG scores are significantly influenced not only by a few key financial indicators, but also by companies' overall financial structure, strategic spending preferences, and balance sheet components. The analysis results reveal that variables such as intangible assets, period-specific effects, earnings per share, operational expense items, and long-term debt play a decisive role in predicting ESG scores. This situation demonstrates that ESG performance is not only dependent on environmental or social policies, but is also closely linked to companies' resource utilization, investment strategy, and financial sustainability.

This thesis study provides both theoretical and methodological contributions to the literature on the prediction of ESG scores. This research, based on a long-term and sector-diverse sample, particularly focused on Turkey as an emerging market, contributes to the development of decision support systems for investors, policy makers, and company managers.

Keywords: ESG, ESG Score Prediction, Financial Indicators, Machine Learning, SHAP Analysis

1. INTRODUCTION

In recent years, companies' ESG performance has emerged as a crucial evaluation tool reflecting sustainability, social responsibility, and adherence to ethical values (Artiach et al., 2010). The ESG concept further deepens the understanding of sustainability by holistically addressing a company's environmental impact, social responsibilities, and governance quality (Adams & Abhayawansa, 2022). The three core dimensions of ESG can be summarized as follows (Rajesh, 2020; de Souza Barbosa et al., 2023; Çek & Eyüpoğlu, 2020): The environmental dimension encompasses issues such as carbon emissions, energy usage, climate risk, natural resource utilization, pollution, and waste management. The social dimension includes aspects such as employee rights, labor issues, diversity, community engagement, and product and information security. The governance dimension covers principles such as transparency, ethical management, stakeholder rights, internal control systems, and board structure (Ferriani & Natoli, 2021; Przychodzen et al., 2016). Within this framework, ESG data are predominantly non-financial in nature and have become vital tools for evaluating companies' long-term sustainability (Ferriani & Natoli, 2021; Arora & Sharma, 2022). Such data, often referred to as "non-accounting information," complement traditional financial reporting by capturing dimensions that conventional financial metrics cannot fully explain, thereby providing additional value for investors and stakeholders (Krishnamoorthy, 2021).

The comprehensive structure of ESG is reshaping the business world's understanding of sustainability. Companies with high ESG scores are not only perceived as having strong corporate reputations but are also becoming increasingly attractive to investors and consumers (Wang et al., 2024). These firms fulfill both their financial and social responsibilities more effectively through sustainable business practices (Wang et al., 2024). Organizations face growing demands to act more responsibly toward the environment, society, and their stakeholders. These demands are due to a large number of stakeholders, including shareholders, customers, regulators, employees, suppliers, activist groups and the media. Global challenges, especially climate change, air pollution, and social inequalities, are intensifying the need for ESG strategies (Lee et al., 2022). These factors represent dynamic forces that impact the long-term success of companies and compel them to reconstruct their business models (Adams & Abhayawansa, 2022).

Companies with strong ESG strategies view sustainability not only as a legal obligation but also as a competitive advantage and an opportunity for corporate responsibility (Lee et al., 2022).

These firms implement exemplary practices that go beyond regulatory compliance, aiming to reduce their environmental and social impacts and strengthen stakeholder relationships. The importance of ESG performance becomes even more pronounced in the context of pressing environmental challenges, such as climate change, air pollution, and biodiversity crises, as well as social issues, including human rights, equality, and broader societal concerns.

Additionally, volatile economic environments, particularly those experienced during the recent global financial crisis, have compelled companies to adopt more sustainable approaches in their strategic decision-making processes (Elgergeni et al., 2018; Garcia-Sanchez et al., 2014). In the post-crisis period, unethical corporate behavior, weaknesses in risk oversight, and a lack of accountability have heightened the importance of the ESG concept (Sultana et al., 2018). In this context, ESG extends beyond mere regulatory compliance; it is also developed as a means to strengthen companies' relationships with stakeholders, reduce their environmental and social impacts, and enhance corporate responsibility (Adams & Abhayawansa, 2022).

In recent years, the increasing preference of investors and consumers for sustainability-oriented companies has significantly enhanced the strategic importance of ESG criteria. A primary contributing factor is the growing tendency of consumers and investors to favor companies that demonstrate sustainability and social responsibility (Krappel et al., 2021). Investors are now turning not only to financial returns but also to companies with ethical and sustainable values (Adams and Abhayawansa, 2022). This trend shows that sustainability awareness and social responsibility consciousness are influential in investment decisions (Krappel et al., 2021). ESG performance reflects not only environmental and social sensitivity but also incorporates the evaluation of non-financial factors such as carbon emissions, energy consumption, labor rights, and corporate transparency (Przychodzen et al., 2016). Therefore, investors regard ESG as a strategic indicator that both mitigates financial risks and supports long-term value creation (D'Amato et al., 2022; Sultana et al., 2018)

For investors, ESG factors are becoming increasingly important in understanding a company's long-term financial stability, reputational value, and operational risks. ESG criteria not only reflect environmental and social sensitivity but also provide insights into the quality of corporate governance. Institutional investors, in particular, tend to make their portfolios more sustainable by preferring companies with high ESG performance. This approach enables investors to act responsibly not only with respect to financial returns but also regarding social and environmental impacts (Przychodzen et al., 2016; de Souza Barbosa et al., 2023). With the integration of the sustainability concept into investment processes, ESG evaluations have

become an indispensable component of strategic investment decisions (Ortas et al., 2015; Nevado-Peña et al., 2015).

In the modern economic landscape, the growing importance of ESG factors has increased the demand for accurate ESG scoring and prediction models to better assess corporate performance. The rising interest of investors in ESG data has made it imperative to improve the measurement and rating systems in this field. However, since ESG scores are calculated using different methodologies by various institutions, inconsistencies and uncertainties may arise in evaluations. This, in turn, can negatively affect the accuracy and comparability of investment decisions (Del Vitto et al., 2023). Amid this complexity, the need for ESG professionals and advanced analytical methods is steadily increasing. In this context, statistical and machine learning-based modeling techniques for ESG score prediction offer investors a more consistent, objective, and transparent evaluation framework. These techniques go beyond predictions based solely on financial ratios and enable a more holistic assessment of a company's long-term potential (Krishnamoorthy, 2021). Supporting ESG data with advanced prediction models contributes to more informed and prudent investment decisions (Gupta et al., 2021).

Accordingly, the accurate prediction and interpretation of ESG scores has become increasingly important in both academic literature and practical applications. In particular, studies that predict ESG scores based on financial indicators help clarify the relationship between sustainability and financial performance. These studies demonstrate that companies can be assessed not only based on their financial statements but also based on their environmental and social responsibility levels. Machine learning (ML) techniques developed in the past few years have become effective tools for estimating ESG scores due to their ability to model complex relationships using multidimensional data. Compared to traditional statistical models, these methods offer higher predictive accuracy and are more effective in capturing patterns within financial data. In this context, machine learning-based modeling studies that predict ESG scores using financial ratios pave the way for a more objective and data-driven approach to investment decision-making.

In the existing literature, various studies have employed different ML algorithms to predict ESG scores (D'Amato et al., 2022; Raza et al., 2022; Cini & Ferrari, 2025; Assael, 2023). These studies involve forecasting ESG scores through regression or classification using financial data, as well as conducting comparative performance analyses of different algorithms. Moreover, some have focused on model interpretability, analyzing variable contributions using methods such as SHAP (Assael, 2023; Cots, 2024). However, the majority of these studies primarily aim

to identify the most accurate prediction model and do not offer a sufficiently detailed and systematic analysis of the financial variables influencing ESG performance. In contrast, this thesis not only seeks to determine the optimal predictive model but also provides a comprehensive examination of how financial indicators within the model affect ESG scores, thereby contributing to the literature. Additionally, data limitations regarding ESG scoring in developing markets such as Turkey have resulted in a relatively small number of studies in this area.

This thesis seeks to develop a comprehensive model that can most accurately predict ESG performance and to determine how important the financial indicators included in the model are in predicting ESG scores. To achieve this, the predictive performances of various ML algorithms are compared, and SHAP values are analyzed using the best-performing algorithm to assess the contribution of each feature. Accurately predicting ESG performance not only contributes to the academic literature but also holds critical importance for various stakeholders, including investors, shareholders, consumers, and policymakers.

This study seeks to contribute to both theory and practice by developing a model that predicts ESG performance. The developed model aims to contribute to relevant theoretical frameworks, particularly stakeholder theory, legitimacy theory, and theory of slack resources, by enabling a better understanding of the relationship between ESG factors and financial performance.

From a practical standpoint, the model enables investors to more effectively assess companies with high ESG performance, thereby supporting long-term and sustainable investment decisions. Likewise, considering the growing importance of ESG factors for policymakers, the study offers a valuable tool to support the development of policies that prioritize environmental and social responsibility.

The research sample comprises 95 companies listed on Borsa Istanbul for which complete data required for the study model were available. The full list of these companies is presented in Table D.1 in the appendix section. The dataset includes ESG scores of these companies from 2008 to 2023, alongside financial data reflecting multiple dimensions such as profitability, liquidity, operational efficiency, leverage, and cash flow. All data were accessed via the Thomson Reuters Eikon database. The analysis process was conducted using the Python programming language. In the ML applications, data cleaning, feature selection, modeling, hyperparameter optimization, and result evaluation steps were systematically carried out. For predicting ESG scores, models were constructed using Gradient Boosting Regressor, RF

Regressor, XGBoost Regressor, LightGBM Regressor, Ridge Regressor, and KNN Regressor algorithms, and their performances were compared based on predefined evaluation metrics. To better understand the decision-making processes of the model and interpret the effects of variables on ESG scores, the SHAP method was applied. This approach identified the most influential financial variables in predicting ESG scores and enhanced the model's internal consistency.

This thesis is structured into six main chapters. The first chapter, *Introduction*, presents the background of the study, the research problem, the objectives and significance of the study, as well as its overall scope.

The second chapter, *Conceptual and Theoretical Framework*, elaborates on the conceptual foundations of ESG performance, detailing its components and measurement methods. Additionally, the theoretical background explaining the impact of corporate financial structure on ESG scores is provided. Within this context, the relationship between ESG performance and financial indicators is discussed through the lens of Stakeholder Theory, Legitimacy Theory, and Resource Slack Theory.

The third chapter, *Literature Review*, examines existing studies on the prediction of ESG performance using ML techniques. The strengths and weaknesses of previous studies are analyzed, highlighting some unanswered questions and explaining how this research contributes to the field.

The fourth chapter, *Research Framework*, outlines the research model in detail. It presents the hypotheses, describes the dataset structure and cleaning processes, explains the feature selection methods, the ML algorithms used, and the performance evaluation techniques. Furthermore, the explainability process using SHAP values is discussed in detail to interpret model results.

The fifth chapter, *Findings*, presents the performance metrics of the constructed models, identifies key factors in ESG score prediction, and evaluates the reliability and consistency of the developed models.

The final chapter, *Discussion and Conclusion*, provides a comparative analysis of the findings in light of existing literature, offers policy recommendations for practitioners, discusses the limitations of the study, and presents suggestions for future research.

2. CONCEPTUAL AND THEORETICAL FRAMEWORK

2.1. Conceptual Foundations of ESG Performance

The ESG concept became important when companies started to focus more on profits and paid less attention to environmental issues, especially during the industrialization period (Chowdhury et al., 2023). The roots of ESG lie in the concepts of Corporate Social Responsibility (CSR) and sustainable investing (D'Amato et al., 2021). In today's world, the advancement of information technologies has significantly facilitated access to information. Consumers of goods and services, who now have easier access to information, have become more aware of the environmental and social impacts of companies. There is a growing demand among consumers for products that meet ethical and environmental standards throughout their supply chain. Consequently, companies are evaluated by society not only on the basis of their balance sheets but also on their understanding of sustainability.

The concept of sustainability has evolved over the years, influenced by economic, environmental, and social changes, before gaining its current significance. Technological advancements that accelerated after the 1950s led to significant transformations in production processes, resulting in issues such as environmental pollution and depletion of natural resources (Soytaş et al., 2017). The rise of environmental movements in the 1970s and 1980s contributed to increased social awareness of the environmental impacts of businesses (D'Amato et al., 2021). In this context, the declaration adopted at the United Nations Conference on the Human Environment held in Switzerland in 1972 marked an important turning point in the international discussion of sustainability.

The widespread adoption of the concept in both academic and practical domains occurred with the publication of the report *Our Common Future* by the World Commission on Environment and Development in 1987 (Brundtland, 1987). According to the report, sustainable development means meeting current needs while protecting the ability of future generations to meet their own, and this definition has gained wide acceptance in the literature (Dağıstanlı & Çelik, 2023). During the 1990s, the concept of CSR began to gain prominence, and national and international pressures for companies to act more transparently on environmental and social issues increased (D'Amato et al., 2021). Notable initiatives from this period, such as the CERES Principles (1989), the Earth Summit (1992), and the Kyoto Protocol (1997), made significant contributions to the institutionalization of sustainability principles (Tokgöz & Önce, 2009).

In the early 2000s, rising awareness of environmental and social issues led investors and other stakeholders to begin assessing companies' sustainability performance (Wang et al., 2024). The term "ESG" emerged in 2004 with the publication of the *Who Cares Wins* report by the United Nations Global Compact (UNGC) (Zhang & Zhang, 2024). This was followed by the launch of the United Nations' Principles for Responsible Investment (PRI) in 2006, and the adoption of the 2030 Agenda for Sustainable Development in 2015 (Wang et al., 2024; Friede, 2019). These developments led to the global recognition of the ESG concept, supported the integration of environmental, social, and governance-based investments into mainstream financial strategies, and positioned ESG as a key tool for long-term value creation and sustainability at the corporate level (Tsalis et al., 2023).

The growing significance of ESG has become especially apparent during the COVID-19 pandemic (Adams & Abhayawansa, 2022; Krishnamoorthy, 2021). ESG is an assessment index that evaluates a company's non-financial performance (Lee et al., 2022). This evaluation serves as a key metric that reflects a firm's contribution to sustainable development and social well-being, forming the foundation of the "sustainable investment" paradigm for companies (Chowdhury et al., 2023; Raza et al., 2022). The term ESG is structured to address specific issues within each of its three pillars and aims to integrate the core components of ethical finance: Environmental (E), Social (S), and Governance (G) (Cini & Ferrari, 2025). These are typically categorized as ESG dimensions, respectively (Krappel et al., 2021) ;

- The Environmental pillar (E) evaluates the extent to which a company develops eco-friendly materials and processes, and how effectively it manages environmental risks through best management practices (Kulali, 2022; Crespi et al.). It also encompasses the company's overall impact on natural systems, including air, soil, water, and broader ecosystems. This pillar focuses on issues such as climate change, pollution, emissions, energy and resource efficiency, and waste management efforts undertaken by the company (De Lucia et al., 2020; Cini & Ferrari, 2025).
- The Social pillar (S) quantitatively reflects a company's sensitivity toward human capital, and more specifically, its ability to foster trust and loyalty among stakeholders (Crespi & Migliavacca, 2020). It encompasses aspects such as gender policies, labor practices, human rights, community engagement, and product responsibility (De Lucia et al., 2020). Additionally, factors like workplace safety,

employee satisfaction, customer loyalty, and interaction with society are also evaluated within this component (Peng & Isa, 2020).

- The Governance pillar (G) encompasses the systems and processes through which a company's management structure ensures the protection of long-term shareholder interests. This dimension reflects the company's ability to guide and control rights and responsibilities in a balanced manner, as well as the incentive mechanisms and oversight procedures implemented to create long-term value (Kulalı, 2022). The Governance pillar focuses on issues such as business ethics, corporate governance practices, and broader governance-related discussions (Köse, 2021).

In the world of finance and investment, evaluating companies' ESG performance has become an increasingly important criterion. In this context, financial institutions and investors use the scores provided by various ESG rating agencies to measure ESG performance (Wang et al., 2024; D'Amato et al., 2022; Edhrabooh, 2024). The methods used to evaluate ESG performance typically involve a combination of quantitative and qualitative data, and the accuracy of these assessments largely depends on the quality of the data sets employed and the effectiveness of the analytical techniques applied (Lin & Hsu, 2023).

ESG rating agencies determine ESG scores by analyzing publicly disclosed company data, sustainability reports, financial information, and other relevant disclosures. Among the leading providers in this domain are MSCI, Refinitiv, Bloomberg, and Sustainalytics each of which differs in how it approaches ESG topics and the methodologies it employs in the rating process (Krappel et al., 2021). These agencies collect and analyze data based on specific parameters they have defined.

