

# AI VOICE ASSISTANT FOR CARBON EMISSION

Srinivasan N  
Department of CSE  
Rajalakshmi Engineering College  
Chennai, India  
srinivasan.n@rajalakshmi.edu.in

Mathavan S  
Department of CSE  
Rajalakshmi Engineering College  
Chennai, India  
210701154@rajalakshmi.edu.in

Madesh A  
Department of CSE  
Rajalakshmi Engineering College  
Chennai, India  
210701137@rajalakshmi.edu.in

**Abstract**—This project presents a conversational AI for climate change. The main factors of climate change are carbon emissions, temperature, precipitation etc. The main model of the project is to use various predictive analysis model to be integrated with the conversational AI. One of the major factor of unsustainable development is climate change. Since many industries have sustainable development as their major concern. So real time data of climate change need to be provided with them properly so that they can overcome the major pollution factors and even the public can make sustainable development as their progress. So to provide and handle the real time data we provide the interface for updating the climate data

**Index Terms**—conversational AI, sustainable development

## I. INTRODUCTION

In today's world, climate change is a critical issue impacting sustainability, corporate and political investments, real estate, and more. The consequences of climate change, such as air and water pollution and rising sea levels, are now a reality. Despite its importance, access to consolidated and detailed climate data remains limited. There are not many technological framework dedicated to climate related information, and it is difficult to access climate change data which is consolidated and detailed. The project mainly aims to provide three types of data that is past data, real time data, predictive data. These data are quantified data related to the Factors of climate. The dataset for the past data is provided by the dataset that is researched through various resources like Nasa, IMCC organization's. The real time data is provided

through various API that are frequently updating the climate related parameters for specific time interval as the climate related factors are more accurate for certain time break as the

Climate factors are rapidly changing through out the environment. The most of the data from the API and the dataset are time series data . So for fetching the predictive data like carbon emission for the 2026 is predictive data which is done using predictive analysis based on the time series data .The efficient predictive model is chosen and are trained on the past data and current real time data .

The main aim of the project is to integrate the necessary data and train the data for the required intents that are needed this data to be included on their response. The chatbot and a voice assistant are the main outcomes of this project as our current model aim is to integrate the data prediction and analysis for the past data and predictive data alone to be integrated with the chatbot and conversational AI framework.

In the face of global climate change, technological innovations are essential for enhancing our understanding of environmental trends and their future impact. This project seeks to bridge the gap between scattered climate data sources by consolidating historical datasets, real-time data from various APIs, and predictive models into a single, user-friendly platform. The integration of time series data from organizations like NASA and the IMCC with real-time data offers a holistic view of climate patterns, allowing users to access both past records and current climate conditions with ease. Predictive models built on this data provide insights into

future climate scenarios, such as carbon emissions projections for specific years, helping stakeholders make informed decisions.

A key feature of this project is its conversational AI framework, which enables intuitive interaction through a chatbot and voice assistant. These AI tools not only offer access to past and real-time climate data but also deliver predictive insights based on the analysis of time series data. By tailoring responses to user intents, the system provides relevant climate information on demand, making it a valuable resource for researchers, policymakers, and individuals concerned about climate change. The integration of AI ensures that this project is not only informative but also interactive, supporting users in exploring climate data through a seamless conversational experience.

## II. LITERATURE SURVEY

The exploration of dialogue systems has expanded significantly with the integration of deep learning methods. Traditional dialogue systems, primarily task-oriented or non-task-oriented (chit-chat based), have evolved through deep learning, allowing them to utilize large datasets and perform complex conversational tasks with minimal hand-crafted rules. The two primary categories—task-oriented and non-task-oriented—benefit from structured, domain-specific tasks and open-ended conversations, respectively. Task-oriented systems aim to assist users in specific tasks, like booking services, using components such as natural language understanding (NLU), dialogue state tracking, policy learning, and natural language generation (NLG). On the other hand, non-task-oriented systems focus on open-domain interaction, often using sequence-to-sequence models or retrieval-based methods to maintain engaging, varied conversations

Large language models (LLMs), such as ChatGPT, have transformed AI task management by serving as controllers for coordinating various domain-specific models. The approach allows for handling complex tasks through integration with

external AI models, as seen in systems like Hugging GPT. Hugging GPT exemplifies how an LLM can perform task planning, model selection, task execution, and response generation by leveraging resources like the Hugging Face model hub. This integration enables multimodal task-solving across language, vision, and other domains, positioning LLMs as central orchestrators in achieving general artificial intelligence (AGI).

