Facultad de Ciencias Económicas y Empresariales

THE IMPORTANCE OF ESG FACTORS ON SOVEREIGN BONDS IN THE EURO ZONE

The application of tree-based machine learning algorithms to predict the spread of sovereign bonds in the Euro Zone countries using ESG performance measures.

Autor: Elena Medina Lazcano
Tutor: María Coronado Vaca

MADRID | June 2024

ABSTRACT

This paper investigates the impact of Environmental, Social, and Governance (ESG) performance, measured by the Sustainable Development Goals (SDG) Scores, on the spread of government bonds, focusing on the founding countries of the European Economic and Monetary Union during the post-European debt crisis period. Leveraging both random forest and boosting algorithms, the study explores the problem from two perspectives: as a regression task to predict the bond spread and as a classification task to forecast the direction of the spread over time. The findings indicate that random forests outperform boosting models in regression tasks, with SDG Scores identified as the most influential variable. Conversely, in classification tasks, boosting algorithms surpass random forests, which suffer from significant overfitting, and factors such as inflation, real effective exchange rate, and current account ratio to gross domestic product (CA/GDP) emerge as the primary determinants of spreads. The research underscores the relevance of ESG performance in predicting government bond spreads and highlights the need for unified ESG performance measures to enhance the consistency and comparability of results across studies.

Keywords: ESG performance, government bonds, random forest, boosting algorithm, European Economic and Monetary Union, sustainable investment, machine learning, spread prediction.

RESUMEN

El presente trabajo investiga el impacto que tiene el desempeño de un país en aspectos medioambientales, sociales y de gobernanza (ESG) en el diferencial de sus bonos de Estado. El nivel de rendimiento ESG de un país se mide según la puntuación que el mismo ha recibido según su nivel de cumplimiento de los Objetivos de Desarrollo Sostenible (ODS) de la ONU. Además, el presente trabajo se centra en el estudio de los países fundadores de la Unión Económica y Monetaria Europea durante el período posterior a la crisis económica de 2009. En este trabajo se emplean dos algoritmos de *machine learning*: *random forest* y *boosting*; para resolver un problema de predicción desde dos perspectivas distintas: un problema de regresión que busca predecir el diferencial del bono y un problema de clasificación que busca predecir la dirección del mismo con el paso del tiempo. Los resultados indican que el algoritmo de *random forest* es superior al de *boosting* en el problema de regresión, identificando las Puntuaciones ODS como la variable más importante. Por otro lado, cuando se plantea un problema de clasificación, los algoritmos *boosting* muestran una mejor capacidad de predicción, al tener el *random forest* una clara tendencia al sobreajuste. En este caso, factores como la inflación, el tipo de cambio efectivo real, y la ratio cuenta corriente/PIB se muestran como los principales determinantes del resultado. El presente trabajo trata la importancia de considerar el desempeño ESG al predecir la evolución de los bonos del estado y subraya la necesidad de una medida unificada de dicho desempeño para asegurar la coherencia y comparabilidad de los resultados obtenidos en distintos los distintos estudios realizados en este ámbito.

Palabras clave: Rendimiento ESG, bonos del estado, *random forest*, algoritmos *boosting*, Unión Económica y Monetaria Europea, inversión sostenible, *machine learning*, predicción del diferencial.

List of Acronyms

UNGC	United Nations Global Compact
UNEP FI	United Nations Environment Programme Finance Initiative
PRI	Principles for Responsible Investment
ESG	Environmental, Social, Governance
EU	European Union
EMU	Economic and Monetary Union
ECB	European Central Bank
UN	United Nations
GDP	Gross Domestic Product
REER	Real Effective Exchange Rate
SDSN	UN Sustainable Development Solutions Network
SDG	Sustainable Development Goals
ANN	Artificial neural network
RF	Random Forest
RMSE	Root Mean Squared Error
ME	Mean Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
CA	Current Account

Table of contents

Table of illustrations

1. INTRODUCTION

1.1. Research objectives

This paper delves into the possibility of applying different machine learning models to determine the effects of a country's ESG performance in the spread of its government bonds. In order to achieve that goal and evaluate the importance of ESG performance, it looks to predict the spread of government bonds, implementing non-linear machine learning algorithms, particularly decision tree-based ensembles, which allow for some level of interpretability of the underlying relationships among the variables. Furthermore, the present research seeks to determine a proxy measure for the ESG performance of countries, given the lack of standardized criteria for evaluating the ESG practices of entities – encompassing both corporations and sovereign states.

1.2. Justification of the subject matter of study

In 2006, the United Nations Global Compact (UNGC) and the United Nations Environment Programme Finance Initiative (UNEP FI) launched a new project: the Principles for Responsible Investment (PRI), which aims to "promote the consideration of environmental, social and governance (ESG) issues by institutional investors" (Gond et al., 2012). Since then, many investors have joined the PRI, which are currently signed by a total of 5,391 institutions (Principles for Responsible Investment, 2023), among which we can find both investors and service providers.

The PRI is only one of many initiatives that are currently driving an increasing interest in ESG integration in financial decision making. ESG refers to different environmental, social and governance factors used to measure how sustainable an organization is. Among other things, these factors are increasingly being used as criteria to make financial decisions, formally referred to as sustainable investing, or ESG investing. However, these ESG investing strategies (negative screening, positive screening, best-in-class, among others (Chen, 2023)) were initially considered to be preventing investors from maximizing their financial returns, which were sacrificed in favor of ethical values. This initial position has changed in recent times, as portfolios constructed according to ESG criteria have been shown to outperform their traditional counterparts and even endure periods of economic crisis rendering positive results (Friede et al., 2015; Crifo et al., 2017).

The popularity of ESG integration, however, has had a more important role in stock markets and even corporate bonds, being just recently introduced in the realm of sovereigns. As explained by the Principles for Responsible Investment (2019), "because sovereign debt was considered traditionally a risk-free asset class, there has been a tendency to underestimate the importance of ESG integration relative to other fixed income classes".

In spite of this, investors in sovereign bonds are adapting their ESG integration frameworks to the particular characteristics of this type of assets, as there is a "more complex transmission mechanism of the sovereign debt asset class to sustainable outcomes in the real economy" (Gratcheva et al., 2020).

Moreover, there are countries that have put a pressing focus on ESG concerns, which has had an impact both in governmental policies and corporations. It is the case of the European Union (EU), which has developed a great normative reform to promote sustainable policies and ESG requirements for corporations in all Member States during the last years (Chen, 2023). The most recent changes have been the regulation that requires all corporations to disclose sustainability information as of 2024 (Directive EU 2022/2464), and the agreement of the European Union Council to start negotiations in order to regulate ESG ratings (Store, 2023). The importance given to ESG concerns in Europe has manifested in the fact that, up to 2016, Europe presented the higher proportion of sustainable investing (Serafeim, 2020), and that ESG integrated indexes "generally show higher returns, especially in Europe" (Chen, 2023).

Within the European Union, Member States created in 1999 the Economic and Monetary Union (EMU or Eurozone), as an initiative to pursue greater economic cooperation (European Comission, 2024). Hence, they not only share a currency – the euro – but also economic policies under the guidance of the European Central Bank (ECB). Despite their differences, the close economic relationship between these countries, including Spain, makes them an ideal geographical area for studying the effects of ESG performance on their sovereign bonds.

Furthermore, following the sovereign debt economic crisis in 2009, there has been research that supports the idea that there are factors, beyond traditional fiscal and macroeconomic ones, that need to be considered when addressing the yield of bonds (Capelle-Blancard et al., 2016). The inclusion of ESG factors, among others such as the monetary policies applied by the ECB, in the analysis of eurozone bonds will help identify additional extra-financial factors that should be taken into account when looking to explain recent economic events (Pineau et al., 2022).

Consequently, the study of sovereign bonds in the Eurozone, which includes several developed economies, and how these are impacted by different ESG factors can provide insights that not only investors use in their decision making, but which can also be used as a guide by other countries searching to increase their ESG performance and improve their creditworthiness.

Given the increasing importance of ESG matters in the EU area, investors have shown concern on how different ESG strategies implemented by both corporations and governments will affect the market. Looking to integrate these concerns, I study herein the importance that ESG performance has on determining the credit spread of sovereign bonds in the Euro Zone.

1.3. Methodology

The methodology conducted during this research has three main phases. Firstly, I carried out a review of the existing literature related to the subject of this paper – covering literature on both ESG investing, focusing on the review of that related to measuring ESG performance, fixed income, and government bonds; and on the implementation of machine learning algorithms to predict the performance of financial assets.

Second, I conducted an empirical analysis to determine the importance of ESG factors in predicting the spread of a sovereign bond in the eurozone. The study will analyze the years after the 2009 crisis, using macroeconomic and financial variables of the eurozone countries as well as the SDG scores granted since 2000 by the United Nations (UN) for each country as a measure of their ESG performance. This analysis will be carried out using tree-based ensembles: random forest and boosting model, for predicting bond spreads and determining the importance of the ESG variable in them. Lastly, after explaining the results obtained, the conclusions reached throughout the research will be detailed at the end of this work.

1.4. Structure

The structure of this paper is as follows. First, we will review the existing literature on the matter of ESG integration, focusing on sovereign bonds; and different machine learning techniques applied in financial decision-making. Secondly, we will explain the empirical analysis conducted through this research, focusing on the data used, the algorithms applied, and analyzing the results obtained. Lastly, we will summarize the research conducted and the conclusions reached.

2. LITERATURE REVIEW

The increasing interest in ESG and sustainable investing has led scholars and institutions to undertake research in order to determine the role that ESG can play in financial returns. As a result, and against what was thought initially, several studies have shown that ESG investing leads to higher returns, and that portfolio construction on ESG criteria may redeem better financial performance, even in moments of economic stress (Chen, 2023).

