Mapping the Landscape of Artificial Intelligence in Climate Change Research: A Meta-Analysis on Impact and Applications

Anonymous Author(s)
Affiliation
Address
email

Abstract

This proposal advocates a comprehensive and systematic analysis aimed at mapping and characterizing the intricate landscape of Artificial Intelligence and Machine Learning applications and their impacts within the domain of climate change research, both in adaption and mitigation efforts. Notably, a significant upswing in this interdisciplinary intersection has been observed since 2020. Utilizing advanced topic clustering techniques and qualitative analysis, we have discerned 12 distinct macro areas that supplement, enrich, and expand upon those identified in prior research. The primary objective of this undertaking is to furnish a data-rich panoramic view and informative insights regarding the functions and tools of the mentioned disciplines. Our intention is to offer valuable guidance to the scholarly community and propel further research endeavors, encouraging meticulous examinations of research trends and gaps in addressing the formidable challenges posed by climate change and the climate crisis.

1 Introduction

Artificial Intelligence (AI) is largely considered a vehicle for climate action and several initiatives have been started [The Global Partnership on Artificial Intelligence, 2022] [BCG, 2022]. Rolnick et al. [2019] offers a comprehensive analysis of this topic, encompassing fundamental concepts, applications, and challenges in using Machine Learning (ML) to combat climate change. Similarly, Nishant et al. [2020] provides a thorough literature review and research agenda, contributing to the discourse on the role of AI in sustainability.

Our motivation to present this proposal is rooted in the conspicuous gap that exists in the academic literature. To the best of our knowledge, there is a notable absence of a comprehensive and systematic overview of the work laying at the intersection between AI and climate change. This need was distinctly highlighted by Rolnick et al. [2019] in their assertion that "[m]any ML practitioners wish to act, but are uncertain how." With our proposal, our aim is to bridge this gap, complement existing studies, and offer the scholarly community a robust and holistic understanding of this crucial intersection. Taking a more quantitative approach, our proposal strives to augment knowledge of ongoing efforts and complement the primarily qualitative work of Rolnick et al. [2019]. By doing so, we seek to strengthen this area of research and provide informative characterizations of the diverse roles that AI can assume in addressing climate change, as well as its potential positive and negative impacts. Inspired by Rolnick et al. [2019], our goal was to characterize the intersection with keywords and descriptive terminology.

Submitted to Tackling Climate Change with Machine Learning: workshop at NeurIPS 2023.

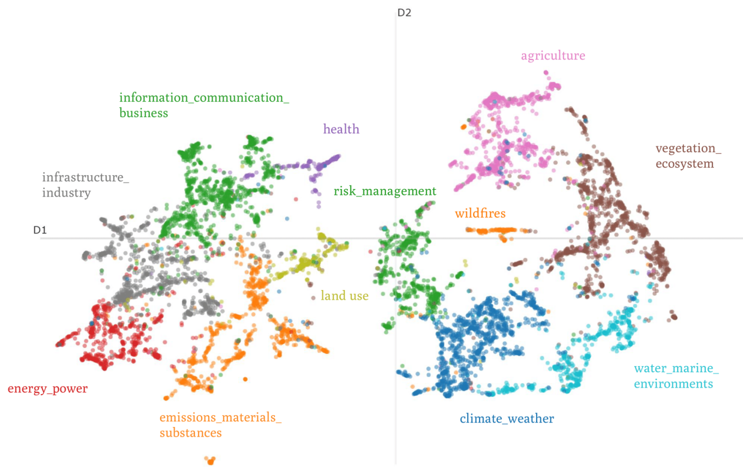

Figure 1: Topic visualizations in a two-dimensional vector space using dimensionality reduction on the document embeddings. This supported interpretable results and manual labelling. Visualizations were produced with the default *BERTopic* colors.

