From Data to Grassroots Initiatives: Leveraging Transformer-Based Models for Detecting Green Practices in Social Media

Anna Glazkova
Carbon Measurement Test Area in Tyumen' Region (FEWZ-2024-0016),
University of Tyumen
Tyumen, Russia
a.v.glazkova@utmn.ru

Olga Zakharova
Carbon Measurement Test Area in Tyumen' Region (FEWZ-2024-0016),
University of Tyumen
Tyumen, Russia
o.v.zakharova@utmn.ru

Abstract

Green practices are everyday activities that support a sustainable relationship between people and the environment. Detecting these practices in social media helps track their prevalence and develop recommendations to promote eco-friendly actions. This study compares machine learning methods for identifying mentions of green waste practices as a multi-label text classification task. We focus on transformer-based models, which currently achieve state-of-the-art performance across various text classification tasks. Along with encoder-only models, we evaluate encoder-decoder and decoder-only architectures, including instruction-based large language models. Experiments on the GreenRu dataset, which consists of Russian social media texts, show the prevalence of the mBART encoder-decoder model. The findings of this study contribute to the advancement of natural language processing tools for ecological and environmental research, as well as the broader development of multi-label text classification methods in other domains.

1 Introduction

Growing environmental challenges and climate change have led governments to develop adaptation and mitigation policies. These policies are expected to influence people's behavior, shaping what are known as social practices (Giddens, 1984). However, it is unclear whether these practices are becoming more eco-friendly or how they can be improved to better address the environmental crisis.

Green practices are social actions aimed at harmonizing the relationship between people and the environment by reducing resource consumption, waste, pollution, and emissions (Zakharova et al., 2021). Studying the prevalence of green waste practices is crucial to give people new ideas for promoting and expanding these actions (Lamphere and Shefner, 2018; van Lunenburg et al., 2020). Despite this, awareness of these practices in society remains limited.

To fill this gap, researchers need to collect and analyze large amounts of data on green waste practices. Social media provides a rich repository of environmental information, but manually reviewing posts is time-consuming and inefficient. Automated approaches, such as deep learning and content analysis, can contribute to solving this problem. However, to date, only a limited number of studies have used big data tools to investigate green waste practices (Haines et al., 2023; Zakharova et al., 2023; Sivarajah et al., 2020).

In this work, we explore the possibilities of natural language processing (NLP) tools for detecting mentions of green waste practices in social media. This task is framed as a multi-label text classification problem. Since large language models (LLMs) demonstrate superior performance across various NLP tasks, the focus of our research is on applying pre-trained language models (PLMs) to detect mentions of green waste practices. We seek to answer the following research questions (RQs):

- How effective can PLMs be in detecting mentions of green waste practices in social media?
- Which transformer-based model architectures are the most effective for this task?

The contributions of this paper can be summarized as follows. To address RQs, we present the first large-scale comparison of encoder-only, encoder-decoder, and decoder-only transformer-based models for the task of detecting mentions

Proceedings of the 1st Workshop on Ecology, Environment, and Natural Language Processing (NLP4Ecology2025), pages 1–9
March 2, 2025

of green waste practices in social media. Several label descriptors to represent data for generative models were evaluated. The presented evaluation has revealed that encoder-decoder models, namely mBART, can outperform both encoder-only and decoder-only models for detecting mentions of green waste practices. The obtained results provide insights into the potential of NLP to address environmental challenges. Our findings can also be used in other similar NLP tasks related to multi-label text classification.

2 Related Work

Modern approaches to multi-label text classification are mainly based on the use of encoder-only PLMs. Existing research often utilizes Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2019) and other transformer-based models. In particular, BERT-based approaches to multi-label text classification were used by Zahera et al. (2019); Chalkidis et al. (2020); Yarullin and Serdyukov (2021). Chalkidis et al. (2021) were the first to use the T5 model (Raffel et al., 2020) for multi-label classification. However, their approach utilized only the encoder component of the model, omitting the use of the model's decoder.

To date, there are several studies that used the encoder-decoder models fine-tuned for multi-label text classification in a generative manner. Kementchedjhieva and Chalkidis (2023) analyzed four methods for multi-label classification based on T5 and evaluated several types of label descriptors. Savci and Das (2024) compared multi-label BART (Lewis et al., 2020) and BERT; however, the results of BART were lower.

