Journal of Healthcare Engineering 2017 vol.2017

Combination of Deep Recurrent Neural Networks and Conditional Random Fields for Extracting Adverse Drug Reactions from User Reviews

Tutubalina E., Nikolenko S.
Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© 2017 Elena Tutubalina and Sergey Nikolenko. Adverse drug reactions (ADRs) are an essential part of the analysis of drug use, measuring drug use benefits, and making policy decisions. Traditional channels for identifying ADRs are reliable but very slow and only produce a small amount of data. Text reviews, either on specialized web sites or in general-purpose social networks, may lead to a data source of unprecedented size, but identifying ADRs in free-form text is a challenging natural language processing problem. In this work, we propose a novel model for this problem, uniting recurrent neural architectures and conditional random fields. We evaluate our model with a comprehensive experimental study, showing improvements over state-of-the-art methods of ADR extraction.

http://dx.doi.org/10.1155/2017/9451342

References

[1] I. Segura-Bedmar, P. Martínez, R. Revert, and J. Moreno-Schneider, "Exploring Spanish health social media for detecting drug effects," BMC Medical, vol. 15, Supplement 2, p. S6, 2015.

[2] M. S. Boguski, K. D. Mandl, and V. P. Sukhatme, "Repurposing with a difference," Science, vol. 324, no. 5933, pp. 1394-1395, 2009.

[3] P. Avillach, J.-C. Dufour, G. Diallo et al., "Design and validation of an automated method to detect known adverse drug reactions in MEDLINE: A contribution from the EU-ADR project," Journal of the American Medical Informatics Association, vol. 20, no. 3, pp. 446-452, 2013.

[4] Z. S. Miftahutdinov, E. V. Tutubalina, and A. E. Tropsha, Identifying Disease-Related Expressions in Reviews Using Conditional Random Fields, Komp'juternaja Lingvistika i Intellektual'nye Tehnologii, 2017.

[5] A. N. Jagannatha and H. Yu, "Bidirectional RNN for medical event detection in electronic health records," Proceedings of the Conference Association for Computational Linguistics, vol. 2016, pp. 473-482, 2016.

[6] A. Nikfarjam, A. Sarker, K. O'Connor, R. Ginn, and G. Gonzalez, "Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features," Journal of the American Medical Informatics Association, vol. 22, no. 3, pp. 671-681, 2015.

[7] O. Irsoy and C. Cardie, "Opinion mining with deep recurrent neural networks," in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014): A meeting of SIGDAT, a Special Interest Group of the ACL, pp. 720-728, Doha, Qatar, October 2014.

[8] P. Liu, S. R. Joty, and H. M. Meng, "Fine-grained Opinion Mining with Recurrent Neural Networks and Word Embeddings," in Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP 2015), pp. 1433-1443, Lisbon, Portugal, September 2015.

[9] W. Wang, "Mining adverse drug reaction mentions in twitter with word embeddings," in Proceedings of the Social Media Mining Shared Task Workshop at the Pacific Symposium on Biocomputing, 2016.

[10] H.-C. Lee, Y.-Y. Hsu, and H.-Y. Kao, "An enhanced CRFbased system for disease name entity recognition and normalization on BioCreative V DNER task," in Proceedings of the Fifth BioCreative Challenge Evaluation Workshop, pp. 226-233, 2015, http://www.biocreative.org.

[11] Q. Wei, T. Chen, R. Xu, Y. He, and L. Gui, "Disease named entity recognition by combining conditional random fields and bidirectional recurrent neural networks," Database, vol. 2016, 2016.

[12] R. Gareev, M. Tkachenko, and V. Solovyev, Introducing Baselines for Russian Named Entity Recognition, 2013.

[13] V. A. Mozharova and N. V. Loukachevitch, "Combining knowledge and CRF-based approach to named entity recognition in Russian," in Analysis of Images, Social Networks and Texts, pp. 185-195, Springer, Cham, 2016.

[14] Z. Huang, W. Xu, and K. Yu, Bidirectional LSTM-CRF Models for Sequence Tagging, 2015, arXiv preprint arXiv:1508.01991.

[15] J. Chiu and E. Nichols, Named Entity Recognition with Bidirectional LSTM-CNNs, 2015, arXiv preprint arXiv:1511.08308.

[16] X. Ma and E. Hovy, End-to-End Sequence Labeling via Bi-Directional LSTM-CNNs-CRF, 2016, arXiv [cs.LG].

