Molecular Pharmaceutics 2017 vol.14 N9, pages 3098-3104

DruGAN: An Advanced Generative Adversarial Autoencoder Model for de Novo Generation of New Molecules with Desired Molecular Properties in Silico

Kadurin A., Nikolenko S., Khrabrov K., Aliper A., Zhavoronkov A.
Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© 2017 American Chemical Society. Deep generative adversarial networks (GANs) are the emerging technology in drug discovery and biomarker development. In our recent work, we demonstrated a proof-of-concept of implementing deep generative adversarial autoencoder (AAE) to identify new molecular fingerprints with predefined anticancer properties. Another popular generative model is the variational autoencoder (VAE), which is based on deep neural architectures. In this work, we developed an advanced AAE model for molecular feature extraction problems, and demonstrated its advantages compared to VAE in terms of (a) adjustability in generating molecular fingerprints; (b) capacity of processing very large molecular data sets; and (c) efficiency in unsupervised pretraining for regression model. Our results suggest that the proposed AAE model significantly enhances the capacity and efficiency of development of the new molecules with specific anticancer properties using the deep generative models.

http://dx.doi.org/10.1021/acs.molpharmaceut.7b00346

Keywords

adversarial autoencoder, deep learning, drug discovery, generative adversarial network, variational autoencoder

References

[1] Munos, B. H.; Chin, W. W. How to revive breakthrough innovation in the pharmaceutical industry Sci. Transl. Med. 2011, 3 (89) 89cm16 10.1126/scitranslmed.3002273

[2] Mignani, S.; Huber, S.; Tomas, H.; Rodrigues, J.; Majoral, J. P. Why and how have drug discovery strategies in pharma changed? What are the new mindsets? Drug Discovery Today 2016, 21 (2) 239-249 10.1016/j.drudis.2015.09.007

[3] Yu, M. J. Druggable chemical space and enumerative combinatorics J. Cheminf. 2013, 5 (1) 19 10.1186/1758-2946-5-19

[4] Vanhaelen, Q.; Mamoshina, P.; Aliper, A. M.; Artemov, A.; Lezhnina, K.; Ozerov, I.; Labat, I.; Zhavoronkov, A. Design of efficient computational workflows for in silico drug repurposing Drug Discovery Today 2017, 22 (2) 210-222 10.1016/j.drudis.2016.09.019

[5] LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning Nature 2015, 521 (7553) 436-444 10.1038/nature14539

[6] Mamoshina, P.; Vieira, A.; Putin, E.; Zhavoronkov, A. Applications of Deep Learning in Biomedicine Mol. Pharmaceutics 2016, 13 (5) 1445-1454 10.1021/acs.molpharmaceut.5b00982

[7] Putin, E.; Mamoshina, P.; Aliper, A.; Korzinkin, M.; Moskalev, A.; Kolosov, A.; Ostrovskiy, A.; Cantor, C.; Vijg, J.; Zhavoronkov, A. Deep biomarkers of human aging: Application of deep neural networks to biomarker development Aging 2016, 8 (5) 1021-1033 10.18632/aging.100968

[8] Ozerov, IV; Lezhnina, K. V.; Izumchenko, E.; Artemov, A. V.; Medintsev, S.; Vanhaelen, Q.; Aliper, A.; Vijg, J.; Osipov, A. N.; Labat, I.; West, M. D.; Buzdin, A.; Cantor, C. R.; Nikolsky, Y.; Borisov, N.; Irincheeva, I. et al. In silico Pathway Activation Network Decomposition Analysis (iPANDA) as a method for biomarker development Nat. Commun. 2016, 7, 13427 10.1038/ncomms13427

[9] Gawehn, E.; Hiss, J. A.; Schneider, G. Deep Learning in Drug Discovery Mol. Inf. 2016, 35 (1) 3-14 10.1002/minf.201501008

[10] Aliper, A.; Plis, S.; Artemov, A.; Ulloa, A.; Mamoshina, P.; Zhavoronkov, A. Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data Mol. Pharmaceutics 2016, 13 (7) 2524-2530 10.1021/acs.molpharmaceut.6b00248

[11] Artemov, A. V.; Putin, E.; Vanhaelen, Q.; Aliper, A.; Ozerov; Zhavoronkov, A. Integrated deep learned transcriptomic and structure-based predictor of clinical trials outcomes bioRxiv 2016, 10.1101/095653

[12] Kadurin, A.; Aliper, A.; Kazennov, A.; Mamoshina, P.; Vanhaelen, Q.; Khrabrov, K.; Zhavoronkov, A. The cornucopia of meaningful leads: Applying deep adversarial autoencoders for new molecule development in oncology Oncotarget 2017, 8 (7) 10883-10890 10.18632/oncotarget.14073

[13] Hinton, G. E.; Osindero, S.; Teh, Y. W. A fast learning algorithm for deep belief nets Neural computation 2006, 18 (7) 1527-1554 10.1162/neco.2006.18.7.1527

[14] Salakhutdinov, R.; Hinton, G. Deep Boltzmann Machines PMLR 2009, 5, 448-455

[15] Salakhutdinov, R. Learning Deep Generative Models Annu. Rev. Stat. Its Appl. 2015, 2, 361-385 10.1146/annurev-statistics-010814-020120

[16] Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting J. Mach. Learn. Res. 2014, 15, 1929-1958

[17] Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift PMLR 2015, 37, 448-456

[18] Kingma, D. P.; Ba, J. L. Adam: A method for stochastic optimization arXiv 2014, 1412.6980

[19] Duchi, J.; Hazan, E.; Singer, Y. Adaptive subgradient methods for online learning and stochastic optimization J. Mach. Learn. Res. 2011, 12, 2121-2159

[20] Radford, A.; Metz, L.; Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks arXiv 2015, 1511.06434

[21] Rezende, D. J.; Mohamed, S.; Danihelka, I.; Gregor, K.; Wierstra, D. One-shot generalization in deep generative models arXiv 2016, 1603.05106

[22] Kingma, D. P.; Welling, M. Auto-encoding variational Bayes arXiv 2013, 1312.6114

[23] Doersch, C. Tutorial on variational autoencoders arXiv 2016, 1606.05908

[24] Rezende, D. J.; Mohamed, S.; Wierstra, D. Stochastic backpropagation and approximate inference in deep generative models arXiv 2014, 1401.4082

[25] Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets Adv. Neural Inf. Process. Syst. 2014, 2672-2680

[26] Salimans, T.; Goodfellow, I.; Zaremba, W.; Cheung, V.; Radford, A.; Chen, X. Improved techniques for training GANs Adv. Neural Inf Process. Syst. 2016, 2226-2234

[27] Makhzani, A.; Shlens, J.; Jaitly, N.; Goodfellow, I.; Frey, B. Adversarial autoencoders arXiv 2015, 1511.05644

[28] Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S. H. PubChem Substance and Compound databases Nucleic Acids Res. 2016, 44 (D1) D1202-1213 10.1093/nar/gkv951

[29] Delaney, J. S. ESOL: estimating aqueous solubility directly from molecular structure Journal of chemical information and computer sciences 2004, 44 (3) 1000-1005 10.1021/ci034243x

[30] Savjani, K. T.; Gajjar, A. K.; Savjani, J. K. Drug solubility: importance and enhancement techniques ISRN Pharm. 2012, 2012, 195727 10.5402/2012/195727

[31] Goodfellow, I. NIPS 2016 Tutorial: Generative Adversarial Networks arXiv 2016, 1701.00160