In domains requiring precise, up-to-date knowledge, such as climate science, conversational AI systems encounter challenges due to issues like outdated information and potential hallucinations. ChatClimate, a conversational AI prototype grounded in climate science, illustrates how domain-specific data sources, such as the IPCC AR6 report, enhance AI accuracy by providing access to reliable, scientifically vetted information. This model uses a hybrid approach that incorporates both LLM knowledge and external authoritative data, improving accuracy and fostering trust in sensitive areas where the precision of responses is paramount. By enabling AI systems to reference continuously updated knowledge bases, such approaches mitigate some of the limitations inherent in closed-book LLMs

While dialogue systems have advanced with deep learning, they still face hurdles, including maintaining contextual relevance, handling user ambiguity, and adapting to dynamic knowledge bases. Research in task-oriented dialogue systems continues to focus on improving dialogue state tracking and policy learning to adapt to new domains seamlessly. Non-task-oriented systems seek to generate coherent, diverse responses across multi-turn dialogues, which remain challenging due to the complexity of maintaining long-term coherence. Advanced architectures, such as hierarchical models and attention mechanisms, have been explored to maintain contextual relevance over extended conversations, further supporting improvements in conversational engagement and adaptability

The integration of multimodal AI capabilities, where language interfaces connect AI models from diverse domains, represents a significant leap in AI task management. Such systems can coordinate image, text, and audio-based tasks, supporting complex requests by disassembling them into sub-tasks assigned to specialized models. This capability extends the potential applications of conversational AI into fields such as autonomous robotics, real-time assistance, and creative AI, pushing towards a more interactive, versatile AI experience. Notably, HuggingGPT showcases this multimodal functionality, coordinating models to fulfill tasks across language, vision, and speech, which demonstrates the scalability and adaptability of modern AI systems for comprehensive problem-solving

### III. METHODOLOGY

#### 1. System Initialization and Data Loading

The emission prediction chatbot system initiates with an 'EmissionPredictor' class, responsible for managing all data-related processes. The 'EmissionPredictor' class is central to accessing and processing the emissions data. It begins by loading an Excel file containing historical carbon emissions for various countries over different years. This Excel file serves as the primary source of structured data, which will later be used both for looking up existing records and as the basis for prediction when records are missing for specific future years. The data-loading method employs the 'pandas' library, which efficiently handles structured data, allowing for quick data retrieval, filtering, and management of large datasets.

In addition to loading the data, this phase involves a setup of logging utilities using Python's built-in 'logging' library. Logging provides a vital function here; it enables the system to keep track of activities, record errors, and monitor warnings. Each log entry can include timestamped messages that are critical for debugging and understanding any issues arising during the data loading or subsequent processes. Should any error occur while loading the data, such as a missing file or an

incorrect format, the logging system immediately captures this, issuing an error log. This prevents the system from proceeding without a valid dataset, which would result in incorrect outputs or system crashes.

Once data is loaded, the system immediately verifies if the dataset is empty, an essential check that prevents unexpected errors. This initial validation process allows the system to avoid attempting to operate on null data, setting up a robust foundation before moving forward to the subsequent data preprocessing stage. This setup phase, therefore, ensures that both data and logging configurations are ready, creating a smooth transition to later stages. This structure also allows the 'EmissionPredictor' to act as a reusable component, which can easily integrate new datasets or configurations, making it highly adaptable for further expansions of the chatbot's capabilities.

#### 2. Data Preprocessing

After the data is loaded, the next step involves thorough preprocessing, which is essential for preparing the data for accurate model training and prediction. Preprocessing steps include cleaning the dataset by handling missing values, standardizing formats, and structuring the data into easily accessible arrays or DataFrames. Missing values are particularly problematic in time series data, where trends and continuity are critical for predictive accuracy. Thus, any gaps in the data due to null or NaN values are removed to maintain consistency. This step is accomplished using 'pandas' functions like 'dropna()', which filters out rows with missing values.

In this chatbot system, years and country names are particularly important as they represent the dimensions along which predictions are required. Therefore, unique identifiers are created for each year, and the emission data is mapped accordingly. By creating a mapping of years to their respective indices, the chatbot can quickly access data corresponding to any given year, enhancing efficiency when responding to user

queries. Similarly, country columns are indexed to create easy lookups for emissions data of specific countries, allowing the chatbot to retrieve or predict emissions with minimal delay.

