International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s,2025
https://theaspd.com/index.php

Assessing the Impact of ESG Factors on Financial Performance Using an AI-Enabled Predictive Model

Md Nazmuddin Moin Khan[1], Nudrat Fariha[2], Md Iqbal Hossain[3], Sajib Debnath[4], Md Abdullah Al Helal[5], Uzzal Basu[6], Md Kauser Miah[7], Nisha Gurung[8]
[1]Analytics and System, University of Bridgeport, Bridgeport, CT, USA
[2]Business analytics, University of Bridgeport, Bridgeport, CT, USA
[3]Business analytics, University of Bridgeport, Bridgeport, CT, USA
[4]Computer Science, Western Illinois University
[5]Master of Science in Business Analytics, Trine University
[6]Engineering Technology, Western Carolina University
[7]MCIS Data Science, Gannon University, Erie, PA
[8]MBA Business Analytics, Gannon University, Erie, PA, USA

Abstract
This study looks at how ESG factors, environmental, social, and governance, relate to the financial performance of U.S. companies, using a machine learning approach built in three parts. We started by looking at each ESG pillar on its own. Using several regression models, we measured how well each one could predict basic financial outcomes. That gave us a sense of their individual signals. Then we shifted gears. In the second phase, we combined all the ESG features into a single model. The goal was to see how a unified view compares to looking at each piece separately. The final phase took a different angle: we used unsupervised clustering to group firms into ESG profiles based on how they behave across all three pillars. These profiles were then used as inputs in our predictive models. What stood out is that governance metrics tended to offer the most consistent signal on their own. But the models improved noticeably when we included all ESG factors together, and even more so when we added the ESG profiles from the clustering step. Those profiles captured deeper, often hidden, patterns that individual metrics couldn't. Here's the thing: looking at ESG factors one by one can give you some insight, but it doesn't really tell the whole story. These elements, environmental, social, and governance, don't operate on their own. They bump into each other, influence outcomes in ways that aren't always obvious at first glance. What we found is that when you step back and treat ESG as a connected system, the relationship to financial performance becomes a lot clearer. It's less about ticking boxes and more about understanding how a company functions. The three-part approach we used isn't some abstract model for the sake of theory. It's a way to get practical answers. For investors, it can help show where a company stands today and how ready it might be for what's coming. For companies, it's a way to see how their decisions in one area might ripple across others. Especially now, as ESG standards keep evolving, having that kind of perspective matters.
Keywords: *ESG Profile, Clustering, Predictive Modeling, Financial Return, Corporate Sustainability*

INTRODUCTION

Background and Motivation

Over the last decade or so, ESG(Environmental, Social, and Governance) considerations have moved from the sidelines into the core of investment strategy. It's no longer just about financial statements, as investors are paying closer attention to the less tangible factors that can shape a company's long-term value. This shift is not just random. Gompers et al. (2003) were among the first to draw a clear link between sound governance practices and firm performance, and Bebchuk et al. (2009) built on that by showing how factors like independent boards and well-structured executive pay can lower agency costs and increase returns for shareholders [12], [2]. On the environmental front, the pressure to respond to climate risks has consequently pushed many firms to disclose data on emissions and renewable energy targets. Khan, Serafeim, and Yoon (2016) showed that those disclosures are not just good PR for they can actually flag future regulatory or legal trouble before it hits [28]. Still, for all the talk about ESG transparency, the reality is messier. Major ESG rating agencies like MSCI, Sustainalytics, and Refinitiv each have their own rules and formulas for how they score companies. Their systems aren't always clear, and more often than not, they tend to produce conflicting ratings for the same firm, prompting investors to consider additional metrics for decision making. That sort of inconsistency throws a wrench into investment decisions and weakens trust in the whole framework [24].

Another vital matter that doesn't get enough attention is that ESG isn't one-size-fits-all. What matters in one industry might barely register in another with a similar or different structure. In heavy industries like energy or manufacturing, environmental indicators tend to dominate, but governance, according to Klapper and Love (2004), tends to correlate strongly with profitability and financing costs across the board [22]. This is where machine learning interestingly comes in, not as a silver bullet, but as a tool to help make sense of the ESG noise. AI models, from ensemble approaches to deep neural networks, have already demonstrated they can handle tangled, nonlinear relationships in financial data. Jakir et al. (2023), for instance, used gradient boosting methods to detect fraud with impressive results, while Hasan et al. (2024) applied classification models to predict customer churn in e-commerce, and both studies show how adaptable these methods can be across business domains [20, 14]. But when it comes to ESG, the use of AI is still pretty new. Part of the problem is data, as proprietary ESG scores are often locked behind paywalls, and the information that is public tends to be patchy or irregular, which makes it hard to train reliable models. This study is trying to work through that by creating, collecting, and combining a comprehensive ESG dataset that's aligned with real-world benchmarks, like the CDP's reports on emissions and water use, to sidestep the usual data constraints. The aim here isn't to replace human judgment but to give it sharper tools. ESG variables also tend to evolve over time, and their impact isn't always immediate, and a company's green investments might not pay off for years. Social policies can also ripple outward through things like employee satisfaction or public perception. These delayed, often subtle effects don't show up clearly in standard linear models, but machine learning has the potential to surface patterns and interactions that would otherwise be conventionally missed.

Importance Of This Research

Getting a reliable read on how ESG scores shape financial performance has become more than an academic exercise, as institutional investors, who collectively steer trillions, are no longer treating ESG considerations as side notes. Regulatory frameworks like the European Union's Sustainable Finance Disclosure Regulation (SFDR) have made these metrics part of the investment baseline, while growing public pressure ensures they remain front and center. For companies, this creates both a challenge and an opportunity, which is, if you can anticipate where your ESG rating is headed before the official scorecard lands, you can make deliberate, informed choices, shifting capital toward sustainability efforts with actual financial upside, tweaking disclosures, and, when necessary, getting ahead of any signals that might unsettle shareholders. From the boardroom's perspective, predictive ESG models don't just clarify where you stand; they offer a way to act with precision. Take corporate governance as an example. We have for a long time known that things like stronger audits or tighter risk oversight can bring down equity risk premiums (Edmans 2011 [9]), but firms still grapple with identifying which reforms yield the best return. Instead of guesswork, a robust model gives decision-makers a comparative lens, but what's actually moving the needle? And in industries confronting climate pressure, these tools become even more vital. If you're weighing the ROI(Return On Investment) of solar arrays versus carbon capture, intuition might not cut it, meaning only the data can speak, that is if you're listening.

On the research front, this work steps into two longstanding debates. First, it tackles the often-raised question: which of the three ESG pillars, Environmental, Social, or Governance, matters most for predicting financial outcomes? Past work, including studies by Klapper & Love (2004) and Bebchuk et al. (2009) [22, 2], leans heavily toward governance, but the evidence has mostly come from case-specific or regional studies. Here, we try something broader. By normalizing inputs across sectors, we get a more generalizable view of how each component performs. Second, rather than isolating metrics, we explore whether a firm's overall ESG "signature" tells a deeper story. Using unsupervised clustering, we test whether the shape of a firm's ESG profile, that is its strengths and blind spots taken together, adds more explanatory power than the sum of its parts. There's precedent for this idea, where in energy forecasting, Shovon et al. (2025) showed that models blending convolutional and recurrent layers capture richer dynamics than any one method alone [29]. ESG, we suspect, may follow a similar pattern as there's signal in the structure, not just the individual points.

1.3 Research Objectives and Contributions

This study is driven by two main questions. First, when you look at Environmental, Social, and Governance (ESG) pillars separately, which one says the most about a company's financial performance? Second, can we learn more by looking at a company's overall ESG profile, not as a set of disconnected metrics, but as a cohesive pattern of behavior? To answer this, we build machine learning models that not only measure how much each ESG pillar contributes on its own, but also test whether integrated ESG profiles, formed through clustering, give us a stronger signal. The approach unfolds in three main parts. First, we collect and build a synthetic dataset designed to mirror how ESG data typically shows up in real-world company reports, especially where high-quality public data is lacking. This lets us run controlled experiments that would be hard to carry out otherwise. Second, we train a range of regression models, that is from simple linear ones to more complex neural networks, directed to comparing how well individual ESG pillars predict financial outcomes versus how well they do when taken together. Third, we use unsupervised learning to form clusters of ESG behavior, treating these clusters as new features in our predictive models. This lets us move beyond surface-level metrics and look at deeper patterns in how companies behave. Rather than treating ESG factors as isolated checkboxes, the goal here is to understand the interplay between them, in order to capture the shape of a company's ethical and operational stance, not just its scores. In doing so, this work offers a more layered view of ESG's role in financial performance, one that may prove useful to both researchers and practitioners looking to make sense of the increasingly noisy ESG landscape.

LITERATURE REVIEW

Traditional ESG Impact Analysis in Finance

The connection between ESG performance and financial outcomes has been picked over by financial scholars for years. Gompers et al. (2003) were among the early voices in the conversation, examining shareholder rights and firm valuation [12.] They found that firms with stronger shareholder protections tended to be valued more highly, a finding that didn't go unnoticed. Their work sparked a wave of studies that zeroed in on governance, and over time, it became clear that governance was the ESG pillar with the most consistent and measurable links to financial performance. Bebchuk et al. (2009) built on this as well by showing how features like board independence and executive pay alignment weren't just symbolic, as they shaped stock performance and long-term firm value, largely by reducing agency costs [2.]. Klapper and Love (2004) added a broader global dimension to the picture, showing that good governance tends contextually go hand in hand with stronger operational metrics and better market valuation, particularly in emerging markets where governance quality varies more widely [22.].

But once we move beyond governance, the story gets murkier as environmental and social factors don't show the same consistent impact across the board. Khan et al. (2016) tried to make sense of this by introducing the idea of "materiality." In short, they posited that ESG factors matter, but only if they're relevant to the business [21.]. A firm in a high-emission industry, for example, will see investors pay attention to its environmental disclosures and emissions control efforts. That same metric, however, means much less in a software company. Their study found that when ESG efforts align with what actually matters in a given sector, the payoff is visible, in terms of better returns and better accounting performance. Still, despite the growing interest in ESG as a strategic tool, there's a real problem with how it's measured. The big rating agencies use black-box methods that vary wildly, encompassing different data, different priorities, and different definitions. Chatterji et al. (2016) pointed out how a company might earn a glowing ESG score from one agency and a mediocre one from another, depending on which levers the agency chooses to pull [5.]. This lack of consistency makes it tough to build reliable models or draw firm conclusions from cross-company comparisons, and comparatively, you're left trying to measure relationships using a ruler that keeps changing length.

