

RESEARCH

Open Access



# Development of models for classification of action between heat-clearing herbs and blood-activating stasis-resolving herbs based on theory of traditional Chinese medicine

Zhao Chen<sup>1,2</sup> , Yanfeng Cao<sup>1,2</sup>, Shuaibing He<sup>1</sup> and Yanjiang Qiao<sup>1,2\*</sup>

## Abstract

**Background:** Action (“*gongxiao*” in Chinese) of traditional Chinese medicine (TCM) is the high recapitulation for therapeutic and health-preserving effects under the guidance of TCM theory. TCM-defined herbal properties (“*yaoxing*” in Chinese) had been used in this research. TCM herbal property (TCM-HP) is the high generalization and summary for actions, both of which come from long-term effective clinical practice in two thousands of years in China. However, the specific relationship between TCM-HP and action of TCM is complex and unclear from a scientific perspective. The research about this is conducive to expound the connotation of TCM-HP theory and is of important significance for the development of the TCM-HP theory.

**Methods:** One hundred and thirty-three herbs including 88 heat-clearing herbs (HCHs) and 45 blood-activating stasis-resolving herbs (BAHRHs) were collected from reputable TCM literatures, and their corresponding TCM-HPs/ actions information were collected from Chinese pharmacopoeia (2015 edition). The Kennard–Stone (K–S) algorithm was used to split 133 herbs into 100 calibration samples and 33 validation samples. Then, machine learning methods including supported vector machine (SVM), k-nearest neighbor (kNN) and deep learning methods including deep belief network (DBN), convolutional neural network (CNN) were adopted to develop action classification models based on TCM-HP theory, respectively. In order to ensure robustness, these four classification methods were evaluated by using the method of tenfold cross validation and 20 external validation samples for prediction.

**Results:** As results, 72.7–100% of 33 validation samples including 17 HCHs and 16 BASRHs were correctly predicted by these four types of methods. Both of the DBN and CNN methods gave out the best results and their sensitivity, specificity, precision, accuracy were all 100.00%. Especially, the predicted results of external validation set showed that the performance of deep learning methods (DBN, CNN) were better than traditional machine learning methods (kNN, SVM) in terms of their sensitivity, specificity, precision, accuracy. Moreover, the distribution patterns of TCM-HPs of HCHs and BASRHs were also analyzed to detect the featured TCM-HPs of these two types of herbs. The result showed that the featured TCM-HPs of HCHs were cold, bitter, liver and stomach meridians entered, while those of BASRHs were warm, bitter and pungent, liver meridian entered.

**Conclusions:** The performance on validation set and external validation set of deep learning methods (DBN, CNN) were better than machine learning models (kNN, SVM) in sensitivity, specificity, precision, accuracy when predicting the actions of heat-clearing and blood-activating stasis-resolving based on TCM-HP theory. The deep learning

\*Correspondence: yjqiao@263.net

<sup>2</sup> Research Center of TCM Information Engineering, Beijing University of Chinese Medicine, Yangguang South Avenue, Fangshan District, Beijing 102488, China

Full list of author information is available at the end of the article

classification methods owned better generalization ability and accuracy when predicting the actions of heat-clearing and blood-activating stasis-resolving based on TCM-HP theory. Besides, the methods of deep learning would help us to improve our understanding about the relationship between herbal property and action, as well as to enrich and develop the theory of TCM-HP scientifically.

**Keywords:** Heat-clearing herbs (HCHs), Blood-activating stasis-resolving herbs (BASRHs), Machine learning, Deep learning, Herbal property, Traditional Chinese medicine (TCM)

## Background

Traditional Chinese medicines (TCM) is one of the great herbal medicine systems worldwide, which plays an important role in current health care system in many countries. In the view of TCM theory, Yin-yang and five-elements theory is the central theory, which is used to explain how the world and body work [1]. The action of TCM is the high recapitulation for its therapeutic and health-preserving effect under the guidance of TCM theory [2]. TCM-HP, is the basic property of TCM and the high recapitulation of its functional characteristics [2]. The classic concept of TCM-HPs defines four fundamental characters (cold, cool, warm and hot), five fundamental tastes (salty, sour, bitter, sweet and pungent), four toxic states (toxic, nontoxic, very toxic, and slightly toxic), 12 meridians (bladder, spleen, large intestine, stomach, small intestine, liver, lung, heart, kidney, gallbladder, xin bao or pericardium and san jiao) [3]. TCM-HP also provides strong evidence to guide the clinical application of TCM. Many bioinformatics and pharmacological approaches were applied to study TCM-HP [1, 4, 5]. The research on the relationship between TCM-HP and action has been our great concern in the field of TCM.

The strategy of studying the relationship between TCM-HP and action have changed over the past decades. TCM-HP can not only be limited to a single property, but also need to be considered as a whole [6]. An A-priori algorithm was employed for producing association rules that described the intrinsic relationships between herbal property (qi, flavor and their combinations) and herbal efficacy [7]. However, the confirmed 120 resulting rules were dispersed and single property research could hardly characterize the whole effects of TCM. Hence the four fundamental characters, five fundamental tastes, and meridians need to be as a whole. Only in this way could we reveal the relationship between action and herbal property [8]. Multidimensional property of TCM is inherent basis of multiple action and is a collection of many herbal properties that determine the characteristics for efficiency of TCM [6]. Subsequently, property combination patterns for TCM [9–11] were proposed to reveal the relationship of TCM-HP and action from a holistic view of TCM.

Machine learning and deep learning methods have been widely applied in pharmaceutical research [12–17]. In life sciences, machine learning is often used to explain phenomena that are not completely theoretically understood [13]. Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction [18].

Research on TCM-HP and action is helpful to discover the inherent relation between TCM-HP and action of TCM [7], which can illustrate the connotation holistic view of TCM-HP theory.