Up to now, there are only several approaches to perform multi-label text classification using decoder-only models. Peña et al. (2023); Siddiqui et al. (2024) performed fine-tuning of a pre-trained GPT2-model (Radford et al., 2019) with different prompt formats. Peskine et al. (2023) analyzed the performance of GPT-3 (Brown et al., 2020) for fine-grained multi-label tweet classification using zero-shot labeling. Vithanage et al. (2024) revealed that few-shot learning consistently outperforms zero-shot learning.

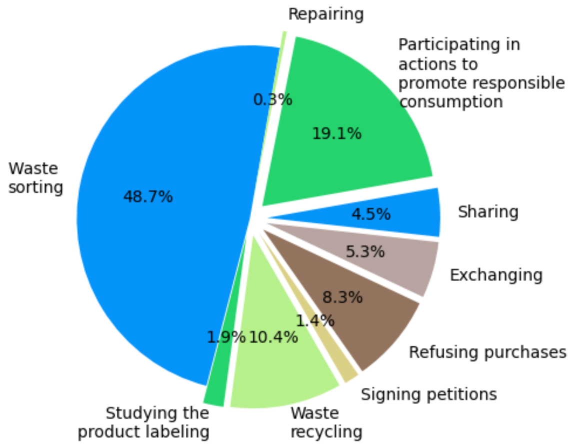

Figure 1: The distribution of mentions of green waste practices in GreenRu.

3 Data

This study uses the GreenRu[1] dataset (Zakharova and Glazkova, 2024) for detecting mentions of green waste practices in Russian social media texts. GreenRu consists of 1,326 posts in the Russian language with an average length of 880 symbols collected from online green communities. The posts have a sentence-level multi-label markup indicating green waste practices mentioned in them. The average length of a sentence is 110 symbols. Nine types of green waste practices (Zakharova et al., 2022) were used for the annotation of GreenRu: 1) **waste sorting**, i.e. separating waste by its type; 2) **studying the product labeling** to indicate product packaging as a type of waste; 3) **waste recycling**, i.e. transforming waste materials into reusable resources for future production.; 4) **signing petitions** to influence the authorities; 5) **refusing purchases** to reduce consumption and environmental footprint; 6) **exchanging** an unnecessary item or service for a desired one; 7) **sharing** things with other people for a fee or free of charge; 8) **participating in actions to promote responsible consumption**, including workshops, festivals, lessons, etc.; 9) **repairing** things as an alternative to throwing them away. The distribution of mentions of green waste practices in the dataset is presented in Figure 1. GreenRu is pre-split into training and test sets, with their characteristics presented in Table 1.

[1] `https://github.com/green-solutions-lab/GreenRu`

Characteristic		Training set	Test set
Total number of posts		913	413
Total number of sentences with multi-label markup		2442	1058
Distribution of green practice mentions			
1	Waste sorting	1275	560
2	Studying the product labeling	55	17
3	Waste recycling	272	121
4	Signing petitions	22	31
5	Refusing purchases	236	75
6	Exchanging	146	52
7	Sharing	109	62
8	Participating in actions to promote responsible consumption	510	209
9	Repairing	10	3

Table 1: The statistics of GreenRu.

4 Models

In this study, we compared several approaches to multi-label text classification to detect mentions of green waste practices in social media. Alongside encoder-only PLMs, which are traditionally used for multi-label text classification, we also employed encoder-decoder and decoder-only models. All transformer-based PLMs were implemented using the Simple Transformers[2] and Transformers (Wolf et al., 2020) libraries. The overview of models is shown in Table 2. In addition to fine-tuned PLMs, we evaluated the effectiveness of prompt-based learning and two traditional machine learning baselines.

4.1 Encoder-only Models

- ruBERT, a version of BERT (Devlin et al., 2019) for the Russian language. We used two versions of this model, namely **ruBERT-base**[3] (Kuratov and Arkhipov, 2019) and **ruBERT-large**[4] (Zmitrovich et al., 2024).