2 Methods

We considered Deep Learning (DL) and ML as integral components of AI. We base our hypothesis on the premise that ML and DL play analogous roles within climate change research and applications. We employed several Natural Language Processing methods to explore the research landscape of AI and climate change. We searched the full *Scopus* scholarly database, extracting articles that included "artificial intelligence" or "machine learning" or "deep learning" **and** "climate change" or "climate crisis" or "global warming" or "climate action" (or "sustainab*" and "climate") in the title, abstracts, or authors' keywords. We selected keywords through a preliminary literature review with the aim of encompassing the widest possible range of research. We kept conference proceedings and articles written in English to ensure a rigorous comparison. Our final dataset consisted of *8,191* items. The dataset is available in our github repository[1].

For the preliminary analysis, we concatenated the titles and abstracts but excluded the authors' keywords to keep the original context for each word in the abstract. We used *BERTopic* [Grootendorst, 2022] to perform *topic modelling*. The model generated 143 *micro* topics which were manually reviewed and labelled by the authors through a qualitative analysis. This process resulted in a clustering of *micro* topics into 12 *macro* categories, represented in Figure 1. We observed that the model generated semantically significant embeddings, as demonstrated by the inter-cluster distances. Finally, we applied additional *keyword* and *n-gram* extraction, targeting a more precise and informative representation for each *macro* class. The use of *dependency parsing* techniques to review the prevalence of common verbs, nouns, and adjectives, allowed us to outline the multifaceted roles and functions of AI in this domain.

3 Preliminary results

In our preliminary results, we present an analysis of the labeled clusters generated by *BERTopic*, including their document counts and keyword representations. These findings (see Table 1 in the Appendix), reviewed in the light of Rolnick et al. [2019][2], reveal several similarities and some differences. Notably, our analysis unveiled variations in cluster compositions. While the **health** cluster did not emerge as a distinct entity in Rolnick et al. [2019], it is important to consider that the surge of this body of works could connect with the *COVID-19* outbreak. The **vegetation and ecosystems** cluster in our work appeared more significant. Our study expanded the scope of the

[1] `https://github.com/jchburmester/ai_cc.git`

[2] We acknowledge the authors' work on ML and climate change, which served as a valuable reference for our research in AI, ML, and DL due to the limited existing research in this field.

water and **marine environments** cluster, encompassing additional subtopics like **snow cover**, **coral reefs**, **marine ecosystems**, **glacial lakes**, and **lake areas**. In our analysis, we identified a distinct and homogeneous **wildfires** cluster, characterized by high compactness.

Evaluating the results obtained through dependency parsing (see the appendix for the complete table), we have identified several prominent adjectives (500+ occurrences) that paint a vivid picture of AI's attributes. AI proves invaluable on both **spatial** and **temporal** features, addressing a spectrum of **complex**, **real-time**, **human-related** challenges. These encompass **physical**, **meteorological**, **solar**, **climatic**, **remote**, and **hydrological** domains, as well as **industrial**, **green**, **renewable**, **agricultural**, **ecological**, **economic**, **social**, and **public** sectors. The integration of **smart**, **hybrid** solutions results in **efficient** and **predictive** functionalities. Simultaneously, verbs (500+) highlight AI's actions and contributions. The ability to **predict**, **support**, **model**, **monitor**, **sense**, **optimize**, **forecast**, **simulate**, **classify**, and **measure** underscores AI's versatility. Additional terms like *reduce*, *decrease*, *grow*, and *increase* provide intriguing nuances, warranting further exploration into the specifics of these alterations. Nouns (500+) encompass **energy**, **water**, **floods**, **forests**, **land**, **fields**, **soil**, and **regions**, among others. This expansive lexicon spans vital concepts like **vegetation**, **power**, **species**, **surface**, **crops**, **carbon**, **temperature**, **health**, **weather**, and **emissions**. Within these domains, targets extend to **production**, **consumption**, **growth**, **resources**, and **risk mitigation**. Key applications in **prediction**, **classification**, **detection**, **monitoring**, **forecasting**, **decision-making**, and **support** underscore AI's role. The dimensions of **scale**, **time**, **images**, **resolution**, **scenarios**, and **satellite** technology enrich this domain, reflecting the multi-faceted nature of AI's involvement.