TCM-HP and action of TCM are both summary of clinical practices, so the classification of actions based on the holistic concept of TCM-HP is consistent with clinical practices. Machine learning methods have been applied to TCM researches with some success, for instance, for the TCM syndrome classification [19, 20], the relationship of TCM-HP and action [1, 3, 12, 13, 21–26].

In this work, 88 HCHs and 45 BASRHs were collected as our research objects and four different artificial intelligence methods were used to develop the classification models of TCM actions based on TCM-HP theory. In order to ensure robustness, these four constructed classification models (kNN, SVM, CNN, DBN) were tested with external validation set, including 15 HCHs and 5 BASRHs.

Heat clearing herbs, which tend to have cold characters, have been found to produce some combination of antimicrobial [27], anti-toxic [28], anti-inflammatory [29, 30], antipyretic [28], antioxidant [27, 31], platelet aggregation inhibition, sedative, immunomodulatory [32], and hepatoprotective activities [33]. The HCHs own anti-inflammatory and antimicrobial effects, and their potential mechanisms of action contributing to their anti-inflammatory and antimicrobial activity may be related to their action of removing heat and counteracting toxicity [34]. *Radix Salviae Miltiorrhizae* (*Dan Shen*) is an example of BASRHs, and the compound Tanshinone IIA isolated from it could be a promising agent to improve blood viscosity and microcirculation and to prevent cardiovascular diseases [35]. Salvianolic acid B is clinically effective because of its ability to change the gene expression profile of endothelial cells thereby

preventing vascular events [36]. The Blood-Activating and Stasis-Resolving herb—Chuanxiong Hort may treat headache and has potential to be an agent for treating headache [37]. Because of the important clinical significance of these two kinds of typical TCM, they can lay the foundation for elucidating the relationship between TCM-HP and action.

**Methods**

**Selection of HCHs and BASRHs**

A total of 133 well established HCHs and BASRHs, given in Additional file 4: Table S1, were collected from reputable TCM literatures [38, 39], which were composed of 88 HCHs and 45 BASRHs. Their TCM-HPs and actions were collected from Chinese pharmacopoeia (2015 edition). External validation set (Additional file 4: Table S1) including 15 HCHs and 5 BASRHs were collected from a reputable TCM literature [39].

**Digital representation of TCM and feature selecting for TCM-HPs**

The classic concept of TCM herbal properties (TCM-HPs) defines four fundamental characters (cold, cool, neutral, warm and hot), five fundamental tastes (salty, sour, bitter, sweet and pungent), four toxic states (toxic, nontoxic, very toxic, and slightly toxic), and 12 meridians (bladder, spleen, large intestine, stomach, small intestine, liver, lung, heart, kidney, gallbladder, xin bao or pericardium and san jiao) [3]. With the continuous understanding of TCM-HP, the TCM-HPs of bland and astringent were added into five fundamental tastes in Chinese pharmacopoeia (2015 edition), and it would further enrich and develop TCM-HP theory. Four toxic states were considered the toxic effects of TCM, and the toxic action will cause injury of organs and tissues, functional impairment, pathological changes and even death [2]. So in this research, we only considered the relationship between

TCM actions and TCM-HPs (four fundamental character, five fundamental tastes, 12 meridians).

TCM-HPs are thus divided into three classes: character (C), taste (T), meridian (M) for this research and there are totally 24 TCM-HPs in the class of C, T, M, respectively. The 24 herbal properties can be further divided into 5, 7 and 12, which fall into C, T and M, respectively, as shown in Table 1. This study is based on the holism concept of TCM-HP to determine the classification for actions of TCM. Other 3 herbal properties (neutral, bland, astringent) were added based on classic concept of TCM-HPs. The strength of each TCM-HP is not considered, as this study is a qualitative one. The value of a specific TCM-HP is 1 if the herb possesses the corresponding property, and it is 0 if the herb does not possess the property [3].

For instance, the heat-clearing herb Huang Qin (*Scutellariae Radix*) has cold character, bitter taste; lung, gallbladder, spleen, large intestine and small intestine meridians entered. So the TCM-HP vector of Huang Qin (*Scutellariae Radix*) is  $V = [V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9, V_{10}, V_{11}, V_{12}, V_{13}, V_{14}, V_{15}, V_{16}, V_{17}, V_{18}, V_{19}, V_{20}, V_{21}, V_{22}, V_{23}, V_{24}] = [1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0]$  according to the order in Table 1.

**Traditional machine learning and deep learning methods**

Machine learning explores the study and construction of algorithms that can learn from and make predictions on data [40]. The field of machine learning, which aims to build model from an example training set of input observations, and then make data-driven predictions or decisions expressed as outputs. The methods of machine learning hold promise to enable computers to assist humans in the analysis of large, complex data sets [41], and they are not following strictly static program instructions. Machine learning methods have been applied to a broad range of areas within genetics and genomics [7], drug discovery [42–44], medicinal and biomedical properties identification [45, 46], tracking literature [47],

**Table 1 Chinese herbs properties' binarization table of some HCHs and BASRHs**

CHMs	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>	V <sub>6</sub>	V <sub>7</sub>	V <sub>8</sub>	V <sub>9</sub>	V <sub>10</sub>	V <sub>11</sub>	V <sub>12</sub>	V <sub>13</sub>	V <sub>14</sub>	V <sub>15</sub>	V <sub>16</sub>	V <sub>17</sub>	V <sub>18</sub>	V <sub>19</sub>	V <sub>20</sub>	V <sub>21</sub>	V <sub>22</sub>	V <sub>23</sub>	V <sub>24</sub>
Zu Ye	1	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	1	1	0	0	0
Qin Pi	1	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	1	0	0	1	0	0
Lian Qiao	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0
Zhi Zi	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1
Qing Hao	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0
Huang Qin	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	1	0	1	0	0
Dan Shen	1	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0