ESG scores are calculated by applying specific formulas to data collected from various sources in accordance with predefined parameters. These scores are typically reported in either numerical format (ranging from 0 to 100) or as letter grades (e.g., D- to A+). In the assessment of ESG performance, predefined percentage weights are assigned to ten key indicators across the environmental, social, and governance dimensions. Based on these weights, separate scores for each ESG component are computed, followed by the calculation of an overall ESG score (Şişman & Çankaya, 2021).

2.2. Theoretical Foundations of The Impact of Financial Structure on ESG Scores

This section discusses the effects of financial structure on ESG scores within the framework of various theoretical approaches. Stakeholder Theory, Legitimacy Theory, and Resource

Slack Theory are frequently employed in the literature to explain the factors influencing firms' ESG performance. These theories generally posit that strong financial characteristics enhance the extent of information disclosure related to ESG practices (Şahin & Acar, 2023).

2.2.1. Stakeholder theory

Stakeholder theory asserts that companies hold responsibilities not only to their shareholders, but also to all other stakeholders, including investors, employees, customers, suppliers, governments, and society at large (Freeman, 1984). This theory builds upon Freeman's (1984) stakeholder approach and the work of Ullmann (1985), who linked stakeholder perspectives to CSR reporting and disclosure processes (Gümrah & Tanç, 2018). It argues that businesses should aim to maximize the interests of all stakeholders in a balanced way, rather than focusing solely on shareholder value (Freeman, 1984; Ricart et al., 2005). Accordingly, firms are expected to be responsive to the needs of all stakeholders when shaping their operations (Ceran, 2017).

There is a strong connection between Stakeholder theory and ESG components. To improve their financial performance, companies must fulfill their responsibilities to all stakeholders. In this regard, firms increasingly allocate financial resources to ESG related activities to meet stakeholder expectations. Stakeholder theory provides a valuable framework for understanding and managing corporate ESG practices, emphasizing that these initiatives can positively influence the long-term success of companies (Wang C., 2024). By investing in ESG activities, companies aim to satisfy their stakeholders and ensure long-term sustainability (Kulalı, 2022; Dereköy, 2015). In this context, large firms benefit from economies of scale, allowing them easier access to financial resources and the ability to allocate more funds to social responsibility projects, thereby addressing stakeholder expectations more effectively (D'Amato et al., 2021).

Strong ESG performance offers significant advantages for firms. It reduces corporate risks and supports organizational sustainability (Garcia et al., 2017). Moreover, when a company's social contributions are featured in the media, they enhance corporate image, communicate social responsibility to stakeholders, increase public recognition, and indirectly contribute to improved financial performance (Wang C., 2024). Particularly, environmental sustainability practices help companies fulfill social responsibilities, respond to consumer and societal demands, and reduce compliance costs associated with environmental regulations ultimately attracting greater investor interest. Consequently, firms are able to

alleviate financing constraints and enhance their overall corporate performance (De Lucia et al., 2020).

2.2.2. Legitimacy theory

Legitimacy theory posits that organizations must operate in accordance with societal values, norms, and expectations in order to sustain their existence. According to this theory, for businesses to maintain long-term sustainability, they must act in ways that are consistent with societal norms and seek public approval (Çakır, 2009; Dereköy, 2015). Legitimacy, therefore, requires companies not only to generate economic profit but also to fulfill social expectations (Yavuz, 2023). Lindblom (1994) defines legitimacy as a condition in which an organization is perceived as congruent with the broader social system it is a part of, while Suchman (1995) explains legitimacy as a situation in which the actions of an entity are perceived as desirable and appropriate within a socially constructed system of norms, values, beliefs, and definitions. In this context, legitimacy theorists argue that organizations must be accepted by society to be considered legitimate.

Sustainability reporting emerges as a response to the desire for organizational practices to align with societal expectations (Owen, 2008). As one of the most dominant theoretical frameworks in the fields of corporate, social, and environmental reporting, legitimacy theory asserts that companies are not solely driven by economic goals but are also obligated to generate social value (Deegan et al., 2002; Yavuz, 2023).

The relationship between legitimacy theory and ESG factors is primarily centered on the notion that companies enhance their legitimacy through ESG disclosures. Financial characteristics such as profitability, firm size, leverage, ownership structure, foreign investor share, and analyst coverage all contribute to increased corporate visibility, thereby reinforcing the need for legitimacy (Şahin & Acar, 2023). Profitability and size are powerful drivers in line with legitimacy theory, supporting the view that larger firms are more likely to engage in CSR reporting due to greater accountability and public scrutiny (Crespi & Migliavacca, 2020).

ESG reports help companies share their environmental and social impacts clearly with the public and meet society's expectations about sustainability (Mahajan et al., 2023). This process also allows public companies to demonstrate their contributions to societal welfare and meet the expectations of various stakeholder groups (Greiling & Grüb, 2014, p. 211). To

legitimize their operations in the public eye, firms engage in ESG-related information sharing, thereby responding to social and political pressures (Ng & Rezaee, 2015).

Through ESG reporting, companies help secure their legitimacy by showcasing their commitment to sustainable management practices. By disclosing non-financial information, firms fulfill their social contract and respond to stakeholder expectations (Ng & Rezaee, 2015). Furthermore, firms with strong liquidity and greater public visibility tend to be more inclined to publish ESG reports as a means of gaining societal legitimacy (Şahin & Acar, 2023). Considering that legitimacy is a right granted by stakeholders, it is evident that a strong connection exists between legitimacy theory and stakeholder theory (Küçükaycan, 2022). In sum, legitimacy theory explains how ESG reporting helps organizations communicate their sustainability commitments to the public, thereby maintaining legitimacy at both institutional and societal levels.

2.2.3. Theory of slack resources

Slack Resources theory explains how firms utilize their unused or excess resources in alignment with their strategic objectives. According to this theory, financial resources such as cash flow, retained earnings, and similar assets enable a firm to allocate resources beyond its current operational activities towards various social and environmental initiatives (Chen et al., 2021; Duque-Grisales & Aguilera-Caracuel, 2021; Özçelik et al., 2014). With increasing institutionalization, sustainability issues have gained greater priority among firms' slack resource allocation decisions (Xiao et al., 2018). Accordingly, companies demonstrating stronger financial performance tend to possess more financial resources available for ESG activities, which in turn enhances their flexibility to invest in CSR initiatives (El Khoury et al., 2023; Cheng et al., 2014).

A firm's sustainability performance, particularly in achieving high sustainability outcomes, requires a robust financial structure because the effective execution of such performances depends on resources derived from financial success (Fauzi & Idris, 2009; Şahin & Acar, 2023). An increase in financial slack resources is regarded as an encouraging factor for firms to invest more in ESG practices (Chams et al., 2021). Indeed, the positive relationship between ESG reporting and liquidity as well as firm size aligns with slack resources theory, supporting the notion that firms with strong cash flows and effective working capital management tend to achieve higher ESG scores (Waddock & Graves, 1997; Scholtens, 2006; Surroca et al., 2010).

Within this framework, due to the high costs and uncertainties associated with sustainability investments, only firms with sufficient slack resources can undertake such investments, and profitable firms are more capable of improving their ESG scores (Wood & Jones, 1995; Tate et al., 2010; Lin et al., 2020; Drempetic et al., 2020). Firms endowed with adequate slack resources move beyond symbolic sustainability efforts towards more concrete actions, such as investing in new technologies that reduce energy consumption, pollution, and waste levels (Perez-Batres et al., 2012; Soytaş et al., 2017; Xiao et al., 2018).



3. LITERATURE REVIEW

In today's context, investment decisions are based not only on financial indicators but also on companies' ESG performance. ESG scores, which quantitatively represent a company's sustainability performance, play an increasingly critical role in both investors' decision-making processes and the strategic orientation of firms. This growing significance has positioned ESG scores as a central element in strategic decision-making (Cini & Ferrari, 2025). However, the calculation of ESG scores remains a problematic process due to methodological discrepancies, lack of transparency, and data inconsistencies (Kotsantonis & Serafeim, 2019; Berg et al., 2022). The use of differing methodologies by ESG rating agencies results in low correlations among scores and leads to information asymmetries, which adversely affect investors' decision-making processes (Adams & Abhayawansa, 2022; Zhang & Zhang, 2024).

The ability to predict ESG performance holds strategic importance not only from a sustainability perspective but also in terms of financial success. A comprehensive meta-analysis conducted by Friede et al. (2015) revealed a strong and positive relationship between ESG and corporate financial performance. Similarly, a study by Alshehhi et al. (2018), which reviewed 132 peer-reviewed academic papers, reported that 78% of these studies identified a positive correlation between ESG and financial performance. These findings indicate that considering ESG data in decision-making processes is a rational and strategic approach for both investors and corporate managers. Financial ratios, particularly balance sheet components and core metrics, emerge as effective predictors of ESG scores (D'Amato et al., 2021; Alsayyad & Fadel, 2025). Moreover, the literature emphasizes that the main variables influencing ESG scores are financial indicators and firm size, with a particularly strong relationship observed between profitability ratios, such as Return on Assets (ROA) and Return on Equity (ROE), and ESG scores (D'Amato et al., 2021; Wang et al., 2024; Lee et al., 2022; Nijhof, 2023). Therefore, ESG score prediction is of critical importance not only for investors but also for firms seeking to shape and optimize their financial strategies (Dhaliwal et al., 2011; Bassen et al., 2006).

In conclusion, the prediction of ESG scores has become a crucial tool that addresses the shortcomings of existing rating systems, supports investor decision-making, and guides companies toward long-term sustainability strategies. Data-driven approaches such as ML enhance the robustness and reliability of this process, offering a powerful solution for both academic research and practitioners. However, the structure and calculation of ESG scores have yet to reach a universal standard. Methodological discrepancies, lack of transparency, and data

inconsistencies among ESG rating agencies are frequently subject to criticism (Kotsantonis & Serafeim, 2019; Berg et al., 2022). Therefore, accurately, comparably, and predictably modeling ESG scores is of paramount importance for effective corporate risk management and ensuring that investors can access reliable information. ESG scores are increasingly becoming one of the most significant criteria in strategic investment decision-making (Cini & Ferrari, 2025).

In this context, the prediction of ESG scores, particularly through the use of ML algorithms, has become an essential necessity. With rapid technological advancements, ML and artificial intelligence algorithms are now widely employed in the analysis of ESG data and the prediction of ESG scores (Lee et al., 2022; Wang et al., 2024). The integration of ML algorithms into this domain significantly enhances the accuracy and consistency of ESG score predictions. Algorithms such as RF and XGBoost help to overcome the limitations of traditional approaches, making ESG evaluations more transparent, reproducible, and cost-effective (Lin & Hsu, 2023; Assael, 2023). This is particularly beneficial for non-public companies or firms operating in emerging markets, where ESG scores are often unavailable and difficult to obtain (Garcia et al., 2017). Models based on ESG score prediction reduce information asymmetry for investors, enabling them to make more informed and sustainable decisions, thereby gaining increasing strategic importance. A detailed examination of studies in the literature enables a deeper understanding of the potential use of ML techniques for ESG score estimation.

Accordingly, several prominent studies that contribute to the understanding and prediction of ESG scores through ML techniques are summarized below:

D'Amato et al. (2021) aimed to examine the impact of financial ratios on the prediction of ESG scores and to identify the key determinants among these ratios. For this purpose, the study included Bloomberg ESG scores of 109 companies listed in the Stoxx Europe 600 Index for the period 2014–2018, along with financial indicators such as net income to sales, sales-to-assets ratio, and debt-to-total assets ratio. The authors employed the RF algorithm and the traditional Generalized Linear Model (GLM) to build predictive models and evaluated their performance using Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) metrics. The results showed that the RF algorithm outperformed the GLM models by providing higher predictive accuracy. Furthermore, the analysis revealed that net income to sales, debt-to-total assets ratio, and sales-to-assets ratio were the most influential variables in predicting ESG scores.

Krappel et al. (2021) aimed to develop a predictive model for estimating companies' ESG scores using ML algorithms. The study utilized ESG scores from 7,413 firms across 84 countries covering the period from 2002 to 2019, along with firm-level variables such as historical financial performance, firm size, ROA, and earnings per share (EPS). To build the prediction model, the authors employed three different ML algorithms: Feedforward Neural Networks, XGBoost, and Categorical Gradient Boosting (CatBoost). The performance of these models was evaluated and compared based on selected performance metrics. The findings indicated that a heterogeneous ensemble model outperformed the individual algorithms. Additionally, firm size and financial indicators were identified as the most significant predictors of ESG scores.

Raza et al. (2022) employed ML algorithms to evaluate how balance sheet and income statement data affect the ESG scores of publicly listed non-financial companies in the United States, the United Kingdom, and Germany. The study analyzed accounting data from a total of 300 firms (100 from each country) covering the period from 2008 to 2020. To predict the ESG pillar score, market- and sales-related variables were included in the model. The authors tested several ML algorithms, including KNN, Polynomial Regression, Naive Bayes, RF, Artificial Neural Networks (ANN), and Support Vector Machines (SVM) with both linear and RBF kernels. The results indicated that ANN, KNN, and SVM with the RBF kernel outperformed the other models, demonstrating the lowest RMSE and MAE values.

Lee et al. (2022) aimed to analyze ESG data using ML and Deep Learning algorithms, with a particular focus on identifying the effects of adversarial attacks in ESG datasets and how these can be detected. The study utilized data obtained from the Kaggle platform and introduced Gaussian noise into the financial dataset to simulate adversarial attacks. The Local Outlier Factor (LOF) method was employed to detect these anomalies. The analysis revealed that after the introduction of Gaussian noise, the LightGBM demonstrated higher performance, suggesting that noisy data significantly impacts the accuracy of ML models. To predict ESG scores, financial indicators such as ROA, ROE, net profit margin, and nominal return were used. Several ML models were tested, including Decision Tree, Linear Regression, Gradient Boosting, LightGBM, XGBoost, and RF. Among these, Gradient Boosting and XGBoost achieved the highest performance. The study highlights the significance of data noise in ESG datasets and demonstrates that ESG scores can be effectively predicted using a relatively simple ML model, with LOF playing a key role in anomaly detection.

D'Amato et al. (2022) aimed to evaluate the contribution of corporate financial structure indicators on the prediction of ESG scores. The study included ESG scores for 401 companies listed in the STOXX Europe 600 Index over the period 2009–2019, along with financial variables such as net sales, Earnings Before Interest and Taxes (EBIT), ROE, price-to-earnings ratio, and dividend yield. Both the RF algorithm and the traditional GLM were employed to compare predictive performance. The results showed that the most significant predictors of ESG scores were net sales and EBIT, followed by ROE. The study highlighted that the RF algorithm outperformed linear models by better capturing non-linear relationships and effectively modeling complex interactions among variables. This demonstrated the superiority of ML methods like RF in ESG score prediction.

Chowdhury et al. (2023) aimed to develop a predictive model using ML algorithms to estimate companies' ESG scores. The study incorporated ESG scores from 6,166 firms across 73 countries for the period 2005–2019, including firm-level variables such as historical ESG scores, firm size, financial leverage ratio, EPS, total assets, as well as country-level variables Gross Domestic Product growth rate and unemployment rate. Six different ML algorithms were applied and compared: RF Classifier, ANN Classifier, Bagging Classifier (BGC), KNN Classifier, Naive Bayes Classifier (NBC), and SVM Classifier. Based on the performance metrics used, the RF algorithm demonstrated the best performance. The most important predictors for ESG score estimation were identified as historical ESG scores, firm size, and financial leverage ratio.

Nijhof (2023) investigated the use of financial data in predicting ESG ratings. For this purpose, ESG ratings and financial data from 11,070 companies covering the period 2019–2021 were obtained from the Refinitiv database. Additionally, variables such as number of employees and board size were included in the analysis model. The study employed Multiple Linear Regression (MLR), RF, and Support Vector Regression (SVR) algorithms to compare model performance. The findings revealed that both RF and SVR significantly outperformed the MLR model. Although RF and SVR showed almost equal performance in terms of RMSE and MAE, the RF model exhibited the highest overall performance compared to the other two. To assess how much each variable contributed to the model's outcome, feature relevance was analyzed. It was observed that employee count and revenue figures played the most prominent role in both RF and SVR models. Furthermore, dividend yield emerged as a key variable in the RF model, whereas operating income was found to be more prominent in the SVR model.

