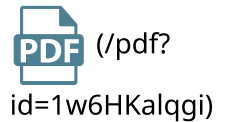


# Contrastive Classification via Linear Layer Extrapolation



Mayukh Sharma (/profile?id=~Mayukh\_Sharma1), Sean O'Brien (/profile?id=~Sean\_O%27Brien1), Julian McAuley (/profile?id=~Julian\_McAuley1)

15 Jun 2024 (modified: 02 Jul 2024) ACL ARR 2024 June Submission Everyone Revisions (/revisions?id=1w6HKalqgi) BibTeX CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>)

## Abstract:

Early-exiting predictions in a deep Transformer network evolve from layer to layer in a somewhat smooth process. This has been exploited in language modeling to improve factuality (Chuang et al., 2023), with the observation that factual associations emerge in later layers. We find a similar process multiway emotion classification, motivating Linear Layer Extrapolation, which finds stable improvements by recasting contrastive inference as linear extrapolation. Experiments across multiple models and emotion classification datasets find that Linear Layer Extrapolation outperforms standard classification on fine-grained emotion analysis tasks.

**Paper Type:** Short

**Research Area:** Sentiment Analysis, Stylistic Analysis, and Argument Mining

**Research Area Keywords:** applications, sentiment analysis, language modeling

**Contribution Types:** Model analysis & interpretability, NLP engineering experiment, Publicly available software and/or pre-trained models

**Languages Studied:** English

**Reviewing Volunteers For Emergency Reviewing:** N/A, no volunteers were provided in the previous question.

**Reviewing No Volunteers Reason:** All qualified authors are already involved in the reviewing process in some capacity (as Area Chairs, as Senior Area Chairs, etc.).

**TLDR:** Contrasting across layers improves classifier performance on fine-grained emotion detection

**Previous URL:** /forum?id=FHxXTO\_tAI (/forum?id=FHxXTO\_tAI)

**Response PDF:** pdf (/attachment?id=1w6HKalqgi&name=response\_PDF)

**Reassignment Request Action Editor:** No, I want the same action editor from our previous submission and understand that a new action editor may be assigned if the previous one is unavailable

**Reassignment Request Reviewers:** No, I want the same set of reviewers from our previous submission and understand that new reviewers may be assigned if any of the previous ones are unavailable

**Preprint:** yes

**Preprint Status:** We are considering releasing a non-anonymous preprint in the next two months (i.e., during the reviewing process).

**Preferred Venue:** EMNLP 2024

**Consent To Share Data:** yes

**Consent To Share Submission Details:** On behalf of all authors, we agree to the terms above to share our submission details.

**Author Submission Checklist:** I confirm that the paper is anonymous and that all links to data/code repositories in the paper are anonymous., I confirm that the paper has proper length ( Short papers: 4 content pages maximum, Long papers: 8 content pages maximum, Ethical considerations and Limitations do not count toward this limit), I confirm that the paper is properly formatted (Templates for \*ACL conferences can be found here: <https://github.com/acl-org/acl-style-files> (<https://github.com/acl-org/acl-style-files>)).

**A1 Limitations Section:** This paper has a limitations section.

**A2 Potential Risks:** No

**A2 Elaboration:**  Our method is benchmarked on widespread academic benchmarks and is a mechanistic modification of an existing task that does not cost significant compute; there are not many foreseeable new risks that this work introduces.

**A3 Abstract And Introduction Summarize Claims:**  Yes

**B Use Or Create Scientific Artifacts:**  No

**B1 Cite Creators Of Artifacts:**  Yes

**B2 Discuss The License For Artifacts:**  N/A

**B3 Artifact Use Consistent With Intended Use:**  No

**B4 Data Contains Personally Identifying Info Or Offensive Content:**  No

**B5 Documentation Of Artifacts:**  N/A

**B6 Statistics For Data:**  Yes

**C Computational Experiments:**  Yes

**C1 Model Size And Budget:**  Yes

**C2 Experimental Setup And Hyperparameters:**  Yes

**C3 Descriptive Statistics:**  Yes

**C4 Parameters For Packages:**  Yes

**D Human Subjects Including Annotators:**  No

**D1 Instructions Given To Participants:**  N/A

**D2 Recruitment And Payment:**  N/A

**D3 Data Consent:**  N/A

**D4 Ethics Review Board Approval:**  N/A

**D5 Characteristics Of Annotators:**  N/A

**E Ai Assistants In Research Or Writing:**  Yes

**E1 Information About Use Of Ai Assistants:**  No

**E1 Elaboration:**  AI assistants were only used for reference in formatting LaTeX.

**Association For Computational Linguistics - Blind Submission License Agreement:**  On behalf of all authors, I agree

**Submission Number:** 3832

## Discussion (/forum?id=1w6HKalggi#discussion)

Filter by reply type  Filter by author  Search keywords...

