Evaluation of UN SDG-related formal learning activities in a university common core curriculum

Chi-Un Lei
Teaching and Learning Innovation Centre, The University of Hong Kong, Hong Kong, China

Wincy Chan
Common Core Office, The University of Hong Kong, Hong Kong, China and Department of Pathology, The University of Hong Kong, Hong Kong, China, and

Yuyue Wang
The University of Hong Kong, Hong Kong, China

Received 12 February 2023
Revised 15 August 2023
3 November 2023
Accepted 7 November 2023

Abstract

Purpose – Higher education plays an essential role in achieving the United Nations sustainable development goals (SDGs). However, there are only scattered studies on monitoring how universities promote SDGs through their curriculum. The purpose of this study is to investigate the connection of existing common core courses in a university to SDG education. In particular, this study wanted to know how common core courses can be classified by machine-learning approach according to SDGs.

Design/methodology/approach – In this report, the authors used machine learning techniques to tag the 166 common core courses in a university with SDGs and then analyzed the results based on visualizations. The training data set comes from the OSDG public community data set which the community had verified. Meanwhile, key descriptions of common core courses had been used for the classification. The study used the multinomial logistic regression algorithm for the classification.

The authors would like to thank Mr Jeremy T.D. Ng, Professor Gray Kochhar-Lindgren, Dr Jack Tsao and Ms Xinyi Liang for their time and effort on commenting the study. The comments have led to a substantial improvement on the quality of the report. The authors also want to express their gratitude to Ms Xiaodong Hu and Ms. Xue Qian for their help in training data collection and modeling. In addition, we highly appreciate Mr Chun Yin Cham for their support in sending the teachers' surveys to the teaching staff.

Geolocation information: Pokfulam Road, Hong Kong.

Disclosure statement: The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding details: This research was supported by the University Grants Committee (UGC) of Hong Kong under the UGC Special Grant for Strategic Development of Virtual Teaching and Learning and UGC Teaching Development and Language Enhancement Grant.

Statement on ethics: Informed consent was obtained from all study participants. All collected data were treated confidentially. This research has been approved by the Human Research Ethics Committee of the University of Hong Kong (EA210317).

Data availability statement: Coding and data for classifications can be found in the Github repository (https://github.com/HKU-SDG-Classification/). Results of teachers' surveys can be provided upon request.

Data disposition: Coding and data for classifications can be found in the Github repository (https://github.com/HKU-SDG-Classification/).

International Journal of Sustainability in Higher Education

1467-6370
DOI 10.1108/IJSHE-02-2023-0050

Descriptive analysis at course-level, theme-level and curriculum-level had been included to illustrate the proposed approach's functions.

Findings – The results indicate that the machine-learning classification approach can significantly accelerate the SDG classification of courses. However, currently, it cannot replace human classification due to the complexity of the problem and the lack of relevant training data.

Research limitations/implications – The study can achieve a more accurate model training through adopting advanced machine learning algorithms (e.g. deep learning, multioutput multiclass machine learning algorithms); developing a more effective test data set by extracting more relevant information from syllabus and learning materials; expanding the training data set of SDGs that currently have insufficient records (e.g. SDG 12); and replacing the existing training data set from OSDG by authentic education-related documents (such as course syllabus) with SDG classifications. The performance of the algorithm should also be compared to other computer-based and human-based SDG classification approaches for cross-checking the results, with a systematic evaluation framework. Furthermore, the study can be analyzed by circulating results to students and understanding how they would interpret and use the results for choosing courses for studying. Furthermore, the study mainly focused on the classification of topics that are taught in courses but cannot measure the effectiveness of adopted pedagogies, assessment strategies and competency development strategies in courses. The study can also conduct analysis based on assessment tasks and rubrics of courses to see whether the assessment tasks can help students understand and take action on SDGs.

Originality/value – The proposed approach explores the possibility of using machine learning for SDG classifications in scale.

Keywords Sustainable development goals, SDGs, Sustainable development education classification, Curriculum analysis, Machine-learning classification, Classification

Paper type Research paper

1. Introduction

1.1 United Nations sustainable development goals

In 2015, the United Nations (UN) established 17 sustainable development goals (SDGs) targeting sustainable economic growth and social development by the year of 2030. The 17 SDGs cover five critical sustainability areas (see Appendix): people, planet, prosperity, peace and partnership (United Nations General Assembly, 2015). The aim of the goals is to end poverty, promote all factors that make the planet more habitable and make sure every person can enjoy an equitable, peaceful and prosperous life (Morton *et al.*, 2017; Rieckmann, 2017). Although the 17 goals have different targets and indicators, it is essential to understand that they are intimately interconnected. The interlinkages among the goals show a well-consulted, scientifically robust, politically acceptable and publicly intuitive framework. Therefore, different approaches and technologies, including artificial intelligence (Allen *et al.*, 2021; Vinuesa *et al.*, 2020; Porciello *et al.*, 2020), have been adopted to promote and achieve the SDGs.

1.2 Sustainable development goal education in higher education institutions

Higher education plays an essential role in promoting and achieving the SDGs. Higher education institutions (HEIs) are significant influencers and agents of systematically and sustainably transforming societies and serving the greater public good (Waas *et al.*, 2010). In this sense, HEIs should be responsible for enabling students to consider issues raised by sustainability and guide them to search for solutions and alternatives to the problems (Hesselbarth and Schaltegger, 2014). Once students know more about the SDGs, they can develop their knowledge, value, belief and perspectives on sustainable development, which can help them develop the skills to solve emerging world problems. Embedding the SDGs within and across the curriculum will not only enhance human capital but also stimulate students to take action to live sustainably, which can ultimately help to construct a

sustainable society (Leal Filho *et al.*, 2018). By 2030, the UN envisions all learners acquire the knowledge and skills needed to promote sustainable development (Target 4.7) (Leicht *et al.*, 2018). As such, the HEIs must include topics related to global citizenship and sustainable development in their course curricula.

