OR2025 Presentation Submission

The 20th International Conference on Open Repositories, June 15-18 2025, Chicago, Illinois, USA

SDG-Classify: Automating the classification of research outputs into UN SDGs

Suchetha Nambanoor Kunnath, KMi, The Open University. suchetha.nambanoor-kunnath@open.ac.uk

Matteo Cancellieri, KMi, The Open University. matteo.cancellieri@open.ac.uk

Petr Knoth, KMi, The Open University. petr.knoth@open.ac.uk

Abstract

This paper presents SDG-Classify, a novel AI model for multi-label classification of research papers based on UN Sustainable Development Goals (SDGs), along with its integration into the CORE Dashboard, helping Higher Education Institutions (HEIs) to better understand how the content held in their repository contributes to SDGs. Using a few-shot, two-stage contrastive learning approach, the method generates contextual embeddings from publication metadata, including titles and abstracts, to train a classification head, leveraging an out-of-domain (OOD) multi-label SDG dataset from news articles. While the two-stage fine-tuned model performs effectively in OOD settings, incorporating additional context through label descriptions significantly enhances the model's performance in the in-domain evaluations. Additionally, integrating SDG-Classify into the CORE Dashboard streamlines the monitoring of SDG contributions for HEIs and supports research managers in targeted resource allocation and impact-driven decision-making.

Keywords

UN SDGs, multi-label classification, AI, few-shot learning

Audience

repository managers, research managers

Acknowledgements

We acknowledge the support of other members of the CORE team in the development of the SDG integration in CORE and the CORE Dashboard.

Introduction

The United Nations Sustainable Development Goals (SDGs), adopted in 2015 [1], provide a global framework comprising 17 objectives to address critical challenges such as poverty, inequality, and environmental sustainability. These goals are designed to foster inclusive development, drive economic growth, and ensure environmental protection, with a collective commitment to leaving no one behind by 2030. Classification of research outputs into UN SDGs is important for efficiently monitoring and aligning research efforts at Higher Educational Institutions (HEIs) with specific SDGs, targeted resource allocation and facilitating impact-driven decision-making. According to Körfgen et al., "HEIs can support an implementation of SDGs through: (1) Performing problem-oriented real-world research; (2) critically reflecting the SDGs and associated measures; (3) educating future decision-makers by fostering critical and systems-thinking; (4) offering best-practice examples for sustainable development on campus; and (5) strengthening the nexus at the policy-science-society interfaces" [2]. Manually classifying SDGs is time-consuming and costly, leaving most HEIs unaware of their research contributions to SDGs, thus highlighting the need for tools to automate and streamline this process.

Available automated SDG classification solutions predominantly rely on keyword-based searches [3], ignoring the broader context in which keywords are used, leading to misclassifications. Supervised classification approaches typically utilise datasets, such as the OSDG community dataset, which feature a single SDG label annotated for text excerpts from policy documents [4]. However, a single research paper is frequently relevant to multiple SDGs. As a result, the lack of a multi-label gold standard dataset presents significant challenges in developing effective multi-label SDG classification models. While Large Language Models (LLMs) can deliver good SDG classification performance, scaling their use to papers across the global network of repositories is currently computationally and financially infeasible for CORE. This underscores the need for lightweight models that can deliver comparable performance to existing LLMs while being more resource-efficient.

We present SDG-Classify, (a) the first multi-label Machine Learning (ML) model for SDG classification applied and evaluated on a research outputs dataset and (b) a new SDG Classification Module in the CORE Dashboard [7] for the monitoring and management of research outputs with respect to UN SDGs at a repository level. More specifically, the model automates the classification of research outputs by SDGs using a novel multi-label ML model, thus eliminating manual effort. By streamlining the mapping process, SDG-Classify provides actionable insights, enabling institutions to align research with global sustainability targets, support informed decision-making, and prioritise efforts in line with international agendas. This improved visibility enhances institutional accountability and highlights contributions to global challenges.

Method

The proposed AI-powered SDG classification model addresses existing gaps in the field by providing a scalable framework capable of mapping research publications to one or more SDG labels. The method consists of the following five steps:

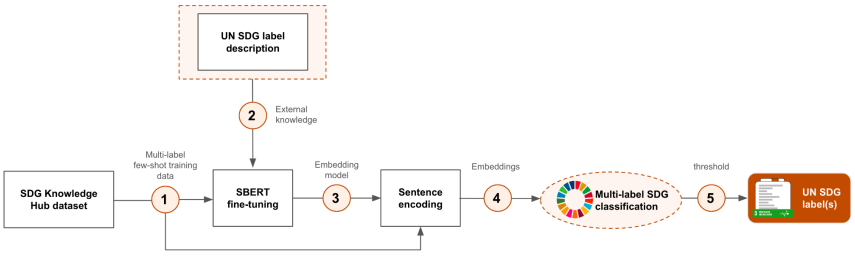

Figure 1 - Proposed method

1. **Multi-label few-shot training data**: The supervised multi-label SDG-Classify model leverages an out-of-domain (OOD) dataset of news articles, manually tagged for multiple SDGs sourced from the SDG Knowledge Hub [5].