Then there's the causality puzzle. Even if you see that firms with higher ESG scores perform better financially, you're stuck with the chicken-or-egg problem. Are they succeeding because of good ESG practices, or can they afford better ESG practices because they're already successful? Researchers have tried to tease this apart using lagged models or matched samples, but without longitudinal data or natural experiments, these methods are more educated guesswork than definitive proof. A more promising turn in recent literature has been the move toward working with direct, observable ESG inputs, what firms actually do, rather than composite scores that obscure the raw details. Sharfman and Fernando (2008), for example, focused on tangible environmental risk management practices and found that firms managing those risks well tend to enjoy lower capital costs [28.]. Similarly, Eccles et al. (2014) tracked companies that voluntarily adopted sustainability policies and compared them to similar firms that didn't [8.]. Over nearly two decades, the sustainability-minded firms outperformed the rest. What this body of research leaves us with is a sense of both promise and friction. ESG does seem to offer a real lens for evaluating corporate strategy, especially when it comes to governance, but measuring environmental and social impact meaningfully requires nuance, context, and far better data. This study attempts to sidestep some of those pitfalls by starting not with pre-packaged ESG ratings, but with a modeling approach built from the ground up. Instead of learning on scores that often come with baked-in biases and unclear assumptions, it works directly with measurable inputs to test how predictive these inputs are of financial outcomes. It is a way of honoring the empirical tradition of past studies while eschewing the haze of third-party scoring systems that too often obscure more than they reveal.

2.2 Behavior-Driven Clustering Approaches

Over the past few years, there's been a noticeable shift in how researchers are thinking about ESG analysis because more and more, focus is turning to machine learning architecture and methods that can be made use of to group companies based on patterns in their behavior, instead of relying solely on predefined categories or linear models, which might at times prove inefficient. Traditional regression architectures have their place in ESG analysis, but they display a tendency to fall short when dealing with the messy, nonlinear nature of ESG data, while considering things like overlapping influences or subtle interdependencies among environmental, social, and governance variables. To navigate that complexity, many are leaning techniques like clustering, neural nets, and ensemble models. Take Jakir et al. (2023), for example; they built a fraud detection system using gradient boosting that picked up on intricate, non-obvious transaction patterns with remarkable accuracy. Their results made it clear that nonlinear models can tease out signals in dense financial data far more effectively than rule-based systems [20]. Similar breakthroughs are showing up elsewhere, too, where Hasan et al. (2024) found that customer behavior–based models in e-commerce churn prediction consistently beat out the standard classifiers by a wide margin [14].

So, what does this mean for ESG modeling? It invites us to think about ESG not as a checklist of factors, but as a complex behavioral fingerprint. Clustering methods like KMeans and HDBSCAN can reveal composite ESG profiles, with groupings like "strong environmental performance but weak governance" or "balanced across all areas." These aren't just academic curiosities as they can also feed directly into predictive models, enriching them with more context and nuance. It's not unlike what Islam et al. (2025) explored in their market simulation work, where synthetic clusters of consumer behavior helped fine-tune recommendation systems in e-commerce [19]. Oddly enough, even though this kind of approach seems tailor-made for ESG, it's still not widely adopted in the field, but its potential is hard to ignore. Shovon et al. (2025) showed how hybrid CNN-LSTM models could handle both immediate and delayed patterns in electricity demand forecasting. Their point? When data is sequential and interdependent, like ESG behavior often is, you need models that can account for that layered structure [29]. ESG events don't happen in a vacuum as a governance shake-up might follow an environmental scandal. An abrupt increase in social spending could be a response to reputational damage, and it is important to comprehend that these aren't isolated datapoints, they're signals in a narrative, and if your model isn't wired to catch that interplay, it's bound to miss something important.

There's also a philosophical advantage to this cluster-based thinking, which is that it sidesteps the rigidity baked into most ESG scores. Instead of hardcoding the idea that environmental issues always carry, say, a 30% weight, clustering lets real-world patterns speak for themselves. What matters in one sector or country might be totally different in another. For example, in places with tough environmental laws but loose oversight, governance might not be a strong predictor, but flip that in a shareholder-driven market, and governance could be the main signal. That's the spirit behind this study. Rather than treating ESG data as a static table of numbers, we're using clustering as a way to capture how companies behave, which is how to find signals that traditional metrics miss. It's an approach that's in line with work like Ray et al. (2025), who used unsupervised learning to map patterns of economic resilience across U.S. cities, drawing from mobility and financial data [27]. In the same way, this study uses unsupervised methods to surface ESG behavior profiles that often precede major financial shifts, either for better or worse.

GAPS AND CHALLENGES

Despite years of research into the relationship between ESG factors and financial performance, we're still missing some crucial pieces. A major stumbling block has been the heavy dependence on third-party ESG ratings. On the surface, these scores seem helpful, but when you look closer, they're often all over the place as different providers use different definitions, assign different weights, and fill in missing data however they see fit. Studies like those by Khan and Chatterji have pointed this out clearly [21, 5]. The result? When ESG ratings don't line up, it leaves investors and researchers without much to stand on. That kind of inconsistency makes it harder to figure out which ESG factors really have an impact on a company's financial future, and to add it gets worse when the ratings are based on black-box methods, proprietary systems that don't concisely show their work. That lack of transparency means researchers can't replicate results or build models that help us understand what's actually driving the scores behind the good.

Another issue is the outdated modelling approaches, where much of the existing work still leans on linear and logistic regression, methods that are straightforward and interpretable, sure, but also quite limited. They assume clean, simple relationships between variables that rarely exist in practice. Think about it: does a company's return really rise or fall in a neat line with every small change in emissions? Or is the impact of governance completely disconnected from how the company manages risk? That kind of logic flattens complex realities. More nuanced techniques, like the hybrid CNN-LSTM models used by Shovon and colleagues in forecasting power system behaviour [29], have proven they can handle messy, layered interactions, yet ESG research has been slow to adopt them. Then there's the problem of how ESG components are handled. Many studies either treat environmental, social, and governance metrics in isolation or mash them together into a single score, but firms, if not most, all, don't operate like that. Their ESG profiles are often uneven; a company might be strong on environmental practices but lag in governance, or vice versa. Treating these dimensions separately misses the full picture, and behavioural clustering offers a way out of this blind spot. Ray et al. (2025) showed how clustering based on shared behavioural traits helped make sense of urban economic resilience [27]. That same logic could help in ESG by grouping companies not just by scores but by strategic or behavioural archetypes.

Timing is one of those issues in ESG research that keeps getting overlooked, even though it really shouldn't be. Different ESG factors unfold on different timelines. Take environmental investments, they can take years before you see any meaningful return. On the other hand, governance reforms might shake things up in the short term before eventually settling into something more stable. But many studies skip over that complexity. Instead of looking at how things evolve, they rely on static snapshots, which can completely throw off the interpretation. This gets even more problematic when you think about climate risk. Long-term changes might look slow and steady on paper, but they can suddenly tip into crises that hit hard and fast. And yet, most ESG datasets don't include enough detailed, time-stamped data to support any kind of serious temporal analysis. Without that, we're flying half-blind when it comes to figuring out cause and effect or making predictions that hold up. Then there's the whole area of synthetic data, which ESG research barely touches. Other fields, fraud detection, e-commerce, cybersecurity, routinely use synthetic datasets to train and test models when real-world data is spotty or limited. It's not perfect, but it's incredibly useful. ESG, though, hasn't really taken that step. That's a missed chance, especially when you're dealing with small companies or regions where disclosures are thin or all over the place. Building solid proxy datasets could open the door to much better experimentation and smarter modelling. It's not a silver bullet, but it's a practical way forward.

So, where does that leave us? The field is being held back by a few persistent habits: an overreliance on obscure third-party scores, simplistic modelling assumptions, a fragmented view of ESG behaviour, poor handling of time, and an underuse of synthetic data. These limitations aren't just academic; they have real consequences for how we understand corporate responsibility and financial risk. There's a clear opportunity to move toward richer, AI-driven methods that are better suited to the complexity of the problem. What's needed now is less reliance on what's easy and more willingness to rethink how we approach ESG from the ground up.

METHODOLOGY

Data Sources and Description

For this study, we worked with a synthetically collected and generated and combined dataset designed to mirror the kind of ESG indicators and financial metrics you'd typically see reported by publicly listed companies. Given the lack of open-source ESG datasets that offer both methodological clarity and consistency, we opted to compile our own, giving us complete control over how features were structured, how much variability they showed, and how they interacted with one another, all while staying grounded in established ESG materiality frameworks, including the MSCI ESG Ratings Methodology. The dataset constitutes 500 firms, with each record blending a range of environmental, social, and governance features with three core financial outcomes: return on equity (ROE), profit margin, and financial return. On the environmental front, we took the initiative to track things like carbon emissions, percentage of energy from renewables, waste management effectiveness, water use intensity, and biodiversity impact. Social indicators in our dataset covered employee satisfaction, workplace diversity, supply chain ethics, philanthropic spending, and staff turnove, whereas the Governance variables included the executive pay ratio, board independence, proxy voting alignment, transparency of audits, and the robustness of internal risk and compliance structures.

To keep the data grounded in plausible realities, we set value ranges and added noise based on real-world ESG disclosures, drawing inspiration from sources like CDP filings and sustainability reports aligned with SASB standards. Importantly, we avoided constructing overly neat or linear relationships. Each feature was sampled independently, within boundaries, but we wove in partial correlations across some dimensions to mimic real-life dependencies, for instance, a company with high carbon emissions typically shows lower reliance on renewable energy. We also created an internal esg_score by combining the environmental, social, and governance sub-scores using equal weighting. This wasn't used as a modeling target, since doing so would risk introducing circular logic; instead, it served as a useful reference point when checking correlation patterns or making visualizations. Lastly, we applied unsupervised clustering across the full ESG feature space to surface underlying company archetypes, which are latent ESG profiles that we later used as higher-order inputs during modeling.

DATA PREPROCESSING

The dataset underwent a series of preparation steps to make it ready for machine learning, with care taken to maintain the integrity of the information throughout. Missing values were examined at the outset by scanning for nulls column by column. No values were imputed or discarded at this stage, since the data was collected, generated, and concatenated with utmost care and precision. To guard against multicollinearity, an issue that can quietly destabilize models, the Variance Inflation Factor (VIF) scores were calculated for all ESG-related variables. All VIF values appeared to be very close to 1, with the highest being around 1.026, indicating very low to no multicollinearity among the dataset's ESG features. This is good news for the models, as it means the features are relatively independent of each other and there is no need to worry about multicollinearity, which would negatively impact the results. Standardization followed, with all numeric ESG attributes being normalized using z-scores, placing them on a comparable scale. This was done to ensure that no single feature overpowered the rest, particularly in models where distance or gradient calculations are sensitive to magnitude. Binary variables in our dataset, such as data_privacy_compliance, audit_transparency, and excluded_sector, were retained in integer-encoded form, objectively to preserve their categorical nature without distortion. The dataset was then split into training and testing sets using an 80:20 ratio. To prevent skewed representations, stratified sampling was applied based on key financial performance outcomes. This helped maintain similar distributions of the target variable across both subsets, reducing the risk of biased evaluation. The standardized matrix was saved to ensure that the same version of the data would be used throughout the modeling process. By the end of preprocessing, the dataset had been sharpened as collinearity was addressed, values were standardized, and the structure was set for rigorous model development.

