Leveraging political alignment information for stance detection

Matheus Camasmie Pavan and Ivandré Paraboni
School of Arts, Sciences and Humanities
University of São Paulo,
Av. Arlindo Bettio, 1000. São Paulo, Brazil
{matheus.pavan,ivandre}@usp.br

Abstract

Stance detection is the task of determining whether an input text expresses a stance in favour of or against a given target topic. This, in a standard supervised fashion, will typically require a new set of labelled training examples for each test topic. As an alternative to full supervision (or costly LLM-based methods), this study leverages political alignment information by assuming that stances on related moral or political issues tend to co-occur (e.g., support for a right-wing politician correlating with support for the death penalty or opposition to abortion). This alignment, presently treated as a form of distance labelling, enables stance inference without constructing new corpora and is evaluated against standard cross-domain and prompt-based methods using a large corpus of stances in the Portuguese language.

1 Introduction

Stance detection is the task of determining whether an input text expresses a stance in favour or against a given target topic (Kucuk and Can, 2020; Aldayel and Magdy, 2021; Alturayeif et al., 2023). For example, the text '*vaccines save lives*' may be interpreted as expressing a stance in favour of the target topic 'vaccination'. Stance detection models can be developed in a standard *in-domain* manner, that is, by employing as a knowledge source a set of training examples expressing stance toward the same topic (or domain). Under this approach, and assuming sufficient data quality and quantity, in-domain strategies may arguably achieve optimal performance. However, such models require a new topic-specific training dataset for each topic under evaluation. This limitation, compounded by the fact that training examples are often manually annotated, may render general-purpose in-domain stance detection impractical.

To overcome the need for continuously expanding labelled datasets, *cross-domain* stance detection (Khiabani and Zubiaga, 2025) seeks to reuse existing stance examples toward previously unseen topics as training data. For instance, one might attempt to detect stance toward a political figure by training a model with labelled examples expressing stance toward another politician or, more generally, a different topic. In this suboptimal setting, however, certain limitations of the in-domain approach may persist since a sufficiently large set of labelled examples – albeit concerning a topic other than the target – remains necessary.

To bridge the knowledge gap between in-domain and cross-domain methods, previous research – predominantly focused on English – has explored the use of external knowledge sources such as common sense networks (Liu et al., 2021; Liang et al., 2022; Chunling et al., 2023) which, in the case of Portuguese, remain largely unavailable. More recently, as in other areas of NLP, a shift towards the use of large language models (LLMs) has been observed (Xu et al., 2022; Lan et al., 2024; Zhang et al., 2025). The extensive knowledge embedded in LLMs can effectively mitigate the scarcity of training data, but achieving top performance typically requires closed-source models and involves substantial financial and computational costs.

As a cost-effective alternative to conventional cross-domain stance detection, the present study investigates a method that neither depends on labelled data nor relies on resource-intensive models. Specifically, it is assumed that stances towards multiple moral or political issues may *align* in the sense that supporting one target x may increase the likelihood of supporting another target y (for instance, favouring a right-wing politician may correlate with supporting the death penalty or, conversely, opposing abortion). This alignment information is then used to generate additional, low-cost training data for cross-target stance detection through distant labelling.

In this work, we propose a method for cross-

Proceedings of the 17th International Conference on Computational Processing of Portuguese (PROPOR 2026)
Vol. 1, pages 20–29

target stance detection that uses political alignment information as a substitute for manually labelled stance data. We evaluate the approach through a series of experiments comparing it with standard cross-domain models and prompt-based methods employing large language models. In addition, we examine whether any observed performance gains can be specifically attributed to the incorporation of political alignment information rather than simply to the availability of additional training data, and whether there is an upper limit to the amount of distantly labelled data that can be effectively utilised.

The remainder of this article is organised as follows. Section 2 reviews existing work on cross-target stance detection. Section 3 describes the corpus employed as training and test data for our experiments and the computational models used. Section 4 presents the experimental results, and Section 5 offers concluding remarks and outlines directions for future research.

2 Related work

Cross-domain and zero-shot stance detection are often treated as synonymous in the stance detection field. Table 1 summarises a number of recent studies of this kind devoted to the English and Portuguese languages. These are categorised according to the datasets used (VAST (Allaway and McKeown, 2020), SemEval-2016 (Mohammad et al., 2016), UstanceBR (Pereira et al., 2026) and *=others), language (En=English, Pt=Portuguese) and main computational methods (GCNN=Graph Convolutional Neural Networks, ADDA=Adversarial Discriminative Domain Adaptation, *=others, etc.). Details are discussed below.