Several studies have been conducted on monitoring how universities achieve the SDGs via campus infrastructure development (Omazic and Zunk, 2021) and research (Bautista-Puig *et al.*, 2021; Goodall and Moore, 2019). However, only scattered studies have reported on how universities promote the SDGs through such activities explicitly:

- Aleixo *et al.* (2020) examined the implementation of the SDGs in undergraduate and master's courses of 33 Portuguese public HEIs using content analysis. The data was collected in the 2017/2018 school year. The authors found that the most frequent SDGs covered in the Portuguese university courses were SDG 15 (Life on land) and SDG 7 (Affordable and clean energy). Moreover, only a small number of university courses directly addressed the SDGs. The authors claimed that this may be because Portuguese HEIs do not have sufficient time to respond to the UN publication of the SDGs in 2015.
- University College London (UCL) reported the classification of the SDGs in their undergraduate taught modules in the UCL SDGs Report 2020–2021 (UCL, 2021). The results showed that most of the taught modules focused on SDG 3 (good health and well-being) and SDG 11 (sustainable cities and communities). For example, in response to the COVID-19 pandemic in the UK, UCL's final-year medical students were deployed to the frontline and took up the post of a Foundation Interim Year 1 doctor, which further promoted SDG 3 (e.g. Indicator 3.3, "Ending the epidemics of communicable diseases").
- The University of South Florida promoted the SDGs at the undergraduate level through the Global Citizens Project (Hansen *et al.*, 2021). In this project, students signed up for a one-week, cocurricular experience focusing on global issues related to the 17 SDGs. They could also take Global Citizens Courses that provide the understanding, skills and values students need to cooperate in resolving the interconnected challenges of the 21st century, such as climate change, poverty, hunger and equity.

It could be challenging to measure and monitor how the SDGs are taught in undergraduate courses in a systematic manner. In higher education, globally, embedding the SDGs into the curriculum may be encouraged but not often statutory, making any relevant teaching and learning initiatives hard to trace. At the course level, teachers may design courses that illustrate topics related to the SDGs or sustainability development, without making direct references to the SDG policy documents from the UN. Furthermore, it is too much of a demand to expect individual teachers to understand the 17 SDGs that cover multiple disciplines and implement them accurately and effectively into their courses. Meanwhile, at the program level, it can be time-consuming for curriculum consultants to analyze every piece of course content in hundreds or thousands of courses. It can also be unrealistic to achieve the measurable indicators targeted by the SDGs in a semester course, for example, making a significant difference in the prevalence of undernourishment (Indicator 2.1.1).

Since 2010, researchers have started to adopt technologies for machine-based SDG classification. The majority of previous efforts to classify the 17 SDGs fall into the following four main approaches, i.e. ontology building, supervised machine-learning, unsupervised

machine-learning and a combination of the above approaches. First, for ontology building, Bautista-Puig and Mauleón (2019) constructed a data set from a corpus of documents that had attached SDG labels, and then they built an otology of keywords for each SDG based on the keywords extracted from the data set and the definitions of SDGs. Second, the utilization of supervised machine learning, which works by taking a set of input data together with the output labels and using algorithms to train a model that can be used to make predictions on new unseen data. For instance, Sovrano *et al.* (2020) created a deep learning model based on several terms in SDG-related documents and used the model to detect UN General Assembly resolutions implicitly related to SDGs. Third, unsupervised machine learning represents an approach that is a middle ground between the two previous approaches. Although it relies on dedicated algorithms to identify obvious features in data, it does not predict a specific label, instead, it finds the underlying structure or clusters in the data. An example of such effort is OECD (2020) developed the SDG Pathfinder project to create a meaningful list of topics linked to SDGs as well as a model to detect these topics in new text, thus linking new documents to SDGs. Finally, the hybrid approaches combine elements from the previous three approaches. An example of this is Nakamura *et al.* (2019) utilized both ontology building and unsupervised machine learning approaches to identify a set of SDG papers.

In addition to researchers' works on SDGs classification, UCL classified its 6,113 taught modules an open-access tool (OSDG, UNDP IICPSD SDG AI Lab and PPMI, 2021) with a set of SDG-relevant terminology (UCL, 2021), and this study was based on OSDG to build the training model. The OSDG tool searched for keywords in the module descriptions and tagged an SDG to them if the descriptions contained two or more keywords for that SDG. The approach mainly relied on the existence of SDG keywords in the descriptions, but the tool cannot cover all SDG-relevant terminologies and their variations. The tool cannot classify multidisciplinary courses, which likely simultaneously teach multiple SDGs.

1.3 Manuscript organizations

Technical details of the computer-based SDG classifications are discussed in Section 2. Section 3 discusses the findings from the exploratory study and the comparison with the human-based classifications. Findings of computer-based and human-based classifications are shown to teachers, and their feedback is shown in Section 4. Section 5 discusses how the framework can be adopted and elaborated in the future.

2. Objectives and research questions

The main purpose of this study was to investigate the connection between common core courses and SDG-related education in a comprehensive research university in Hong Kong, where sustainable development education for undergraduate students has been attached to greater importance. The common core facilitates the teaching of the trans- and interdisciplinary undergraduate curriculum offered by all 10 faculties of the university. As the "crossroads" of the undergraduate curriculum, the core provides the key holistic learning experience for all undergraduates in this university, so it is necessary to incorporate SDGs-related teaching content into the courses and examine how many SDGs have been involved in these courses.

By analyzing the extent of coverage of various SDGs in the courses, it would then be possible to adjust courses to ensure a well-rounded SDG curriculum is provided to students. Through the investigation, instructors and university administrators can understand more about how much or how well courses in the university have been aligned with the SDGs. They can then be more informed to develop and evaluate courses and mini-curriculum (such as thematic clusters or micro-majors) considering the SDGs. Students can also benefit by

selecting relevant core courses for studying based on their interests and ambitions in sustainable development.

