Investigating the Robustness of Modelling Decisions for Few-Shot Cross-Topic Stance Detection: A Preregistered Study

Myrthe Reuver*, Suzan Verberne$, Antske Fokkens*^
*Vrije Universiteit Amsterdam, $Universiteit Leiden, ^Eindhoven University of Technology

Introduction

Democratic societies require citizens to be **aware of multiple viewpoints o**n salient issues in societies.. **News recommenders** can lead to situations were people get in so-called **filter-bubbles,** leading to increased fragmentation.

Research Questions

1. How do **different modelling choices** (task definitions and architecture differences) **affect few-shot classification performance on different stance datasets**?

2. To what extent do these modelling choices affect few-shot cross-topic robustness?

Methods

RoBERTa trained on Same Side Stance (Stein et. al, 2021) or Pro/Con Stance, and with bi or cross encoding and NLI or not, on 100 random shots per stance dataset.

Preregistration
Van Miltenburg et. al. (2021): making explicit expectations and experiments, and registering them before running them. *Why?*
- systematically comparing
- also reporting mixed results

Results

- Effects of Same Side Stance definition and also bi vs cross-encoding on performance **differs per dataset;**
- No clear relationship between number of training topics;
- NLI training gives considerable improvement, but inconsistent.

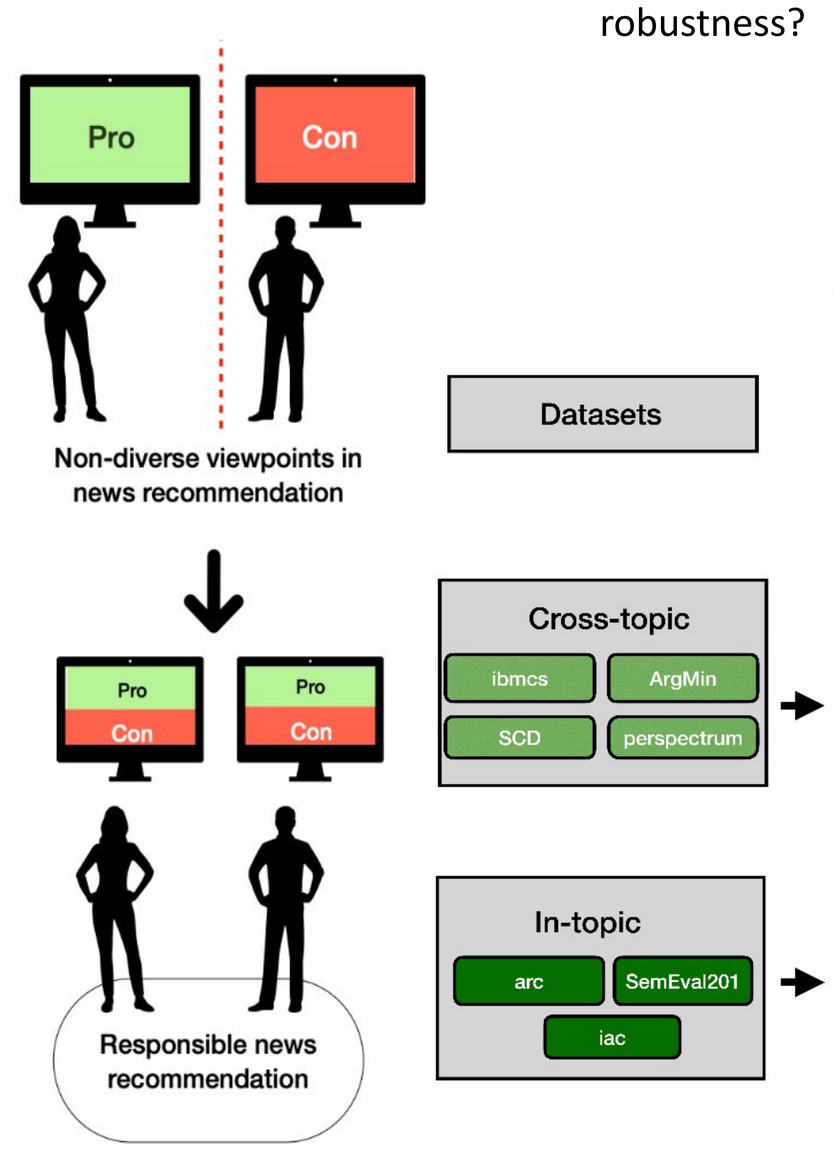

Figure 1: stance for responsible news recommendation

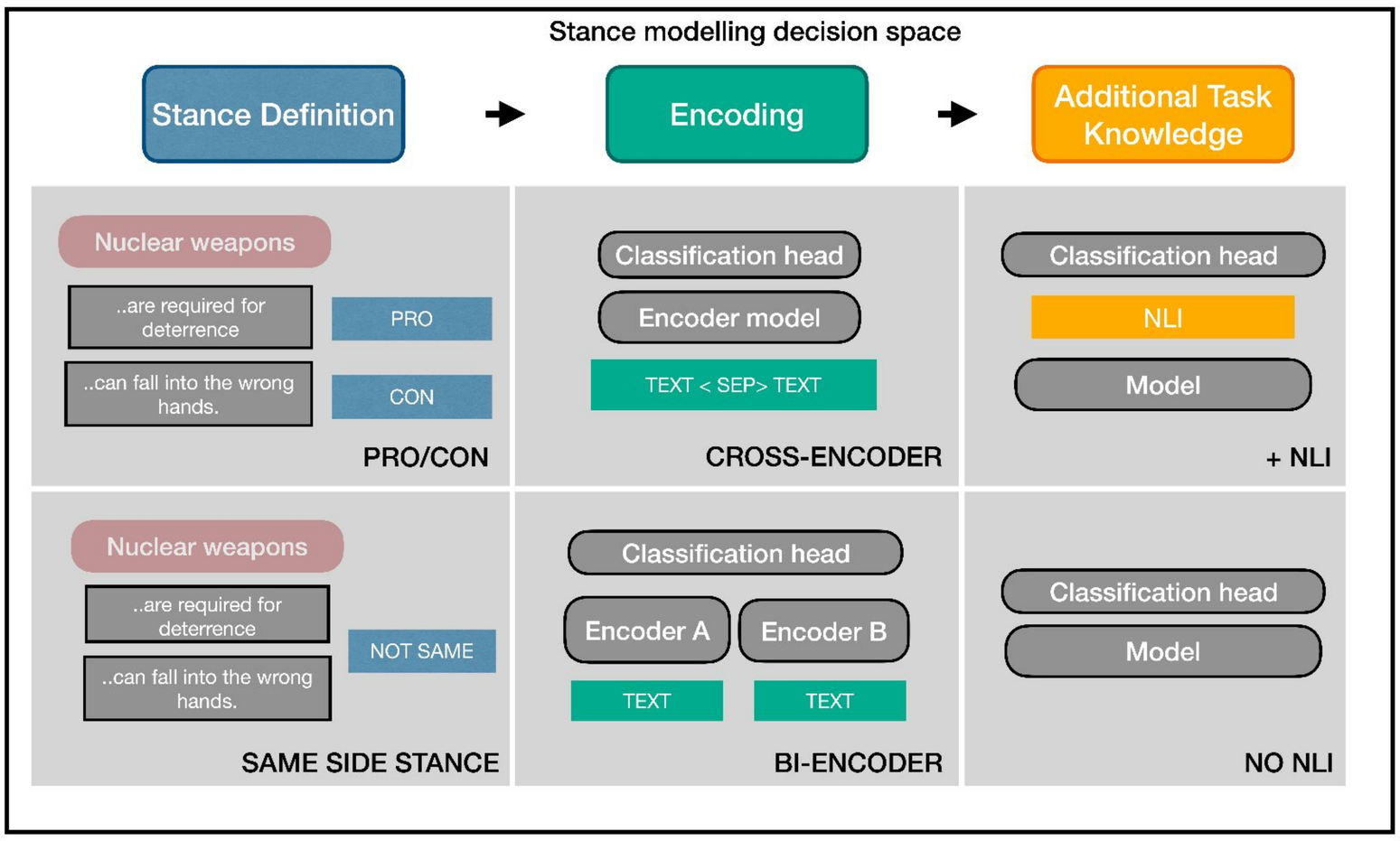

Figure 2: investigated modelling decisions

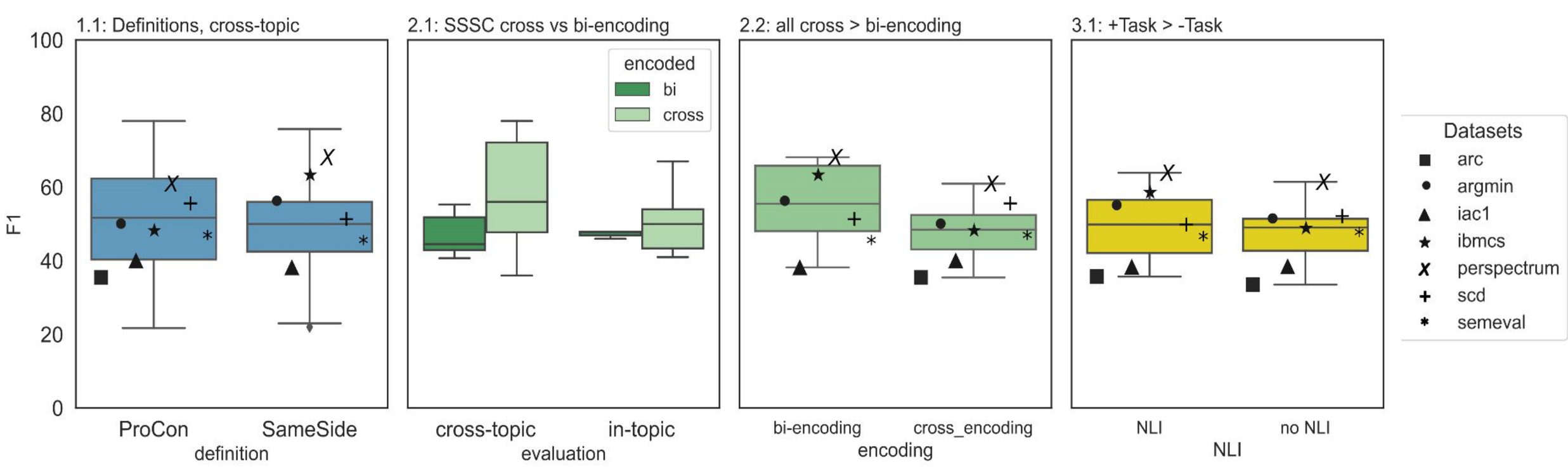

Figure 3: results over hypotheses

Paper:

References

Stein, Benno, et al. "Same side stance classification." Same Side Shared Task 2019 Proceedings.
Emiel van Miltenburg, Chris van der Lee, and Emiel Krahmer. 2021. Preregistering NLP research. NAACL 2022