Table 1 indicates that existing work in stance detection has predominantly focused on two major English-language datasets: the SemEval-2016 corpus (Mohammad et al., 2016) and VAST (Allaway and McKeown, 2020). The SemEval-2016 dataset contains 4,870 tweets covering six topics and was originally developed for in-domain stance classification, though it has also been used in some cross-domain studies (Xu et al., 2018). VAST (Allaway and McKeown, 2020), by contrast, was specifically designed for cross-target stance detection and comprises 23,525 stances across 5,634 targets, which implies that many targets are under-represented. This motivates the use of generalised topic or target representations discussed in (Allaway and McKeown, 2020).

Regarding the computational methods under consideration, cross-target stance detection employs a variety of strategies to address the scarcity of training data for the target topic. These include the aforementioned generalised topic representations (Allaway and McKeown, 2020), adversarial learning approaches (Allaway et al., 2021; Liang et al., 2022; Pavan and Paraboni, 2022; Zhao et al., 2022; Chunling et al., 2023) and large language model prompting (Zhang et al., 2023; Li et al., 2023; Pavan and Paraboni, 2024; Zhang et al., 2024a; Lan et al., 2024; Zhang et al., 2024b, 2025), among others.

For our target language, Portuguese, the only cross-target studies identified are those based on the UstanceBR corpus (Pereira et al., 2026). This corpus, which will be the focus of the present work as well, is discussed in Section 3.1.

Using an earlier version of UstanceBR, the work in (Pavan and Paraboni, 2022) addressed cross-target stance classification with BERT classifiers combined with adversarial learning and knowledge distillation. More recently, the work in (Pavan and Paraboni, 2024) explored a range of zero-shot methods, including commercial LLMs such as GPT, based on a small subset of the corpus comprising only 100 instances. Although both studies are relevant to the present work, their results are not directly comparable due to these differences in dataset size and definition.

3 Materials and methods

We envisage a method for cross-target stance detection that relies on political alignment information as a substitute for manually labelled stance data. The comparison between our main method and a number of relevant baseline systems aims to address the following research questions.

- Q1 How does the incorporation of distantly labelled data based on political alignment information compare with standard methods for stance detection?
- Q2 Are the results genuinely attributable to the political alignment information?
- Q3 Does the inclusion of additional distantly labelled data consistently yield improvements in accuracy?

To address question $Q1$, we compare our main approach – which leverages political alignment information – against standard cross-target and prompt-based methods for stance detection, with

Reference	Corpus	Lang.	Main methods
(Allaway and McKeown, 2020)	VAST	En	topic-related network
(Liu et al., 2021)	VAST	En	GCNN
(Allaway et al., 2021)	SemEval	En	adversarial network
(Liang et al., 2022)	SemEval, VAST, *	En	contrastive learning
(Xu et al., 2022)	SemEval, VAST, *	En	entailment, GPT, BERT
(Luo et al., 2022)	VAST	En	common sense graph
(Pavan and Paraboni, 2022)	UstanceBR r1	Pt	BERT ADDA
(Zhao et al., 2022)	SemEval, VAST, *	En	BERT+contrastive learning
(Chunling et al., 2023)	SemEval, *	En	graph+BERT ADDA
(Zhang et al., 2023)	SemEval, *	En	ChatGPT
(Wen and Hauptmann, 2023)	VAST	En	Conditional generation
(Li et al., 2023)	SemEval	En	ChatGPT, *
(Zhang et al., 2024a)	SemEval	En	ChatGPT
(Lan et al., 2024)	SemEval, VAST, *	En	GPT
(Zhang et al., 2024b)	VAST, *	En	BART
(Pavan and Paraboni, 2024)	UstanceBR r2	Pt	BERT, Llama, GPT
(Zhang et al., 2025)	SemEval, VAST, *	En	Qwen, Llama, Mistral
(Wang et al., 2025)	SemEval, *	En	GCNN, BERT
(Yan et al., 2025)	SemEval, VAST, *	En	BERT, GPT

Table 1: Recent work in cross-target stance detection based on English and Portuguese text.

the aim of demonstrating that the proposed method outperforms these baseline alternatives.

For question $Q2$, we contrast two data augmentation scenarios: one in which the additional data are labelled according to political alignment (as in our main approach), and another in which the data are assigned random labels. This comparison is intended to show that any observed performance gains stem from political alignment information rather than merely from an increased volume of training data.

Finally, to explore question $Q3$, we evaluate models trained with varying amounts of augmented data in order to determine whether an upper bound exists, that is, a point beyond which additional data no longer improve, or even degrade, model accuracy.

