ДОКЛАДЫ РОССИЙСКОЙ АКАДЕМИИ НАУК. МАТЕМАТИКА, ИНФОРМАТИКА, ПРОЦЕССЫ УПРАВЛЕНИЯ, 2023, том 514, № 2, с. 417–430

ESGify: АВТОМАТИЗИРОВАННАЯ КЛАССИФИКАЦИЯ ЭКОЛОГИЧЕСКИХ, СОЦИАЛЬНЫХ И УПРАВЛЕНЧЕСКИХ РИСКОВ

© 2023 г. А. Казаков[1], С. Денисова[1], И. Барсола[1], Е. Калугина[1], И. Молчанова[1], И. Егоров[1], А. Костерина[1], Е. Терещенко[1], Л. Шутихина[1], И. Дорощенко[1], Н. Сотириади[1], С. Будённый[1,2,*]

Представлено академиком РАН А. А. Шананиным
Поступило 24.08.2023 г.
После доработки 15.09.2023 г.
Принято к публикации 24.10.2023 г.

Растущее признание значимости экологических, социальных и управленческих (ESG) факторов при принятии финансовых решений стимулирует потребность в эффективных и всеобъемлющих инструментах оценки ESG-рисков. В этом исследовании мы представляем модель обработки естественного языка (NLP) с открытыми весами, “ESGify”, основанную на архитектуре MPNet и предназначенную для классификации текстов об ESG-рисках. Мы также представляем иерархическую методологию классификации ESG-рисков, использующую опыт специалистов по ESG и лучшие мировые практики. ESGify разработана для автоматизации классификации ESG-рисков в соответствии с установленной методологией. Модель обучена на основании набора данных с несколькими метками, составленного вручную и состоящего из 2000 новостных статей, и адаптирована к ESG-тематике с использованием текстов отчетов об устойчивом развитии. Мы сравниваем методы аугментации, основанные на обратном переводе и больших языковых моделях (LLMs) для улучшения качества модели и достижения метрики качества, взвешенной F1 метрики, на уровне 0.5 на наборе данных с 47 классами. Для простых подсказок этот результат превосходит ChatGPT 3.5. Веса модели и документация размещены на Github https://github.com/sb-ai-lab/ESGify под лицензией Apache 2.0.

Ключевые слова: ESG, устойчивое развитие, LLM, NLP
DOI: 10.31857/S2686954323601525, **EDN:** CXCFVV

1. ВВЕДЕНИЕ

Повестка дня в области охраны окружающей среды, социальной защиты и управления (ESG-повестка) в последние годы набрала значительный импульс благодаря повышению осведомленности, прежде всего в финансовом секторе, о нефинансовых рисках и их интеграции в процесс принятия финансовых решений. Сегодня от компаний ожидается, что они будут более устойчивыми, принимая во внимание не только финансовые показатели, но и ESG-показатели. По мере роста интереса к воздействию бизнеса на окружающую среду, общество и экономику участники рынка предпринимают попытки оценить такие риски и их влияние на финансовую стабильность. В связи с этим возникает индустрия ESG-рейтингов и различных инструментов для оценки ESG-рисков.

Кроме того, конкретные ESG-риски (такие как последствия изменения климата и изменение поведенческих моделей) становятся существенными в краткосрочной перспективе. Своевременное выявление этих рисков может привести к значительной экономии средств. В некоторых отраслях с развитой инфраструктурой и пространственно распределенной деятельностью некорректное определение рисков может привести к остановке работы [1].

В этом контексте особую ценность представляют структурированные количественные и качественные ESG-данные, как производимые компанией и предоставляемые в виде регулярной нефинансовой отчетности и ESG Data Book, так и собранные независимо, например, на основе анализа открытых данных о новостных ESG-событиях. Многие методологии ESG-рейтингов учитывают так называемые резонансные события (controversies), существенные негативные факты о компании, собранные в основном из новостных источников.

[1]*Сбер, Москва, Россия*
[2]*Институт искусственного интеллекта AIRI, Москва, Россия*
*****E-mail: sanbudenny@sberbank.ru*

Однако ручная обработка и анализ таких текстов являются сложной задачей из-за огромного объема и сложности данных. В результате растет интерес к использованию методов обработки естественного языка (NLP) для автоматизации анализа ESG-текстов. NLP может упростить обработку ESG-текста, позволяя компьютерам понимать и анализировать человеческий язык. Методы NLP, такие как анализ настроений, тематическое моделирование и распознавание именованных сущностей, могут помочь извлечь релевантную информацию из ESG-текстов и выявить закономерности и тенденции.

Более того, NLP может помочь классифицировать ESG-риски, что важно для компаний при определении приоритетов и устранении наиболее критических рисков. Классификация ESG-рисков важна по нескольким причинам. Во-первых, это помогает компаниям выявлять и расставлять приоритеты в отношении наиболее существенных ESG-рисков, позволяя эффективно распределять ресурсы. Во-вторых, это позволяет заинтересованным сторонам, включая инвесторов, клиентов и сотрудников, принимать обоснованные решения на основе ESG-показателей. В-третьих, это способствует увеличению прозрачности и подотчетности, побуждающих компании со временем улучшать свои показатели в области ESG.

В этой статье представлена модель NLP, которая классифицирует ESG-риски по данному тексту. Модель ESGify[1, 2] направлена на использование потенциала неструктурированной текстовой информации в виде новостей, связанных с устойчивым развитием. Основная задача, которую выполняет модель, – это классификация неструктурированного новостного потока по тематическим группам. Мы описываем методологию, использованную для разработки модели, включая набор данных, методы предварительной обработки и архитектуру модели. Мы представляем результаты оценки модели и обсуждаем возможные применения наших выводов для классификации ESG-рисков. Мы также выделяем сильные стороны и ограничения модели и предлагаем будущие направления исследований в этой области.

Вклад. Вклад нашей статьи заключается в двух аспектах:

• Во-первых, мы вводим многоуровневую иерархию классов ESG для определения рисков.

• Во-вторых, мы представляем *ESGify*, библиотеку Python с открытым исходным кодом для автоматической классификации ESG-рисков по тексту на основе модели MPNet-base [2].

Статья организована следующим образом. В разделе 2 мы рассматриваем существующие решения для классификаторов ESG-рисков и описываем их отличие от нашей библиотеки. В разделе 3 представлен метод расчетов, в разделе 4 показаны результаты моделирования с использованием методов аугментации. В разделе 5 мы подводим итоги нашей работы.