Along with the conclusions reached, which showed that an ESG driven portfolio may, in fact, yield higher returns than its non-ESG counterparts, financial experts have also looked to explain the effect that ESG has on financial assets. Up to 2012, most of the research had been done on the financial performance of funds. Although this has changed in the later years, it is safe to say that, currently, most of the research conducted around ESG investing has been based in determining the effect that the different ESG factors have on the stock market (Slimane et al., 2019). The rationale behind the extensive literature on ESG and stock markets (i.e. (Serafeim, 2020)) accompanies the interest of ESG investors, who "pursue two different goals when they consider equities and bonds. They invest in stocks with good ESG ratings in order to avoid extra-financial long-term risks, whereas they consider that fixed income is the field of impact investing. This explains the high demand for green and social bonds, and this also explains why ESG screening is less widely implemented in fixed income markets than in equity markets" (Slimane et al., 2019).

However, this doesn't mean that ESG integration in the fixed income market has not been studied. Apart from the abovementioned green and social bonds, studies on ESG integration in bonds originated on the possibility that ESG could help determine the spread of securities. Again, in this case research has unfolded in two different ways: one, studying corporate bonds and how ESG can affect the cost of capital of corporations (Adriaens et al., 2023); and the other, the effect of ESG on sovereign debt (i.e. Pineau et al., 2022). And again, research has been more focused on the former, rather than the latter.

Consequently, literature on ESG investing relating solely to sovereign bonds is relatively scarce compared to the research that has been conducted on sustainable investing in relation to other corporate financial assets. In spite of this reality, "there is a growing case and appetite for structured ESG integration, since traditional sovereign credit risk analysis appears to inadequately reflect emerging pressures" (Principles for Responsible Investment, 2019). While ESG frameworks have been developed in order to invest in corporate assets, whether stocks or bonds, there are no such frameworks relating to sovereign bonds, which has triggered an increase in research and framework design done in-house by institutional investors.

In line with the explanation above, an increased interest in ESG investing in sovereigns has been accompanied by research on the possible relationship between ESG factors and sovereign credit rates, on one hand; and ESG factors and credit ratings, on the other.

However, as the issuers of sovereign debt are not companies but governments, and thus, they present different characteristics; ESG integration on this type of fixed-income instrument is different from what initially has been studied on stocks, bonds, and portfolios. An example of such differences can be deduced from the importance of the Social pillar of ESG and the Environmental pillar, which is lower to that of the Governance pillar. As explained by Crifo, et al. (2017), political factors, which would be included in the latter, have been studied in greater detail than social or environmental factors, and have been considered to have a greater effect on the risk of default of a country. Thus, the importance and the role that different ESG factors play may change when considering sovereigns.

When studying the relationship between ESG and credit ratings, Pineau et al. (2022) assess quantitatively the importance of ESG and non-ESG factors in determining a country's creditworthiness, that is, its sovereign credit rating. Their findings confirm that ESG performance has a significant impact on sovereign ratings.

Moreover, when considering ESG performance of different countries, we have to differentiate between regions. The main reason resides in the fact that ESG pillars may have varying importance based on the reality of a specific country. In the corporate universe, the importance of ESG factors mainly varies depending on the industrial sector that the company participates in, "because different environmental, social and governance risks are identified as material in different industries" (Adriaens et al., 2023). Although countries cannot be classified by sector, the main activities that drive the economy do affect the importance that ESG factors can have on sovereigns, as well as other characteristics such as demographics. Thus, studies relating to this field tend to focus on a region, grouping similar countries together or dividing their data based on geographical criteria in order to better understand the results. For example, even though the "E" factors tend to have less effect at the sovereign level compared to the corporate level – as they are considered to be less material to determine ESG scores (Gratcheva et al., 2020) -, they were proven to be the most important in determining the credit rating of sovereigns in the Sub-Saharan African region (Estran et al., 2022). This is explained by the fact that these countries have an agricultural economy, and thus environmental policies have a greater effect on the economic well-being of the nations included in such region.

Due to this reality, when studying the effect of ESG on sovereign bond characteristics, research involving different bonds should also distinguish between regions, being the main distinction done through research between developing and developed economies, as this usually entails clear differences in creditworthiness (Adriaens et al., 2023; Principles for Responsible Investment). Not only have studies on sovereign bonds used this distinction, but many financial research papers tend to focus on the American and the European markets as examples of developed economies (Chen, 2023; Viviers et al., 2012).

The studies focused on European countries also differentiate between European Union Member States and non-EU countries; and even in the EU, there is a differentiation of the EMU countries. Especially after the sovereign debt crisis, research focused on trying to establish the factors that determined the spread of bonds (Capelle-Blancard et al., 2016). In their study, Capelle-Blancard, et al., show that the particular characteristics of these countries had different effects than expected in the face of the crisis: "[t]he fundamentals one would expect to be the most important determinants for the price of sovereign risk – public debt, fiscal deficit and current account – actually do explain very little of the pricing of risk in European countries before the crisis, but have much more explanatory power for sovereign risk in other advanced economies. […] Other factors than fundamentals may have been key determinants of sovereign debt in Europe" (2016). The inclusion of ESG factors, among others, in the analysis of Eurozone bonds is considered an advance in the literature that can help determine which extra-financial factors explain the economic events of the last decade (Pineau et al., 2022).

Additionally, and due to the recent interest in sustainable and ESG investing, all throughout the existing literature, there is one prevalent idea: the lack of standardized criteria to measure ESG performance (Beckaert et al., 2023; Buss, 2021 and PRI, 2019; among others). This has led to the creation of many different rating systems and ESG integration frameworks, mainly developed by service providers and rating agencies; using numerous and diverse factors to construct their ESG scores. This has resulted in not only diverse methodologies that are being implemented, but also substantially different results (Beckaert et al., 2023). Moreover, there is additional complexity in the construction of these ESG scores in the fact that, the evaluation of ESG performance is different for companies than for countries (Gratcheva et al., 2020). Consequently, different ESG measures and methodologies have been used throughout academic and institutional research; for example, Beckaert et al., suggested, when studying bonds, the use of the SDG Index scores constructed by the United Nations, in order to address the existing gap in relation to ESG scores.

Having reviewed the research conducted on ESG investing, particularly in sovereign bonds, it is also necessary to understand the state of the art in the various methodologies and models used by researchers in their work (Jabeur et al., 2019; Belly et al., 2023; Yuankang et al., 2019).

Focusing on studies in sovereign debt and ESG, the models used in quantitative research are regression and other statistical models. However, there has been increasing research on the financial applications of machine learning since 2015 (Nazareth et al., 2023). When comparing the different types of models that could be implemented, the initial premise of all investigations is Wolppert's "no free lunch theorem", which states that "it is impossible to know in advance which machine-learning model is best suited to a particular dataset" (Wolppert, 1996). Taking this into consideration, it is true that some machine learning models are considered black boxes (Chen, 2021), and although they show greater performance in most scenarios compared to other statistical methods (Nazareth et al., 2023), their lack of interpretability requires researchers to decide the models implemented not only based on prediction and performance criteria, but also on the capability of the model to convey information on the underlying correlations of the variables, in Chen's words: "the choice between conventional linear methods and machine learning alternatives hinges on this balance between accuracy and interpretability" (2021).

Although regression models have been the ones used traditionally, in conjunction with stepwise methods, which reduce the dimensionality of problems; other algorithms, such as KNN, random forests, or neural networks; have been proven to increase the accuracy of predictions. Zhang et al. (2020), for example, used neural networks for portfolio optimization, looking to optimize the Sharpe Ratio. On the other hand, Sardosky applied random forests and other tree-based algorithms as a prediction tool for stock prices (2021a) and the price direction of commodities (2021b). Moreover, Belly et al. (2023) applied different machine learning algorithms to forecast sovereign risk in the Euro area, showing superior results of boosting ensembles in performing this task.

This way, the use of machine learning to make financial decisions has increased during the last years. Apart from portfolio optimization, it has been used to predict stock prices, as well as commodities' prices. In particular, Sadorsky (2021a, 2021b) has shown on various occasions the utility of decision trees over other, simpler algorithms to predict prices and the expected direction of prices. Another example is the use of the Fast Large-scale Almost Matching Exactly (FLAME) algorithm to determine the effect of ESG factors on credit ratings (Adriaens et al., 2023). All these, among others, serve as examples of the wide implementation of machine learning and artificial intelligence models in the financial field.

3. EMPIRICAL ANALYSIS

3.1. Data

In this section I will describe the data used in this research, both dependent and independent variables. As explained before, the scope of this investigation is limited to the period between 2010 and 2022 and focuses on the founding countries of the European Economic and Monetary Union. In particular, research has been conducted on data from the following countries: Austria, Belgium, Spain, Finland, France, Greece, Ireland, Italy, the Netherlands, and Portugal[1]. The decision to focus on these countries was made under the rationale that other EMU countries don't provide data consistently (Government 10-year yield values and other measures, such as GDP or inflation rate, were not available in some cases), which would critically affect the results. Additionally, these are the countries on which previous research on the Euro Zone was based (Belly et al., 2023; Afonso et al., 2015; Eijiffinger et al., 2023; Gómez-Puig et al., 2014).

