

# An analysis of annual reports from the sustainable development goals perspective

Erk Hacıhasanoğlu, Ömer Faruk Ünlüsoy and Fatma Selen Madenoğlu

## Abstract

**Purpose** – The sustainable development goals (SDGs) are introduced to guide achieving the sustainable goals and tackle the global problems. United Nations members may perform activities to achieve the predetermined goals and report on their SDG activities. The comprehension and commitment of several stakeholders are essential for the effective implementation of the SDGs. Countries encourage their stakeholders to perform and report their activities to meet the SDGs. The purpose of this study is to investigate the extent to which corporations' annual reports address the SDGs to assess and comprehend their level of commitment to, priority of and integration of SDGs within their reporting structure. This research makes it easier to evaluate corporations' sustainability performance and contributions to global sustainability goals by looking at the extent to which they address the SDGs.

**Design/methodology/approach** – In the study, it is revealed to what extent the reports meet the SDGs with the multilabel text classification approach. The SDG classification is carried out by examining the report with the help of a text analysis tool based on an enhanced version of gradient boosting. The implementation of a machine learning-based model allowed it to determine which SDGs are associated with the company's operations without the requirement for the report's authors to perform so. Therefore, instead of reading the texts to seek for "SDG" evidence as typically occurs in the literature, SDG proof was searched in relevant texts.

**Findings** – To show the feasibility of the study, the annual reports of the leading companies in Turkey are examined, and the results are interpreted. The study produced results including insights into the sustainable practices of businesses, priority SDG selection, benchmarking and business comparison, gaps and improvement opportunities identification and representation of the SDGs' importance.

**Originality/value** – The findings of the analysis of annual reports indicate which SDGs they are concerned about. A gap in the literature can be noticed in the analysis of annual reports of companies that fall under a particular framework. In addition, it has sparked the idea of conducting research on a global scale and in a time series. With the aid of this research, decision-making procedures can be guided, and advancements toward the SDGs can be achieved.

**Keywords** Multilabel text classification, Extreme gradient boosting, Annual report, SDGs

**Paper type** Research paper

Erk Hacıhasanoğlu, Ömer Faruk Ünlüsoy and Fatma Selen Madenoğlu are all based at the Department of Business Administration, Abdullah Gül Üniversitesi, Kayseri, Turkey.

## 1. Introduction

The United Nations (UN, 2015) released the important document "Transforming our world: the 2030 agenda for sustainable development" in 2015. For the period 2015–2030, the document contains 17 sustainable development goals (SDGs) and 169 targets. The first published agenda expresses globally the attraction of the activities toward economic, environmental and social sustainability (Moldan *et al.*, 2012). The agenda aims to address a wide range of sustainable development issues in both developed and developing countries with all actors at all levels including government, civil societies, the private sector and the science communities for attaining expected consequences (Virto, 2018).

Received 14 April 2023  
Revised 28 August 2023  
Accepted 14 September 2023

As one of the key stakeholders, international companies have played a crucial role in reaching SDGs. They can also identify activities that can direct toward practices that are compatible with the SDGs by establishing sustainable frameworks. They report on their sustainability efforts in either their annual reports or sustainability reports, which are up to the specific companies.

Specifying influential involvement to reach the SDGs is necessary to establish the methods for monitoring and measuring the activities and progress alignment with the SDGs. A wide variety of indicators are required to measure and show progress toward the targets to the policymakers and the public (Moldan *et al.*, 2012). The composite indexes are generated to evaluate the countries' performances in policy generation and public responsibility. These indices aggregate the single indicators to make the assessment broadly in a multidimensional framework (Nardo *et al.*, 2015). The composite indices are used to make the comparison of countries easily and trigger the competition. Some indices quantify the progress of a specific SDG, and others ensure the combined measurement of multiple SDGs. The 232 global indicators which are introduced by UN Commission attribute to field data which reflects the current situation. The fact that many factors and multiple stakeholders are taken into account in progress measures based on field data shows how they affect the performance of a country as a whole. Researchers have recognized one of the most important scientific challenges: developing effective methods for visualizing the status and changes in SDG indicators and communicating the results to policymakers and other end users.

With the help of the technologies, big data are formed in many areas. These data sets are not processed and analyzed in an unstructured way. The data sets include photo, video, audio and text such as e-mail, review, Web pages, report and document. The analysis process on data sets in text format is carried out with the help of text mining, which is a subfield of data mining. In this way, information and meaning are extracted from text contents that are not in structural format. In recent years, studies have been published in which the SDGs are examined using machine learning models or specialized software with the use of data mining. Researchers have turned to artificial intelligence-supported methodologies because human research techniques do not handle the model in a consistent and inclusive manner (Tóth *et al.*, 2022). It is now possible to uncover the texts' hidden principles by using models backed by artificial intelligence (Zhou *et al.*, 2021). The quantity and objectivity of the data that must be evaluated present the researchers studying this topic with their first challenge. A large amount of text can be tested quickly using artificial intelligence-based models (Lafleur, 2019). These methods allow for the association of texts with a single SDG or a number of SDGs. In addition, it is possible to gauge the degree of linkage with related SDGs (Guisiano and Chiky, 2021). Multiple label classification organizes texts into categories and clearly illustrates how closely they relate to the SDGs (Sovrano *et al.*, 2020).

The purpose of this study is to detect and evaluate how the SDGs and the annual reports of publicly traded corporations for a single year aligned. In line with the purpose, it was determined which SDGs these enterprises prioritized and addressed to evaluate their sustainability practices and their contributions to sustainable development. Instead of examining the published texts about the SDGs, it was preferred to examine the annual reports, taking into account the gaps in the literature and to determine how much each company deliberately referred to each SDG. Companies also specify their sustainability-related activities in their mandatory annual reports. The SDGs are not necessarily being actively addressed or that any steps have been done to incorporate them into operations just because they are included in their annual report (Song *et al.*, 2022). As a result, rather than seeking out direct conclusions about the SDGs, the texts were examined in light of the SDGs and were linked to the SDGs. Assessing the degree of correlation between annual reports and particular SDGs offers important insights into how well a corporation performs and contributes in terms of sustainability. While facilitating comparative analyses and

materiality assessments, it encourages transparency, accountability and stakeholder participation. By coordinating corporate efforts with the SDGs, this study aids in identifying focus areas, evaluating risks and fostering positive change toward sustainable development.

Türkiye has shown a significant commitment to implementing the SDGs, as seen by its institutional structures, national development goals and active involvement in international forums. The SDGs are included into Türkiye's national agenda in the 11th Development Plan (2019–2023), which was created by the Ministry of Development and lays out precise objectives that are in line with the global agenda (Ministry of Development of the Republic of Turkey, 2019). Institutional procedures have been set up by the Ministry of Development, the Strategy and Budget Directorate and the Presidency of the Republic of Türkiye to coordinate and oversee SDG implementation, guaranteeing a thorough and integrated approach (Sustainable Development Goals Evaluation Report 2016-2019). Türkiye presented its first voluntary national review (VNR) in 2017 at the UN High-Level Political Forum in accordance with the global reporting mechanism, providing a thorough evaluation of its progress toward the SDGs and reiterating its commitment to monitoring and reporting on SDG implementation (United Nations Economic and Social Council, 2017). The SDGs for eradicating poverty, promoting gender equality, renewable energy, sustainable agriculture and inclusive economic growth have all been taken into account by Türkiye in its policies and plans. These legislative actions demonstrate Türkiye's commitment to resolving the complex issues of sustainable development (Republic of Turkey Ministry of Foreign Affairs, 2021). BIST30 companies are selected because they are the most valuable and leading businesses in Türkiye.