- ruELECTRA (Zmitrovich et al., 2024), a model is based on the ELECTRA architecture (Clark et al., 2020). In this study, **ruELECTRA-base**[5] and **ruELECTRA-large**[6] were utilized.

[2]https://simpletransformers.ai/
[3]https://huggingface.co/DeepPavlov/rubert-base-cased
[4]https://huggingface.co/ai-forever/ruBert-large
[5]https://huggingface.co/ai-forever/ruElectra-base
[6]https://huggingface.co/ai-forever/ruElectra-large

All encoder-only PLMs were fine-tuned for five epochs using the AdamW optimizer, a learning rate of 4e-5, a batch size of eight, and a maximum sequence length of 256 tokens. The fine-tuning procedure was performed in a multi-label setting, with a transformer-based classifier outputting n binary labels. This study used n equal to nine in accordance with the number of green waste practices in the dataset.

4.2 Encoder-decoder Models

- **ruT5**[7] (Zmitrovich et al., 2024), a text-to-text transformer pre-trained only on Russian-language textual data and designed analogically to T5 (Raffel et al., 2020).

- **mBART**[8] (Tang et al., 2021), a sequence-to-sequence machine translation model built on the baseline architecture of BART (Lewis et al., 2020). It was pre-trained on more than 50 languages using a combination of span masking and sentence shuffling techniques.

ruT5 and mBART were fine-tuned for 20 epochs. We explored several alternative forms of label descriptors, some of which were previously introduced in (Kementchedjhieva and Chalkidis, 2023), while others were proposed for the first time in this study. The following descriptors were used: **original** label descriptors in the Russian language; **simplified** one-word versions of original label descriptors; **numbers** assigned to green

[7]https://huggingface.co/ai-forever/ruT5-base
[8]https://huggingface.co/facebook/mbart-large-50

Model	Version	Architecture	Params	Data source
ruBERT	rubert-base-cased	encoder-only	180M	Wikipedia, news texts
	rubert-large		427M	Wikipedia, news texts, Librusec, C4, OpenSubtitles
ruELECTRA	ruelectra-medium		85M	
	ruelectra-large		427M	
ruT5	rut5-base	encoder-decoder	222M	
mBART	mbart-large-50		680M	Common Crawl (CC25), mono-lingual data from XLMR
ruGPT	rugpt-3-medium	decoder-only	355M	Wikipedia, news texts, Librusec, C4, OpenSubtitles
T-lite	t-lite-instruct-0.1		8B	Open Source English-language datasets, translations of English-language datasets, synthetic grounded QA contexts

Table 2: Overview of transformer-based models.

waste practices according to Table 1; **special tokens** added to the model and corresponding to green waste practices; **one-hot** label presentation. Since mBART is a multi-lingual model designed for machine translation, we also evaluated original and simplified label descriptors translated into the English language (**original-Eng**, **simplified-Eng**) for the mBART model. The examples of label descriptors are given in Table 3.

4.3 Decoder-only Models

- **ruGPT**[9] (Zmitrovich et al., 2024), a Russian equivalent of GPT-3, uses its architecture (Brown et al., 2020) and the GPT-2 code base from the Transformers library (Radford et al., 2019; Wolf et al., 2020).
- **T-lite**[10], an open-source instruction-based LLM with 85% of its pre-training data in Russian. For text generation, a temperature value was set to 1.

ruGPT was fine-tuned with a causal language modeling objective with a maximum sequence length of 1024 tokens for ten epochs. The input text was presented as follows: *text* + ”Категории: ” (*”Categories: ”*) + *label descriptors*. The same list of label descriptors was used for ruGPT as for ruT5.

For T-lite, prompt-based learning was implemented using the Transformers library (Wolf et al., 2020). The models were tasked with analyzing the text, identifying mentions of green waste practices, and selecting one or more categories from the list of labels. Then, ten examples of texts and their corresponding labels were provided. We used two variations of a few-shot prompt. In the first case, the list of labels was provided without explanations. In the second case, each label was accompanied by a description (for example, Перерабатывать отходы: преобразование отходов в перерабатываемые материалы для дальнейшего использования в производстве, *Waste recycling: converting waste materials into reusable materials for further use in the production of something*).