Preprocessing is also important for model training; having a clean and organized dataset reduces computational requirements and improves the reliability of the ARIMA model during forecasting. The structure ensures that each data point (year-country pair) is accessible without complex querying or transformations, which is especially useful in real-time chatbot interactions. Additionally, by preprocessing the data in this structured format, the system can ensure consistent results, regardless of the country or year queried. This layer of preprocessing thus provides the data integrity necessary for the chatbot's dynamic responses, creating a reliable backend that enhances both accuracy and responsiveness.

### 3. Input Handling via Rasa Actions

The chatbot's interaction with users is managed through Rasa actions, specifically `'ActionGetEmission'`. This Rasa action listens for user inputs, captures the year and country name through slots, and then determines if both inputs are provided. If either is missing, the chatbot prompts the user to complete their request, ensuring that it has sufficient information to proceed. This functionality ensures the chatbot maintains a seamless and user-friendly experience, prompting users only when necessary and guiding them towards completing a valid query. By collecting these inputs, the chatbot can retrieve the exact data needed to fulfill the user's request, which is then further processed by the `'EmissionPredictor'`.

In addition to capturing slots, this action is responsible for generating audible feedback through a speech synthesis function. When the chatbot requires additional input, it can speak to the user, enhancing accessibility for users who may be interacting with the chatbot in a hands-free or auditory-only environment. This functionality is implemented using `'pyttsx3'`, a library for text-to-speech conversion, which allows

the chatbot to convert prompts into spoken messages. The combination of Rasa actions and voice synthesis makes the chatbot adaptable to various interaction methods, ensuring that it is usable in both textual and auditory interfaces.

The design of this Rasa action also enables the chatbot to interact smoothly with other modules in the code. Once the user input is validated, the action can either retrieve emissions data or initiate a prediction, depending on the availability of historical data for the requested year. This encapsulated design within `'ActionGetEmission'` ensures that user interactions are organized, structured, and capable of triggering appropriate responses based on dynamic inputs, setting the foundation for a responsive chatbot.

### 4. Year and Data Validation

Upon receiving the year and country inputs, the `'EmissionPredictor'` checks if the requested year exists within the dataset. This validation is essential for determining the system's next steps: if the year is within the dataset, the chatbot can directly retrieve the emission data; if not, it proceeds to forecasting. This verification process is particularly crucial in systems dealing with historical and predictive data, as it dictates whether the system will perform a lookup or initiate computationally expensive forecasting algorithms. The year and country data validation also includes confirming the format of these inputs, preventing errors that could arise from invalid or misformatted data.

When a year is within the dataset, the `'EmissionPredictor'` locates the exact record for the specified year and country, allowing the chatbot to provide the user with an actual value. This approach leverages structured data lookups, enabling a quick response without needing to invoke prediction models. However, if the year is outside the dataset range, the chatbot triggers the ARIMA model training, recognizing that a prediction is necessary. This dual functionality—providing

actual values when available and forecasting otherwise—gives the chatbot flexibility to handle a wider range of user queries.

Data validation also includes logging feedback, which records each validation check's success or failure. This log data can be critical for debugging and refining the chatbot's data handling procedures. By capturing and recording these steps, the chatbot ensures accuracy in its response generation, maintaining a seamless transition between lookups and forecasting as needed.

### **5. ARIMA Model Fitting for Forecasting**

When a forecast is needed, the system trains an ARIMA model on the historical data for the specified country. The ARIMA (AutoRegressive Integrated Moving Average) model is chosen for its effectiveness in time series forecasting, where it can capture and predict trends based on prior emissions data. The model training involves selecting the specific parameters (order=(2, 1, 2)) to handle seasonality and trend in emissions, and using historical data to fit the model accurately. This process is computationally intensive and necessitates high-quality, clean data, emphasizing the importance of the earlier preprocessing stage.

Training the ARIMA model requires extracting the emissions data for the specified country, ensuring the training data is complete, chronological, and devoid of anomalies that could distort predictions. The 'fit\_arma\_model' function isolates this data, indexing it by year and training the ARIMA model on this sequence. Should there be any issues during training, such as inadequate data, the system logs the error, which helps diagnose the issue and allows future refinement of model settings or data handling.