Sort: Newest First

Everyone  Submission3832...  Submission3832 Area...  Submission3832 Authors 8 / 8 replies shown

Submission3832...  Program Chairs  Submission3832...  Submission3832...

Submission3832...  Submission3832...

Add:

**Author-Editor Confidential Comment**

**Withdrawal**



## Meta Review of Submission3832 by Area Chair fcnU

Meta Review Area Chair fcnU 06 Aug 2024, 15:34 (modified: 22 Aug 2024, 15:45)

Senior Area Chairs, Area Chairs, Authors, Reviewers Submitted, Program Chairs, Commitment Readers

Revisions (/revisions?id=x5swS00bIL)

### Metareview:

This paper uses layer-wise contrastive decoding methods in the scope of fine-grained emotion and sentiment analysis tasks to address underrepresented classes at inference time and improve overall classifier (FLAN-T5, DeBERTa) performance. The key idea is to contrast the predictions from the final layer (expert) against an earlier layer (amateur) of the same model. The

results beat standard training approaches on four representative datasets (EmpatheticDialogue, tweetHate, tweetEmotion, goEmotions).

#### Summary Of Reasons To Publish:

- The finding regarding the early layer's influence on prediction certainty is interesting and the extrapolation approach is principled and fresh in the context of emotion and sentiment analysis (**7m4T**, **puvi**, **35QN**);
- The study is conducted across various models and emotion classification datasets, ensuring the robustness of the proposed method and showing clear empirical gains (**puvi**, **35QN**, **7m4T**);
- The experimental section is targeted and elaborates on the improvements made for ambiguous samples (**7m4T**);
- The paper is well-embedded into related work and offers a range of contributions that fits the requirements of a short paper.

#### Summary Of Suggested Revisions:

- The main paper is too concerned with higher-level results (recall,  $F_1$ ), while ignoring a discussion on precision trade-offs (**35QN**) and hiding more interesting, fine-grained observations about the improvements in the Appendix (e.g., App. B);
- The Abstract does not clearly state the objective and motivation of the paper, is hard to read in general (e.g., "We find a similar process multiway emotion classification"), while the Title is leaving out the domain, which might raise false expectations for interested readers. Since the evaluation is targeted specifically at emotion classification and the novelty or main draw is neither the "Linear Layer Extrapolation" or the "Contrastive Classification", it might be worthwhile to add "Emotion Analysis" in the Title;
- The paper lacks clarity in various aspects and sections. Reviewers pointed out that a visualization of cases about how to change the wrong emotion tag with extrapolation (**7m4T**), making the used symbols less chaotic (**puvi**), mentioning what the final optimization object is (**puvi**) and what the use of the classification head for early-exiting is (**puvi**) would help a lot;
- Theoretical justification for why the linear extrapolation approach should work needs to be improved (**35QN**). It is poorly integrated and under-explained in the context of the methodology section (§3). Rather than overwhelming the reader with a mathematic formalization, I would propose to add one or two sentences at the start of the paragraph explaining the purpose, expected outcome and the approach on a higher level and how it connects to the rest of the methodology;
- Sensitivity analysis of hyperparameters like the extrapolation layer  $l_t$  could be added (**35QN**);
- Source code should be publicly available (**puvi**).

While the consensus is only slightly positive, all three reviewers agree that the change from a long (ARR 2024 February version ([https://openreview.net/forum?id=FHxXTO\\_tAI](https://openreview.net/forum?id=FHxXTO_tAI))) to a short paper made sense and most issues were addressed by the authors. While the list of remaining issues is not short, it should be possible to mitigate them in the commitment / camera-ready submission.

