Cross-topic Stances and How to Find Them (in the News)

Myrthe Reuver - PhD candidate @ CLTL, VU Amsterdam

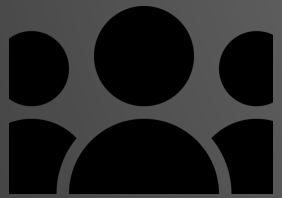
Talk at University of Groningen, 24-02-2023

VU VRIJE UNIVERSITEIT AMSTERDAM

Who am I?

Myrthe Reuver, PhD candidate at CLTL at VU Amsterdam.

→ Supervisors: Antske Fokkens (CLTL @ VU), Suzan Verberne (LIACS @ Leiden).

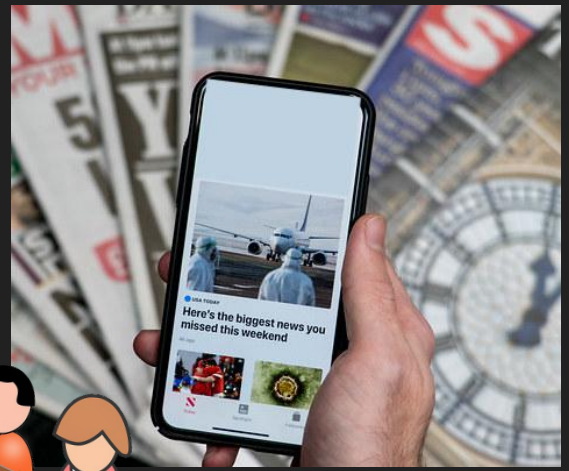


Computational linguist in an **interdisciplinary project** on diversity in news recommendation.

Social scientists, philosophers, and RecSys/computer scientists.

What am I doing?

- News recommendation provide **more of the same**
- Theoretic models of **democracy** (Vrijenhoek et. al., 2021)
  - deliberative model
  - critical model
- One key idea: **diverse viewpoints on issues**



How to operationalize this?

- maybe **stance?**
- Stances are **positional claims about topics** (e.g. gun control, immigration, abortion). They indicate a position: **pro, against, or neutral.**

What is (going on with) stance?

Common definition: **classification task** with labels Pro, Con, Neutral towards an issue or topic.

“Abortion is a sin, and should never be practiced.”

Topic: Abortion, Stance: Con

For online news recommendation:

New topics and issues continuously appear online!

Joseph Mucira @ Pixabay, Simplified Pixabay License

Cross-topic, cross-domain stance

Can we detect stance (pro, con)

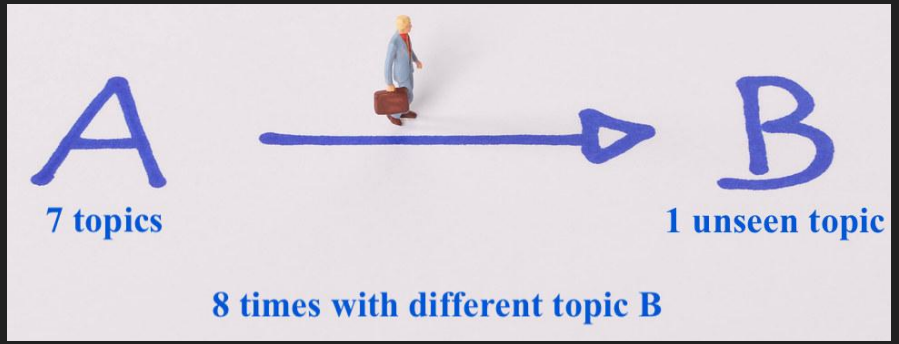
on discussion **topics or issues unseen** in training?

Reimers et. al. (2019)

→ Train: 7 topics, test: 8th topic

→ Fine-tuning BERT (base & large)

→ Findings:



Marco Verch @ Flickr, Creative Commons 2.0.
https://foto.wuestenigel.com/businessman-walking-from-a-to-b-point/

- avg. F1 (10 seeds) = .633
- +.20 over reference model (LSTM)
- Results are ***"very promising and stress the feasibility of the task"*** (Reimers et al. 2019, p. 575)

Dataset: UKP Dataset (Stab et. al., 2018)

25,492 arguments on 8 topics, in 3 classes:

- **For** or **against** “the use, adoption, or idea” of the topic, or **no argument**
- **8 controversial debate topics** from the internet: *abortion, cloning, death penalty, gun control, marijuana legalization, minimum wage, nuclear energy* and *school uniforms*.

Results

Model	UKP Dataset				
mean (stdev) 10 seeds	F1	P pro	P con	R pro	R con
Reimers et al. (2019) biclstm+BERT	.424	.267	.389	.281	.403
Reimers et al. (2019) BERT base	.613 (-)	.505 (-)	.531 (-)	.470 (-)	.576 (-)
Reimers et al. (2019) BERT large	**.633** (-)	.554 (-)	.584 (-)	.505 (-)	.560 (-)
SVM+tf-idf	.517	.418	.460	.414	.423
Reproduction BERT-base	**.617 (.006)**	.519 (.011)	.538 (.007)	.464 (.029)	.581 (.019)
Repr. BERT-large - all seeds	.596 (.043)	.483 (.057)	.527 (.057)	.464 (.058)	.516 (.063)
Repr. BERT-large - 5 evenly performing seeds	.636 (.007)	.532 (.014)	.578 (.016)	.515 (.016)	.567 (.022)

Reimers et. al. (2019) provided **excellent preliminaries for reproducibility:** documented, shared, working code (through a GitHub repository) + available for questions.

