Implicit Reasoning over Temporal Relations in Evidence-Based Fact-Checking

Marlon Saelens, Marie-Francine Moens, Ruben Cartuyvels, and Liesbeth Allein*

Department of Computer Science, KU Leuven, Celestijnenlaan 200A, 3001 Leuven, Belgium

Keywords: automated fact-checking · temporal reasoning · temporal semantics

1 Introduction

This study[1] introduces a novel method for incorporating temporal reasoning in an evidence-based fact-checking model. Inspired by Allein et al. [1] who enforced *explicit* temporal relations during model training, we let a model *implicitly* reason over temporal relations between a given claim and its accompanying set of evidence documents. The fact-checking model presented by Augenstein et al. [2] serves as baseline, which we enrich with temporal reasoning abilities. This is done by modifying the contextualised representations of the claim, h_c, and the evidence documents, h_e, using learned time embeddings.

2 Implicit Reasoning over Temporal Relations

Grounding claims and evidence in time We divide a claim and its evidence documents into buckets according to the time difference between the claim and evidence document in terms of publication date *(E1)* and/or time references available in their text *(E2)* (see Figure 1). Publication dates are available in the metadata (for claims) or deduced from the beginning of the text (for evidence). Claims and evidence documents lacking a publication date are assigned to special buckets. We use HeidelTime [8] for extracting time references in both claim and evidence text. Every time bucket k implies a time attribution vector which has the same size as h_c and h_e, and which is updated during model training.

Expansion 1: Publication date of the evidence (E1) Time-aware claim or evidence representations are obtained by taking the weighted sum of h_c or h_e and the attribution vector of the time bucket where the claim or evidence p_k belongs to, with hyperparameter α:

$$h_c^* || h_e^* = h_c || h_e + \alpha \, p_k \ . \tag{1}$$

Expansion 2: Time references in claim and evidence text (E2) In this expansion, a claim or evidence can belong to zero or multiple time buckets as multiple time references can be extracted from the text. To obtain time-aware representations,

* From left to right: thesis student, promotor, advisor 1, advisor 2.
[1] Abstract of master thesis, graded at department of Computer Science, KU Leuven.

the weighted sum is taken of h_c or h_e and the average of the set of S attribution-vectors t_k of the buckets the claim or evidence are part of:

$$h_c^*||h_e^* = h_c||h_e + \beta \, \frac{\sum_{k \in S} t_k}{|S|} \; . \tag{2}$$

Expansion 3: Combining E1 and E2 Both expansions are combined using the following weighted sum with hyperparameters α and β:

$$h_c^*||h_e^* = h_c||h_e + \alpha \, p_k + \beta \, \frac{\sum_{k \in S} t_k}{|S|} \; . \tag{3}$$

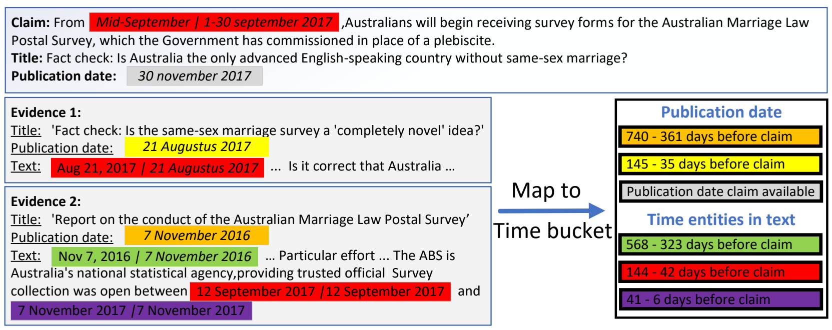

Fig. 1. Example of the division of the claim and evidence in time buckets according to the publication date (E1) and time-entities in the text (E2)

3 Results

Table 1. Aggregated test results for the veracity prediction task in Micro and Macro F1, with improvement difference according the base model as subscript.

