NAACL 2025

The 5th Workshop on Insights from Negative Results in NLP

Proceedings of the Workshop

May 4, 2025

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-240-4

Introduction

Publication of negative results is difficult in most fields, and the current focus on benchmark-driven performance improvement exacerbates this situation and implicitly discourages hypothesis-driven research. As a result, the development of NLP models often devolves into a product of tinkering and tweaking, rather than science. Furthermore, it increases the time, effort, and carbon emissions spent on developing and tuning models, as the researchers have little opportunity to learn from what has already been tried and failed.

The mission of the workshop on Insights from Negative Results in NLP is to provide a venue for many kinds of negative results, with the hope that they could yield useful insights and provide a much-needed reality check on the successes of deep learning models in NLP. In particular, we solicit the following types of contributions:

- broadly applicable recommendations for training/fine-tuning, especially if X that didn't work is something that many practitioners would think reasonable to try, and if the demonstration of X's failure is accompanied by some explanation/hypothesis;
- ablation studies of components in previously proposed models, showing that their contributions are different from what was initially reported;
- datasets or probing tasks showing that previous approaches do not generalize to other domains or language phenomena;
- trivial baselines that work suspiciously well for a given task/dataset;
- cross-lingual studies showing that a technique X is only successful for a certain language or language family;
- experiments on (in)stability of the previously published results due to hardware, random initializations, preprocessing pipeline components, etc;
- theoretical arguments and/or proofs for why X should not be expected to work;
- demonstration of issues with under-reporting of training details of pre-trained models, including test data contamination and invalid comparisons.

The fifth iteration of the Workshop on Insights from Negative Results attracted 23 submissions and 2 from ACL Rolling Reviews. We accepted 16 papers, resulting in 64% acceptance rate. We hope the workshop will continue to contribute to the many reality-check discussions on progress in NLP. If we do not talk about things that do not work, it is harder to see what the biggest problems are and where the community effort is the most needed

Table of Contents

Program

Sunday, May 4, 2025

09:00 - 09:10 *Opening Remarks*

09:10 - 09:50 *Technical session 1*

09:50 - 10:30 *Technical session 2*

10:30 - 11:30 *Coffee Break*

11:30 - 12:00 *Invited Talk 1*

12:30 - 14:00 *Lunch*

14:10 - 14:50 *Technical session 3*

14:50 - 15:30 *Technical session 4*

15:30 - 16:00 *Coffee Break*

16:00 - 16:30 *Invited Talk 2*

16:30 - 17:30 *Poster Session*