SciForum
MOL2NET

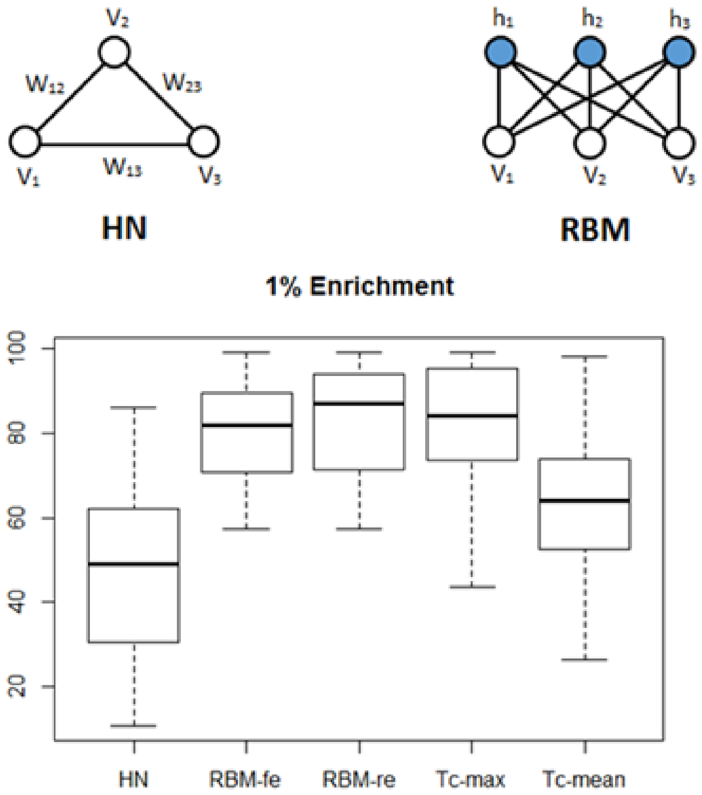
The Use of Energy-Based Neural Networks for Similarity-Based Virtual Screening

Igor I. Baskin *, Nelly I. Zhokhova
Faculty of Physics, M.V.Lomonosov Moscow State University, Moscow 119991, Russia
***** Author to whom correspondence should be addressed; E-Mail: igbaskin@gmail.com

Abstract: Energy-based learning [1] is a common framework for building models in which dependencies between variables are captured by means of a scalar function conventionally called “energy”. The energy-based learning is implemented mostly in recurrent neural networks with symmetric connections between neurons. In this study for the first time, the energy-based neural networks were applied to build structure-activity models. The Hopfield Networks (HN) [2] and the Restricted Boltzmann Machines (RBM) [3,4] were used to build one-class classification models for conducting similarity-based virtual screening [5,6]. The AUC (Area Under Curve) score for ROC curves and 1%-enrichment rates were compared for 20 targets taken from DUD repository. Five different scores were used to assess similarity between each the tested compounds and the training sets of active compounds: the mean and the maximum values of Tanimoto coefficients (Tc-mean and Tc-max, respectively), the energy for Hopfield Networks (HN), the free energy and the reconstruction error for Restricted Boltzmann Machines (RBM-fe and RBM-rec, respectively). The latter score was shown to provide the superior mean predictive performance. Additional advantages of using energy-based neural networks for similarity-based virtual screening over the state-of-the-art similarity searching based on Tanimoto coefficients are: computational efficacy and scalability of prediction procedures, the ability to implicitly reweight structural features and consider their interactions, their “creativity” and compatibility with modern deep learning and artificial intelligence techniques (see reviews in the use of neural networks and deep learning in drug discovery [7-10]).
.
.
.

Keywords: neural networks, Hopfield nets, Restricted Boltzmann Machines, similarity searching, virtual screening, one-class classification

Graphical Abstract:



References:

1. LeCun, Y.; Chopra, S.; Hadsell, R.; Ranzato, M.A.; Huang, F.J. Energy-based models. In *Predicting structure data*, Bakir, G.; Hofmanb, T.; Schoelkopf, B.; Smola, A.J.; Taskar, B.; Vishwanathan, S.V.N., Eds. The MIT Press: Cambridge, Massachusetts, 2007; pp 191-246.
2. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 2554-2558.
3. Smolensky, P. Information processing in dynamical systems: Foundations of harmony theory. In *Parallel distributed processing*, Rumelhart, D.E.; McClelland, J.L., Eds. MIT Press: Cambridge, 1986; Vol. 1, pp 194-281.
4. Hinton, G.E. A practical guide to training restricted boltzmann machines. In *Neural networks: Tricks of the trade*, Springer: 2012; pp 599-619.
5. Baskin, I.I.; Kireeva, N.; Varnek, A. The one-class classification approach to data description and to models applicability domain. *Mol. Inf.* **2010**, *29*, 581-587.
6. Karpov, P.V.; Osolodkin, D.I.; Baskin, I.I.; Palyulin, V.A.; Zefirov, N.S. One-class classification as a novel method of ligand-based virtual screening: The case of glycogen synthase kinase 3oi inhibitors. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 6728-6731.
7. Halberstam, N.M.; Baskin, I.I.; Palyulin, V.A.; Zefirov, N.S. Neural networks as a method for elucidating structure-property relationships for organic compounds. *Russ. Chem. Rev.* **2003**, *72*, 629-649.
8. Baskin, I.I.; Palyulin, V.A.; Zefirov, N.S. Chapter 8. Neural networks in building qsar models. In *Artificial neural networks: Methods and protocols*, Livingstone, D.S., Ed. Humana Press, a part of Springer Science + Business Media: 2008; pp 139-160.
9. Baskin, I.I.; Winkler, D.; Tetko, I.V. A renaissance of neural networks in drug discovery. *Expert opinion on drug discovery* **2016**, *11*, 785-795.
10. Gawehn, E.; Hiss, J.A.; Schneider, G. Deep learning in drug discovery. *Mol. Inf.* **2016**, *35*, 3-14.