2. ОБЗОР ЛИТЕРАТУРЫ

В этой главе мы даем обзор того, как была разработана методология ответственного инвестирования и анализа ESG-рисков. Далее мы кратко опишем существующие модели с открытым исходным кодом для классификации новостей ESG и приведем краткое описание сравнений.

2.1. Практика классификации ESG-рисков

Одной из отправных точек для ESG-движения стали Принципы ответственного инвестирования, инициатива, основанная в 2005–2006 гг., которая предложила шесть принципов в форме высокоуровневых подходов к инвестированию [3]. По мере того как внедрение принципов ответственного инвестирования набирало обороты, появились различные инициативы по классификации и оценке ESG-рисков. Были разработаны стандарты и руководящие принципы для выявления и оценки рисков в сфере финансов, особенно проектного финансирования, включая стандарты деятельности МФК [4]; экологические и социальные стандарты Всемирного банка [5], набор инструментов ЕБРР для управления экологическими и социальными рисками [6], и т.д.

Эти принципы повлияли на разработку стандартов нефинансовой отчетности, главным образом GRI [7] и SASB [8]. GRI предоставляет списки тем и показателей устойчивости, связанных с наибольшим воздействием, для некоторых отраслей в рамках разработанных отраслевых стандартов, в то время как SASB фокусируется на темах и показателях, существенных для финансовой устойчивости, которые могут представлять интерес для инвесторов. SASB также предоставляет собственный инструмент для определения существенности[3]. Стандарты Международного совета по стандартам устойчивого развития (ISSB), наиболее новые на момент написания статьи, основаны на финансовой существенности, однако они не дают ясности в отношении того, что считается финансово существенным для каждой отрасли. Ожидается, что компания определит свои собственные существенные риски и показатели, по которым будет отчитываться.

Аналитические агентства также учитывают ESG-принципы в своих рейтинговых методологиях и создают рейтинговые ESG-продукты. Наи-

[1] https://huggingface.co/ai-lab/ESGify
[2] https://github.com/sb-ai-lab/ESGify
[3] https://sasb.org/standards/materiality-finder/

Таблица 1. Характеристики моделей анализа текста в рамках ESG-повестки

Название модели	ESG-BERT	FinBERT-ESG	FinBERT-esg-9-categories	FinBERT-ESG-Refinitive	ESGify
Общие сведения					
Дата выхода	2020	2022	2022	2020	2023
Лицензия	Apache 2.0	N/A	N/A	Not Open-Sourced	Apache 2.0
Архитектура	BERT-base	FinBERT-base	FinBERT-base	FinBERT-base	MPNet-base
Позиционирование					
	✓	✓	✓	✓	✓
Выявление рисков	✓	✓	✓	✓	✓
Предсказанные сущности					
Подсчет количества меток	26	4	9	20 / 9	47
Тип классификации	multi-class	multi-class	multi-class	multi-class/ multi-label	multi-label
Поддержка иерархии меток					✓
Тренировочные данные					
Набор данных	N/A	2000 аннотированных предложений	14000 аннотированных предложений	31600/12600 новостей	14000 аннотированных текстов содержащих более 35000 предложений
Адаптация предметной области	ESG-отчеты	Финансовые новости	Финансовые новости	Финансовые новости	ESG-отчёты

более распространенными и прозрачными с точки зрения ESG-риков и тем являются рейтинги MSCI ESG Ratings [9] (включая MSCI ESG Industry Materiality Map [4]) и S&P Global Corporate Sustainability Assessment [10].

2.2. Обзор решений с открытым исходным кодом

Популяризация моделей глубокого обучения в финансовой индустрии привела к развитию тематического анализа ESG-новостей и отчетов с использованием моделей NLP. Использование моделей NLP для анализа ESG-новостей позволяет финансовым экспертам и специалистам по ESG получить более глубокое представление о влиянии событий, тенденций и достижений, связанных с ESG, на компании и отрасли. Тематический анализ превосходит простой анализ настроений, концентрируясь на распознавании и классификации конкретных тем и сюжетов, связанных с экологическими, социальными аспектами и управлением, на основе обширных текстовых данных. Модели глубокого обучения, особенно нейронные сети, продемонстрировали высокую эффективность в автоматическом обнаружении закономерностей и связей в текстах, что делает их подходящими для проведения тематического анализа [11, 12].

На рынке уже есть модели, которые классифицируют ESG-новости. Мы приводим сравнительную табл. 1, которая позволяет проанализировать различия между ними.

[4] https://www.msci.com/our-solutions/esg-investing/esg-industry-materiality-map

Был разработан ряд библиотек для классификации текстов по категориям ESG.

ESG-BERT[5] это текстовый классификатор с открытыми весами. Эта библиотека позиционируется авторами как специфичная для области устойчивого развития модель BERT (Bidirectional Encoder Representations from Transformers) для интеллектуального анализа текста.

FinBERT-ESG[6] основан на модели FinBERT [13]. FinBERT-ESG был доработан на основе 2000 аннотированных вручную предложений из ESG-отчетов и годовых отчетов компаний. Модель была обучена классифицировать тексты по четырем ESG-темам (E, S, G или текст другого типа).

FinBERT-esg-9-categories[7] дополняет модель FinBERT-ESG и была разработана теми же авторами. Модель была доработана примерно на 14 000 предложениях, аннотированных вручную, и классифицирует тексты по девяти ESG-темам: Изменение климата, Природный капитал, Загрязнение окружающей среды и отходы, Человеческий капитал, Ответственность за качество продукции, Общественные отношения, Корпоративное управление, Деловая этика и ценности, Не связанные с ESG.

FinBERT-ESG-Refinitive [14] была предобучена по новостям, связанным с финансами. Существует два различных контекста: задача с несколькими метками и многоклассовая классификация с 20 и 9 классами соответственно. Акцент в этих классах в рамках модели делается в первую очередь на социальных и управленческих аспектах, причем только один класс охватывает все экологические явления.