[17] S. Karimi, A. Metke-Jimenez, M. Kemp, and C. Wang, "Cadec: A corpus of adverse drug event annotations," Journal of Biomedical Informatics, vol. 55, pp. 73-81, 2015.

[18] C.-H. Wei, Y. Peng, R. Leaman et al., "Overview of the BioCreative V chemical disease relation (CDR) task," Proceedings of the Fifth BioCreative Challenge Evaluation Workshop, pp. 154-166, 2015, http://www.biocreative.org.

[19] J. Gu, L. Qian, and G. Zhou, "Chemical-induced disease relation extraction with various linguistic features," Database, vol. 2016, 2016.

[20] Z. Zhang, J.-Y. Nie, and X. Zhang, "An ensemble method for binary classification of adverse drug reactions from social media," Proceedings of the Social Media Mining Shared Task Workshop at the Pacific Symposium on Biocomputing.

[21] S. Spangler, A. D. Wilkins, B. J. Bachman et al., "Automated hypothesis generation based on mining scientific literature," Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1877-1886, ACM, New York, NY, USA, 2014.

[22] E. Yom-Tov and E. Gabrilovich, "Postmarket drug surveillance without trial costs: discovery of adverse drug reactions through large-scale analysis of web search queries," Journal of Medical Internet Research, vol. 15, no. 6, article e124, 2013.

[23] C.-C. Huang and Z. Lu, "Community challenges in biomedical text mining over 10 years: success, failure and the future," Briefings in Bioinformatics, vol. 17, no. 1, pp. 132-144, 2016.

[24] A. M. Cohen and W. R. Hersh, "A survey of current work in biomedical text mining," Briefings in Bioinformatics, vol. 6, no. 1, pp. 57-71, 2005.

[25] D. L. Ngo, N. Yamamoto, V. A. Tran et al., "Application of word embedding to drug repositioning," Journal of Biomedical Science and Engineering, vol. 9, p. 7, 2016.

[26] M. Rastegar-Mojarad, H. Liu, and P. Nambisan, "Using social media data to identify potential candidates for drug repurposing: A feasibility study," JMIR Research Protocols, vol. 5, no. 2, article e121, 2016.

[27] S. N. Deftereos, C. Andronis, E. J. Friedla, A. Persidis, and A. Persidis, "Drug repurposing and adverse event prediction using high-throughput literature analysis," Wiley Interdisciplinary Reviews: Systems Biology and Medicine, vol. 3, no. 3, pp. 323-334, 2011.

[28] T. C. Carter and M. M. He, "Challenges of identifying clinically actionable genetic variants for precision medicine," Journal of Healthcare Engineering, vol. 2016, Article ID 3617572, 14 pages, 2016.

[29] A. Helmer, M. Lipprandt, T. Frenken, M. Eichelberg, and A. Hein, "3DLC: A comprehensive model for personal health records supporting new types of medical applications," Journal of Healthcare Engineering, vol. 2, no. 3, pp. 321-336, 2011.

[30] S. Saboor, A. Hoerbst, and E. Ammenwerth, "Quality of electronic health records-coverage of potential information weaknesses by major EHR quality seals," Journal of Healthcare Engineering, vol. 2, no. 3, pp. 365-388, 2011.

[31] J. Bian, U. Topaloglu, and F. Yu, "Towards large-scale twitter mining for drug-related adverse events," Proceedings of the 2012 International Workshop on Smart Health and Wellbeing, pp. 25-32, ACM, New York, NY, USA, 2012.

[32] A. Sarker, A. Nikfarjam, and G. Gonzalez, "Social media mining shared task workshop," Pacific Symposium Biocomputing, vol. 21, pp. 581-592, 2016.

[33] A. Metke-Jimenez and S. Karimi, Concept Extraction to Identify Adverse Drug Reactions in Medical Forums: A Comparison of Algorithms, 2015, arXiv [cs.AI].

[34] C. K. Reddy and C. C. Aggarwal, Healthcare Data Analytics, CRC Press, 2015.

[35] S. Gupta, D. L. MacLean, J. Heer, and C. D. Manning, "Induced lexico-syntactic patterns improve information extraction from online medical forums," Journal of the American Medical Informatics Association, vol. 21, pp. 902-909, 2014.

[36] A. Benton, L. Ungar, S. Hill et al., "Identifying potential adverse effects using the web: A new approach to medical hypothesis generation," Journal of Biomedical Informatics, vol. 44, pp. 989-996, 2011.