The research questions are as follows:

RQ1. How can common core courses be classified according to the SDGs, through machine learning algorithms?

RQ2. How have the SDGs been currently taught in the university common core courses, from a course, curriculum and theme cluster perspective?

RQ3. From teachers' perspectives, what are the differences between human SDG classifications and machine SDG classifications?

3. Computer-based classifications: research method

The studied university is a public research university in East Asia, with nearly 30,000 students. Its common core office (CC, or general education office) is committed to responding purposefully to the challenges outlined by SDGs through its teaching and learning experiences. In particular, the common core office addresses the SDGs through our comprehensive undergraduate courses taught by the 10 faculties, SDG competitions (such as the SDG game and photo competition) and transdisciplinary research opportunities for students. The office also collaborates with internal units (e.g. libraries, e-learning development center) and external units (e.g. Hong Kong Sustainability Campus Consortium) on promoting SDGs in the university learning environment.

In the academic year 2021–2022, the common core office ran 167 core courses. One of the courses had not submitted a detailed course syllabus and hence was excluded from this study. As this is a new initiative, the study did not have sufficient data to conduct model training. Therefore, we used a public policy-related data set as the training set (Section 3.1). The trained model is then used to evaluate the common core data set (Section 3.2).

3.1 Training data collection and model training

The training data set came from the OSDG community data set (OSDG, UNDP IICPSD SDG AI Lab and PPMI, 2021). The data set was based on publicly available documents, including reports, policy documents and publication abstracts. These documents were mainly from the UN and often already have SDG labels associated with them. The OSDG community extracted records from these documents. Currently, there are around 32,000 records of text comprised of three to six sentences. More than 1,000 community volunteers then validated the records on the relevance to SDGs. The data set only includes SDGs between 1 and 15 because SDGs 16 and 17 are overarching goals that might pop up in almost all kinds of texts. Furthermore, classified records are not evenly distributed across SDGs, ranging from hundreds (SDG 12) to thousands (SDG 4).

Before the training process, the study extracted and selected features as well as performed classifications. For feature extraction, the training set was transformed into a vector matrix representing term frequency – inverse document frequency values. For feature selection, Scikit-learn, a useful library for machine learning in Python, was used. The Scikit-learn package was used for the classification with the following setting:

- "newton-cg" was used for the solver.
- A penalty parameter l2 had been included to avoid overfitting.

- Because the amount of training data in each category was unbalanced, the study set the weighting parameter to "balanced" to increase the weight of categories with fewer data.

For ease of computations, only the top 5,000 features were retained for the model. After classifications, an SDG probability (relative relevance score) distribution from SDG 1 to SDG 15 will be calculated for each course.

3.2 Test data collection and model application

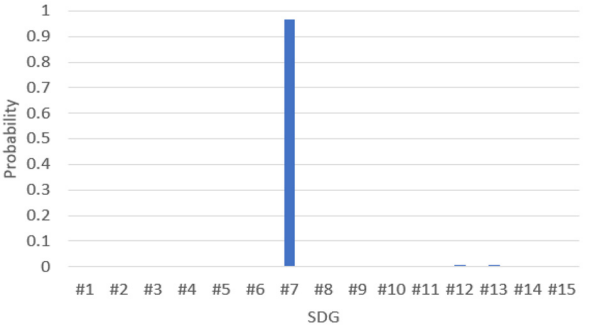
In response to the United Nations' call for action, the common core office began to assign the SDGs to all of its undergraduate courses to make sustainable development a more visible concept in its curriculum. The assignment of SDGs was conducted by an independent review of each of the courses by a curriculum expert, who is familiar with both the common core curriculum structure and the SDGs. The course outline (including the weekly topics, learning activities, assessment tasks and required readings, where available) was examined carefully and compared against the UNESCO teaching development guide for sustainable development (Rieckmann, 2017). The SDGs were assigned to each of the courses accordingly, with a maximum of four goals (the most predominant) to be posted on the course webpage. In the 2021–2022 academic year, the mapping of 166 common core courses was completed. Each course was tagged to 2.19 SDGs on average (range: 0–4 SDGs). The mapping was used as the test data in this study.

Meanwhile, the study obtained the syllabus of 166 courses from the common core office to extract the data set. Although there was no rigid structure for each course syllabus, the study extracted course overview, suggested reading list and weekly topics of each course for the classification. Meanwhile, learning outcomes and assessment activities were not extracted because those items contained too many pedagogy terminologies (e.g. "education," "curriculum," "learning" and "lessons"), which can misclassify courses to SDG Goal 4 "Quality Education."

4. Computer-based classifications: results

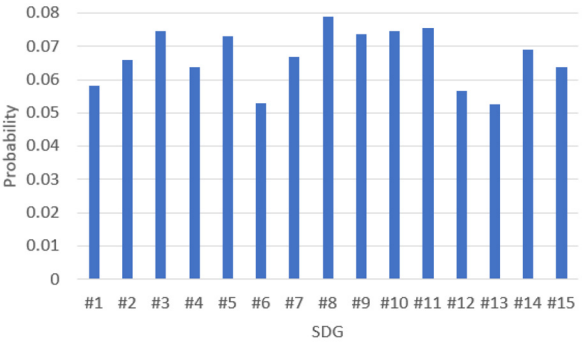
4.1 Descriptive analysis: courses

Figures 1 and 2 show the calculated classification SDG probabilities in two courses. For CCST9016 (Energy: Its Evolution and Environmental Impacts) in Figure 1, the calculated probabilities of this course being tagged to SDG 7 (affordable and clean energy) and other