Lin and Hsu (2023) aimed to evaluate the impact of financial structure indicators on the prediction of ESG scores for non-financial firms in Taiwan, as well as the effect of COVID-19 era data on ESG score predictions. For this purpose, the study utilized Taiwan ESG Sustainable Development Index data developed by the Taiwan Economic Journal for the period 2018–2021, incorporating financial variables such as pre-tax and interest ROA (ROA(A)), post-tax ROE (ROE(A)), and fixed asset turnover ratio. The study applied and compared the performance of RF, Extreme Learning Machines (ELM), SVM, and XGBoost algorithms. The findings demonstrated that all four models showed high predictive ability for ESG scores across both training and testing phases, regardless of the time period. Specifically, ELM, XGBoost, and SVM models achieved R^2 values above 0.98 during the testing phase. During non-pandemic periods, ELM and SVM algorithms outperformed the others in terms of predictive accuracy. However, when focusing on the pandemic period, the RF model delivered stronger results than its peers during the pandemic. The study also revealed that while RF and XGBoost had slightly lower predictive performance than ELM and SVM, the pandemic period did not cause any significant deterioration in the overall prediction capabilities of the algorithms.

Si (2023) aimed to build a model to estimate corporate ESG scores using ML algorithms. In the study, the prediction model incorporated variables such as ESG data, firm size, financial performance indicators, and ESG subcomponents (ESG dimensions). Accordingly, RF was applied, and its results were analyzed in relation to those of other ML models such as ANN, Deep Neural Networks (DNN), and SVM. The results indicated that, based on the performance metrics used, the RF algorithm achieved the highest predictive accuracy. Furthermore, the most influential variables in ESG score prediction were identified as environmental sustainability and human rights.

Alsayyad and Fadel (2025) conducted a study aimed at predicting the ESG scores of firms in the oil and gas sector. The research analyzed ESG scores and various financial performance indicators of 497 oil and gas companies over 12 years. The study employed a total of 14 different ML algorithms, including decision trees, ensemble algorithms (boosted, bagged, voting, and RUSBoosted), XGBoost, LightGBM, RF, Linear SVM, Extreme Random Trees, Linear, Robust Linear, and Elastic Net models. Both classification and regression analyses were applied, and the classification approach was found to outperform regression in terms of predictive accuracy. The Minimum Redundancy Maximum Relevance feature selection method was used, which ranks features sequentially and supports both categorical and continuous data types. The feature importance analysis revealed that the country-based variable representing the

firm's headquarters location had the greatest impact on ESG scores. The results demonstrated that ML algorithms significantly outperformed linear models in ESG score prediction, with RF and XGBoost delivering particularly strong performance. The study aims to provide a practical framework for investors and decision-makers to assess ESG performance more accurately, thus enabling more informed and sustainable investment decisions.

Huotari (2024) aimed to predict the ESG ratings of S&P 500 companies based on financial data. In this context, financial indicators and ESG ratings of S&P 500 firms from 2012 to 2023 were analyzed. The RF algorithm was employed as the predictive model. The findings revealed that ESG ratings can be predicted with relatively high accuracy using only financial metrics. In particular, profitability indicators were identified as having the most significant influence on the predictions. Model performance was evaluated using RMSE and MAE metrics. The RF model achieved a relatively low MAE and was able to explain 26.9% of the variance in ESG scores. The study concluded that financial indicators are effective predictors of ESG performance; however, incorporating additional variables could further enhance the model's predictive accuracy.

Wang et al. (2024) aimed to develop a novel Stacked Generalization Model based on Bayesian Ridge Regression (SGM-BRR) to predict the ESG scores of companies operating in China. To achieve this, a prediction model was constructed using the ESG scores and financial data of 968 Chinese firms listed on A-shares between 2012 and 2020. Variables such as firm size, financial expenditures, debt-to-asset ratio, EPS, and market capitalization were included in the analysis model. The proposed SGM-BRR is a hybrid model in which RF, Gradient Boosting Decision Tree (GBDT), XGBoost, and LightGBM serve as base learners, while Bayesian Ridge Regression (BRR) is used as the meta-learner. Model performance was assessed based on evaluation metrics including RMSE, MAE, and R^2 and was compared to seven other baseline models. The analysis showed that the SGM-BRR model reduced RMSE by 18.4%, MAE by 15.4%, and increased R^2 by 2% compared to the competing models. Furthermore, when tested across 27 different industries, the model achieved the lowest RMSE and highest R^2 values, with the best R^2 performance observed in the energy, materials, and industrial sectors, respectively.

Cini and Ferrari (2025) aimed to evaluate the impact of firms' balance sheet ratios and systemic risk variables on ESG score prediction. To this end, the study utilized ESG scores of companies listed in the STOXX Europe 600 Index for the period 2016–2021, incorporating financial variables such as total revenue, EBITDA, ROE, net debt, intangible assets, and Conditional Beta (BETA) into the prediction model. The RF classification algorithm was employed to

predict the ESG score class of the following year. The findings indicated that the most influential predictors in ESG score classification were Total Equity, ROE, EBITDA, and BETA.

Edhrabooh et al. (2024) aimed to contribute to the existing body of knowledge by identifying key research gaps in the estimation of ESG scores using ML and by proposing innovative approaches to address them. In their analysis, six core studies also referenced in this thesis were examined and synthesized using the Miles Theoretical Research Gap Model. These studies, published between 2021 and 2023, include works by Chowdhury et al. (2023), Lin & Hsu (2023), Raza et al. (2022), Lee et al. (2022), D'Amato et al. (2022) and D'Amato et al. (2021). The analysis revealed major research gaps, such as the limited inclusion of relevant ESG features, underutilization of advanced ML models, lack of diverse datasets and sectoral variations, and insufficient use of up-to-date data.

In summary, while studies on the prediction of ESG scores using ML algorithms provide valuable insights, they also exhibit various methodological and empirical limitations. Many studies are constrained by specific time periods, sectors, or geographic regions and tend to reach conclusions without conducting systematic comparisons across algorithms. Moreover, the models employed often focus exclusively on predictive accuracy, overlooking interpretability components such as feature importance analysis. Additionally, the scarcity of datasets covering recent years and the lack of comparative analyses between developed and emerging markets are among the frequently noted gaps in the literature (Edhrabooh et al., 2024).

In this context, future research on ESG score prediction should not only focus on forecasting performance, but also emphasize the comparative evaluation of different ML algorithms, and the use of comprehensive and up-to-date datasets. Furthermore, it is essential for such models to identify the key variables influencing ESG scores, thereby enabling the generation of interpretable and policy-relevant insights for decision-makers.

This study builds on the research gaps outlined by Edhrabooh et al. (2024) by utilizing a long-term dataset (2008–2023), constructing a sample with sectoral variety from Turkish firms, and applying various ML algorithms to estimate ESG scores. The models were then compared to determine the best-performing ones. Furthermore, prediction accuracy was enhanced through the implementation of feature selection, hyperparameter optimization, and model validation procedures.

4. RESEARCH FRAMEWORK

4.1. Sample and Variables

This study aims to predict companies' ESG scores, quantitative indicators of their sustainability performance, by employing ML algorithms.

In the initial phase of the study, ESG scores of 574 companies listed on BIST between 2008 and 2023, with available ESG data, were collected. Subsequently, ESG scores were matched with financial indicators covering the same period. Companies operating in the financial sector were excluded from the sample due to differing financial reporting standards; additionally, observations with missing values in either financial or ESG data were excluded from the analysis. Following these steps, the final dataset used for regression analyses consisted of 95 companies and 512 firm-year observations.

In the research model, the dependent variable is the ESG score of companies. These scores are continuous variables ranging from 0 to 100, making them suitable for regression analysis. The ESG score is calculated based on three main dimensions: environmental (e.g., emissions, resource use, environmental innovation), social (e.g., employee rights, community involvement, product responsibility), and governance (e.g., management structure, shareholder rights, corporate social responsibility). The composite of these three dimensions forms the final ESG score, which represents a company's overall sustainability performance. The ESG scores used in this study were obtained from the Thomson Reuters Eikon Refinitiv database. Refinitiv provides one of the most comprehensive ESG datasets available, with data extending back to 2002. Its ESG methodology is based on a detailed assessment framework comprising three pillar scores (Environmental, Social, and Governance), covering 10 main themes and more than 870 individual ESG metrics (LSEG, 2024).

The independent variables consist of 31 financial indicators representing the companies' capital structure and operational characteristics, including working capital, EBIT, ROE, intangibles, advertising expense, and basic EPS, among others. Detailed descriptions of all variables used in the model are presented in Appendix Table A.1.

Additionally, sector and year dummy variables were included in the model to control sectoral and temporal effects (Crespi & Migliavacca, 2020; Silva & Mascena, 2024).

4.2. Research Hypotheses

ESG scores have become a key indicator for measuring corporate sustainability performance. Numerous studies have demonstrated that companies' financial structures, particularly indicators such as profitability, leverage, and operational efficiency, play a significant role in shaping ESG performance. For example, profitability is expected to have a positive effect on ESG scores, as companies with greater financial resources are more capable of making the necessary investments to enhance their ESG performance (Drempetic et al., 2020). Furthermore, operational efficiency metrics, such as asset turnover ratios, have been directly associated with performance, especially in the environmental dimension (D'Amato et al., 2021).

Within the scope of this study, the hypothesis that financial indicators have a significant impact on ESG scores is tested. In this context, the research hypothesis developed is presented below:

H1: Firms' financial indicators have a statistically significant impact on their ESG scores.

In line with this hypothesis, the research model treats ESG scores as the dependent variable, while financial and operational structure indicators are included as independent variables. The model has been estimated using various ML algorithms, and the predictive performance of these models has been evaluated through cross-validation and performance metrics such as RMSE, MAE, and R^2 .

4.3. Data Cleaning and Preprocessing

Data cleaning and preprocessing is the process of transforming raw data into a consistent and meaningful format suitable for analysis by ML algorithms. The primary motivation behind this step lies in the fact that real-world data often contains missing, inconsistent, erroneous, or non-systematic structures, making it ineffective to expect ML algorithms to learn directly from such data (Si, 2023). Prior to the application of ML models, various preprocessing steps were undertaken to improve data quality and render the dataset suitable for analysis.

First, observations with completely missing ESG score information were excluded from the dataset. Subsequently, rows with more than 90% missing values were eliminated on the grounds that they did not contribute meaningful information. These steps were carried out to enhance analytical consistency and reinforce methodological reliability (Hsu et al., 2025).

In the data set used for analysis, missing values ranging from 10% to 30% were identified for certain variables. Tree-based algorithms are particularly tolerant of such missingness (Batista

& Monard, 2003). After imputing missing values with the mean, the Winsorization method was applied to reduce the impact of outliers on the model. Through this sequential approach, imbalances in the data were addressed, and mean imputation was adopted as a practical and efficient strategy. This method ensured the preservation of relevant features without introducing significant bias or compromising model performance (Acuna & Rodriguez, 2004).

Following the handling of missing values, another critical data cleaning step (winsorization) was applied. Winsorization is a technique that transforms statistics by capping extreme values to reduce the influence of outliers without removing them entirely from the dataset (Nyitrai & Virag, 2019; Demir & Sahin, 2023). In this study, winsorization was applied using the 1st and 99th percentiles as thresholds (Chowdhury et al., 2023). This technique aims to preserve the overall distributional structure by pulling outlier values closer to acceptable boundary values rather than excluding them. In doing so, extreme observations that could negatively affect model performance were mitigated, and the statistical properties of the data were maintained. Outliers and minor perturbations in the data hinder the model's ability to learn effectively, as the model may treat both distorted and clean data with equal importance, especially during the initial training phase (Bastounis et al., 2021; Ilyas et al., 2019). Therefore, winsorization, as a preprocessing method, enables the model to learn more reliably by reducing the influence of outliers before training begins. This, in turn, helps preserve the relationship between the target variable and independent variables and enhances the model's generalization capacity (Sharma & Chatterjee, 2021).

After data cleaning was completed, the dataset was split into 80% training and 20% test sets for the purpose of developing and evaluating ML models (Park et al., 2024). This division allowed for testing on previously unseen data, thereby enabling a more objective evaluation of the models' predictive performance.

Before transitioning to the modeling phase, hyperparameter optimization was conducted. In ML models, hyperparameters are settings that must be defined manually before the training process and directly influence the model's structure, learning speed, and generalization capability (Barysheva & Vasilevsky, 2023). These parameters may include tree depth, learning rate, and minimum leaf size, among others, and play a decisive role in the model's accuracy, generalization ability, and computational efficiency. Proper hyperparameter tuning not only improves the model's accuracy but also enhances its generalizability across different datasets (Barysheva & Vasilevsky, 2023). Particularly in deep learning and tree-based models, the

appropriate selection of hyperparameters helps prevent issues such as overfitting and underfitting (Chowdhury et al., 2023; Aue et al., 2025).

For each regression algorithm used in the modeling phase, hyperparameter optimization was individually performed using the RandomizedSearchCV method. This method samples random combinations within predefined parameter ranges and aims to identify the optimal parameter set that improves model performance (Huotari, 2024). Since RandomizedSearchCV only explores a subset of the hyperparameter space, it is particularly advantageous for reducing computational costs when working with high-dimensional and large-scale datasets (Akodia et al., 2024). The hyperparameters identified through RandomizedSearchCV are presented in Table B.1 in the Appendix.

In conclusion, this optimization process significantly improved both the accuracy and overall reliability of the employed models, contributing to the generation of more robust, balanced, and high-predictive-power outcomes.

4.4. Feature Selection and Modeling

The dataset employed in this study exhibits a high-dimensional structure due to the inclusion of numerous financial indicators. This structure increases the risk of overfitting in ML models; therefore, a feature selection process was undertaken. Moreover, since the performance of ML algorithms is largely dependent on the quality and relevance of the selected variables, both statistical and model-based approaches were adopted for feature selection.

4.4.1. Feature selection process

In ML models, high correlation among variables can lead to multicollinearity, which may negatively affect model performance. Accordingly, the first step in the analysis involved examining the correlation matrix among the independent variables. Out of the 62 features included in the dataset, variables with a correlation coefficient of 70% or higher with one another were screened, and only one from each highly correlated pair was retained. As a result, 31 variables were selected for the modeling process. The Pearson correlation matrix illustrating the correlations among these variables is presented in Figure C.1 in the Appendix.

Following this correlation-based preliminary screening, the final variable selection was performed using SHAP values. SHAP analysis identified the financial indicators that had the most significant impact on the prediction of ESG scores, thereby enhancing both the interpretability and predictive power of the model (Zhou & Wang, 2025).

4.4.2. Modeling approach

ML offers significant advantages over traditional statistical methods due to its capacity to identify patterns within large and complex datasets and to make high-accuracy predictions (Dipierro et al., 2025; Palynska et al., 2024). ML algorithms are particularly preferred in cases such as ESG score prediction, where numerous variables interact and nonlinear relationships are prevalent. This is because they are capable of capturing complex, nonlinear interdependencies between variables, offering advantages such as flexibility, automatic feature learning, and strong generalization capabilities (D'Amato et al., 2022).

In this context, supervised learning algorithms used for regression problems are highly suitable for predicting ESG scores, which are continuous variables. These algorithms provide a more robust approach in dealing with challenges frequently encountered in financial data, such as noise, multicollinearity, and nonlinearity (D'Amato et al., 2021). Furthermore, ML algorithms allow for hyperparameter optimization, enabling systematic enhancement of model performance; the proper tuning of these hyperparameters directly impacts the overall accuracy and generalizability of the model (Assael, 2023; Nijhof, 2023).

In this study, various regression algorithms were implemented to predict ESG scores. These included tree-based models such as RF Regressor, Gradient Boosting Regressor, LightGBM Regressor, and XGBoost Regressor; linear-based Ridge Regression, and instance-based KNN Regressor. These models were selected based on their differing abilities to capture both linear and nonlinear relationships. Moreover, thanks to their openness to hyperparameter tuning, these models adapt well to different data structures and allow for more accurate predictions of complex target variables such as ESG scores.

The models used in the analysis are explained individually in the following sub-sections.

4.4.2.1. Random forest regressor

RF is an ensemble learning method that involves training multiple models simultaneously for the same problem (Palynska et al., 2024). Each decision tree within the forest is trained on a different subset of the data, and the final prediction is obtained by averaging the individual tree outputs.

Advantages (Palynska et al., 2024):

- It is robust to outliers and missing data.
- It provides a ranking of feature importance.