The annual reports of companies have analyzed the contribution of their activities to SDGs by researchers. The proposed approach links many reports to the 17 SDGs and measures the degree of correspondence of the report in a consistent way. In general, the country's progress in the SDG process can be represented, as these leading firms will reflect their contribution to the country's overall SDG situation. As far as we know, there has not yet been a study that reviews at the annual reports of the leading companies in the country to assess how well the 17 SDGs are being met. This study offers helpful insights into the sustainability practices of prominent firms in Türkiye by examining their annual reports. It reveals which SDGs these businesses prioritize and the degree of dedication they show to accomplishing particular objectives. The public, governments, investors and other stakeholders may find this information helpful in evaluating the sustainability performance of businesses and holding them responsible for their contributions to the SDGs. In their reports, businesses outline their efforts and accomplishments in sustainability. Te Liew *et al.* (2014) introduced the sectoral approach by examining the sustainability reports of businesses in particular industries with the aim of identifying the sustainability trends and practices of organizations using text analysis.

The following part of this paper is organized as follows. The related literature review is presented in Section 2. The research method is explained in Section 3. The results of the analysis are given in Section 4. Finally, Section 5 concludes the paper.

## 2. Literature review

The SDGs urge corporations to adopt sustainable practices and incorporate sustainability information into their reporting cycle to solve the most urgent sustainability challenges facing the entire globe (Pizzi *et al.*, 2021). By examining the features of engagement with the SDGs of 1,370 organizations from 97 countries, one study looked at how organizations were interacting with them. The majority of organizations only had a cursory understanding of the SDGs, which suggests a “SDG-washing” process (Heras-Saizarbitoria *et al.*, 2022). Another study looked into how far the SDGs have been incorporated into the disclosure processes of Italian-listed companies. According to the study's findings (Izzo *et al.*, 2020),

the majority of highly traded, liquid and capitalized Italian corporations have included the SDGs into their transparency and storytelling processes. A company's commitment to the SDGs was examined in relation to internal corporate governance. Results show that companies are more likely to report on the SDGs in their sustainability reports when their corporate governance is stronger, including aspects like chief executive officer independence, board composition and board attendance (Martínez-Ferrero and García-Meca, 2020). Elalfy *et al.* (2021) advanced knowledge of the variables affecting SDG reporting. According to their findings, larger organizations, publicly traded companies and sectors with higher implications on sustainability are more likely to include the SDGs in their reporting. Understanding the various factors that influence the adoption of SDGs by organizations and how these factors fluctuate across various sectors and countries is an area that needs more research. In addition, Hummel and Szekely (2022) observed the standard of SDG disclosure in companies' annual reports and find a progressive development. They also draw attention to the absence of quantitative and prognostic information, highlighting a research hole in our knowledge of the thoroughness and caliber of SDG disclosure in annual reports. In their analysis of corporate sustainability reporting methods in the banking industry, they identify the major reporting issues, frameworks and difficulties banks encounter when revealing their sustainability performance and fostering accountability. Ehnert *et al.* (2016) aimed to examine and evaluate the human resource management (HRM) component of company paperwork. According to the human resource perspective of the SDGs, the study illustrates the contrasts between coordinated and liberal economy corporate reporting. Results indicate that businesses engaged in a coordinated economy are more likely to report than in liberal economy. Performance linked to HRM is on par with environmental performance in terms of SDG achievement. (Van Zanten and Van Tulder, 2018) examined how the private sector may effectively implement foreign policy goals. The SDGs were selected as global policy. By using an exploratory survey of the managers at various levels of the companies, the findings about the incorporation of the SDGs were discovered. The stakeholders would not be happy with the financial reporting alone, according to Al Amosh and Khatib (2021), who looked at the new and voluntary approach to sustainability activities. The study implemented content analysis to ascertain the environmental, social and governance (ESG) strategy used by the companies listed on the Amman Stock Exchange. Caratas *et al.* (2021), in their study of 335 health-care companies for their research carried out within the framework of ESG, extended to the empirical literature by expressing that sustainability at the institutional level will contribute to risk avoidance and value creation. Cosma *et al.* (2020) manually used content analysis of materials provided by 262 European banks in an effort to gauge the amount of banks' contributions to the SDG and various interbank SDG approaches. Priority SDGs for banks were determined and then assessed based on factors such as country, year and business model.

The studies on sustainability with the text mining method are reviewed in this part more specifically. In general, the reports published by the companies or research papers are examined in detail in terms of meeting SDGs. Wang *et al.* (2020) demonstrated how the maritime industry contributes to the SDGs. They did this by analyzing the content of sustainability reports provided by container shipping lines and terminal operators between 2016 and 2019. They used an annual text classification to classify the sustainability report depending on 17 SDGs and automatic text mining to emphasize the contribution of the marine industry in each SDG. Juventia *et al.* (2020) presented the text mining technique to analyze the agrobiodiversity index and commitments scores for nine countries. With the help of text analysis methods, Lee *et al.* (2020) created an urban sustainability activity index for their study, which used the C40 city data set to identify four underlying issues. Within the context of sustainability, the activities of the cities were determined in accordance with this index. Sebestyén *et al.* (2020) investigated the report of VNRs about SDG and used network-based comparative analysis. They analyzed more than 100 voluntary reviews. Text

mining algorithms turn the report into keyword networks to find country-specific subject areas of the strategies and cluster countries with similar challenges and development strategies. The results depicted that SDG 5 (gender equality) and SDG 8 (decent work and economic growth) are the most examined goals, and SDG1 (no poverty) and SDG 10 (reduced inequalities) are the least examined goals. [Sovrano et al. \(2020\)](#) developed the deep learning-based automated multilabel model with the goal of developing software to categorize documents in accordance with the SDGs to provide solutions beyond the difficult task and inconsistency issue that emerges in human-based content analysis. Using unsupervised text mining via the latent Dirichlet allocation (LDA) algorithm, [Zhou et al. \(2021\)](#) investigated the latent information of 4,735 sustainability report paragraphs collected from top-ranked container shipping businesses. They made a contribution to the maritime literature by providing a unified sustainability disclosure methodology for this sector. [Hwang et al. \(2021\)](#) examined SDG 13.3, which aims to “improve education, awareness-raising and human and institutional capacity on climate change mitigation, adaptation, impact reduction, and early warning” among social groups in Korea from a text mining approach. LDA and co-occurrence network analysis are used as text analysis techniques. [Bordignon \(2021\)](#) analyzed the important and widely applied field for each SDGs using a text mining approach. [Jang \(2021\)](#) reviewed the academic papers which are published by Springer to specify policy issues related to the education gap using text mining approach. [Roy et al. \(2022\)](#) presented the recent trends in SDG 6 (clean water and sanitation) for six years. They used bibliometric analysis based on text mining to support SDG 6. That pointed out that the top three countries are the USA, UK and The Netherlands, considering research publication, and a few countries play role in SDG 6 research. Although there are models that have been developed through machine learning methods that take into account the SDGs, no article has been discovered that explains how to assess a company's target success by looking at the company's annual reports. Therefore, a model appropriate for the goal of the pertinent study was constructed by developing machine learning techniques discovered in the literature.

### 3. Research methodology

The general framework of the research model has two stages. The training data set consists of texts associated with the SDGs. As a first step, a preprocessing process was applied to these texts. Using Natural Language ToolKit (NLTK), a tool for natural language processing (NLP), the specified training data set was made appropriate for the research model. The science of NLP is concerned with the development and application of computer systems whose primary purpose is to analyze, comprehend, interpret and produce natural language. Functions were used to stem the words to their root word and remove the stop words and punctuations from the texts. The second stage involved using the XGBoost Machine Learning Model to train the text set with class labels to develop a model that automatically learns the characteristics of the class that each document should adhere to [Tantuğ \(2016\)](#). The model was assessed using accuracy, precision, recall and f1 score in the final phase. The proportion of association with the SDGs was calculated after processing the annual report texts in the model ([Baker and Korhonen, 2017](#)).