I have decided to focus my research on the years after the sovereign debt crisis (2009) for two main reasons. On one hand, even though the spread of sovereigns was traditionally explained by macroeconomic circumstances, especially the fiscal position of a country;

[1] Luxembourg is excluded from this research due to its low level of debt outstanding, a criteria used in other research (Gómez-Puig et al., 2014) and Greece is included because, despite not being a founder, it became a part of the EMU in 2001 and adopted the Euro at the same time as the rest in 2002.

after the crisis, the ECB's monetary policy and other factors were starting to be considered in research, as bond spread tendencies were changing due to unconventional tools used by the ECB (Eijiffinger et al., 2023). Eijiffinger et al. showed in their work that there is great difficulty in using a general model for all years since 1999, concluding that better predictive results are obtained when a differentiation is made between the pre-crisis and post-crisis periods. On the other hand, ESG concerns are a rather recent matter. On one hand, the PRI, which is considered by many the starting point of ESG investing, was a recent initiative when the sovereign debt crisis started; in 2013 it had little over a thousand signatories (Caplan et al., 2013), while currently over 5,000 have signed it (Principles for Responsible Investment, 2023) and, on the other, the SDGs were created in 2015. Taking these circumstances into consideration, the timeline of this research can be considered reasonable in order to try and determine whether ESG performance of countries has any impact on their bonds' spread.

The frequency of the data is quarterly due to the fact that many of the variables are published either annually – like SDG scores – and/or quarterly – like GDP. To avoid errors due to missing data (Eijiffinger et al., 2023), I have opted for this frequency, assuming that countries have the same SDG score for all quarters of the year, and gathering bond spread data quarterly available on FactSet.

3.1.1. Dependent variable: the spread of government bonds

The "measure of a country's borrowing cost in international capital markets is its yield spread, which is defined as a market's measure of a country's risk of default" (Crifo et al., 2017). As explained before, some believe that this variable – a country's risk of default – may be diminished when the country presents a positive ESG performance, showing a negative relationship between the two variables (Capelle-Blancard et al., 2016).

This paper aims to predict the spread of sovereigns using, among other variables, the ESG performance of the country, to determine the importance that the latter has on the former. Consequently, the dependent variable of this study is the spread of government bonds, which is calculated as the difference between the 10-year maturity government yield of each country and the yield of the same maturity of Germany.

The 10-year maturity is a benchmark maturity, which provides good quality data and allows for a country's debt dynamics to show the effects of the macroeconomic circumstances. Additionally, "Germany is considered the benchmark issuer by the majority of market participants and its bonds are therefore the best proxy for the riskless rate among the Eurozone countries" (Eijiffinger et al., 2023).

3.1.2. Independent variables

In order to conduct the study, I have gathered different variables to consider and that, as research shows, they are good predictors of government bond spread. Among these macroeconomic and financial variables detailed below, I have included an ESG variable that serves as a proxy for ESG performance.

3.1.2.1. Macroeconomic and financial variables

According to Crifo et al. (2017), there are three types of risks that affect bond spreads, and that need to be considered when trying to predict them: credit risk, determined by the fiscal and macroeconomic position; liquidity risk, determined by the size of the bond market; and international risk aversion, which refers mainly to investors' sentiment towards these assets. Following these criteria, I have gathered a set of 8 variables, named below and described in Table 1 in the Annex.

On one hand, the macroeconomic situation of a country will be described by the following variables: first, the gross domestic product (GDP), which indicates the situation of the corresponding economy and its output production (Castellani et al., 2006), and gives us a sense of the country's ability to comply with its financial obligations (Crifo et al., 2017); second, the unemployment rate (Unemployment), which can be considered a proxy to capture the country's growth potential (Gomez Puig et al., 2014 & Eijiffinger et al., 2023); and third, the Consumer Price Index (Inflation), that measures the inflation rate and "reveals sustainable monetary rate policies" (Crifo et al., 2017).

On the other hand, when considering the liquidity risk of sovereigns, and the credit worthiness of a country, two variables have been included: the debt to GDP ratio and Credit Ratings. The depth of the debt market in a country, and its default risk, is measured through the debt-to-GDP ratio (Debt_GDP), as a higher level of debt is expected to increase default risk (Capelle-Blancard et al., 2016 & Escaleira Marques da Fonte, 2021), which would in turn mean an increase in the spread of bonds. In this line of thought, credit ratings (Ratings) also provide a measure of financial distress, as they are constructed by rating agencies in order to inform investors of the existing default risk, in this case, of a country (Afonso et al., 2015; Capelle-Blancard et al., 2016 & Jabeur et al., 2019; among others).

Lastly, in order to reflect international risk aversion, we will use both the real effective exchange rate (REER) (Eijiffinger et al., 2023) and the current account to GDP ratio (CA_GDP), which is "used as a proxy of the foreign debt and the net position of the country vis à vis the rest of the world" (Gómez Puig et al., 2014).

The variables enumerated above are the main variables used in literature in relation to the yield of government bonds, as they serve as a proxy for the economic and financial conditions a country.

3.1.2.2. SDG score: a measure of ESG performance

The main focus hereof is to study the impact of a country's ESG performance on its bond spreads. ESG integration has been one of the most researched topics, even though there is not a standard measure of ESG performance yet. There are many institutions (rating agencies, data providers, financial institutions or service providers) that have constructed their own ESG scores and criteria in order to somehow classify companies and/or countries to make financial decisions. In research, there are different methodologies applied in order to determine ESG performance: some use already constructed ESG scores or indexes – such as Vigeo (Crifo et al., 2017), Refinitiv (Escaleira Marques da Fonte, 2021), or the MSCI ESG momentum scores (Beckaert et al. 2023); and others construct their own scores and indices based primarily on public data (Pineau et al., 2022; & Capelle-Blancard, 2016).

Among these efforts, the UN Sustainable Development Solutions Network (SDSN), with other public institutions, has been presenting since 2019 a *Sustainable Development Report*, in which they assess the performance of all European countries towards the SDGs. The Sustainable Development Goals (SDG) are part of the UN 2030 Agenda for Sustainable Development, signed in 2015 (World Health Organization, 2024). They "are designed to end poverty, hunger, AIDS and discrimination" (United Nations Development Program, 2024) and require environmental, social and economic sustainability. Consequently, SDG goals are only reached through practices that can be considered ESG-friendly. The level of completion of these goals is measured by the United Nations using an SDG Score, used to rank countries.

The present research is limited to countries, and as such, it considers the close correlation existing between ESG performance of countries and SDGs (Schieler, 2017): as many of the indicators used to construct the SDG Index can be classified as determinants of the environmental, social or governance factor in ESG. For example, CO_2 and Nitrogen emissions and access to renewable energy or natural resources like water can be classified as environmental factors (E); the PISA score – related to education -, unemployment rate, and gender wage gap can be included as social factors (S); and the Corruption Perception Index, timeliness of administrative procedures and lawful and adequately compensated expropriations can all be included in the governance factor (G)[2] (Principles for Responsible Investment, 2019).

All this considered, the ESG component of the present investigation is obtained (as described in Table 1 in the Annex) from the official SDG Index scores that each of the countries analyzed has received. This will not only be used to measure their performance on ESG matters, but it will also provide, in a field where there is no consensus on how to quantify ESG performance, an official measure provided by international authorities[3].

3.2. Methodology and algorithms

Having explained the data that will be used herein, the main objective of this paper is to determine the role of ESG performance of countries in the spread of bonds – that is, the impact on the creditworthiness of the country and, thus, its cost of financing.

Research on government bonds mainly focuses on predicting their spread, which would require applying a regression algorithm (Castellani et al., 2006). However, other studies that apply machine learning to prediction problems involving financial assets, look to study the direction of prices (Sadorsky, 2021b) and spreads (Yuankang et al., 2019) – an alternative that will be considered later on.

Despite the different possibilities, in this paper I have decided to follow the general tendency in research and look to predict the spread of government bonds in the Euro Zone. One of the main focuses in the field of economics is the prediction of values, and the most popular method "is linear regression via the ordinary least square (OLS) method" (Chen, 2021). However, and given the greater performance of machine learning (Sadorsky, 2021a; Nazareth et al., 2023) and its capability to detect nonlinear relationships between

[2] The specific examples provided are based on the indicators used to construct the SDG Index detailed by the SDSN in their Sustainable Development Report of 2023 (Sustainable Development Solutions Network, 2024)

[3] It is important to note that using SDG scores as a proxy for ESG scores would not be possible when analyzing corporate performance, as SDGs refer to Goals to be achieved at an estate level. There has been however, research on the relationship of corporate ESG scores and how they impact and contribute to the completion of the different SDGs of the countries where they conduct their activity (Beckaert et al., 2023).

variables (Belly et al., 2023), the research conducted herein focuses on the use of machine learning algorithms in order to complete this prediction task.

There are different machine learning algorithms that can be used to solve the present problem, including both classical models such as KNN, decision trees or random forests; and other deep learning algorithms, among which artificial neural network (ANN) can be considered (Nazareth et al., 2023). The latter have been experimenting increased popularity due to their improved performance and ability to imitate the human brain. However, "the primary limitation of machine learning is its lack of interpretative clarity" (Chen, 2021). Thus, we can see that more complex algorithms, such as ANN, are considered black boxes, which entails sacrificing any interpretation of the relationship, influence, and impact that each variable has on the final result.

Since the research conducted looks to analyze the importance of ESG performance in determining the spread of government bonds, deep learning models cannot be considered optimal to reach this goal. Instead, I turn to tree-based models, which provide a more interpretable result (Yuankang et al., 2019), although this interpretability is limited in comparison to regression models (Chen, 2021).

A decision tree can be defined as a decision support tool that uses a tree-like model of decisions and their possible consequences (Yuankang et al., 2019), the basic training principle of decision trees is "the recursive partitioning of the feature space using a tree structure where each child node is split until pure nodes are achieved" (Basaka et al., 2019). However, the use of decision trees tends to be insufficient, as these algorithms exhibit a tendency of overfitting (Basaka et al., 2019) – that is, they adjust almost perfectly to the data and the noise, making nearly perfect predictions in training datasets, but showing poor generalization capabilities when new data is introduced.

