Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2018 vol.10716 LNCS, pages 3-15

Automated detection of adverse drug reactions from social media posts with machine learning

Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© Springer International Publishing AG 2018. Adverse drug reactions can have serious consequences for patients. Social media is a source of information useful for detecting previously unknown side effects from a drug since users publish valuable information about various aspects of their lives, including health care. Therefore, detection of adverse drug reactions from social media becomes one of the actual tools for pharmacovigilance. In this paper, we focus on identification of adverse drug reactions from user reviews and formulate this problem as a binary classification task. We developed a machine learning classifier with a set of features for resolving this problem. Our feature-rich classifier achieves significant improvements on a benchmark dataset over baseline approaches and convolutional neural networks.

http://dx.doi.org/10.1007/978-3-319-73013-4_1

Keywords

Adverse drug reactions, Deep learning, Health social media analytics, Machine learning, Text mining

References

[1] Onakpoya, I.J., Heneghan, C.J., Aronson, J.K.: Post-marketing withdrawal of 462 medicinal products because of adverse drug reactions: a systematic review of the world literature. BMC Med. 14(1), 10 (2016)

[2] Pirmohamed, M., James, S., Meakin, S., Green, C., Scott, A.K., Walley, T.J., Farrar, K., Park, B.K., Breckenridge, A.M.: Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ 329(7456), 15–19 (2004)

[3] Classen, D.C., Pestotnik, S.L., Evans, R.S., Lloyd, J.F., Burke, J.P.: Adverse drug events in hospitalized patients: excess length of stay, extra costs, and attributable mortality. JAMA 277(4), 301–306 (1997)

[4] Lazarou, J., Pomeranz, B.H., Corey, P.N.: Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 279(15), 1200–1205 (1998)

[5] Bates, D.W., Cullen, D.J., Laird, N., Petersen, L.A., Small, S.D., Servi, D., Laffel, G., Sweitzer, B.J., Shea, B.F., Hallisey, R., et al.: Incidence of adverse drug events and potential adverse drug events: implications for prevention. JAMA 274(1), 29– 34 (1995)

[6] Sloane, R., Osanlou, O., Lewis, D., Bollegala, D., Maskell, S., Pirmohamed, M.: Social media and pharmacovigilance: a review of the opportunities and challenges. Br. J. Clin. Pharmacol. 80(4), 910–920 (2015)

[7] Tutubalina, E., Nikolenko, S.: Automated prediction of demographic information from medical user reviews. In: Prasath, R., Gelbukh, A. (eds.) MIKE 2016. LNCS, vol. 10089, pp. 174–184. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58130-9 17

[8] Solovyev, V., Ivanov, V.: Knowledge-driven event extraction in Russian: corpus-based linguistic resources. Comput. Intell. Neurosci. 2016, 16 (2016)

[9] Sayfullina, L., Eirola, E., Komashinsky, D., Palumbo, P., Karhunen, J.: Android malware detection: building useful representations. In: 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 201–206, December 2016

[10] Ivanov, V., Tutubalina, E., Mingazov, N., Alimova, I.: Extracting aspects, sentiment and categories of aspects in user reviews about restaurants and cars. In: Proceedings of International Conference Dialog, vol. 2, pp. 22–34 (2015)

[11] Murff, H.J., Patel, V.L., Hripcsak, G., Bates, D.W.: Detecting adverse events for patient safety research: a review of current methodologies. J. Biomed. Inform. 36(1), 131–143 (2003)

[12] Sarker, A., Ginn, R., Nikfarjam, A., O'Connor, K., Smith, K., Jayaraman, S., Upadhaya, T., Gonzalez, G.: Utilizing social media data for pharmacovigilance: a review. J. Biomed. Inform. 54, 202–212 (2015)

[13] Lardon, J., Abdellaoui, R., Bellet, F., Asfari, H., Souvignet, J., Texier, N., Jaulent, M.C., Beyens, M.N., Burgun, A., Bousquet, C.: Adverse drug reaction identification and extraction in social media: a scoping review. J. Med. Internet Res. 17(7), e171 (2015)

[14] Harpaz, R., Callahan, A., Tamang, S., Low, Y., Odgers, D., Finlayson, S., Jung, K., LePendu, P., Shah, N.H.: Text mining for adverse drug events: the promise, challenges, and state of the art. Drug Saf. 37(10), 777–790 (2014)

[15] Harpaz, R., DuMouchel, W., Shah, N.H., Madigan, D., Ryan, P., Friedman, C.: Novel data-mining methodologies for adverse drug event discovery and analysis. Clin. Pharmacol. Ther. 91(6), 1010–1021 (2012)

[16] Sarker, A., Gonzalez, G.: Portable automatic text classification for adverse drug reaction detection via multi-corpus training. J. Biomed. Inform. 53, 196–207 (2015)

[17] Karimi, S., Metke-Jimenez, A., Kemp, M., Wang, C.: Cadec: a corpus of adverse drug event annotations. J. Biomed. Inform. 55, 73–81 (2015)