The temporal analysis of articles revealed that research at the AI and climate change intersection has witnessed substantial growth in recent years, particularly from 2020 onward. The **water** and **marine environments** cluster emerged as the earliest areas of research interest, dating back to 1993. Subsequently, the **climate** and **weather**, the **agriculture**, and the **health** clusters gained prominence in 1994, 1996, and 1997, respectively.

4 Discussion and future directions

The quantitative analysis of scientific works remains challenging, especially when seeking meaningful insights. Our study provides a snapshot of this vast domain, focusing on research findings. To delve deep into this expansive domain and solidify our methodology, we acknowledge that this research would benefit from the involvement of climate change experts (e.g. to refine keyword selection and analysis).

We envision to streamline knowledge and information related to the synergy of AI and climate change in a way to support responsible research and action. Our plan is to compile a catalogue of *go-to* methods and strategies for the community and industry, i.e. researchers and practitioners who are willing to tackle climate change problems. This may result in an instrument enabling people to navigate AI-relevant climate change interventions across different areas. Here, we intend to add to existing efforts[3]. We further want to finely assess where AI solutions and tools can be most impactful, given problem-specific constraints and conditions (e.g. data availability). To this aim, we plan to consider additional methods for the automated analysis (e.g. *Named Entity Recognition*), other features of existing archive resources (Scopus *subject areas* to highlight the diversity of disciplines involved), and the creation of a dashboard for results visualization.

We aim to not only acknowledge the potential benefits but also the ethical considerations and environmental responsibilities associated with AI technologies. The proposal aligns with global initiatives from organizations such as the OECD [OECD, 2021] and the EU [Horizon Magazine, 2019] and previous work highlighting the need for a concerted effort to harness AI for climate change mitigation [Kaack et al., 2022]. We recognize that AI systems can play a positive role in monitoring emissions and energy consumption. However, they can also present environmental challenges due to their energy consumption and emissions. By taking a meticulous approach to our analysis, we hope to provide guidance for future studies.

[3]For example, the Climate Change AI Wiki: `https://wiki.climatechange.ai/wiki/Welcome_to_the_Climate_Change_AI_Wiki`.

4.1 Appendix

Table 1: Clusters representations and magnitudes (in descending order)

