The 4th Iranian Conference on
Systems Biology

Deep neural network in QSAR studies using deep belief network

--------one line distance ----------

Fahimeh Ghasemi1*, Alireza Mehridehnavi1

1 Department of Bioinformatics and Systems Biology, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Hezar-Jerib Ave., Isfahan, IR Iran, 81746 73461.

**Correspondence to: Fahimeh Ghasemi (f_ghasemi@amt.mui.ac.ir)*

Abstract

There are two major challenges in the current high throughput screening drug design: the large number of descriptors which may also have autocorrelations and, proper parameter initialization in model prediction to avoid over-fitting problem. Deep architecture structures have been recommended to predict the compounds biological activity. In this paper, the mean and variance of squared correlation for the proposed model and deep neural network are 0.618 ±0.407e-4 and 0.485±4.82e-4, respectively. The outputs of this model seem to outperform those of the models obtained from deep neural network.

Keywords. Deep belief network, QSAR, Drug design

Introduction

Deep architecture structures have been recommended to predict the compounds biological activity. Performance of deep neural network is not always acceptable in QSAR studies. This study tries to find a solution to this problem focusing on primary parameter computation. Deep belief network has been getting popular as a deep neural network model generation method in other fields such as image processing.

Materials and Methods

Data Sets. All fifteen targets of the Kaggle database competition held by Merck sponsor (2012) were chosen as an input network to consider the proposed network efficacy (www.kaggle.com)

Method. Deep belief network is one of the classes of deep generative models composed of l stacks of restricted Boltzmann machine. The main aim of DBN is the weight initialization of a deep neural network to produce optimum model in comparison to the model by random weights. This approach makes the predictions extremely effective. Alternatively, DBN can be effectively used to perform layer by layer pre-training intended to initialize training of a back propagation algorithm.

Results

In order to assess the performance of the obtained model, the most common standard metrics, correlation coefficient (R), was used with regard to the predicted and experimental activities in test sets. To evaluate the effect of initialization on learning network parameters, the results of DNN and some of the shallow learning algorithms, MLR, RF and ANN are reported. (Table 1)

Table 1. R2 for MLR, RF, ANN, DNN and DBN-DNN.

Method	MLR		RF		ANN		DNN		DBN-DNN	
Index	R2 ± std	F-test	R2 ± std	F-test	R2 ± std	F-test	R2 ± std	F-test	R2 ± std	F-test
Mean	0.039±0.0343	0.346	0.558±1.7e-3	1.56	0.322±0.038	0.890	0.485±4.8e-4	1.33	0.618±4.1e-4	1.69

Discussion, Conclusion and Suggestions

The main purpose of this research was to examine the proper parameter initialization in deep neural networks using deep belief networks. In order to come across with the best decision in choosing some key parameters,

To investigate the efficiency of the proposed model, the results were calculated through three different metrics and finally compared with some of the shallow learning algorithm, MLR, RF, ANN and DNN with random initial parameters. The results indicated that DBN-DNN model performed well, and the mean correlation of all data sets were greater than mentioned algorithms while the range of the changes was smaller than others.

The results revealed that an optimization in initialization will improve DNN potential to provide high quality predicting models. In fact, time saving and possibility of replacing high performance computer or cluster by a mediocre personal computer proved that there would be no need to use a very complex network after changing the weights and biases, whereas the output of such model demonstrates significant superiority over the output of the deep neural network.

References

Ghasemi, F, AR Mehridehnavi, A Fassihi and H Pérez-Sánchez. "Deep Neural Network in Biological Activity Prediction Using Deep Belief Network." Applied Soft Computing, (2018).