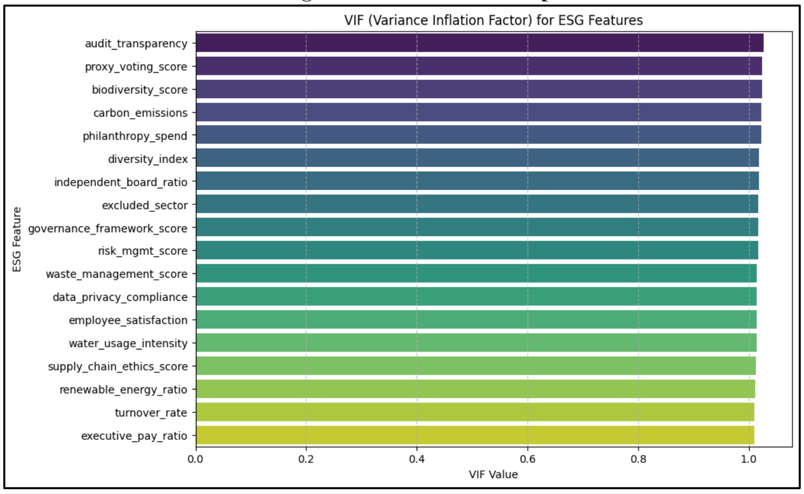

Fig.1. VIF plot for multicollinearity

Exploratory Data Analysis

In order to better understand the structure and relationships within the dataset, exploratory data analysis was conducted. In the heatmap analysis of the correlation between various ESG (Environmental, Social, and Governance) features and key financial performance metrics: financial return, return on equity (ROE), and profit margin, carbon emissions stand out with a consistently negative correlation across all three financial indicators. This relationship is most evident with ROE, where the correlation dips to −0.36. In those industries heavily reliant on carbon, energy, manufacturing, and transport, such a pattern comes as little surprise, because conventionally, emissions often carry financial baggage, with regulatory costs, inefficiencies, and reputational drag adding to the weight. On the other hand, factors like renewable energy usage and employee satisfaction show a faint but positive relationship, particularly with ROE and overall financial return. These aren't monumental shifts, but neither can they be categorized as noise. The data mirrors what's been seen in real markets, where companies leaning into clean energy or cultivating strong internal cultures tend to be rewarded, subtly but steadily. In sectors like consumer goods or tech, where brand trust and talent retention are inseparable from margin growth, such correlations aren't merely academic.

Among governance measures, board independence stands out with a positive correlation with ROE (0.21), reflecting what corporate governance literature has long noted: boards with independent voices tend to exercise better oversight, restrain managerial excesses, and guard shareholder value more effectively. The signal stands out a lot, especially in regions where investor rights are well-protected, like the United States or northern Europe, where there is a consistent balance of power between executives and the board is more carefully maintained throughout. A wide swath of ESG variables hovers near zero correlation, hinting at either minimal financial impact or a complexity that simple linear models fail to capture. Take supply chain ethics, for instance; it may barely register in aggregate, yet in sectors like apparel or extractives, where labor practices or local community friction can lead to reputational fallout or operational delays, the stakes are much higher. ESG's financial relevance, in other words, refuses to be generalized. What is observed from the heatmap isn't a clean narrative, but that's arguably the point. Whereas some ESG factors matter, some don't, and many operate in shades of grey, meaning they are dependent on certain contexts, which are specific to certain sectors, and are often non-linear. The temptation to reduce ESG to a single score or metric risks flattening a much more textured landscape. That is why the data used for this study supports a more tailored approach, especially one that considers not just what a company does, but where, how, and under what conditions.

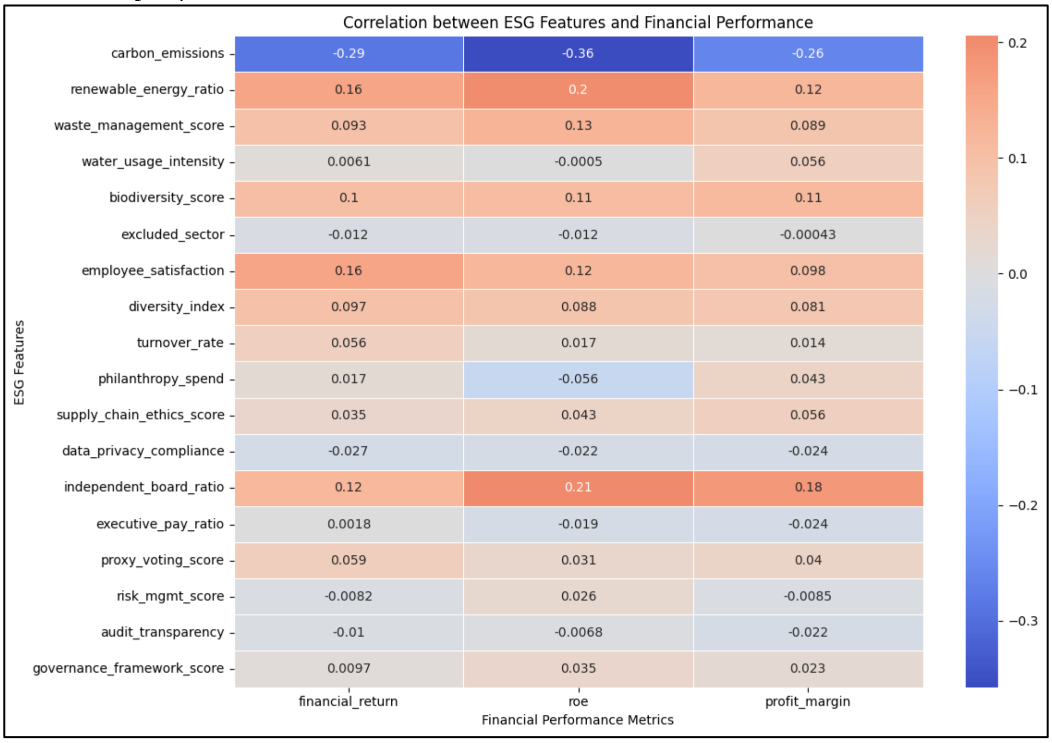

Fig.2. Correlation between ESG features and financial performance

When we broke down the ESG dataset, some interesting patterns started to emerge across the environmental, social, and governance categories. Each category displayed its range of characteristics, and companies do not behave uniformly across them, meaning the result might vary per company, with some differences being subtle, and others more outstanding. On the environmental side, things were all over the place. Carbon emissions stood out right away. Most companies fall into a moderate range, but there's a small group, probably from sectors like energy, manufacturing, or transport, that emits far more than the rest. The distribution here is skewed hard to the right. On the other hand, renewable energy usage tells the opposite story. A lot of firms are still lagging, barely using renewables at all. The data skews left, and it's a reminder that despite all the public talk, the actual shift to clean energy is moving slowly in many areas, likely due to infrastructure gaps or cost barriers. The social metrics were a bit more nuanced. Employee satisfaction looked more balanced, roughly following a normal distribution. That could be a sign of shared expectations around workplace quality, or maybe it's the result of external pressure to meet certain HR standards. But turnover told a very different story. Here, we saw a clear split, some companies keep their employees for the long haul, others see a lot of people coming and going. That bimodal shape probably reflects differences in job type more than anything else, think office jobs versus labor-intensive roles.

Diversity scores stayed pretty tightly packed, which was surprising. It might mean there's not much variation in how companies approach inclusion, or perhaps the measurement itself isn't catching the full picture. Philanthropic spending, meanwhile, was heavily skewed. Most firms gave very little, while a few gave a lot, which pulled up the average and masked how low the baseline is. Governance looked more consistent across the board. Things like board independence, executive compensation ratios, and risk management didn't show much spread. Risk management especially seemed tightly clustered, probably because of regulation or shared frameworks that companies are expected to follow. Executive pay, though, still showed that familiar long right tail. A handful of companies, likely in tech or finance, had huge gaps between C-suite pay and what regular employees make. Financial performance followed patterns you'd expect. Profit margins and returns were skewed right, which is a common shape in competitive industries. A few firms do well, while the majority are just okay, or worse. It reflects a "winner-takes-most" reality, where the biggest gains go to those with scale, capital, or some edge in innovation. If you look at everything together, environmental metrics were the most scattered. That makes sense, they're heavily influenced by what a company does and the resources it relies on. Governance was more stable, probably due to rules and shared norms. Financial outcomes reminded us that the playing field isn't level, and never really has been. These patterns matter, especially when it comes to building predictive models. For instance, skewed data and split behaviors mean your modeling approach needs to be flexible, and a one-size-fits-all method will smooth over the very differences you need to capture.