Label	Number of abstracts (percentage)	BERTopic keyword representation
Vegetation, ecosystem	1,200 (14.7%)	[’forest’, ’species’, ’soil’, ’climate’, ’change’, ’land’, ’learning’, ’carbon’, ’using’, ’vegetation’, ’machine’, ’climate change’, ’machine learning’, ’based’, ’cover’, ’used’, ’distribution’, ’classification’, ’spatial’, ’remote’, ’sensing’, ’remote sensing’, ’area’, ’forests’, ’tree’, ’use’, ’high’, ’areas’, ’global’, ’changes’]
Climate, weather	1,072 (13.1%)	[’climate’, ’water’, ’learning’, ’temperature’, ’precipitation’, ’change’, ’machine’, ’using’, ’groundwater’, ’based’, ’rainfall’, ’machine learning’, ’climate change’, ’drought’, ’prediction’, ’used’, ’river’, ’future’, ’time’, ’forecasting’, ’deep’, ’high’, ’hydrological’, ’downscaling’, ’global’, ’term’, ’neural’, ’mean’, ’network’, ’long’]
Information, communication, business	955 (11.7%)	[’ai’, ’sustainable’, ’intelligence’, ’learning’, ’social’, ’artificial’, ’artificial intelligence’, ’development’, ’environment’, ’sustainability’, ’industry’, ’climate’, ’technology’, ’education’, ’digital’, ’change’, ’business’, ’students’, ’based’, ’paper’, ’technologies’, ’climate change’, ’new’, ’manufacturing’, ’environmental’, ’machine’, ’human’, ’approach’, ’using’, ’process’]
Infrastructure, industry	892 (10.9%)	[’energy’, ’building’, ’smart’, ’based’, ’environment’, ’learning’, ’sustainable’, ’design’, ’buildings’, ’iot’, ’systems’, ’paper’, ’proposed’, ’performance’, ’cities’, ’consumption’, ’traffic’, ’urban’, ’using’, ’development’, ’intelligence’, ’city’, ’energy consumption’, ’network’, ’machine’, ’sustainability’, ’decision’, ’artificial’, ’time’, ’machine learning’]
Emissions, substances, materials	859 (10.5%)	[’emissions’, ’carbon’, ’air’, ’co2’, ’learning’, ’waste’, ’based’, ’machine’, ’emission’, ’energy’, ’machine learning’, ’pollution’, ’using’, ’environmental’, ’prediction’, ’environment’, ’used’, ’quality’, ’air quality’, ’development’, ’gas’, ’global’, ’climate’, ’neural’, ’time’, ’sustainable’, ’performance’, ’air pollution’, ’high’, ’concrete’]
Agriculture	835 (10.2%)	[’crop’, ’yield’, ’agriculture’, ’climate’, ’agricultural’, ’production’, ’food’, ’learning’, ’using’, ’based’, ’crops’, ’machine’, ’change’, ’machine learning’, ’farmers’, ’climate change’, ’plant’, ’yields’, ’prediction’, ’used’, ’rice’, ’irrigation’, ’soil’, ’farming’, ’water’, ’management’, ’crop yield’, ’disease’, ’different’, ’time’]
Risk management	674 (8.2%)	[’water’, ’flood’, ’decision’, ’management’, ’climate’, ’support’, ’risk’, ’change’, ’based’, ’decision support’, ’climate change’, ’areas’, ’use’, ’information’, ’development’, ’land’, ’systems’, ’using’, ’assessment’, ’dss’, ’learning’, ’urban’, ’used’, ’paper’, ’resources’, ’disaster’, ’planning’, ’flooding’, ’natural’, ’machine’]
Energy, power	595 (7.3%)	[’energy’, ’power’, ’solar’, ’wind’, ’renewable’, ’electricity’, ’pv’, ’based’, ’renewable energy’, ’generation’, ’learning’, ’systems’, ’consumption’, ’using’, ’forecasting’, ’grid’, ’demand’, ’photovoltaic’, ’proposed’, ’load’, ’smart’, ’optimization’, ’electric’, ’paper’, ’sources’, ’machine’, ’prediction’, ’control’, ’deep’, ’used’]
Water, marine environments	540 (6.6%)	[’sea’, ’ice’, ’ocean’, ’learning’, ’water’, ’climate’, ’marine’, ’machine’, ’based’, ’using’, ’coastal’, ’change’, ’machine learning’, ’sea ice’, ’deep’, ’surface’, ’high’, ’global’, ’climate change’, ’snow’, ’temperature’, ’images’, ’resolution’, ’coral’, ’deep learning’, ’satellite’, ’permafrost’, ’arctic’, ’used’, ’spatial’]
Land use	244 (3.0%)	[’urban’, ’built’, ’land’, ’environment’, ’cities’, ’spatial’, ’sustainable’, ’city’, ’learning’, ’heat’, ’built environment’, ’development’, ’landscape’, ’street’, ’areas’, ’based’, ’machine’, ’machine learning’, ’using’, ’planning’, ’lst’, ’urbanization’, ’area’, ’temperature’, ’climate’, ’use’, ’surface’, ’high’, ’ecological’, ’used’]
Health	177 (2.2%)	[’health’, ’disease’, ’climate’, ’risk’, ’covid’, ’covid 19’, ’19’, ’cases’, ’learning’, ’diseases’, ’based’, ’transmission’, ’incidence’, ’machine’, ’factors’, ’change’, ’machine learning’, ’patients’, ’virus’, ’using’, ’climate change’, ’distribution’, ’medical’, ’temperature’, ’infection’, ’prediction’, ’vector’, ’surveillance’, ’areas’, ’used’]
Wildfires	148 (1.8%)	[’wildfire’, ’forest’, ’fires’, ’wildfires’, ’smoke’, ’forest fires’, ’learning’, ’severity’, ’detection’, ’using’, ’climate’, ’area’, ’areas’, ’based’, ’machine’, ’risk’, ’machine learning’, ’change’, ’occurrence’, ’deep’, ’prediction’, ’used’, ’climate change’, ’forests’, ’high’, ’deep learning’, ’burn’, ’images’, ’network’, ’human’]
total	*8,191*	