NLP serves as the theoretical foundation for this essay on multilabel text classification in business annual reports. The study makes use of the fundamental ideas of NLP, which concentrate on the extraction of pertinent material from enormous volumes of textual data and entail the computational analysis and understanding of human language ([Feldman and Sanger, 2007](#)). The idea of text classification, which covers the classification of documents into preset classes or labels, is specifically the foundation upon which this research is built. The analysis of annual reports, which are crucial disclosure documents that offer insights into a company's general business operations, also shows the sustainability priorities of organizations ([Schütze et al., 2008](#)). This study seeks to provide an efficient and reliable

methodology by using these theoretical foundations. This research seeks to create a reliable and efficient system for multilabel text classification in corporate annual reports, improving reader understanding and interpretation of these important publications.

### **3.1 Data acquisition**

The purpose of this study is to obtain SDGs-related deductions from annual report texts. Annual reports, which businesses may compile and display their annual activities, are used in this study instead of the sustainability report, which is left to the initiative of the companies to prepare. The annual reports of the companies also provide information about their sustainability efforts; thus, this situation does not pose a barrier to the study's fulfillment (Deloitte ve TKYD, p. 4). The research was focused on BIST30 companies, which are required by Turkish Commercial Code No. 6102 to prepare annual reports (Deloitte ve TKYD, p. 5). From a wider viewpoint, the BIST30 companies were selected because they are the most valuable and leading businesses in Türkiye (Borsa İstanbul, p. 12). The reports of 27 out of 30 firms have been obtained because the BIST30 index has changed over time and several companies no longer maintain the websites on which they publish their annual reports. 2020 was selected as the year having the least amount of incompleteness in the annual reports of the BIST30 companies. Annual reports are downloaded from companies' websites with a length of 50 pages on average, thus acquiring big data for this study. The reports have been used in the literature for this purpose (Boiral and Henri, 2017; Al Amosh and Khatib, 2021; López-Santamaría *et al.*, 2021).

### **3.2 Data preprocessing**

The process of transforming a raw data set into a model format is known as data preprocessing. Data preprocessing is a crucial stage in data mining to increase the effectiveness level of data. The data cleaning procedure finds and eliminates the flaws and discrepancies in the data, hence enhancing its quality. Misspellings made during data storage, scraping, missing values or any other erroneous data cause problems about data quality.

First of all, all uppercase letters are converted to lowercase ones. This makes it easy to read the text data by the method. The NLTK module, which was created for the Python programming language and includes numerous NLP packages, is a substantial toolbox. Using NLTK paragraphs include framing sentences, separating words, determining parts of speech, highlighting themes and assisting machine learning models in comprehending the content of texts. This module is used in the following step. The following step is tokenization. The text data which is lowercased is splitted to separate words because these words can express the meaning of data. For example, "no progress were available" is tokenized into the words "no," "progress," "were" and "available." Next, stop words and punctuation are taken out of the original text data. Afterwards, the tokenized text data turns words into their root form. This process is named as lemmatization. For example, the tokenized text data "no," "progress," "were" and "available" is transformed into a lemmatized words "no," "progress," "be" and "available." Finally, we obtained the normalized text data to incorporate into the model.

### **3.3 Text mining**

The classification methods used in machine learning and deep learning are examined in three groups. The process of classifying data into one of two categories to solve classification problems is known as binary classification. This categorization technique yields one of two outputs for a given input. In classification issues where there are at least two classes, multiclass classification is used. It is a categorization technique that yields more than two results from a single input. In other words, an input can be categorized into

two outputs simultaneously. The input and all outputs may match concurrently in some circumstances. When there are multiple inputs in a classification problem, multilabel classification is used. This categorization method has more than one label in contrast to the others. According to the multiclass classification method, these inputs may correspond to more than two outputs (Meier *et al.*, 2021; Baker and Korhonen, 2017; Liu *et al.*, 2017).

To show the differences among classification methods, Table 1 presents an example of movie classification and categorizes three movies. The binary classification method divides movies into two categories as good and bad. The multiclass classification method rates the movies between 1 and 10. The movies are categorized using the multilabel classification method by assigning various values between 0 and 5 to different layers.

The multilabel classification method is used in this study to determine the degree of correlation between the texts in the companies' annual reports and the 17 SDGs. The annual reports' texts were examined, and each SDG's level of relevance was determined independently. All reports in the model are linked to several SDGs. Some SDGs were highly correlated in reports, while others were completely unrelated.

Machine learning has sparked a lot of interest in social science in recent years. In an effort to assess and solve the research issues, a variety of machine learning models, including linear/logistic regression, support vector machines, neural networks and tree-based models, are being attempted and implemented. Researchers have discovered that some models outperform other machine learning models in terms of success rate. One such machine learning model that has been highly praised by machine learning researchers is XGBoost (Gumus and Kiran, 2017). XGBoost (eXtreme Gradient Boosting), proposed by Chen and Guestrin (2016), is known as an enhanced version of gradient boosting, which is an effective machine learning method developed by Friedman (2001). XGBoost is an approach to machine boosting, or more specifically, the use of boosting to machines for tree boosting algorithms. XGBoost assists in making use of all available hardware and memory resources. It offers the advantages of improving algorithms, fine-tuning models and deployment in computing settings. The algorithm is very efficient in cutting down on computation time and making the most effective use of memory (Mitchell and Frank, 2017). High prediction accuracy, the capability to avoid overlearning, the ability to manage empty data and velocity are the algorithm's key characteristics of the method (Ogunleye and Wang, 2019). It has proven possible to achieve better results with fewer resources by using software and hardware efficiency approaches (Chen and Guestrin, 2016).

For the following reasons, XGBoost is chosen as the classification model of choice when addressing difficulties in real word classification problems:

- XGBoost is around ten times faster than existing approaches on a single platform, solving the problem of time consumption, particularly when preprocessing of the network data is carried out.
- XGBoost offers the benefit of parallel processing, which uses all of the processor cores on the system it is executing on. With the use of algorithmic optimization procedures, distributed or parallel computation, and a high degree of scalability, it may generate billions of examples while consuming the fewest resources possible. The classification

**Table 1** An example of classification methods

Movie name	Binary classification	Multiclass classification	Multilabel classification						
	(Good or Bad)	(Rating 1–10)	Action	Adventure	Crime	Drama	Romance	Sci-Fi	War
X	Good	9	0	0	5	4	0	0	0
Y	Good	8	0	0	0	5	4	0	4
Z	Bad	6	1	0	0	3	3	0	0

Source: Authors' own work

of data and high-level preprocessing of data are two difficulties that it is quite effective at handling.

- The classification model produces fewer false alarms and labels data easily and correctly due to XGBoost's capacity to turn a weak learner into a strong learner (boosting) through its optimization step for each new tree that attaches.
- XGBoost is well suited to identify missing values and handle them.
- XGBoost is a flexible classifier because it allows the user the ability to customize the goal function by modifying the model's parameters. Along with addressing regression, classification and ranking issues, it also enables user-defined assessment measures (Dhaliwal *et al.*, 2018).

The XGBoost approach was found to be more effective in solving classification problems by Hendrawan *et al.* (2022) in their study on e-commerce product reviews. Elmitwally (2020) claimed in his research that he used XGBoost to accomplish the classification problem's highest prediction performance. The XGBoost approach is an efficient text classification algorithm as well as a useful prediction tool, according to Qi (2020), who used the theft crime data for text classification and prediction. Finding some potential explanations for why tree boosting is so successful is intriguing for a variety of reasons. First, it might help us better understand how tree boosting techniques function internally. Second, it might contribute to furthering and improving the existing tree boosting method. Third, by comprehending the fundamental ideas behind tree boosting that give it its versatility, we may be able to create entirely new teaching strategies that combine those fundamental ideas (Nielsen, 2016). The XGBoost method was determined to be suitable for the text classification and prediction challenge according to the literature.