[37] R. Feldman, O. Netzer, A. Peretz, and B. Rosenfeld, "Utilizing text mining on online medical forums to predict label change due to adverse drug reactions," Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1779-1788, ACM, New York, NY, USA, 2015.

[38] S. Karimi, C. Wang, A. Metke-Jimenez, R. Gaire, and C. Paris, "Text and data mining techniques in adverse drug reaction detection," ACM Computing Surveys, vol. 47, pp. 56:1-56:39, 2015.

[39] A. Sarker, R. Ginn, A. Nikfarjam et al., "Utilizing social media data for pharmacovigilance: A review," Journal of Biomedical Informatics, vol. 54, pp. 202-212, 2015.

[40] V. Solovyev and V. Ivanov, "Knowledge-driven event extraction in Russian: corpus-based linguistic resources," Computational intelligence and neuroscience, vol. 2016, Article ID 4183760, 11 pages, 2016.

[41] R. Leaman, L. Wojtulewicz, R. Sullivan, A. Skariah, J. Yang, and G. Gonzalez, "Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks," Proceedings of the 2010 Workshop on Biomedical Natural Language Processing, pp. 117-125, Association for Computational Linguistics, Stroudsburg, PA, USA, 2010.

[42] C. C. Yang, H. Yang, L. Jiang, and M. Zhang, "Social media mining for drug safety signal detection," Proceedings of the 2012 International Workshop on Smart Health and Wellbeing, pp. 33-40, ACM, New York, NY, USA, 2012.

[43] S. Gupta and C. D. Manning, Improved Pattern Learning for Bootstrapped Entity Extraction, 2014, CoNLL.

[44] G. Stanovsky, D. Gruhl, and P. N. Mendes, Recognizing Mentions of Adverse Drug Reaction in Social Media Using Knowledge-Infused Recurrent Models, 2017, EACL.

[45] J. L. Elman, "Finding structure in time," Cognitive Science, vol. 14, pp. 179-211, 1990.

[46] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," IEEE Transactions on Neural Network, vol. 5, no. 2, pp. 157-166, 1994.

[47] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," IEEE Transactions on Neural Networks and Learning Systems, 2016.

[48] K. Cho, B. van Merrienboer, C. Gulcehre et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1724-1734, Association for Computational Linguistics, Stroudsburg, PA, USA, 2014.

[49] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," IEEE Transactions on Signal Processing, vol. 45, pp. 2673-2681, 1997.

[50] A. Graves, S. Fernández, and J. Schmidhuber, "Bidirectional LSTM networks for improved phoneme classification and recognition," Artificial Neural Networks: Formal Models and Their Applications-ICANN 2005, pp. 799-804, Springer, Berlin, Heidelberg, 2005.

[51] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM networks," Proceedings 2005 IEEE International Joint Conference on Neural Networks, vol. 4, pp. 2047-2052, 2005.

[52] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: probabilistic models for segmenting and labeling sequence data," Proceedings of the Eighteenth.

[53] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," Journal of Machine Learning Research, vol. 15, pp. 1929-1958, 2014.

[54] J. Wiebe, T. Wilson, and C. Cardie, "Annotating expressions of opinions and emotions in language," Language Resources and Evaluation, vol. 39, pp. 165-210, 2006.

[55] O. Irsoy and C. Cardie, "Opinion mining with deep recurrent neural networks," Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 720-728, Association for Computational Linguistics, Stroudsburg, PA, USA, 2014.

[56] N. Loukachevitch, P. Blinov, E. Kotelnikov, Y. Rubtsova, V. Ivanov, and E. Tutubalina, "SentiRuEval: testing objectoriented sentiment analysis systems in Russian," Proceedings of International Conference Dialog, pp. 3-13, 2015.

[57] E. F. Tjong Kim Sang and F. De Meulder, "Introduction to the CoNLL-2003 shared task: language-independent named entity recognition," Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003-Volume 4, pp. 142-147, Association for Computational Linguistics, Stroudsburg, PA, USA, 2003.

[58] D. Kinga and J. B. Adam, "A method for stochastic optimization," International Conference on Learning Representations (ICLR).

[59] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on imagenet classification," in Proceedings of the IEEE International Conference on Computer Vision, pp. 1026-1034, Santiago, Chile, 2015.

[60] E. Tutubalina and S. Nikolenko, "Automated prediction of demographic information from medical user reviews," International Conference on Mining Intelligence and Knowledge Exploration, pp. 1-11, Springer, 2016.