3.1 Data

Our experiments make use of the corpus UstanceBR r2 described in (Pereira et al., 2026). This comprises 46.8 k Twitter/X publications made in the Portuguese language and manually labelled with for/against stance information on six target topics of political nature (two Brazilian presidents, two measures discussed during the Covid-19 pandemic and two local institutions). This choice is motivated by the observation that, in this corpus, each pair comprises a topic that tends to be favoured by the political right and another that tends to be favoured by the left, which makes the data potentially ideal for the present study in the use of political alignment information for cross-target stance detection.

As in (Gohring et al., 2021), the corpus data includes both explicit and implicit (or less explicit) stances towards each topic. An explicit stance occurs when the given target is the actual linguistic topic under discussion in the text, whereas the more challenging less explicit stance occurs when the linguistic topic does not coincide with the intended target, but since the two are related, we may infer that the stance still reflects positively or negatively upon it. For instance, assuming the target to be a president of Brazil, in '*The president was wrong*' there is an explicit stance against the said president. By contrast, in '*The economy is worse than ever*', there is a stance towards a target other than the linguistic topic, that is, towards the economic situation, but since this indirectly reflects upon the president as well, it is also considered as a stance against the president in this corpus.

In our experiments, we used the standard train and test datasets provided by the corpus. Descriptive statistics are presented in Table 2. More details are available from (Pereira et al., 2026).

	Train		Test	
Target	Against	For	Against	For
Bolsonaro	4,173	2,887	1,392	962
Lula	3,385	2,855	1,129	951
Hydrox.	2,983	3,013	995	1,004
Sinovac	3,043	2,936	1,015	979
Globo TV	2,505	2,004	836	668
Church	2,654	2,698	885	900

Table 2: Data descriptive statistics, adapted from (Pereira et al., 2026).

3.2 Models

In the following sections, we present a computational approach to cross-target stance detection based on political alignment information, alongside the relevant baseline systems. This approach extends the pilot study on political alignment reported in (Pavan and Paraboni, 2024), expanding the original analysis to include the complete test data from (Pereira et al., 2026) and employing a more robust neural architecture for the task. The original study in (Pavan and Paraboni, 2024), by contrast, only examined a small sample of 100 test instances and employed a logistic regression classifier with a bag-of-words text representation.

3.2.1 Stance detection using political alignment information

Assuming that a target topic x may be aligned with a particular political ideology, this information can be leveraged to overcome some of the limitations of conventional stance prediction. In particular, we may use a non-stance corpus known to exhibit ideological bias (e.g., a collection of texts linked to the political right or left) to extract *unlabelled texts* that simply mention x, and use these as proxies for stance towards x. This procedure is illustrated in Figure 1.

In the example, given the target *Sinovac* and the task of predicting the stance expressed in '*Sinovac does not work*', the input text is processed by an augmented model trained on standard cross-target data (e.g., stances towards a different topic such as *Church*) together with distantly labelled data drawn from an auxiliary, non-stance corpus (illustrated on the right-hand side of the figure). The latter corpus consists of unlabelled statements that capture only the distinction between texts produced by government supporters and critics. Specifically, we extract from this auxiliary corpus a set of references to the target *Sinovac* and use it as a proxy for stance detection.

Although many of these proxy statements are factual in nature (e.g., *'They are offering Sinovac today'*) and may occasionally be mislabelled when, for instance, a government supporter expresses a stance in favour of vaccination, they nevertheless include genuine stances that exhibit the expected alignment with left- or right-leaning groups. As illustrated in the example, texts produced by government supporters include an oppositional stance towards the target in *'Authorities say Sinovac is safe. Really?'*, whereas those written by government critics express a favourable stance in *'Took my Sinovac today. I am so happy.'* When political alignment is strong, instances of this kind are likely to be frequent enough that the potential drawbacks of including factual (i.e., non-stance) or mislabelled data are outweighed by the benefits of incorporating additional distantly labelled examples.

As a computational implementation of this method, we use a neural architecture that integrates bi-directional long short-term memory (BiLSTM) networks and multi-head self-attention mechanisms with a BERT model pre-trained on Twitter/X data in Brazilian Portuguese (da Costa et al., 2023). The resulting model, hereby called BiLSTM, comprises an embedding layer of 128 to 256 dimensions, a recurrent layer containing 16 to 128 LSTM units, an attention depth of 8 or 32, and 2 or 4 attention heads. The specific configuration adopted for each task is determined through grid search.

In our experiments, the BiLSTM model is evaluated both as a baseline system in its own right (cf. the next section) and as the foundation for an extended variant, hereafter referred to as BiLSTM+, which represents our main proposed approach. BiLSTM+ augments the training data with a set of distantly labelled text instances extracted from a non-stance corpus that is assumed to be politically or ideologically aligned with the target topic under evaluation.