V<sub>1</sub>: cold, V<sub>2</sub>: cool, V<sub>3</sub>: neutral, V<sub>4</sub>: warm, V<sub>5</sub>: hot, V<sub>6</sub>: sour, V<sub>7</sub>: bitter, V<sub>8</sub>: sweet, V<sub>9</sub>: pungent, V<sub>10</sub>: bland, V<sub>11</sub>: astringent, V<sub>12</sub>: salty, V<sub>13</sub>: liver, V<sub>14</sub>: heart, V<sub>15</sub>: spleen, V<sub>16</sub>: lung, V<sub>17</sub>: kidney, V<sub>18</sub>: xin bao or pericardium, V<sub>19</sub>: gallbladder, V<sub>20</sub>: small intestine, V<sub>21</sub>: stomach, V<sub>22</sub>: large intestine, V<sub>23</sub>: bladder, V<sub>24</sub>: san jiao, respectively, each of which includes 5, 7 and 12 TCM-HPs. The total number of unique TCM-HP vector for all TCM is 5 + 7 + 12 = 24

cancer risk prediction and diagnosis [48], wind power prediction [49], etc.

However, the success of machine learning systems often requires a large amount of labeled data which is expensive to obtain and significant manual feature engineering. These feature representations are often hand-designed, require significant amounts of domain knowledge and human labor, and do not generalize well to new domains [50].

Deep learning (DL), a concept closely associated with artificial neural networks (ANNs), is in principle the learning of layered concepts. Thus, a model could describe higher and lower-level concepts at different layers of its structure [51]. Deep learning discovers intricate structure in large data sets by using the back propagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer [46]. Deep learning had been applied in adapting advanced neural network architectures for pharmaceutical research [14, 15, 52], predicting drug-induced liver injury (DILI) [17]. While deep learning and particularly unsupervised deep learning is still in its infancy, particularly in biological applications [53]. Moreover, deep learning network predicted drug property and activity with a relative accuracy improvement of approximately 14% over Merck’s in-house systems and resulted in an article in *The New York Times* [50, 54].

The research of TCM-HP will be accelerated by using the deep learning methods, which will promote the intelligent study of TCM-HP. With the deep neural network architectures, we can excavate the underlying regularities

and rules from the data recorded in ancient literature. As it is well-known that the data of traditional Chinese medicine is highly nonlinear, it is an inevitable trend to find out the inherent rules by using deep neural network (Fig. 1).

**Supported vector machine (SVM)**

The support vector machine (SVM) is based on the statistical learning theory of VC dimension (Vapnik–Chervonenkis Dimension) and the risk structure of the minimum principle [55–58] and is also one of the most popular and successful binary classification methods. Its basic idea is to find a hyperplane in the feature space which separates the training data perfectly into two classes [59]. Moreover, SVM is a classifier that performs classification tasks by constructing hyperplanes in a multidimensional space that separates cases of different class labels [60]. Least square support vector machine (LS-SVM) is an extension of standard support vector machines and it changes the error of optimization from first order into secondary order which covers a problem of solving quadratic programming into a problem of solving a set of linear equations [61, 62]. The method has been applied widely in biomedicine [63, 64].

**K-nearest neighbor (KNN)**

K-nearest neighbor algorithm is the most widely used classification and clustering algorithm. The k-nearest neighbors (kNN) algorithm is one of the simplest machine-learning methods to understand and explain, and the principle being that an instance is classified by a majority vote of its neighbors [65]. It provides a simple



**Fig. 1** Interpretation of the scientific connotation behind the theory of TCM by deep learning methods. After TCM-HPs being converted to digital representation, they were entered as input vectors into the multi-layer neural networks. The output layer is the action classification with multiple processing layers to learn representations of TCM-HPs. We can excavate the underlying regularities and rules between TCM-HPs and actions with the deep neural networks architectures

and intuitive rule for pattern discrimination, which has resulted in its extensive use in a variety of applications and gains a high classification rate [66, 67].

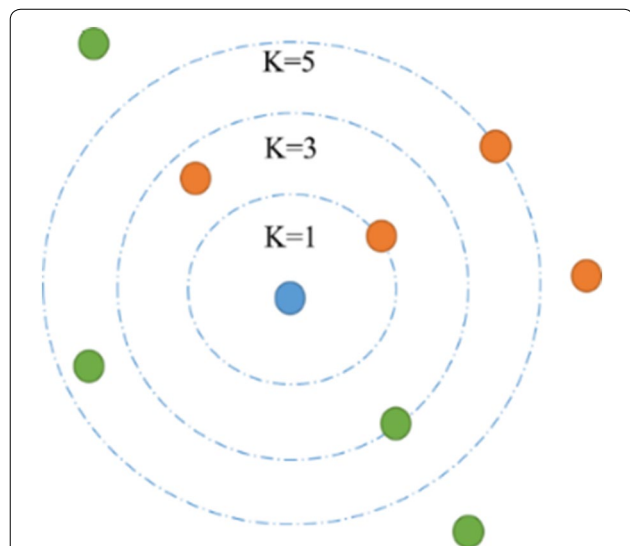
Each test sample is predicted to belong to the class most commonly found amongst its  $k$  closest neighbors, where  $k$  is a positive integer (Fig. 2). The Chinese herbal action classification is typically based on TCM-HP. TCM-HP is described as position vectors in the feature space that is usually of high dimensionality. Neighbors are identified on the basis of distance in the feature space. This is usually taken to be the Euclidean distance, though other metrics such as the Jaccard distance could be used. The minimum distance between the vectors gives the closest neighbor, so it is predicted that it belongs to the same class with the test object which the testing samples in the dataset are assigned to the class target value by a majority of its  $k$  nearest neighbor in the training set [68].