2. **UN SDG description**: As there is currently, to our knowledge, no multi-label in-domain SDG corpus, we utilise SDG label descriptions, in the form of goal and target definitions, to incorporate additional context to our model.

3. **SBERT Fine-tuning**: The method involves fine-tuning of Sentence BERT (SBERT) through few-shot contrastive learning, following Tunstall et al [6]. SBERT fine-tuning is performed in two-stages as shown in Figure 1. During the first stage, SBERT is fine-tuned using SDG label descriptions, which includes the definitions for goals and targets. In the second stage, it is further refined using the OOD dataset.

4. **Sentence Encoding**: Once fine-tuned, SBERT generates contextual embeddings from research paper metadata (titles and abstracts).

5. **Multi-label SDG Classification**: The embeddings from the previous steps are then fed into a lightweight linear classifier (Logistic Regression) for multi-label classification, providing both predicted labels and their confidence scores, based on a threshold value (0.5).

Results

We evaluate three classification models based on the above approach: (1) Multi-label SBERT fine-tuning + LR - Single-stage SBERT fine-tuning using OOD dataset followed by classification using LR, (2) Label description SBERT fine-tuning + LR - Single-stage label description fine-tuning of SBERT followed by classification using LR and (3) Two-stage SBERT fine-tuning + LR - Label description fine-tuning is followed by a second stage few-shot OOD dataset fine-tuning of SBERT.

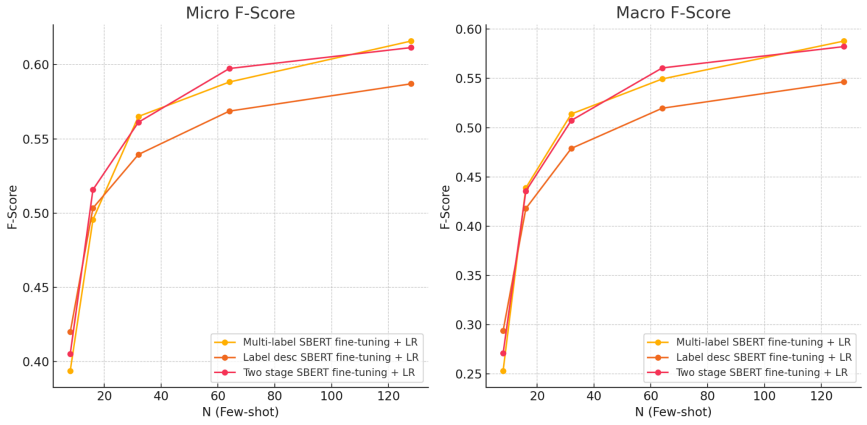

Figure 2 - Micro and macro f-scores for few-shot examples from OOD dataset. N is the number of positive/negative examples used for contrastive learning

The three models are evaluated on both the OOD dataset (Figure 2) and a manually annotated in-domain subset from Open Research Online (ORO) (Table 1), which will serve as a benchmark for real-world applicability.

In the **out-of-domain** OOD dataset experiment, the two-stage SBERT fine-tuned model achieves the best performance out of the three tested models. Not surprisingly, there is a steady improvement in performance in the case of single-stage SBERT fine-tuning using OOD dataset as the number of examples increases.

Model	**Multi-label SBERT fine-tuning + LR**	**Label desc SBERT fine-tuning + LR**	**Two stage SBERT fine-tuning + LR**
micro f-score	0.5571	0.5702	**0.5840**
macro f-score	0.4643	**0.4828**	**0.4812**

Table 1 - Performance of SDG classifiers on in-domain dataset (from ORO) (|N|=128)

Table 1 shows the performances for the manually annotated **in-domain** ORO dataset of research papers (276 papers manually annotated for multi-label SDGs). When interpreting the received performances, one

needs to consider that the inter-annotator agreement of SDG classification on research papers we acquired is 0.405 (augmented kappa [8]). This indicates the difficulty of the task and suggests that the performance of the ML model is applicable. Again, the two-stage SBERT fine-tuning + LR model performed the best. As a result we picked this model for integration with CORE. The classification is highly scalable, with inference taking approximately 1ms for 100 papers on an average server machine without requiring the use of GPU resources.