Как видно из табл. 1, большинство моделей предназначены для решения задачи многоклассовой классификации. Однако задача с несколькими метками в анализе новостей предпочтительнее задачи с единственной меткой среди нескольких классов [15]. Это важно при анализе ESG-новостей из-за сложного и взаимосвязанного характера экологических, социальных вопросов и вопросов управления. Задача с несколькими метками позволяет охватить несколько актуальных ESG-тем в каждой новостной статье, устраняя двусмысленность, поддерживая всесторонний анализ и позволяя расставлять приоритеты на основе релевантности и значимости. Такой подход обеспечивает большую точность и гибкость, что делает его идеальным для понимания многогранных аспектов ESG-тем в финансовой индустрии.

[5] https://github.com/mukut03/ESG-BERT
[6] https://huggingface.co/yiyanghkust/finbert-esg
[7] https://huggingface.co/yiyanghkust/finbert-esg-9-categories

В контексте ESG-новостей, где статьи часто охватывают многогранные темы и взаимосвязанные аспекты, иерархическая классификация обеспечивает тонкий подход для эффективной категоризации и анализа различных аспектов воздействия на окружающую среду, социальной динамики и практики управления. Организовывая темы в иерархию классов, этот метод обеспечивает более детальный анализ новостного контента, позволяя исследователям, аналитикам и заинтересованным сторонам различать не только крупные темы, но и тонкие нюансы внутри каждой из них.

3. МЕТОДЫ ИССЛЕДОВАНИЯ

Разработка модели ESGify включала в себя три последовательных этапа: формирование классов ESG-рисков, подготовку корпуса текстов для обучения модели и обучение классификатора. Первый этап направлен на подготовку достаточно полной методологии, отражающей практику управления ESG-рисками. Следующий этап включает в себя подготовку данных, методологии и инфраструктуры для организации процесса разметки. Последним этапом является разработка классификатора, который состоит из адаптации языковой модели к предметной области и тренировки полносвязной классификационной модели на скрытых состояниях языковой модели.

3.1. Разработка методологии классификации ESG-рисков

Как видно из раздела 2.1, единой методологии классификации ESG-рисков не существует. Все перечисленные методологии содержат различные наборы ESG-рисков или тем и метрик для их измерения. Каждая методология может отличаться подходом к существенности (существенность воздействий или финансовая существенность) и степенью детализации (группировки) рисков. Например, экологические риски для отрасли воздушного транспорта включают 2 категории рисков в методологиях SASB и MSCI и 8 типов рисков в руководстве ЕБРР.

Проблема расхождения ESG-рейтингов начинается с различий между иерархиями тем и показателей, которые используются для оценки ESG-эффективности.

Классификация ESG-рисков, лежащая в основе модели ESGify, представляет собой детализированную систему, которая может применяться для измерения как воздействий, так и финансово значимых аспектов. При разработке системы классификации группа экспертов по ESG проанализировала методологии международных стандартов и рейтингов, упомянутых выше. В методологии были учтены знания и опыт экспертной группы по

Рис. 1. Визуализация методологии иерархической классификации ESGify.

оценке ESG-рисков для различных отраслей в рамках due diligence для принятия финансовых решений.

В нашей методологии мы постарались найти баланс между полнотой, детализацией и представлением событий. Мы уделили особое внимание существенности рисков, все классы которых оказывают финансовое влияние, либо воздействие на реципиента в области устойчивого развития. Методология классификации разработана с единой степенью детализации рисков. Мы также стремимся избегать пересечений между классами, чтобы упростить дальнейшее использование использование в автоматизированных системах. Мы показываем иерархическое представление типов событий с агрегированием, чтобы проиллюстрировать нашу логику.

В процессе разработки категоризации группа из 10 экспертов по ESG, имеющих опыт анализа и оценки ESG-рисков, использовала следующие подходы:

• Анализ лучших практик – методологий международных стандартов и рейтингов, упомянутых выше;

• Обобщение экспертных знаний и опыта экспертной группы по оценке ESG-рисков в различных сферах деятельности;

• Классификация рисков на основе фактора риска, источника риска и реципиента риска, т.е. компонента окружающей среды или группы заинтересованных лиц, на которую влияет наступление риска;

• Кластеризация ESG-рисков в зависимости от их возникновения с учетом существенности последствий событий, связанных с рисками.

Система классификации имеет три уровня иерархии: аспекты (экологический, социальный, управленческий), классовая группа (тематическая группа) и классы (темы). Общее количество классов составляет 47 (см. табл. 4, табл. 5, табл. 6, рис. 1).

Наша конечная цель – разработать универсальную классификацию, применимую к компаниям с различными видами деятельности. Модель ESGify основана на первой версии такой классификации.

Например, чтобы охватить все экологические риски, эксперты включили различные категории для каждого компонента окружающей среды (вода, воздух, почва и т.д.). Чтобы учесть сложный характер и регулирование выбросов парниковых газов и климатических рисков, они были выделены в отдельный класс “Климатические риски и воздействия”, в то время как все вещества окру-

жающей среды, загрязняющие воздух, рассматриваются в категории "Загрязнение атмосферного воздуха".

Некоторые категории разделены на несколько подклассов, чтобы отразить различные значимые риск-события. В области управления водными ресурсами были созданы три подкласса: потребление воды, управление сточными водами и загрязнение поверхностных вод (которое может быть вызвано чрезвычайной ситуацией, а не сбоем системы управления сточными водами). Существует также отдельный класс "Загрязнение почвы и подземных вод". Для каждого инцидента, связанного с водой, обычно существуют разные нормы экологического законодательства и разные методы оценки ущерба окружающей среде. Таким образом, эти события должны быть представлены в различных классах. С помощью такого подхода любые существенные новости о рисках, связанных с водой, будут проанализированы и классифицированы.

Мы также создали категории для восстановления земель, которым часто пренебрегают компании защиты животных, поскольку эта проблема вызывает большой интерес у общественности, управления окружающей средой, касающегося оценки воздействия, лицензий на эксплуатацию и т.д.

Есть также темы, которые разделены между аспектами, такие как "Отведение земель и переселение", которые можно найти в экологическом и социальном аспектах, означающих различные типы рисков, – в экологическом аспекте это связано с переводом земли из одной категории в другую, в то время как в социальном аспекте это относится к экономическому переселению.

Двумя ключевыми ограничениями текущей версии методологии ESGify являются ограничения экспертов, поскольку они в основном использовали свой опыт при анализе российских компаний с акцентом на экологические аспекты, присущие производственной деятельности; и ограничения предварительных условий, поскольку классификатор настроен только для негативных (рискованных) событий, которые представляют наибольший интерес – в будущем его можно расширить до позитивных событий.

