International Conference on Artificial Intelligence in Medicine

AIME 2021: **Artificial Intelligence in Medicine** pp 114–123

Transformers for Multi-label Classification of Medical Text: An Empirical Comparison

Vithya Yogarajan, Jacob Montiel, Tony Smith & Bernhard Pfahringer

Conference paper | First Online: 08 June 2021

Part of the Lecture Notes in Computer Science book series (LNAI,volume 12721)

Abstract

Recent advancements in machine learning-based multi-label medical text classification techniques have been used to help enhance healthcare and aid better patient care. This research is motivated by transformers' success in natural language processing tasks, and the opportunity to further improve performance for medical-domain specific tasks by exploiting models pre-trained on health data. We consider transfer learning involving fine-tuning of pre-trained models for predicting medical

codes, formulated as a multi-label problem. We find that domain-specific transformers outperform state-of-the-art results for multi-label problems with the number of labels ranging from 18 to 158, for a fixed sequence length. Additionally, we find that, for longer documents and/or number of labels greater than 300, traditional neural networks still have an edge over transformers. These findings are obtained by performing extensive experiments on the semi-structured eICU data and the free-form MIMIC III data, and applying various transformers including BERT, RoBERTa, and Longformer variations. The electronic health record data used in this research exhibits a high level of label imbalance. Considering individual label accuracy, we find that for eICU data medical-domain specific RoBERTa models achieve improvements for more frequent labels. For infrequent labels, in both datasets, traditional neural networks still perform better.

Keywords

Multi-label **Fine-tuning** **Medical text**

Transformers **Neural networks**

- DOI: 10.1007/978-3-030-77211-6_12
- Chapter length: 10 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

- ISBN: 978-3-030-77211-6
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

- ISBN: 978-3-030-77210-9
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Notes

1. https://github.com/huggingface/transformers.

References

1. Alsentzer, E., et al.: Publicly available clinical BERT embeddings. In: Proceedings of the 2nd Clinical Natural Language Processing Workshop, pp. 72–78 (2019)

2. Amin, S., Neumann, G., Dunfield, K., Vechkaeva, A., Chapman, K.A., Wixted, M.K.: MLT-DFKI at CLEF eHealth 2019: multi-label classification of ICD-10 Codes with BERT. In: CLEF (Working Notes) (2019)

3. Amin-Nejad, A., Ive, J., Velupillai, S.: Exploring transformer text generation for medical dataset augmentation. In: Proceedings of The 12th Language Resources and Evaluation Conference, pp. 4699–4708 (2020)

4. Beltagy, I., Peters, M., Cohan, A.: Longformer: the long-document transformer. arXiv preprint arXiv:2004.05150 (2020)

5. Cho, K., van Merrienboer, B., Bahdanau, D., Bengio, Y.: On the properties of neural machine translation: Encoder-decoder approaches. In: Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation (SSST-8), 2014 (2014)

6. Dai, Z., Yang, Z., Yang, Y., Carbonell, J., Le, Q.V., Salakhutdinov, R.: Transformer-XL: attentive language models beyond a fixed-length context. In: ACL (2019)

7. Devlin, J., Chang, M., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. In: NAACL-HLT (2019)

8. Goldberger, A.L., et al.: PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation **101**(23), e215–e220 (2000)

9. Gu, Y., et al.: Domain-specific language model pretraining for biomedical natural language processing. arXiv preprint arXiv:2007.15779 (2020)

10. Gururangan, S., et al.: Don't stop pretraining: adapt language models to domains and tasks. In: Proceedings of ACL (2020)

11. Johnson, A.E., et al.: MIMIC-III, a freely accessible critical care database. Sci. Data **3**, 160035 (2016)

12. Kim, Y.: Convolutional neural networks for sentence classification. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1746–1751. Association for Computational Linguistics (2014)

13. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. In: International Conference on Learning Representations (ICLR) (2015)

14. Liu, Y., et al.: RoBERTa: a robustly optimized BERT pretraining approach. arXiv preprint arXiv:1907.11692 (2019)

15. Mikolov, T., Grave, E., Bojanowski, P., Puhrsch, C., Joulin, A.: Advances in pre-training distributed word representations. In: Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018) (2018)

16. Moons, E., Khanna, A., Akkasi, A., Moens, M.F.: A comparison of deep learning methods for ICD coding of clinical records. Appl. Sci. **10**(15), 5262 (2020)

17. Mullenbach, J., Wiegreffe, S., Duke, J., Sun, J.,

Eisenstein, J.: Explainable prediction of medical codes from clinical text. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1. ACL: New Orleans, LA, USA (2018)

18. Pollard, T.J., Johnson, A.E.W., Raffa, J.D., Celi, L.A., Mark, R.G., Badawi, O.: The eICU Collaborative Research Database, a freely available multi-center database for critical care research. Sci. Data **5**, 180178 (2018)

19. Sänger, M., Weber, L., Kittner, M., Leser, U.: Classifying german animal experiment summaries with multi-lingual BERT at CLEF eHealth 2019 Task 1. In: CLEF (Working Notes) (2019)

20. Schäfer, H., Friedrich, C.: Multilingual ICD-10 code assignment with transformer architectures using MIMIC-III discharge summaries. In: CLEF 2020 (2020)

21. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems, vol. 30, pp. 5998–6008 (2017)

22. Yogarajan, V., Gouk, H., Smith, T., Mayo, M., Pfahringer, B.: Comparing high dimensional word embeddings trained on medical text to bag-of-words for predicting medical codes. In: Nguyen, N.T., Jearanaitanakij, K., Selamat, A., Trawiński, B., Chittayasothorn, S. (eds.) ACIIDS 2020. LNCS (LNAI), vol. 12033, pp. 97–108. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-41964-6_9

Author information

Authors and Affiliations

Department of Computer Science, University of Waikato, Hamilton, New Zealand

Vithya Yogarajan, Jacob Montiel, Tony Smith & Bernhard Pfahringer

Corresponding author

Correspondence to Vithya Yogarajan .

Editor information

Editors and Affiliations

Brunel University London, Uxbridge, UK

Allan Tucker

University of Coimbra, Coimbra, Portugal

Pedro Henriques Abreu

University of Porto, Porto, Portugal

Jaime Cardoso

University of Porto, Porto, Portugal

Pedro Pereira Rodrigues

Universitat Rovira i Virgili, Tarragona, Spain

David Riaño

Rights and permissions

Copyright information

About this paper

Cite this paper

Yogarajan, V., Montiel, J., Smith, T., Pfahringer, B. (2021). Transformers for Multi-label Classification of Medical Text: An Empirical Comparison. In: Tucker, A., Henriques Abreu, P., Cardoso, J., Pereira Rodrigues, P., Riaño, D. (eds) Artificial Intelligence in Medicine. AIME 2021. Lecture Notes in Computer Science(), vol 12721. Springer, Cham. https://doi.org/10.1007/978-3-030-77211-6_12

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-3-030-77211-6_12

Published	Publisher Name	Print ISBN
08 June 2021	Springer, Cham	978-3-030-77210-9

Online ISBN	eBook Packages
	Computer Science

978-3-030-77211-6 Computer Science (R0)