CORE Dashboard integration

Figure 3 depicts the CORE Dashboard with a new SDG-Classify module. For each institutional repository, the module will show the yearly and overall distribution of individual SDG contributions from the institutions. The module includes a search feature (Figure 4) that allows users to query SDG-related publications within the repository. The search results display SDG labels along with their confidence scores and provide options to view the papers through CORE or the institutional repository.

Figure 3 - SDG module on CORE Dashboard

As a result, SDG-Classify offers institutions fast and efficient insights into their contributions to sustainability objectives, enabling data-driven decision-making and strategic alignment through a scalable solution.

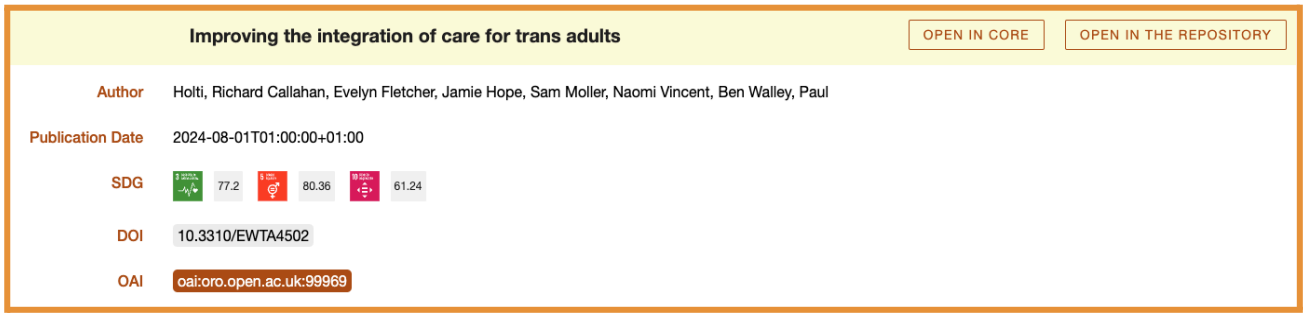

Figure 4 - Search result for a paper with SDGs on the dashboard

Conclusion

1. We developed **SDG-Classify**, which consists of:
    a. a scalable AI-based multi-label model to identify research article's contributions to the UN SDGs. The model uses a few-shot contrastive learning approach, trained on an OOD dataset and label descriptions for feature extraction, followed by a simple linear classifier.
    b. SDG Classification module for the CORE Dashboard. This can help HEIs in better understanding their contributions to UN SDGs.
2. As a global indexing service, CORE can scale out automated SDG classification across the repositories network, delivering new business intelligence and enabling comparisons of SDG research profiles across HEIs.

References

[1] United Nations. (2015). Transforming our world: The 2030 agenda for sustainable development. Resolution adopted by the General Assembly on 25 September 2015, 42809, 1-13.

[2] Körfgen, A.; Förster, K.; Glatz, I.; Maier, S.; Becsi, B.; Meyer, A.; Kromp-Kolb, H.; Stötter, J. It's a Hit! Mapping Austrian Research Contributions to the Sustainable Development Goals. *Sustainability* **2018**, *10*, 3295. https://doi.org/10.3390/su10093295

[3] Roberge, Guillaume; Kashnitsky, Yury; James, Chris (2022), "Elsevier 2022 Sustainable Development Goals (SDG) Mapping", Elsevier Data Repository, V1, doi: 10.17632/6bjy52jkm9.1

[4] OSDG, UNDP IICPSD SDG AI Lab, and PPMI. 2024. OSDG Community Dataset

(OSDG-CD). https://doi.org/10.5281/zenodo.10579179

[5] Dirk U. Wulff and Dominik S. Meier. 2023. SDG Knowledge Hub Dataset of SDG-labeled News Articles. https://doi.org/10.5281/zenodo.7523032

[6] Lewis Tunstall, Nils Reimers, Unso Eun Seo Jo, Luke Bates, Daniel Korat, Moshe Wasserblat, and Oren Pereg. 2022. Efficient Few-Shot Learning Without Prompts. arXiv:2209.11055 [cs.CL] https://arxiv.org/abs/2209.11055

[7] CORE Dashboard: https://core.ac.uk/services/repository-dashboard

[8] Marian Marchal, Merel Scholman, Frances Yung, and Vera Demberg. 2022. Establishing Annotation Quality in Multi-label Annotations. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 3659–3668, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.