The Euclidean distance is often used to measure the similarity between two samples and more generally the distance between two  $p$ -dimensional vectors [69–71]. So the distance between these two TCM-HP vectors is computed as the length of the difference TCM-HP vector  $V_r - V_s$ , denoted by

$$d(V_r, V_s) = |V_r - V_s|$$

$$= \sqrt{(V_{r1} - V_{s1})^2 + (V_{r2} - V_{s2})^2 \dots + (V_{r24} - V_{s24})^2}$$

where  $V_r, V_s$  denotes the vectors of TCM  $r$  and  $s$ . The 24 TCM-HPs were considered as 24-dimensional vectors



**Fig. 2** Illustration of a kNN classification model. For  $k = 3$ , the blue one will be assigned to the red class, this time by a 2-1 vote; however, the blue one will be classified into the green class by a 3-2 majority. The 24 TCM-HPs were considered as 24-dimensional vectors and Euclidean distance were used to compute any two Chinese herbal vectors distance. Chinese herbal actions classification are typically based on TCM-HPs and we can classify the two kinds herbs based on 24-dimensional vectors with the kNN

and Euclidean distance were used to compute any two Chinese herbal vectors distance.

**Deep belief network (DBN)**

The deep belief network (DBN) is a neural network constructed from many layers of probabilistic model called restricted Boltzmann machines (RBMs) [72, 73]. The training process of DBN can be achieved layer-by-layer from low to high layer to train these multilayer RBMs. Each RBM layer is trained by using the previous layer’s hidden units ( $h$ ) as input/visible units ( $v$ ). Moreover, one RBM has a single layer of hidden units which are not connected to each other and have undirected, symmetrical connections to a layer of visible units. Contrastive divergence [74] based pre-training of these RBM layers is carried out to initialize the weights of DBN. Then, using the gibbs sampling method, the unbiased sample set could be got.

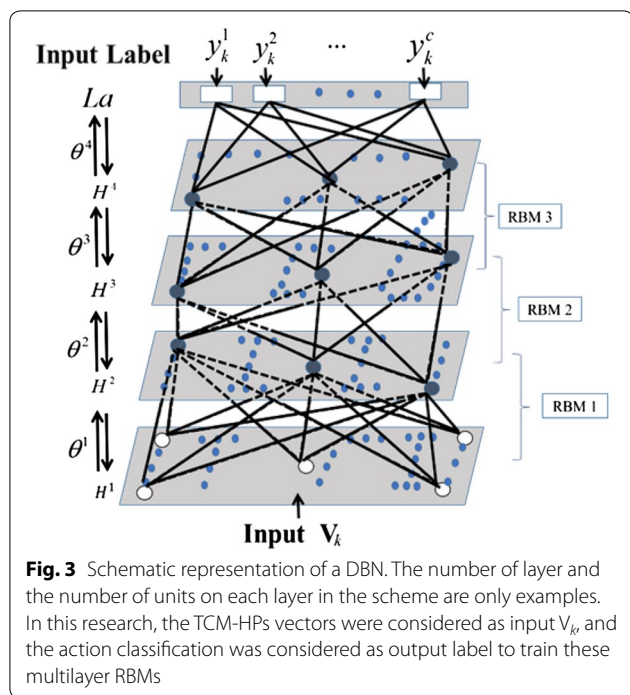
Hinton et al. [73] used complementary priors and derived a fast, greedy algorithm that could learn deep, directed belief networks one layer at a time, provided the top two layers form an undirected associative memory. The fast, greedy algorithm was used to initialize a slower learning procedure that fine-tunes the weights using a contrastive version of the wake-sleep algorithm.

In order to solve the problem of scaling full-sized, high-dimensional in images recognition, multiresolution deep belief networks [75] and convolutional deep belief networks [76] were constructed. In this research, the TCM-HP vectors were considered as input  $V_k$ , and the action classification was considered as output label as given in Fig. 3.

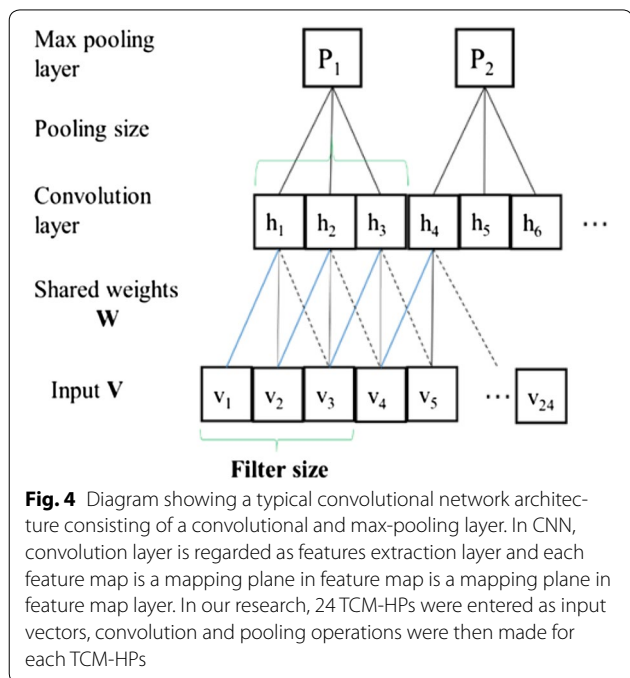
**Convolutional neural network (CNN)**

Convolutional networks combine three architectural ideas to ensure some degree of shift and distortion invariance local receptive fields, shared weights or weight replication and sometimes spatial or temporal subsampling. Convolutional process is the biologically inspired variant of multilayer perceptions (MLPs), which exploits the spatially local correlation by enforcing a local connectivity pattern [77, 78]. The classical convolutional network is composed of alternating layers of convolution and pooling (i.e. subsampling). The aim of the first convolutional layer is to extract patterns found within local regions of the input images that are common throughout the dataset [79].