4.4 Baselines

- K-nearest Neighbors classifier (**KNN**) with a number of neighbors equal to three and the weight points obtained by the inverse of their distance.
- Multi-layer Perceptron (**MLP**), a feed-forward neural network with a single hidden layer of size 100 and a hyperbolic tangent (tanh) activation function.

Both models were implemented using Scikit-Learn (Pedregosa et al., 2011) and the paraphrase-multilingual-MiniLM-L12-v2 model (Reimers and Gurevych, 2020) as a sentence embedder.

4.5 Evaluation Metric

The multi-label F1-score was used as an evaluation metric. This metric was calculated by determining the F1-score for each class individually and then averaging the results.

[9] https://huggingface.co/ai-forever/rugpt3medium_based_on_gpt2

[10] https://huggingface.co/AnatoliiPotapov/T-lite-instruct-0.1

Label descriptor	Example
Text: Мой муж возит меня на сортировку с мешками вторсырья и не ворчит, неидентифицируемую упаковку складывает кучкой в кухне (*My husband takes me to the waste sorting center with the bags of recyclables without complaining and neatly stacks unidentifiable packaging in a corner of the kitchen*)	
Original	сортировать отходы, изучать маркировку товаров
Simplified	сортировка, маркировка
Numbers	1, 2
Special tokens	P1, P2
One-hot	110000000
Original-Eng	waste sorting, studying the product labeling
Simplified-Eng	sorting, labeling

Table 3: Label descriptors.

5 Results and Discussion

The results are presented in Table 4. The scores of baselines were 43.03% and 59.75% in terms of the multi-label F1-score for KNN and MLP respectively. The scores that outperform both baselines are underlined. The dotted line underlines the scores that surpass the KNN baseline. The highest value of the multi-label F1-score is shown in bold.

Encoder-only PLMs demonstrated relatively high results. All four PLMs outperformed baselines. The highest result of 67.88% in terms of the multi-label F1-score was shown by ruBERT-large.

Encoder-decoder PLMs mostly achieved the results above baselines. The best scores for ruT5 were obtained using simplified and original label descriptors (62.51% and 60.54% respectively). The use of the numbers, special tokens, and one-hot label descriptors did not increase the MLP results. The one-hot label descriptors did not even surpass KNN (34. 95%), indicating that the ruT5 model struggles to interpret this method of label representation. mBART demonstrated the highest score using simplified label descriptors (69.76%). The second and third highest scores were obtained with the original and simplified-Eng labels descriptors (69.49% and 69%). The use of the numbers and original-Eng label descriptors also improved the results of encoder-only PLMs (68.91% and 68.53%). The one-hot and special token label descriptors demonstrated the multi-label F1-score of 67.12% and 65.71% respectively which did not surpass ruBERT-large but outperformed baselines.

In general, decoder-only PLMs demonstrated

Model	F1-score, %
Encoder-only models	
ruBERT-base	66.53
ruBERT-large	67.88
ruELECTRA-base	65.28
ruELECTRA-large	65.69
Encoder-decoder models	
ruT5 + original	60.54
ruT5 + simplified	62.51
ruT5 + numbers	59.16
ruT5 + special tokens	52.60
ruT5 + one-hot	34.95
mBART + original	69.49
mBART + simplified	**69.76**
mBART + numbers	68.91
mBART + special tokens	65.71
mBART + one-hot	67.12
mBART + original-Eng	68.53
mBART + simplified-Eng	69.00
Decoder-only models	
ruGPT + original	46.66
ruGPT + simplified	51.08
ruGPT + numbers	33.07
ruGPT + special tokens	38.96
ruGPT + one-hot	41.29
T-lite$_{few-shot}$	42.04
T-lite$_{few-shot+explanations}$	47.77
Baselines	
KNN	43.03
MLP	59.75

Table 4: Results.

the lowest results in comparison to encoder-only and encoder-decoder PLMs. The highest result of ruGPT was obtained using the simplified label descriptors (51.08%). The use of the numbers, special tokens, and one-hot label descriptors showed the results below the KNN baseline. The instruction-based T-lite model also did not demonstrate high results. The use of prompt-based learning obtained 42.04% and 47.77% in terms of the multi-label F1-score. Despite the fact that incorporating explanations of green waste practices led to a more than 5% improvement in performance, T-lite failed to outperform the MLP baseline.