Nevertheless, this overfitting tendency can be avoided when implementing ensemble models. There are several types of tree-based ensemble algorithms that can be used to solve prediction problems; and that avoid the overfitting derived from using single decision trees (Belly et al., 2023). Taking this into consideration, the present research has been conducted through two of these ensemble models: random forests and boosting.

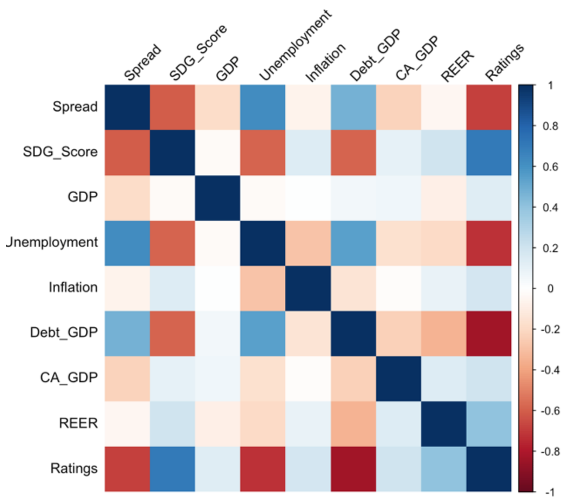
Having explained the different models that will be applied to the dataset, I have first constructed a descriptive model – based on a logarithmic regression – in order to have a better understanding of the relationship between the variables studied (Capelle-Blancard et al., 2016). Secondly, I partition the dataset into training and test subsets, in order to examine the prediction performance of the models. Lastly, I have trained the models described below to conduct the prediction task set in this paper and analyzed their performance – based on RMSE criteria – and the importance that they attribute to SDG Scores.

All the models described below have been constructed on R, using the caret and RandomForest packages and caret, ranger and gbm libraries.

3.2.1. Descriptive analysis

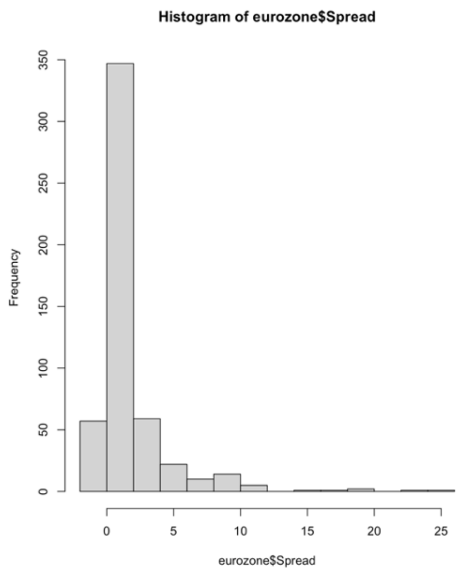
In order to have a better understanding of the variables in the dataset, I have conducted a statistical analysis. This, although it may not be as accurate in prediction tasks, it will enable a preliminary approach on the relationship between the variables being analyzed (Capelle-Blancard et al., 2016). This is particularly important when considering the lack of interpretability that more complex algorithms present (Castellani et al., 2006). In the case of tree-based models, used herein, it is important to consider that, despite the information they provide on variable importance, they do not indicate the degree and direction of correlation between variables (Chen, 2021).

First, a correlation matrix was constructed. As shown in Graph 1, SDG Scores, unemployment rates and Credit Ratings present high correlations (greater than 0.5) with the target variable. This could lead to collinearity in the model, which is avoided by centering the variables – that would also solve the right skewness that the distribution shows (Graph 2).



Graph 1: Correlation matrix (Source: own elaboration)

Upon normalizing the variables, the model presents a low Variance Inflation Factor, which indicates that there is no significant collinearity in the model (Akiwande et al., 2015). However, further research could be done on the relationship between SDG Scores and Credit Ratings, as they present a very high correlation. This could be explained by the fact that Credit Ratings partly convey information similar to the ESG metrics used to determine SDG Scores, so “ESG scores by their nature should measure aspects which other metrics already reflect” (Gratcheva et al., 2020). Although there is evidence that ESG performance has an impact on credit ratings, research showing causation between the two is mixed, so we decide to not eliminate credit ratings as a variable in this study (Gratcheva et al., 2022). Additionally, Graph 1 above shows a negative and strong correlation between spreads and ESG Scores, meaning that an improvement in ESG performance could possibly imply a decrease in the country’s credit spread, and thus a lower cost of debt and a lower risk of default.



Graph 2: Variable Distribution (Source: own elaboration)

3.2.2. Data partition and hyperparameter tuning

In order to apply the model to the dataset and test its prediction performance, the dataset shall be partitioned into training and test subsets (Sadorsky, 2021b). Different criteria can be used to make this partition. On one hand, the first decision is whether to maintain the chronological order of the data, or to shuffle all data available and make a random partition, ignoring chronological considerations (Sadorsky, 2021b). Research on panel data in the finance field has been done using both chronological (Belly et al., 2023) and out-of-order data (Olson et al., 2021), and results have shown no significant differences regarding prediction made by tree-based models.

On the other hand, some studies analyze the performance of models in specific countries, in order to determine whether machine learning methods show better performance on countries with particular characteristics. Although we could consider partitioning data on a geographical criterion, there is not enough evidence that suggests any improvement on model performance (Belly et al., 2023), so I decide to disregard this idea. Taking all this into consideration, I have randomly partitioned the data in the following manner: 70% of the data into the training subset, and 30% of the data into the test subset.

Moreover, when using bagging and boosting ensembles there are several hyperparameters that need to be tuned in order to establish the criteria under which the model will work. This tuning process is conducted using the repeated cross-validation method (Chen, 2021)

– with 10 folds, and 3 repetitions. Alongside cross-validation, grid search is used in order to conduct hyperparameter tuning (Belly et al., 2023; Chen, 2021). There are a great number of hyperparameters that can be tuned in these types of algorithms. However, if we introduce “too much complexity in the training phase it could lead to overfitting” (Nuta et al., 2021). Thus, there will be some hyperparameters fixed, while others will be tuned through grid search method.

3.2.3. Random Forest

The first ensemble model applied to the data set is a random forest (RF), which is a bagging method used for both regression and classification problems, based on the principle that “weak learners combined can form a strong predictor model” (Belly et al., 2023). Bagging (which comes from Bootstrap Aggregating) enables a decrease in the variance of the prediction (Yuankang et al., 2019) through “non-sequential learning that draws, with replacement, a random subset of data from the training dataset. These draws are not correlated in any way, but displays the same distribution” (Nuta et al., 2021).

The random forest applied in this research has 500 trees and determines variable importance and decisions based on impurity measures. In this model, there are 3 hyperparameters being tuned through cross-validation: minimum size of nodes, number of variables used in each split, and the rule to make splits.

3.2.4. Boosting model

On the other hand, and in order to explore the possibility that boosting models present greater generalization capabilities than random forests (Yuankang et al., 2019), I have conducted further research using a boosting ensemble of decision trees. Boosting models consist of “embedding many weak learners into one efficient regression or classification algorithm” (Nuta et al., 2021). In this case, again, the weak learner used in this ensemble is the decision tree, as it has been proven to be more efficient than other models (Belly et al., 2023).

In this model, there are four hyperparameters determined: the maximum depth of the tree – considering values of 1, 5 and 9 trees -, the number of trees – from 50 to 1500 in steps of 50, same as before -, the shrinkage – that is, the contribution of each tree to the final prediction, fixed at 0.1 – and the minimum number of observations required in each node – fixed at 20 -. From these, the first two will be tuned through a grid search method, as mentioned before.

3.3. Results

Having explained the methodology implemented to predict the spread of government bonds, I will analyze below the results obtained.

3.3.1. Descriptive analysis

Prior to applying the described models, I have conducted a descriptive analysis to have a previous understanding of the existing relationships among the variables. To do so, I have used a regression model, which can indicate the relationship of the different predictors with the dependent variable. This linear model provides information on the sign of the coefficients and correlation of the variables that will not be conveyed in other nonlinear models – such as the ensembles implemented (Chen, 2021).

```
Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.791e+00  5.115e-01   3.501 0.000504 ***
SDG_Score    -3.935e-02  4.374e-03  -8.996  < 2e-16 ***
GDP           3.719e-06  2.379e-05   0.156 0.875848
Unemployment  2.643e-02  4.485e-03   5.894 6.87e-09 ***
Inflation     1.277e-02  7.126e-03   1.791 0.073812 .
Debt_GDP     -4.046e-03  8.615e-04  -4.696 3.42e-06 ***
CA_GDP       -1.382e-02  2.397e-03  -5.765 1.41e-08 ***
REER          3.436e-02  4.181e-03   8.219 1.70e-15 ***
Ratings      -2.070e-02  1.966e-03 -10.530  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3574 on 511 degrees of freedom
Multiple R-squared:  0.7423,    Adjusted R-squared:  0.7383
F-statistic:   184 on 8 and 511 DF,  p-value: < 2.2e-16
```

Graph 3: Statistical Results (Source: own elaboration)

The regression model (Graph 3) shows that all variables, except for GDP and inflation rate, have a significant relationship to the Spread. Particularly, and taking into consideration the main object of the present paper, results show that SDG Scores have a significant negative relationship with the spread of government bonds and a coefficient of 0.0392. This means that an increase of 1 point in the SDG Score of the country causes a 0.0392 decrease on the spread of its 10-year maturity government bond. This would initially confirm the hypothesis that increase ESG performance of a country would decrease the spread of its bonds, which means a lower risk of default.

3.3.2. Random forest

After the tuning process, the hyperparameters of the random forest were tuned to the following optimal values: 8 variables are used to make each split (parameter *mtry*), the rule taken into consideration in order to make a split is “extratrees”, which means that splits are made at random, rather than based on the optimal split, and the minimum node size is 5 observations. The criteria used to determine the hyperparameter was the root mean squared error (RMSE), so that the model was determined to that which presented a lower value of RMSE.