Source: Authors' own creation

Figure 1. Calculated probability of CCST9016



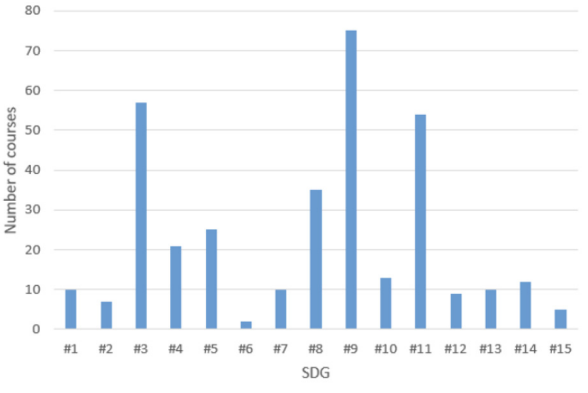
Source: Authors' own creation

Figure 2. Calculated probability of CCHU9068

SDGs are about 0.9689, and nearly 0, respectively. So, this course should be tagged to SDG 7 only. Meanwhile, from the course title, it is intuitive that the course's teaching content is about SDG 7. On the other hand, for CCHU9068 (Shaping our World: Liberalism, Socialism and Nationalism) in Figure 2, the probability (SDG relevance score) of this course is relatively average for each SDG. Therefore, the course cannot be classified to any SDGs (SDGs 1–15). This judgment is also aligned with the human judgment ("cannot be classified to one of the SDGs [SDGs 1–15]"). In general, because courses are often multidisciplinary, courses can be tagged with multiple SDGs.

4.2 Descriptive analysis: theme clusters and overall curriculum

Instead of using SDG probabilities for the analysis, we selected "Probability of the SDG > 0.079" as the criteria for classifying whether the course is teaching the particular SDG. With these criteria, the algorithm tagged 2.29 SDGs per course (Human: 2.19 SDGs on average). Figure 3 shows the distribution of taught SDGs in courses in the overall common core curriculum through the classification. Among courses, SDG 3 (Good Health and Well-being), SDG 9 (Industry, Innovation and Infrastructure) and SDG 11 (Sustainable Cities and Communities) are the most mentioned SDGs in courses. The distribution is expected because teachers aimed to teach topics addressing contemporary social issues and personal growth.



Source: Authors' own creation

Figure 3. The distribution of taught SDGs in courses

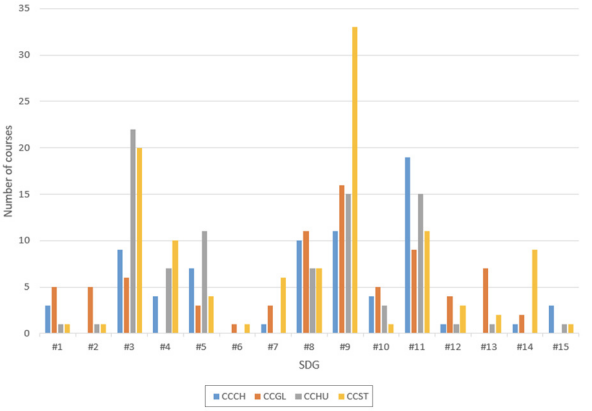
The Faculty of Medicine offered a significant number of medical health courses. Meanwhile, humanity teachers also offered courses that teach about well-being and mental health. On the other hand, SDG 6 (Clean water and sanitation) and SDG 15 (Life on Land) are the least two mentioned SDGs in courses.

Figure 4 shows the distribution of the taught SDGs in courses through the classification, according to their "Area of Inquiry" (AoI, theme cluster). Besides the common theme aforementioned, the figure shows different AoI focus on teaching several SDGs. For example, most courses in "China" AoI and "Global Issues" AoI discussed SDG 8 (Decent Work and Economic Growth). Meanwhile, most courses in "Arts and Humanities" AoI covered topics in SDG 5 (Gender Quality). "Science, Technology and Big Data" AoI strongly focuses on SDG 9 (Industry, Innovation and Infrastructure).

The study also conducted a hierarchical agglomerative clustering for each AoI and the CC curriculum to examine the similarity between courses. We used their SDG probability for profiling each course of performing the clustering. Dendrograms of all courses in CC and courses in China AoI are shown in Figure 5(a) and (b), respectively. The figure shows that there are courses that teach similar SDG contents (courses that have a smaller difference ("a shorter distance") from each other). Meanwhile, some courses differ from other CC courses based on the SDG probability. For example, in the China AoI, most courses in the China AoI often focused on the development of Chinese society. Meanwhile, there were three courses that were different from the others. For example, CCCH9039 "Curing the Chinese: Medicine and Society in Modern China" focused on Chinese medicine, rather than the China society. This indicates the diversity of courses in the CC curriculum.

4.3 Evaluating the performance of the algorithm: curriculum level

The study evaluated the performance of machine learning (ML) classifications by comparing results from human classifications, with the assumption of the human classification being the model standard. Because the training data set and test data set were from different origins, the algorithm's performance cannot be evaluated via cross-validation. The study evaluated the performance of the algorithm from an SDG perspective through different criteria (Note: Each course had been tagged to 2.19 SDGs on average by human.):