To avoid having different learning outcomes for each execution, the model was developed using the same test and learning sets. Given that the problem was one of classification, the largest number of iterations – which maximized the number of decision trees that would grow in the model – was selected. The model was developed by freely assigning values to each tree's weight estimation. A random split of 70% of the data set was used for learning data. In total, 30% of the data were used for testing after the learning phase. The high ability for learning and prediction of the model led to the selection of these parameters. The model produced highly accurate predictions with these configurations.

### 3.4 Data sources

A GitHub repository (GitHub "SDG-s-classification" Project, 2022) data set and "OSDG Community Dataset," which was made available to the public in 2021, were used in this study as training and test data. OSDG is conducting a number of SDG-labeled text data sets to participate in the international effort to accomplish the SDGs. The OSDG data set includes tens of thousands of text samples that community volunteers assessed in terms of the SDGs. Volunteers provided positive or negative feedback on whether the relevant texts were related to the associated SDGs. OSDG applies the set of data to perform multilabel text classification. The data set file has 35,605 assigned labels and 39,696 text excerpts. The data set consists of paragraph-length text extracts from publicly accessible sources, such as reports, policy documents and publication abstracts. A substantial portion of the documents originate from United Nations-affiliated resources like SDG-Pathfinder and SDG Library. Documents with SDG labeling are frequently found in these sources. Each text has approximately three to six sentences and, on average, 90 words (OSDG, UNDP IICPSD SDG AI Lab, & PPMI, 2022). For instance, "From a gender perspective, Paulgaard points out that the labor markets of the fishing [...]," the SDG 5 is linked to the given sentence. While one respondent thought this judgment was inaccurate, seven volunteers indicated that this association was accurate in their comments. Equation (1) was used to determine

the value of the agreement. There are a lot of entries in the data set where the value of the negative labels is greater than the value of the positive labels. These data have an extremely low agreement value. These entries are useful for training the model; however, they actually render it less accurate. Therefore, only data with an agreement value of at least 0.75 were used. The value of 0.75 was used as a criterion in this study to provide more accurate results, even though the agreement value of 0.6 is thought to be sufficient in similar projects (GitHub “UN-SDG-Classifier,” 2021):

$$agreement = \frac{|labels_{positive} - labels_{negative}|}{labels_{positive} + labels_{negative}} \quad (1)$$

### 3.5 Model evaluation

The number of actual and expected classifications made by the classifier is contained in the confusion matrix, which is given in Table 2. When the model predicted true, and the actual value was true, this is referred to as a true positive (TP). False positive (FP) is a term used to describe situations when the model predicted true, but the actual value was false. When the model predicted false, but the actual value was true, the result is referred to as a false negative (FN). When the model predicted false, and the actual value was false, this is expressed as true negative (TN).

The following performance indicators are used to evaluate the ensemble classifier: The performance of the classifier is shown using a confusion matrix based on the four parameters described above – TP, FP, TN and FN. Precision, recall and f1-score are calculated using TP, FP, TN and FN. Equations (2)–(5) are used to determine the accuracy of the model in terms of precision, recall and f1-score (Ludwig, 2019; İş and Tuncer, 2019).

Precision is the ratio of the TP samples in the data set that have been accurately sorted to the total of the TP and FN. Equation (2) is used to calculate the ratio:

$$precision = \frac{TP}{(TP + FN)} \quad (2)$$

Recall, which is the ratio of correctly separated TP samples to the total of TP and FN in the data set, demonstrates how successfully positive conditions have been analyzed. Equation (3) is used to calculate the ratio:

$$recall = \frac{TP}{(TP + FP)} \quad (3)$$

The division of the successfully separated samples in the data set by the overall sample count is referred to as accuracy. Equation (4) is used to determine accuracy values:

$$accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (4)$$

**Table 2** Confusion matrix

	Actual positive	Actual negative
Predictive positive	Ture positive (TP)	False positive (FP)
Predictive negative	False negative (FN)	True negative (TN)

Source: Created by İş and Tuncer (2019)

The harmonic mean of precision and recall is provided by the f1-score. In comparison to all other classes, the scores assigned to each class indicate how accurately the classifier categorized the data points inside that class (Matsui *et al.*, 2022). This score accounts for both FP and FN. Although f1-score is generally more significant than accuracy, it is not as intuitively clear as accuracy (Hajikhani and Suominen, 2022):

$$f1 - score = \frac{precision * recall}{precision + recall} \quad (5)$$

After machine learning training and testing process, for each SDG, individual precision scores were obtained. Table 3 displays the results of correlating the data set with the SDGs. Accuracy based on SDGs is between 0.9398 and 0.9861.

Then, as input to the model, the texts of the companies' annual reports were added. For reports from 27 firms, the methodology computed the SDGs' percentage of the relation. The results were combined according to the percentage of SDGs that each report's content related to.

#### 4. Results and discussions

The alignment of the reports with the SDGs was evaluated by the research using a multilabel text classification method and a text analysis technique built on an improved version of gradient boosting. The study's findings shed light on the particular SDGs that businesses prioritize and support through their sustainability initiatives. The study filled a gap literature by concentrating on the examination of annual reports within a specific framework and doing so by evaluating the annual reports of top corporations in Türkiye.

The SDGs ratios of firms are evaluated to assess which SDGs are most mentioned by firms' overall activity. Figure 1 shows the SDG reporting status of companies according to the model. The demonstration of SDG's ratios is given in Figure 2. Figure 3 shows a Sankey diagram of BIST 30 Companies to cover SDGs.

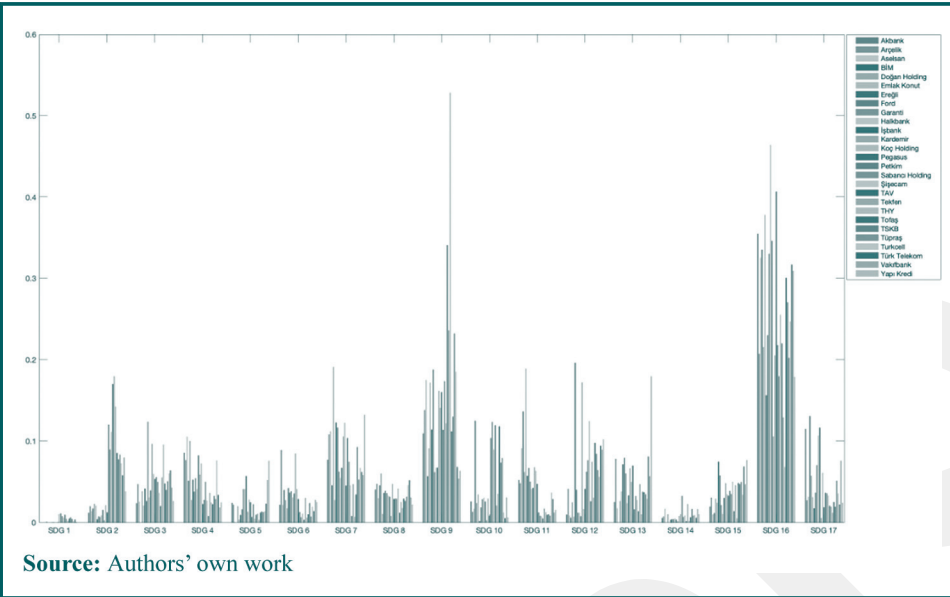
Then, as input to the model, the texts of the companies' annual reports were added. For reports from 27 firms, the methodology computed the SDGs' percentage of the relation. The results were combined according to the percentage of SDGs that each report's content related to. The SDGs ratios of firms achieving are evaluated to assess which SDGs are most