Having tuned the hyperparameters, the results of the model, conveyed in Graph 4, show that the most important variable in determining the spread of government bonds is SDG scores, followed by credit ratings and the unemployment rate.

The prediction capacity of the model can be analyzed through different error measures, detailed in Table 1.

```
              Overall
SDG_Score     100.000
Ratings        85.098
Unemployment   20.395
Debt_GDP       17.383
Inflation      10.577
GDP             8.417
REER            5.768
CA_GDP          0.000
```

Graph 4: Variable Importance in regression according to Random Forest. (Source: own elaboration).

Measure	ME	RMSE	MAE	MAPE
Training	0.0041	0.5846	0.277	7.046
Test	-0.077	0.6775	0.4854	-18.404

Table 1: Performance of Random Forest (Source: own elaboration)

When considering the performance of the model, as described in Table 1, the RMSE shows that the model makes more errors in the test dataset than in the training. This would be considered normal, as the model tends to be more accurate in the subset that it has been trained on. Although this could result in the small difference (0.9) shown above, it is true that a higher difference among errors would be a sign of overfitting, which would require further variable analysis and increase of observations to correct it.

3.3.3. Boosting model

On the other hand, due to this risk of overfitting that exists on random forests when applied to panel data (Yang et al., 2024); I decided to additionally implement a boosting model under the hypothesis that it would present a lower tendency to overfit. In this case, after the hyperparameter tuning, the hyperparameters used by the model are: 450 trees, with a depth of 5, a learning rate of 0.1 (shrinkage parameter) and a minimum of 20 observations per node.

The results of this model (Graph 5) show that the most important variable, again, is SDG Scores, followed by the debt-to-GDP ratio, GDP and unemployment rate of the country. This presents interesting results because, although both models coincide on the most important variable in predicting the spread of government bonds, according to the boosting model, macroeconomic variables play an important role in the prediction of spreads, reflecting a diminished influence of credit ratings.

```
        var   rel.inf
  SDG_Score 46.666632
   Debt_GDP 13.655004
        GDP 10.286817
Unemployment 10.127938
     CA_GDP  6.648966
    Ratings  5.562364
  Inflation  4.214634
       REER  2.837645
```

Graph 5: Variable importance in regression according to Boosting Model (Source: own elaboration)

If we analyze the performance of the model, we have to consider the following measures (Table 2):

Measure	ME	RMSE	MAE	MAPE
Training	-0.013	0.6887	0.3106	11.78
Test	-0.26	1.06	0.698	19.77

Table 2: Performance of Boosting Model (Source: own elaboration)

In this case, the test dataset shows a higher error than the training, again. However, the RMSE is nearly double the value of this error measure. This indicates overfitting, which means that the algorithm is adjusting not only to the information but also to the noise of the dataset. Consequently, there is a lack of generalization that causes a significant increase in error when introducing new data.

3.4. The possibility of applying a classification problem

As introduced before, there has been some research done not based on the prediction of a numeric value – which would be done through a regression model, such as the ones explained above – but focused on the direction in which the price or spread of the asset would move. In this case, the model implemented would be a classification problem, being the classes in which the observation would be classified: up – if the spread increased – or down – if it decreased. When applying this model to the problem at hand, the direction variable is calculated based on the difference in spread from the previous quarter to the next.

Using the same models as before, I have implemented decision tree-based ensembles: both random forest and boosting model, to address the new question that was posed. The analysis conducted is briefly synthesized below, to analyze whether changing the dependent variable would have any consequences on the models implemented.

Regarding the random forest, in this case the tuned hyperparameters are different: 7 variables are used when doing a split, being the split rule "gini" which is commonly used in classification problems in order to make the split that reduces the entropy of the dataset the most; and the minimum node size is 1. The most important variables to determine whether the spread increases or decreases every quarter, according to this model (Graph 6), are inflation, CA-to-GDP ratio and the real effective exchange rate. Relating to the matter object of this study, it is important to note that SDG Scores are the second least important variable, being the least important credit ratings. Additionally, the model presents an accuracy of 1 in training and 0.6 in test. This may be a clear indicator that the model is overfitting, as it adjusts perfectly to the training data – including the noise.

	Overall
Inflation	100.00
CA_GDP	62.67
REER	53.57
Unemployment	40.18
GDP	39.04
Debt_GDP	37.67
SDG_Score	26.17
Ratings	0.00

Graph 6: Variable importance in classification according to Random Forest (Source: own elaboration)

On the other hand, when implementing a boosting model, the hyperparameters are tuned as follows: the optimal number of trees is 50, whose depth is 1, the learning rate is 0.1 (shrinkage) and the minimum number of observations in each node is 20. This model reiterates the importance of inflation, real effective exchange rate and CA-to-GDP ratio (Graph 7). However, SDG Scores seem to have greater importance when determining the direction of spread than in the random forest model. Additionally, the accuracy of the model goes from 0.648 in training and 0.59 in test. The difference in accuracy measures in training and test shows that, as expected, the model shows a worse prediction capability when introducing new data. However, the difference does not seem to be a signal of overfitting. Consequently, it seems that when the task is established as a classification problem, then boosting models tend to have improved performance compared to random forests.

	var	rel.inf
Inflation	Inflation	38.814735
REER	REER	17.454053
CA_GDP	CA_GDP	14.145709
GDP	GDP	8.928725
SDG_Score	SDG_Score	8.713434
Unemployment	Unemployment	5.983757
Debt_GDP	Debt_GDP	4.691119
Ratings	Ratings	1.268467

Graph 7: Variable importance in classification according to Boosting model (Source: own elaboration)

4. CONCLUSION

In recent times, numerous initiatives such as the Principles for Responsible Investment show the increasing interest in ESG (Environmental, Social and Governance) matters in the financial field. Among other concerns, sustainable investment has been pursued by many institutions, seeking to optimize their returns while consciously investing in ESG-friendly corporations and States.

Even though there has been extensive research conducted on corporate financial assets, mainly stocks but also bonds; literature relating to sovereigns remains limited. Given the current state of literature, this work focuses on determining whether ESG factors have any influence on the spread of government bonds.

To achieve this, I employed machine learning algorithms, which have demonstrated a higher level of accuracy than traditional models. However, this better performance usually entails sacrificing any capability to interpret the relationships between variables, which is why machine learning models are commonly known as "black boxes", particularly the most complex ones – such as artificial neural networks, among other deep learning algorithms -. Therefore, the present investigation uses different decision-tree ensembles, which can be considered "grey-box" models in terms of their balance between accuracy and interpretability, as they provide some insight into the significance of each variable in determining the outcome. Consequently, both random forest and boosting models were used to predict the influence of a country's ESG performance on the spread of its bonds.

Despite the increasing research on ESG investing, there remains no standardized measure of ESG performance. Recent strategies developed in the field of sustainable investing have been accompanied by the creation of proprietary indexes, scores, and other various measures of ESG performance. This issue is not limited to corporations, as there is not an ESG performance measure in relation to countries. Consequently, a significant challenge in this research was determining how to quantify countries' ESG scores. To address this issue, I have used in the present paper the SDG Scores provided annually by the United Nations as a proxy for ESG performance of countries, as they most accurately convey information relating to the performance of each country across all ESG factors.

The application of the described models to assess the influence of such SDG Scores in determining the spread of sovereigns has been focused on sovereigns from founding countries of the European Economic and Monetary Union, throughout the post European debt crisis period. The rationale behind this decision is based on two main arguments: firstly, the European Union can be considered one of the most engaged regions globally with the 2030 Agenda – in which the SDG objectives where established – and exhibits prolific legislative activity regarding different directives and regulations in pursuit of a greater and improved sustainability and ESG performance of both European corporations and Member States. Secondly, the chosen timeline is based on existing research indicating a shift in the factors that determine spreads, as the sovereign debt crisis revealed the increasing importance of extra-financial aspects – traditionally overlooked – in analyzing the evolution of government bonds. Consequently, this research focuses on the spread of government bonds of the EMU founding members (including Greece) and seeks to determine whether their SDG scores are a determining factor.

Random forest and boosting algorithms have been applied in two different ways in the present research: as regression models and as classification models. The problem was firstly and mainly framed as a regression problem, aiming to predict spreads. Upon solving the regression model, it was concluded that random forests exhibit superior performance compared to boosting models, which show a tendency to overfitting. Additionally, SDG Scores emerged as the most important variable in predicting the spread of government bonds. Secondly, a classification problem was considered, wherein both random forest and boosting algorithms were employed to predict the direction of the spread of government bonds (either increasing or decreasing) over time. In this scenario, boosting algorithms outperformed random forests, as the latter showed extreme overfitting. Furthermore, results indicated that SDG Scores were less relevant in determining the direction of government bond's spread, being the most important variables the inflation rate, real effective exchange rate and current account to GDP ratio.

All in all, it can be concluded that ESG performance is relevant in predicting the spread of government bonds. Considering this, future research could explore the relationship between ESG performance measures and other variables, particularly credit ratings, to mitigate variable-related issues that could lead to overfitting. Additionally, all research in this field highlights the lack of unified ESG performance measures, which would allow for comparable interpretations and results, that are currently non-existent among researchers due to differing methodologies. Despite the existing shortcomings in this field of research, there is no doubt that ESG performance should be considered when making financial decisions, even in the context of sovereigns, to pursue a more sustainable investment strategy.

Statement of Use of Generative Artificial Intelligence Tools in Bachelor's Thesis

DISCLAIMER: The University considers that tools like ChatGPT or similar AI tools are highly useful in academic life, although their use remains solely the responsibility of the student, as the answers they provide may not be accurate. In this regard, their use to generate code for the Bachelor's Thesis is NOT permitted because these tools are not reliable for this task. Even if the code works, there is no guarantee that it is methodologically correct, and it is highly likely that it is not.