**Overall Assessment:** 4 = There are minor points that may be revised

**Suggested Venues:** EMNLP

**Best Paper Ae:** No

**Ethical Concerns:**



There are no concerns with this submission

**Needs Ethics Review:** No

**Author Identity Guess:** 1 = I do not have even an educated guess about author identity.

Add: **Author-Editor Confidential Comment**

## Official Review of Submission3832 by Reviewer 35QN

Official Review  Reviewer 35QN  19 Jul 2024, 17:11 (modified: 22 Aug 2024, 15:45)

 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors, Reviewer 35QN, Commitment Readers

 Revisions (/revisions?id=4jpdIvBOO)

**Paper Summary:**

This paper proposes a Linear Layer Extrapolation method for improving fine-grained emotion classification using large language models. The key idea is to perform contrastive classification by contrasting the predictions from the final layer (expert) against an earlier layer (amateur) of the same model. The contrastive strength is dynamically determined using a linear extrapolation technique, stabilizing performance across different models and datasets. Experiments on multiple emotion classification datasets show that this contrastive approach, especially with dynamic contrastive strength selection, improves recall and F1 scores compared to standard classification, particularly for underrepresented emotion classes.

Disclaimer: I reviewed this paper in a previous cycle.

#### Summary Of Strengths:

- Novel application of contrastive decoding methods to the fine-grained emotion classification task, an important problem in NLP.
- The proposed linear layer extrapolation provides a principled way to set the contrastive strength, improving dynamic performance stability.
- Thorough experimental evaluation across multiple models (FlanT5, DeBERTa) and datasets, with clear empirical gains.

#### Summary Of Weaknesses:

1. The improvements are modest and mainly focused on recall/F1; there is little discussion of precision tradeoffs.
2. Sensitivity analysis of hyperparameters like the extrapolation layer  $\lambda$  is limited.
3. Theoretical justification for why the linear extrapolation approach should work needs to be improved.
4. Evaluation is limited to text classification tasks; applicability to other NLP tasks needs to be clarified.

The revised version has somewhat clarified point 2, so I am increasing the soundness of the paper.

#### Comments Suggestions And Typos:

NA

**Confidence:** 4 = Quite sure. I tried to check the important points carefully. It's unlikely, though conceivable, that I missed something that should affect my ratings.

**Soundness:** 3 = Acceptable: This study provides sufficient support for its major claims/arguments. Some minor points may need extra support or details.

**Overall Assessment:** 2.5

**Best Paper:** No

**Needs Ethics Review:** No

**Reproducibility:** 3 = They could reproduce the results with some difficulty. The settings of parameters are underspecified or subjectively determined, and/or the training/evaluation data are not widely available.

**Datasets:** 1 = No usable datasets submitted.

**Software:** 1 = No usable software released.

**Knowledge Of Or Educated Guess At Author Identity:** No

**Knowledge Of Paper:** N/A, I do not know anything about the paper from outside sources

**Knowledge Of Paper Source:** N/A, I do not know anything about the paper from outside sources

**Impact Of Knowledge Of Paper:** N/A, I do not know anything about the paper from outside sources

Add: **Author-Editor Confidential Comment**



### Official Comment by Authors

Official Comment

Authors ( Sean O'Brien (/profile?id=~Sean\_O'Brien1), Julian McAuley (/profile?id=~Julian\_McAuley1), Mayukh Sharma (/profile?id=~Mayukh\_Sharma1))

30 Jul 2024, 01:11 (modified: 22 Aug 2024, 15:45)

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors, Reviewer 35QN, Commitment Readers

Revisions (/revisions?id=cgVKZmCMKW)

#### Comment:

Thank you for your thorough review and feedback! We have tried to address the points you raised and welcome further discussion.

Response to weaknesses:

**The improvements are modest and mainly focused on recall/F1; there is little discussion of precision tradeoffs.**

In this work our main focus was to extend the idea of contrastive methods as purely runtime inference optimization techniques for boosting the performance of the model. Our proposed method does not require any additional training/fine-tuning. The results show a general increase in recall and F1 scores without much harm to the precision. The idea can be extended to further study the use of intermediate layer features and how to boost the model performance without using any external models/data. The major advantage of inference time boosting methods as proposed in our work is that they do not require any additional resources. The additional hyperparameters that we proposed can also be easily swapped as per the task requirements without any training/extra data.

