SlotGAN: Detecting Mentions in Text via Adversarial Distant Learning
Extended Abstract*

Daniel Daza[1,2,3], Michael Cochez[1,3], and Paul Groth[2,3]

[1] Vrije Universiteit Amsterdam
[2] University of Amsterdam
[3] Discovery Lab, Elsevier, The Netherlands
{d.dazacruz,m.cochez}@vu.nl, p.groth@uva.nl

Abstract. Keywords: Information extraction · Mention detection · Distant supervision.

Introduction Detecting mentions of entities in text is a crucial component when solving the task of information extraction. An effective Mention Detection (MD) module is essential for downstream tasks such as named entity recognition [11,13], entity linking [14,1], and relation extraction [3,17]

Recent results on the use of neural networks trained via supervised learning show that the task can be solved successfully, provided enough labeled data [11,15,13]. However, obtaining such data often entails the use of costly domain expertise and a lengthy annotation process.

An alternative that addresses this issue is distant supervision, which encompasses a broad range of methods that in general, create labeled datasets automatically rather than through human annotators [9,5,16,8,10,4]. While distant supervision greatly reduces the cost of generating training data, additional noise is incorporated into the data, for example, in the form of false negatives (mentions of entities that are missed by the automatic annotation procedure, and are labeled as negatives).

SlotGAN Towards methods that reduce annotation costs, and also avoid training with false negatives, we propose SlotGAN: a method based on Generative Adversarial Networks (GANs, [2]) that only requires unlabeled text, and a list of entity names for training (also known as an *gazetteer*). SlotGAN consists of a generator network G that takes as input a sentence, and outputs a set of spans containing words that it deems as named entities. Furthermore, a discriminator D is in charge of determining whether a span of words comes from G, or from the gazetteer (see Fig. 1).

There are a number of challenges arising in SlotGAN that are usually not present in application of GANs. First, the spans extracted by the generator are

* This work was presented at the ACL 2022 Workshop on Structured Prediction for NLP, available at http://dx.doi.org/10.18653/v1/2022.spnlp-1.4.

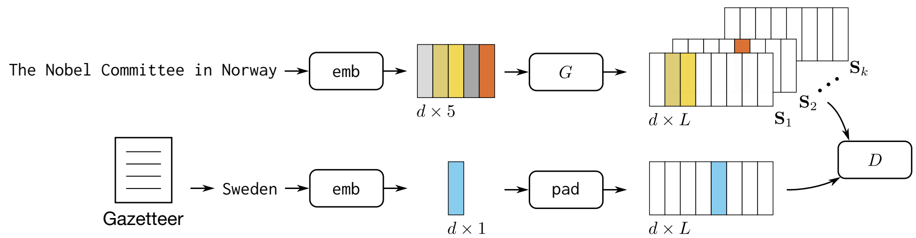

Fig. 1. SlotGAN consists of a generator G that extracts spans from an input sentence, and a discriminator D that determines if a span was generated from G or from a gazetteer.

Table 1. Mention detection results evaluated via exact match precision (P), recall (R), and F1 score; and overlap metrics (preceded with O). The "Data" column indicates what is required to train the model in addition to a corpus.

Method	Data	P	R	F1	OP	OR	OF1
String matching	Gazetteer	76.2	54.0	63.2	57.4	61.3	58.6
ACE [13]	Gold labels	96.0	97.1	96.5	98.3	98.1	98.1
AutoNER [10]	Type dictionary	88.4	94.2	91.2	97.4	97.2	96.9
Unsupervised [7]	Domain concepts	80.0	72.0	76.0	—	—	—
SlotGAN - no pretraining	Gazetteer	55.9	66.1	60.6	82.9	79.5	82.9
SlotGAN - pretrained		60.1	71.1	65.2	93.2	83.0	84.7

discrete selections of words at the input, whose count is not known in advance. We design a mechanism based on Slot Attention [6] to group input words into distinct *slots* that represent a span. The second challenge is that of generating valid spans, which can also be empty if the input sentence does not contain any mentions. To address this, we randomly zero out spans from the gazetteer before passing them to the discriminator, and modify the training objective with a constraint that penalizes spans that are not contiguous.

Experimental results We evaluate the model on the CoNLL dataset for named entity recognition [12], and compare it with a string matching baseline, and supervised, distantly supervised, and unsupervised methods. Results are shown in Table 1. The overlap metrics show that when SlotGAN predicts a span, it overlaps well with gold labels. However, it does not perform well when evaluating exact boundary match. Furthermore, we observe that SlotGAN tends to generate more spans than actually present in the input sentence. This indicates that an additionaly filtering mechanism is likely to improve performance. Directions of future work include an error analysis of the architecture in the fully-supervised setting, and its applications as an end-to-end differentiable architecture for learning to extract spans for information extraction.

Acknowledgements This project was funded by Elsevier's Discovery Lab.