As a non-stance corpus of political information, BiLSTM+ uses the GovBR corpus described in (da Silva and Paraboni, 2023), a collection of complete Twitter/X timelines from supporters and critics of the former conservative government in Brazil. From this corpus, we extracted sets of unlabelled messages that mentioned any of the topics of interest within the present domain (cf. the previous section). This procedure yielded a large number of references to each topic made by either government

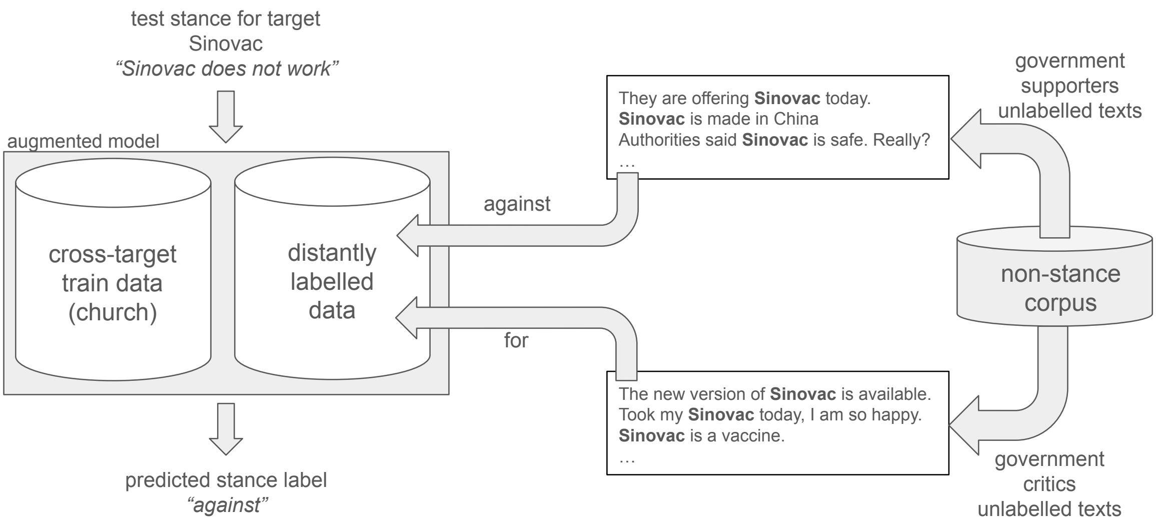

Figure 1: Training data expansion.

supporters or members of the opposition (541 k and 485 k unlabelled texts, respectively).

Assuming that the political orientation represented in the GovBR corpus correlates with the topics of the stance corpus, distant stance labels are assigned accordingly. For the polarised topics in the UStanceBR corpus, we assume that support for the conservative government correlates positively with support for *Bolsonaro*, *Hydroxychloroquine*, and *Church*, and negatively with support for *Lula*, *Sinovac*, and *Globo TV*.

3.2.2 Baseline systems

Our main approach BiLSTM+ will be compared with two types of baseline system. The first consists of the best train-test topic combinations obtained in cross-target fashion using the original BiLSTM model. Optimal results were obtained by using *Lula* as train data for the *Bolsonaro*, *Sinovac*, and *Globo TV* targets, using *Sinovac* as train data for the *Lula* and *Hydroxychloroquine* targets, and using *Hydroxychloroquine* data for the *Church* target.

The second baseline is a zero-shot method built using the open-source Alpaca LLM, a fine-tuned version of LLama-7B (Touvron et al., 2023). This consists of a prompt-based method that instructs the LLM to assess each test instance as in

> '*Read the following text and give a score between 0 and 10 where 0 means that the text is totally against the target and 10 means the text is totally in favour of the target*'.

This instruction was followed by the text to be evaluated, taken from the test portion of the corpus. Scores lower than 5 are subsequently converted to 'against' labels, whereas all others are converted to 'for' labels.

4 Evaluation

Our main approach BiLSTM+ and the two baseline systems BiLSTM and Alpaca were evaluated on the test portion of the corpus in (Pereira et al., 2026) by computing F1 scores. In the case of BiLSTM baseline, all possible train-test topic combinations were evaluated, but only the results from the top-performing configurations are presently reported.

4.1 Q1: Comparison with standard cross-domain and prompt-based methods

Our first research question Q1 concerns the comparison between the model enriched with distantly labelled political alignment information (BiLSTM+) with Alpaca and BiLSTM. Results are summarised in Table 3, in which the best scores for each class are highlighted.