[18] Kim, Y.: Convolutional neural networks for sentence classification. arXiv preprint arXiv:1408.5882 (2014)

[19] Sarker, A., Nikfarjam, A., Gonzalez, G.: Social media mining shared task workshop. In: Proceedings of the Pacific Symposium on Biocomputing, pp. 581–592 (2016)

[20] Rastegar-Mojarad, M., Komandur Elayavilli, R., Yu, Y., Hiu, H.: Detecting signals in noisy data-can ensemble classifiers help identify adverse drug reaction in tweets. In: Proceedings of the Social Media Mining Shared Task Workshop at the Pacific Symposium on Biocomputing (2016)

[21] Zhang, Z., Nie, J., Zhang, X.: An ensemble method for binary classification of adverse drug reactions from social media. In: Proceedings of the Social Media Mining Shared Task Workshop at the Pacific Symposium on Biocomputing (2016)

[22] Ofoghi, B., Siddiqui, S., Verspoor, K.: Read-BioMed-SS: adverse drug reaction classification of microblogs using emotional and conceptual enrichment. In: Proceedings of the Social Media Mining Shared Task Workshop at the Pacific Symposium on Biocomputing (2016)

[23] Jonnagaddala, J., Jue, T.R., Dai, H.: Binary classification of twitter posts for adverse drug reactions. In: Proceedings of the Social Media Mining Shared Task Workshop at the Pacific Symposium on Biocomputing, pp. 4–8 (2016)

[24] Egger, D., Uzdilli, F., Cieliebak, M., Derczynski, L.: Adverse drug reaction detection using an adapted sentiment classifier. In: Proceedings of the Social Media Mining Shared Task Workshop at the Pacific Symposium on Biocomputing (2016)

[25] Ginn, R., Pimpalkhute, P., Nikfarjam, A., Patki, A., O'Connor, K., Sarker, A., Smith, K., Gonzalez, G.: Mining twitter for adverse drug reaction mentions: a corpus and classification benchmark. In: Proceedings of the Fourth Workshop on Building and Evaluating Resources for Health and Biomedical Text Processing. Citeseer (2014)

[26] Yang, M., Wang, X., Kiang, M.Y.: Identification of consumer adverse drug reaction messages on social media. In: PACIS, vol. 193 (2013)

[27] Bian, J., Topaloglu, U., Yu, F.: Towards large-scale twitter mining for drug-related adverse events. In: Proceedings of the 2012 International Workshop on Smart Health and Wellbeing, pp. 25–32. ACM (2012)

[28] Patki, A., Sarker, A., Pimpalkhute, P., Nikfarjam, A., Ginn, R., O'Connor, K., Smith, K., Gonzalez, G.: Mining adverse drug reaction signals from social media: going beyond extraction. In: Proceedings of BioLinkSig 2014, pp. 1–8 (2014)

[29] Gurulingappa, H., Mateen-Rajpu, A., Toldo, L.: Extraction of potential adverse drug events from medical case reports. J. Biomed. Semant. 3(1), 15 (2012)

[30] Liu, X., Liu, J., Chen, H.: Identifying adverse drug events from health social media: a case study on heart disease discussion forums. In: Zheng, X., Zeng, D., Chen, H., Zhang, Y., Xing, C., Neill, D.B. (eds.) ICSH 2014. LNCS, vol. 8549, pp. 25–36. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08416-9 3

[31] Huynh, T., He, Y., Willis, A., Rüger, S.: Adverse drug reaction classification with deep neural networks. In: COLING (2016)

[32] Gurulingappa, H., Rajput, A.M., Roberts, A., Fluck, J., Hofmann-Apitius, M., Toldo, L.: Development of a benchmark corpus to support the automatic extraction of drug-related adverse effects from medical case reports. J. Biomed. Inform. 45(5), 885–892 (2012)

[33] Nikfarjam, A., Sarker, A., O'Connor, K., Ginn, R., Gonzalez, G.: Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. J. Am. Med. Inform. Assoc. 22(3), 671–681 (2015)

[34] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al.: Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830 (2011)

[35] Kiritchenko, S., Zhu, X., Mohammad, S.M.: Sentiment analysis of short informal texts. J. Artif. Intell. Res. 50, 723–762 (2014)

[36] Baccianella, S., Esuli, A., Sebastiani, F.: Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: LREC, vol. 10, pp. 2200–2204 (2010)

[37] Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, pp. 347–354. Association for Computational Linguistics (2005)

[38] Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 168–177. ACM (2004)

[39] Miftahutdinov, Z., Tutubalina, E., Tropsha, A.: Identifying disease-related expressions in reviews using conditional random fields. Komp'juternaja Lingvistika i Intellektual'nye Tehnologii 1(16), 155–166 (2017)

[40] Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15(1), 1929–1958 (2014)

[41] Kingma, D., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

[42] He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1026–1034 (2015)

[43] Chollet, F., et al.: Keras (2015). https://github.com/fchollet/keras