- It tends to outperform other ML models, particularly in complex situations involving large datasets with many features, due to its ability to reduce overfitting, improve generalization, and produce robust predictions.

4.4.2.2. Gradient boosting regressor

Gradient Boosting is a ML algorithm that falls under the category of ensemble learning methods, similar to Random Forest (Srivalli & Sumanthi, 2025; Palynska et al., 2024; Barysheva & Vasilevsky, 2023). The algorithm follows a sequential structure in which each new model attempts to learn the errors of the previous one. Initially, a simple decision tree is constructed as the base model, and its residual errors are calculated. Subsequent models are then built to correct these residuals. This process is repeated multiple times, with each tree being trained to minimize the remaining errors. The final prediction is obtained by combining the predictions of all models, weighted according to their ability to correct the errors (Palynska et al., 2024; Barysheva & Vasilevsky, 2023).

Advantages:

- Provides high predictive accuracy.
- Capable of capturing complex patterns through iterative learning.
- Offers flexibility through hyperparameter tuning.

4.4.2.3. LightGBM regressor

LightGBM is an algorithm optimized for large datasets and designed to operate faster than traditional gradient boosting methods. One of its core features is the use of a leaf-wise growth strategy, as opposed to the level-wise approach used in conventional boosting algorithms.

Advantages:

- Offers practical features such as increased speed and improved accuracy (Assael, 2023).
- Naturally handles categorical variables without the need for extensive preprocessing.
- Demonstrates high performance on large-scale datasets (Chang et al., 2025).

4.4.2.4. XGBoost regressor

XGBoost is an optimized version of gradient boosting algorithms that utilizes second-order derivatives of the loss function to minimize prediction errors (Alsayyad & Fadel, 2025). The model's core objective is to prevent overfitting through regularization mechanisms.

Advantages:

- Offers fast learning and high predictive accuracy due to its exceptional capability to handle large-scale datasets (Khan et al., 2024; Park et al., 2024).
- Applies both L1 (Lasso) and L2 (Ridge) regularization to control model complexity and reduce the risk of overfitting (Park et al., 2024).
- Demonstrates superior performance in classification tasks with high-dimensional and imbalanced data (Chang et al., 2025).

4.4.2.5. Ridge regression

Ridge Regression is a variant of Linear Regression that typically applies L2 regularization (penalty) to reduce model complexity and address issues such as multicollinearity (Chicco et al., 2021; Dağıstanlı & Çelik, 2023; Burman, 2022).

Advantages:

- Effective against multicollinearity.
- Produces simple and interpretable results.

4.4.2.6. KNN regressor

The KNN algorithm is a non-parametric, instance-based regression method used to predict a continuous target variable for a new data point (Chowdhury et al., 2023). In the KNN algorithm, no explicit model is built during the training phase; instead, data points are simply stored, and the modeling process occurs during the prediction stage (De Lucia et al., 2020).

For a new observation to be predicted, the algorithm identifies the 'k' closest neighbors from the training dataset. This neighborhood is typically defined using distance metrics such as Euclidean or Manhattan distance (Ramezan et al., 2021; De Lucia et al., 2020). In the context of regression, the prediction is made by calculating either the arithmetic mean or the distance-weighted average of the target values of these 'k' neighbors (De Lucia et al., 2020; Chowdhury et al., 2023). Due to its non-reliance on linear assumptions and its ability to capture local patterns, KNN offers a flexible solution, particularly for complex and structurally non-linear datasets.

Advantages:

- Flexible due to its non-parametric structure.
- Effective for small datasets.

4.5. Model Performance Evaluation Methods

To evaluate the effectiveness of ML models, model accuracy and generalizability were assessed using four main metrics: RMSE, MAE, MSE, and R^2 .

4.5.1. RMSE (Root Mean Squared Error)

RMSE serves as one of the most common metric to assess prediction accuracy in regression-based models (Dipierro et al., 2025; Si, 2023). RMSE shows how much the predicted values typically differ from the actual values, giving a clear sense of the model's average error (Raza et al., 2022; Alsayyad & Fadel, 2025). It helps in assessing how different models handle prediction errors for specific features, making it suitable for model comparison tasks (Lin & Hsu, 2023).

The mathematical Formula (4.1) for RMSE is given below (Lin & Hsu, 2023; Assael, 2023; Chicco et al., 2021):

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$
(4.1)

n : Total number of observations

y_i : Actual (observed) value

$\hat{y}_i$: Predicted value

$(y_i - \hat{y}_i)^2$: Squared errors

$\sum_{i=1}^n$: Sum of all squared errors

$\sqrt{}$: Square root of the MSE

RMSE is computed by first squaring the differences between predicted and actual values, averaging these squared differences, and then taking the square root of that average (Chicco et

al., 2021; Alsayyad & Fadel, 2025). Since each error is squared, larger deviations are penalized more heavily; for this reason, RMSE is often preferred in contexts where large errors are particularly critical.

One of the main advantages of RMSE is that the resulting value is expressed in the same unit as the predicted variable (Assael, 2023). This property allows for a more intuitive and interpretable understanding of the magnitude of prediction errors. The RMSE value ranges from 0 to positive infinity, where values approaching zero indicate high predictive accuracy (Chicco et al., 2021; Assael, 2023). As such, RMSE serves as an effective and practical reference for comparing the predictive accuracy of different models (Lin & Hsu, 2023).

RMSE belongs to the family of distance-based error metrics and is closely related to MAE and MSE (Mean Squared Error) (Alsayyad & Fadel, 2025; Chicco et al., 2021). While MSE is more sensitive to large errors, RMSE, being the square root of MSE, offers a more interpretable reflection of that sensitivity and typically yields similar model rankings (Chicco et al., 2021).

Nonetheless, RMSE also has several limitations. Its most significant drawback lies in its high sensitivity to outliers. Due to the squaring of errors, extreme values in the data can disproportionately inflate the total error (Nijhof, 2023). This characteristic implies that RMSE, when used alone, may not fully represent model performance in heterogeneous data sets. Therefore, RMSE is often evaluated alongside complementary metrics such as MAE or MSE to provide a more comprehensive assessment (Alsayyad & Fadel, 2025; Chicco et al., 2021).

In conclusion, RMSE is a powerful metric for regression problems aimed at predicting continuous numerical values, as it captures both the magnitude and severity of prediction errors. However, in order to ensure a holistic evaluation of model performance, it is crucial to consider RMSE's advantages and limitations in conjunction.

4.5.2. R^2 (Determination Coefficient):

R^2 , also known as the Coefficient of Determination, indicates how much of the variation in the target variable is accounted for by the explanatory variables (Lai et al., 2020). In other words, the R^2 value indicates how much of the total variability in the data is captured by the model. This value ranges between 0 and 1. The nearer R^2 gets to 1, the more accurately the model explains the data and makes predictions. If R^2 is close to 0, it means the model doesn't do a good job in explaining the dependent variable. R^2 measures how much of the total variance (i.e., the extent to which the actual values deviate from their mean) is accounted for by the values

predicted by the model. For instance, an R^2 value of 0.85 indicates that the model explains 85% of the total variation in the dependent variable.

The mathematical Formula (4.2) of the R^2 coefficient is as follows;

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (4.2)$$

y_i : Actual (observed) value

$\hat{y}_i$: Predicted value

$\bar{y}$: Mean of the actual values

SS_{res} : Residual Sum of Squares

SS_{tot} : Total Sum of Squares

A high R^2 value indicates strong model prediction performance (Jumin et al., 2020); however, a high R^2 does not necessarily imply a good model. This is because R^2 does not provide direct information about the magnitude of prediction errors, and although it shows the percentage of variance explained, it does not fully reflect the model's accuracy (Priya, 2021).

To sum up, although R^2 is a valuable indicator of how well the model accounts for the dependent variable, it should be considered together with error metrics like MAE, MSE, and RMSE to obtain a fuller understanding of model performance.

4.5.3. MAE (Mean Absolute Error):

MAE is a commonly used error metric for evaluating the prediction accuracy of regression models (Chicco, 2021). This metric shows whether the predicted results are close enough to the actual outcomes. It is calculated by taking the average of the absolute values of the individual prediction errors, thereby measuring the average deviation regardless of the direction (positive or negative) of the errors. MAE represents the mean absolute difference between predicted and actual values, and its value ranges from 0 to positive infinity (Lin & Hsu, 2023).

MAE is typically calculated using the following Formula (4.3) :

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4.3)$$

Here:

y_i : Actual values (i.e., observed values)

$\hat{y}_i$: Predicted values by the model

n : Number of observations in the dataset

$|y_i - \hat{y}_i|$: Absolute difference between actual and predicted values

MAE has several fundamental characteristics. First, it has a simple structure; it is easy to implement and interpret. MAE is always a positive value since it uses absolute differences in its calculation. It treats all errors equally regardless of their magnitude or direction (Huotari, 2024). This means that large errors do not have a disproportionately higher impact than small ones.

Among its advantages, MAE stands out for being easy to interpret, as it expresses errors in the same units as the original variable, making the error level more intuitive. Additionally, because it treats each error equally, MAE is more robust to outliers and does not excessively penalize large deviations (Assael, 2023; Chicco et al., 2021). On the other hand, MAE also has limitations. For instance, it does not provide information about the distribution of errors, meaning it cannot distinguish whether the errors are frequent or large in magnitude.

MAE stands out as a useful metric when assessing how accurately a model performs across all predictions. However, due to its equal weighting of all errors, in some cases more sensitive metrics such as MSE or RMSE may be preferred (Lin & Hsu, 2023; Lee et al., 2022; Wang et al., 2024; Alsayyad & Fadel, 2025).

4.5.4. MSE (Mean Squared Error):

MSE is a common way to check how close a model's predictions are to the actual results in regression analysis. This metric calculates the average of the squared differences between the predicted values and the actual values (Priya, 2021). Its primary purpose is to quantify the extent of the model's errors, with particular sensitivity to larger errors.

The mathematical Formula (4.4) for MSE is as follows (Alsayyad & Fadel, 2025):

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4.4)$$

n : Refers to the number of observations (samples)

y_i : Represents the actual (observed) value

$\hat{y}_i$: Denotes the value predicted by the model

$(y_i - \hat{y}_i)^2$: The squared error for each prediction

$\sum_{i=1}^n$: The total of these squared differences across all observations

$\frac{1}{n}$: The sum of squared errors is divided by the number of observations to obtain the average

The formula for MSE shows the total of the squared differences between predicted and actual values. MSE is the result of dividing this total by the number of observations (n) (Si, 2023). By squaring the errors, it prevents positive and negative differences from canceling each other out. This ensures that MSE reflects not only the presence of errors but also their magnitude.

MSE has several important properties. First, it is always positive because squaring eliminates negative signs (Tatachar, 2021). Additionally, it is more sensitive to large errors; that is, when a prediction error is large, it disproportionately affects the MSE value. This characteristic can be seen both as an advantage and a disadvantage.

Among its advantages is its strong ability to highlight large errors. Since squaring emphasizes bigger discrepancies, the model is implicitly encouraged to avoid them (Chai & Draxler, 2014). Moreover, MSE is mathematically differentiable, making it suitable for use in many optimization algorithms, such as gradient descent.

However, MSE also has certain limitations. The most notable is its high sensitivity to outliers. Since squaring large errors significantly inflates the overall MSE, the presence of outliers can distort the evaluation of model performance (Huotari, 2024). Furthermore, the result of MSE is often expressed in squared units, which differ from the original scale of the variable and can complicate interpretation.

In conclusion, MSE is a performance metric that gives more weight to the magnitude of prediction errors, making it particularly effective during model training. However, it is less interpretable than MAE due to its scale and greater sensitivity to extreme values. While MAE is commonly used as an evaluation metric, MSE is typically employed as a loss function during model training (Krappel et al., 2021). Therefore, depending on the model's purpose and the nature of the dataset, it is recommended to assess model performance using not only MSE but also other error metrics such as MAE (Si, 2023).

5. FINDINGS

Summary statistics of the 31 variables used in the modeling process are presented in detail in Appendix Table C.1 of the study. In order to evaluate the temporal evolution of the dependent variable, ESG scores, between 2008 and 2023, annual average ESG scores were calculated and visualized in Figure 5.1.

Upon examining Figure 5.1, it is evident that ESG scores exhibited a consistent upward trend over the years. The average score, which was 28.54 in 2008, rose significantly to 71.49 by 2023. This notable increase can be attributed to the growing adoption of corporate sustainability practices, heightened regulatory pressure, and a shift in investor sentiment. The upward trend becomes particularly pronounced after 2015. Between 2015 and 2023, ESG scores increased by approximately 45%, a rise that can be explained by the rising awareness of corporate sustainability, intensifying regulatory demands, and the evolving expectations of stakeholders. The sharp increase observed particularly after 2020 is especially noteworthy. In 2020, the average ESG score was 57.93; driven by the impact of the COVID-19 pandemic and a surge in interest toward ESG-oriented policies, the score reached 71.49 by 2023. Conversely, a decline was observed in 2018, when the average score dropped to 49.86. Such fluctuations may be explained by macroeconomic and political volatility, sector-specific crises, or inconsistencies in corporate reporting behavior.

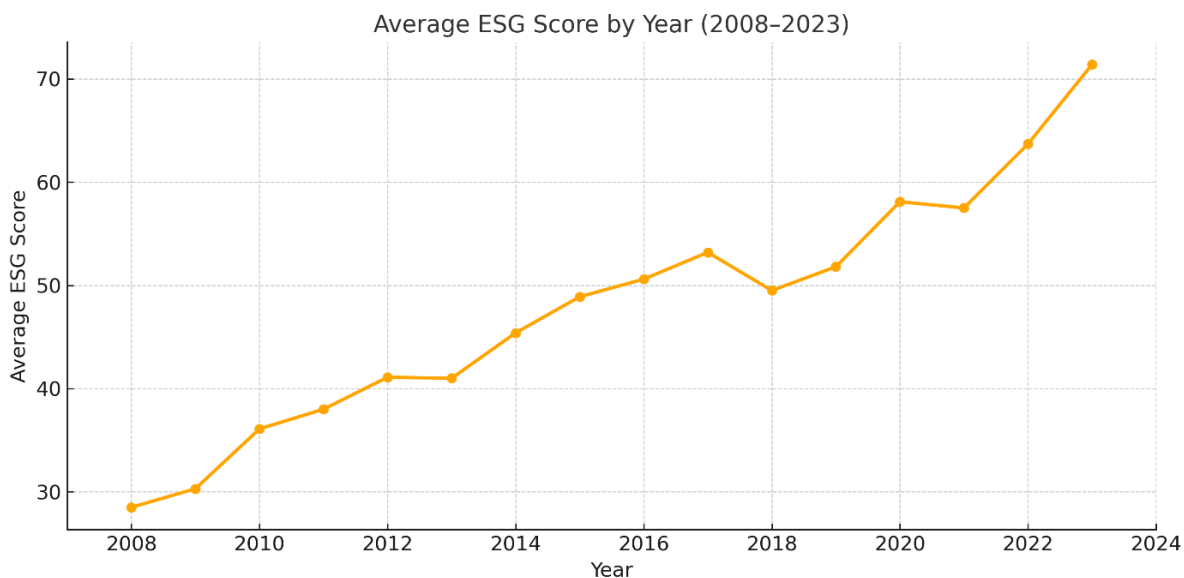


Figure 5.1. Yearly evolution of ESG scores

5.1. Model Performance Results

As detailed in the section titled *Data Cleaning and Preprocessing*, baseline modeling was first conducted for each algorithm using the optimal hyperparameters obtained through the hyperparameter tuning process. At this stage, the dataset was split into training and testing sets; each model was trained with the specified hyperparameters, and their performance was evaluated accordingly.

Subsequently, in order to assess the generalizability and robustness of the models, a 5-fold cross-validation procedure was implemented. Through this method, the prediction performance of each model was re-evaluated and systematically compared.

Following all stages of analysis, XGBoost emerged as the highest-performing algorithm. Accordingly, SHAP analysis was conducted based on this model to identify the most influential financial variables contributing to the prediction of ESG scores.

5.1.1. Baseline modeling

The optimal hyperparameter combinations obtained through the tuning process were employed during the baseline modeling stage. At this stage, the data was partitioned into training (80%) and testing (20%) sets. The models were trained on the training portion and then tested to evaluate their prediction performance. The size of the training sample plays a significant role in determining the accuracy of the models (Ramezan et al., 2021). This baseline modeling provides a fundamental reference point for assessing the extent to which ESG scores can be predicted using financial data.