Hereby, I, Elena Medina Lazcano, a student of the Double Degree in Law and Business Analytics (E-3 Analytics) at Pontificia Comillas University, declare that in presenting my Bachelor's Thesis titled "The importance of ESG factors on sovereign bonds in the Euro Zone: the application of tree-based machine learning algorithms to predict the spread of sovereign bonds in the Euro Zone countries using ESG performance measures," I have used the generative artificial intelligence tool ChatGPT or similar IAG tools only in the context of the activities described below:

1. **Code Interpreter**: To conduct preliminary data analysis.
2. **Literary and Language Style Corrector**: To enhance the linguistic and stylistic quality of the text.
3. **Translator**: To translate texts from one language to another.

The information and content presented in this work are the product of my individual research and effort, except where otherwise indicated and credit has been given (I have included appropriate references in the Bachelor's Thesis and have explicitly stated the use of ChatGPT or similar tools). I am aware of the academic and ethical implications of submitting non-original work and accept the consequences of any violation of this statement.

Date: June 19, 2024

Signature: Elena Medina Lazcano

5. REFERENCES

Adriaens, P., & Li, D. (2023). Deconstructing the Impact of ESG Ratings on US Corporate Bond Credit Spreads: Evidence of Information Channeling. *Journal of Management in Engineering*, 40(1). https://doi.org/10.1061/JMENEA.MEENG-55

Afonso, A., Arghyrou, M., & Kontonikas, A. (2015). The determinants of sovereign bond yield spreads in the EMU. *ECB Working Papers*, (1781), 1-37. https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1781.en.pdf.

Akinwande, M., Dikko, H., & Samson, A. (2015). Variance Inflation Factor: As a Condition for the Inclusion of Suppressor Variable(s) in Regression Analysis. *Open Journal of Statistics*, 5 (7), 754-767. https://www.scirp.org/pdf/OJS_2015122416050944.pdf.

Basaka, S., Karb, S., Sahaa, S., Khaidema, L., & Roy, S. (2019). Predicting the direction of stock market prices using tree-based T classifiers. *The North American Journal of Economics and Finance*, 7, 552-567. https://doi.org/10.1016/j.najef.2018.06.013.

Beckaert, G., Rothenberg, R., & Noguer, M. (2023). Sustainable investment - Exploring the linkage between alpha, ESG and SDGs. *Sustainable Development*, 31 (5), 3831-3842. https://doi.org/10.1002/sd.2628

Belly, G., Boeckelmann, L., Caicedo Graciano, C., Di Iorio, A., Istrefi, K., Siakoulis, V., & Stalla-Bourdillon, A. (2023). Forescasting sovereign risk in the Euro area via machine learning. *Journal of Forescasting. Special Issue: Advances in Forecasting in Macroeconomics and Financial Markets*, 42 (3), 657-684. https://doi.org/10.1002/for.2938

Buss, A. (2021). Solving the puzzle: ESG assessment in sovereign bond portfolios. *Environmental Finance: Sustainable Investment Awards*, 1-5. https://globalfundsearch.com/wp-content/uploads/2021/06/aegonam_solving_the_puzzle_esg_assessment_sovereign_bond_portfolios-2.pdf

Capelle-Blancard, G., Crifo, P., Diaye, M., Scholtens, B., & Oueghlissi, R. (2016). Environmental, Social and Governance (ESG) Performance and Sovereign Bond

Spreads: an Empirical Analysis of OECD Countries. Available at SSRN 2874262. https://ssrn.com/abstract=2874262

Caplan, L., Griswold, J., & Jarvis, W. (2013). *From SRI to ESG: The Changing World of Responsible Investing*. Commonfund Institute. https://files.eric.ed.gov/fulltext/ED559300.pdf

Castellani, M., & Santos, E. (2006). Forecasting long-term government bond yields: an application of statistical and AI models. *Department of Economics at the School of Economics and Management (ISEG), Technical University of Lisbon*, Working Paper. https://depeco.iseg.ulisboa.pt/wp/wp042006.pdf

Chen, J. (2021). An introduction to Machine Learning for Panel Data. *International Advances in Economic Research*, 27. http://dx.doi.org/10.2139/ssrn.3717879

Chen, J.I. (2023). *The Drivers of ESG Index Outperformance: a Transatlantic Analysis of US and European Markets*. Master's thesis [Master of Science in Management Studies, Massachusetts Institute of Technology]. MIT. https://dspace.mit.edu/bitstream/handle/1721.1/151576/chen-irisc207-msms-sloan-2023-thesis-unsigned.pdf?sequence=1&isAllowed=y

Crifo, P., Diaye, M. A., & Oueghlissi, R. (2017). The effects of countries' ESG ratings on their sovereign borrowing costs. *The Quarterly Review of Economics and Finance*, 66, 13-20. https://doi.org/10.1016/j.qref.2017.04.011.

Council of the European Union (2023 December 20). *Sustainable finance: Council agrees negotiating mandate on ESG ratings* [Press release]. https://www.consilium.europa.eu/en/press/press-releases/2023/12/20/sustainable-finance-council-agrees-negotiating-mandate-on-esg-ratings/

Directive (EU) 2022/2464 of the European Parliament and of the Council of 14 December 2022 amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting. http://data.europa.eu/eli/dir/2022/2464/oj

Eijiffinger, S., & Pieterse-Bloem, M. (2023). Eurozone government bond spreads: a tale of different ECB policy regimes. *Journal of International Money and Finance*, 139: 102965. https://doi.org/10.1016/j.jimonfin.2023.102965

Escaleira Marques da Fonte, R. (2021). *Integrating ESG in Factor-based investing in Sovereign Bond Markets: a Size Approach*. [Work Project of the Master's degree in Finance, Nova School of Business and Economics]. NSBE: Nova SBE – MA Dissertations. https://run.unl.pt/bitstream/10362/142227/3/2021-22_fall_44171_rui-fonte.pdf.txt

Estran, R., Le, P., & Pineau, E. (2022). Importance of ESG factors in sovereign credit ratings. *Finance Research Letters*, 49: 102966. https://doi.org/10.1016/j.frl.2022.102966.

European Commission. (2024, April 4). What is the euro area. *Economy and Finance*. https://economy-finance.ec.europa.eu/euro/what-euro-area_en

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance and Investment*, 5 (4), 210-233. https://doi.org/10.1080/20430795.2015.1118917

Gómez Puig, M., Sosvilla-Rivero, S., & Ramos-Herrera, M. (2014). An update on EMU sovereign yield spread drivers in time of crisis: a panel data analysis. *UB Riskcenter*, Working Paper Series 2014/04, 1-37. http://www.ub.edu/rfa/research/WP/UBriskcenterWP201404.pdf

Gond, J.-P., & Piani, V. (2012). Enabling Institutional Investors' Collective Action: The Role of the Principles for Responsible Investment Initiative. *Business & Society*, 52 (1), 64-104. https://doi.org/10.1177/0007650312460012

Gratcheva, E. M., Emery, T., & Wang, D. (2020). Demystifying Sovereign ESG. *Equitable Growth, Finance and Institutions Insight - Finance*. Washington DC: World Bank. https://documents1.worldbank.org/curated/en/842671621238316887/pdf/Demystifying-Sovereign-ESG.pdf

Gratcheva, E., Gurhy, B., Skarnuli, A., Stewart, F., & Wang, D. (2022). Credit Worthy: ESG factors and sovereign credit ratings. *Equitable Growth, Finance and Institutions Insight - Finance*. Washington DC: World Bank. https://documents1.worldbank.org/curated/en/812471642603970256/pdf/Credit-Worthy-ESG-Factors-and-Sovereign-Credit-Ratings.pdf

Hua Fan, J., & Michalski, L. (2020). Sustainable factor investing: where doing well meets doing good. *International Review of Economics and Finance*, 70, 230-256. https://doi.org/10.1016/j.iref.2020.07.013

Jabeur, S., Sadaaoui, A., Sghaier, A., & Aloui, R. (2020). Machine Learning models and cost-sensitive decision trees for bond rating prediction. *Journal of the operational research society*, 71(8), 1161-1179. https://doi.org/10.1080/01605682.2019.1581405

Nazareth, N., & Ramana Reddy, Y. (2023). Financial applications of machine learning: a literature review. *Expert Systems With Applications*, 219: 119640. https://doi.org/10.1016/j.eswa.2023.119640

Nuta, F., Nuta, A., Zamfir, C., Petrea, S.-M., Munteanu, D., & Cristea, D. (2021). National Carbon Accounting - analyzing the impact of urbanization and energy related factors upon CO2 emissions in Central-Eastern European Countries by Using Machine Learning Algorithms and Panel Data Analysis. *Energies*, 14(10): 2775. https://doi.org/10.3390/en14102775

Olson, L., Qi, M., Zhang, X., & Zhao, X. (2021). Machine Learning loss given default for corporate debt. *Journal Of Empirical Finance*, 144-169. https://ssrn.com/abstract=3970125

Pineau, E., Phuong, L., & Estran, R. (2022). Importance of ESG factors in sovereign credit ratings. *Finance Research Letters*, Elsevier, 49: 102966. https://doi.org/10.1016/j.frl.2022.102966

Principles for Responsible Investment. (2019). *A practical guide to ESG integration in sovereign debt*. https://www.unpri.org/download?ac=9696

Principles for Responsible Investment. (2023). *2022-23 Annual Report*. https://dwtyzx6upklss.cloudfront.net/Uploads/z/s/n/pri_ar2023_smaller_file_8875.pdf

Sachs, J., Lafortune, G., Fuller, G., & Drumm, E. (2023). Implementing the SDG Stimulus. *Sustainable Development Report 2023*. Paris: SDSN, Dublin: Dublin University Press. https://www.tara.tcd.ie/handle/2262/102924