In CNN, convolution layer is regarded as features extraction layer and each feature map is a mapping plane in feature map layer. The fully connected layers aggregate the local information learned in the convolutional layers to do class discrimination and fully-connected network like DNNs, each hidden activation  $h_i$  is computed



by multiplying the entire input  $V$  by weights  $W$  in that layer [80]. The weights  $W$  are then shared across the entire input space, as indicated in Fig. 4. In our research, 24 TCM-HPs were entered as input vectors, convolution and pooling operations were then made for each TCM-HP.



### Methods for evaluating prediction performance of deep learning and traditional machine learning methods

For a binary classification exercise, predictions can be classed as true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). Cross-validation is also a popular strategy, and still allows models to be tested on data unseen in their generation.

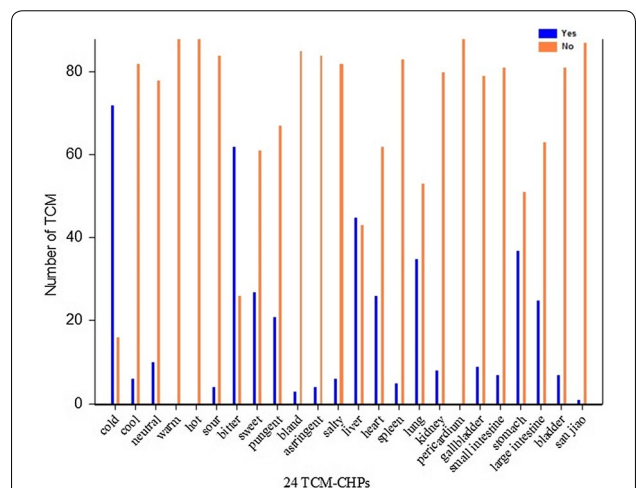
As in the case of all discriminative methods, the performance of deep learning and traditional machine learning methods can be evaluated by the quantity of true positive or TP (correctly classified HCHs), true negative or TN (correctly classified BASRHs), false positive or FP (BASRHs falsely classified as HCHs), and false negative or FN (HCHs falsely classified as BASRHs) respectively. Sensitivity (P+),  $SEN = TP / (TP + FN)$  and specificity (P-),  $SPE = TN / (TN + FP)$  are the prediction accuracy for HCHs and BASRHs, respectively. The overall prediction accuracy,  $ACC = [(TP + TN) / (TP + TN + FP + FN)]$ , and precision,  $PRE = TP / (TP + FP)$ . The overall prediction accuracy and precision are used to measure the overall prediction performance.

The minimum standards of reporting checklist contains details of the experimental design, and statistics, and resources used in this study (Additional file 1).

### Results

#### Distribution patterns of TCM-HPs of two kinds of herbs and their characteristics

According to holistic view of the TCM-HPs, the properties of 88 known HCHs are predominantly cold characters, bitter taste; liver and stomach meridians entered, respectively, which are given in Fig. 5. The properties of 45 known BASRHs are predominantly warm characters,



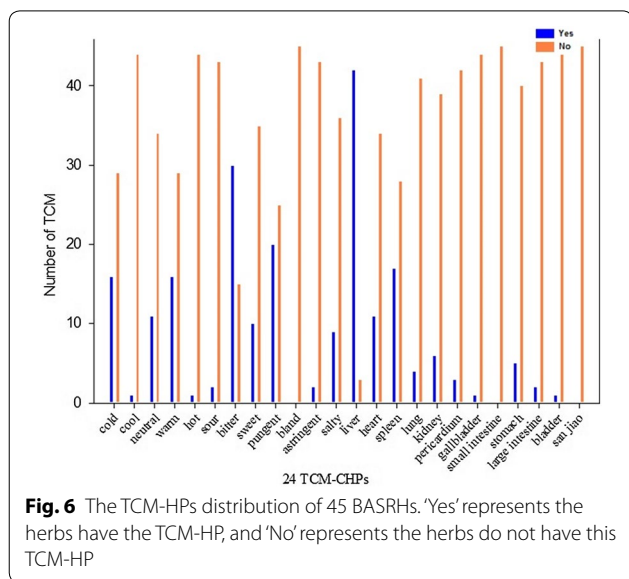
**Fig. 5** The TCM-HPs distribution of 88 HCHs. 'Yes' represents the herbs have the TCM-HP, and 'No' represents the herbs do not have this TCM-HP

bitter and pungent taste; liver meridian entered, respectively, which are given in Fig. 6.

Figures 5, 6 showed the common distribution patterns of two kinds of herbs were bitter taste; liver meridian entered.

The TCM-HP rates of HCHs and BASRHs were compared as given in Fig. 7. From the herbal properties rate distribution, we knew that significant TCM-HP of BASRHs are bitter, pungent; liver entered and their rates were 66.7, 44.4, 93.3%, respectively. The prominent TCM-HP features were cold (81.8%), bitter (70.5%); liver

(51.1%) and stomach (42.0%) entered in the 88 HCHs. Both of bitter and heart property rates in the two types of herbs were close proximity. However, the absolute value of differences for seven TCM-HP rates between HCHs and BASRH differed considerably as given in Table 2. If thirty percent of absolute value of difference was considered as setting value, the TCM-HP features were cold, warm character; spleen, liver and stomach meridians entered. Cold (81.8%)-bitter (70.5%)-liver (51.1%) combination could distinguish HCHs from BASRHs and warm (35.6%)-bitter (72.73%)/pungent (44.4%)-liver (93.3%) combination could distinguish BASRHs from HCHs.