	BiLSTM		DistilRoBERTa	
	Micro F1	*Macro F1*	*Micro F1*	*Macro F1*
baseline (no time)	.5520	.3239	.6952	.5532
+ publication date (E1)	.6006 $_{+8.8\%}$	.4271 $_{+31.9\%}$	**.6973** $_{+0.3\%}$	.5608 $_{+2.7\%}$
+ in-text time references (E2)	.6089 $_{+10.3\%}$	.4425 $_{+36.6\%}$	.6882 $_{-1.0\%}$	**.5744** $_{+3.8\%}$
+ E1 and E2	**.6417** $_{+16.3\%}$	**.4743** $_{+46.4\%}$	.6947 $_{-0.1\%}$	.5739 $_{3.7\%}$

We conduct experiments on the MultiFC dataset [2], which contains a large variety of fact-checked claims accompanied by evidence documents from the Web. We also experiment with two encoding architectures, namely a bistacked BiLSTM or a DistilRoBERTa[2] model. Table 1 shows the average test results across all fact-check domains for the veracity prediction task. For both the BiLSTM and DistilRoBERTa model, the evidence documents sharing the same time bucket have similar label preferences. This is shown by the fact that the Spearman's rank correlation coefficient between the label ranking of the evidence in the same bucket is higher than those not in the same bucket for both models.

These results show that implicit reasoning over temporal relations further improves fact-checking in comparison with explicit methods [1, 3]. We hope this encourages others to apply other temporal reasoning methods [4, 5].

[2] Fine-tuned starting from pretrained weights [7, 9, 6].

References

1. Allein, L., Augenstein, I., Moens, M.F.: Time-aware evidence ranking for fact-checking. Journal Of Web Semantics **71**(C) (nov 2021). https://doi.org/10.1016/j.websem.2021.100663
2. Augenstein, I., Lioma, C., Wang, D., Chaves Lima, L., Hansen, C., Hansen, C., Simonsen, J.G.: MultiFC: A real-world multi-domain dataset for evidence-based fact checking of claims. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). pp. 4685–4697. Association for Computational Linguistics, Hong Kong, China (Nov 2019). https://doi.org/10.18653/v1/D19-1475, https://aclanthology.org/D19-1475
3. Halpin, T.: Temporal modeling and orm. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Lecture Notes in Computer Science, vol. 5333, pp. 688–698. Springer Berlin Heidelberg, Berlin, Heidelberg (2008)
4. Leeuwenberg, A., Moens, M.F.: Temporal information extraction by predicting relative time-lines. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. pp. 1237–1246. Association for Computational Linguistics, Brussels, Belgium (Oct-Nov 2018). https://doi.org/10.18653/v1/D18-1155, https://aclanthology.org/D18-1155
5. Leeuwenberg, T., Moens, M.F.: Towards extracting absolute event timelines from english clinical reports. Transactions on Audio, Speech and Language Processing **28**, 2710–2719 (2020). https://doi.org/10.1109/TASLP.2020.3027201
6. Reimers, N., Gurevych, I.: Sentence-bert: Sentence embeddings using siamese bert-networks. In: Inui, K., Jiang, J., Ng, V., Wan, X. (eds.) Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019, Hong Kong, China, November 3-7, 2019. pp. 3980–3990. Association for Computational Linguistics (2019). https://doi.org/10.18653/v1/D19-1410
7. Sanh, V., Debut, L., Chaumond, J., Wolf, T.: Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. In: Thirty-third Conference on Neural Information Processing Systems (NeurIPS 2019). Neural Information Processing Systems Foundation, Ithaca (2019)
8. Strötgen, J., Gertz, M.: Multilingual and cross-domain temporal tagging. Language Resources and Evaluation **47**(2), 269–298 (2013). https://doi.org/10.1007/s10579-012-9179-y, http://www.jstor.org/stable/42636370
9. Zhuang, L., Wayne, L., Ya, S., Jun, Z.: A robustly optimized BERT pre-training approach with post-training. In: Proceedings of the 20th Chinese National Conference on Computational Linguistics. pp. 1218–1227. Chinese Information Processing Society of China, Huhhot, China (Aug 2021), https://aclanthology.org/2021.ccl-1.108