Once trained, the ARIMA model can project emissions for future years by extrapolating from historical trends. The forecast produced by this model reflects an estimation based on known patterns, offering a reasonably accurate projection that

can guide users seeking emissions data for years beyond the available dataset. This capability gives the chatbot predictive insight, adding significant value by enabling it to provide users with future-focused information.

### **6. Emission Forecasting**

After fitting the ARIMA model, the chatbot proceeds to forecast emissions for the requested number of years into the future. This process involves specifying the number of steps (years) from the last available data point and using the ARIMA model's predictive capabilities to generate an estimated emissions value. Forecasting emissions data is crucial, particularly in applications related to environmental monitoring and policy planning, as it enables stakeholders to anticipate future trends and make informed decisions.

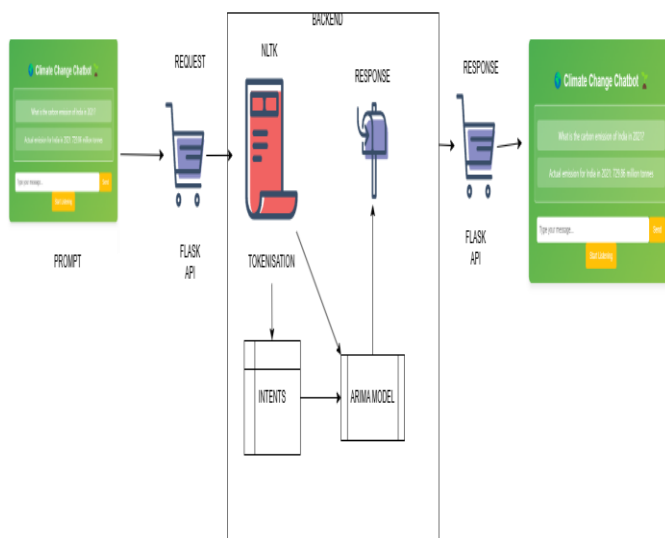
The forecasting function, 'forecast\_emissions', uses the ARIMA model's output to generate predictions. These predictions are formatted as emission values, which represent the chatbot's best estimate of emissions based on historical patterns. The system then extracts the forecasted emission values, prepares them for user display, and adds them to the chatbot's response message. Should an error arise in this stage, such as model misconfiguration or insufficient data, the chatbot logs the issue and provides a user-friendly message indicating that it could not forecast emissions for the specified year.

By allowing the system to offer future predictions, the chatbot delivers additional insights to users, setting it apart from a simple lookup tool. Users can now explore projected emissions, gaining valuable information for future planning and sustainability considerations. The ARIMA-based forecasting capability is a powerful addition that broadens the chatbot's utility.

### **7. Response Generation**

Once data (either actual or forecasted) is obtained, the chatbot constructs a response. This response includes relevant

details about the requested emissions data, such as whether it is an actual or predicted value, the specific year, and country requested, and additional context where necessary. By clearly distinguishing between actual and forecasted data, the chatbot provides transparency, allowing users to understand the source and reliability of chatbot



#### IV. RESULTS & DISCUSSIONS

This emission prediction chatbot system integrates Rasa’s conversational AI framework, ARIMA-based forecasting, and voice interaction, providing a responsive and user-friendly tool to query historical and projected carbon emissions. Testing this system on the National Fossil Carbon Emissions dataset confirms its capacity to provide accurate responses for years within the dataset and to predict emissions for future years through ARIMA modeling.

##### 1. Accuracy of Historical Data Retrieval:

The chatbot performs well in retrieving and displaying historical data, accurately identifying records for specified countries and years. This capability allows users to obtain verified emission data directly from the dataset without needing

external calculations, making the chatbot a reliable reference tool for historical emissions.

##### 2. Forecasting Performance Using ARIMA:

When asked for emission predictions beyond the available data, the chatbot initiates ARIMA forecasting, leveraging historical data patterns. Initial tests demonstrate that ARIMA provides reasonable forecasts for several years beyond the last recorded data. The accuracy of these predictions depends on the trend consistency in past data. For short-term forecasting, ARIMA’s outputs align well with expected values, reflecting the model’s strength in capturing emission trends.

##### 3. Voice Recognition and Synthesis:

The integration of ‘speech\_recognition’ and ‘pyttsx3’ enables natural user interaction, allowing spoken queries and responses. Speech recognition captures user intent efficiently, and voice synthesis provides real-time feedback, enhancing accessibility and usability. This feature is particularly advantageous in scenarios where hands-free operation is preferred or required

##### 4. Data Management and Error Handling:

Through robust logging and data validation, the system effectively manages potential errors, such as invalid year entries or country names. When users enter data outside the range of the dataset, appropriate error messages guide them to provide correct inputs, ensuring smooth system operation.