**Theoretical justification for why the linear extrapolation approach should work needs to be improved.**

We establish the interpretation of DoLa as linear extrapolation, as opposed to the original theoretical justification for contrastive decoding that casts the problem as maximizing mutual distinguishability between an expert and amateur model. For approaches that utilize dynamic layer selection, the mutual distinguishability formulation treats all amateur layers as equivalent, while the linear extrapolation approach reduces the contrastive penalty for amateur layers closer to the final layer. We demonstrate empirically that this interpretation leads to more stable performance than a fixed beta.

**Evaluation is limited to text classification tasks; applicability to other NLP tasks needs to be clarified.**

Contrastive methods in the inference space have been mostly explored with respect to text generation in LLMs where they have shown to improve factuality and reduce hallucinations. We built on that existing work using DoLA and contrastive decoding and extended it to fine-grained classification tasks. We observed similar behavior in the probability distribution of pre-final and final layers as used in other contrastive methods for fine-grained classification tasks. Additionally, we introduced dynamic contrastive strength as linear layer extrapolation. In future works, we aim to extend this idea to broader NLP tasks. In this work as a short manuscript we wanted to focus on a targeted task introducing contrastive action as linear layer extrapolation.

**The revised version has somewhat clarified point 2, so I am increasing the soundness of the paper.**

Thank you for the constructive feedback – we tried our best to address the sensitivity analysis of our extrapolation layer hyperparameter as part of our results in the Effect of amateur layer selection section.

Add:

**Author-Editor Confidential Comment**

## Official Review of Submission3832 by Reviewer puvi

Official Review  Reviewer puvi  17 Jul 2024, 20:11 (modified: 22 Aug 2024, 15:45)

 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors, Reviewer puvi, Commitment Readers

 Revisions (/revisions?id=ZxIBzFGKGR)

### Paper Summary:

The paper introduces an innovative approach to enhance the performance of deep Transformer networks in fine-grained emotion classification tasks. By recasting contrastive inference as linear extrapolation, the method aims to achieve stable improvements in emotion classification. Extensive experiments across multiple models and datasets demonstrate that this approach outperforms standard classification methods. Compared to the previous versions, there has been an enhancement in terms of conciseness and interpretability; however, the article still leaves certain formal issues unresolved.

### Summary Of Strengths:

- The paper presents a novel method of contrastive classification that employs linear layer extrapolation, offering a fresh perspective in the domain of emotion analysis.

- The study is conducted across various models and emotion classification datasets, ensuring the robustness of the proposed method.
- The paper introduces a dynamic method for selecting the strength of contrastive penalties, allowing the model to adjust automatically based on the chosen amateur layer, thus enhancing robustness.

#### Summary Of Weaknesses:

- It is unclear whether contrastive classification is used only for inference. What is the final optimization object?
- The contrastive decoding is based on the Early-Existing Prediction method. Is there a classifier at each layer that predicts the probability of classification?
- In the Methods section, the used symbols are somewhat chaotic. It is suggested to further enhance the readability. How to get L\_valid on page 2, line 142? Is it the L on page 3, line 197?
- Regarding formatting issues, (1) the context following equations should be with no indent; (2) all equations should be numbered for reference. (3) Non-standardized headings such as 3.2, 4.2, B, and C should be revised.
- Source code is required to be submitted and publicly available.

#### Comments Suggestions And Typos:

See Weaknesses.

**Confidence:** 4 = Quite sure. I tried to check the important points carefully. It's unlikely, though conceivable, that I missed something that should affect my ratings.

**Soundness:** 3.5

**Overall Assessment:** 3 = Good: This paper makes a reasonable contribution, and might be of interest for some (broad or narrow) sub-communities, possibly with minor revisions.

**Best Paper:** No

**Needs Ethics Review:** No

**Reproducibility:** 3 = They could reproduce the results with some difficulty. The settings of parameters are underspecified or subjectively determined, and/or the training/evaluation data are not widely available.