Figure 2 shows the performance growth using different label descriptors in comparison to the MLP baseline. The figure reveals that the labels descriptors based on text representation (original, simplified, original-Eng, and simplified-Eng) show higher results than the labels descriptors based on numerical and special token representation. For all three models (ruT5, mBART, ruGPT) the best results were achieved using the simplified label descriptors.

The RQs were aimed to evaluate the effectiveness of PLMs in detecting mentions of green waste practices on social media and to determine which transformer-based model architectures are the most effective for this task. Our experiments demonstrated that the performance of PLMs varies depending on their architecture and model type. Encoder-only models achieved the multilabel F1 score values between 65.28% and 67.88%, showing consistent and relatively strong performance. This supports their common use in multi-label classification tasks. However, the best result in our experiments was achieved by the mBART model (69.76%), highlighting the strong potential of encoder-decoder models for multi-label classification. Label descriptors greatly affect encoder-decoder models; for example, ruT5 results vary from 34.95% to 62.51%. Decoder-only models, including instruction-based ones, showed the poorest performance in our experiments. However, the results indicate that incorporating explanations into the prompt can enhance the performance of instruction-based models.

6 Conclusion

In this work, we explored the efficiency of PLMs for detecting mentions of green waste practices in social media. To address RQs, we compared encoder-only, encoder-decoder, and decoder-only PLMs. Our findings showed that encoder-only and encoder-decoder models generally outperformed decoder-only models. mBART achieved the best performance and revealed the most suitable label descriptors for generative PLMs in the multi-label text classification task.

This current study is limited by the use of only one data set to detect green waste practices. This is due to the fact that, to the best of the authors' knowledge, GreenRu is currently the only freely available dataset specifically annotated for this task. A potential future direction for this research could involve applying transfer learning techniques and generating texts to train models for other languages. Another possible limitation of this study is the use of general-domain models. Further research can investigate the role of in-domain pre-training for this task. Future research directions can additionally include exploring additional multilingual models beyond mBART and the MLP baseline to expand comparative insights. Investigating models with billion-scale parameters while incorporating PEFT (Parameter-Efficient Fine-Tuning) approaches could also enhance performance and efficiency.

The results obtained in this study allowed us to identify the most effective models for searching for green waste practices on social networks. Using these models, the following management tasks can be solved:

1. The prevalence of green waste practices in the text of posts from individual communities can be used to identify the specific activities of a particular community and select the most appropriate solutions when organizing companies to combat plastic pollution or solve the problem of food waste by organizing food sharing.

2. The most popular practices can be found in formalizing this activity through the form of standards for organizing green waste practices and their subsequent replication through training and information for eco-activists. For example, organizing separate waste collection in the yards of apartment blocks.

3. The least popular practices include organizing support for these practices, if they are seen as important behavioral changes to reduce anthropogenic climate impacts. For ex-