Sadorsky, P. (2021a). A Random Forests Approach to Predicting Clean Energy Stock Prices. *Journal of Risk and Financial management*, 14(2):48. https://doi.org/10.3390/jrfm14020048

Sadorsky, P. (2021b). Predicting Gold and Silver Price Direction Using Tree-Based Classifiers. *Journal of Risk and Financial Management*, 14(5):198. https://doi.org/10.3390/jrfm14050198

Schieler, M. (2017). Country Sustainability Update May 2017 - Sweden re-takes the lead. ROBECOSAM Report. Available at: https://www.researchgate.net/publication/319502799_Country_Sustainability_Ranking_Update_May_2017_-_Sweden_re-takes_the_lead

Serafeim, G. (2020). Public Sentiment and the Price of Corporate Sustainability. *Financial Analysts Journal*, 76(2), 26-46. https://doi.org/10.1080/0015198X.2020.1723390

Slimane, M. B., Brard, E., Le Guenedal, T., Roncalli, T., & Sekine, T. (2019). ESG Investing and Fixed Income: It's Time to cross the Rubicon. Discussion Paper DP-45-2020. Paris: Amundi Asset Management. https://research-center.amundi.com/files/nuxeo/dl/8d140ac3-868e-4aef-9260-d791577b26c1

Sustainable Development Solutions Network. (2024). Methodology. *Sustainable Development Report*. https://dashboards.sdgindex.org/chapters/methodology#table-a-4-5-indicators-included-in-the-sustainable-development-report-2023

United Nations Development Program. (2024). The SDGs in action. *Sustainable Development Goals*. https://www.undp.org/sustainable-development-goals

Viviers, S., & Eccles, N. (2012). 35 years of socially responsible investing (SRI) research: general trends over time. *South African Journal of Business Management*, 43(4), 1-16. https://doi.org/10.4102/sajbm.v43i4.478%0A

Wolppert. (1996). The lack of a priori distinctions between learning algorithms. *Neural Computation*, 8(7), 1341-1390. https://doi.org/10.1162/neco.1996.8.7.1341

World Health Organization, Regional Office for Europe (2024). Sustainable Development Goals. *World Health Organization – Europe.* https://www.who.int/europe/about-us/our-work/sustainable-development-goals#:~:text=The%20Sustainable%20Development%20Goals%20(SDGs,no%20one%20is%20left%20behind.)

Yang, B., Long, W., & Cai, Z. (2024). Machine learning based panel data models. (No. 202402). University of Kansas, Department of Economics. https://www2.ku.edu/~kuwpaper/2024Papers/202402.pdf

Yuankang, R., Lu, Y., Xu, X., & Yin, Y. (2019). Forecasting Credit Spreads: A Machine Learning Approach. https://www.iaqf.org/resources/Documents/2019%20Student%20Competition/UMich_Submartingale_Web.pdf

Zhang, Z., Zohren, S., & Roberts, S. (2020). Deep learning for portfolio optimization. *The Journal of Financial Data Science*, 2(4): 8-20. https://doi.org/10.3905/jfds.2020.1.042

6. ANNEX

6.1. Description of the variables

Variable	Description	Source
Gov_Bond	Government 10-year yield	Factset. Missing data completed with IMF data.
German_GB	German government 10-year yield	Factset
Spread	Difference between Gov-Bond and German_GB	
SDG_Score	SDG Score of each country (over 100)	SDSN
GDP	Gross Domestic Product (billions $)	Factset
Unemployment	Unemployment rate (%) as increase from the year before annualized	Factset. Missing values retrieved from OECD dataset.
Inflation	Inflation rate (%)	Factset
Debt_GDP	Debt to GDP ratio (debt as % of GDP)	Factset
CA_GDP	Current account to GDP ratio (CA as % of GDP)	Factset
REER	Real Effective Exchange Rate	Factset
Ratings	S&P credit ratings for each country	S&P

6.2. R Code

6.2.1. Regression model

```
#PREDICTION OF GOVERNMENT BOND SPREADS

#install necessary packages

install.packages("ggplot2")

install.packages("dplyr")

install.packages("lubridate")

install.packages("e1071")

install.packages("caret")

install.packages("ROCR")

install.packages("Metrics")

install.packages("caret")

install.packages("rpart")

install.packages("rpart.plot")

install.packages("randomForest")

install.packages("ranger")

install.packages("corrplot")

install.packages("PerformanceAnalytics")

install.packages("car")

#Install necessary libraries

library(ggplot2)

library(dplyr)

library(lubridate)

library(readxl)

library(e1071)

library(caret)
```

```
library(ROCR)
library(Metrics)
library(rpart)
library(rpart.plot)
library(ranger)
library(corrplot)
library(MASS)
library(car)
#Set working directory before importing dataset
#Import the dataset
eurozone <- read_excel("Datos Fundadores Eurozona.xlsx",
                       col_types = c("text", "text", "text", "date",
                                     "numeric", "numeric", "numeric",
                                     "numeric", "numeric", "numeric",
                                     "numeric", "numeric", "numeric",
                                     "numeric", "numeric"))
warnings()
View(eurozone)
#Check for missing values in dataset.
summary(eurozone)
#There is no NAs
#Study of correlation between variables
#correlation matrix
cor_matrix<-cor(eurozone[,-1:-6])
View(cor_matrix)
#graph of correlation matrix
```

```
corrplot(cor_matrix, method="shade",
        shade.col=NA, tl.col="black",
        tl.srt = 45)


library(PerformanceAnalytics)
chart.Correlation(eurozone[7:15],
                histogram=TRUE,
                method="pearson")
#statistical distribution of the target variable
hist(eurozone$Spread)
#the variable exhibits right skewness
#Explicative Model (log): use non-linear in order to normalize distribution - due to right
skewness of the variable)
model_log<-lm(log1p(Spread)~
SDG_Score+GDP+Unemployment+Inflation+Debt_GDP+CA_GDP+REER+Ratings,
            data=eurozone)
summary(model_log)
#risk of collinearity. We study the Variance Inflation Factor (VIF) of the model.
vif(model_log)


#Partition Dataset. 70% train and 30% test.
RNGkind("Super", "Inversion", "Rounding")
set.seed(4321)
trainIndex <- createDataPartition(eurozone$Spread, p=0.7, list=FALSE)
training <- eurozone[ trainIndex, ]
test <- eurozone[-trainIndex, ]
#FIRST MODEL (TRIAL): DECISION TREE
```

```
numbers<-10

rep<-3

grid<-expand.grid(cp=c(seq(0.01,0.3,by=0.001)))

#uses the grid search method to tune the cp hyperparameter (complexity of the tree - how many nodes)

#values will go from 0.01 to 0.3 in increments of one thousandth.

RNGkind("Super","Inversion", "Rounding")

set.seed(123)

x<-trainControl(method="repeatedcv", #repeated cross validation

                number = numbers, #folds (10)

                repeats = rep, #repetitions (3)

                verboseIter = TRUE)

#Train

tree<-
train(Spread~SDG_Score+GDP+Unemployment+Inflation+Debt_GDP+CA_GDP+REER+Ratings,

          data=training,

          method="rpart",

          trControl = x,

          tuneGrid = grid)

#model summary

tree

plot(tree)

rpart.plot(tree$finalModel)  #graphic representation of the tree

rpart.rules(tree$finalModel) #rules used by the tree

varImp(tree) #importance of variables

#model performance
```

```
#in training
pred_train_tree<-predict(tree,newdata=training)
error_train_tree<-training$Spread-pred_train_tree
h<-nrow(training) # numero de observaciones en el test
ME<-sum(error_train_tree)/h
RMSE<-sqrt(sum(error_train_tree^2)/h)
MAE<-sum(abs(error_train_tree))/h
MAPE<-sum (abs(error_train_tree)/training$Spread)/h*100
"In training:"; "ME";ME;"RMSE";RMSE;"MAE";MAE;"MAPE";MAPE
#prediction
prediction_tree<-predict(tree, newdata = test)
#calculate error
error_tree<-test$Spread-prediction_tree
#calculate measure
h<-nrow(test) #numero observaciones en el test
ME<-sum(error_tree)/h
RMSE<-sqrt(sum(error_tree^2)/h)
MAE<-sum(abs(error_tree))/h
MAPE<-sum(abs(error_tree)/test$Spread)/h*100
"In test:"; "ME";ME;"RMSE";RMSE;"MAE";MAE;"MAPE";MAPE
#Given overfitting we try ensemble algorithm.
#MODEL 1: RANDOM FOREST
numbers<-10
rep<-3
x<-trainControl(method="repeatedcv",
        number=numbers,
```

```
              repeats=rep,
              verboseIter = TRUE)

#RF estimation
forest<-
train(Spread~SDG_Score+GDP+Unemployment+Inflation+Debt_GDP+CA_GDP+RE
ER+Ratings,
             data=training,
             method="ranger",
             trControl=x,
             #tuneGrid=tgrid,
             tuneLength = 8,
             num.trees=500,
             importance="impurity")
#summary of random forest
forest
plot(forest)
varImp(forest)
#performance analysis of the model
#in training
pred_train_forest<-predict(forest,newdata=training)
error_train_forest<-training$Spread-pred_train_forest
h<-nrow(training) #numero de observaciones en el test
ME<-sum(error_train_forest)/h
RMSE<-sqrt(sum(error_train_forest^2)/h)
MAE<-sum(abs(error_train_forest))/h
MAPE<-sum (abs(error_train_forest)/training$Spread)/h*100
```

```
"In training:"; "ME";ME;"RMSE";RMSE;"MAE";MAE;"MAPE";MAPE
#prediction
prediction_forest<-predict(forest, newdata = test)
#calculate error
error_forest<-test$Spread-prediction_forest
#calculate measure
h<-nrow(test) #numero observaciones en el test
ME<-sum(error_forest)/h
RMSE<-sqrt(sum(error_forest^2)/h)
MAE<-sum(abs(error_forest))/h
MAPE<-sum(abs(error_forest)/test$Spread)/h*100
"In test:"; "ME";ME;"RMSE";RMSE;"MAE";MAE;"MAPE";MAPE
#MODEL 2: BOOSTING
library(gbm)
gbmGrid <- expand.grid(interaction.depth = c(1, 5, 9),
                    n.trees = (1:30)*50,
                    shrinkage = 0.1,
                    n.minobsinnode = 20)
x<-trainControl(method="repeatedcv",
             number=numbers,
             repeats=rep,
             verboseIter = TRUE)
#Boosting estimation
modelo_boost_trees <-
train(Spread~SDG_Score+GDP+Unemployment+Inflation+Debt_GDP+CA_GDP+RE
ER+Ratings,
                     data = training,
```

```
                     method = "gbm",
                     trControl = x,
                     verbose = FALSE,
                     tuneGrid = gbmGrid)
#summary of boosting model
 summary(modelo_boost_trees)
#performance analysis of the model
#in training
pred_train_boosting<-predict(modelo_boost_trees,newdata=training)
error_train_boosting<-training$Spread-pred_train_boosting
h<-nrow(training) # numero de observaciones en el test
ME<-sum(error_train_boosting)/h
RMSE<-sqrt(sum(error_train_boosting^2)/h)
MAE<-sum(abs(error_train_boosting))/h
MAPE<-sum (abs(error_train_boosting)/training$Spread)/h*100
"In training:"; "ME";ME;"RMSE";RMSE;"MAE";MAE;"MAPE";MAPE
#prediction
predicciones_boosting<-predict(modelo_boost_trees, newdata=test, type="raw")
errores_boosting<-test$Spread-predicciones_boosting
h<-nrow(test) # numero de observaciones en el test
ME<-sum(errores_boosting)/h
RMSE<-sqrt(sum(errores_boosting^2)/h)
MAE<-sum(abs(errores_boosting))/h
MAPE<-sum (abs(errores_boosting)/test$Spread)/h*100
"In test: ";"ME";ME;"RMSE";RMSE;"MAE";MAE;"MAPE";MAPE
```