**Datasets:** 1 = No usable datasets submitted.

**Software:** 1 = No usable software released.

**Knowledge Of Or Educated Guess At Author Identity:** No

**Knowledge Of Paper:** N/A, I do not know anything about the paper from outside sources

**Knowledge Of Paper Source:** N/A, I do not know anything about the paper from outside sources

**Impact Of Knowledge Of Paper:** Not at all

Add: **Author-Editor Confidential Comment**



### Official Comment by Authors

Official Comment

Authors ( Sean O'Brien (/profile?id=~Sean\_O'Brien1), Julian McAuley (/profile?id=~Julian\_McAuley1), Mayukh Sharma (/profile?id=~Mayukh\_Sharma1))

30 Jul 2024, 01:07 (modified: 22 Aug 2024, 15:45)

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors, Reviewer puvi, Commitment Readers

Revisions (/revisions?id=YKEKgtv9xN)

#### Comment:

Thank you for your thorough review and feedback! We have tried to address the points you raised and welcome further discussion.

#### Response to weaknesses:

##### It is unclear whether contrastive classification is used only for inference. What is the final optimization object?

We use contrastive classification as an inference time intervention requiring no additional finetuning or extra resources to extract maximum performance from the model. Previous works on contrastive methods mainly focus on runtime decoding, which has been shown to reduce hallucinations in text generation. We are the first to leverage contrastive inference methods for fine-grained emotion classification, demonstrating their effectiveness



in improving performance. Additionally, we introduce contrastive classification as linear layer extrapolation, enhancing the stability of contrastive methods and reducing the sensitivity to additional introduced hyperparameters. Contrastive classification does not require any additional training/optimization. The idea is to use different layers of the same network to boost model performance. Our work shows that for different emotions, models can make an early/late decision in their layers. Based on this observation we combine the idea of DoLA and contrastive decoding to fine-grained classification tasks.

**The contrastive decoding is based on the Early-Existing Prediction method. Is there a classifier at each layer that predicts the probability of classification?**

There are no separate classifiers at each layer. Following previous early-exiting work, we apply the pretrained classification head of the final network layer to the hidden activations of earlier layers without any changes. In other words, the same classification head is reused across all layers(amateur layers) to obtain the probability distribution over output labels. Lines 70–76 explain and cite the idea of early exiting as used in our work. As advised, we will make the use of the classification head clearer in the final version of the manuscript.

**In the Methods section, the used symbols are somewhat chaotic. It is suggested to further enhance the readability. How to get L\_valid on page 2, line 142? Is it the L on page 3, line 197?**

As advised we will make sure to improve the use of symbols to further enhance readability.

L\_valid on page 2 line 142 is the set of pre final layers which are used for amateur layer selection. As pointed out it is the same as L on page 3, line 197. This is a hyperparameter which is used to define the search space for amateur layer selection. We will revise the final version to address this and improve the readability of the manuscript. Appendix A further discusses the impact of this hyperparameter and advantage of using dynamic beta as proposed in our work.

**Regarding formatting issues, (1) the context following equations should be with no indent; (2) all equations should be numbered for reference. (3) Non-standardized headings such as 3.2, 4.2, B, and C should be revised.**



As advised we will ensure that we make these changes in the final version of our manuscript making it more concise and readable.

**Source code is required to be submitted and publicly available.**

We will make the source code publicly available on GitHub as requested.

Add: **Author-Editor Confidential Comment**

## Official Review of Submission3832 by Reviewer 7m4T

Official Review  Reviewer 7m4T  02 Jul 2024, 19:35 (modified: 22 Aug 2024, 15:45)

 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors, Reviewer 7m4T, Commitment Readers

 Revisions (/revisions?id=dhQO2YJ8eu)

### Paper Summary:

From the observations on fine-grained emotion recognition tasks, the authors obtained some findings: certain samples had their predicted emotions determined in the early layers, while others were decided in the final layers. After that, the authors propose the use of layer contrast for extrapolation to enhance some ambiguous emotional samples. Within the model, the authors present a new dynamic parameter design.

### Summary Of Strengths:

The discovery regarding the early layer's influence on prediction certainty is both interesting and innovative. I'd like to keep a strong interest in seeing these findings being shared at a conference.

The extrapolation approach sounds promising.

The experimental section is targeted and elaborates on the improvements made for ambiguous samples.