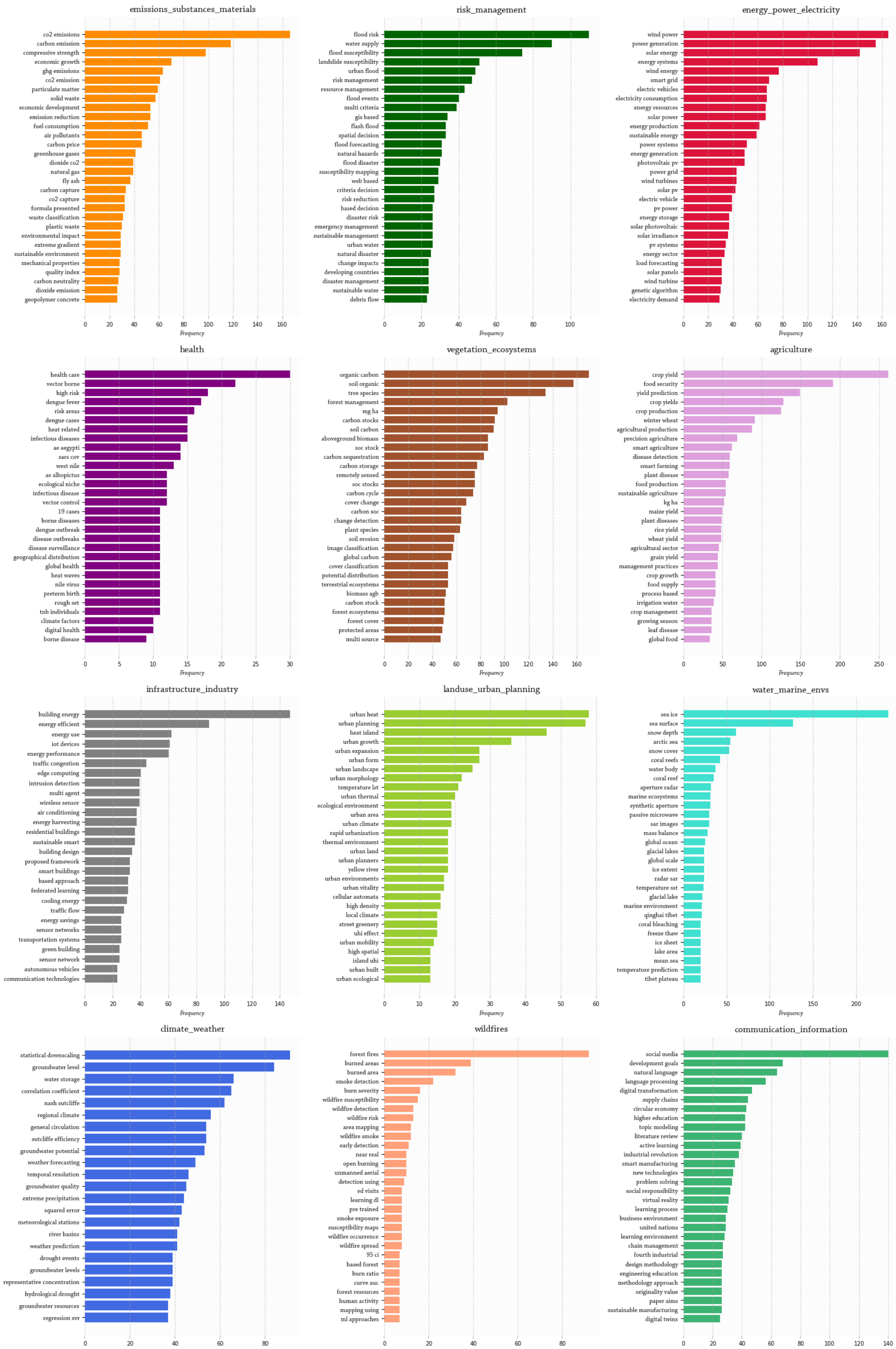

Figure 2: Most frequent Bigrams per cluster.

