[L14] Deep Learning: What Makes Neural Networks Great Again?

Igor I. Baskin

Group of Chemoinformatics and Materials Informatics, Faculty of Physics, Lomonosov Moscow State University, Leninskie Gory 1/2, 119991, Moscow, Russia

This lecture surveys the application of deep learning in chemoinformatics (see reviews [1-4]). It starts with the consideration of classical “shallow” neural networks and their applications in chemoinformatics (see reviews [5-6]). Then, the major drawbacks of the “shallow” neural networks arising from the problem of vanishing gradients are mentioned, and several ways to solve this problem are discussed. Thanks to them, it has become possible to train neural networks with multiple hidden layers of neurons - deep neural networks, which have attained human and even superhuman performance in several important domains (image and speech recognition, natural language processing, games, *etc*) and revolutionized the whole field of artificial intelligence.

Deep learning can be defined as the application of artificial neural networks with multiple hidden layers that form multiple levels of representations corresponding to different levels of abstraction (see [7-9]). This definition is discussed in the lecture in the context of chemoinformatics applications. The importance of representation learning is underlined. Several benchmarking studies comparing deep *vs* shallow learning for building QSAR models are also considered.

After that, several important techniques rooted in deep learning are discussed in the context of their application in chemoinformatics: auto-encoders, convolutional neural networks, recurrent neural networks, generative adversarial networks, and deep reinforcement learning. Where possible, the connection with earlier studies (e.g., [10-12]) is discussed.

The last part of the lecture is devoted to the use of neural networks to generate new chemical structures with desired activities/properties, including the ability to perform ligand-based *de novo* drug design. A novel approach based on combining auto-encoders based on recurrent neural networks with the generative topographic mapping (GTM) technique [13-15] is discussed, and its advantages over the use of variational auto-encoders or stand-alone recurrent neural networks for the same purpose are explained.

Bibliography:

[1]. I.I.Baskin, D.Winkler, I.V.Tetko. Expert Opin. Drug Discov. 11 (2016) 785-795.
[2]. E.Gawehn, J.A.Hiss, G.Schneider. Mol. Inf. 35 (2016) 3-14.
[3]. G.B.Goh, N.O.Hodas, A.Vishnu. J. Comp. Chem. 38 (2017) 1291-1307.
[4]. H.Chen, et al. Drug Discov. Today (2018) in press
[5]. N.Halberstam, et al. Russ. Chem. Rev. 72 (2003) 629-649.
[6]. I.I.Baskin, V.A.Palyulin, N.S.Zefirov. Methods Mol. Biol. 458 (2008) 137-158.
[7]. G.E.Hinton, R.R.Salakhutdinov. Science 313 (2006) 504-507.
[8]. Y.LeCun, Y.Bengio, G.Hinton. Nature 521 (2015) 436-444.
[9]. I.Goodfellow, Y.Bengio, A.Courville. Deep Learning. MIT Press, 2016.
[10]. I.I.Baskin, V.A.Palyulin, N.S.Zefirov. J. Chem. Inf. Comput. Sci. 37 (1997) 715-721.
[11]. A.Varnek, et al. J. Chem. Inf. Model. 49 (2009) 133-144.
[12]. P.V.Karpov, et al. Bioorg. Med. Chem. Lett. 21 (2011) 6728-6731.
[13]. C.M.Bishop, M.Svensen, C.K.I.Williams. Neural Comput. 10 (1998) 215-234.
[14]. N.Kireeva, et al. Mol. Inf. 31 (2012) 301-312.
[15]. H.A.Gaspar, et al. J. Chem. Inf. Model. 55 (2015) 2403-2410.