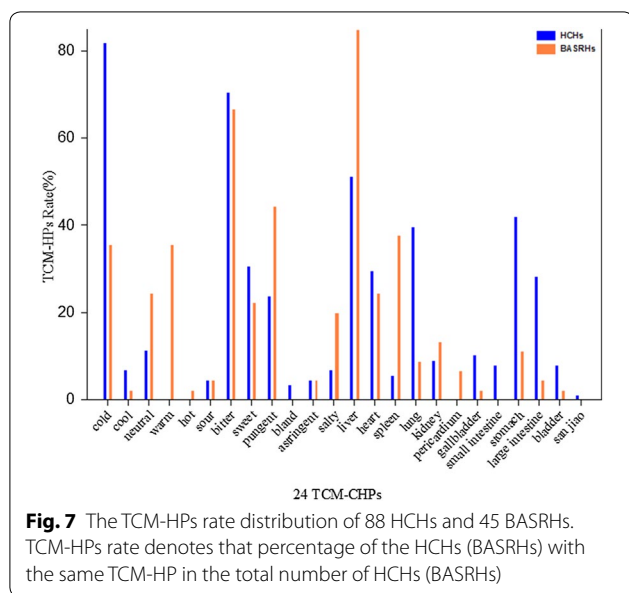


**Models analysis of TCM-HPs for distinguishing HCHs from BASRHs**

A cross-validation study was conducted to determine whether the traditional machine learning method SVM is able to separate HCHs and BASRHs based on their TCM-HPs. In this research, a SVM on the calibration set was constructed with setting coef 0 to 10, the highest number of polynomial kernel functions to 1. Polynomial was chosen as kernel function of this model and set tenfold cross validation when training this model. The sensitivity for the SVM model was 94.4%, and the specificity for this model was 72.4%, respectively. The overall prediction accuracy was 88.0%. The results on the validation set and external validation set were given in Table 3.

The best overall accuracy by kNN method was achieved 72.7% with k = 4 using the Euclidean distance on validation set and achieved 85.0% on external validation set.

A 100 hidden units RBM were constructed in this research, with setting batchsize to 5, learning rate to 2. After 10,000 iterations, the results showed that its sensitivity, specificity, precision, accuracy were 100.00, 100.0, 100.0, 100.0% on the validation set. On the external validation set, the same result were obtained with this DBN model.



**Table 2 Seven TCM-HP rates of HCHs and BASRHs and their absolute values of difference between HCHs and BASRHs**

TCM-HPs	HP rates of HCHs (%)	HP rates of BASRHs (%)	Absolute value of difference (%)
Cold	81.8	24.24	62.65
Warm	0.00	35.6	35.6
Pungent	23.9	44.4	20.5
Liver	51.1	93.3	42.2
Spleen	5.7	37.8	32.1
Stomach	42.0	11.1	30.9
Large intestine	28.4	4.4	24.0

**Table 3 Binary classification results with traditional machine learning and deep learning methods**

Data set	Type of models	Sensitivity (%)	Specificity (%)	Precision (%)	Accuracy (%)
Calibration set	SVM	94.4	72.4	89.3	88
	DBN	/	/	/	/
	kNN	91.5	75.9	90.3	87
	CNN	/	/	/	/
Validation set	LS-SVM	82.4	81.3	82.4	81.8
	DBN	100	100	100	100
	kNN	82.4	62.5	70.0	72.7
	CNN	100	100	100	100
External validation set	SVM	93.3	75.0	87.5	85.0
	DBN	100	100	100	100
	kNN	86.7	80.0	92.9	85.0
	CNN	100	100	100	100

In this research, we trained a 2 convolution layers and 2 subsampling layers convolutional neural network with setting batchsize to 5, learning rate to 0.001. In the first convolution layer, we set 3 output units and 2 kernels; and in the second convolution layer, we set 6 output units and 2 kernels. After 10,000 iterations, the model’s sensitivity, specificity, precision, accuracy were 100.00, 100.0, 100.0, 100.0% on the validation set and the external validation set.

The two deep learning method DBN, CNN achieved 100.0% accuracy on the validation set and external validation set than two other traditional machine learning methods was given in Table 3. It illustrated that deep learning methods had stronger predictive ability than two machine learning methods. These results also suggest that HCHs and BASRHs can be separated by deep learning and traditional machine learning methods based on their TCM-HPs with different accuracy and precision.

**Discussion**

The herbal properties distribution of HCHs and BASRHs showed their differences on TCM-HPs. The TCM-HPs of 88 known HCHs are predominantly cold, bitter; liver and stomach meridians entered. The TCM-HPs of 45 known

BASRHs are predominantly warm, bitter, pungent; liver meridian entered. The traditional machine learning methods and deep learning methods were adopted to construct the actions classification models based on the TCM-HP theory. The traditional machine learning methods SVM and kNN achieved 87.5, 92.9% overall prediction accuracy on external validation set. Furthermore, deep learning method DBN, CNN achieved 100.0% overall prediction accuracy. Two Chinese herbs (San Qi, Yin Xing Ye) were falsely classified as HCHs and Yu Gan Zi were falsely classified as BASRHs using SVM methods. In kNN model, two herbs (Niu Huang, Yin Xian Ye) were classified as HCHs and Yu Gan Zi (*Phyllanthus emblica* L.) was classified as BASRHs. Error classification CHMs with SVM and kNN on external validation set were given in Tables 4, 5.

The 4 falsely classified CHMs using the traditional machine learning method SVM and kNN have the same characteristic in common—sweet. Furthermore, they are mostly liver and heart meridians entered. These results suggested that deep learning method DBN and CNN are capable of dividing known HCHs from known BASRHs and the TCM-HPs of the known HCHs contain useful information for distinguishing them from BASRHs.