**Table 3** Accuracy of The Model

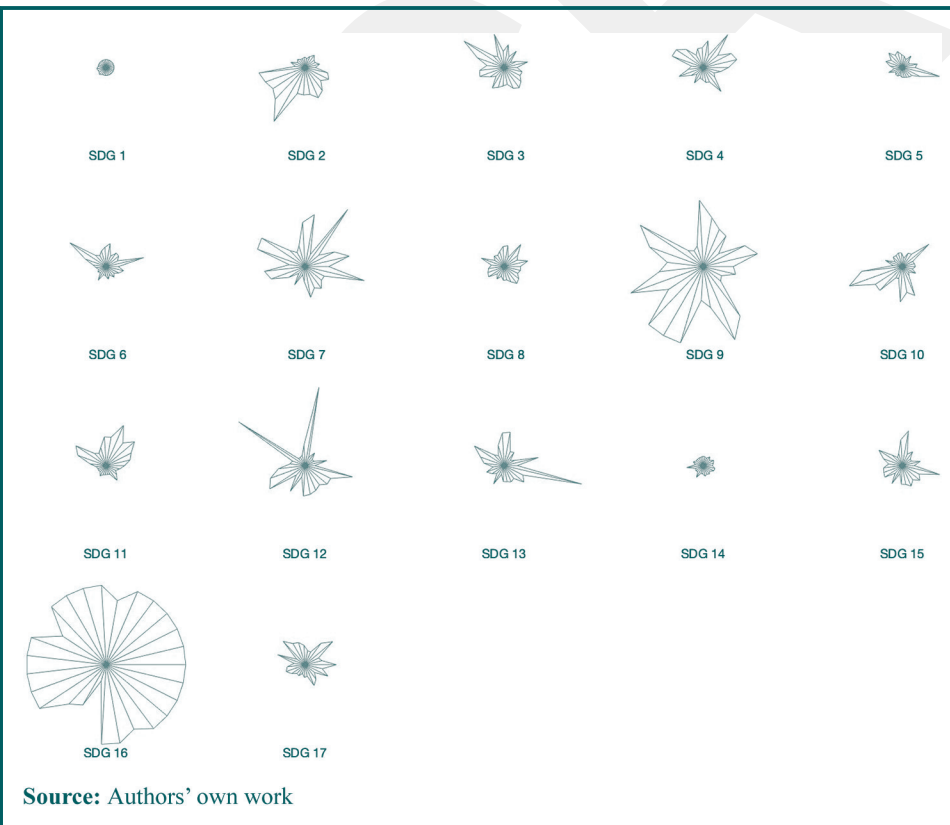
SDGs	Training data	Precision	Recall	f1-score	Accuracy
SDG 1 No poverty	1514	0.83	0.83	0.83	0.9656539
SDG 2 Zero hunger	1514	0.85	0.84	0.84	0.9537649
SDG 3 Good health and well-being	1514	0.89	0.87	0.88	0.9517834
SDG 4 Quality education	1514	0.88	0.92	0.90	0.9557464
SDG 5 Gender equality	1514	0.91	0.86	0.89	0.9630119
SDG 6 Clean water and sanitation	1514	0.94	0.85	0.89	0.9729194
SDG 7 Affordable and clean energy	1514	0.87	0.87	0.87	0.9669749
SDG 8 Decent work and economic growth	1514	0.73	0.73	0.73	0.9445178
SDG 9 Industry, innovation and infrastructure	1514	0.85	0.78	0.81	0.9623514
SDG 10 Reduced inequalities	1514	0,87	0,77	0.81	0,9649934
SDG 11 Sustainable cities and communities	1514	0.79	0.79	0.79	0.9451783
SDG 12 Responsible consumption and production	1514	0.85	0.75	0.79	0.9821664
SDG 13 Climate action	1514	0.87	0.86	0.87	0.9709379
SDG 14 Life below water	1514	0.91	0.87	0.89	0.9861295
SDG 15 Life on land	1514	0.85	0.88	0.86	0.9735799
SDG 16 Peace, justice and strong institutions	1514	0.79	0.78	0.78	0.9398943
SDG 17 Partnerships for the goals	1514	0.87	0.69	0.75	0.9689564

Source: Authors' own work

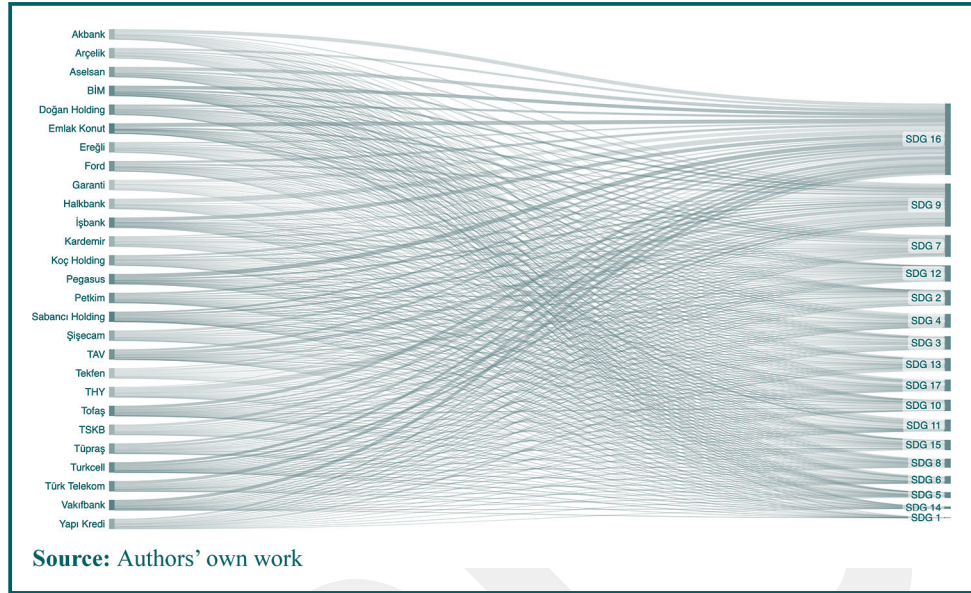
**Figure 1** SDGs performance of companies



**Figure 2** Demonstration of SDG's ratios



**Figure 3** Sankey diagram of BIST 30 companies to cover the SDGs

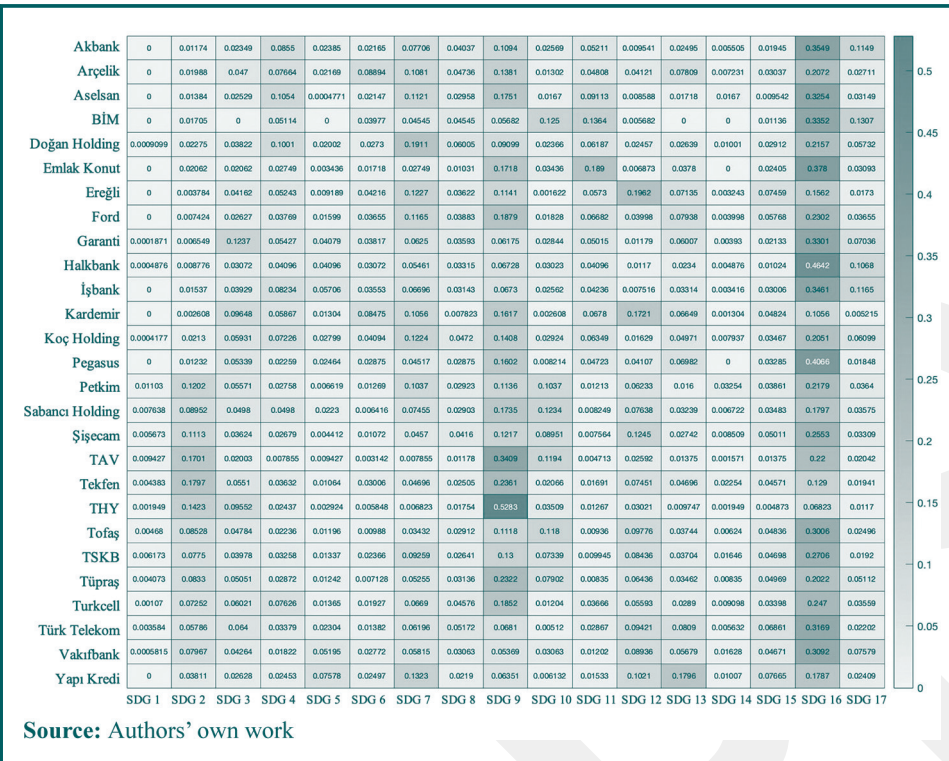


supported by firms' overall activity. Figure 1 shows the SDG fulfillment status of companies according to the reports. The demonstration of SDG's ratios is given in Figure 2. Figure 3 shows a Sankey diagram of BIST 30 Companies to cover SDGs.