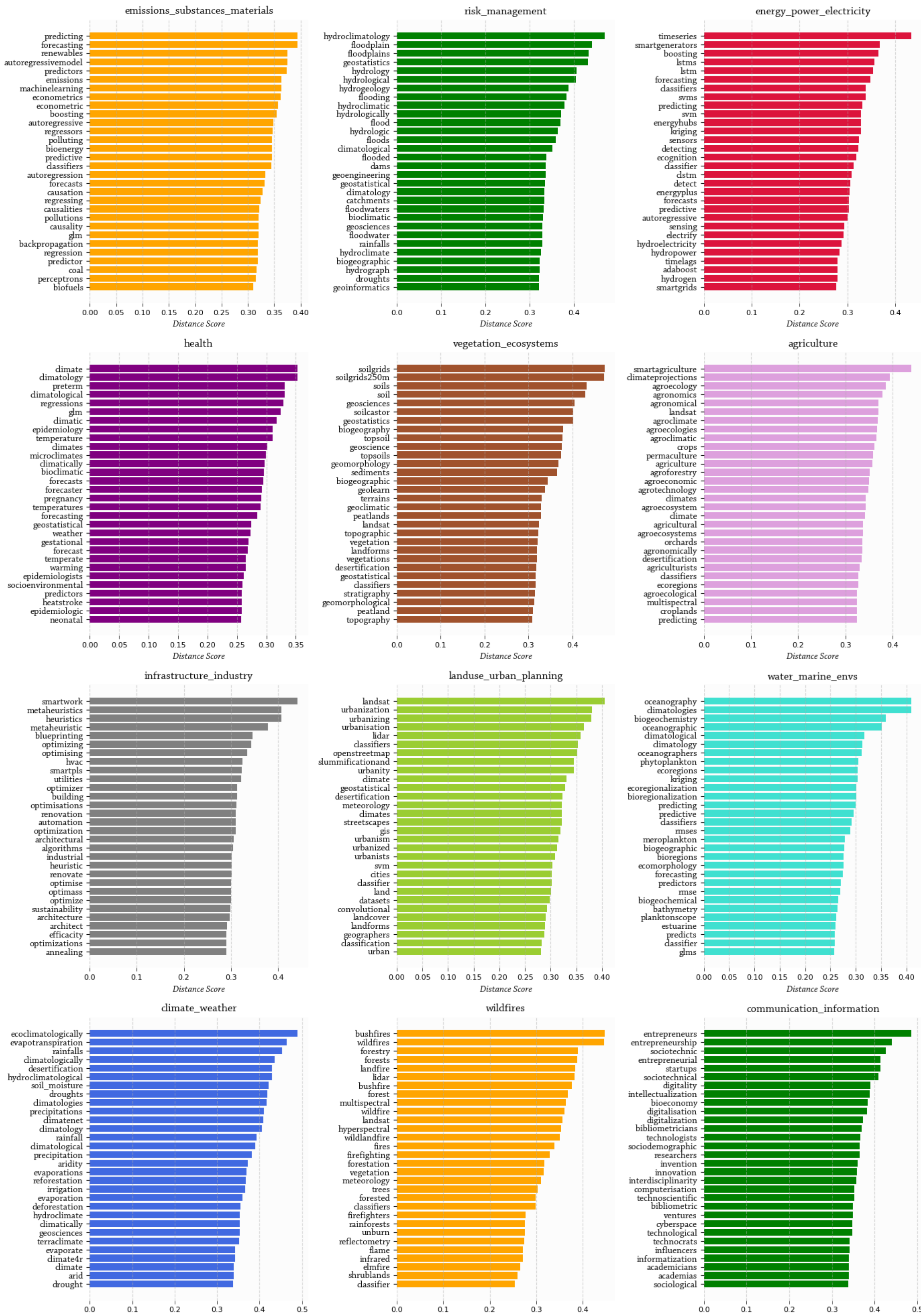

Figure 3: Most relevant keywords per cluster, extracted using *KeyBERT*.