References

1. Cao, N.D., Izacard, G., Riedel, S., Petroni, F.: Autoregressive entity retrieval. In: 9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021. OpenReview.net (2021), https://openreview.net/forum?id=5k8F6UU39V
2. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A.C., Bengio, Y.: Generative adversarial nets. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada. pp. 2672–2680 (2014), https://proceedings.neurips.cc/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html
3. Katiyar, A., Cardie, C.: Going out on a limb: Joint extraction of entity mentions and relations without dependency trees. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). pp. 917–928. Association for Computational Linguistics, Vancouver, Canada (Jul 2017). https://doi.org/10.18653/v1/P17-1085, https://aclanthology.org/P17-1085
4. Li, J., Ding, H., Shang, J., McAuley, J., Feng, Z.: Weakly supervised named entity tagging with learnable logical rules. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). pp. 4568–4581. Association for Computational Linguistics, Online (Aug 2021). https://doi.org/10.18653/v1/2021.acl-long.352, https://aclanthology.org/2021.acl-long.352
5. Lison, P., Barnes, J., Hubin, A., Touileb, S.: Named entity recognition without labelled data: A weak supervision approach. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. pp. 1518–1533. Association for Computational Linguistics, Online (Jul 2020). https://doi.org/10.18653/v1/2020.acl-main.139, https://aclanthology.org/2020.acl-main.139
6. Locatello, F., Weissenborn, D., Unterthiner, T., Mahendran, A., Heigold, G., Uszkoreit, J., Dosovitskiy, A., Kipf, T.: Object-centric learning with slot attention. In: Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., Lin, H. (eds.) Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual (2020), https://proceedings.neurips.cc/paper/2020/hash/8511df98c02ab60aea1b2356c013bc0f-Abstract.html
7. Luo, Y., Zhao, H., Zhan, J.: Named entity recognition only from word embeddings. In: Webber, B., Cohn, T., He, Y., Liu, Y. (eds.) Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020. pp. 8995–9005. Association for Computational Linguistics (2020). https://doi.org/10.18653/v1/2020.emnlp-main.723, https://doi.org/10.18653/v1/2020.emnlp-main.723
8. Ratner, A.J., Sa, C.D., Wu, S., Selsam, D., Ré, C.: Data programming: Creating large training sets, quickly. In: Lee, D.D., Sugiyama, M., von Luxburg, U., Guyon, I., Garnett, R. (eds.) Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain. pp. 3567–3575 (2016), https://proceedings.neurips.cc/paper/2016/hash/6709e8d64a5f47269ed5cea9f625f7ab-Abstract.html
9. Safranchik, E., Luo, S., Bach, S.H.: Weakly supervised sequence tagging from noisy rules. In: The Thirty-Fourth AAAI Conference on Artificial Intelligence,

AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020. pp. 5570–5578. AAAI Press (2020), https://aaai.org/ojs/index.php/AAAI/article/view/6009

10. Shang, J., Liu, L., Gu, X., Ren, X., Ren, T., Han, J.: Learning named entity tagger using domain-specific dictionary. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. pp. 2054–2064. Association for Computational Linguistics, Brussels, Belgium (Oct-Nov 2018). https://doi.org/10.18653/v1/D18-1230, https://aclanthology.org/D18-1230
11. Straková, J., Straka, M., Hajic, J.: Neural architectures for nested NER through linearization. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. pp. 5326–5331. Association for Computational Linguistics, Florence, Italy (Jul 2019). https://doi.org/10.18653/v1/P19-1527, https://aclanthology.org/P19-1527
12. Tjong Kim Sang, E.F., De Meulder, F.: Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition. In: Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003. pp. 142–147 (2003), https://aclanthology.org/W03-0419
13. Wang, X., Jiang, Y., Bach, N., Wang, T., Huang, Z., Huang, F., Tu, K.: Automated concatenation of embeddings for structured prediction. In: Zong, C., Xia, F., Li, W., Navigli, R. (eds.) Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1: Long Papers), Virtual Event, August 1-6, 2021. pp. 2643–2660. Association for Computational Linguistics (2021). https://doi.org/10.18653/v1/2021.acl-long.206, https://doi.org/10.18653/v1/2021.acl-long.206
14. Wu, L., Petroni, F., Josifoski, M., Riedel, S., Zettlemoyer, L.: Scalable zero-shot entity linking with dense entity retrieval. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). pp. 6397–6407. Association for Computational Linguistics, Online (Nov 2020). https://doi.org/10.18653/v1/2020.emnlp-main.519, https://aclanthology.org/2020.emnlp-main.519
15. Yu, J., Bohnet, B., Poesio, M.: Neural mention detection. In: Proceedings of the 12th Language Resources and Evaluation Conference. pp. 1–10. European Language Resources Association, Marseille, France (May 2020), https://aclanthology.org/2020.lrec-1.1
16. Zhao, X., Ding, H., Feng, Z.: GLaRA: Graph-based labeling rule augmentation for weakly supervised named entity recognition. In: Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. pp. 3636–3649. Association for Computational Linguistics, Online (Apr 2021), https://aclanthology.org/2021.eacl-main.318
17. Zhong, Z., Chen, D.: A frustratingly easy approach for entity and relation extraction. In: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. pp. 50–61. Association for Computational Linguistics, Online (Jun 2021). https://doi.org/10.18653/v1/2021.naacl-main.5, https://aclanthology.org/2021.naacl-main.5