In Figure 4, the ratios of 27 companies' activities to support each SDG are given in detail in numbers. Because the ratios of SDG 9 and SDG 16 goals are higher than the others, companies have carried out more activities that contribute to these targets. In the framework of sustainable development, SDG 16 (peace, justice and strong institutions) and SDG 9 (industry, innovation and infrastructure) have a beneficial relationship. SDG 16 strives to advance inclusive and peaceful communities, ensure everyone has access to justice and create inclusive institutions at all levels of government. SDG 9 is concerned with supporting innovation, encouraging inclusive and sustainable industry and creating resilient infrastructure. Together, these two objectives can contribute to the development of resilient and sustainable societies that value both social justice and technical innovation. For example, ensuring people have access to justice and institutions operate effectively helps create an atmosphere where infrastructure growth and innovation thrive without endangering people or the environment. Investments in infrastructure and innovation can also assist bolster institutions and increase access to justice. Highlighting the beneficial connection between SDG 16 and SDG 9 in an annual report may show a commitment to sustainable development, in terms of both social justice and technical advancement. It can also demonstrate how these objectives are related and can support one another in producing results for sustainable development.

The banking industry is one of a company's key stakeholders in putting long-term objectives into practice and achieving internationalization strategies that always adapt to the current global environment. By engaging in socially good initiatives that fall under the purview of SDGs, businesses can enhance their credibility with banks. As Having international partners would likely have an effect on the annual reports of BIST30 firms, including Akbank, BM, Ereli, Ford, Garanti, Kardemir, Koç Holding, Petkim, Şişecam, TAV, Tekfen, THY, Tofaş, Tüpraş, Turkcell and Yapı Kredi. Once the study findings are examined in the context of internationalization, it is evident that BIST30 enterprises with foreign partners report all SDGs more adequately, with the exception of SDGs 4, 5 and 11, which are less

**Figure 4** Ratios of BIST 30 companies to cover the SDGs



Source: Authors' own work

capably reported. While the difference is insignificant for SDGs 4 and 11, it can be claimed that SDG 5 is the only one that distinguishes out. It has been shown that BIST30 enterprises with international partners report on SDGs 1, 2, 10 and 12 more consistently than other SDGs.

According to Figure 3, it is seen that the distribution of SDG 1, SDG 5, SDG 6, SDG 8 and SDG 14 goals is relatively low, and it is essential to intensify efforts and activities toward these SDGs. It has been found that the relevant SDGs are appraised at a lower rate than the others in reporting, even though it may not follow that companies give these SDGs less importance.

This study uses a subjective assessment of the SDGs while taking into account the reports of commercial enterprises with a particular geographic focus. With the idea that the SDGs will help businesses, it engages in activities and reports these efforts (Van Zanten and Van Tulder, 2018). According to a study that analyzed the Web pages of 500 Fortune companies, the percentage of businesses that have adopted the SDGs framework is encouraging, but many of those adoptions are superficial. The most popular SDGs among corporations are SDG 8 and SDG 13. The continents of Europe, the Middle East and Africa are used as comparison points. Across all industries, the engagement rate with the overall SDGs framework is highest in the information and technology industry and lowest in the health-care sector. The SDGs offer possibilities for a society that is more habitable as well as opportunities for businesses to forge close ties and acquire a competitive edge (Song et al., 2022).

The SDGs provide a framework for analyzing corporate annual reports, which is extremely important because it provides insightful information about a company's contributions to sustainable development. Researchers can assess a company's alignment with the SDGs using the ESG practices detailed in annual reports. By supporting studies on the integration of sustainability practices, evaluating the impact of businesses on SDG targets and identifying the best strategies for achieving sustainable development (Eccles et al., 2014),

this analysis adds to the body of knowledge. Practically speaking, reviewing annual reports enables investors and stakeholders to evaluate a company's adherence to sustainable business practices and advancement toward SDG-related goals. This understanding aids in the decision-making process for investments and motivates businesses to enhance their ESG performance (Khan *et al.*, 2021; Grewal *et al.*, 2019). In addition, annual reports operate as a mechanism for accountability and transparency that enables society to hold businesses accountable for their social and environmental repercussions. Civil society organizations and the general public can assess a company's contributions to the SDGs, promoting dialogue and bringing about positive change by looking at yearly reports (Grewal *et al.*, 2019; Eccles *et al.*, 2014). In the end, analyzing annual reports of companies in the context of the SDGs advances research, directs sustainable business practices and encourages societal understanding and engagement in the pursuit of sustainable development.

The banking industry is one of a company's key stakeholders in putting long-term objectives into practice and achieving internationalization strategies that always adapt to the current global environment (Nilsson and Persson, 2021). By engaging in socially good initiatives that fall under the purview of SDGs, businesses can enhance their credibility with banks. Once the study findings are examined in the context of internationalization, it is evident that BIST30 enterprises with foreign partners report all SDGs more adequately, with the exception of SDGs 4, 5 and 11, which are less capably reported. While the difference is insignificant for SDGs 4 and 11, it can be claimed that SDG 5 is the only one that distinguishes out. It has been shown that BIST30 enterprises with international partners report on SDGs 1, 2, 10 and 12 more consistently than other SDGs.

## 5. Conclusion

This study presented a clear mapping from explanatory views to SDGs. All significant sustainability-related concerns need to be characterized uniformly, and annual reports may cover each one. As part of the execution of the SDGs, this research set out to provide proof of the factors that influence corporate nonfinancial reporting. Finding solution techniques for sustainable management is a significant difficulty in many industries. The findings of this article offer useful opportunities for researchers, collaborators from business institutions and organizations and policymakers for a variety of reasons. The degree to which each of the 17 SDGs is covered in the annual report provides insight into which topics are extensively covered and which topics receive little attention. Determining how to optimize the effect of annual reports and how to enhance the alignment between the industrial perspective and SDG implementation can be helpful for decision-makers.

Monitoring of the sustainability performance of corporations through annual reports is intended to represent the current situation of the country in accordance to the SDGs in this study. Text analyses based on multilabel classification were used as a promising methodology to identify the ratio of SDGs from the varied perspectives of various groups. In the study, a text analysis method based on an improved version of gradient boosting is used to examine the report and classify SDGs. To maximize the impact of yearly reports and improve the alignment between the industrial perspective and SDG implementation, more study, collaboration and policymaking is needed.

Companies may demonstrate their commitment to sustainability, adhere to rules, promote sustainable practices, increase transparency and accountability and strengthen their reputation and brand value by disclosing their SDG-related actions. It is anticipated that the study's findings will increase the existing body of knowledge in sustainability science, and that the solutions and key elements that have been identified will encourage policymakers and researchers to work together and perform more in-depth research. The opportunity to evaluate more data more quickly for comparable research is provided by the use of

machine learning enabled models. Furthermore, it avoids inconsistency problems that could occur during content analysis.