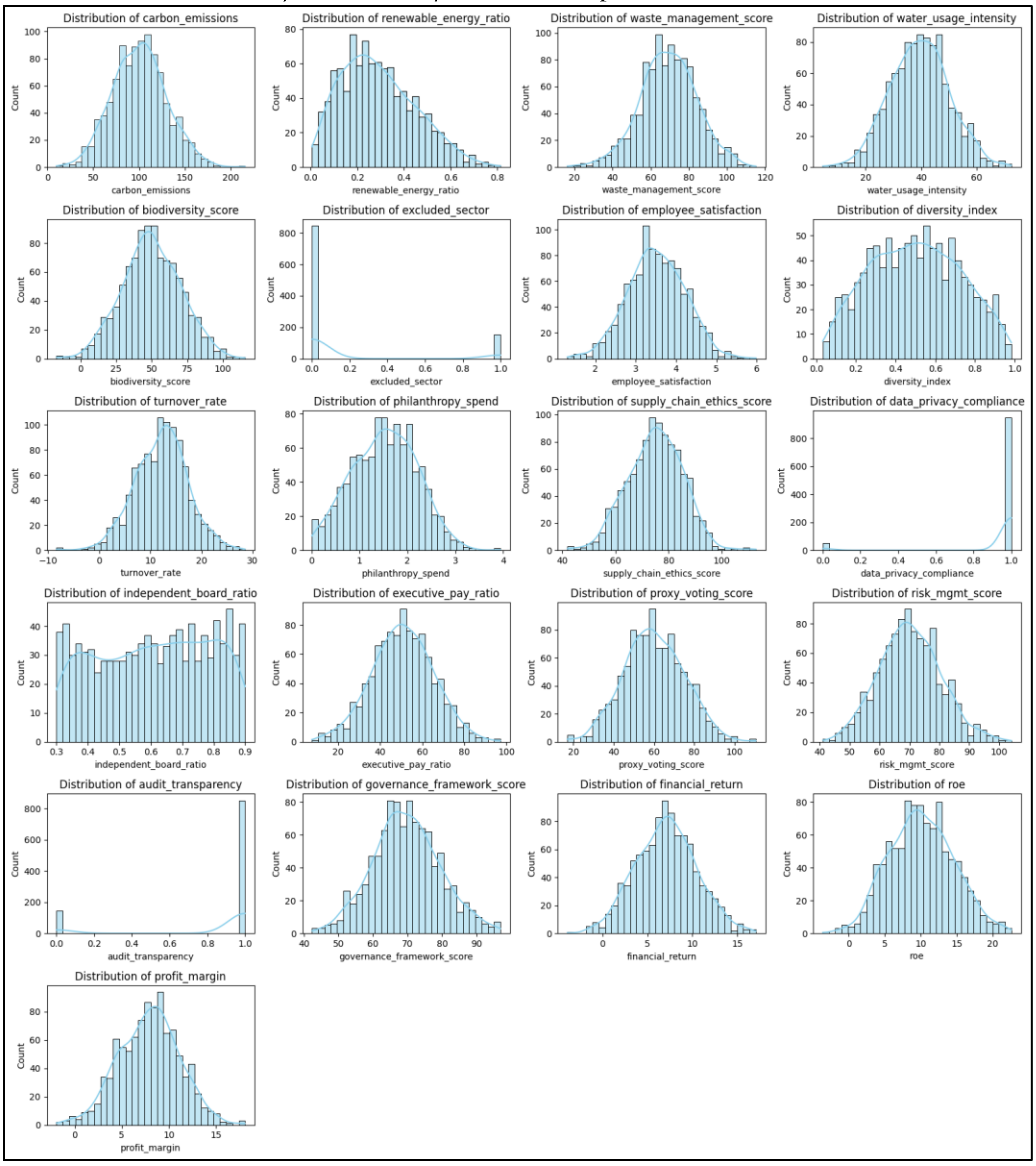

Fig.3. ESG dataset distribution analysis

When looking closely at how financial performance stacks up against certain ESG compliance indicators, specifically data privacy compliance, audit transparency, and industry exclusion, a set of boxplots was used to map out the relationships. These charts help make something pretty clear: when a company takes ESG seriously, especially data privacy, it tends to show up in the numbers. Firms that followed privacy regulations generally performed better financially. Their returns and ROE weren't just higher on average; they were more consistent too, with less variation between the best and worst performers. That kind of stability matters. The companies that didn't comply told a different story. Their results were all over the place, and more of them landed at the bottom end of the scale. That could be the result of fines, damaged reputation, or the inefficiencies that come with falling behind on privacy protections. Profit margins showed a similar pattern. The takeaway here isn't complicated: handling data privacy well may signal something deeper about how a firm runs its operations, and the market seems to pick up on that. Audit transparency stood out even more, as companies rated as transparent in their auditing practices outperformed their less transparent peers across all three financial measures. The differences weren't subtle either, higher medians, tighter spreads, shorter whiskers, each point toward a link between internal clarity and external confidence. These patterns echo long-standing theories in governance research, which outline that strong internal controls and open disclosure often pave the way for steadier financial outcomes. In markets, firms with transparent books are less likely to rattle investors and more likely to secure favorable capital terms, and the data here aligns with that logic.
The industry exclusion variable, which flags businesses operating in areas like fossil fuels, tobacco, or weapons, also showed marked effects. Companies outside these sectors generally saw better median returns and stronger profit margins. The companies in the excluded sectors showed more unpredictable results, often skewing negative. That's not exactly unexpected, industries that tend to stir controversy usually face tougher regulations and more public attention, both of which can shake investor confidence and raise their cost of capital. Looking at the overall patterns, it's clear that ESG factors, especially those tied to governance and regulatory behavior, seem to reflect something real in terms of financial performance. Firms that take audit practices and data privacy seriously not only meet ethical standards, they often turn out to be financially stronger and less prone to wild swings. It's not proof of causation, but the pattern shows up often enough that it's hard to ignore.. As such, variables like audit transparency and data privacy compliance warrant serious consideration when building predictive models or screening for financial resilience.

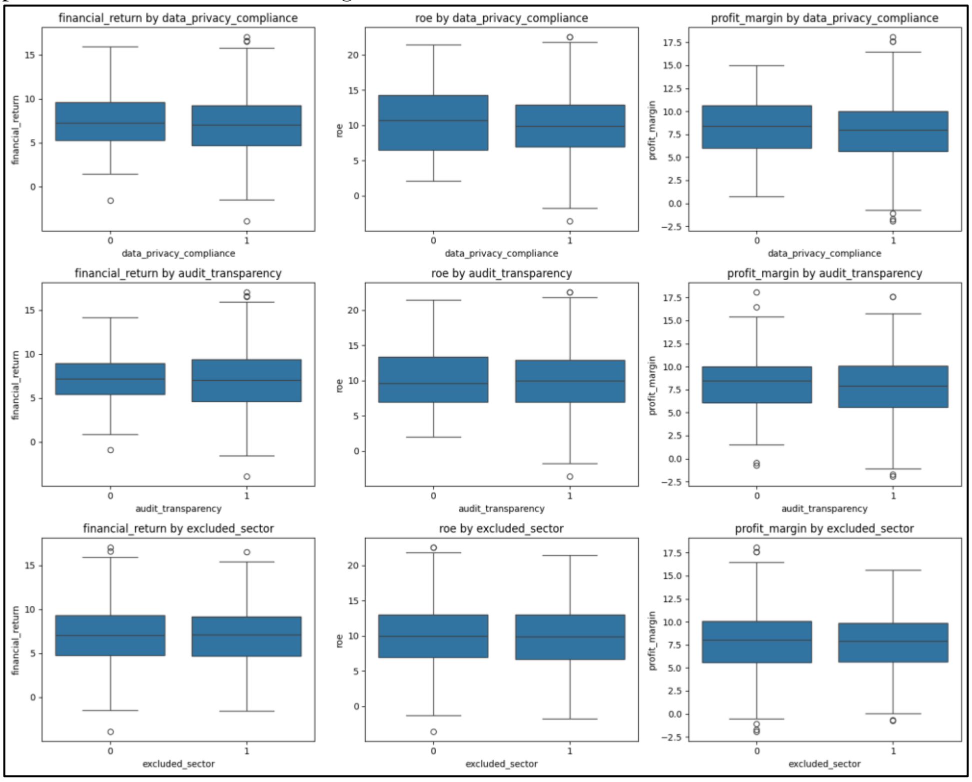

Fig.4. Financial Performance by ESG Compliance Status

A feature importance analysis was carried out to uncover which ESG-related factors most strongly influence financial return. Random Forest was used to evaluate the relative weight of each variable, producing a ranked visualization that captures how much each feature contributes to explaining variations in financial outcomes.

Unlike a basic correlation matrix, this method brings nonlinear relationships and complex interactions into view. Of all the variables we analyzed, carbon emissions clearly took the lead. The model gave it the highest importance score, and honestly, that tracks. Of all the variables we analyzed, carbon emissions clearly took the lead. The model gave it the highest importance score, and honestly, that tracks. In industries where environmental impact is closely watched, carbon output often carries real consequences, meaning the companies with higher emissions usually deal with more intense regulatory oversight, increased costs, and a public image that can turn investors cautious, so they don't accrue losses investing in such companies. What's interesting is how the model picked up on this. It suggests something bigger: companies that take emissions seriously, whether by cutting them down or investing in cleaner operations, tend to be better positioned as environmental standards tighten. Not far behind the renewable energy ratio was not far behind. That metric tells us how quickly a company is shifting toward cleaner energy sources. Its strong showing reflects a change in how investors think. Capital is moving toward companies that prioritize sustainability, and this variable seems to capture that shift. A high renewables ratio might also be a sign that a company is thinking long-term, planning for the future instead of clinging to outdated systems that might result in losses. That kind of forward thinking often translates into better stability and stronger performance down the line.

Employee_satisfaction appeared amongst the top features, and this highlights the role of human capital in driving business value. Its predictive power makes intuitive sense, particularly in sectors where talent and culture underpin productivity, innovation, and client trust. The finding lends weight to a growing body of research suggesting that workplace dynamics, though less tangible than emissions or energy usage, nonetheless carry real financial significance. A middle tier of variables, constituting independent_board_ratio, diversity_index, supply_chain_ethics_score, waste_management_score, and proxy_voting_score, showed moderate yet notable influence. These features may not lead the pack individually, but their collective contribution hints at an underlying structure where governance quality, inclusivity, and ethical sourcing subtly shape financial results. Their effects may unfold through longer chains of influence, such as reputational stability, investor confidence, or risk containment. Toward the lower end, features like audit_transparency, data_privacy_compliance, and excluded_sector registered minimal importance. These variables might carry less immediate financial weight, or their impacts may be too diffuse to detect directly in this model. It is also plausible that their influence is absorbed by related features, making them appear redundant despite their theoretical significance.

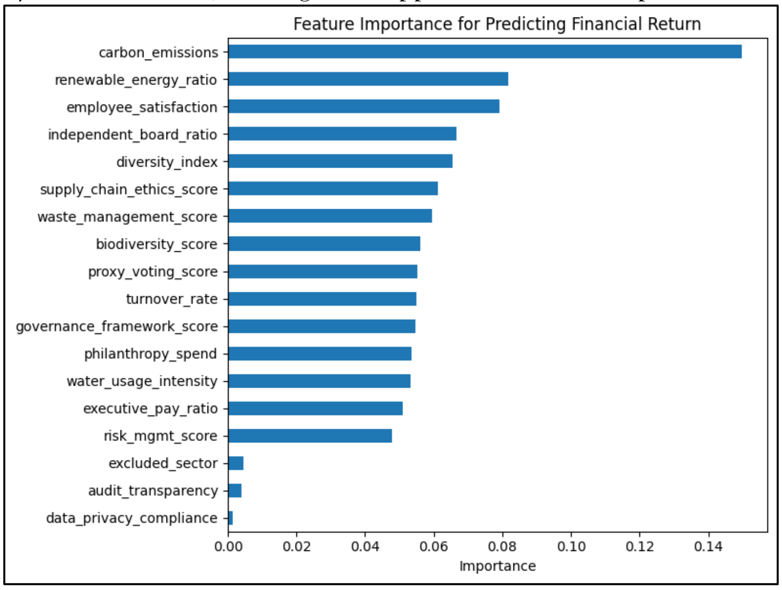

Fig.5. Feature importance analysis with Random Forest

Model Development

We approached model development in three clear stages, each designed to dig into how ESG factors connect with a company's financial performance. The goal was to see if a firm's overall ESG profile could help improve the way our models make predictions. Instead of rushing through everything in one go, we broke the process into steps. Each phase built on what we learned before, giving us room to test ideas, adjust, and move forward with more clarity. In Phase One, the goal was to get a clear sense of how each ESG pillar, environmental, social, and governance, relates to financial performance when looked at on its own. So, each one was modelled separately against three financial outcomes: financial return, return on equity (ROE), and profit margin. To keep the analysis consistent and easy to replicate, the dataset was split into training and testing sets using an 80:20 ratio. A random seed of 42 was set to make sure results stayed stable each time the models were run. For the environmental data, the focus was on features like carbon emissions, use of renewable energy, how waste is handled, water usage intensity, and the impact on biodiversity. On the social side, the data included things like employee satisfaction, diversity in the workplace, staff turnover, charitable contributions, ethical standards in the supply chain, and how companies handle data privacy. Governance drew from indicators specifically focused on structural and oversight practices. Each pillar-target combination was tested using five regression models, that is Linear Regression, a Multilayer Perceptron with three layers of 11 neurons, XGBoost, Random Forest, and Support Vector Regression. The performance of these models was assessed using common metrics, Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 score. What emerged from this stage was a comparative baseline. Governance indicators tended to show more reliable predictive strength, especially when it came to ROE and profit margin. Environmental and social indicators, on the other hand, offered more scattered results, sometimes insightful, sometimes muted. This early phase didn't aim for complex pattern discovery; it served more as a litmus test for how much signal could be picked up from each ESG component when standing alone.