### Summary Of Weaknesses:

This is my second time reviewing this manuscript. Given the last discussion with authors, I'm glad to see the modifications to its format to make it clearer. All in all, there are no concerns from my side, except for some suggestions:

1, This paper would be more exciting with more visualization of cases about how to change the wrong emotion tag with extrapolation.

2, How about conducting some experiments on LLMs? I'm interested in this.

### Comments Suggestions And Typos:

above

**Confidence:** 4 = Quite sure. I tried to check the important points carefully. It's unlikely, though conceivable, that I missed something that should affect my ratings.

**Soundness:** 4 = Strong: This study provides sufficient support for all of its claims/arguments. Some extra experiments could be nice, but not essential.

**Overall Assessment:** 4 = This paper represents solid work, and is of significant interest for the (broad or narrow) sub-communities that might build on it.

**Best Paper:** No

**Needs Ethics Review:** No

**Reproducibility:** 4 = They could mostly reproduce the results, but there may be some variation because of sample variance or minor variations in their interpretation of the protocol or method.

**Datasets:** 3 = Potentially useful: Someone might find the new datasets useful for their work.

**Software:** 3 = Potentially useful: Someone might find the new software useful for their work.

**Knowledge Of Or Educated Guess At Author Identity:** No

**Knowledge Of Paper:** N/A, I do not know anything about the paper from outside sources

**Knowledge Of Paper Source:** N/A, I do not know anything about the paper from outside sources

**Impact Of Knowledge Of Paper:** Somehow

Add: **Author-Editor Confidential Comment**



### Official Comment by Authors

Official Comment

Authors ( Sean O'Brien (/profile?id=~Sean\_O'Brien1), Julian McAuley (/profile?id=~Julian\_McAuley1), Mayukh Sharma (/profile?id=~Mayukh\_Sharma1))

30 Jul 2024, 01:03 (modified: 22 Aug 2024, 15:45)

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors, Reviewer 7m4T, Commitment Readers

Revisions (/revisions?id=xMQAiGUmID)

#### Comment:

We thank the reviewer for their constructive feedback which helped us further improve the quality and readability of our manuscript.

#### Response to Weaknesses:

**This paper would be more exciting with more visualization of cases about how to change the wrong emotion tag with extrapolation.**

As advised, we will try to add more visualizations in the final version of our manuscript. We agree that it will further help in conveying the advantage of our proposed approach and help the readers better understand it.

**How about conducting some experiments on LLMs? I'm interested in this.**



In this work we restricted ourselves to relatively small models(500 million to 3 billion) as the datasets available for fine-grained classification tasks are limited in size and fine tuning them with smaller models resulted in more stable performance. Contrastive methods with larger models have been explored in the decoding space with good performance. We aim to extend the idea of dynamic contrastive strength(dynamic  $\beta$ ) in the decoding space as part of our future work where data size is not a limitation.




Add: **Author-Editor Confidential Comment**



## response

Official Comment  Reviewer 7m4T  31 Jul 2024, 01:20 (modified: 22 Aug 2024, 15:45)

 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors, Reviewer 7m4T, Commitment Readers

 Revisions (/revisions?id=0BciGQsls9)

### Comment:

Thank you for your reply. Break a leg.

Add: **Author-Editor Confidential Comment**

[About OpenReview \(/about\)](/about)

[Hosting a Venue \(/group?id=OpenReview.net/Support\)](/group?id=OpenReview.net/Support)

[All Venues \(/venues\)](/venues)

[Sponsors \(/sponsors\)](/sponsors)

[Frequently Asked Questions](#)

[\(https://docs.openreview.net/getting-started/frequently-asked-questions\)](https://docs.openreview.net/getting-started/frequently-asked-questions)

[Contact \(/contact\)](/contact)

[Feedback](#)

[Terms of Use \(/legal/terms\)](/legal/terms)

[Privacy Policy \(/legal/privacy\)](/legal/privacy)

[OpenReview \(/about\)](/about) is a long-term project to advance science through improved peer review, with legal nonprofit status through [Code for Science & Society \(https://codeforscience.org/\)](https://codeforscience.org/). We gratefully acknowledge the support of the [OpenReview Sponsors \(/sponsors\)](/sponsors). © 2024 OpenReview