**Table 4 Error classification CHMs with SVM and kNN on external validation set**

Category of CHMs	Type of models	Accuracy (%)	Error classification CHMs
HCHs and BASRHs	SVM	85.0	Yu Gan Zi ( <i>Phyllanthus emblica</i> L.), San Qi ( <i>Notoginseng radlx Et Rhizoma</i> ), Yin Xian Ye ( <i>Ginkgo folium</i> )
	kNN	85.0	Niu Huang ( <i>Bovis calculus</i> ), Yu Gan Zi ( <i>Phyllanthus emblica</i> L.), Yin Xian Ye ( <i>Ginkgo folium</i> )



**Table 5 The herbal property of error classification CHMs**

Error classification CHMs	TCM-HPs
Yu Gan Zi ( <i>Phyllanthus emblica</i> L.)	Cool; sour, sweet, astringent; spleen, lung, stomach meridian entered
San Qi ( <i>Notoginseng Radix Et Rhizoma</i> )	Warm, bitter, sweet, liver and stomach meridian entered
Yin Xing Ye ( <i>Ginkgo folium</i> )	Neutral, bitter, sweet, astringent; heart, lung meridian entered
Niu Huang ( <i>Bovis calculus</i> )	Cool, sweet; heart, liver meridians entered

## Conclusions

With deep learning methods and machine methods, we could understand the nonlinear relationship between TCM-HPs and actions. Moreover, the deep learning classification models would had better accuracy and generalization ability than machine learning in predicting actions of TCMs based TCM-HP theory. The distribution patterns of TCM-HPs between HCHs and BASRHs were analysed. HCHs were mainly cold, bitter; liver and stomach meridians entered. BASRHs are predominantly warm, bitter, pungent; liver meridian entered, respectively. Traditional machine learning and deep learning methods classification studies showed that HCHs could be distinguished from BASRHs based on TCM-HP theory.

Future work for elucidating the regularity of TCM formula compatibility using the deep learning methods based on TCM-HP theory was on going. This work will contribute to more specific and deeper understanding of the traditional Chinese medicine system. The expectation is that we can utilize the deeper regularity to guide the discovery of TCM, drug design and clinical treatment in the future.

## Additional files

**Additional file 1.** Minimum standards of reporting checklist.

**Additional file 2.** Raw data in Tables 1, 2, 3, 4 and 5.

**Additional file 3.** The raw data in Figs. 5, 6 and 7.

**Additional file 4.** The TCM-HP for HCHs and BARSHs.

**Additional file 5.** The codes of kNN, SVM, DBN, CNN.

## Abbreviations

TCM: traditional Chinese medicine; HPs: herbal properties; TCM-HPs: traditional Chinese medicine herbal properties; SVM: supported vector machine; kNN: k-nearest neighbor; DBN: deep belief network; CNN: convolutional neural network; HCHs: heat-clearing herbs; BASRHs: blood-activating stasis-resolving herbs; DILI: drug-induced liver injury; DL: deep learning; ML: machine learning.

## Authors' contributions

ZC and YJQ substantial contributions to the conception or design of the work. YFC and SBH analysis and interpretation of data for the work. CZ and YFC writes all the codes. YJQ drafting the work and revising it critically for

important intellectual content. ZC and YJQ final approval of the version to be published. All authors read and approved the final manuscript.

## Author details

<sup>1</sup> School of Chines Materia Medica, Beijing University of Chinese Medicine, Yangguang South Avenue, Fangshan District, Beijing 102488, China. <sup>2</sup> Research Center of TCM Information Engineering, Beijing University of Chinese Medicine, Yangguang South Avenue, Fangshan District, Beijing 102488, China.

## Acknowledgements

This work is supported and sponsored by the Natural Science Foundation of China (No. 81430094) and 2013 Co-operation Programs of Beijing Municipal Education Commission: Research on data fusion supporting platform of BT/IT/CT for traditional Chinese medicine.

## Competing interests

The authors declare that they have no competing interests.

## Availability of data and materials

The readers can use data and materials in this manuscript by quotation of author names and Journal of Chinese Medicine. Raw data in Tables 1, 2, 3, 4 and 5 were summarized as Additional file 2. Availability of raw data for Figs. 5, 6 and 7 were summarized as Additional file 3. Availability of raw data for the TCM-HP for HCHs and BARSHs were summarized as Additional file 4. These four methods' codes were summarized as Additional file 5.

## Consent for publication

All of authors consent to publication of this study in *Journal of Chinese Medicine*.

## Ethics approval and consent to participate

Not applicable.

## Funding

This work was supported by 2013 Co-operation Programs of Beijing Municipal Education Commission: Research on data fusion supporting platform of BT/IT/CT for traditional Chinese medicine; National Natural Science Foundation of China (No. 81430094).

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 13 October 2017 Accepted: 14 February 2018

Published online: 27 February 2018

## References

- Liang F, Li L, Wang ML, Niu XY, Zhan JP, He XJ, Yu CY, Jiang M, Lu AP. Molecular network and chemical fragment-based characteristics of medicinal herbs with cold and hot properties from Chinese medicine. *J Ethnopharmacol.* 2013;148(3):770–9.
- Zhang T. Chinese materia medica. Beijing: High Education Press; 2008.
- Ung CY, Li H, Kong CY, Wang JF, Chen YZ. Usefulness of traditionally defined herbal properties for distinguishing prescriptions of traditional Chinese medicine from non-prescription recipes. *J Ethnopharmacol.* 2007;109(1):21–8.
- Zhang YL, Wang Y, Qiao YJ. Research on material base of TCM property based on pharmaco-phore. *World Sci Technol.* 2009;11(5):735–8.
- Zhou J, Xie G, Yan X. Encyclopedia of traditional Chinese medicines. *Isolat Compound AB.* 2011;1:455.
- Liang Q, Xie M. Identification of herbal properties from pharmaceutical multidimensional. *Shanghai J Trad Chin Med.* 2007;41:45–6 (In Chinese).
- Jin R, Lin Z, Xue C, Zhang B. An improved association-mining research for exploring TCM-HP theory: based on data of the Shennong's classic of materia medica. *Chin J Integr Med.* 2013;11:352–65.