Source: Authors' own creation

Figure 4.
The distribution of taught SDGs in courses according to their AoI

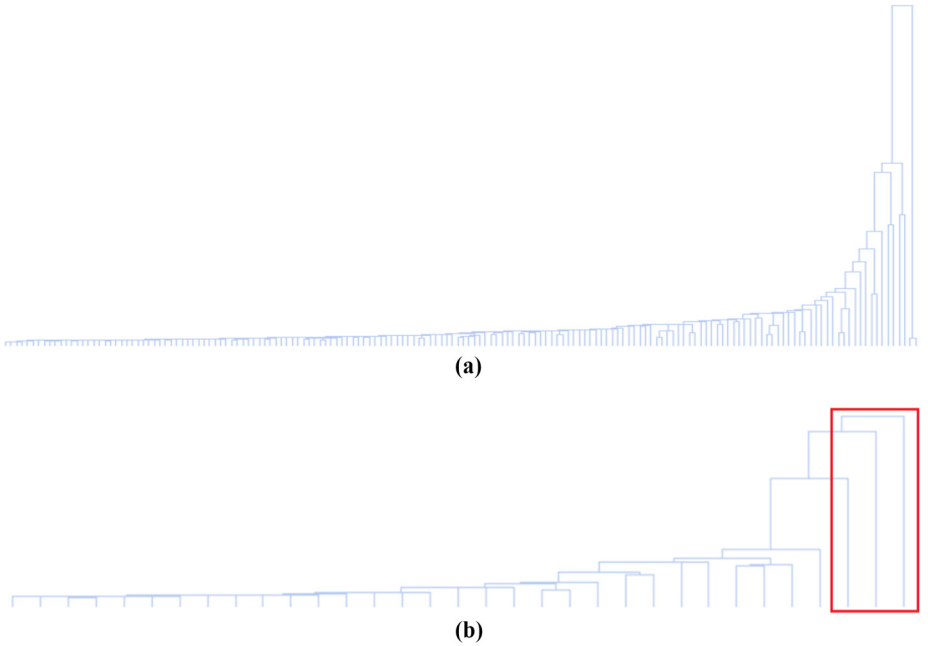


Figure 5.
The dendrograms of clusters of courses in terms of their SDG probability

Notes: (a) Overall; (b) China AoI
Source: Authors' own creation

- If each course is tagged to one SDG by the algorithm (i.e. choosing the SDG with the highest probability), the overall accuracy and F1-measure are 0.8855 and 0.4174, respectively.
- If each course is tagged to two SDGs by the algorithm (i.e. choosing two SDGs with the two highest probabilities), the overall accuracy and F1-measure are 0.8638 and 0.4621, respectively.
- If each course is tagged to SDGs with its "class probability > 0.079," the algorithm has tagged 2.29 SDGs per course. The overall accuracy and F1-measure are 0.8667 and 0.4711, respectively. Results show that 54 courses (32.53% of total courses) had been tagged with more than 2 SDGs.

In general, the obtained F1-measure of machine classification is relatively low. Teaching multiple SDGs in a course is common, as CC courses are often multidisciplinary. The complicated situation makes the classification challenging. Furthermore, the context of the OSDG training data set, which mainly were documents from the UN, may not be ideally the same as the style and context of the course syllabus (test data set). The generalizability of adopting a developed machine-learning model to a new population with a different context also made the problem challenging (Baker, 2019). Meanwhile, some teachers also used metaphoric expressions or ambiguous descriptions in course documents, which may not be easily interpreted by the machine. Based on teachers' feedback, we observed that the manual SDG classification results may not be objectively accurate.

The performance of the proposed approach was also compared to an ontology-based classification approach (Pukelis *et al.*, 2020), which had been adopted by the UCL (UCL, 2021).

The overall accuracy and F1-measure of the ontology-based classification approach are 0.8474 and 0.3705, respectively. The result indicates the limitation of only using keywords to classify SDG taught in courses.

4.4 Evaluating the performance of the algorithm: sustainable development goals level

Table 1 shows the performance of the algorithm if each course is tagged to SDGs with its "Class probability > 0.079." Based on the results, we have several observations:

- Tagging on SDGs 3, 4, 8, 9, 10 and 11 has an accuracy lower than average. In particular, more courses were tagged with SDG 9 by the algorithm than humans. We observed that course documents excessively emphasized the relevance of innovation, society development and/or pedagogical design but may not describe the contents taught in the course. Meanwhile, fewer courses had been tagged with SDG 10 by the algorithm than humans.
- Tagging on SDGs 4, 6, 10, 12, 14 and 15 has an F1 measure lower than average. The issue may be due to a small number of course samples identified by both humans and the algorithm. For example, SDGs 6, 14 and 15 are highly specific to mechanical engineering or biological sciences. As a result, only a small number of teachers can teach these SDGs.
- The performance of classifying SDGs 4 and 10 is not satisfactory. We observed that excessive teaching and learning terminologies (e.g. "curriculum," "learning" and "lessons") may also affect the classification of SDG 4. Meanwhile, discussions of SDG 10 are often related to politics and international affairs. Therefore, some teachers often used metaphors for indirect illustrations. These terminologies may affect the classification of SDG 10.

4.5 Forming course clusters according to a target set of sustainable development goals

In the university, students should study at least one general course in each cluster (AoI). Therefore, a study explored how courses can be clustered to identify mini-SDG-curriculum or concentration. The result can help students determine which courses in each AoI to choose if they want to study in-depth on an SDG. For example, if they are interested in SDG 3 (Good Health and Well Being), the recommended course selection can be Journey into Madness (CCHU), Health Literacy: Things to Know Before Consulting Dr. Google (CCST), Suicide: Risks, Research, and Realities (CCGL) and Curing the Chinese: Medicine and Society in Modern China (CCCH).

To further identify courses that teach a common set of SDGs, the study also conducted a K-mean clustering according to SDGs. We used K = 20 as the clustering parameter.

Table 1. Performance of the algorithm if each course is tagged to SDGs

SDG goal	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
Accuracy	0.95	0.98	0.77	0.82	0.88	0.96	0.95	0.79	0.68	0.83	0.69	0.86	0.94	0.93	0.97
F1-measure	0.47	0.67	0.68	0.32	0.57	0.40	0.60	0.49	0.58	0.42	0.55	0.08	0.62	0.35	0.29
A	10	7	57	21	25	2	10	35	75	13	54	9	10	12	5
B	7	5	64	23	21	8	10	34	50	35	59	16	16	5	2

Notes: A = number of courses that had been tagged by the algorithm; B = number of courses that had been tagged by the human
Source: Authors' own creation

The algorithm identified 13 clusters that have 4 or more than 4 courses with a common set of SDG(s). Among them, there are five possible clusters that teach the same set of two SDGs (i.e. cross-discipline):

- SDG 3/9 (e.g. technologies for well-being);
- SDG 5/8 (e.g. inclusive development);
- SDG 5/10 (e.g. reduced inequalities);
- SDG 8/9 (e.g. technology-facilitated economic growth); and
- SDG 9/11 (e.g. infrastructure for sustainable cities).