This step was conducted to quickly observe the overall accuracy level of the models and to detect any early signs of overfitting or underfitting. It also served as a benchmark for evaluating the degree of improvement achieved through more advanced modeling techniques.

Four key performance metrics were used for model evaluation: R^2 , RMSE, MAE, and MSE.

Table 5.1 below presents the performance metrics obtained from the baseline modeling process:

Table 5.1. Comparison of baseline model performance

Model	RMSE (baseline)	MAE (baseline)	R^2 (baseline)	MSE (baseline)
RandomForest	14.395	11.020	0.564	207.223
GradientBoosting	15.270	11.511	0.509	233.174
XGBoost	15.524	12.351	0.493	241.018

Table 5.1. (continued) Comparison of baseline model performance

Model	RMSE (baseline)	MAE (baseline)	R ² (baseline)	MSE (baseline)
LightGBM	15.602	11.877	0.488	243.450
KNN	19.779	15.530	0.177	391.241
Ridge	20.186	16.143	0.143	407.480

According to the performance metrics presented in Table 5.1, the highest performance was achieved by the RF Regressor model. This model demonstrated the greatest explained variance with an R² value of 0.564, while also exhibiting the lowest error rates with RMSE (14.39), MAE (11.02), and MSE (207.22) values. As shown in Table 5.1, the Gradient Boosting and XGBoost models also exhibited relatively high performance; however, their explained variance was inferior to that of the RF model. Conversely, Table 5.1 indicates that simpler models such as KNN and Ridge Regression were insufficient for predicting ESG scores, as evidenced by their low R² values and high error rates.

These findings reveal that ESG scores are structurally too complex to be explained solely by simple linear relationships. Consequently, despite the limited success of linear models, tree-based algorithms capable of capturing nonlinear relationships and modeling multilayered interactions among variables have demonstrated superior performance. This indicates that model flexibility and interactive learning capacity provide significant advantages in predicting ESG scores. The prominence of algorithms such as RF and Gradient Boosting confirms that the structural complexity inherent in ESG data can be more effectively analyzed using methods beyond traditional regression techniques

5.1.2. Modeling with 5-fold cross-validation

In evaluating ML models, it is insufficient for a model to perform well solely on the training data; it is also expected to demonstrate robust performance on new, unseen data. In this context, the k-fold cross-validation method is a widely used and reliable validation technique to assess a model's generalizability. This method partitions the dataset into k equally sized subsets; in each iteration, one subset is used as the test set while the remaining k-1 subsets serve as the training set. This process is repeated k times, ensuring that each observation is used once for testing and k-1 times for training (David et al., 2024; Wang et al., 2024).

To evaluate the overall model performance, 5-fold cross-validation was employed. This technique divides the dataset into five equal parts, sequentially setting aside one part as the test set and training the model on the remaining four. Consequently, the model's performance across different data subsets was measured, and its robustness against overfitting was tested (Chaklader et al., 2024; Alsayyad & Fadel, 2025). Table 5.2 below summarizes the average performance metrics obtained from cross-validation for each model ;

Table 5.2. 5-Fold cross-validation results

Model	RMSE (CV)	MAE (CV)	R ² (CV)	MSE (CV)
Random Forest	18.417	14.585	0.375	340.288
Gradient Boosting	18.668	14.805	0.351	349.073
XGBoost	18.250	14.835	0.379	334.335
LightGBM	18.501	14.782	0.360	343.031
KNN	22.692	18.581	0.033	515.446
Ridge	23.572	18.243	-0.041	561.638

According to Table 5.2, the highest R² (0.379) value was obtained with the XGBoost algorithm. This model also exhibited the lowest MSE (334.33) and the lowest RMSE (18.25) values. RF and LightGBM models demonstrated comparable performance levels, closely following XGBoost. The strong performance of RF is attributed to its robust results under high variance conditions, achieved through the ensemble of numerous decision trees. Although LightGBM offers a faster and lighter structure, it delivered a competitive outcome. These three models stand out as the most reliable algorithms for predicting ESG scores.

Conversely, Gradient Boosting, along with its optimized variants such as XGBoost and LightGBM, lagged behind with a lower-than-expected R² value (0.351) and higher error metrics. This indicates that the classic GBDT algorithm may have limitations, particularly in terms of scalability and efficiency. KNN and Ridge models exhibited poor performance, characterized by both low R² values and high error metrics. The negative R² value of the Ridge model indicates its failure to explain the variance in the data. These findings support the conclusion that due to the structural complexity of ESG scores, linear models have limited capacity, whereas nonlinear and tree-based algorithms yield more stable results.

5.1.3. Comparison of training-test and 5-fold cross-validation performance

There are significant differences between baseline modeling based on a single training-test split and cross-validation methods conducted on multiple subsets of the data during model evaluation. In this study, both baseline modeling and 5-fold cross-validation were applied for all models using the hyperparameter values obtained after optimization, and the results were analyzed comparatively.

Table 5.3 below presents the performance metrics produced by both evaluation strategies side by side. Upon comparison, it is observed that all models demonstrate higher R^2 values and lower RMSE, MAE, and MSE values in the baseline modeling results. For example, the RF model shows an R^2 of 0.564 in baseline modeling, which decreases to $R^2 = 0.375$ under cross-validation. Similarly, for XGBoost, this drop is from 0.493 to 0.379. Similar notable performance decreases are observed in other models as well.

Table 5.3. Train-test and 5-fold cross-validation performance metrics

Model	RMSE (Baseline)	RMSE (CV)	MAE (Baseline)	MAE (CV)	R^2 (Baseline)	R^2 (CV)	MSE (Baseline)	MSE (CV)
Random Forest	14.395	18.417	11.020	14.585	0.564	0.375	207.223	340.288
Gradient Boosting	15.270	18.668	11.511	14.805	0.509	0.351	233.174	349.073
XGBoost	15.524	18.250	12.351	14.835	0.493	0.379	241.018	334.335
LightGBM	15.602	18.501	11.877	14.782	0.488	0.360	243.450	343.031
KNN	19.779	22.692	15.530	18.581	0.177	0.033	391.241	515.446
Ridge	20.186	23.572	16.143	18.243	0.143	-0.041	407.480	561.638

The primary reason for this discrepancy is that the test data used in the training-test modeling tends to share similar characteristics with the distribution the model has previously seen. This situation can lead the model to learn data-specific patterns and produce more "optimistic" results on the test set. However, such outcomes do not accurately reflect the model's generalization performance in real-world scenarios.

On the other hand, cross-validation evaluates the model's generalization capacity in a more unbiased and balanced manner by splitting the dataset into different combinations of training and test sets. This approach ensures more reliable and robust results, especially when the dataset is imbalanced or when the distribution of variables varies over the years.

In this context, the lower performance metrics obtained from cross-validation compared to baseline modeling indicate that the model's risk of overfitting is controlled and that the real-world predictive performance is more accurately represented. Therefore, in this study, cross-validation metrics were prioritized in the model selection and variable analysis processes.

5.2. Key Factors Determining ESG Scores

It is not sufficient for ML models to merely achieve high prediction accuracy; it is also critically important, especially for complex constructs like sustainability performance, that the model can explain which variables it considers and to what extent in its decision-making process. In this context, the SHAP method was employed in this study to enhance the transparency and interpretability of the model's decision mechanism.

SHAP is a technique used to explain the predictions or outputs of ML models (D'Amato et al., 2022; Zhou & Wang, 2025). By quantifying each feature's contribution to the model, SHAP reveals how influential each variable is on the predictions. Particularly well-suited for tree-based models, SHAP provides a detailed explanation by balancing individual and interaction effects of variables.

In this study, SHAP analysis was conducted on the highest-performing XGBoost algorithm to generate a variable importance ranking. The calculation of SHAP values was based on the results of 5-fold cross-validation. This analysis was utilized to understand which ESG factors exert the most significant impact on ESG performance predictions (Cots, 2024). Furthermore, the relationships between feature values and SHAP values were examined to elucidate how specific feature values influence the predictions (Assael, 2023).

As shown in Figure 5.2, the most influential variable in predicting ESG scores was “Intangibles Net.” This variable represents the value of the company's intangible fixed assets (e.g., patents, trademarks, software, R&D outputs) and reflects the long-term value creation aspect of the ESG score. Following this, the “Year” variable's high SHAP value indicates a systematic increasing trend in ESG scores over time.

The third-ranked variable, “Basic EPS Excluding Extraordinary Items,” demonstrates a direct relationship between earnings per share from ordinary operations and ESG performance. This suggests that financial solidity and operational efficiency are perceived as reflections of corporate sustainability.

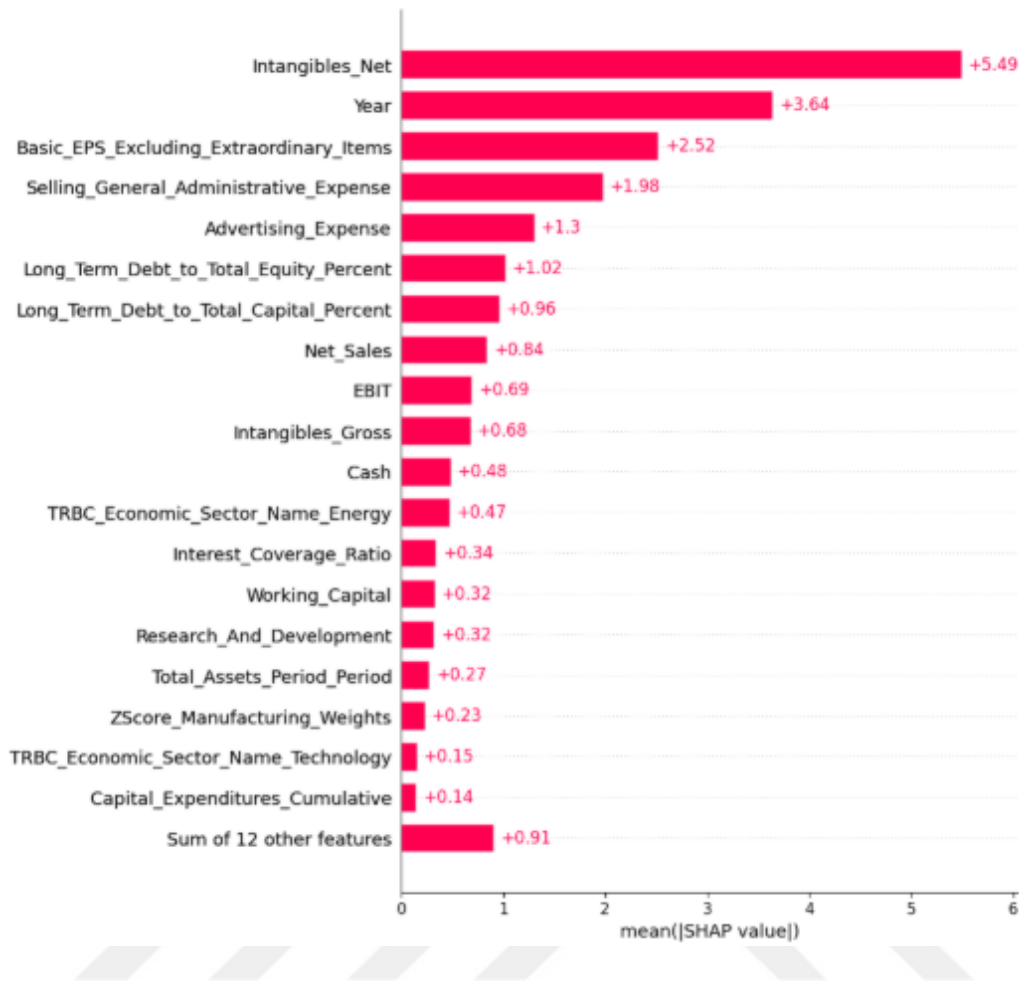


Figure 5.2. SHAP value ranking of variables determining ESG scores

Expense items such as “Selling General Administrative Expense” and “Advertising Expense,” which follow, indicate that the company’s activities in corporate communication, marketing, and transparency affect ESG scores. Moreover, financial structure indicators like “Long Term Debt to Total Equity Percent” reveal the relationship between capital structure and sustainability.

Although ranked lower, classical financial performance indicators such as “Net Sales,” “EBIT,” “Cash,” and “R&D Expenditures” were also observed to have significant effects on ESG scores. This finding highlights that ESG performance is directly linked not only to social or environmental indicators but also to the fundamental financial structure of the company.

As a result of this analysis, it was determined that the key factors influencing ESG scores consist both of indicators related to value-creation capacity, such as intangible assets, and financial structure elements including indebtedness, profitability, and operational expenses. SHAP

analysis not only made the model's decision-making process more transparent but also unveiled the multidimensional nature of ESG performance.

5.3. Model Reliability and Consistency

For ML models to be confidently utilized in academic and applied studies, achieving high prediction accuracy alone is insufficient. The model is also expected to maintain similar levels of success across different data subsets, demonstrating consistent and generalizable performance. In this context, the reliability of the model has been examined multidimensionally through both the methodological framework and the results obtained.

First, hyperparameter optimization was performed for all models using the RandomizedSearchCV method. This ensured that models were evaluated not with random configurations, but with the best parameter combinations determined for each (Barysheva & Vasilevsky, 2023). This step not only facilitated a fair comparison among models but also allowed each algorithm to maximize its capacity.

Second, through 5-fold cross-validation, each model was tested on different training-test splits, thus measuring the model's generalizability without relying on a single random test split. The results revealed that especially the XGBoost, RF, and LightGBM models exhibited consistent performance across folds. These models showed no significant deviations in R^2 values and produced similar error rates across different folds. This indicates that the models are not overly sensitive to data distribution and are likely to replicate comparable success in real-world scenarios.

Furthermore, the SHAP analysis conducted to test the model's intrinsic reliability demonstrated that the decision-making processes of the model are based on logical and economically meaningful foundations. The fact that the variables most influential in predicting ESG scores consist of theoretically expected indicators, such as the company's intangible assets, periodic earnings, and leverage ratios, indicates that the model possesses not only statistical but also semantic consistency (Zhou & Wang, 2025).

Considering all these factors together, it can be stated that the developed modeling approach provides not only high accuracy but also high consistency, low variance, and interpretability. This makes these models reliable tools for predicting ESG performance, which is a complex and strategic multidimensional concept.

6. DISCUSSION AND CONCLUSION

6.1. Evaluation of Findings

This study reveals a significant and explanatory relationship between firms' financial indicators and their ESG scores. Among the various ML algorithms employed, the highest predictive performance was achieved by the XGBoost model, which attained an average R^2 value of 0.379 after cross-validation. This finding demonstrates that ESG scores can be predicted to a certain extent through financial variables, highlighting the potential of financial fundamentals to reflect corporate sustainability practices.

To enhance model transparency and interpretability, SHAP analysis was conducted, which detailed the relative importance of variables influencing ESG scores. The variables with the greatest impact included intangible fixed assets (Intangibles Net), year (Year), basic earnings per share excluding extraordinary items (Basic EPS Excluding Extraordinary Items), selling, general and administrative expenses (SG&A Expense), advertising expenses (Advertising Expense), and long-term debt ratios. This outcome underscores the critical role of structural factors such as a firm's asset composition, expenditure patterns, and financial risk profile in determining ESG scores, going beyond traditional cash-flow-based approaches.

Most of the top-ranked variables identified by the SHAP analysis have been extensively examined in the literature with empirical support linking them to ESG performance. For instance, Joshi and Joshi (2024) also found a positive influence of intangible fixed assets (Intangibles Net) on ESG scores. The determinant role of EPS and operating expenses on ESG scores has been reported by Lin and Hsu (2023). Similarly, D'Amato et al. (2021) documented the positive effects of advertising expenditures (Advertising Expense) on overall ESG scores as well as on the E, S, and G components. Moreover, the prominence of financial leverage ratios in ESG predictions has been empirically validated by Chowdhury et al. (2023). These findings indicate that the variables prominent in the SHAP analysis are structurally associated with ESG scores.

In the heterogeneous ensemble model developed by Krappel et al. (2021), which includes XGBoost, CatBoost, and feedforward neural networks, intangible assets and sectoral effects emerged as key predictors of ESG scores, consistent with the similar results obtained in the SHAP analysis of this thesis. Additionally, Velte (2017) emphasized the interaction of the governance dimension with financial performance, and Friede et al. (2015), in their meta-analysis, identified an overall positive relationship between ESG and financial success. Within

this context, the findings of this study align with prior research emphasizing the importance of financial variables influencing ESG scores.