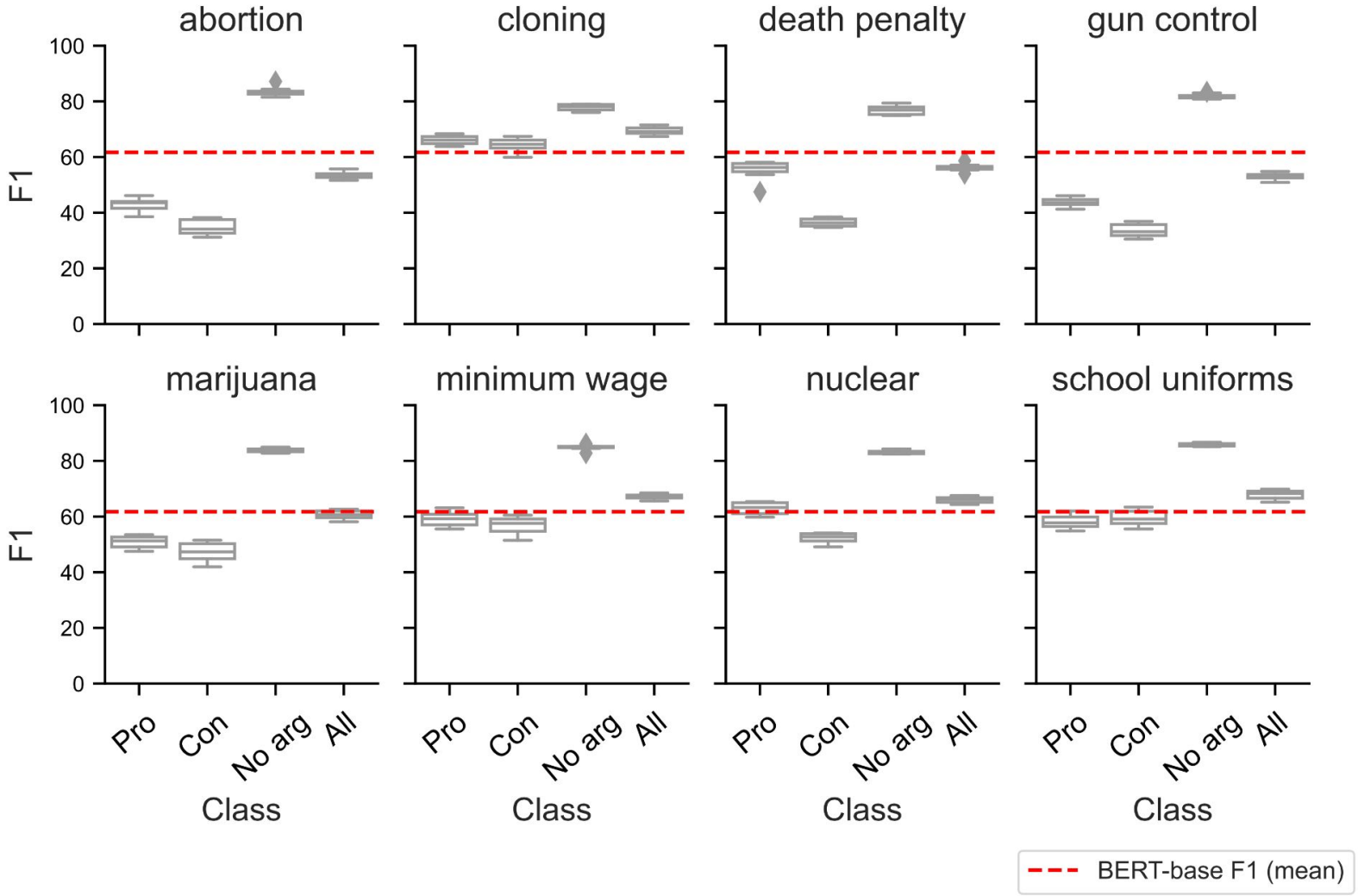
Results: further details

Model	UKP Dataset				
mean (stdev) 10 seeds	F1	P pro	P con	R pro	R con
Reimers et al. (2019) biclstm+BERT	.424	.267	.389	.281	.403
Reimers et al. (2019) BERT base	.613 (-)	.505 (-)	.531 (-)	.470 (-)	.576 (-)
Reimers et al. (2019) BERT large	**.633** (-)	.554 (-)	.584 (-)	.505 (-)	.560 (-)
SVM+tf-idf	.517	.418	.460	.414	.423
Reproduction BERT-base	**.617 (.006)**	.519 (.011)	.538 (.007)	.464 (.029)	.581 (.019)
Repr. BERT-large - all seeds	.596 (.043)	.483 (.057)	.527 (.057)	.464 (.058)	.516 (.063)
Repr. BERT-large - 5 evenly performing seeds	.636 (.007)	.532 (.014)	.578 (.016)	.515 (.016)	.567 (.022)