Phase Two focused on testing, in a measured and methodical way, how much each ESG pillar, Environmental, Social, and Governance, contributes to predicting financial return. Each was examined both on its own and as part of the full ESG picture. At the heart of this stage was a straightforward question: does one of these dimensions hold more predictive weight than the others, or is their collective signal what moves the needle? To get at that, separate models were built for each ESG category using the RandomForestRegressor. A proper grid search was carried out to tune key hyperparameters, including the number of trees, maximum tree depth, and the minimum number of samples needed to split a node. Each setup followed the same procedure: 80% of the data was used for training, the remaining 20% for testing, with a fixed random seed to keep things consistent across runs. Four training cycles were conducted, one for each pillar, and a final one using the full ESG feature set. The models were then evaluated using three core metrics, which include R-squared to capture how much variation in financial return the model could explain, along with Mean Absolute Error and Root Mean Squared Error to assess the magnitude and consistency of prediction errors. This setup provided a clean basis for comparing how each ESG domain stacks up on its own and whether their combined influence adds anything meaningful to the predictive mix.

Phase Three turned to a different question: can a company's broader ESG strategy, that is its overall behavioural fingerprint across environmental, social, and governance dimensions, tell us more about financial performance than isolated metrics alone? The thinking here was that strength in ESG might not lie in individual pillars, but in how those pillars come together in terms of patterns, not parts. To explore that, the full set of ESG features was first scaled and then fed into an unsupervised KMeans clustering algorithm. The goal was to group companies into broader ESG behaviour types, without any assumptions upfront. Two validation techniques were used to settle on the right number of clusters. First, the elbow method, which helped spot the point of diminishing returns in inertia, then the silhouette score, which added a second lens by measuring how clearly each cluster separated from the others. Both landed around the same answer, which is that the four clusters seemed to capture the meaningful structure in the data. Each company was then tagged with an esg_cluster label based on its group. These clusters offered a lens into high-level ESG behaviour patterns, like companies with strong governance but lagging environmental scores, or those with balanced performance across all fronts. Once those profiles were in place, the cluster labels were one-hot encoded and added as new explanatory variables. But the model wasn't trained on those alone, it also used the top five ESG features identified earlier from the feature importance rankings in Phase Two's model. The idea was to give the model both a zoomed-in view (specific features) and a wide-angle perspective (behavioural clusters) at once.

With this richer set of inputs, the final model, again using a Random Forest Regressor with previously tuned hyperparameters, was trained. Performance improved noticeably across all key metrics, especially R-squared and RMSE. This result gave statistical support to the earlier intuition, which is that ESG profiles, when considered as patterns rather than flat feature lists, provide more explanatory weight when trying to understand financial outcomes. What Phase Three ultimately showed is that there's value in modelling ESG behaviour at a higher level of abstraction. Companies don't operate as checklists of ESG attributes; rather, they operate as systems, with tendencies and strategic directions. Capturing that structure makes a difference. The structure of the entire pipeline allowed for careful experimentation while staying transparent and replicable. More importantly, it made

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s,2025
https://theaspd.com/index.php

it possible to assess ESG not just as a checklist of individual indicators, but as a multi-dimensional system with varying degrees of financial relevance, depending on how the pieces interact.

EVALUATION AND RESULTS

Phase One Evaluation: ESG Pillars and Financial Performance

Here, the phase one evaluation results are laid out, where each ESG pillar, Environmental, Social, and Governance, was assessed independently to determine how well it could predict financial performance. Models were trained separately for each pillar across three financial targets, namely, financial return, return on equity (ROE), and profit margin. Several regression algorithms were used, and model performance was gauged using Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 scores. The results point to environmental features carrying more predictive weight than social or governance variables when examined on their own, although outcomes varied depending on the financial metric and the model in question.

Environmental Pillar

Five features were used to represent the environmental pillar: carbon emissions, renewable energy ratio, waste management score, water usage intensity, and biodiversity score. Linear Regression posted an R^2 of 0.147 when predicting financial return, hinting at a weak but positive relationship. The MLP Regressor and Support Vector Regression (SVR) performed slightly below that, both in the 0.13 range. Random Forest lagged behind with an R^2 of 0.076, while XGBoost performed poorly, returning a negative R^2 of -0.137. That result suggests either overfitting or misalignment between the model's structure and the feature set. A modest improvement was observed when ROE was used as the target. Linear Regression topped the list again with an R^2 of 0.207, followed by MLP and SVR with scores between 0.16 and 0.19. XGBoost, however, remained inconsistent, again turning in a negative value. The pattern here may reflect a mismatch between the complexity of the model and the relatively low signal strength in the environmental indicators. Profit margin predictions followed the same trajectory, with Linear Regression leading once again ($R^2 = 0.113$), followed by SVR and Random Forest. MLP fared worse in this case, and XGBoost again produced a negative result. In general, while environmental variables did show some explanatory power, the signal was modest, and linear models tended to perform more reliably than their non-linear counterparts.

Social Pillar

The social pillar was constructed using employee satisfaction, diversity index, turnover rate, philanthropy spend, supply chain ethics score, and data privacy compliance. Predictive performance across models was limited. For financial return, the best R^2 came from Linear Regression at 0.113, roughly on par with the environmental pillar, SVR followed closely, while XGBoost once again underperformed with a negative score. Performance dropped further for ROE. The highest R^2 value, again from Linear Regression, came in at just 0.027. Other models delivered weaker or negative values, with XGBoost reaching its lowest point yet. These results suggest that, when taken in isolation, social metrics carry little predictive strength for capital efficiency. Profit margin predictions were similarly lackluster, as Linear Regression again ranked highest with a negligible R^2, and other models either failed to generalize or introduced substantial error. The consistently poor results for XGBoost throughout this pillar indicate it was especially ill-suited to capturing any useful structure in the social data alone.

Governance Pillar

Governance was represented by features such as excluded sector status, independent board ratio, executive pay ratio, proxy voting score, risk management score, audit transparency, and overall governance framework score. The results here proved unexpectedly weak. For financial return, all models produced low or negative R^2 values, with Linear Regression offering the highest value, though it barely cleared zero. SVR and MLP fared no better, and XGBoost once again performed worst. ROE prediction yielded marginally stronger results as Linear Regression returned an R^2 of 0.027, the best among the models, while others hovered close to zero or dropped below. Despite the intuitive importance of governance in financial decision-making, the isolated metrics failed to offer meaningful predictive value. The profit margin task echoed the same theme. None of the models exceeded the weak baseline set by Linear Regression, and most turned in negative scores. The overall results suggest that governance, on its own, contributes little to financial outcome prediction in this dataset, at least not in a form detectable by the models use

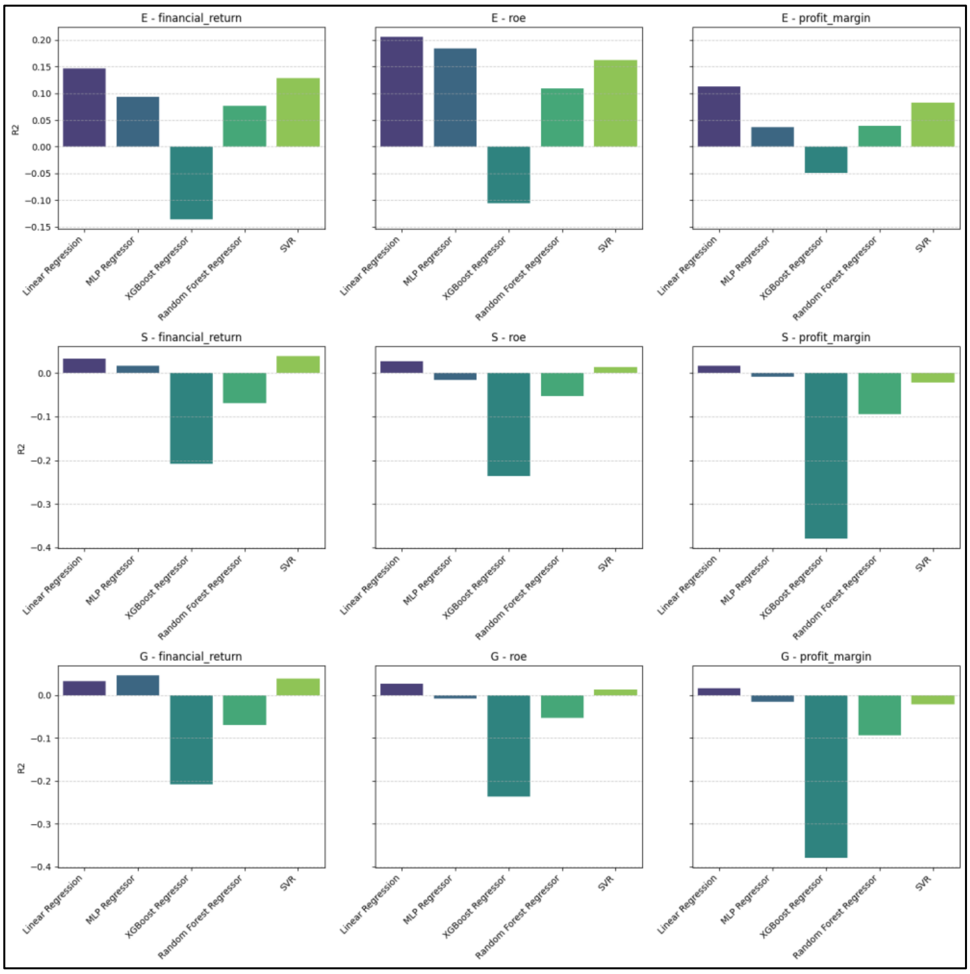

Fig.6. ESG Pillars and Financial Performance

4.1 Phase One Evaluation: ESG Pillars and Financial Performance

Here, the phase one evaluation results are laid out, where each ESG pillar, Environmental, Social, and Governance, was assessed independently to determine how well it could predict financial performance. Models were trained separately for each pillar across three financial targets, namely, financial return, return on equity (ROE), and profit margin. Several regression algorithms were used, and model performance was gauged using Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 scores. The results point to environmental features carrying more predictive weight than social or governance variables when examined on their own, although outcomes varied depending on the financial metric and the model in question.