Compared to other ML based ESG prediction studies in the literature, this research exhibits several structural parallels. For example, Si (2023) reported significant effects of leverage ratios, sales revenues, and asset composition on ESG scores using RF algorithms. Likewise, the heterogeneous ensemble models (CatBoost, XGBoost, and neural network combinations) developed by Krappel et al. (2021) provided an important alternative to enhance ESG score prediction accuracy, highlighting the critical roles of intangibles, EPS, and sectoral effects. The SHAP analysis indicating significant impacts from sectors such as energy and technology further confirms the consistency of this study's findings with the relevant literature.

In this thesis, the best performing model according to the selected performance metrics was the tree-based ensemble model, XGBoost. Similarly, Lin and Hsu (2023) compared algorithms including XGBoost, RF, ELM, and SVM, concluding that tree-based and ensemble models offer higher prediction accuracy and advantages in interpretability over linear models in ESG predictions.

Furthermore, the results of this study offer not only technical but also theoretical insights. While SHAP analysis renders the model's decision-making process transparent, it also reveals the directionality and weight of variables affecting ESG scores. The effects of indicators such as intangibles, EPS, SG&A expenses, advertising costs, and long-term debt ratios on ESG scores appear to relate directly not only to financial performance but also to the firm's strategic sustainability approach.

These findings are consistent with Stakeholder Theory, which posits that firms bear responsibility not only to their shareholders but to all stakeholders, managing resources accordingly. Intangibles such as brand value and human capital directly affect stakeholder relationships and corporate reputation; hence, a positive association with high ESG scores is expected. The prominent SHAP impact of the intangibles variable in this study supports this theoretical framework.

However, the relatively lower influence of some expected indicators (e.g., net income or total assets) suggests that ESG scoring is shaped less by financial size and more by qualitative strategic choices. This can also be explained by Legitimacy Theory, which suggests that corporate ESG performance may sometimes be symbolic or superficial. According to this

theory, firms develop visible ESG actions in certain areas to maintain social legitimacy, which may cause some financial variables to exert less influence than anticipated.

6.2. Recommendations for Policies and Practices

The findings of this research demonstrate that ESG performance is strongly dependent not only on environmental and social factors but also on the financial structure of firms. This implies that sustainability strategies should not be limited solely to environmental and social objectives; rather, financial management practices must also be integrated into these strategies. For corporate managers, these findings necessitate the development of strategic approaches toward financial sustainability, especially considering that indicators such as free cash flow, leverage ratios, and operational profitability can directly influence ESG scores.

Moreover, the adoption of transparent, standardized, and comparable systems for reporting financial data related to ESG will not only enhance investor confidence but also positively impact the consistency of evaluations conducted by ESG rating agencies. In this context, decision support systems developed using explainable artificial intelligence tools like SHAP can contribute to the management of ESG strategies in a more data-driven and traceable manner. Thus, managers can optimize internal financial dynamics while simultaneously strengthening sustainability scores.

From the perspective of policymakers and regulators, regulations aimed at improving ESG performance should not be confined only to environmental and social parameters but should be expanded to encompass the overall financial resilience of companies. In this regard, tax incentives, financing facilities, and credit policies that encourage sustainable investments and high ESG performance gain critical importance. Furthermore, to ensure the consistency and transparency of methodologies used in ESG scoring, it is essential to establish clear standards regarding how and to what extent financial indicators should be utilized within these systems. Accordingly, ESG assessment processes should consider not only the aggregate scores but also the key variables contributing to these scores, enabling investors to make more informed decisions.

6.3. Limitations of the Study

As with any scientific inquiry, this study is subject to certain limitations stemming from both the characteristics of the dataset and the methodological approaches employed.

First, the analysis was conducted solely using data compiled on an annual basis. This constraint limited the ability to examine the short-term fluctuations in ESG scores or the immediate impacts of policy shifts in detail. Instantaneous effects of economic crises, abrupt regulatory changes, or sectoral shocks on ESG scores could not be analyzed with temporal sensitivity. Consequently, the relationships examined in this study were necessarily focused on long-term trends.

Second, the scope of the study was confined to companies operating in Turkey. Therefore, the generalizability of the findings to other countries or economic systems may be limited. Turkey's unique regulatory framework, market dynamics, and ESG implementation practices reduce the transferability of the results to different institutional contexts. This restricts the study's capacity to make cross-national or internationally comparative inferences.

Furthermore, the modeling process relied exclusively on financial indicators. By its very nature, ESG is a multidimensional construct encompassing environmental, social, and governance dimensions. However, due to data constraints, several variables that directly influence ESG performance, such as environmental emission levels, labor practices, social policies, or corporate governance structures, could not be included in the model. This shortcoming may have resulted in predictions that reflect only a financial perspective and may have limited the potential accuracy of the model.

6.4. Recommendations for Future Research

This study focused solely on predicting ESG scores based on financial indicators and achieved a meaningful level of predictive performance through the XGBoost model. However, given the multidimensional nature of the ESG concept and the structural limitations of the available data, it is recommended that future research develops more comprehensive and holistic approaches to overcome these constraints.

First, in order to more realistically model the factors influencing ESG scores, the inclusion of non-financial indicators in the analysis would be appropriate for broadening the model's scope. In particular, incorporating qualitative sustainability-related data, such as carbon emissions, energy consumption, workforce diversity, employee satisfaction, and corporate governance practices, into existing model structures by quantifying them could enhance predictive performance. As these variables constitute direct components of ESG scores, their use alongside financial ratios may yield stronger and more meaningful results.

Additionally, this study utilized only annual data, thereby limiting the ability to thoroughly analyze the effects of temporal variations. Therefore, future studies conducted using datasets organized on a quarterly or monthly basis would enable a more accurate investigation of short-term fluctuations in ESG scores and the impacts of sudden economic or political developments.

Diversifying data sources represents another critical area for improvement in future studies. ESG scores used in this thesis were obtained from a single data provider. However, the literature acknowledges methodological discrepancies and score inconsistencies among ESG ratings produced by different agencies. In this context, future research could comparatively analyze ESG scores from multiple rating providers, thereby generating extensive insights into rating consistency, methodological transparency, and inter-provider variations.

Lastly, as this study is limited to companies operating in Turkey, the findings are constrained to a specific context. In the future, applying similar modeling frameworks to firms in different countries and examining the relationship between ESG scores and financial indicators in light of cultural, institutional, and regulatory variations may contribute to a deeper understanding of the international dynamics of ESG. This would clarify to what extent ESG performance is influenced not only by internal corporate factors but also by external environmental conditions.

REFERENCES

- Acuna, E., & Rodriguez, C. (2004, February). The treatment of missing values and its effect on classifier accuracy. In *Classification, Clustering, and Data Mining Applications: Proceedings of the Meeting of the International Federation of Classification Societies (IFCS), Illinois Institute of Technology, Chicago*, 15–18 July 2004 (pp. 639-647).
- Adams, C. A., & Abhayawansa, S. (2022). Connecting the COVID-19 pandemic, environmental, social and governance (ESG) investing and calls for ‘harmonisation’ of sustainability reporting. *Critical Perspectives on Accounting*, 82, 102309.
- Alsayyad, M., & Fadel, S. M. (2025). Predicting ESG Scores Using Firms' Financial Indicators: A Machine Learning Regression Approach. DOI: <https://doi.org/10.20944/preprints202503.0096.v1>
- 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), 494.
- Akodia, J. A., Dzidonu, C. K., Boison, D. K., & Kisembe, P. (2024). Application of Random Search Methods in the Determination of Learning Rate for Training Container Dwell Time Data Using Artificial Neural Networks. *Intelligent Control and Automation*, 15(4), 109-124.
- Arora, A., & Sharma, D. (2022). Do environmental, social and governance (ESG) performance scores reduce the cost of debt? Evidence from Indian firms. *Australasian Accounting, Business and Finance Journal*, 16(5), 4-18.
- Artiach, T., Lee, D., Nelson, D., & Walker, J. (2010). The determinants of corporate sustainability performance. *Accounting & Finance*, 50(1), 31-51.
- Assael, J. (2023). Machine learning for ESG data in the financial industry (Doctoral dissertation, Université Paris-Saclay).
- Aue, T., Jatowt, A., & Färber, M. (2025, May). Predicting Company ESG Ratings from News Articles Using Multivariate Timeseries Analysis. In *Companion Proceedings of the ACM on Web Conference 2025* (pp. 1774-1780).
- Barysheva, I., & Vasilevsky, K. (2023). Time series forecasting using gradient boosting algorithms. In A. Nedzved & A. Belotserkovsky (Eds.), *Pattern Recognition and Information Processing (PRIP'2023). Artificial Universe: New Horizons: Proceedings of the 16th International Conference, Belarus, Minsk, October 17–19, 2023* (pp. 176–179). Minsk: Belarusian State University.
- Bassen, A.; Meyer, K.; Schlange, J. The Influence of Corporate Responsibility on the Cost of Capital; Working paper series; University of Hamburg: Hamburg, Germany, 2006.
- Bastounis, A., Hansen, A. C., & Vlačić, V. (2021). The mathematics of adversarial attacks in AI-Why deep learning is unstable despite the existence of stable neural networks. *arXiv preprint arXiv:2109.06098*. DOI: <https://doi.org/10.48550/arXiv.2109.06098>

- Batista, G. E., & Monard, M. C. (2003). An analysis of four missing data treatment methods for supervised learning. *Applied Artificial Intelligence*, 17(5-6), 519-533.
- Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 26(6), 1315-1344.
- Burman, A. F. L. B. (2022). *Integrating Emission Indicators in Investment Decisions: An Evaluation of OLS Regression, KNN and Gradient Boosting Classification Approaches* (Master's thesis, Universidade Catolica Portuguesa (Portugal)).
- Brundtland, G. H. (1987). World Commission on Environment and Development. (1987). *Our common future* [Report]. United Nations.
<https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf>
- Ceran, E. B. (2017). Kurumsal sürdürülebilirlik kavramı ve ölçümüne ilişkin bir ön çalışma. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 46, 59-70.
- Chams, N., García-Blandón, J., & Hassan, K. (2021). Role reversal! financial performance as an antecedent of ESG: The moderating effect of total quality management. *Sustainability*, 13(13), 7026.
- Chang, R., Liu, X., & Deng, W. (2025). Prediction of corporate default risk considering ESG performance and unbalanced samples. *Applied Soft Computing*, 171, 112864.
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247-1250.
- Cheng, B., Ioannou, I., & Serafeim, G. (2014). Corporate social responsibility and access to finance. *Strategic Management Journal*, 35(1), 1-23.
- Chen, L., Yuan, T., Cebula, R. J., Shuangjin, W., & Foley, M. (2021). Fulfillment of ESG responsibilities and firm performance: a zero-sum game or mutually beneficial. *Sustainability*, 13(19), 10954.
- Chaklader, B., Chaklader, G., & Shrivastav, S. K. (2024). Impact of environmental, social and governance (ESG) on foreign institutional investments in India: a machine learning perspective. *Benchmarking: An International Journal*. DOI: <https://doi.org/10.1108/BIJ-12-2023-0910>
- Chowdhury, M. A. F., Abdullah, M., Azad, M. A. K., Sulong, Z., & Islam, M. N. (2023). Environmental, social and governance (ESG) rating prediction using machine learning approaches. *Annals of Operations Research*, 1-25
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peerj Computer Science*, 7, e623.
- Cini, F., & Ferrari, A. (2025). Towards the estimation of ESG ratings: A machine learning approach using balance sheet ratios. *Research in International Business and Finance*, 73, 102653.

- Cots, M. E. D. *A Study Of Classification Modeling of Esg Performance On European Firms and Unveiling Of Influential Factors* (Doctoral dissertation, Tilburg University).
- Crespi, F., & Migliavacca, M. (2020). The determinants of ESG rating in the financial industry: The same old story or a different tale?. *Sustainability*, 12(16), 6398.
- Cek, K., & Eyupoglu, S. (2020). Does environmental, social and governance performance influence economic performance?. *Journal of Business Economics and Management (JBEM)*, 21(4), 1165-1184.
- Çakır, T. (2009). Kurumsal itibar-örgütsel doku ilişkisine kuramsal yaklaşım-I: Paydaş Teorisi. *İstanbul Üniversitesi İletişim Fakültesi Dergisi*, (37), 49-64.
- Dağıstanlı, H. A., & Çelik, İ. (2023). Sürdürülebilirlik raporlaması ve firma performansı: BIST Sürdürülebilirlik Endeksi üzerine bir uygulama. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, (76), 1-16.
- D'Amato, V., D'Ecclesia, R., & Levantesi, S. (2021). Fundamental ratios as predictors of ESG scores: A machine learning approach. *Decisions in Economics and Finance*, 44(2), 1087-1110.
- D'Amato, V., D'Ecclesia, R., & Levantesi, S. (2022). ESG score prediction through random forest algorithm. *Computational Management Science*, 19(2), 347-373.
- David, L. K., Wang, J., Angel, V., & Luo, M. (2024). China's ESG scorecard: A predictive machine learning model. *Corporate Social Responsibility and Environmental Management*, 31(4), 3468-3486.
- Deegan, C., Rankin, M., & Tobin, J. (2002). An examination of the corporate social and environmental disclosures of BHP from 1983-1997: A test of legitimacy theory. *Accounting, Auditing & Accountability Journal*, 15(3), 312-343.
- Del Vitto, A., Marazzina, D., & Stocco, D. (2023). ESG ratings explainability through machine learning techniques. *Annals of Operations Research*, 1-30.
- Demir, S., & Sahin, E. K. (2023). Application of state-of-the-art machine learning algorithms for slope stability prediction by handling outliers of the dataset. *Earth Science Informatics*, 16(3), 2497-2509.
- Dereköy, F. (2015). Kurumsal yönetim ve temel dinamikleri. *Ekonomi ve Yönetim Araştırmaları Dergisi*, 4(1), 31-51.
- 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.
- Dipierro, A. R., Barrionuevo, F. J., & Toma, P. (2025). Predicting ESG controversies in banks using machine learning techniques. *Corporate Social Responsibility and Environmental Management*, 32, 3525-3544. DOI: <https://doi.org/10.1002/csr.3146>
- Drempetic, S., Klein, C., & Zwergel, B. (2020). The influence of firm size on the ESG score: Corporate sustainability ratings under review. *Journal of Business Ethics*, 167(2), 333-360.

- Duque-Grisales, E., & Aguilera-Caracuel, J. (2021). Environmental, social and governance (ESG) scores and financial performance of multilatinas: Moderating effects of geographic international diversification and financial slack. *Journal of Business Ethics*, 168(2), 315-334.
- de Souza Barbosa, A., da Silva, M. C. B. C., da Silva, L. B., Morioka, S. N., & de Souza, V. F. (2023). Integration of Environmental, Social, and Governance (ESG) criteria: their impacts on corporate sustainability performance. *Humanities and Social Sciences Communications*, 10(1), 1-18.
- De Lucia, C., Paziienza, P., & Bartlett, M. (2020). Does good ESG lead to better financial performances by firms? Machine learning and logistic regression models of public enterprises in Europe. *Sustainability*, 12(13), 5317.
- Dwivedi, D., Batra, S., & Pathak, Y. K. (2023). A machine learning based approach to identify key drivers for improving corporate's ESG ratings. *Journal of Law and Sustainable Development*, 11(1). DOI: <https://doi.org/10.37497/sdgs.v11i1.242>
- Edhrabooh, K. M., Alromaihi, M., Binzaiman, F., & AlShammari, M. (2024, October). ESG Score Prediction Using Machine Learning, Gap Analysis. In *2024 5th International Conference on Data Analytics for Business and Industry (ICDABI)* (pp. 123-127).
- El Khoury, R., Nasrallah, N., & Alareeni, B. (2023). The determinants of ESG in the banking sector of MENA region: a trend or necessity?. *Competitiveness Review: An International Business Journal*, 33(1), 7-29.
- Elgergeni, S., Khan, N., & Kakabadse, N. K. (2018). Firm ownership structure impact on corporate social responsibility: evidence from austerity UK. *International Journal of Sustainable Development & World Ecology*, 25(7), 602-618.
- Fauzi, H., & Idris, K. (2009). The relationship of CSR and financial performance: New evidence from Indonesian companies. *Issues in Social and Environmental Accounting*, 3(1). 66–87
- Ferriani, F., & Natoli, F. (2021). ESG risks in times of Covid-19. *Applied Economics Letters*, 28(18), 1537-1541.
- Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Cambridge University Press.
- Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), 210-233.
- Friede, G. (2019). Why don't we see more action? A metasynthesis of the investor impediments to integrate environmental, social, and governance factors. *Business Strategy and the Environment*, 28(6), 1260-1282.
- Garcia-Sanchez, I. M., Cuadrado-Ballesteros, B., & Sepulveda, C. (2014). Does media pressure moderate CSR disclosures by external directors?. *Management Decision*, 52(6), 1014-1045.