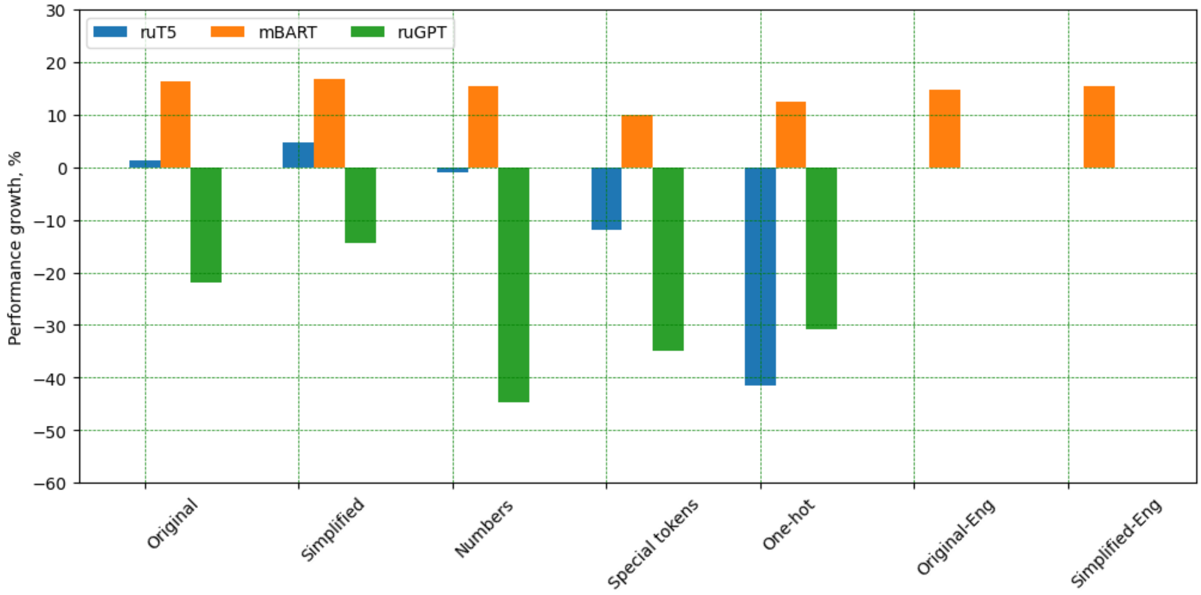

Figure 2: The performance growth using different label descriptors in comparison to the MLP baseline.

ample, organizing enlightening lectures on sustainable fashion.

4. The dynamics of mentions of each green practice can be studied to further explore the ways in which it is scaled up or the factors influencing green social innovation.

5. Communities can be found that do not position themselves as green, but organize eco-friendly activities to develop interactions between activists and provide mutual support for promoting green waste practices.

The information obtained through PLMs can be used by authorities, eco-businesses, and activists to promote behavioral change, support green innovation and promote sustainable social practices. The models for automatically detecting mentions of green waste practices make researching these practices easier and cheaper as they replace experts in dealing with textual information. Additionally, these methods allow processing large amounts of textual data that are not accessible to expert analysis.

Acknowledgment

This study was supported by the Ministry of Science and Higher Education of the Russian Federation within the framework of the Carbon Measurement Test Area in Tyumen' Region (FEWZ-2024-0016).

We are grateful to Nadezhda Zhuravleva (Center for Academic Writing "Impulse", University of Tyumen) for her assistance with the English language.

References

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. In *Advances in Neural Information Processing Systems*, volume 33, pages 1877–1901. Curran Associates, Inc.

Ilias Chalkidis, Manos Fergadiotis, and Ion Androutsopoulos. 2021. MultiEURLEX-A multi-lingual and multi-label legal document classification dataset for zero-shot cross-lingual transfer. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 6974–6996.

Ilias Chalkidis, Manos Fergadiotis, Sotiris Kotitsas, Prodromos Malakasiotis, Nikolaos Aletras, and Ion Androutsopoulos. 2020. An empirical study on large-scale multi-label text classification including few and zero-shot labels. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7503–7515.

Kevin Clark, Minh-Thang Luong, Quoc V. Le, and Christopher D. Manning. 2020. ELECTRA: Pre-training text encoders as discriminators rather than generators. In *ICLR*.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Anthony Giddens. 1984. The constitution of society: Outline of the theory of structuration. *Polity*.

Shelley Haines, Omar H Fares, Myuri Mohan, and Seung Hwan Lee. 2023. Social media fashion influencer eWOM communications: understanding the trajectory of sustainable fashion conversations on YouTube fashion haul videos. *Journal of Fashion Marketing and Management: An International Journal*, 27(6):1027–1046.

Yova Kementchedjhieva and Ilias Chalkidis. 2023. An exploration of encoder-decoder approaches to multi-label classification for legal and biomedical text. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 5828–5843.

Yuri Kuratov and Mikhail Arkhipov. 2019. Adaptation of deep bidirectional multilingual transformers for Russian language. In *Komp'juternaja Lingvistika i Intellektual'nye Tehnologii*, pages 333–339.

Jenna A Lamphere and Jon Shefner. 2018. How to green: Institutional influence in three us cities. *Critical Sociology*, 44(2):303–322.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880.