These provide insights to the common core office offering creative mini-curriculum on cross-discipline sustainable education.

5. Differences between human classification and machine mapping: teachers' perspectives

The study sent a survey to invite teaching staff who taught the common core courses to learn about their perceptions about using machine learning to map SDGs involved in the courses. The survey contained multiple-choice questions, five-point Likert scale questions and open-ended questions and was designed totally by the research team. In total, 122 common core courses' teaching staff finished the survey, and 81 courses' teaching staff consented to include their responses in the study.

5.1 The differences between the level of agreement on sustainable development goals classification results by human and machine

To investigate the teaching staff's level of agreement toward the results of SDGs classification by human and by machine, the survey incorporated two of the five-point Likert scale questions (1 – strongly disagree, 2 – somewhat disagree, 3 – neither agree or disagree, 4 – somewhat agree, 5 – strongly agree):

(1) How much do you agree or disagree with the above findings by human?

(2) How much do you agree or disagree with the above findings by machine?

As Table 2 shows, an independent-sample *t*-test was performed to compare the level of agreement toward the results by human and by machine. Results show that there was a significant difference [$t(160) = 2.050$, $p = 0.042 < 0.05$] in the level of agreement on the results mapping by human (M = 4.234, SD = 0.965) and by machine (M = 3.914, SD = 1.027). This indicates that compared to SDGs classification results by machine, the teaching staff preferred to refer to the results by human.

To better understand why teaching staff agreed more with the results by human than by machine, the survey also incorporated an open-ended question asking why they disagreed

	Human		Machine				
	M	SD	*M*	SD	df	*t*	*p*
Level of agreement	4.235	0.965	3.914	1.027	160	2.050	0.042

Source: Authors' own creation

Table 2. Results of independent-samples *t*-test ($N = 81$)

with classification results by human or machine. Two reasons emerged based on the teaching staff's answers on why they disagreed with classification results by machine:

(1) Reason 1: the performance of human classification was better than machine classification. After showing the results both by human and machine to the teaching staff, some of the teaching staff found that the classification results by human were more accurate based on their own acknowledgment of the course, as stated by:

> *Agree better with the Human classification instead.* (a teaching staff who taught two common core courses).
>
> *According to the machine classification, 4, 9, and especially 14 are not relevant to the course. SDG16 should be added.* (a teaching staff who strongly agreed with classification results by human)
>
> It seems odd that quality education (4) is not included, as a good portion of the content is related to mathematics education and the factors that facilitate or restrict access by individuals. The inclusion of "sustainable cities and communities" (11) is also a bit puzzling, as it seems that some connection is being made there that the teaching team doesn't see. (a teaching staff who somewhat agreed with classification results by human)

(2) Reason 2: course content was not relevant to the SDGs. Some of the teaching staff did not find there was a connection between the teaching content and SDGs, believing it was meaningless to use machine learning to tag the SDGs involved in their courses, as stated:

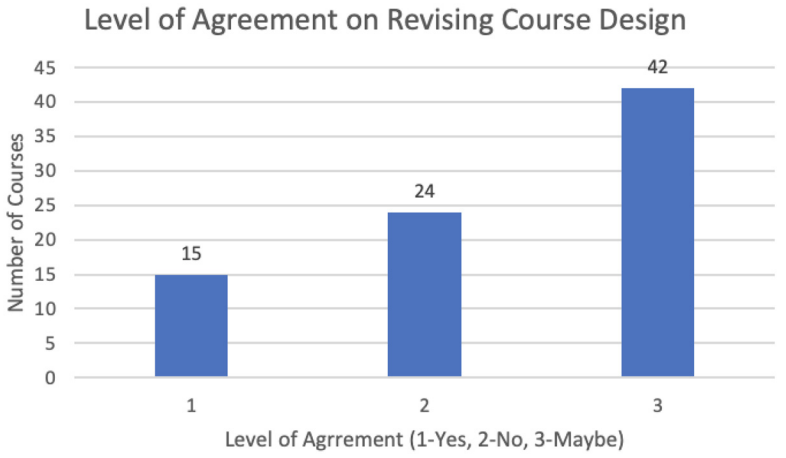
> Course content is not directly relevant to any of these goals. If a vague and distant connection is all that is required for a SDG, then these badges would become meaningless. (a teaching staff who strongly disagreed with both the classification results by human and machine)

5.2 The level of agreement on revising course design based on sustainable development goals classification results

To provide references for future SDGs-related education, this study also investigated the teaching staff's attitude toward whether they would be inspired to revise their course design based on the human-/computer-based SDGs classification results. As Figure 6 shows, 15 out of 81 courses' teaching staff had the intention to revise their courses based on the study's findings, 24 out of 81 courses' teaching staff did not think it was necessary to revise the course design and 42 out of 81 courses' teaching staff were not sure whether they would revise the course design in future.