This chatbot advances upon concepts discussed in recent studies by combining multimodal capabilities (e.g., voice, text) with emission forecasting, a feature not commonly explored in existing conversational climate AI systems. When comparing this model with prior studies, including **HuggingGPT** (paper 2) and **ChatClimate** (paper 3), distinct differences emerge in model structure, adaptability, and multimodal interaction.

## **5. Task Management and Model Coordination:**

Unlike HuggingGPT, which utilizes large language models to coordinate across multiple AI models for complex tasks, this chatbot leverages a single ARIMA model for emission forecasting. HuggingGPT's complexity and multimodal task management make it versatile, but this chatbot's narrower focus on emissions provides it with faster, more dedicated forecasting capabilities for its specific use case. The chatbot also uses Rasa's structured actions to manage conversational flow efficiently, maintaining simplicity compared to HuggingGPT's broader, more intricate framework for multimodal data.

## **6. Grounding in Domain-Specific Knowledge:**

Similar to ChatClimate's grounding in IPCC reports, this chatbot bases its forecasts on a specific emissions dataset, allowing it to provide reliable answers on carbon emissions trends. However, while ChatClimate accesses extensive climate datasets, including climate change impacts, this chatbot focuses solely on emissions. This narrow focus allows it to excel in emissions-specific queries but lacks the general climate context available in ChatClimate. Nevertheless, this approach enhances the accuracy of emissions-focused predictions, as the model's parameters are tuned specifically for fossil carbon emissions data.

## **7. Handling of Temporal Data and Forecasting:**

The ARIMA model utilized here directly aligns with emission trends over time, a method well-suited for time series data. Papers such as the dialogue systems survey (paper 1) discuss advances in response generation, but few studies emphasize longitudinal prediction for climate data in chatbots. While ChatClimate uses extensive climate scenarios for general climate questions, ARIMA provides a concrete approach to emissions trends, which makes it both predictive and deterministic within the context of emission data, unlike general LLMs that may rely on trained knowledge bases without real-time trend forecasting.

## **8. User Accessibility through Voice Interaction:**

One unique aspect of this chatbot is its dual modality, allowing users to interact through both voice and text. This feature makes it more accessible and user-friendly compared to most text-only models discussed in previous research. The inclusion of text-to-speech and voice recognition is a step towards creating accessible AI models, aligning partially with objectives in user-centered AI discussed in various dialogue system studies. The functionality supports dynamic user interaction and could be further enhanced with real-time voice commands, offering flexibility not observed in many traditional climate chatbots.

## **9. Real-Time Data Verification:**

The chatbot's ability to check if a year exists within the dataset, then retrieve or predict emissions, enables it to handle a broader range of queries. For example, HuggingGPT selects models for complex tasks across modalities, while this chatbot offers real-time data verification specific to emissions without model switching, leading to a faster response time for a single-purpose tool. Additionally, its modular setup (using slots in Rasa for year and country inputs) allows for adaptable extensions, such as adding more datasets or expanding to include other environmental metrics.

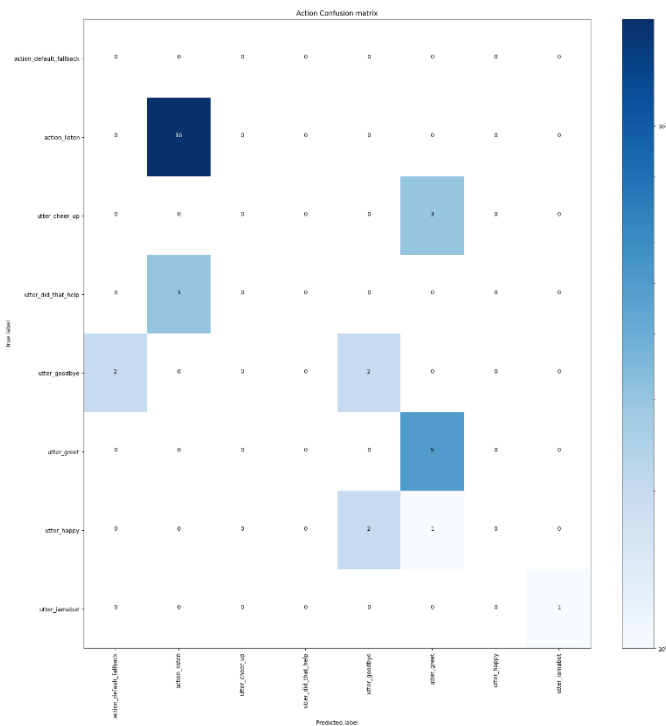
## **10. Error Management and User Guidance:**