- Garcia, A. S., Mendes-Da-Silva, W., & Orsato, R. J. (2017). Sensitive industries produce better ESG performance: Evidence from emerging markets. *Journal of Cleaner Production*, 150, 135-147.
- Greiling, D., & Grüb, B. (2014). Sustainability reporting in Austrian and German local public enterprises. *Journal of Economic Policy Reform*, 17(3), 209-223.
- Gupta, A., Sharma, U., & Gupta, S. K. (2021, December). The role of ESG in sustainable development: An analysis through the lens of machine learning. In *2021 IEEE international humanitarian technology conference (IHTC)* (pp. 1-5).
- Gümrah, A., & Tanç, Ş. G. (2018). Sürdürülebilirlik Raporlarının İçerik Kalitesi: Bist Sürdürülebilirlik Endeksinde Bir Uygulama. *Muhasebe Bilim Dünyası Dergisi*, 20, 334-357.
- Huotari, E. (2024). *Predicting ESG ratings for S&P 500 companies using financial metrics* (Master's thesis). Lappeenranta–Lahti University of Technology LUT.
- Hsu, W. L., Lin, Y. L., Lai, J. P., Liu, Y. H., & Pai, P. F. (2025). Forecasting Corporate Financial Performance Using Deep Learning with Environmental, Social, and Governance Data. *Electronics*, 14(3), 417.
- Ilyas, A., Santurkar, S., Tsipras, D., Engstrom, L., Tran, B., & Madry, A. (2019). Adversarial examples are not bugs, they are features. *Advances in Neural Information Processing Systems*, 32.
- Jumin, E., Zaini, N., Ahmed, A. N., Abdullah, S., Ismail, M., Sherif, M., ... & El-Shafie, A. (2020). Machine learning versus linear regression modelling approach for accurate ozone concentrations prediction. *Engineering Applications of Computational Fluid Mechanics*, 14(1), 713-725.
- Joshi, B., & Joshi, H. (2024). Financial determinants of environmental, social and governance performance: Empirical evidence from India. *Investment Management & Financial Innovations*, 21(1), 13-24.
- Khan, M. H., Alabdeen, Z. Z., & Anupam, A. (2024). Firm-level climate change risk and adoption of ESG practices: a machine learning prediction. *Business Process Management Journal*, 30(6), 1741-1763
- Kotsantonis, S., & Serafeim, G. (2019). Four things no one will tell you about ESG data. *Journal of Applied Corporate Finance*, 31(2), 50-58.
- Köse, E. (2021). Kurumsal yönetim özelliklerinin çevresel, sosyal ve yönetim (ESG) açıklamalarına etkisi. *Uluslararası İşletme, Ekonomi ve Yönetim Perspektifleri Dergisi*, 5, 463–474.
- Küçükaycan, D. (2022). Sürdürülebilirlik muhasebesi ve raporlamanın kamu mali yönetimindeki öneminin teorik açıdan incelenmesi. *Denetim*, (25), 116-130.
- Kulalı, G. (2022). Çevresel, sosyal ve kurumsal yönetim (ESG) performansının piyasa değeri üzerindeki etkisi: Firma büyüklüğünün rolü. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 17(3), 787-809.

- Krappel, T., Bogun, A., & Borth, D. (2021). Heterogeneous ensemble for ESG ratings prediction. DOI: <https://doi.org/10.48550/arXiv.2109.10085>
- Krishnamoorthy, R. (2021). Environmental, social, and governance (ESG) investing: Doing good to do well. *Open Journal of Social Sciences*, 9(7), 189-197.
- Lai, J. P., Chang, Y. M., Chen, C. H., & Pai, P. F. (2020). A survey of machine learning models in renewable energy predictions. *Applied Sciences*, 10(17), 5975
- Lin, H. Y., & Hsu, B. W. (2023). Empirical Study of ESG Score Prediction through Machine Learning—A Case of Non-Financial Companies in Taiwan. *Sustainability*, 15(19), 14106.
- Lin, W. L., Ho, J. A., Ng, S. I., & Lee, C. (2020). Does corporate social responsibility lead to improved firm performance? The hidden role of financial slack. *Social Responsibility Journal*, 16(7), 957-982.
- Lee, O., Joo, H., Choi, H., & Cheon, M. (2022). Proposing an integrated approach to analyzing ESG data via machine learning and deep learning algorithms. *Sustainability*, 14(14), 8745.
- LSEG. (2024). *Refinitiv ESG Scores Methodology* [PDF]. https://www.lseg.com/content/dam/data-analytics/en_us/documents/methodology/lseg-esg-scores-methodology.pdf
- Mahajan, R., Lim, W. M., Sareen, M., Kumar, S., & Panwar, R. (2023). Stakeholder theory. *Journal of Business Research*, 166, 114104.
- Nevado-Peña, D., López-Ruiz, V. R., & Alfaro-Navarro, J. L. (2015). The effects of environmental and social dimensions of sustainability in response to the economic crisis of European cities. *Sustainability*, 7(7), 8255-8269.
- Nijhof, S.(2023) *ESG rating prediction using multiple linear regression, random forest and support vector regression* (Doctoral dissertation, Tilburg University).
- Ng, A. C., & Rezaee, Z. (2015). Business sustainability performance and cost of equity capital. *Journal of Corporate Finance*, 34, 128-149.
- Nyitrai, T., & Virág, M. (2019). The effects of handling outliers on the performance of bankruptcy prediction models. *Socio-Economic Planning Sciences*, 67, 34-42.
- Ortas, E., Álvarez, I., & Garayar, A. (2015). The environmental, social, governance, and financial performance effects on companies that adopt the United Nations Global Compact. *Sustainability*, 7(2), 1932-1956.
- Owen, D. (2008). Chronicles of wasted time? A personal reflection on the current state of, and future prospects for, social and environmental accounting research. *Accounting, Auditing & Accountability Journal*, 21(2), 240-267.
- Özçelik, F., Öztürk, B. A., & Gürsakal, S. (2014). Investigating the relationship between corporate social responsibility and financial performance in Turkey. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 28(3).

- Palynska, M., Medda, F., Caivano, V., Di Stefano, G., & Scalese, F. (2024). The impact of the ESG factor on industrial performance: An analysis using machine learning techniques. *CONSOB Sustainable Finance Series, No. 4*. <https://ssrn.com/abstract=4910370>
- Park, J., Na, H. J., & Kim, H. (2024). Development of a success prediction model for crowdfunding based on machine learning reflecting ESG information. *IEEE Access*, 12, 197275–197289. DOI: [10.1109/ACCESS.2024.3519219](https://doi.org/10.1109/ACCESS.2024.3519219)
- Peng, L. S., & Isa, M. (2020). Environmental, social and governance (ESG) practices and performance in Shariah firms: agency or stakeholder theory?. *Asian Academy of Management Journal of Accounting & Finance*, 16(1).
- Perez-Batres, L. A., Doh, J. P., Miller, V. V., & Pisani, M. J. (2012). Stakeholder pressures as determinants of CSR strategic choice: Why do firms choose symbolic versus substantive self-regulatory codes of conduct?. *Journal of Business Ethics*, 110, 157-172.
- Priya, K. S. (2021). Linear regression algorithm in machine learning through MATLAB. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 9(12), 989-995.
- Przychodzen, J., Gómez-Bezares, F., Przychodzen, W., & Larreina, M. (2016). ESG Issues among fund managers—Factors and motives. *Sustainability*, 8(10), 1078.
- Rajesh, R. (2020). Exploring the sustainability performances of firms using environmental, social, and governance scores. *Journal of Cleaner Production*, 247, 119600.
- Ramezan, C. A., Warner, T. A., Maxwell, A. E., & Price, B. S. (2021). Effects of training set size on supervised machine-learning land-cover classification of large-area high-resolution remotely sensed data. *Remote Sensing*, 13(3), 368.
- Raza, H., Khan, M. A., Mazliham, M. S., Alam, M. M., Aman, N., & Abbas, K. (2022). Applying artificial intelligence techniques for predicting the environment, social, and governance (ESG) pillar score based on balance sheet and income statement data: A case of non-financial companies of USA, UK, and Germany. *Frontiers in Environmental Science*, 10, 975487.
- Ricart, J. E., Rodríguez, M. Á., & Sánchez, P. (2005). Sustainability in the boardroom: An empirical examination of Dow Jones Sustainability World Index leaders. *Corporate Governance: the International Journal of Business in Society*, 5(3), 24-41.
- Scholtens, B. (2006). Finance as a driver of corporate social responsibility. *Journal of Business Ethics*, 68(1), 19-33.
- Si, D. (2023). An efficient predicative approach of ESG invest scoring using random forest algorithm. *BCP Business & Management*, 45, 382–392.
- Silva, F. É. D., & Mascena, K. M. C. D. (2024). The relationship between ESG and binancial performance in brazilian companies. *Revista de Administração da UFSM*, 17(4), e3.
- Soytaş, M. A., Ersoy, İ., Uşar, D. D., & Denizel, M. (2017). Sürdürülebilirlik yatırımlarının finansal performansla etkisi: Türkiye örneği. *Yönetim ve Ekonomi Araştırmaları Dergisi*, 15, 140–162.

- Sultana, S., Zulkifli, N., & Zainal, D. (2018). Environmental, social and governance (ESG) and investment decision in Bangladesh. *Sustainability*, 10(6), 1831.
- Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, 20(3), 571-610.
- Surroca, J., Tribó, J. A., & Waddock, S. (2010). Corporate responsibility and financial performance: The role of intangible resources. *Strategic Management Journal*, 31(5), 463-490.
- Sharma, S., & Chatterjee, S. (2021). Winsorization for robust bayesian neural networks. *Entropy*, 23(11), 1546.
- Srivalli, K. S., & Sumanthi, D. (2025, March 23). Enhancing financial risk assessment through explainable AI: A SHAP-based approach for transparent decision-making. In *Proceedings of the International Conference on Innovative Computing & Communication (ICICC 2024)*. DOI: <https://doi.org/10.2139/ssrn.5190418>
- Şahin, A., & Acar, E. (2023). ESG skorunun firmaya özgü belirleyicileri, Borsa İstanbul örneği. *Muhasebe ve Denetime Bakış*, 23(70), 111-128.
- Şişman, M. E., & Çankaya, S. (2021). Çevresel, sosyal ve kurumsal yönetim (ESG) verilerinin firmaların finansal performansına etkisi: Hava yolu sektörü üzerine bir çalışma. *Çukurova Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 25(1), 73-91.
- Tatachar, A. V. (2021). Comparative assessment of regression models based on model evaluation metrics. *International Research Journal of Engineering and Technology (IRJET)*, 8(09), 2395-0056.
- Tate, W. L., Ellram, L. M., & Kirchhoff, J. F. (2010). Corporate social responsibility reports: a thematic analysis related to supply chain management. *Journal of Supply Chain Management*, 46(1), 19-44.
- Tokgöz, N., & Önce, S. (2009). Şirket sürdürülebilirliği: geleneksel yönetim anlayışına alternatif. *Afyon Kocatepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 11(1), 249-275.
- Tsalis, T. A., Terzaki, M., Koulouriotis, D., Tsagarakis, K. P., & Nikolaou, I. E. (2023). The nexus of United Nations' 2030 Agenda and corporate sustainability reports. *Sustainable Development*, 31(2), 784-796.
- Velte, P. (2017). Does ESG performance have an impact on financial performance? Evidence from Germany. *Journal of Global Responsibility*, 8(2), 169-178.
- YAVUZ, A. E. (2023). Esg Skorlarının Firma karlılığı üzerindeki etkisi: Borsa İstanbul örneği. *Üçüncü Sektör Sosyal Ekonomi Dergisi*, 58(3), 2686-2701.
- Zhang, A. Y., & Zhang, J. H. (2024). Renovation in environmental, social and governance (ESG) research: the application of machine learning. *Asian Review of Accounting*, 32(4), 554-572.

- Zhou, G., & Wang, S. (2025). Enhancing credit risk decision-making in supply chain finance with interpretable machine learning model. *IEEE Access*, 13, 14239–14251. <https://doi.org/10.1109/ACCESS.2025.3530433>
- Waddock, S. A., & Graves, S. B. (1997). The corporate social performance–financial performance link. *Strategic Management Journal*, 18(4), 303-319.
- Wang, Z., Wang, X., Liu, X., Zhang, J., Xu, J., & Ma, J. (2024). A Novel Stacked Generalization Ensemble-Based Hybrid SGM-BRR Model for ESG Score Prediction. *Sustainability*, 16(16), 6979.
- Wang, C. (2024). The Relationship between ESG Performance and Corporate Performance-Based on Stakeholder Theory. In *SHS Web of Conferences* (Vol. 190, p. 03022). EDP Sciences.
- Wood, D. J., & Jones, R. E. (1995). Stakeholder mismatching: A theoretical problem in empirical research on corporate social performance. *The International Journal of Organizational Analysis*, 3(3), 229-267.
- Xiao, C., Wang, Q., van Donk, D. P., & van der Vaart, T. (2018). When are stakeholder pressures effective? An extension of slack resources theory. *International Journal of Production Economics*, 199, 138-149.



APPENDICES

APPENDIX A

Table A.1 Descriptions of Independent Variables Used in the Modeling

Variables	Calculation	Source
ESG Score	<i>Overall Environmental Social and Governance Score</i>	Thomson Reuters Eikon Database
Working Capital	<i>the difference between current assets and current liabilities for the fiscal period</i>	
Cash	<i>cash on hand, balances on demand deposits, and near-term savings and time deposits</i>	
Advertising Expense	<i>Represents the cost of advertising media and promotional expenses. Advertising Expense may include outsourced advertising expenses for marketing.</i>	
Basic EPS Excluding Extraordinary	<i>Represents Income Available to Common Excluding Extraordinary Items divided by Basic Weighted Average Shares.</i>	
Capital Expenditures Cumulative	<i>Represents the sum of: Purchase of Fixed Assets, Purchase/Acquisition of Intangibles and Software Development Costs.</i>	
Cash Dividends Paid Common	<i>Represents cash dividends paid to common stockholders.</i>	
EBIT	<i>EBIT is computed as Total Revenues for the fiscal year minus Total Operating Expenses plus Operating Interest Expense, Unusual Expense/Income [SUIE] and Non-Recurring Items, Supplemental, Total [SUIT] for the same period. This definition excludes non-operating income and expenses.</i>	
EBITDA, %Prd/Prd	<i>This is the percent change in the Earnings Before Interest, Taxes, Depreciation & Amortization as compared to the same period one year ago.</i>	
Intangibles Gross	<i>Consists of patents, copyrights, franchises, goodwill, trademarks, trade names, secret processes, and organization costs. Intangibles, Gross represents the gross amount of intangibles before being reduced by Accumulated Intangible Amortization.</i>	
Intangibles Net	<i>Represents Intangibles, Gross reduced by Accumulated Intangible Amortization. Intangibles, Net is utilized when the company does not provide Intangibles, Gross.</i>	
Interest Coverage Ratio	<i>Measures the number of times within a fiscal period the company generates enough operating income to meet its interest payments.</i>	

Long Term Debt to Total Capital Percent	<i>This is the ratio of Long Term Debt divided by Total Capital at the end of the fiscal period and is expressed as percentage.</i>	Thomson Reuters Eikon Database
Long Term Debt to Total Equity Percent	<i>This is the ratio of Total Debt for the most recent fiscal period divided by Total Shareholder Equity for the same period and is expressed as percentage.</i>	
Net Sales	<i>Represents sales receipts for products and services, less cash discounts, trade discounts, excise tax, and sales returns and allowances.</i>	
Oper Profit, %Prd/Prd, FY	<i>This is the percent change in annual Operating Profit as compared to the same period one year ago.</i>	
ROE Common Equity %	<i>This value is calculated as the Income Available to Common Excluding Extraordinary Items for the fiscal period divided by the same period Average Common Equity and is expressed as a percentage.</i>	
Research And Development	<i>Represents expenses for research and development of new products and services by a company in order to obtain a competitive advantage.</i>	
Research and Development Expenses To Total Revenue, %	<i>This ratio is calculated as the Research and Development Expenses divided by Total Revenue for the same period and is expressed as percentage.</i>	
Selling General Administrative Expense	<i>Represents all of the costs of operating a business other than the costs of readying a product for sale.</i>	
Tot Revenue, % Period/Period, FY	<i>This is the percent change in annual Total Revenue as compared to the same period one year ago.</i>	
Total Assets, % Period/Period	<i>This is the percent change in the Total Assets as compared to the previous fiscal period. It is calculated as Total Assets for the fiscal period minus Total Assets for the previous fiscal period divided by the Total Assets for the previous fiscal period, multiplied by 100.</i>	
ZScore, Manufacturing Weights	<i>Manufacturing Z-Score is calculated using data in its model for the fiscal period. The Z-score is a multivariate formula that measures the financial health of a company and predicts the probability of bankruptcy within two years.</i>	

TRBC Economic Sector Name	<i>Dummy variables representing the economic sector of each firm based on TRBC (Thomson Reuters Business Classification). Each variable takes the value of 1 if the firm belongs to that sector, and 0 otherwise.</i>	Thomson Reuters Eikon Database
Year	<i>Indicates the fiscal year corresponding to the financial indicators reported for each firm. Used to capture time-based effects or trends in the modeling process.</i>	



APPENDIX B

Table B.1 Best Hyperparameter Sets for All Models Using RandomizedSearchCV

Model	Best Hyperparameters
Random Forest	n_estimators=300, min_samples_split=5, min_samples_leaf=4, max_depth=20
Gradient Boosting	subsample=1.0, n_estimators=100, max_depth=3, learning_rate=0.05
XGBoost	subsample=0.7, n_estimators=200, max_depth=3, learning_rate=0.01, colsample_bytree=1.0
Ridge Regression	alpha=215.44
K-Nearest Neighbors	weights='uniform', p=1, n_neighbors=7
LightGBM	subsample=0.8, num_leaves=31, n_estimators=300, max_depth=5, learning_rate=0.01, colsample_bytree=0.8

APPENDIX C

Table C.1 Descriptive statistics of the independent variables in the cleaned dataset (n=...)