Несмотря на ограничения, мы считаем, что методология ESGify обладает высоким потенциалом для использования в качестве универсальной системы классификации ESG-рисков.

3.2. Подготовка тренировочного корпуса текстов

3.2.1. Данные. Мы собрали ESG-отчеты и отчеты об устойчивом развитии, совокупный объем которых превысил 2000 страниц, чтобы разработать набор данных для адаптации к предметной области без учителя. Затем мы извлекли более 17 000 релевантных ESG-новостей, новостей из платформы RepRisk[8], чтобы сделать аннотацию. RepRisk – это аналитическая платформа, которая проводит ESG-оценки компаний и собирает новости о них со всего мира. Поскольку корпус RepRisk содержит только новости, связанные с ESG, мы также использовали корпус текстов, не относящихся к ESG-тематике, из набора данных CNN/Daily Mail [16] для достижения способности модели классифицировать тексты как связанные с ESG и не связанные с ESG.

3.2.2. Процесс аннотирования. При разработке модели мы придерживались парадигмы с несколькими метками, где каждый текст может быть помечен (аннотирован) более чем одной меткой. Эта парадигма позволяет создавать более точные текстовые описания. Однако это требует больше времени, чем многоклассовая маркировка.

10 экспертов по ESG были вовлечены в процесс аннотирования для достижения высокого качества набора данных. Чтобы обеспечить согласованность аннотаций и уменьшить количество человеческих ошибок, мы организовали процесс аннотирования таким образом, чтобы по крайней мере 15% текстов было обработано 3 специалистами. Мы также разработали клиент-серверное приложение, учитывающее специфику описания иерархического класса и ускоряющее процесс аннотирования. После разметки мы также перепроверили размеченный корпус. В общей сложности мы разметили (аннотировали) и проверили 2000 текстов, относящихся к ESG-рискам.

3.3. Обучение классификатора

3.3.1. Адаптация к предметной области. Обучение языковых моделей требует большого объема текстов для достижения достаточного качества языкового моделирования [17]. Это одна из главных причин, по которой модели NLP обычно обучаются без связи с какой-либо областью знаний. Поскольку модели обучаются на общих наборах данных, они часто испытывают трудности, когда сталкиваются с конкретными областями. Адаптация моделей к целевой области может значительно повысить их производительность [13] по мере того, как они становятся более приспособленными к конкретным лингвистическим конструкциям, терминологии и контекстуальным подсказкам, уникальным для данной предметной области. В нашем исследовании мы адаптировали общие модели к предметной области (перенесли на предметную область), используя данные 23 отчетов об устойчивом развитии (объ-

[8] https://www.reprisk.com

емом до 2000 страниц), полученные от компаний, работающих в различных секторах экономики.

Процесс адаптации (подстройки под предметную область) состоит из решения задачи восстановления маскированных слов (MLM). Мы использовали предварительно обученную модель MPNet-base [2] в качестве отправной точки и дообучаем ее в течение трех эпох с размером пакета 128 и скоростью обучения 5e-5. Общее количество обучающих примеров составило 31 432 с максимальной длиной 100 токенов, а вероятность маскирования слова составила 0.15 при использовании маски из целого слова. Количество проверочных примеров с теми же параметрами составило 15 482.

3.3.2 Классификация. Адаптированная к предметной области модель не может напрямую ответить на основную задачу, но уже понимает информацию, связанную с ESG. Чтобы решить задачу классификации ESG-текстов, нам нужно обучить дополнительный слой поверх модели сжатия.

Прежде, чем обучить основную модель, мы подготовили набор данных, состоящий из новостей. Для различных секторов экономики ESG-риски могут иметь значительные диспропорции. Более того, некоторые компании занимают доминирующие позиции в секторе, и это вызывает ложные ассоциации между названиями компаний и риском. Чтобы решить эту проблему, мы замаскировали обучающий набор данных.

Мы использовали Flair Named Entities Recognition model [18] для поиска и маскирования следующих классов: организация, дата, страна и человек. Мы используем специальные токены для каждого класса и добавляем их в словарь токенизатора.

Как уже упоминалось, набор данных состоит из 2000 аннотированных текстов. Поскольку мы также помечаем новости как не относящиеся к ESG, мы добавили в набор данных 120 не относящихся к теме новостей с целью расширить границы применимости нашей модели. В тестовом наборе используется 400 аннотированных текстов. Другие (остальные) тексты были использованы в процессе обучения.

Мы разделили обучающий корпус текстов на обучающие и валидационные подмножества в пропорции 2 : 1 и провели перекрестную валидацию. Мы используем постановку задачи с несколькими метками, и каждый образец может быть помечен более чем одной меткой. Соответственно, для разделения на обучение/валидацию мы используем алгоритм итеративной стратификации [19, 20] для корректного распределения выборок.

Результатом работы модели сжатия является набор векторов, размер которого зависит от исходного текста. Поскольку мы решаем задачу классификации, необходимо получить одинаковую размерность для всех выборок. Мы используем операцию усреднения над набором векторов сжатого пространства для получения векторов одинакового размера. Мы также ограничиваем максимальное количество входных токенов до 512. Далее мы классифицируем текст с помощью двух полносвязных слоев, вставив ReLU функцию активации и случайное зануление части весов между первым и вторым слоем нейросети.

Тренировочный процесс состоит из двух этапов. Первый – это обучение линейных слоев с постоянными весами модели сжатия. В качестве функции потерь мы выбрали взвешенную двоичную кросс-энтропийную потерю. Выбранные веса соответствуют обратной частоте классов в обучающем наборе. Мы делаем это, чтобы лучше учитывать редкие классы. Второй этап – это обучение с более низкой скоростью обучения без ограничений на изменение весов для всех моделей.

Мы также дополняем набор обучающих данных двумя различными способами. Первый метод заключается в применении дополнения к исходным текстам. Мы заменяем случайные слова из текста синонимами с вероятностью 0.15. Далее мы используем технику обратного перевода. Этот метод состоит из двух последовательных переводов на французский и английский язык после этого. Из-за неточностей в процессе перевода мы получаем синонимизированный текст с той же центральной идеей, что и в исходном тексте. Мы используем библиотеку textaugment [21].