The main factor for some of the teaching staff hoping to revise the course based on the classification results was that after the study introduced the importance of SDGs to them, they realized teaching SDGs-related topics was beneficial to students. As the teaching staff claimed:

> It is better to teach our students for this good goal. (a teaching staff whose course involved 12 SDG goals)
>
> I would like to encourage students to consider their roles in society and the impact they have on resources. (a teaching staff whose course involved SDG4-quality education and SDG9-industry, innovation and infrastructure)



Source: Authors' own creation

Figure 6.
Level of agreement on revising course design based on SDGs classification results

On the contrary, about a quarter of courses' teaching staff did not plan to revise their courses based on the classification results. Most of them thought their courses already incorporated necessary SDGs for students' learning, and the classification results also matched their course design. However, some of the teaching staff thought the SDGs goals should not be involved in their teaching, so they would not polish their course based on it. As stated by:

> The mapping results align with the course intention. (a teaching staff who strongly agreed with both the classification results by human and machine)

> The UN scheme is incompatible with real-world policymaking. (a teaching staff who knew well about UN SDGs)

Moreover, above half of courses' teaching staff were not sure whether they would revise their course design in the future. This may be because they were not familiar with the SDGs and still needed time to learn more about the related topics and how to apply SDGs to their daily teaching, as stated:

> Although our course covered several targets, we did not intentionally do that. The course design will be considered to be modified to fit into more focus on the SDG targets. (a teaching staff who learned a lot about SDGs from the study)

This study suggests that the teacher support center in HEI can introduce more content about UN SDGs to teaching staff, helping them to realize the importance of teaching students this topic and designing courses based on SDGs.

6. Discussions

6.1 Adopting the framework for the future sustainable development goals-related education

By 2030, the UN envisions all learners should acquire the knowledge and skills needed to promote sustainable development (Target 4.7) (Leicht *et al.*, 2018). This study presents an empirical machine learning analysis of SDG education and the role of HEIs in its development. The study can support university administrators and teachers in designing

strategies, curricula and courses to promote learning for the SDGs. The framework cannot suggest how SDGs can be taught in courses, but it can identify the relevance of SDGs taught in a course or a curriculum. Given the holistic SDG distribution of courses, educators may deepen their awareness and understanding of SDG concepts. Consequently, they can be stimulated to adapt relevant SDGs to fit concrete learning contexts.

The analysis can also be implemented in school/university/national curricula in developing countries, such that they can measure the progress on the education of sustainable development (SDG Target 17.19). Furthermore, the framework can help identify the strengths and weaknesses of curricula in different universities and regions on SDG education, which could encourage global cross-university and cross-region partnerships for sharing knowledge resources to support the learning of the SDG and eventually the achievement of SDGs in all countries, in particular developing countries (SDG Target 17.16).

Besides analyzing curriculum in HEIs, the framework can also be adopted for analyzing how SDGs can be taught in other educational settings, such as high school curricula (pre-university learning) and massive open online courses (MOOCs) (post-university learning) (Wang *et al.*, 2022). By understanding how SDGs have been taught before and after university curricula, HEIs can also design an SDG-relevant university curriculum.

6.2 Limitations of the work and future work of the study

The study can achieve a more accurate model training through adopting advanced machine learning algorithms (e.g. deep learning, multioutput multiclass ML algorithms); developing a more effective test data set by extracting more relevant information from syllabus and learning materials; expanding the training data set of SDGs that currently have insufficient records (e.g. SDG 12); and replacing the existing training data set from OSDG by authentic education-related documents (such as course syllabus) with SDG classifications. The performance of the algorithm should also be compared to other computer-based and human-based SDG classification approaches for cross-checking the results, with a systematic evaluation framework.

Furthermore, the study can be analyzed by circulating results to students and understanding how they would interpret and use the results for choosing courses for studying. Furthermore, the study mainly focused on the classification of topics that are taught in courses but cannot measure the effectiveness of adopted pedagogies, assessment strategies and competency development strategies in courses. The study can also conduct analysis based on assessment tasks and rubrics of courses to see whether the assessment tasks can help students understand and take action on SDGs.

7. Conclusion

Universities around the world support the achievement of SDGs through all areas of their activities on research, teaching and learning, knowledge exchange and campus operation. This study serves as a kickstart to institutionally analyze how SDGs have been taught in courses through machine learning. Descriptive analysis at course-level, theme-level and curriculum-level has been provided. The results indicate that the machine-learning classification approach can significantly accelerate the SDG classification of courses. However, currently, it cannot replace human classification due to the complexity of the problem and the lack of relevant training data. The proposed approach explores the possibility of using machine learning for SDG classifications in scale. We hope that with more technical advancements, the future study can inform and empower stakeholders in the studied university to support SDGs in the longer term.

References

Aleixo, A.M., Azeiteiro, U.M. and Leal, S. (2020), "Are the sustainable development goals being implemented in the Portuguese higher education formative offer?", *International Journal of Sustainability in Higher Education*, Vol. 21 No. 2, pp. 336-352.

Allen, C., Smith, M., Rabiee, M. and Dahmm, H. (2021), "A review of scientific advancements in datasets derived from big data for monitoring the sustainable development goals", *Sustainability Science*, Vol. 16 No. 5, pp. 1701-1716.

Baker, R.S. (2019), "Challenges for the future of educational data mining: the baker learning analytics prizes", *Journal of Educational Data Mining*, Vol. 11 No. 1, pp. 1-17.

Bautista-Puig, N. and Mauleón, E. (2019), "Unveiling the path towards sustainability: is there a research interest on sustainable goals?", *ISSI*, pp. 2770-2771.

Bautista-Puig, N., Aleixo, A.M., Leal, S., Azeiteiro, U. and Costas, R. (2021), "Unveiling the research landscape of sustainable development goals and their inclusion in higher education institutions and research centers: major trends in 2000–2017", *Frontiers in Sustainability*, Vol. 2, p. 12.

Goodall, M. and Moore, E. (2019), "Integrating the sustainable development goals into teaching, research, operations, and service: a case report of Yale university", *Sustainability: The Journal of Record*, Vol. 12 No. 2, pp. 93-96.

Hansen, B., Stiling, P. and Uy, W.F. (2021), "Innovations and challenges in SDG integration and reporting in higher education: a case study from the university of South Florida", *International Journal of Sustainability in Higher Education*, Vol. 22 No. 5, pp. 1002-1021.