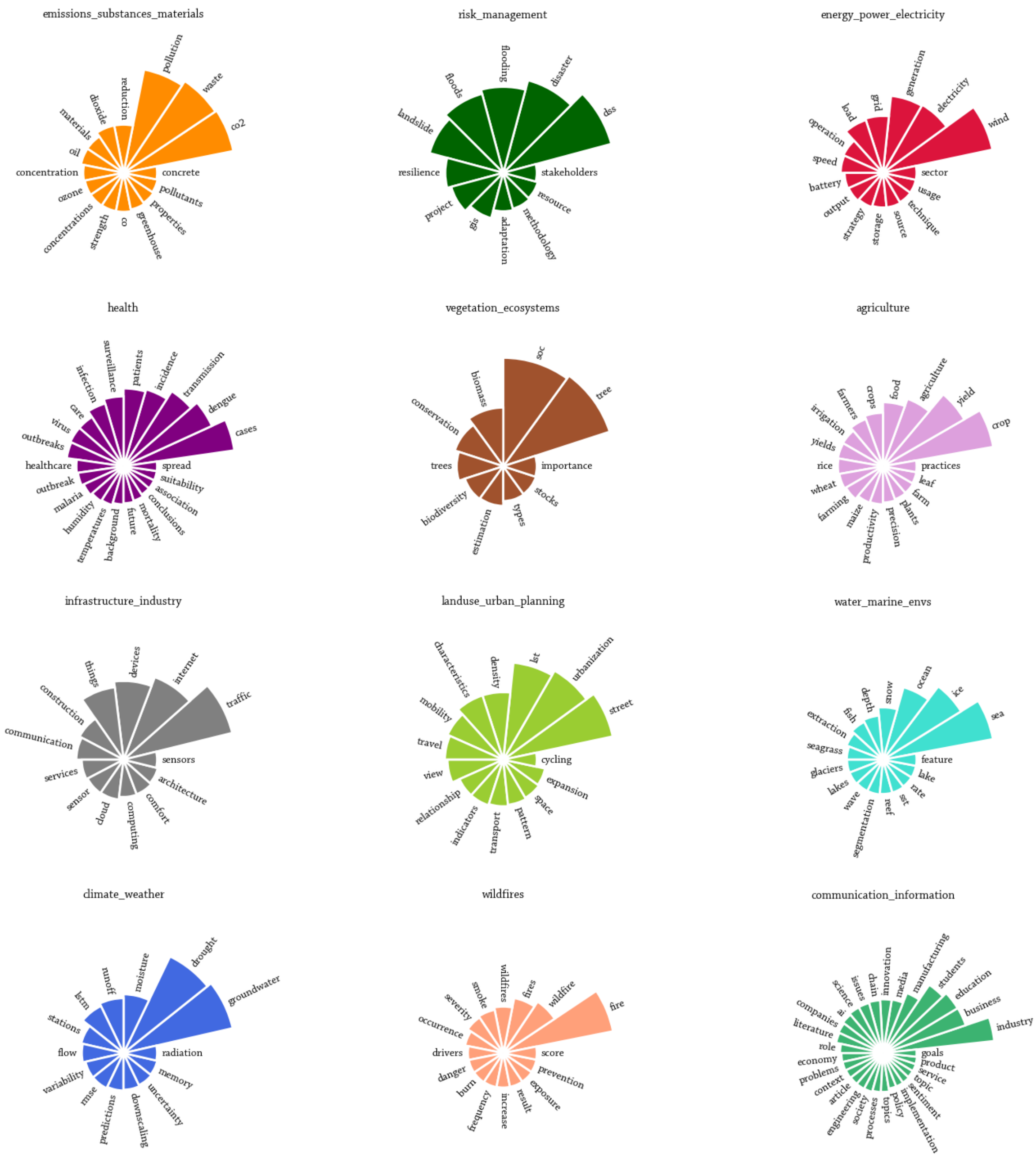

Figure 4: Unique nouns among the most frequent 100 nouns for each cluster.

References

The Global Partnership on Artificial Intelligence. A responsible AI strategy for the environment, 2022. URL https://gpai.ai/projects/responsible-ai/environment/.

BCG. How AI Can Be a Powerful Tool in the Fight Against Climate Change, 2022. URL https://cdn-assets.inwink.com/de3a211c-d522-4534-8cb8-65999ca21469/c62eeb95-4510-4dbc-9256-41cecffec46d?sv=2018-03-28&sr=b&sig=%2F3XoAqqSuHOFciTEys66oQCOz%2FkobYBDOMdYd3CwjF8%3D&se=9999-12-31T23%3A59%3A59Z&sp=r&rscd=inline%3B%20filename%3D%22AI%2520for%2520the%2520Planet%2520x%2520BCG%2520report%2520-%2520July%25202022%2520-%2520vFinal2.pdf%22.

David Rolnick, Priya L. Donti, Lynn H. Kaack, Kelly Kochanski, Alexandre Lacoste, Kris Sankaran, Andrew Slavin Ross, Nikola Milojevic-Dupont, Natasha Jaques, Anna Waldman-Brown, Alexandra Sasha Luccioni, Tegan Maharaj, Evan D. Sherwin, Surya Karthik Mukkavilli, Konrad Paul Kording, Carla P. Gomes, Andrew Y. Ng, Demis Hassabis, John C. Platt, Felix Creutzig, Jennifer T. Chayes, and Yoshua Bengio. Tackling climate change with machine learning. *ACM Computing Surveys (CSUR)*, 55:1 – 96, 2019. URL https://api.semanticscholar.org/CorpusID:189762063.

Rohit Nishant, Mike Kennedy, and Jacqueline Corbett. Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *Int. J. Inf. Manag.*, 53:102104, 2020. URL https://api.semanticscholar.org/CorpusID:219087919.

Maarten R. Grootendorst. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *ArXiv*, abs/2203.05794, 2022. URL https://api.semanticscholar.org/CorpusID:247411231.

OECD. A Responsible AI Strategy for Climate Action, 2021. URL https://oecd.ai/en/wonk/a-responsible-ai-strategy-for-climate-action.

Horizon Magazine. AI Can Help Us Fight Climate Change, but It Has an Energy Problem Too, 2019. URL https://ec.europa.eu/research-and-innovation/en/horizon-magazine/ai-can-help-us-fight-climate-change-it-has-energy-problem-too.

Lynn H. Kaack, Priya L. Donti, Emma Strubell, George Kamiya, Felix Creutzig, and David Rolnick. Aligning artificial intelligence with climate change mitigation. *Nature Climate Change*, 12:518 – 527, 2022. URL https://api.semanticscholar.org/CorpusID:249524697.