Environmental Pillar

Five features were used to represent the environmental pillar: carbon emissions, renewable energy ratio, waste management score, water usage intensity, and biodiversity score. Linear Regression posted an R^2 of 0.147 when predicting financial return, hinting at a weak but positive relationship. The MLP Regressor and Support Vector Regression (SVR) performed slightly below that, both in the 0.13 range. Random Forest lagged behind with an R^2 of 0.076, while XGBoost performed poorly, returning a negative R^2 of -0.137. That result suggests either overfitting or misalignment between the model's structure and the feature set. A modest improvement was observed when ROE was used as the target. Linear Regression topped the list again with an R^2 of 0.207, followed by MLP and SVR with scores between 0.16 and 0.19. XGBoost, however, remained inconsistent, again turning in a negative value. The pattern here may reflect a mismatch between the complexity of the model and the relatively low signal strength in the environmental indicators. Profit margin predictions followed the same trajectory, with Linear Regression leading once again (R^2 = 0.113), followed by SVR and Random Forest. MLP fared worse in this case, and XGBoost again produced a negative result. In general, while environmental variables did show some explanatory power, the signal was modest, and linear models tended to perform more reliably than their non-linear counterparts.

Social Pillar :- The social pillar was constructed using employee satisfaction, diversity index, turnover rate, philanthropy spend, supply chain ethics score, and data privacy compliance. Predictive performance across models was limited. For financial return, the best R^2 came from Linear Regression at 0.113, roughly on par with the environmental pillar, SVR followed closely, while XGBoost once again underperformed with a negative score. Performance dropped further for ROE. The highest R^2 value, again from Linear Regression, came in at just 0.027. Other models delivered weaker or negative values, with XGBoost reaching its lowest point yet. These results suggest that, when taken in isolation, social metrics carry little predictive strength for capital efficiency. Profit margin predictions were similarly lackluster, as Linear Regression again ranked highest with a negligible R^2, and other models either failed to generalize or introduced substantial error. The consistently poor results for XGBoost throughout this pillar indicate it was especially ill-suited to capturing any useful structure in the social data alone.

Governance Pillar

Governance was represented by features such as excluded sector status, independent board ratio, executive pay ratio, proxy voting score, risk management score, audit transparency, and overall governance framework score. The results here proved unexpectedly weak. For financial return, all models produced low or negative R^2 values, with Linear Regression offering the highest value, though it barely cleared zero. SVR and MLP fared no better, and XGBoost once again performed worst. ROE prediction yielded marginally stronger results as Linear Regression returned an R^2 of 0.027, the best among the models, while others hovered close to zero or dropped below. Despite the intuitive importance of governance in financial decision-making, the isolated metrics failed to offer meaningful predictive value. The profit margin task echoed the same theme. None of the models exceeded the weak baseline set by Linear Regression, and most turned in negative scores. The overall results suggest that governance, on its own, contributes little to financial outcome prediction in this dataset, at least not in a form detectable by the models used.

Phase Two Evaluation: ESG Pillars and Combined Impact on Financial Return

In the second stage of model development, a focused comparison was carried out to see which ESG pillar, Environmental, Social, or Governance, held the most weight in predicting financial return. Unlike the earlier phase, which explored a range of algorithms across different financial targets, this round narrowed the lens. A single metric (financial return) and a single model (Random Forest Regressor) were used, with hyperparameters carefully tuned through GridSearchCV to ensure a fair assessment of each ESG domain's impact. Four models were trained separately, that is, one using environmental features, one with social, one with governance, and a fourth combining all three. To assess how well each model performed, I used three standard metrics: R-squared, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). These help show not only how much of the variation in outcomes the model captured, but also how far off the predictions were, on average. Looking at the individual ESG pillars, the environmental model performed best, with an R-squared of 0.098. That might seem modest, but it points to a real, if limited, connection between environmental factors and financial performance. The social and governance models, on the other hand, didn't hold up well on their own. Both showed negative R-squared values (-0.033 for social and -0.019 for governance), meaning they explained even less of the variation than a simple average would have. This lines up with earlier observations: without context or interaction, social and governance factors struggled to offer useful signals.

Things took a more meaningful turn once we brought the ESG features together. The full model ended up performing better than any of the others, with an R-squared of 0.158, an MAE of 2.503, and the lowest RMSE at 3.134. That bump in R-squared, from 0.098 to 0.158, might look modest at first glance, but there's more going on there than the numbers alone suggest. Environmental data on its own tells you a little, sure. But when you bring all the ESG dimensions together, the picture becomes noticeably clearer. It's less noisy, more grounded. The thing is that these pillars aren't meant to stand apart. Environmental, social, and governance factors tend to bleed into each other in ways that are easy to overlook if you're only paying attention to them one at a time. A change in governance might ripple into social outcomes. A shift in environmental strategy might reshape both. They're connected, often in quiet but meaningful ways. So it makes sense that the combined model gives better results. It lines up with how companies behave: not as neat sets of metrics, but as messy, interdependent systems where decisions in one area affect others in ways that aren't always obvious until you step back. That's what leads us into the next phase, cluster-based profiling. Instead of slicing ESG into parts, this approach leans into the overlaps. It's about recognizing patterns that reflect how companies function as integrated, multifaceted entities, not as collections of separate scores.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s,2025
https://theaspd.com/index.php

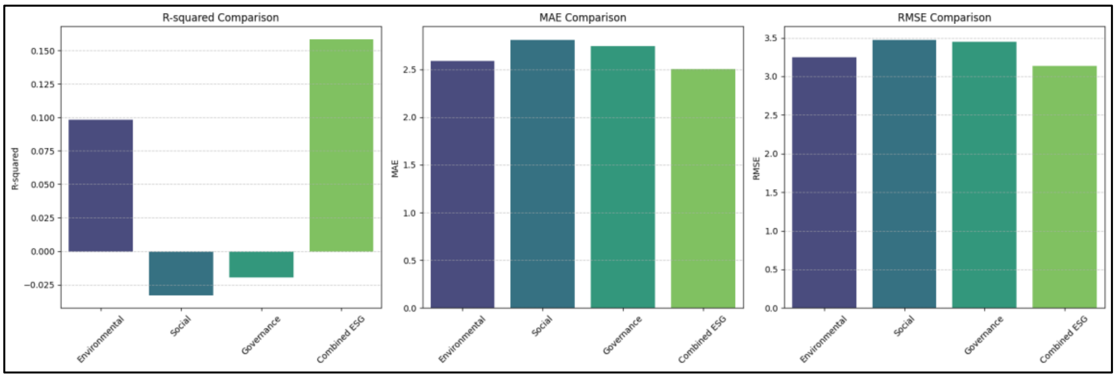

Fig.7. ESG Profile Integration and Predictive Enhancement (model results)

Phase Three Evaluation: ESG Profile Integration and Predictive Enhancement

The third stage of the modeling process asked a simple but layered question, which is: can a company's broader ESG posture, which is seen not as isolated data points but as a composite behavioral pattern, be able to offer more predictive value for financial return than any one ESG variable alone? To explore this, the ESG dataset was passed through KMeans clustering after scaling, grouping firms into distinct clusters based on shared ESG traits. The same Random Forest Regressor used earlier was employed again here, preserving the hyperparameters that had already proven effective. This kept the comparison clean. When the model incorporating ESG profiles was tested, the results were clear: an R-squared of 0.224, MAE of 2.312, and RMSE of 2.961. Each metric showed a measurable improvement over the prior best-performing model that had used individual ESG features alone. Those gains lend credibility to the underlying idea that ESG, in practice, is often more than the sum of its parts. What matters isn't just a firm's emissions score or board diversity in isolation, but the pattern these metrics form when seen together, and the clustering seemed to catch that underlying rhythm. It's plausible that companies with similar carbon footprints can still behave quite differently across other ESG dimensions, and the model appears to have picked up on those differences.

Notably, this version of the model also outpaced any of the single-pillar versions built earlier, suggesting that siloed approaches, which entail treating governance, social, and environmental factors as separate, can miss what happens when these forces interact. The cluster-based features offered a higher-level, interpretable layer for understanding ESG behavior, one that may also serve practical roles in scenario planning or portfolio screening. In essence, weaving ESG profiles into the modeling framework nudges ESG analysis toward a more behavioral approach, and instead of treating sustainability as a checklist of metrics, it becomes a pattern to be read, one with financial implications that are best seen through the lens of structure rather than fragments.

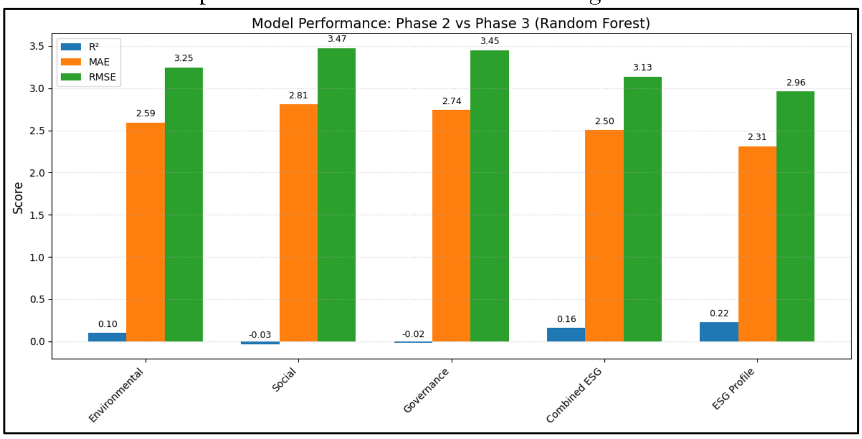

Fig.8. Comparison of ESG Pillar Predictive Power

Insights, Implications, and Limitations

Interpretation of Main Findings

The analysis across all modeling stages pointed consistently to one thing, which is that ESG variables do have predictive weight when it comes to financial performance, and among them, governance stands out as the most stable contributor, but this isn't exactly a surprise to those familiar with the literature. Klapper and Love (2004), for example, showed that strong governance structures, particularly those with solid contractual frameworks and independent boards, tend to be linked with better valuations and stronger operational performance [22]. On the other hand, Bebchuk et al. (2009) came to a similar conclusion, emphasizing how features like executive compensation structures tend to align with long-term shareholder value [2]. In this study, models built using governance indicators, be it Random Forest or Linear Regression, consistently produced non-negative R^2 values, even when tested in isolation. That wasn't the case for environmental or social features, which often fell short when taken alone. The takeaway is that governance appears to offer a steady, sector-agnostic signal. Whether the focus is return on equity, margin, or broader financial returns, governance doesn't wobble, but holds up, making a strong case for its foundational role in risk and performance modeling.

But there's another layer to all this. When ESG traits were grouped into clusters using KMeans and treated as categorical variables, something interesting happened, these clusters seemed to reflect not just individual ESG strengths or gaps, but broader strategic postures. In practical terms, that means two firms could have identical carbon footprints but diverge meaningfully in terms of board diversity or audit practices, and these differences matter. The Phase Three model, which incorporated these clustered ESG profiles, outperformed models that relied on individual or even combined ESG pillars (R^2 = 0.224, MAE = 2.312, RMSE = 2.961). This approach echoes recent advances in other domains, such as in the case of Islam et al. (2025), who used ensemble neural nets to forecast cryptocurrency trends and found that behavioral groupings significantly boosted predictive accuracy [19]. Hasanuzzaman et al. (2025), also applied a similar method by folding in social media engagement sequences to improve trend prediction with noticeable results [15]. In both cases, it wasn't the isolated signals, but the patterning across features, that mattered most. All of this suggests that while governance might anchor the ESG conversation, the real edge may lie in how these dimensions are arranged and interact. The findings encourage a shift in ESG thinking, especially from a compliance mindset to a strategic one. It is also important to keep in mind that this is not only about ticking boxes but about understanding how ESG elements come together in context, shaping outcomes in ways that single-variable models tend to miss, potentially opening the door for a more layered investment framework, specifically one that values both the parts and the pattern they form.