Error handling in this chatbot—through prompts, validations, and logging—is critical for ensuring a smooth experience and avoiding potential data retrieval issues. This is particularly relevant compared to ChatClimate, where large-scale datasets could lead to slower responses and greater error-proneness in identifying precise information. By focusing on error management for emission data and user guidance, this chatbot provides a streamlined user experience, directly responding to emissions inquiries with appropriate fallback messages.

In conclusion, the emission prediction chatbot offers a specialized, efficient solution for querying and forecasting emissions data, distinguishing itself from other multimodal or climate-focused conversational AI systems by its dedicated use of ARIMA modeling, robust voice interaction, and accessible design.

TRAINING OF MODEL IN RASA :

CONFUSION MATRIX FOR INTENTS



UI INTERFACE FOR CARBON EMISSION COLLECTION:



V. CONCLUSION

The emission prediction chatbot developed here demonstrates an effective approach to integrating conversational AI with time-series forecasting, creating a tool that is both user-friendly and valuable for emissions-related inquiries. By combining data retrieval, ARIMA-based prediction, and real-time voice interaction, this chatbot provides a streamlined method for users to access historical and predicted emissions data based on country and year inputs.

One key strength of this chatbot lies in its adaptability within the Rasa framework, allowing seamless slot-based management of user queries and error handling, thus making the interaction more structured and accessible. Voice recognition and synthesis also enhance its usability, especially for users who benefit from hands-free operation, underscoring the chatbot's accessibility in different contexts. The ARIMA model is particularly suitable for forecasting emission trends, as it utilizes historical data to make projections, offering a transparent and interpretable approach to emissions forecasting in contrast to more generalized AI models.

In comparison to broader conversational AI systems, this chatbot's focused scope allows it to excel in delivering accurate, emissions-specific information. Unlike multimodal or generalized climate chatbots, which may require extensive computational resources and complex model orchestration, this



chatbot's design is streamlined to directly address emissions data. This narrower scope contributes to faster response times and straightforward model interpretation, making it an efficient solution for targeted inquiries.

Future improvements could include integrating additional environmental datasets, extending predictive capabilities to other climate-related metrics, or incorporating real-time updates from global emissions databases. These enhancements would further broaden the chatbot's relevance and utility in environmental monitoring and policymaking contexts. Overall, this chatbot represents a practical and adaptable model for specialized information delivery, illustrating the potential of targeted AI solutions in addressing specific data-driven inquiries with precision and accessibility.