6.2.2. Classification model

```r
# PREDICTION OF DIRECTION OF GOVERNMENT BOND SPREADS

#install necessary packages


install.packages("ggplot2")

install.packages("dplyr")

install.packages("lubridate")

install.packages("e1071")

install.packages("caret")

install.packages("ROCR")

install.packages("Metrics")

install.packages("caret")

install.packages("rpart")

install.packages("rpart.plot")

install.packages("randomForest")

install.packages("ranger")

install.packages("corrplot")

install.packages("PerformanceAnalytics")

#Install necessary libraries

library(ggplot2)

library(dplyr)

library(lubridate)

library(readxl)

library(e1071)

library(caret)

library(ROCR)

library(Metrics)
```

```r
library(rpart)
library(rpart.plot)
library(ranger)
library(corrplot)
library(MASS)
#Set working directory before importing dataset
#Import the dataset
eurozone_clas <- read_excel("clas Datos Fundadores Eurozona.xlsx",
                  col_types = c("text", "text", "text", "date",
                            "numeric", "numeric", "numeric",
                            "text", "numeric", "numeric",
                            "numeric", "numeric", "numeric",
                            "numeric", "numeric", "numeric"))
colnames(eurozone_clas)[colnames(eurozone_clas) == 'up/down'] <- 'direction'
warnings()
View(eurozone_clas)
#Check for missing values in dataset.
summary(eurozone_clas)
#There is no NAs
#Study of correlation between variables
#correlation matrix
cor_matrix<-cor(eurozone_clas[,-1:-7])
View(cor_matrix)
#graph of correlation matrix
corrplot(cor_matrix, method="shade",
      shade.col=NA, tl.col="black",
```

```
    tl.srt = 45)
library(PerformanceAnalytics)
chart.Correlation(eurozone[7:15],
          histogram=TRUE,
          method="pearson")
#Explicative Model (classification)
model_ex <- glm(direction ~ SDG_Score + GDP + Unemployment + Inflation + Debt_GDP + CA_GDP + REER + Ratings,
               data = eurozone_clas,
               family = binomial)
summary(model_ex)
#Partition Dataset. 70% train and 30% test.
RNGkind("Super", "Inversion", "Rounding")
set.seed(4321)
trainIndex <- createDataPartition(eurozone_clas$direction, p=0.7, list=FALSE)
#creamos los sets de entrenamiento y test
entrenamiento <- eurozone_clas[ trainIndex, ]
test <- eurozone_clas[-trainIndex, ]
# MODEL (TRIAL): DECISION TREE
numbers<-10
rep<-3
grid<-expand.grid(cp=c(seq(0.01,0.3,by=0.001)))
# In this case, the values to be tested for the hyperparameter cp range from 0.01 to 0.3 in steps of 0.001.
RNGkind("Super","Inversion", "Rounding")
set.seed(123)
```

```
x<-trainControl(method="repeatedcv", #método de remuestreo (cross validation repetida)
        number = numbers, #número de folds
        repeats = rep, #repetimos 3 veces
        classProbs = TRUE, #se obtienen las probabilidades de clasificación
        summaryFunction = twoClassSummary,
        verboseIter = TRUE)
#Train
tree<-
train(direction~SDG_Score+GDP+Unemployment+Inflation+Debt_GDP+CA_GDP+REER+Ratings,
      data=entrenamiento,
      method="rpart",
      trControl = x,
      tuneGrid = grid)
#model summary
tree
plot(tree)
rpart.plot(tree$finalModel)  #graphic representation of the tree
rpart.rules(tree$finalModel) #rules used by the tree
varImp(tree) #importance of variables
#model performance
#in training
pred_train_tree <- predict(tree, newdata=entrenamiento)
entrenamiento$direction <- factor(entrenamiento$direction, levels = levels(pred_train_tree))
confusion_matrix<-confusionMatrix(pred_train_tree, entrenamiento$direction)
```

```
accuracy<-confusion_matrix$overall["Accuracy"]
"In training:"; accuracy
#prediction
prediction_tree<-predict(tree, newdata = test)
test$direction <- factor(test$direction, levels = levels(pred_train_tree))
confusion_matrix<-confusionMatrix(prediction_tree, test$direction)
accuracy<-confusion_matrix$overall["Accuracy"]
"In test:"; accuracy
#Given overfitting we try ensemble algorithm.
#MODEL 1: RANDOM FOREST
numbers<-10
rep<-3
x<-trainControl(method="repeatedcv", #método de remuestreo (cross validation repetida)
        number = numbers, #número de folds
        repeats = rep, #repetimos 3 veces
        classProbs = TRUE, #se obtienen las probabilidades de clasificación
        summaryFunction = twoClassSummary,
        verboseIter = TRUE)
#RF estimation
forest<-
train(direction~SDG_Score+GDP+Unemployment+Inflation+Debt_GDP+CA_GDP+REER+Ratings,
       data=entrenamiento,
       method="ranger",
       trControl=x,
       #tuneGrid=tgrid,
```

```
              tuneLength = 8,
              num.trees=500,
              importance="impurity")

#summary of random forest
forest
plot(forest)
varImp(forest)
#performance analysis of the model
#in training
pred_train_forest<-predict(forest,newdata=entrenamiento)
entrenamiento$direction <- factor(entrenamiento$direction, levels = levels(pred_train_forest))
confusion_matrix<-confusionMatrix(pred_train_forest, entrenamiento$direction)
accuracy<-confusion_matrix$overall["Accuracy"]
"In training:"; accuracy
#in test
prediction_forest<-predict(forest, newdata = test)
test$direction <- factor(test$direction, levels = levels(prediction_forest))
confusion_matrix<-confusionMatrix(prediction_forest, test$direction)
accuracy<-confusion_matrix$overall["Accuracy"]
"In test:"; accuracy
#we use boosting as another ensemble
#MODEL 2: BOOSTING
library(gbm)
gbmGrid <- expand.grid(interaction.depth = c(1, 5, 9),
```

```
                n.trees = (1:30)*50,
                shrinkage = 0.1,
                n.minobsinnode = 20)
x<-trainControl(method="repeatedcv", #método de remuestreo (cross validation repetida)
           number = numbers, #número de folds
           repeats = rep, #repetimos 3 veces
           classProbs = TRUE, #se obtienen las probabilidades de clasificación
           summaryFunction = twoClassSummary,
           verboseIter = TRUE)
#Boosting estimation
modelo_boost_trees <-
train(direction~SDG_Score+GDP+Unemployment+Inflation+Debt_GDP+CA_GDP+REER+Ratings,
                 data = entrenamiento,
                 method = "gbm",
                 trControl = x,
                 verbose = FALSE,
                 tuneGrid = gbmGrid)
#Summary of boosting model
summary(modelo_boost_trees)
#performance analysis of the model
#in training
pred_train_boost<-predict(modelo_boost_trees,newdata=entrenamiento)
entrenamiento$direction <- factor(entrenamiento$direction, levels = levels(pred_train_boost))
confusion_matrix<-confusionMatrix(pred_train_boost, entrenamiento$direction)
```

```r
accuracy<-confusion_matrix$overall["Accuracy"]

"In training:"; accuracy

#prediction

prediction_boost<-predict(modelo_boost_trees, newdata = test)

test$direction <- factor(test$direction, levels = levels(prediction_boost))

confusion_matrix<-confusionMatrix(prediction_boost, test$direction)

accuracy<-confusion_matrix$overall["Accuracy"]

"In test:"; accuracy

#boosting shows best performance in classification.
```