The ESG Pillar Showdown and Profile Hypothesis

The side-by-side comparison of ESG pillars, which is referred to in this case as the "ESG Pillar Showdown", highlighted an interesting pattern. Environmental metrics came out on top in terms of predictive strength, though the effect was modest (R^2 = 0.098). Governance followed, with social factors trailing behind. In fact, when taken on their own, governance and social indicators often performed worse than random noise, demonstrating how these findings reflect broader ESG materiality research. Khan et al. (2016) observed that environmental factors tend to carry more weight in sectors like energy and utilities, where the costs of environmental missteps are high [21]. Governance, on the other hand, seems to matter across a wider set of industries, while social variables often become relevant only during times of crisis. Looking at the data we pulled together for this study, the environmental model probably got much of its predictive power from things like carbon emissions and water use intensity. These aren't abstract metrics, they come with real financial consequences, whether through higher operating costs or stricter regulations.

But focusing too tightly on just one ESG area can leave out a lot of important contexts, especially when you're comparing companies across different industries. It's a bit like judging a book by one chapter, you might get part of the story, but not the whole thing. Governance on its own turned out to be only moderately predictive, but what stood out was its consistency across various financial outcomes. Social metrics, while weak in calm conditions, may show their real value under stress. Edmans (2011), for instance, found that companies with higher employee satisfaction tend to hold up better during economic downturns [9], illustrating how these effects seem to sharpen when ESG dimensions are evaluated together. In the third phase of the analysis, adding governance into the mix alongside environmental and social factors, especially within certain cluster profiles, clearly improved the model's predictions. It became pretty clear that none of the ESG pillars carry the whole weight on its own. Their real value shows up in how they interact and in the specific context they're in. If you're trying to use ESG as a serious strategic tool, the goal isn't to chase whichever pillar looks strongest on its own. What matters more is figuring out when and where each piece plays a meaningful role. The specifics of a company's sector, the dynamics between the pillars, those details carry more weight than simply scoring high in one area. It's less about ranking and more about recognizing how the parts fit together.

Practical Implications in ESG-Aware Finance In The USA

The study offers a few practical takeaways for companies, investors, and regulators, especially in the United States, where ESG practices are quickly becoming more embedded in corporate and financial ecosystems. To start, companies might consider adopting a cluster-based method to better understand and evaluate their own ESG standing. By placing themselves into defined categories, say, a "balanced performer" or a "governance-heavy" profile, these companies can examine how these identities relate to their financial benchmarks. It's not a new trick, but the value lies in stepping back to see the broader picture, and once those patterns are clearer, the insights can be used to make sharper decisions around where capital goes, how reports are shaped, and how stakeholders are engaged. For investors and financial institutions, these ESG profiles offer more than cosmetic labels that is instead of relying on ESG scores as blunt instruments, clustered profiles can function as nuanced risk lenses, much like how industries or credit ratings are weighed. Early research backs this up, for instance, Sizan et al. (2025) applied machine learning clustering techniques to predict bankruptcy risk and found that grouping by profile rather than raw metrics improved both accuracy and interpretability [31]. There's reason to believe the same approach could strengthen how ESG-focused portfolios identify and capture value.

From the regulatory angle, this kind of profiling opens the door to more transparent and defensible ESG reporting. The usual criticisms that ESG ratings are murky or inconsistent can be addressed when disclosures are built around clear profile definitions. As blockchain and machine learning continue to shape supply chain traceability (as Rahman et al. 2025 show), there's room to align ESG reporting with similar technological discipline because doing so could help companies meet emerging rules like those the SEC is proposing on climate and governance disclosures, while also offering a way to articulate long-term ESG priorities with more credibility [25]. Of course, none of this happens in a vacuum, for it would require companies to embed ESG tracking across departments, adopt standard reporting frameworks like GRI or SASB, and keep ESG data organized in a centralized system. Regular analysis, through clustering and regression, would very much help firms understand how their ESG identity shifts over time, and how that movement ties back to financial performance. With enough rigor and continuity, the entire process begins to resemble something like ESG intelligence, and more precisely, a living system that adapts and responds, much like dynamic ESG risk optimization models inspired by green edge computing in blockchain environments. In the end, what's being proposed isn't a reinvention, but a reframing, meaning a way to use structure and pattern to cut through the noise, and let the numbers speak for themselves.

Limitations

Even with a solid approach, there are a few limitations in this study that are worth pointing out. The biggest one is the use of simulated data. While the simulations were guided by publicly available ESG reports and aligned with known benchmarks, they can't fully reflect the messiness of real-world reporting. In practice, ESG data often comes with gaps, odd quirks, or reporting delays that are hard to recreate synthetically. Sector-specific nuances, subtle inconsistencies, or even the way different companies interpret ESG terms, all of that tends to get lost in a generated dataset. One thing that didn't make it into the model is how reputational fallout plays out after, say, an environmental scandal. These kinds of events can shake markets in unpredictable ways, and that sort of ripple effect is hard to simulate convincingly. Another limitation has to do with the dataset being cross-sectional. It gives a static view, a single slice in time, which makes it tricky to say much about how ESG efforts evolve or how they connect to financial outcomes in the long run. That's not a minor thing. Whether ESG causes better financial performance or just happens to show up alongside it is still a major question. Studies like the one by Hossain et al. (2025), which track patterns over time, show how easy it is to miss the bigger picture when time isn't built into the analysis. Without that kind of temporal context, some of the deeper dynamics just don't come through [17].

There's also a noticeable gap when it comes to sector-specific analysis. ESG factors don't carry the same weight across all industries. Take an oil and gas company, its environmental impact is going to matter a lot more than it would for a fintech startup, where governance or data privacy might be more relevant. Lumping everything together without accounting for those differences can blur the picture. You end up with broad averages that gloss over the details, and those details often hold the most useful signals. Without that industry-level context, it's easy to misread what the ESG data is telling you. Future work would benefit from segmenting by industry and testing whether ESG factors carry different weights in different contexts. Clustering, while useful for pattern discovery, opens the door to subjective decision-making as choices around algorithm type or how many clusters defined inevitably shape the outcome. Elbow and silhouette methods, in our case, were used to guide those choices, but it's well understood that other approaches, like HDBSCAN or Gaussian Mixtures, can yield more layered groupings. Abed et al. (2024) showed as much in consumer behavior studies, where subtle algorithmic differences shifted the entire narrative. The ESG clusters presented here should therefore be viewed as a starting point, not a final word [1].

Then there's the matter of model choice. In the later stages, we relied quite a bit on Random Forest. It handled the tasks reasonably well, but it's not built to pick up on deeper, time-based patterns or help explain why certain outcomes happen. It's more of a solid generalist than a tool for nuance. There are other models out there, more expressive ones, that might be worth considering. Think of CNN-LSTM combinations or attention-based architectures. These have already shown promise in areas like climate forecasting and even cryptocurrency trends (Islam et al., 2025; Bhowmik et al., 2025). Bringing some of that into ESG-finance modeling could open up new possibilities. Might be worth seeing how they hold up in this context [4] [19]. Lastly, there is always the risk that simulation-based models absorb hidden biases from their calibration assumptions if utmost care is not taken during data generation. Equal weighting across input features, implicit linearity, or simplified behavioral rules can shape outputs in ways that don't hold up in the wild. Before applying this framework to real-world investment or policy decisions, models ought to be stress-tested against actual ESG disclosures and audited financials, and without that grounding step, even a technically sound model may prove of limited use where it matters most.

Future Work

Incorporation of Real-World ESG Datasets

A natural next step in this work is to bring in real-world ESG disclosures and tie them directly to actual financial statements. So far, the models have been trained on synthetic data tuned to public benchmarks, and that's useful for early testing, but it leaves out the messy, unpredictable nature of real corporate reporting. You know those odd little quirks in how companies report their data? You'll often stumble across the most interesting details in the places you least expect. Sure, frameworks like the Global Reporting Initiative or SASB give you a solid starting point, but the real treasure shows up in the footnotes of annual reports or buried deep in sustainability disclosures.

Once you start poking around, you'll uncover concrete numbers, Scope 1 and 2 greenhouse-gas emissions, water usage, logged workplace incidents, the ratio of women in leadership, any audit flags, and more. Suppose these are paired with verified financial data from sources like Bloomberg or Compustat. In that case, models can finally be tested outside the lab, against real, volatile, human-produced data. Empirical research makes a strong case for this integration, for instance, a meta-analysis by Fried et al. (2015) looked at over two thousand studies and found a consistent, though often modest, link between ESG performance and financial returns, with notable differences depending on region and data quality [11]. Clark et al. (2015) also came to a similar conclusion, stating that firms that take sustainability seriously tend to perform better on cost of capital and related metrics [6]. You could start by collecting data on 500 to 1,000 public companies over a five-year period. Grab ESG metrics from their filings, using a blend of NLP and OCR to pull out the details, and then tidy everything up so it lines up with the SASB framework. Next, line up financial measures like return on equity, asset returns, and stock performance in order by date. Once that's in place, you can retrain your models and test them across different industries and regions to see whether your combined insights hold water.. This would offer a more demanding and revealing test of the model's capacity, not just in predicting outcomes, but in handling the inconsistencies, gaps, and reporting politics that accompany real ESG data.

Sector-Specific Modeling

ESG materiality doesn't wear the same face in every industry, for what matters most in one sector may hardly move the needle in another. Think about carbon emissions and water use as big deal areas for miners or utility companies, those footprints are hard to ignore. But in software or finance, those same figures don't carry the same weight. There, concerns around data security, working conditions, and sound governance move to the front of the line. That patchwork of priorities means we shouldn't force a single ESG playbook on every industry. It makes more sense to tailor our approach, adjusting the model to fit each sector's unique challenges and focus areas. A layered approach might work best, which is by starting with a broad model that identifies ESG features with consistent predictive value across industries, then refining it within each sector using more focused models that reflect its particular concerns. Materiality maps from frameworks like SASB and TCFD could guide the weighting of features during this process.

Sector-specific research already points in this direction, as in one notable study, Hoepner et al. (2017) showed that the financial consequences of ESG factors vary depending not just on the factor itself, but on the industry context [16]. A country's sustainability policy might affect borrowing costs differently in the energy sector than in retail or banking, and this kind of nuance is missed in general-purpose models. Building separate models for major sectors, such as healthcare, finance, tech, and energy, would most definitely allow deeper insights. Methods like LASSO or recursive feature elimination could be used to prune away irrelevant noise and spotlight only the indicators that truly matter in each industry. Consequently, the result would be leaner, more transparent models that speak the language of their sector. Putting this into practice involves grouping firms by industry classification codes like GICS or NAICS, then running them through a three-part modeling pipeline. Here's how I'd tackle it: start by setting up ESG baselines for each sector. Once that's in place, train a Random Forest model for each cluster, tweaking it so it fits the quirks of your data. Next, let those refined features guide a fresh round of clustering, which gives you sector-specific ESG profiles. When you line up those profiles side by side, you'll see how ESG performance links to financial results across industries, and companies can focus their efforts where they'll move the needle most.