	Count	Mean	Min	Max	Std
Advertising Expense (in million)	512.0	247.5	0.0	7,681.1	523.8
Basic EPS Excluding Extraordinary Items	512.0	3.4	-20.7	128.6	10.9
Capital Expenditures Cumulative (in million)	512.0	-1,634.0	-36,609.9	-0.0	3,626.9
Cash (in million)	512.0	1,061.2	0.0	33,235.4	3,030.0
Cash Dividends Paid Common (in million)	512.0	-820.8	-34,247.5	0.0	2,042.3
EBIT (in million)	512.0	2,630.5	-6,380.0	94,652.8	7,811.5
EBITDA Prd Prd	512.0	0.8	-7.5	55.3	3.1
ESG Score	512.0	54.3	2.4	94.4	23.5
Intangibles Gross (in million)	512.0	4,431.1	0.1	214,804.9	17,339.3
Intangibles Net (in million)	512.0	2,641.7	0.0	107,532.8	9,200.6
Interest Coverage Ratio	512.0	45.8	-12.5	5,283.8	348.1
Long Term Debt to Total Capital Percent	512.0	0.2	0.0	1.0	0.2
Long Term Debt to Total Equity Percent	512.0	0.9	0.0	148.6	6.6
Net Sales (in million)	512.0	21,500.9	0.0	617,298.9	36,018.7
Oper Profit Prd Prd FY	512.0	0.3	-171.3	97.2	9.6
ROE Common Equity	512.0	0.2	-3.5	3.4	0.4
Research And Development (in million)	512.0	89.1	0.0	2,793.5	174.5
Research and Development Expenses To Total Revenue (in million)	512.0	0.0	0.0	0,2	0,02
Selling General Administrative Expense (in million)	512.0	1,489.4	-62.3	55,464.9	4,160.9
Tot Revenue Period Period FY	512.0	0.8	-0.6	26.1	1.7
Total Assets Period Period	512.0	0.7	-0.1	6.8	1.0
Working Capital (in million)	512.0	2,642.7	-34,556.8	67,506.2	8,503.4
ZScore Manufacturing Weights	512.0	3.3	-0.4	46.9	3.3

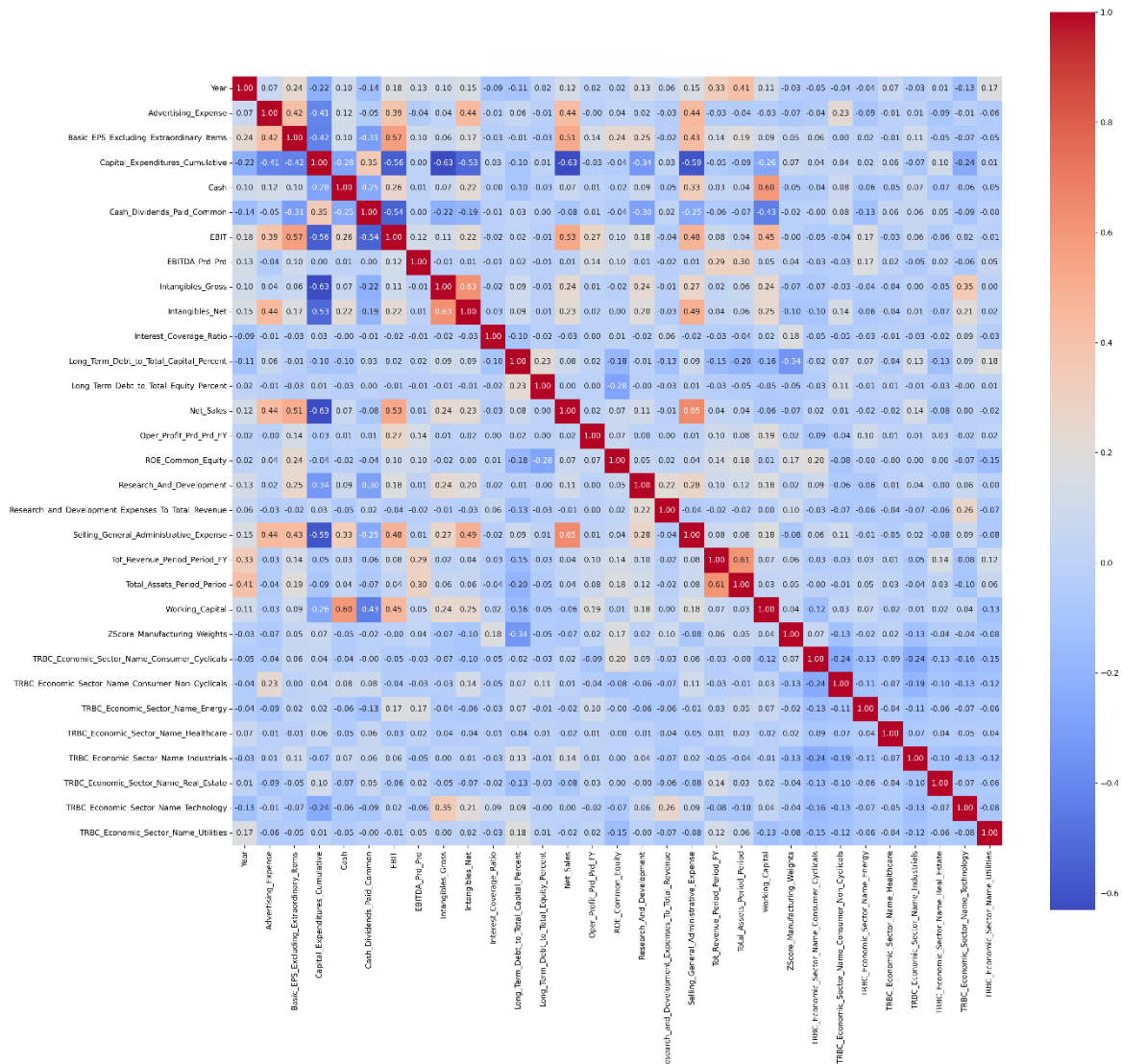


Figure C.1 Pearson Correlation Matrix of the 31 Variables

APPENDIX D

Table D.1 List of the 95 companies analyzed

Identifier (RIC)	Company Name
AEFES.IS	Anadolu Efes Biracilik ve Malt Sanayii AS
AKCNS.IS	Akcansa Cimento Sanayi ve Ticaret AS
AKENR.IS	Akenerji Elektrik Uretim AS
AKFGY.IS	Akfen Gayrimenkul Yatirim Ortakligi AS
AKSA.IS	Aksa Akrilik Kimya Sanayii AS
AKSEN.IS	Aksa Enerji Uretim AS
AKSGY.IS	Akis Gayrimenkul Yatirim Ortakligi AS
ALARK.IS	Alarko Holding AS
ALGYO.IS	Alarko Gayrimenkul Yatirim Ortakligi AS
ALKIM.IS	Alkim Alkali Kimya AS
ANELE.IS	Anel Elektrik Proje Taahhut ve Ticaret AS
ARCLK.IS	Arcelik AS
ASELS.IS	Aselsan Elektronik Sanayi ve Ticaret AS
ASUZU.IS	Anadolu Isuzu Otomotiv Sanayii ve Ticaret AS
AYDEM.IS	Aydem Yenilenebilir Enerji AS
AYGAZ.IS	Aygaz AS
BASGZ.IS	Baskent Dogalgaz Dagitim Gayrimenkul Yatirim Ortakligi AS
BIMAS.IS	BIM Birlesik Magazalar AS
BIOEN.IS	Biotrend Cevre ve Enerji Yatirimlari AS
BIZIM.IS	Bizim Toptan Satis Magazalari AS
BRISA.IS	Brisa Bridgestone Sabanci Lastik Sanayi ve Ticaret AS
BRYAT.IS	Borusan Yatirim ve Pazarlama AS
CANTE.IS	Can2 Termik AS
CCOLA.IS	Coca-Cola Icecek AS
CIMSA.IS	Cimsa Cimento Sanayi ve Ticaret AS
CONSE.IS	Consus Enerji Isletmeciligi ve Hizmetleri AS
CRFSA.IS	CarrefourSA Carrefour Sabanci Ticaret Merkezi AS
DESA.IS	Desa Deri Sanayi ve Ticaret AS
DOAS.IS	Dogus Otomotiv Servis ve Ticaret AS
DYOBY.IS	Dyo Boya Fabrikalari Sanayi Ve Ticaret AS
ECILC.IS	EIS Eczacibasi Ilac Sinai ve Finansal Yatirimlar Sanayi ve Ticaret AS
EKGYO.IS	Emlak Konut Gayrimenkul Yatirim Ortakligi AS
ENJSA.IS	Enerjisa Enerji AS
ENKA.IS	ENKA Insaat ve Sanayi AS
EREGL.IS	Eregli Demir ve Celik Fabrikalari TAS
ESEN.IS	Esenboga Elektrik Uretim AS
FROTO.IS	Ford Otomotiv Sanayi AS
GENIL.IS	Gen Ilac ve Saglik Urunleri Sanayi ve Ticaret AS
GRSEL.IS	Gur Sel Turizm Tasimacilik ve Servis Ticaret AS
GWIND.IS	Galata Wind Enerji Anonim Sirket

HEKTS.IS	Hektas Ticaret TAS
HLGYO.IS	Halk Gayrimenkul Yatirim Ortakligi AS
HTTBT.IS	Hitit Bilgisayar Hizmetleri AS
ISDMR.IS	Iskenderun Demir ve Celik AS
ISGYO.IS	Is Gayrimenkul Yatirim Ortakligi AS
KARSN.IS	Karsan Otomotiv Sanayii ve Ticaret AS
KCAER.IS	Kocaer Celik Sanayi ve Ticaret AS
KERTV.IS	Kerevitas Gida Sanayi ve Ticaret AS
KLKIM.IS	Kalekim Kimyevi Maddeler Sanayi ve Ticaret AS
KLMSN.IS	Klimasan Klima Sanayi ve Ticaret AS
KMPUR.IS	Kimteks Poliuretan Sanayi ve Ticaret AS
KONTR.IS	Kontrolmatik Teknoloji Enerji ve Muhendislik AS
KORDS.IS	Kordsa Teknik Tekstil AS
KOZAA.IS	Koza Anadolu Metal Madencilik Isletmeleri AS
KOZAL.IS	Koza Altin Isletmeleri AS
KRDMD.IS	Kardemir Karabuk Demir Celik Sanayi ve Ticaret AS
LOGO.IS	Logo Yazilim Sanayi ve Ticaret AS
MAGEN.IS	Margun Enerji Uretim Sanayi ve Ticaret AS
MAVI.IS	Mavi Giyim Sanayi ve Ticaret AS
MGROS.IS	Migros Ticaret AS
MPARK.IS	MLP Saglik Hizmetleri AS
NATEN.IS	Naturel Yenilenebilir Enerji Ticaret AS
NETAS.IS	Netas Telekomunikasyon AS
NUHCM.IS	Nuh Cimento Sanayi AS
OTKAR.IS	Otokar Otomotiv ve Savunma Sanayi AS
PETKM.IS	Petkim Petrokimya Holding AS
PETUN.IS	Pinar Entegre Et ve Un Sanayii AS
PGSUS.IS	Pegasus Hava Tasimaciligi AS
PINSU.IS	Pinar Su ve Icecek Sanayi ve Ticaret AS
PNSUT.IS	Pinar Sut Mamulleri Sanayii AS
POLHO.IS	Polisan Holding AS
QUAGR.IS	Qua Granite Hayal Yapi ve Urunleri Sanayi Ticaret AS
SARKY.IS	Sarkuysan Elektrolitik Bakir Sanayi ve Ticaret AS
SASA.IS	SASA Polyester Sanayi AS
SELEC.IS	Selcuk Ecza Deposu Ticaret ve Sanayi AS
SISE.IS	Turkiye Sise ve Cam Fabrikalari AS
SOKM.IS	Sok Marketler Ticaret AS
SUNTK.IS	Sun Tekstil Sanayi Ve Ticaret AS
TATGD.IS	Tat Gida Sanayi AS
TAVHL.IS	TAV Havalimanlari Holding AS
TCELL.IS	Turkcell Iletisim Hizmetleri AS
THYAO.IS	Turk Hava Yollari AO
TKFEN.IS	Tekfen Holding AS
TKNSA.IS	Teknosa Ic ve Dis Ticaret AS
TOASO.IS	Tofas Turk Otomobil Fabrikasi AS
TTKOM.IS	Turk Telekomunikasyon AS
TTRAK.IS	Turk Traktor ve Ziraat Makineleri AS
TUPRS.IS	Turkiye Petrol Rafinerileri AS

ULKER.IS	Ulker Biskuvi Sanayi AS
VERUS.IS	Verusa Holding AS
VESBE.IS	Vestel Beyaz Esva Sanayi ve Ticaret AS
VESTL.IS	Vestel Elektronik Sanayi ve Ticaret AS
VKGYO.IS	Vakif Gayrimenkul Yatirim Ortakligi AS
YUNSA.IS	Yunsa Yunlu Sanayi ve Ticaret AS
ZOREN.IS	Zorlu Enerji Elektrik Uretim AS



ÖZGEÇMİŞ

İlk,orta ve lise öğrenimini Çanakkale’de tamamladı. 2018 yılında girdiği Yalova Üniversitesi İşletme (İngilizce) bölümünden 2023 yılında mezun oldu. 2025 yılında Yalova Üniversitesi İşletme Anabilim Dalında, Yüksek Lisans’ını tamamlamıştır.

TEZDEN TÜRETİLEN YAYIN VE ESERLER

Turğay Ç. & Ateş S. (2025). *ESG PERFORMANCE PREDICTION BY MACHINE LEARNING*, Uluslararası Katılımlı Ekonomi Araştırmaları ve Finansal Piyasalar Kongresi (IERFM), Bursa

CURRICULUM VITAE

She completed her primary, secondary, and high school education in Çanakkale. She graduated from the Department of Business Administration (English) at Yalova University in 2023, which she entered in 2018. She completed her Master’s degree in the Department of Business Administration at Yalova University in 2025.

PUBLICATIONS AND WORKS DERIVED FROM THE THESIS

Turğay Ç. & Ateş S. (2025). *ESG PERFORMANCE PREDICTION BY MACHINE LEARNING*, International Congress on Economic Research and Financial Markets with International Participation (IERFM), Bursa.