



Universiteit
Leiden
The Netherlands

Opinion diversity through hybrid intelligence

Meer, M.T. van der

Citation

Meer, M. T. van der. (2025, March 26). *Opinion diversity through hybrid intelligence*. *SIKS Dissertation Series*. Retrieved from <https://hdl.handle.net/1887/4209024>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4209024>

Note: To cite this publication please use the final published version (if applicable).

II

3

Hybrid Intelligence for NLP

Introducing Part II: Hybrid Intelligence for NLP

We have seen how automated approaches to analyzing online discussions suffer from various limitations when applying them to realistic analysis scenarios. We saw how the generalization of LLMs is contingent on the diversity of training data, and how opinion frequency dictates how well a model can capture it, endangering alienating minority opinions. In Part II of this dissertation, we present our approach to incorporating humans and NLP methods for analyzing opinionated text data to address these limitations. First, we introduce a method for mining diverse arguments from citizen feedback in Chapter 4. Our method, HyEnA, finds more diverse arguments and improves the precision of the argument analysis compared to a manual and an automated approach. HyEnA guides human annotators across three distinct phases supported by LLMs for efficient selection of which opinions to analyze.

In Chapter 5, we further investigate how differences between annotators in subjective tasks, such as interpreting texts for extraction of arguments or personal values, can be modeled more efficiently. Our approach, Annotator-Centric Active Learning (ACAL), steers models to learn diverse label distributions by picking from a large pool of annotators. Central to our work, we create discussion analysis approaches that (1) select samples for human inspection that are interesting to annotate, (2) account for diversity, and (3) seek labels from multiple annotators. The hybrid nature of our methodology leads to **bidirectional gains**, serving the NLP system as well as the humans involved. For instance, we create approaches to capture more diverse interpretations of the arguments in discussions using a crowd of annotators. We observe that an active selection of diverse annotators can inform a model more quickly of the label distribution underlying subjective tasks in cases where the annotator pool is large. In Part III, we will apply our hybrid approach to capturing perspectives, and investigate the role of argumentation in constructing faithful opinion representations.

Part II focuses on the following research question:

Q2 How to combine human intelligence and NLP to effectively capture diverse perspectives?