Второй метод расширения – это создание синтетических примеров новостей с использованием большой языковой модели (LLM). Мы используем модель LLama 2-Chat 13B [22]. С описанием класса и примером для этого класса мы пишем запросы для генерации новых примеров. Пример запроса и ответа приведен в табл. 2.

Чтобы создать расширенный набор данных, мы используем малопредставленные классы новостей из обучающего набора данных. После этого, основываясь на косинусном расстоянии между сгенерированными примерами и исходными новостями, мы исключаем синтетические новости со значениями косинусного сходства более 0.8. С помощью этого метода мы добавили искусственные новости в мультиклассовой парадигме, но эти новости также могут быть дополнительно аннотированы языковой моделью для создания элементов с несколькими метками.

Таблица 2. Пример с запросом для Llama 2

Запрос	Пример	Результат работы Llama 2
“SYS” You are News article creator. Your task is generation News texts with ESG thematics.Write only News. Write News no longer than 400 words.Your News texts must be about Landscape Transformation. It is events of landscape transformation, including fragmentation and violation of landscape integrity, for example, as a result of construction, mining, laying of linear objects (power lines, pipelines, etc.) . Please, highlight the article start by tags <news> .Article finish by </news>. Next will be example of news text. It highlited by <ex> </ex> .You MUST create new News text! New article MUST NOT PARAPHRASE EXAMPLES!!! “/SYS” <ex> </ex> [/INST]	Exxon Mobil subsidiary among companies questioned over quakes in Texas-Exxon Mobil subsidiary EOG Resources is among several companies facing questions about its disposal of wastewater from hydraulic fracturing operations in the wake of numerous earthquakes that have recently hit the Dallas-Fort Worth area Several studies suggest that the use of injection wells to dispose wastewater from fracking and oil-and-gas production may be linked to increased seismic activity While some oil players deny the studies’ claims, the chief executive of ConocoPhillips, a major shale driller, admits there is a link	Oil Spills Threaten Alaskan Wildlife RefugeA recent oil spill in Alaska’s pristine Arctic National Wildlife Refuge has raised concerns about the environmental impact of landscape transformation caused by human activities. The spill, which occurred during the construction of a new oil pipeline, has released thousands of gallons of crude oil into the fragile ecosystem, threatening the habitat of endangered species such as polar bears, caribou, and wolves.The Arctic National Wildlife Refuge, located in northeast Alaska, is home to some of the most diverse and sensitive ecosystems on Earth.

Таблица 3. Таблица содержит усредненные метрики качества для сравниваемых моделей, полученные методом перекрестной проверки. Метрики F1, Precision и Recall показаны для проверочного набора данных с пороговым значением 0.5 для всех меток/для тестового набора данных с пороговым значением 0.5 для всех меток/для тестового набора данных с пороговым значением, оптимизированным максимизацией F1 для проверочного набора данных для каждой метки. Для AUROC показаны результаты для проверочного/тестового набора данных

Модель	F1-Macro	F1-Weighted	Precision-Weighted	Recall-Weighted	AUROC-Weighted
ROBERTA	0.22/0.27/0.40	0.36/0.40/0.48	0.24/0.29/0.46	0.89/0.80/0.54	0.87/0.83
ROBERTA BTR Augm	0.23/0.36/0.37	0.37/0.45/0.46	0.27/0.39/0.43	0.80/0.58/0.54	0.85/0.82
ROBERTA LLama Augm	0.22/0.39/0.41	0.34/0.47/0.49	0.21/0.37/0.46	0.91/0.66/0.57	0.87/0.83
MPNET	0.22/0.28/0.42	0.36/0.42/0.50	0.24/0.30/0.48	0.88/0.79/0.56	0.87/0.84
MPNET BTR Augm	0.24/0.37/0.39	0.38/0.45/0.47	0.28/0.40/0.45	0.80/0.58/0.54	0.86/0.82
MPNET LLama Augm	0.22/0.40/0.44	0.34/**0.48**/**0.50**	0.22/0.40/0.48	0.90/0.65/0.56	0.87/0.84
ESG-BERT	0.22/0.29/0.41	0.37/0.42/0.49	0.25/0.31/0.49	0.87/0.77/0.55	0.86/0.83
ESG-BERT BTR Augm	0.23/0.37/0.40	0.37/0.46/0.48	0.25/0.38/0.46	0.83/0.61/0.53	0.85/0.81
ESG-BERT LLama Augm	0.21/0.39/0.43	0.33/0.47/0.49	0.21/0.38/0.48	0.90/0.65/0.56	0.86/0.82
ChatGPT 3.5 turbo	–/–/0.24	–/–/0.29	–/–/0.35	–/–/0.32	–/–

4. РЕЗУЛЬТАТЫ

В этом разделе мы рассматриваем основные результаты для каждой части модельного конвейера, испытываем архитектуру, оцениваем конечное качество модели и обсуждаем применимость результатов.

4.1. Классификационная модель

Как уже было сказано ранее 3.3.1, мы использовали метод адаптации к предметной области, чтобы улучшить способность наших моделей кодировать тексты, связанные с ESG. Мы адаптировали две предварительно обученные модели:

Таблица 4. Описание экологических классов, сгруппированных по темам и аспектам

Тема	Класс	Описание
Климатические риски и воздействия	Климатические риски	Физические климатические риски (включая климатические изменения, такие как жара, нехватка воды, наводнения, береговая эрозия, лесные пожары, таяние вечной мерзлоты и другие) и переходные климатические риски (риски, связанные с переходом к низкоуглеродной экономике, включая меры, принимаемые правительствами и регулирующими органами стран, подписавших Парижское климатическое соглашение, направленное на предотвращение изменения климата, которые подразделяются на политические, правовые, технологические, рыночные и репутационные риски)
Климатические риски и воздействия	Выбросы парниковых газов	Риски, связанные с увеличением выбросов парниковых газов, несоблюдение требований в отношении сокращения выбросов, раскрытия информации и т.д.
Загрязнение атмосферного воздуха	–	Риски, связанные с загрязнением атмосферного воздуха
Водопотребление и водоотведение	Водопотребление	Риски, связанные с потреблением воды, включая районы, испытывающие нехватку воды
Водопотребление и водоотведение	Управление сточными водами	Риски, связанные с отведением сточных вод и загрязнением водоемов
Водопотребление и водоотведение	Загрязнение поверхностных вод	Риски, связанные с загрязнением водных объектов (включая аварийные разливы, утечки, то есть не связанные состочными водами)
Воздействие на почвы и подземные воды	–	Риски, связанные с негативным воздействием на почвы и подземные воды
Обращение с отходами	–	Риски, связанные с утилизацией отходов (захоронение, переработка, хранение различных видов и опасных классов отходов)
Энергоэффективность и возобновляемые источники энергии	–	Риски, связанные с неэффективным потреблением энергии, низкой энергоэффективностью и отсутствием или низкой долей возобновляемых источников энергии в энергетическом балансе
Обращение с опасными материалами	–	Риски, связанные с управлением опасными материалами, когда несоблюдение требований по обращению с материалами может причинить экологический ущерб
Природные ресурсы	–	Риски, связанные с нерациональным потреблением природных ресурсов, топлива, полезных ископаемых и т.д.
Планировочные ограничения	–	Риски, связанные с нарушением планировочных ограничений (нарушения режимов зон с особыми условиями, включая водоохранные, санитарно-защитные зоны, зоны санитарной охраны источников водоснабжения, ООПТ, населенных пунктов, промышленных зон, аэропортов и т.д.)
Трансформация ландшафтов	–	Риски трансформации ландшафтов, включая фрагментацию и нарушение целостности, например, в результате строительства, добычи, укладки линейных объектов (линии электропередачи, трубопроводы и т.д.)
Восстановление территорий	–	Риски неэффективности или неполноты восстановления/очистки территорий после аварий или из-за общего негативного воздействия компании
Биоразнообразие	–	Риски, связанные с воздействием на биоразнообразие, включая браконьерство, незаконную вырубку лесов, уничтожение редких и краснокнижных видов, нарушение режима ООПТ и т.д.

Таблица 4. Окончание

Тема	Класс	Описание
Защита животных	–	Риски, связанные с правами животных, включая тестирование, бесчеловечное обращение с животными
Чрезвычайные ситуации (экологический аспект)	–	Риски возникновения чрезвычайных ситуаций с негативными экологическими последствиями
Экологический менеджмент	–	Риски, связанные с неэффективностью экологического менеджмента, включая нарушения законодательства, отсутствие или несоблюдение разрешений, нарушения, выявленные в ходе инспекций регулирующих и надзорных органов, штрафы, иски, касающиеся экологического законодательства. Класс включает комплексные воздействия, связанные сразу с несколькими компонентами окружающей среды
Физические воздействия	–	Риски, связанные с физическим воздействием, такие как шум, вибрация, электромагнитное излучение, радиация
Отведение земель (экологический аспект)	–	Риски, связанные с изъятием, включая перевод земель из одной категории в другую, деятельность на землях неподходящей категории
Цепочка поставок (экологический аспект)	–	Экологические риски в цепочке поставок (поставщики и подрядчики)

MPNet-base и Roberta-base. Во время проверки процесса обучения перекрестные потери энтропии уменьшились на 24 процента для модели ROBERTA и на 22 процента для модели MPNet. Это демонстрирует улучшение обеих моделей для реконструкции текстов, связанных с ESG.

Для получения модели с максимальным качеством мы сравниваем MPNet-base, Roberta-base и ESG-Bert (без классификационного слоя) с различными методами аугментации. Более того, мы сравниваем наши модели с ChatGPT на тестовом наборе. В качестве запроса для ChatGPT мы используем простое определение задачи и перечисление классов. Результаты показаны в табл. 3.

Эти показатели включают в себя F1-Macro, F1-Weighted, Precision-Weighted, Recall-Weighted, и AUROC-Weighted. Оценка выполняется для проверочного набора данных с пороговым значением 0.5 для всех меток, тестового набора данных с пороговым значением 0.5 для всех меток и тестового набора данных с пороговым значением, оптимизированным с помощью максимизации F1 в наборе данных проверки для каждой метки.

Среди представленных моделей базовые версии ROBERTA, MPNet и ESG-BERT демонстрируют лучшие метрики качества по показателям F-1 в тестовом наборе, оптимизированном для пороговых значений, чем версии этих моделей с данными, дополненными методом обратного перевода (BTR). Это указывает на то, что метод аугментации, примененный к этим моделям, не улучшает их баланс между показателями точности и полноты. Одной из вероятных причин является сложность корректной аугментации в задачах с несколькими метками при использовании небольших наборов данных и редких меток. Аугментация с помощью LLama 2 изменила эту ситуацию благодаря возможности генерировать качественные синтетические новости и улучшать весь процесс обучения для достижения лучших показателей.

Результаты показывают, что оптимизация порогового значения на основе максимизации F1 в наборе данных проверки для каждой метки может привести к повышению качества на тестовом наборе данных. Такой подход позволяет регулировать баланс между точностью и полнотой в зависимости от конкретных требований задачи.

Также наши модели демонстрируют аналогичную или лучшую производительность по сравнению с одной из самых популярных ESG-моделей. Этот факт косвенно указывает на то, что для адаптации предметной области достаточно небольшого объема текста. Модель, зависящая от предметной области, демонстрирует гораздо лучшую производительность, чем большая языковая модель с подсказкой. Наше исследование не направлено на разработку оптимальных подсказок для больших языковых моделей. Однако результаты эксперимента демонстрируют сравнимую точность и меньшую полноту для большой языковой модели в сравнении с моделью, адаптированной под предметную область.

Одним из важных ограничений этого исследования является ограниченный объем данных, доступных для оценки, что влияет на качество моделей. При более крупном и разнообразном наборе данных модели могли бы демонстрировать иное

