Patronizing and Condescending Language Detection in English and Mandarin

Xinyi Xiang
University of Washington
xinyix7@uw.edu

Catherine Qingyun Wang
University of Washington
qywang@uw.edu

Paula Cortés Lemos
University of Washington
mpcortes@uw.edu

Abstract

1 Introduction

2 Task Description

Our primary task comes from task 4 of SemEval 2022, which focuses on the detection of Patronizing and Condescending Language (PCL). The task describes PCL as language that 'reveals a superior attitude towards others' and is 'often unconscious but harmful and discriminative' (Perez-Almendros et al., 2022). Unlike toxic or offensive language, PCL is not overtly aggressive or inflammatory, but nonetheless furthers power imbalances and harmful stereotypes towards vulnerable communities.

The data set used for this task is the *Don't Patronize Me!* data set presented in Perez-Almendros et al. (2020), an annotated corpus of PCL towards vulnerable communities. The authors define seven subcategories of PCL:

- Unbalanced power relations: the author positions themselves as a savior of those in need

- Shallow solution: the author presents an overly simple solution as though the hardship were superficial

- Presupposition: the author makes assumptions or stereotypes about a vulnerable community

- Authority voice: the author presents advice as an authority about how to overcome under-privileged situations

- Metaphor: the author describes difficult situations in softer, poetic ways

- Compassion: the author uses flowery wording to describe the group with pity

- The poorer, the merrier: the author praises the hardships as beautiful

The SemEval 2022 task has two subtasks: first, the idenfication of PCL, a binary classification task; and second, the identification of which of the seven categories a given instance of PCL belongs to, a multi-class classification task.

Our primary task will focus on subtask 1. This task involves analyzing English text to detect the affect of condescension, a binary classification task. The task will involve text data only, and the genre of said text will be excerpts from news articles.

Our adaptation task will adapt this task for another language, Mandarin Chinese. We will use the CCPC (Chinese Corpus for Patronizing and Condescending Language) corpus, a corpus of discourse from Weibo and Zhihu platforms (Wang et al., 2023). This corpus additionally annotates entries with 'PCL toxicity strength', but can still be used for the binary classification task, distinguishing between entries that are labeled as not toxic and entries that are labeled with any level of toxicity.

3 System Overview

4 Approach

5 Results

6 Discussion

7 Ethical Considerations

References

Carla Perez-Almendros, Luis Espinosa Anke, and Steven Schockaert. 2020. Don't patronize me! An annotated dataset with patronizing and condescending language towards vulnerable communities. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 5891–5902, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Carla Perez-Almendros, Luis Espinosa-Anke, and Steven Schockaert. 2022. SemEval-2022 task 4: Patronizing and condescending language detection. In *Proceedings of the 16th International Workshop*

on Semantic Evaluation (SemEval-2022), pages 298–307, Seattle, United States. Association for Computational Linguistics.

Hongbo Wang, Mingda Li, Junyu Lu, Liang Yang, Hebin Xia, and Hongfei Lin. 2023. CCPC: A hierarchical Chinese corpus for patronizing and condescending language detection. In *Natural Language Processing and Chinese Computing*, pages 640–652, Cham. Springer Nature Switzerland.