Hesselbarth, C. and Schaltegger, S. (2014), "Educating change agents for sustainability–learnings from the first sustainability management master of business administration", *Journal of Cleaner Production*, Vol. 62, pp. 24-36.

Leal Filho, W., Raath, S., Lazzarini, B., Vargas, V.R., de Souza, L., Anholon, R., Quelhas, O.L.G., Haddad, R., Klavins, M. and Orlovic, V.L. (2018), "The role of transformation in learning and education for sustainability", *Journal of Cleaner Production*, Vol. 199, pp. 286-295.

Leicht, A., Combes, B., Byun, W.J. and Agbedahin, A.V. (2018), "From Agenda 21 to Target 4.7: the development of education for sustainable development", *Issues and Trends in Education for Sustainable Development*, UNESCO publishing, Paris, Vol. 25.

Morton, S., Pencheon, D. and Squires, N. (2017), "Sustainable development goals (SDGs), and their implementation: a national global framework for health, development and equity needs a systems approach at every level", *British Medical Bulletin*, Vol. 124 No. 1, pp. 81-90.

Nakamura, M., Pendlebury, D., Schnell, J. and Szomszor, M. (2019), "Navigating the structure of research on sustainable development goals", *Policy*, Vol. 11, p. 12.

OECD (2020), "SDG pathfinder tool", available at: https://sdg-pathfinder.org/about (accessed 12 February 2023).

Omazic, A. and Zunk, B.M. (2021), "Semi-systematic literature review on sustainability and sustainable development in higher education institutions", *Sustainability*, Vol. 13 No. 14, p. 7683.

OSDG, UNDP IICPSD SDG AI Lab and PPMI (2021), "OSDG community dataset (OSDG-CD)", *Zenodo*, available at: https://doi.org/10.5281/zenodo.5550238 (accessed 12 February 2023).

Porciello, J., Ivanina, M., Islam, M., Einarson, S. and Hirsh, H. (2020), "Accelerating evidence-informed decision-making for the sustainable development goals using machine learning", *Nature Machine Intelligence*, Vol. 2 No. 10, pp. 559-565.

Pukelis, L., Puig, N.B., Skrynik, M. and Stanciauskas, V. (2020), "OSDG–Open-Source approach to classify text data by UN sustainable development goals (SDGs)", *arXiv preprint arXiv:2005.14569*.

Rieckmann, M. (2017), *Education for Sustainable Development Goals: Learning Objectives*, UNESCO publishing, Paris.

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.

UCL (2021), "UCL sustainable development goals report 2020–21. UCL sustainable development goals", available at: www.ucl.ac.uk/sustainable-development-goals/sites/sustainable_development_goals/files/sdgis_report_final_version.pdf (accessed 12 February 2023).

United Nations General Assembly (2015), "Transforming our world: the 2030 agenda for sustainable development", available at: https://sustainabledevelopment.un.org/post2015/transformingourworld/publication (accessed 12 February 2023).

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S.D., Tegmark, M. and Fuso Nerini, F. (2020), "The role of artificial intelligence in achieving the sustainable development goals", *Nature Communications*, Vol. 11 No. 1, p. 233.

Waas, T., Verbruggen, A. and Wright, T. (2010), "University research for sustainable development: definition and characteristics explored", *Journal of Cleaner Production*, Vol. 18 No. 7, pp. 629-636.

Wang, Y., Li, Y., Hu, X., Xu, Y., Liang, X. and Lei, C.U. (2022), "Massive open online courses role in promoting United Nations sustainable development goals", *2022 IEEE Learning with MOOCS (LWMOOCS)*, pp. 126-130.

Appendix

List of United Nations Sustainable Development Goals (UN SDGs)
(United Nations General Assembly, 2015)

- Goal 1: End poverty in all its forms everywhere.
- Goal 2: End hunger, achieve food security and improved nutrition and promote sustainable agriculture.
- Goal 3: Ensure healthy lives and promote well-being for all at all ages.
- Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.
- Goal 5: Achieve gender equality and empower all women and girls.
- Goal 6: Ensure availability and sustainable management of water and sanitation for all.
- Goal 7: Ensure access to affordable, reliable, sustainable and modern energy for all.
- Goal 8: Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all.
- Goal 9: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation.
- Goal 10: Reduce inequality within and among countries.
- Goal 11: Make cities and human settlements inclusive, safe, resilient and sustainable.
- Goal 12: Ensure sustainable consumption and production patterns.
- Goal 13: Take urgent action to combat climate change and its impacts.
- Goal 14: Conserve and sustainably use the oceans, seas and marine resources for sustainable development.
- Goal 15: Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification and halt and reverse land degradation and halt biodiversity loss.
- Goal 16: Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels

- Goal 17: Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development.

About the authors

Chi-Un Lei received his PhD in Electrical Engineering at the University of Hong Kong and his EdM in Human Resource Development at the University of Illinois Urbana-Champaign. He has been working on digital learning initiatives in higher education and K12 education across Hong Kong and the Asia-Pacific Region. His current research interests include learning analytics, massive open online course (MOOCs), accessible learning, SDG education and generative AI. He was awarded the Best/Outstanding Paper Award at the IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE) in 2013, 2017 and 2020. He is an IEEE Senior Member and an Advance HE Senior Fellow.

Wincy Chan is a lecturer and pedagogical consultant at the University of Hong Kong. Her work looks into the psychological and learning processes that promote competencies for the 21st century. Her teaching involves blended learning with complement of virtual/mixed reality, chatbot, the metaverse and more.

Ms Yuyue Wang is a doctoral student from the School of Education at Syracuse University. She originally worked in Technology-Enriched Learning Initiative: TELI as a research assistant at the University of Hong Kong. Her research interests are gamification in education and teacher education. Yuyue Wang is the corresponding author and can be contacted at: u3586758@connect.hku.hk