Marion van Lunenburg, Karin Geuijen, and Albert Meijer. 2020. How and why do social and sustainable initiatives scale? a systematic review of the literature on social entrepreneurship and grassroots innovation. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 31(5):1013–1024.

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12:2825–2830.

Alejandro Peña, Aythami Morales, Julian Fierrez, Ignacio Serna, Javier Ortega-Garcia, Iñigo Puente, Jorge Cordova, and Gonzalo Cordova. 2023. Leveraging large language models for topic classification in the domain of public affairs. In *International Conference on Document Analysis and Recognition*, pages 20–33. Springer.

Youri Peskine, Damir Korenčić, Ivan Grubisic, Paolo Papotti, Raphael Troncy, and Paolo Rosso. 2023. Definitions matter: Guiding GPT for multi-label classification. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 4054–4063, Singapore. Association for Computational Linguistics.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners. *OpenAI*. Accessed: 2024-11-15.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of machine learning research*, 21(140):1–67.

Nils Reimers and Iryna Gurevych. 2020. Making monolingual sentence embeddings multilingual using knowledge distillation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Pinar Savci and Bihter Das. 2024. Multi-label classification in text data: An examination on innovative technologies. In *2024 12th International Symposium on Digital Forensics and Security (ISDFS)*, pages 1–4. IEEE.

Muhammad Hammad Fahim Siddiqui, Diana Inkpen, and Alexander Gelbukh. 2024. Instruction Tuning of LLMs for Multi-label EmotionClassification in Social Media Content. *Proceedings of the Canadian Conference on Artificial Intelligence*. Https://caiac.pubpub.org/pub/lezimqvm.

Uthayasankar Sivarajah, Zahir Irani, Suraksha Gupta, and Kamran Mahroof. 2020. Role of big data and social media analytics for business to business sustainability: A participatory web context. *Industrial Marketing Management*, 86:163–179.

Yuqing Tang, Chau Tran, Xian Li, Peng-Jen Chen, Naman Goyal, Vishrav Chaudhary, Jiatao Gu, and Angela Fan. 2021. Multilingual translation from denoising pre-training. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 3450–3466, Online. Association for Computational Linguistics.

Dinithi Vithanage, Chao Deng, Lei Wang, Mengyang Yin, Mohammad Alkhalaf, Zhenyu Zhang, Yunshu Zhu, Alan Christy Soewargo, and Ping Yu. 2024.

Evaluating approaches of training a generative large language model for multi-label classification of unstructured electronic health records. *medRxiv*, pages 2024–06.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2020. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, pages 38–45.

Ramil Yarullin and Pavel Serdyukov. 2021. BERT for sequence-to-sequence multi-label text classification. In *Analysis of Images, Social Networks and Texts*, pages 187–198, Cham. Springer International Publishing.

Hamada M Zahera, Ibrahim A Elgendy, Rricha Jalota, Mohamed Ahmed Sherif, and E Voorhees. 2019. Fine-tuned BERT model for multi-label tweets classification. In *TREC*, pages 1–7.

Olga Zakharova and Anna Glazkova. 2024. GreenRu: A Russian dataset for detecting mentions of green practices in social media posts. *Applied Sciences*, 14(11):4466.

Olga Zakharova, Anna Glazkova, and Lyudmila Suvorova. 2023. Online equipment repair community in Russia: Searching for environmental discourse. *Sustainability*, 15(17):12990.

Olga V Zakharova, Anna V Glazkova, Irina N Pupysheva, and Natalia V Kuznetsova. 2022. The importance of green practices to reduce consumption. *Changing Societies & Personalities. 2022. Vol. 6. Iss. 4*, pages 884–905.

Olga V Zakharova, Tatiana I Payusova, Irina D Akhmedova, and Lyudmila G Suvorova. 2021. Green practices: Approaches to investigation. *Sotsiologicheskie issledovaniya*, (4):25–36.

Dmitry Zmitrovich, Aleksandr Abramov, Andrey Kalmykov, Vitaly Kadulin, Maria Tikhonova, Ekaterina Taktasheva, Danil Astafurov, Mark Baushenko, Artem Snegirev, Tatiana Shavrina, et al. 2024. A family of pretrained transformer language models for Russian. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 507–524.