Таблица 5. Описание социальных классов, сгруппированных по темам и аспектам

Тема	Класс	Описание
Взаимодействие с местным населением	Коренные народы	Риски, связанные с нарушениями прав коренных народов
Взаимодействие с местным населением	Права человека	Риски, связанные с нарушениями прав местных сообществ, которые могут привести к протестам, включая нарушения прав уязвимых социальных групп, нарушения доступности экосистемных услуг для населения (Примеры экосистемных услуг: (1) обеспечивающие услуги: включают продукты питания, пресную воду, древесину, волокна и лекарственные препараты заводы; (2) регулирующие услуги: включают очистку поверхностных вод, хранение и связывание углерода, регулирование климата и защиту от стихийных бедствий; (3) культурные услуги: включают природные зоны, которые являются священными местами, а также территории рекреационного и эстетического значения; (4) поддерживающие услуги: включают почвообразование, круговорот питательных веществ и формирование первичной продукции.)
Взаимодействие с местным населением	Здоровье и безопасность населения	Риски, связанные с воздействием на здоровье и безопасность населения в результате деятельности компании, чрезвычайных ситуаций, включая действия сотрудников службы безопасности компании
Безопасность и качество продукции	–	Риски, связанные с качеством и безопасностью продукции, включая неправильную маркировку
Безопасность данных	–	Риски, связанные с управлением данных о клиентах, сотрудниках и партнерах
Отведение земель и переселение (социальный аспект)	–	Риски, связанные с вынужденным экономическим переселением, компенсациями за изъятие земель или потерю выгоды, нарушениями процесса консультаций с местными жителями, жалобами, судебными исками
Чрезвычайные ситуации (социальный аспект)	–	Риски возникновения чрезвычайных ситуаций, оказывающих негативное воздействие на сотрудников и местные сообщества, социальную инфраструктуру
Здоровье и безопасность сотрудников	–	Риски, связанные с нарушениями здоровья и безопасности сотрудников, включая пожары, электро- и транспортную безопасность. Сюда относятся несчастные случаи, профессиональные и другие заболевания, нарушения в эксплуатации объектов и оборудования, пожары, взрывы. Только в отношении сотрудников компании
Трудовые отношения	Профессиональные объединения	Риски, связанные с нарушениями коллективных договоров и права на организацию, профсоюзную деятельность, забастовки
Трудовые отношения	Принудительный труд	Риски, связанные с использованием принудительного труда, современное рабство
Трудовые отношения	Детский труд	Риски, связанные с использованием детского труда и нарушением минимального возраста работников
Трудовые отношения	Дискриминация	Риски, связанные с дискриминацией и нарушениями равных возможностей
Трудовые отношения	Сокращение штатной численности	Риски, связанные с сокращением штатов (индивидуальные и массовые увольнения)
Трудовые отношения	Управление трудовыми отношениями	Риски, связанные с условиями труда и правами сотрудников, включая заработную плату, пособия и льготы, часы работы и т.д.
Культурное наследие	–	Риски, связанные с воздействием и разрушением культурного наследия, включая религиозные, исторические и археологические объекты
Цепочка поставок (социальный аспект)	–	Социальные риски в цепочке поставок (поставщики и подрядчики)

Таблица 6. Описание классов управления, сгруппированных по темам и аспектам

Тема	Класс	Описание
Экономические преступления	–	Риски, связанные с экономическими преступлениями, включая мошенничество, уклонение от налогов, коррупцию, нарушения антимонопольного законодательства, антиконкурентные нарушения, рыночные манипуляции, фиксацию цен, нарушения закупочной деятельности
Нарушения законодательства и судебные разбирательства	–	Риски, связанные с судебными исками, штрафами, нарушениями закона
Ответственное финансирование и гринвошинг	–	Риски, связанные с несоблюдением стратегий финансовых и других компаний в области ESG, гринвошингом (введение в заблуждение относительно ESG-параметров компаний или финансовых инструментов)
Корпоративное управление и деловая этика	Корпоративное управление	Риски неэффективного корпоративного управления, несоблюдение кодекса корпоративного управления, риски управления, связанные с акционерами, бенефициарами, руководством
Корпоративное управление и деловая этика	Деловая этика	Риски, связанные с несоблюдением корпоративных ценностей и деловой этики
Корпоративное управление и деловая этика	Реализация стратегии	Риски, связанные с ошибками при осуществлении стратегии
Корпоративное управление и деловая этика	Управление рисками и внутренний контроль	Риски, связанные с ошибками в управлении рисками и внутреннем контроле
Корпоративное управление и деловая этика	Раскрытие информации	Риски несоблюдения требований в отношении раскрытия информации, информационных правил и практики, включая недобросовестный маркетинг
Цепочка поставок (экономические/управленческие риски)	–	Экономические и управленческие риски в цепочке поставок (поставщики и подрядчики)

поведение и потенциально достигать более высокого качества. Как было показано, расширение набора данных с помощью модели LLM может улучшить результаты классификации. Поэтому обобщение данных результатов на более масштабные задачи классификации текстов следует производить с осторожностью.

5. ЗАКЛЮЧЕНИЕ

В этом исследовании представлена одна из крупнейших открытых методологий классификации ESG-рисков, основанная на выводах экспертов по ESG и передовой мировой практике. С помощью аннотирования 2000 новостных статей мы подготовили набор данных классификации с несколькими метками, который послужил основой для разработки модели NLP, предназначенной для автоматической классификации рисков в соответствии с установленной методологией. На этапе тщательного обучения наша модель показала достойное значение метрики качества F1, равное 0.5 для тестового набора. Примечательно, что при сравнении с моделью ChatGPT 3.5 с помощью простой подсказки наш подход продемонстрировал лучшее значение метрик качества при меньшем размере модели.

Кроме того, признавая проблемы, связанные с ограниченными наборами данных, мы изучили методы дополнения для повышения эффективности модели. Это исследование выявило потенциал использования больших языковых моделей (LLM) в качестве инструментов для аугментации, в очередной раз подтверждая их полезность для решения проблемы дефицита данных и улучшения результатов классификации.

Мы ожидаем, что достигнутые результаты будут полезны при разработке комплексной системы автоматизированной оценки ESG-рисков на основе NLP и для уточнения отраслевых стандар-

тов. Веса модели и документация размещены на Github https://github.com/sb-ai-lab/ESGify под лицензией Apache 2.0.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

В процессе обучения модели NLP мы произвели 13.26 кг CO_2 на NVIDIA A100 80GB, это было рассчитано с помощью Eco2AI lib [23].