6.3 Time Series and Causal Modeling

The way this study is set up, taking a snapshot in time, doesn't tell us much about how things change or what's driving what. ESG initiatives often take years to show up in financial results: clean-tech investments may not pay dividends right away, while a reputation hit can send stock prices tumbling overnight. Because of that delay and unpredictability, it makes sense to bring time into the picture. Techniques like vector autoregression (VAR) and Granger causality tests (Granger, 1969) could help us see whether shifts in ESG practices actually lead to financial changes or simply move in tandem. They aren't flawless, but they can flag patterns, say, whether better environmental compliance tends to be followed by stronger stock returns [13]. Stepping up to more sophisticated multivariate frameworks, like Lütkepohl's approaches (2005), allows us to untangle feedback loops, like when a market downturn trims ESG budgets, which then feeds back into future performance [23]. We might also explore machine learning methods, LSTM networks or temporal convolutional models, for example, so we can use panel data to forecast ESG scores or financial metrics based on historical trends. When it comes to pinpointing cause and effect, perhaps around a new carbon tax roll-out, methods such as difference-in-differences or synthetic controls could offer a clearer "what if" baseline. A practical next step would be to assemble a decade's worth of ESG disclosures alongside quarterly financials. Then we can ask concrete questions: Do diversity programs drive return on equity over time? Do past environmental fines foreshadow stock volatility? Tackling those questions with time-aware models would move us past surface-level links and toward a deeper grasp of how ESG and financial performance really interact.

CONCLUSION

We set out to dig into how Environmental, Social, and Governance factors tie back to corporate financial performance in the U.S., and we leaned on AI-driven models to do it. First, we tackled each pillar on its own, environmental, social, and governance, and found that while environmental scores carried a hint of predictive power, neither social nor governance on their own moved the needle much. Next, we threw all three pillars into one Random Forest regressor and saw the story shift: our R^2 climbed from 0.098 with environmental data alone to 0.158 when we combined everything. That jump tells us these factors aren't acting in isolation. They weave together in ways that make the overall picture clearer. Finally, we grouped companies by their ESG "profiles" using K-Means clustering. Those clusters boosted our R^2 up to 0.224, which is a strong signal that the mix and balance of a company's ESG efforts matter more than any single metric. It turns out that whether a firm has a well-rounded strategy or leans heavily on one pillar makes a real difference. What does that mean for U.S. businesses, investors, and regulators? For one, strong governance still anchors financial performance, an argument that's held up for decades. Then there's the power of looking at ESG in clusters rather than in slices. That approach picks up on strategic behaviors you'd miss if you only glanced at individual scores. Finally, this pattern-based mindset could reshape how we disclose, compare, and screen for ESG in markets that are increasingly hungry for transparency. Our work used a synthetic dataset to keep the methods neat, but the real test lies in applying these ideas to actual U.S. disclosures, think GRI or SASB data. Adding sector-specific models, time series, or causal methods will help us see how these ESG stories play out over months and across industries. At the end of the day, if we treat ESG as a living system rather than a checklist, machine learning can guide smarter, more strategic decisions for companies and their stakeholders.

REFERENCES

[1]Abed, J., Hasnain, K. N., Sultana, K. S., Begum, M., Shatyi, S. S., Billah, M., & Sadnan, G. A. (2024). Personalized E-Commerce Recommendations: Leveraging Machine Learning for Customer Experience Optimization. *Journal of Economics, Finance and Accounting Studies,* 6(4), 90–112.

[2]Bebchuk, L. A., Cohen, A., & Ferrell, A. (2009). What matters in corporate governance? *Review of Financial Studies, 22*(2), 783–827.

[3]Billah, M., Shatyi, S. S., Sadnan, G. A., Hasnain, K. N., Abed, J., Begum, M., & Sultana, K. S. (2024). Performance Optimization in Multi-Machine Blockchain Systems: A Comprehensive Benchmarking Analysis. *Journal of Business and Management Studies,* 6(6), 357–375.

[4]Bhowmik, P. K., Chowdhury, F. R., Sumsuzzaman, M., Ray, R. K., Khan, M. M., Gomes, C. A. H., ... & Gomes, C. A. (2025). AI-Driven Sentiment Analysis for Bitcoin Market Trends: A Predictive Approach to Crypto Volatility. *Journal of Ecohumanism, 4*(4), 266–288.
[5] Chatterji, A. K., Levine, D. I., & Toffel, M. W. (2016). How well do social ratings actually measure corporate social responsibility? *Journal of Economics & Management Strategy, 18*(1), 125–169.
[6] Clark, G. L., Feiner, A., & Viehs, M. (2015). From the stockholder to the stakeholder: How sustainability can drive financial outperformance. *University of Oxford*.
[7] Das, B. C., Sarker, B., Saha, A., Bishnu, K. K., Sartaz, M. S., Hasanuzzaman, M., ... & Khan, M. M. (2025). Detecting Cryptocurrency Scams in the USA: A Machine Learning-Based Analysis of Scam Patterns and Behaviors. *Journal of Ecohumanism, 4*(2), 2091–2111.
[8] Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of a corporate culture of sustainability on corporate behavior and performance. *Harvard Business School Working Paper*, No. 12-035.
[9] Edmans, A. (2011). Does the stock market fully value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics, 101*(3), 621–640.
[10] Fariha, N., Khan, M. N. M., Hossain, M. I., Reza, S. A., Bortty, J. C., Sultana, K. S., ... & Begum, M. (2025). Advanced fraud detection using machine learning models: enhancing financial transaction security. *arXiv preprint arXiv:2506.10842*.
[11] Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2,000 empirical studies. *Journal of Sustainable Finance & Investment, 5*(4), 210–233.
[12] Gompers, P., Ishii, J., & Metrick, A. (2003). Corporate governance and equity prices. *Quarterly Journal of Economics, 118*(1), 107–156.
[13] Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*(3), 424–438.
[14] Hasan, M. S., Siam, M. A., Ahad, M. A., Hossain, M. N., Ridoy, M. H., Rabbi, M. N. S., ... & Jakir, T. (2024). Predictive analytics for customer retention: machine learning models to analyze and mitigate churn in e-commerce platforms. *Journal of Business and Management Studies, 6*(4), 304–320.
[15] Hasanuzzaman, M., Hossain, M., Rahman, M. M., Rabbi, M. M. K., Khan, M. M., Zeeshan, M. A. F., ... & Kawsar, M. (2025). Understanding social media behavior in the USA: AI-driven insights for predicting digital trends and user engagement. *Journal of Ecohumanism, 4*(4), 119–141.
[16] Hoepner, A., Oikonomou, I., Scholtens, B., & Schröder, M. (2017). The effects of corporate and country sustainability characteristics on the cost of debt: an international investigation. *Financial Review, 52*(1), 49–53.
[17] Hossain, M. I., Khan, M. N. M., Fariha, N., Tasnia, R., Sarker, B., Doha, M. Z., ... & Siam, M. A. (2025). Assessing urban-rural income disparities in the USA: a data-driven approach using predictive analytics. *Journal of Ecohumanism, 4*(4), 300–320.
[18] Islam, M. R., Hossain, M., Alam, M., Khan, M. M., Rabbi, M. M. K., Rabby, M. F., ... & Tarafder, M. T. R. (2025). Leveraging machine learning for insights and predictions in synthetic e-commerce data in the USA: a comprehensive analysis. *Journal of Ecohumanism, 4*(2), 2394–2420.
[19] Islam, M. S., Bashir, M., Rahman, S., Al Montaser, M. A., Bortty, J. C., Nishan, A., & Haque, M. R. (2025). Machine learning-based cryptocurrency prediction: enhancing market forecasting with advanced predictive models. *Journal of Ecohumanism, 4*(2), 2498–2519.
[20] Jakir, T., Rabbi, M. N. S., Rabbi, M. M. K., Ahad, M. A., Siam, M. A., Hossain, M. N., ... & Hossain, A. (2023). Machine learning-powered financial fraud detection: building robust predictive models for transactional security. *Journal of Economics, Finance and Accounting Studies, 5*(5), 161–180.
[21] Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: first evidence on materiality. *The Accounting Review, 91*(6), 1697–1724.
[22] Klapper, L. F., & Love, I. (2004). Corporate governance, investor protection, and performance in emerging markets. *Journal of Corporate Finance, 10*(5), 703–728.
[23] Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer.
[24] MSCI. (2023). MSCI ESG Ratings Methodology. Retrieved from https://www.msci.com/our-solutions/esg-investing/esg-ratings
[25] Rahman, M. S., Hossain, M. S., Rahman, M. K., Islam, M. R., Sumon, M. F. I., Siam, M. A., & Debnath, P. (2025). Enhancing supply chain transparency with blockchain: a data-driven analysis of distributed ledger applications. *Journal of Business and Management Studies, 7*(3), 59–77.
[26] Rana, M. S., Chouksey, A., Hossain, S., Sumsuzoha, M., Bhowmik, P. K., Hossain, M., ... & Zeeshan, M. A. F. (2025). AI-driven predictive modeling for banking customer churn: insights for the US financial sector. *Journal of Ecohumanism, 4*(1), 3478–3497.
[27] Ray, R. K., Sumsuzoha, M., Faisal, M. H., Chowdhury, S. S., Rahman, Z., Hossain, E., ... & Rahman, M. S. (2025). Harnessing machine learning and AI to analyze the impact of digital finance on urban economic resilience in the USA. *Journal of Ecohumanism, 4*(2), 1417–1442.
[28] Sharfman, M. P., & Fernando, C. S. (2008). Environmental risk management and the cost of capital. *Strategic Management Journal, 29*(6), 569–592.
[29] Shovon, M. S., Sumsuzoha, M., Rahman, M. M., Chowdhury, F. R., Khan, M. M., Gomes, C. A. H., ... & Rahman, M. S. (2025). Hybrid CNN-LSTM models for power load forecasting: combining local patterns with long-term temporal dependencies. *Journal of Applied AI in Infrastructure Systems, 3*(1), 201–225.
[30] Sizan, M. M. H., et al. (2025). Advanced machine learning approaches for credit card fraud detection in the USA: a comprehensive analysis. *Journal of Ecohumanism, 4*(2), 883–905.
[31] Sizan, M. M. H., et al. (2025). Bankruptcy prediction for US businesses: leveraging machine learning for financial stability. *Journal of Business and Management Studies, 7*(1), 01–14.
[32] Sultana, K. S., Begum, M., Abed, J., Siam, M. A., Sadnan, G. A., Shatyi, S. S., & Billah, M. (2025). Blockchain-based green edge computing: optimizing energy efficiency with decentralized AI frameworks. *Journal of Computer Science and Technology Studies, 7*(1), 386–408.