These results are further illustrated by the class distributions in Figure 2, in which larger overlap (brown) areas indicate a less clear distinction between the positive and negative classes, most notably in the *Globo TV* and *Church* targets.

The present results show that, generally speaking, BiLSTM+ outperforms both the Alpaca and BiLSTM baselines by a considerable margin. This provides an answer to our research question $Q1$.

	Prompt-based	Best cross-domain	Distant
Target	Alpaca	BiLSTM	BiLSTM+
Bolsonaro	0.60	0.56	**0.86**
Lula	0.64	0.70	**0.81**
Hydrox.	0.42	0.71	**0.84**
Sinovac	0.61	0.74	**0.83**
Globo TV	0.67	**0.72**	0.70
Church	0.30	0.64	**0.71**
Mean	0.55	0.68	**0.79**

Table 3: Stance prediction F1 score results. The best results for each target are highlighted.

4.2 Q2: The role of political alignment information

Research question $Q2$ aims to determine whether the performance gains reported in the previous section stem from the inclusion of political alignment information or merely from the availability of additional training data. To examine this, the training set of the distantly labelled BiLSTM+ model was progressively altered by replacing varying proportions of its instances (ranging from 0% to 100%) with randomly assigned labels, and the model was re-evaluated under each configuration. The results of this analysis are presented in Table 4, where the first row corresponds to the original model without any random label replacement and shows the same results from Table 3 for comparison. The best F1 score for each class is highlighted.

Noise	bo	lu	hy	si	gl	ch
0% (orig.)	**0.86**	**0.81**	**0.84**	**0.83**	**0.70**	**0.71**
25%	0.84	0.78	0.82	0.82	0.63	0.67
50%	0.74	0.71	0.73	0.76	0.60	0.68
75%	0.63	0.59	0.60	0.62	0.55	0.55
100%	0.52	0.51	0.50	0.52	0.51	0.53

Table 4: BiLSTM+ model F1 score results with random substitution of 0 - 100% of train labels.

The results presented in Table 4 show that replacing any proportion of examples with randomly labelled instances leads to a decrease in the performance of the model. This observation may be taken as indirect evidence that the distant labelling strategy used to generate the data originally employed in the model (as shown in the first row of the table) is largely accurate, or at least substantially more reliable than random assignment. This outcome provides an answer to our research question $Q2$.

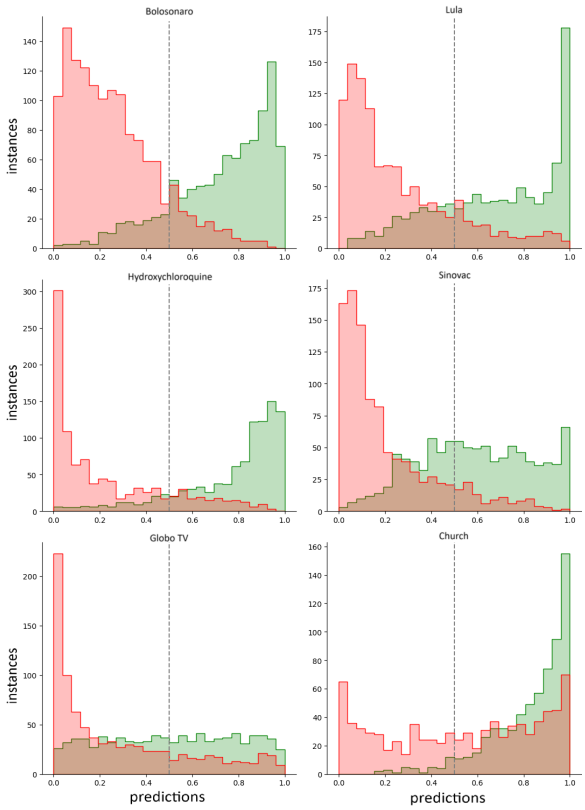

Figure 2: Class distribution per target (red=against, green=in favour, brown=overlap) – best visualised in electronic format.

4.3 Q3: Effects of the training data size

The experiment discussed in the previous section suggests that the data produced through distant labelling are indeed responsible for the improvement in model performance, and naturally raises the question of what the optimal volume of data would be for training under this method, which is the focus of our research question $Q3$.

To investigate this issue, new BiLSTM+ models were trained using the same parameters that yielded the best results in the previous experiment, this time employing all available data from the auxiliary corpus of unlabelled texts. Table 5 presents a comparison of the results obtained by distant labelling with different sizes of training set, in which the best F1 score is highlighted for each class.