СПИСОК ЛИТЕРАТУРЫ

1. *Gary Bell, Francesca, Vuuren*. The impact of climate risk on corporate credit risk. Cogent Economics Finance, 10, 11 2022.
2. *Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, Tie Yan Liu*. Mpnet: Masked and permuted pre-training for language understanding, 2020.
3. What are the principles for responsible investment? https://www.unpri.org/about-us/what-are-the-principles-for-responsible-investment.
4. Ifc performance standards on environmental and social sustainability. https://www.ifc.org/en/types/insights-reports/2012/publications-handbook-pps.
5. World Bank Environmental and Social Framework. World Bank, Washington, DC, 2016.
6. EBRD Environmental, Social Risk Management Toolkit For Financial Intermediaries. The European Bank for Reconstruction and Development (EBRD), 2016.
7. Consolidated Set of the GRI Standards. Global Reporting Initiative, 2023.
8. SASB Standards. IFRS Foundation, 2023.
9. ESG Ratings Methodology. MSCI ESG Research LLC, 2023.
10. CSA Handbook 2023. Corporate Sustainability Assessment. S&P Global, 2023.
11. *Hanseon, Choi Hayoung, Cheon Minjong Lee, Ook, Joo.* Proposing an integrated approach to analyzing esg data via machine learning and deep learning algorithms, 2022.
12. *Woojoo, Jung Sang-Uk Park, Joonbeom, Choi*. Exploring trends in environmental, social, and governance themes and their sentimental value over time. Frontiers in Psychology, 13, 2022.
13. *Dogu Araci*. Finbert: Financial sentiment analysis with pre-trained language models, 2019.
14. *Nicole, Leidner Jochen L. Nugent, Tim, Stelea*. Detecting environmental, social and governance (esg) topics using domain-specific language models and data augmentation. Berlin, Heidelberg. Springer-Verlag, 2021.
15. *Jasmin Bogatinovski, Ljupčo Todorovski, Sašo Džeroski, Dragi Kocev*. Comprehensive comparative study of multi-label classification methods. Expert Systems with Applications. 2022. V. 203. P. 117215.
16. *Abigail See, Peter J. Liu, Christopher D. Manning.* Get to the point: Summarization with pointergenerator networks. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1073–1083, Vancouver, Canada, July 2017. Association for Computational Linguistics.
17. *Miguel, Wanner Leo P'erez-Mayos, Laura, Ballesteros.* How much pretraining data do language models need to learn syntax? arXiv preprint arXiv:2109.03160, 2021.
18. *Duncan, Vollgraf Roland Akbik, Alan, Blythe.* Contextual string embeddings for sequence labeling. In COLING 2018, 27th International Conference on Computational Linguistics. 2018. P. 1638–1649.
19. *Grigorios, Vlahavas Ioannis Sechidis, Konstantinos, Tsoumakas*. On the stratification of multi-label data. Machine Learning, Knowledge Discovery in Databases. 2011. P. 145–158.
20. *Piotr Szymański, Tomasz Kajdanowicz.* A network perspective on stratification of multi-label data. In Luís Torgo, Bartosz Krawczyk, Paula Branco, Nuno Moniz, editor, Proceedings of the First International Workshop on Learning with Imbalanced Domains: Theory, Applications, volume 74 of Proceedings of Machine Learning Research, pages 22–35, ECML-PKDD, Skopje, Macedonia, 2017. PMLR.
21. *Tshephisho Marivate, Vukosi, Sefara*. Improving short text classification through global augmentation methods. In International Cross-Domain Conference for Machine Learning, Knowledge Extraction, Springer, 2020. P. 385–399.
22. *Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Dan Bikel, Lukas Blecher, Cristian Canton Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, Brian Fuller, Cynthia Gao, Vedanuj Goswami andNaman Goyal, Anthony Hartshorn, Saghar Hosseini, Rui Hou, Hakan Inan, Marcin Kardas, Viktor Kerkez, Madian Khabsa, Isabel Kloumann, Artem Korenev, Punit Singh Koura, Marie-Anne Lachaux, Thibaut Lavril, Jenya Lee, Diana Liskovich, Yinghai Lu, Yuning Mao, Xavier Martinet, Todor Mihaylov, Pushkar Mishra, Igor Molybog, Yixin Nie, Andrew Poulton, Jeremy Reizenstein, Rashi Rungta, Kalyan Saladi, Alan Schelten, Ruan Silva, Eric Michael Smith, Ranjan Subramanian, Xiaoqing Ellen Tan, Binh Tang, Ross Taylor, Adina Williams, Jian Xiang Kuan, Puxin Xu, Zheng Yan, Iliyan Zarov, Yuchen Zhang, Angela Fan, Melanie Kambadur, Sharan Narang, Aurelien Rodriguez, Robert Stojnic, Sergey Edunov, Thomas Scialom*. Llama 2: Open foundation and fine-tuned chat models, 2023.
23. *Budennyy S., Korovin A., Plosskaya O., Dimitrov Denis, Akhripkin V., Pavlov I., Oseledets Ivan, Barsola I., Egorov I., Kosterina A., Zhukov Leonid Lazarev, Vladimir, Zakharenko*. eco2ai: Carbon emissions tracking of machine learning models as the first step towards sustainable ai. Doklady Mathematics, 106, 01 2023.

ESGIFY: AUTOMATED CLASSIFICATION OF ENVIRONMENTAL, SOCIAL AND CORPORATE GOVERNANCE RISKS

A. Kazakov[a], S. Denisova[a], I. Barsola[a], E. Kalugina[a], I. Molchanova[a], I. Egorov[a], A. Kosterina[a], E. Tereshchenko[a], L. Shutikhina[a], I. Doroshchenko[a], N. Sotiriadi[a], and S. I. Budennyy[a, b]

[a] *Sber, Moscow, Russian Federation*

[b] *Artificial Intelligence Research Institute (AIRI), Moscow, Russian Federation*

Presented by Academician of the RAS A.A. Shananin

The growing recognition of environmental, social, and governance (ESG) factors in financial decisionmaking has spurred the need for effective and comprehensive ESG risk assessment tools. In this study, we introduce an open-source Natural Language Processing (NLP) model, "ESGify"12, based on MPNet-base architecture and aimed to classify texts within the frames of ESG risks. We also present a hierarchical and detailed methodology for ESG risk classification, leveraging the expertise of ESG professionals and global best practices. Anchored by a manually annotated multilabel dataset of 2,000 news articles and dosmain adaptation with texts of sustainability reports, ESGify is developed to automate ESG risk classification following the established methodology. We compare augmentation techniques based on back translation and Large Language Models (LLMs) to improve the model quality and achieve 0.5 F1-weighted model quality in the dataset with 47 classes. This result outperforms ChatGPT 3.5 with a simple prompt. The model weights and documentation is hosted on Github https://github.com/sb-ailab/ESGify under the Apache 2.0 license.

Keywords: ESG, Sustainability, LLM, NLP