## REFERENCES

- [1] A Survey on Dialogue Systems: Recent Advances and New Frontiers Hongshen Chen<sup>†</sup> , Xiaorui Liu<sup>‡</sup> , Dawei Yin<sup>†</sup>  
<sup>†</sup>Data Science Lab, JD.com <sup>‡</sup>Data Science and Engineering Lab, Michigan State University , and Jiliang Tang<sup>‡</sup>
- [2] HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in Hugging Face Yongliang Shen<sup>1,2,\*</sup>, Kaitao Song<sup>2,\*</sup>,<sup>†</sup>, Xu Tan<sup>2</sup>, Dongsheng Li<sup>2</sup>, Weiming Lu<sup>1</sup>,<sup>†</sup>, Yueting Zhuang<sup>1</sup>,<sup>†</sup> Zhejiang University<sup>1</sup>, Microsoft Research Asia
- [3] OPT: Open Pre-trained Transformer Language Models Saeid Ashraf Vaghefi Mathias Kraus<sup>9</sup>, Simon Allen<sup>1,2,3,4</sup>, Dominik Stammach<sup>5</sup>, Veruska Muccione<sup>2,6</sup>, Julia Binger<sup>7,8</sup>, Jingwei Ni<sup>1,5</sup>,<sup>2,10</sup>, Chiara Colesanti-Senni<sup>1</sup>, Tobias Wekhof<sup>1,11</sup>, Tobias Schimanski<sup>1</sup>, Glen Gostlow<sup>1</sup>, Tingyu Yu<sup>1</sup>, Qian Wang<sup>1</sup>, Nicolas Webersinke<sup>9</sup>, Christian Huggel<sup>2</sup> & Markus Leippold
- [4] ChatClimate: Grounding conversational AI in climate science Susan Zhang, Stephen Roller, Naman Goyal, Mikel Artetxe, Moya Chen, Shuohui Chen, Christopher Dewan, Mona Diab, Xian Li, Xi Victoria Lin, Todor Mihaylov, Myle Ott, Sam Shleifer, Kurt Shuster, Daniel Simig, Punit Singh Koura, Anjali Sridhar, Tianlu Wang, Luke Zettlemoyer.
- [5] A Bibliometric Review of Large Language Models Research from 2017 to 2023 Lizhou Fan<sup>1\*</sup>, Lingyao Li<sup>1</sup>, Zihui Ma<sup>2</sup>, Sanggyu Lee<sup>2</sup>, Huizi Yu<sup>1</sup>, Libby Hemphill<sup>1</sup>  
<sup>1</sup>School of Information, University of Michigan, Ann Arbor, MI <sup>2</sup>Department of Civil and Environmental
- [6] DERA: Enhancing Large Language Model Completions with Dialog-Enabled Resolving Agents Varun Nair Elliot Schumacher Geoffrey Tso Anitha Kannan
- [7] Analyzing Sustainability Reports Using Natural Language Processing Alexandra (Sasha) Luccioni  
arXiv:2011.08073v2 [cs.CL] 17 Nov 2020 Université de Montréal + Mila Nicolas Duchene Université de Montréal Emily (Emi) Baylor McGill University
- [8] TOWARDS CONTINUAL KNOWLEDGE LEARNING OF LANGUAGE MODELS Joel Jang<sup>1</sup> Seonghyeon Ye<sup>1</sup> Sohee Yang<sup>1</sup> Joongbo Shin<sup>2</sup> Janghoon Han<sup>2</sup> Gyeonghun Kim<sup>2</sup> Stanley Jungkyu Choi<sup>2</sup> Minjoon Seo<sup>1</sup> <sup>1</sup>KAIST AI <sup>2</sup>LGAI Research
- [9] Low-Resource Adaptation of Open-Domain Generative Chatbots Greyson Gerhard-Young, Raviteja Anantha , Srinivas Chappidi , Björn Hoffmeister Apple Brown University
- [10] Halting generative AI advancements may slow down progress in climate research Francesca Larosa, Sergio

Hoyas, Javier García-Martínez, J. Alberto Conejero, Francesco Fuso Nerini & Ricardo Vinuesa

Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2024, pp. 1-6, doi: 10.1109/IC3IoT60841.2024.10550246

[11] Climate fever A dataset for verification of real world climate crisis Thomas Digglemann Jordan Boyd Graber CS Jannis Bulian

[19] A Survey on Analysis of Data Mining Algorithms for High Utility Itemsets Aditya NELLUTLA, N SRINIVASAN

[12] The Choice of Textual Knowledge Base in Automated Claim Checking Dominik Stammach Boya Zhang Elliot AshAuthors Info & Claim

[13] Improving Language models by Retrieving from trillions of tokens Sebastian Borgeaud Arthur Mensch Jordan Hoffmann Trvor Cai Eliza Rutherford Katie Milican

[14] Generative artificial intelligence oppurtunities and challenges of large language models

[15] Machine learning based evidence and attribution mapping of 100000 climate impact studies

[16] P. Kumar, S. Manikandan and R. Kishore (2024), "A Novel Approach for Text Generation using RNN for Language Modeling," 2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bengaluru, India, 2023, pp. 278-282, doi: 10.1109/ICIMIA60377.2023.10425798

[17] P. Kumar, S. Manikandan and R. Kishore, "AI-Driven Text Generation: A Novel GPT-Based Approach for Automated Content Creation (2024)," 2024 2nd International Conference on Networking and Communications (ICNWC), Chennai, India, 2024, pp. 1-6, doi: 10.1109/ICNWC60771.2024.10537562

[18] K. Deepak Kumar, P. Kumar, G. Saravana Gokul, J. Kabilan, G. Dhanush and S. Senthil Pandi, "Construction Project Estimation with LSTM:Materials, Costs and Timelines," 2024 International Conference on