The results presented in Table 5 indicate that, on average, the highest F1-macro scores are achieved by models trained with the maximum number of training instances available for each topic. However, the difference between the F1-macro scores obtained from models trained on samples of 15,000 instances and those trained on the entire dataset amounts, on average, to approximately 1%. In ad-

Added inst.	bo	lu	hy	si	gl	ch
500	0.83	0.76	0.80	0.76	0.68	0.65
1,000	0.84	0.77	0.80	0.78	0.67	0.66
5,000	0.85	0.80	0.83	0.80	0.66	0.67
10,000	0.85	0.79	0.84	0.82	0.67	0.66
15,000	0.86	0.81	0.84	**0.83**	0.70	**0.71**
30,000	0.85	0.80	**0.85**		0.68	
60,000	0.86	0.80			0.67	
...	...	...	...	...	...	...
Maximum	**0.87**	**0.82**	**0.85**	0.81	**0.71**	0.64

Table 5: BiLSTM+ model F1 score results with the addition of 500 - 60,000 train instances., or all (maximum) available data for each topic.

dition, models trained with samples of only 500 instances display an average degradation of less than 5% compared to those trained on all available data. Although training with the complete dataset yields models with the best overall performance, increasing the number of training instances – and consequently the computational cost – leads to only marginal improvements, which may not justify the additional expense in resource-constrained scenarios. This provides an answer to our research question $Q3$.

4.4 Error analysis

Table 6 shows randomly selected examples of false positive tests, translated into English from the original Portuguese data. Similarly, Table 7 shows randomly selected examples of false negative tests.

Both the false positive and the false negative examples reveal certain general behavioural patterns of the models for each target. For the target *Bolsonaro*, the selected examples are generally complex. The stance inference is not straightforward and would require an additional reasoning step involving external knowledge or inference of relations between targets. In the first example in Table 6, for instance, a negative stance is directed towards individuals who support Bolsonaro, which could be indirectly inferred as an unfavourable stance towards the target itself, a nuance that the model does not appear to have captured.

Furthermore, instances of irony can be identified in the examples associated with the targets *Chloroquine*, *Coronavac*, and *Lula*, which likely lead the models to assign reversed polarity labels. In the specific case of false positives for the target *Lula*, the model appears to enforce an oppositional relationship between stances on Lula and Bolsonaro, which results in misclassification. This behaviour may stem from the fact that the original GovBR dataset was constructed around the target *Bolsonaro*, while the target *Lula* is frequently associated with the opposite end of the ideological spectrum.

Finally, for the targets *Church* and *Globo*, the texts tend to express clearer and more direct stances, suggesting a weaker correlation with the political alignment information employed by the model.

5 Conclusion

Computational stance detection faces a persistent challenge in its dependence on labelled training data, which can be mitigated through the use of external knowledge bases or, more recently, with the support of large language models (LLMs). However, the knowledge encapsulated in LLMs is not limited to closed-source or commercial solutions. To some extent, similar knowledge may be acquired at a much lower cost from the same underlying sources, that is, from unlabelled text freely available online. This observation motivates the exploration of methods that can exploit this information without relying on extensive supervision or expensive computational resources.

As an example of this approach, the present study investigates the use of political alignment information in cross-target stance detection. By comparing a method based on distant labelling with standard cross-domain models and prompt-based approaches employing LLMs, the results indicate that incorporating unlabelled text – including factual or otherwise non-stance utterances – enhances performance, surpassing both cross-target and prompt-based baselines. Moreover, the experiments suggest that these improvements stem specifically from the use of political alignment information, rather than from the mere inclusion of additional training data.

The present study has several limitations that highlight opportunities for further research. In particular, our experiments focused on an idealised scenario involving topic-target pairs with strong polarisation and considered only six topics. Future work should extend this analysis to a broader and more diverse set of topics to assess the generality of the approach.

More broadly, it remains an open question whether the present method can be generalised to

Target	False positive examples	Score
Bolsonaro	*I unfollowed who follows Bolsonaro*	0.68
	Regardless of arguing with or excluding anyone, I didn't elect Bolsonaro	0.68
Lula	*Lula and Bolsonaro are completely disgusting*	0.67
	Lula out, Bolsonaro out	0.63
Hydrox.	*Hydroxicloroquine is saving the virus*	0.65
	I still don't believe in Hydroxicloroquine	0.68
Sinovac	*Very insteresting, Coronavac is what is left for us*	0.71
	Coronavac is vaccine for cattle	0.74
Globo TV	*I'm very disappointed with you, Globo*	0.65
	So much hate for Globo	0.62
Church	*I boycotted the church a long time ago*	0.85
	Do you miss going to church? – No	0.67

Table 6: False positive examples using BiLSTM+, adapted form the Portuguese original.