- BERT-large under-performs in 50% of seeds
- SVM+tf-idf model

What about different topics?

held-out topic	abortion	cloning	death penalty	gun control	marijuana legalization	minimum wage	nuclear energy	school uniform
SVM+tf-idf	.463	.585	.482	.515	.323	.615	.598	.576
BERT-base	.533 (.011)	.693 (.013)	.562 (.012)	.530 (.013)	.607 (.016)	.670 (.009)	.660 (.011)	.678 (.016)
diff.	+.070	+.108	+.080	*+.028*	**+.283**	+.055	+.0850	+.102


abortion
cloning
death penalty
gun control
marijuana
minimum wage
nuclear
school uniforms
F1
100
80
60
40
20
0
Pro
Con
No arg
All
Class
BERT-base F1 (mean)

Cross-Topic Stance: a Fantastic Beast?

- **Topic matters!** Stance not topic-independent → beyond reporting one avg F1
  - See also: Thorn Jakobsen et. al. (2021)

- A **class/topic interaction effect on performance**

OpenClipArt, Public domain

- **Recent work explores two avenues based on these results;**
  - **adding knowledge** to cross-topic stance detection (Beck et. al., 2023)
  - "ensure a **wide topic coverage** in model training to improve the model's generalizability" > more diverse topic datasets (Ajjour et. al., 2023).

What am I doing now?

Systematic stance detection experiments

I **pre-registered** RQs, hypotheses and analysis plans.

> From AsPredicted.com: *“Would a reader wonder whether a given decision about analysis, data source or hypothesis was made after knowing the results?"*

- **What?** Testing claims on what is more topic-independent, specifically Same Side Stance (SSS) in a **pair-wise classification setting.**

Example SSSC

Topic: '[This house believes] all nations have a right to nuclear weapons'

TASK: are these on the same side?

"Nuclear weapons may lessen a state's reliance on allies for security, thus preventing allies from dragging each other into wars" (**PRO**)

“Nuclear holocaust could result in an end to human life” (**CON**)

→ **Same side stance label: FALSE**

RQ1 hypotheses, example:

Hypothesis: *based on Shnarch et. al. (2022)'s experimental results on topic-dependent versus topic-independent tasks and pre-fine-tuning clustering,* we expect that SSSC models + pre-fine-tune clustering approach improve significantly over SSSC models without the pre-fine-tuning approach, since we consider stance classification a topic-dependent task and topic-dependent tasks responded well to this pre-fine-tuning task.

- Grounding in literature and/or earlier experiments;
- expectation;
- reasons

Annotating Dutch Stances with AnnoTinder (CCSAnnotator)

Dutch stance dataset on **sentences from news texts** on **the 2020 Dutch elections**

Stances in the news on four Issues: *Immigration, Climate measures, taxes, and European Union membership.*

Aim: diversity of stances, actors, issues in news recommendation

VVD komt in opstand tegen stikstofplannen eigen minister

Thank you!

Myrthe Reuver, Vrije Universiteit Amsterdam

myrthe.reuver[at]vu.nl

@myrthereuver

References

Reimers, N., Schiller, B., Beck, T., Daxenberger, J., Stab, C., & Gurevych, I. (2019, July). Classification and Clustering of Arguments with Contextualized Word Embeddings. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 567-578).

Küçük, D., & Can, F. (2020). Stance detection: A survey. ACM Computing Surveys (CSUR), 53(1), 1-37.

Jakobsen, T. S. T., Barrett, M., & Søgaard, A. Spurious Correlations in Cross-Topic Argument Mining. Proceedings of *SEM 2021: The Tenth Joint Conference on Lexical and Computational Semantics.

Reuver, M., Fokkens, A. & Verberne, S. (2021). No NLP Task Should be an Island: Multi-disciplinarity for Diversity in News Recommender Systems. In: Proceedings of the EACL Hackashop on News Media Content Analysis and Automated Report Generation (co-located at EACL 2021, online). Association of Computational Linguistics, p. 45–55.

Shnarch, E., Gera, A., Halfon, A., Dankin, L., Choshen, L., Aharonov, R., & Slonim, N. (2022, May). Cluster & Tune: Boost Cold Start Performance in Text Classification. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 7639-7653).

Stab, C., Miller, T., Schiller, B., Rai, P., & Gurevych, I. (2018). Cross-topic Argument Mining from Heterogeneous Sources. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (pp. 3664-3674).

Reuver, M., Verberne, S., Morante, R., & Fokkens, A. (2021). Is stance detection topic-independent and cross-topic generalizable?-A reproduction study. In Proceedings of the 8th Workshop on Argument Mining (pp. 46-56). Association for Computational Linguistics.