Overall, this study adds to the expanding body of knowledge on sustainable development and offers useful recommendations for businesses, policymakers and other stakeholders looking to assess and promote sustainable practices that are in line with the SDGs. This study promotes continued advancement toward a more sustainable future by addressing the positive and negative aspects of business sustainability reporting.

### Future Work

To acquire a more thorough picture of sustainability performance across industries and geographical contexts, future research might build on this analysis by incorporating a wider range of businesses and areas.

### References

- Al Amosh, H. and Khatib, S.F. (2021), "Corporate governance and voluntary disclosure of sustainability performance: the case of Jordan", *SN Business & Economics*, Vol. 1 No. 12, pp. 1-22.
- Baker, S. and Korhonen, A. (2017), "Initializing neural networks for hierarchical multi-label text classification", *BioNLP 2017*, pp. 307-315.
- Boiral, O. and Henri, J.F. (2017), "Is sustainability performance comparable? A study of GRI reports of mining organizations", *Business & Society*, Vol. 56 No. 2, pp. 283-317.
- Bordignon, F. (2021), "Dataset of search queries to map scientific publications to the UN sustainable development goals", *Data in Brief*, Vol. 34, p. 106731.
- Caratas, M.A., Andreea Trandafir, R., Iftene, C., Cerasela Spataru, E. and Gheorghiu, G. (2021), "The impact of sustainability disclosure on companies' performance in healthcare industry", *Transformations in Business & Economics*, Vol. 20, pp. 593-613.
- Chen, T. and Guestrin, C. (2016), "Xgboost: a scalable tree boosting system", *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785-794.
- Cosma, S., Venturelli, A., Schwizer, P. and Boscia, V. (2020), "Sustainable development and European banks: a non-financial disclosure analysis", *Sustainability*, Vol. 12 No. 15, p. 6146.
- Dhaliwal, S.S., Nahid, A.A. and Abbas, R. (2018), "Effective intrusion detection system using XGBoost", *Information*, Vol. 9 No. 7, p. 149.
- Eccles, R.G., Ioannou, I. and Serafeim, G. (2014), "The impact of corporate sustainability on organizational processes and performance", *Management Science*, Vol. 60 No. 11, pp. 2835-2857.
- Ehnert, I., Parsa, S., Roper, I., Wagner, M. and Muller-Camen, M. (2016), "Reporting on sustainability and HRM: a comparative study of sustainability reporting practices by the world's largest companies", *The International Journal of Human Resource Management*, Vol. 27 No. 1, pp. 88-108.
- Elalfy, A., Weber, O. and Geobey, S. (2021), "The sustainable development goals (SDGs): a rising tide lifts all boats? Global reporting implications in a post SDGs world", *Journal of Applied Accounting Research*, Vol. 22 No. 3, pp. 557-575.
- Elmitwally, N.S. (2020), "Building a multi-class XGBoost model for Arabic figurative language", 2020 2nd International Conference on Computer and Information Sciences (ICCIS), Sakaka, Saudi Arabia, pp. 1-4.
- Feldman, R. and Sanger, J. (2007), *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*, Cambridge University Press, Cambridge.
- Friedman, J.H. (2001), "Greedy function approximation: a gradient boosting machine", *Annals of Statistics*, Vol. 29 No. 5, pp. 1189-1232.
- Grewal, J., Riedl, E.J. and Serafeim, G. (2019), "Market reaction to mandatory nonfinancial disclosure", *Management Science*, Vol. 65 No. 7, pp. 3061-3084.
- Guisiano, J. and Chiky, R. (2021), "Automatic classification of multilabel texts related to sustainable development goals (SDGs)", *TECHENV EGC2021*.

- Gumus, M. and Kiran, M.S. (2017), "Crude oil price forecasting using XGBoost", *2017 International Conference on Computer Science and Engineering (UBMK)*, IEEE, pp. 1100-1103.
- Hajikhani, A. and Suominen, A. (2022), "Mapping the sustainable development goals (SDGs) in science, technology and innovation: application of machine learning in SDG-oriented artefact detection", *Scientometrics*, Vol. 127 No. 11, pp. 1-33.
- Hendrawan, I.R., Utami, E. and Hartanto, A.D. (2022), "Comparison of naïve Bayes algorithm and XGBoost on local product review text classification", *Edumatic: Jurnal Pendidikan Informatika*, Vol. 6 No. 1, pp. 143-149.
- Heras-Saizarbitoria, I., Urbieta, L. and Boiral, O. (2022), "Organizations' engagement with sustainable development goals: from cherry-picking to SDG-washing?", *Corporate Social Responsibility and Environmental Management*, Vol. 29 No. 2, pp. 316-328.
- Hummel, K. and Szekely, M. (2022), "Disclosure on the sustainable development goals – evidence from Europe", *Accounting in Europe*, Vol. 19 No. 1, pp. 152-189.
- Hwang, H., An, S., Lee, E., Han, S. and Lee, C.H. (2021), "Cross-societal analysis of climate change awareness and its relation to SDG 13: a knowledge synthesis from text mining", *Sustainability*, Vol. 13 No. 10, p. 5596.
- İş, H. and Tuncer, T. (2019), "Interaction-based behavioral analysis of twitter social network accounts", *Applied Sciences*, Vol. 9 No. 20, p. 4448.
- Izzo, M.F., Ciaburri, M. and Tiscini, R. (2020), "The challenge of sustainable development goal reporting: the first evidence from Italian-listed companies", *Sustainability*, Vol. 12 No. 8, p. 3494.
- Jang, J.H. (2021), "Analysis of important policy issues relating to affirmative action in education: application of text mining method", *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, Vol. 12 No. 13, pp. 1573-1589.
- Juventia, S.D., Jones, S.K., Laporte, M.A., Remans, R., Villani, C. and Estrada-Carmona, N. (2020), "Text mining national commitments towards agrobiodiversity conservation and use", *Sustainability*, Vol. 12 No. 2, p. 715.
- Khan, P.A., Johl, S.K. and Akhtar, S. (2021), "Firm sustainable development goals and firm financial performance through the lens of green innovation practices and reporting: a proactive approach", *Journal of Risk and Financial Management*, Vol. 14 No. 1, p. 605.
- LaFleur, M. (2019), "Art is long, life is short: an SDG classification system for DESA publications".
- Lee, P., Kleinman, G. and Kuei, C.H. (2020), "Using text analytics to apprehend urban sustainability development", *Sustainable Development*, Vol. 28 No. 4, pp. 897-921.
- Liu, J., Chang, W.C., Wu, Y. and Yang, Y. (2017), "Deep learning for extreme multi-label text classification", *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 115-124.
- López-Santamaría, M., Amaya, N., Hinestroza, M.P.G. and Cuero, Y.A. (2021), "Sustainability disclosure practices as seen through the lens of the signaling theory: a study of companies listed on the Colombian stock exchange", *Journal of Cleaner Production*, Vol. 317, p. 128416.
- Ludwig, S.A. (2019), "Applying a neural network ensemble to intrusion detection", *Journal of Artificial Intelligence and Soft Computing Research*, Vol. 9 No. 3, pp. 177-178.
- Martínez-Ferrero, J. and García-Meca, E. (2020), "Internal corporate governance strength as a mechanism for achieving sustainable development goals", *Sustainable Development*, Vol. 28 No. 5, pp. 1189-1198.
- Matsui, T., Suzuki, K., Ando, K., Kitai, Y., Haga, C., Masuhara, N. and Kawakubo, S. (2022), "A natural language processing model for supporting sustainable development goals: translating semantics, visualizing nexus, and connecting stakeholders", *Sustainability Science*, Vol. 17 No. 3, pp. 969-985.
- Meier, D.S., Mata, R. and Wulff, D.U. (2021), "text2sdg: an open-source solution to monitoring sustainable development goals from text", *arXiv preprint arXiv:2110.05856*.
- Mitchell, R. and Frank, E. (2017), "Accelerating the XGBoost algorithm using GPU computing", *PeerJ Computer Science*, Vol. 3, p. e127.
- Moldan, B., Janoušková, S. and Hák, T. (2012), "How to understand and measure environmental sustainability: indicators and targets", *Ecological Indicators*, Vol. 17, pp. 4-13.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A. and Giovannini, E. (2015), *Handbook on Constructing Composite Indicators: Methodology and User Guide*, In OECD Statistics Working Paper, No. 2005/03 OECD Publishing, Paris.