Target	False negative examples	Score
Bolsonaro	*Bolsonaro's win is the victory for Brazil and Latin America*	0.45
	I just want to know in which way Bolsonaro attacks democracy	0.28
Lula	*Lula is convincing me that he is not a thief*	0.21
	Lula thief, stole my heart	0.24
Hydrox.	*Each person has their own Hydroxicloroquine to believe in*	0.09
	Hydroxicloroquine is not medicine for donkeys	0.43
Sinovac	*Saying bad things about Coronavac is fake news*	0.27
	Let's go Coronavac	0.41
Globo TV	*Exclusive by Globo TV reveals turn of events in Bolsonaro's case*	0.41
	Just stay at home watching Globo TV, that's THE channel	0.21
Church	*It's unbelievable how happy I feel in the church*	0.42
	I want to go to church so badly	0.26

Table 7: False negative examples using BiLSTM+, adapted form the Portuguese original.

less political topics, where alignment information may have weaker or negligible effects. Future research could explore alternative forms of alignment or examine whether the observed phenomena are specific to polarised political discourse.

Finally, another possible future direction is the integration of stance detection models with other NLP tasks such as author profiling (dos Santos and Paraboni, 2022; Zarifi and Naghavi, 2023) and authorship attribution (Custódio and Paraboni, 2021; Huang et al., 2025), among others.

References

Abeer Aldayel and Walid Magdy. 2021. Stance detection on social media: State of the art and trends. *Information Processing & Management*, 58(4):102597.

Emily Allaway and Kathleen R. McKeown. 2020. Zero-shot stance detection: A dataset and model using generalized topic representations. In *EMNLP-2020 proceedings*, pages 8913–8931, Online. Assoc. for Computational Linguistics.

Emily Allaway, Malavika Srikanth, and Kathleen McKeown. 2021. Adversarial learning for zero-shot stance detection on social media. In *ACL-2021 proceedings*, pages 4756–4767, Online. Assoc. for Computational Linguistics.

Nora Alturayeif, Hamzah Luqman, and Moataz Ahmed. 2023. A systematic review of machine learning techniques for stance detection and its applications. *Neural Comput & Applic*, 35:5113–5144.

Wang Chunling, Zhang Yijia, Yu Xingyu, Liu Guantong, Chen Fei, and Lin Hongfei. 2023. Adversarial network with external knowledge for zero-shot stance detection. In *Proceedings of the 22nd Chinese National Conference on Computational Linguistics*, pages 824–835, Harbin, China. Chinese Information Processing Society of China.

José Eleandro Custódio and Ivandré Paraboni. 2021. Stacked authorship attribution of digital texts. *Expert Systems with Applications*, 176:114866.

Pablo Botton da Costa, Matheus Camasmie Pavan, Wesley Ramos dos Santos, Samuel Caetano da Silva, and Ivandré Paraboni. 2023. BERTabaporu: assessing a genre-specific language model for Portuguese NLP. In *Recent Advances in Natural Language Processing (RANLP-2023)*, pages 217–223, Varna, Bulgaria.

Samuel Caetano da Silva and Ivandré Paraboni. 2023. Politically-oriented information inference from text. *Journal of Universal Computer Science*, 29(6):570–595.

Vitor Garcia dos Santos and Ivandré Paraboni. 2022. Myers-briggs personality classification from social media text using pre-trained language models. *Journal of Universal Computer Science*, 28(4):378–395.

Anne Gohring, Manfred Klenner, and Sophia Conrad. 2021. DeInStance: Creating and evaluating a german corpus for fine-grained inferred stance detection. In *17th Conference on Natural Language Processing (KONVENS 2021)*, pages 213–217, Düsseldorf, Germany. KONVENS 2021 Organizers.

Baixiang Huang, Canyu Chen, and Kai Shu. 2025. Authorship Attribution in the Era of LLMs: Problems, Methodologies, and Challenges. *SIGKDD Explor. Newsl.*, 26(2):21–43.

Parisa Jamadi Khiabani and Arkaitz Zubiaga. 2025. Cross-target stance detection: A survey of techniques, datasets, and challenges. *Expert Systems with Applications*, 283:127790.

Dilek Kucuk and Fazli Can. 2020. Stance detection: A survey. *ACM Computing Surveys*, 53(1):1–37.

Xiaochong Lan, Chen Gao, Depeng Jin, and Yong Li. 2024. Stance Detection with Collaborative Role-Infused LLM-Based Agents. *Proceedings of the International AAAI Conference on Web and Social Media*, 18(1):891–903.