Nielsen, D. (2016), "Tree boosting with XGBoost – why does XGBoost win 'every' machine learning competition?", Master's thesis, NTNU.

Nilsson, J. and Persson, M. (2021), "The role of banking relationships in SMEs' internationalization: a systematic literature review", *International Business Review*, Vol. 30 No. 4, p. 101857, doi: [10.1016/j.ibusrev.2020.101857](https://doi.org/10.1016/j.ibusrev.2020.101857).

Ogunleye, A. and Wang, Q.G. (2019), "XGBoost model for chronic kidney disease diagnosis", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol. 17 No. 6, pp. 2131-2140.

Pizzi, S., Rosati, F. and Venturelli, A. (2021), "The determinants of business contribution to the 2030 agenda: introducing the SDG reporting score", *Business Strategy and the Environment*, Vol. 30 No. 1, pp. 404-421.

Qi, Z. (2020), "The text classification of theft crime based on TF-IDF and XGBoost model", *2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, IEEE, pp. 1241-1246.

Roy, A., Basu, A., Su, Y., Li, Y. and Dong, X. (2022), "Understanding recent trends in global sustainable development goal 6 research: scientometric, text mining and an improved framework for future research", *Sustainability*, Vol. 14 No. 4, p. 2208.

Schütze, H., Manning, C.D. and Raghavan, P. (2008), *Introduction to Information Retrieval*, Vol. 39, Cambridge University Press, Cambridge, pp. 234-265.

Sebestyén, V., Domokos, E. and Abonyi, J. (2020), "Focal points for sustainable development strategies – text mining-based comparative analysis of voluntary national reviews", *Journal of Environmental Management*, Vol. 263, p. 110414.

Song, L., Zhan, X., Zhang, H., Xu, M., Liu, J. and Zheng, C. (2022), "How much is global business sectors contributing to sustainable development goals?", *Sustainable Horizons*, Vol. 1, p. 100012.

Sovrano, F., Palmirani, M. and Vitali, F. (2020), "Deep learning based multi-label text classification of UNGA resolutions", *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance*, pp. 686-695.

Tantuğ, A.C. (2016), "Metin sınıflandırma", *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, Vol. 5 No. 2.

Te Liew, W., Adhitya, A. and Srinivasan, R. (2014), "Sustainability trends in the process industries: a text mining-based analysis", *Computers in Industry*, Vol. 65 No. 3, pp. 393-400.

Tóth, Á., Suta, A. and Szauter, F. (2022), "Interrelation between the climate-related sustainability and the financial reporting disclosures of the European automotive industry", *Clean Technologies and Environmental Policy*, Vol. 24 No. 1, pp. 437-445.

Van Zanten, J.A. and Van Tulder, R. (2018), "Multinational enterprises and the sustainable development goals: an institutional approach to corporate engagement", *Journal of International Business Policy*, Vol. 1 Nos 3/4, pp. 208-233.

Virto, L.R. (2018), "A preliminary assessment of the indicators for sustainable development goal (SDG) 14 "conserve and sustainably use the oceans, seas and marine resources for sustainable development", *Marine Policy*, Vol. 98, pp. 47-57.

Wang, X., Yuen, K.F., Wong, Y.D. and Li, K.X. (2020), "How can the maritime industry meet sustainable development goals? An analysis of sustainability reports from the social entrepreneurship perspective", *Transportation Research Part D: Transport and Environment*, Vol. 78, p. 102173.

Zhou, Y., Wang, X. and Yuen, K.F. (2021), "Sustainability disclosure for container shipping: a text-mining approach", *Transport Policy*, Vol. 110, pp. 465-477.

## Further reading

Khan, M.A. (2022), "ESG disclosure and firm performance: a bibliometric and meta analysis", *Research in International Business and Finance*, Vol. 61, p. 101668.

Luo, X. (2017), "A new text classifier based on random forests", *2016 2nd International Conference on Materials Engineering and Information Technology Applications (MEITA 2016)*, Atlantis Press, pp. 290-293.

Rodríguez Medina, S. (2019), "Multi-label text classification with transfer learning for policy documents: the case of the sustainable development goals", Doctoral Thesis.

Shirani-Mehr, H. (2013), *SMS Spam Detection Using Machine Learning Approach*, H. Shirani-Mehr, Stanford.

## Link

Borsa İstanbul (2018), "Bist Pay Endeksleri Temel Kuralları", Borsa İstanbul, Deloitte Türkiye, İstanbul.

Deloitte ve TKYD (2014), Anonim ve Limited Şirketler İçin Yıllık Faaliyet Raporu Hazırlama Kılavuzu.

Ministry of Development of the Republic of Turkey (2019), "11th development plan (2019–2023)", Retrieved from, available at: [www.sbb.gov.tr/wp-content/uploads/2022/07/Eleventh\\_Development\\_Plan\\_2019-2023.pdf](http://www.sbb.gov.tr/wp-content/uploads/2022/07/Eleventh_Development_Plan_2019-2023.pdf)

OSDG, UNDP IICPSD SDG AI Lab, & PPMI (2022), "OSDG community dataset (OSDG-CD) (2022.07) [data set]", Zenodo.

Sustainable Development Goals Evaluation Report (2016-2019), available at: [www.sbb.gov.tr/wp-content/uploads/2022/07/Eleventh\\_Development\\_Plan\\_2019-2023.pdf](http://www.sbb.gov.tr/wp-content/uploads/2022/07/Eleventh_Development_Plan_2019-2023.pdf)

United Nations Economic and Social Council (2019), "Voluntary national review of Turkey", Retrieved from, available at: [https://sustainabledevelopment.un.org/content/documents/23862TC\\_BUTCE\\_BAKANLIGI\\_RAPOR\\_09\\_EN.pdf](https://sustainabledevelopment.un.org/content/documents/23862TC_BUTCE_BAKANLIGI_RAPOR_09_EN.pdf)

UN (2015), *Transforming our world: the 2030 agenda for sustainable development* United Nations, New York, NY.

## Further reading

<https://doi.org/10.5281/zenodo.6831287>

<https://github.com/Marcinter96/SDG-s-classification>

<https://github.com/SURYA-LAMICHANEY/UN-SDG-Classfier>

The Presidency of the Republic of Turkey (2019), "Sustainable Development Goals Evaluation Report 2016-2019", available at: [www.sbb.gov.tr/wp-content/uploads/2020/03/Surdurulebilir-Kalkinma-Amaclari-Degerlendirme-Raporu\\_13\\_12\\_2019-WEB.pdf](http://www.sbb.gov.tr/wp-content/uploads/2020/03/Surdurulebilir-Kalkinma-Amaclari-Degerlendirme-Raporu_13_12_2019-WEB.pdf)

## Corresponding author

Ömer Faruk Ünlüsoy can be contacted at: [omerfaruk.unlusoy@agu.edu.tr](mailto:omerfaruk.unlusoy@agu.edu.tr)

---

For instructions on how to order reprints of this article, please visit our website:

[www.emeraldgroupublishing.com/licensing/reprints.htm](http://www.emeraldgroupublishing.com/licensing/reprints.htm)

Or contact us for further details: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)