Ang Li, Bin Liang, Jingqian Zhao, Bowen Zhang, Min Yang, and Ruifeng Xu. 2023. Stance detection on social media with background knowledge. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 15703–15717, Singapore. Association for Computational Linguistics.

Bin Liang, Qinglin Zhu, Xiang Li, Min Yang, Lin Gui, Yulan He, and Ruifeng Xu. 2022. JointCL: A joint contrastive learning framework for zero-shot stance detection. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 81–91, Dublin, Ireland. Association for Computational Linguistics.

Rui Liu, Zheng Lin, Yutong Tan, and Weiping Wang. 2021. Enhancing zero-shot and few-shot stance detection with commonsense knowledge graph. In *Findings of ACL-IJCNLP-2021*, pages 3152–3157, Online. Assoc. for Computational Linguistics.

Yun Luo, Zihan Liu, Yuefeng Shi, Stan Z. Li, and Yue Zhang. 2022. Exploiting sentiment and common sense for zero-shot stance detection. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 7112–7123, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Saif Mohammad, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. 2016. SemEval-2016 task 6: Detecting stance in tweets. In *10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 31–41, San Diego, California. Assoc. for Computational Linguistics.

Matheus Camasmie Pavan and Ivandré Paraboni. 2022. Cross-target stance classification as domain adaptation. In *Advances in Computational Intelligence - MICAI 2022 - Lecture Notes in Artificial Intelligence vol 13612*, pages 15–25, Cham. Springer Nature Switzerland.

Matheus Camasmie Pavan and Ivandré Paraboni. 2024. A benchmark for portuguese zero-shot stance detection. *Journal of the Brazilian Computer Society*, 30(1):469–479.

Camila Pereira, Matheus Pavan, Sungwon Yoon, Ricelli Ramos, Pablo Costa, Laís Cavalheiro, and Ivandré Paraboni. 2026. UstanceBR: a social media language resource for stance prediction. *Language Resources and Evaluation*, 60(14).

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. 2023. Llama: Open and efficient foundation language models. Cite arxiv:2302.13971.

Hongbin Wang, Kunqiang Zhang, and Yantuan Xian. 2025. Cross-target stance detection with multi-level information fusion. *IEICE Transactions on Information and Systems*, E108.D(8):947–957.

Haoyang Wen and Alexander Hauptmann. 2023. Zero-shot and few-shot stance detection on varied topics via conditional generation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 1491–1499, Toronto, Canada. Association for Computational Linguistics.

Chang Xu, Cecile Paris, Surya Nepal, and Ross Sparks. 2018. Cross-target stance classification with self-attention networks. In *ACL-2018 proceedings*, pages

778–783, Melbourne, Australia. Assoc. for Computational Linguistics.

Hanzi Xu, Slobodan Vucetic, and Wenpeng Yin. 2022. OpenStance: Real-world zero-shot stance detection. In *Proceedings of the 26th Conference on Computational Natural Language Learning (CoNLL)*, pages 314–324, Abu Dhabi, United Arab Emirates (Hybrid). Association for Computational Linguistics.

Yu Yan, Sheng Sun, Zixiang Tang, Teli Liu, and Min Liu. 2025. Collaborative stance detection via small-large language model consistency verification. *arXiv*, 2502.19954.

A. Zarifi and M. Naghavi. 2023. Gender identification of short text author using conceptual vectorization. *Multimed Tools Appl*, 82:17097–17113.

Bowen Zhang, Daijun Ding, and Liwen Jing. 2023. How would stance detection techniques evolve after the launch of chatgpt? *Preprint*, arXiv:2212.14548.

Bowen Zhang, Xianghua Fu, Daijun Ding, Hu Huang, Genan Dai, Nan Yin, Yangyang Li, and Liwen Jing. 2024a. Investigating chain-of-thought with chatgpt for stance detection on social media. *arXiv*.

Zhao Zhang, Yiming Li, Jin Zhang, and Hui Xu. 2024b. LLM-driven knowledge injection advances zero-shot and cross-target stance detection. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 2: Short Papers)*, pages 371–378, Mexico City, Mexico. Association for Computational Linguistics.

ZhaoDan Zhang, Zhao Zhang, Jin Zhang, Hui Xu, and Xueqi Cheng. 2025. MPVStance: Mitigating hallucinations in stance detection with multi-perspective verification. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1053–1067, Vienna, Austria. Association for Computational Linguistics.

Xuechen Zhao, Jiaying Zou, Zhong Zhang, Feng Xie, Bin Zhou, and Lei Tian. 2022. Zero-shot stance detection based on cross-domain feature enhancement by contrastive learning. *arXiv:2210.03380*.