8. Xiao B, Wang Y, Qiao YJ. Study on the Relationship between Chinese Herbal Nature and Function. *Chin J Inf Tradit Chin Med*. 2011;18(1):31–33 **(In Chinese)**.
9. Sun J, Zhang BX, Yan SR. Study on self-similarity of property combination mode of traditional Chinese medicines. *China J Chin Mater Med*. 2014;39(13):2378–81 **(In Chinese)**.
10. Hu Y, Sun J, Wang Y. Property combination patterns of traditional Chinese medicines. *J Tradit Chin Med Sci*. 2016;3(2):110–5.
11. Hu Y, Dong F, Wang Y, Qiao YJ. In silico method for studying property combination of traditional Chinese herbs. *J Tradit Chin Med Sci*. 2016;3(1):37–40.
12. Ung CY, Li H, Cao ZW. Are herb-pairs of traditional Chinese medicine distinguishable from others? Pattern analysis and artificial intelligence classification study of traditionally defined herbal properties. *J Ethnopharmacol*. 2007;111(2):371–7.
13. Balfer J. Development and interpretation of machine learning models for drug discovery [D]. Universitäts-und Landesbibliothek Bonn, Bonn; 2015.
14. 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 Pharm*. 2016;13(7):2524–30.
15. Gawehn E, Hiss JA, Schneider G. Deep learning in drug discovery. *Mol Inform*. 2016;35(1):3–14.
16. Unterthiner T, Mayr A, Klambauer G, Steijaert M, Wegner JK, Ceulemans H, Hochreiter S. Deep learning as an opportunity in virtual screening. In: Proceedings of the deep learning workshop at NIPS. 2014.
17. Xu Y, Dai Z, Chen F, Gao S, Pei J, Lai L. Deep learning for drug-induced liver injury. *J Chem Inf Model*. 2015;55(10):2085–93.
18. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436–44.
19. Li GZ, Sun S, You M, Wang YL, Liu GP. Inquiry diagnosis of coronary heart disease in Chinese medicine based on symptom-syndrome interactions. *Chin Med*. 2012;7(1):9.
20. Yang XB, Liang ZH, Zhang G, Luo YJ, Yin J. A classification algorithm for TCM syndromes based on P-SVM. In: IEEE Proceedings of 2005 international conference on machine learning and cybernetics. 2005; 6:3692–7.
21. Zhao YL, Wang JB, Xiao XH, Zhao HP, Zhou CP, Zhang XR, Jia L. Study on the cold and hot properties of medicinal herbs by thermotropism in mice behavior. *J Ethnopharmacol*. 2011;133(3):980–5.
22. Fusheng Z, Xiaoping L, Xushijie C, Jiannan C, Quanlin CHEN. Methodology for modeling on TCM-HP theory. *World Sci Technol*. 2009;11(2):229–33.
23. Shang E, Ye L, Fan X, Tang Y, Duan J. Discovery of association rules between TCM properties in drug pairs by association mining between datasets and probability tests. *World Sci Technol*. 2010;12(3):377–82.
24. Yao MC, Yuan YM, Ai L. Data mining and its application in the modernization of traditional Chinese medicine and traditional Chinese pharmacy. *J Beijing Univ Tradit Chin Med*. 2002;25(5):20–3 **(In Chinese)**.
25. Yao MC, Qiao YJ, Yuan YM, Ai L. Basis on artificial neural network method Chinese herbal function-efficacy classification research. *China J Chin Mater Medica*. 2003;28(7):689–91 **(In Chinese)**.
26. Yao MC, Zhang YL, Yuan YM. Study on the prediction of the effect attribution of the deficiency-nourishing drugs based on the quantification of TCM drug properties. *J Beijing Univ Tradit Chin Med*. 2004;27:7–18 **(In Chinese)**.
27. Sieniawska E, Baj T, Los R, Skalicka-Wozniak K, Malm A, Glowinski K. Phenolic acids content, antioxidant and antimicrobial activity of *Ligusticum mutellina* L. *Nat Prod Res*. 2013;27:1108–10.
28. China medical science and technology press. Pharmacopoeia of the People's Republic of China. 2015th ed. Beijing: China medical science and technology press; 2015.
29. Or TCT, Yang CLH, Law AHY, Li JCB, Lau ASY. Isolation and identification of anti-inflammatory constituents from *Ligusticum chuanxiong* and their underlying mechanisms of action on microglia. *Neuropharmacology*. 2011;60:823–31.
30. Huang J, Lu X, Zhang C, Lu J, Li G, Lin R, et al. Anti-inflammatory ligustilides from *Ligusticum chuanxiong* Hort. *Fitoterapia*. 2013;91:21–7.
31. Jeong JB, Ju SY, Park JH, Lee JR, Yun KW, Kwon ST, Jeong HJ. Antioxidant activity in essential oils of *Cnidium officinale* makino and *Ligusticum chuanxiong* hort and their inhibitory effects on DNA damage and apoptosis induced by ultraviolet B in mammalian cell. *Cancer Epidemiol*. 2009;33(1):41–6.
32. Bell CM, Simmonds MS, Appiah SS, Howes MJR. Plant substances as alternatives for animal products in traditional medicines. Report submitted to the Department for Environment Food and Rural Affairs. 2006.
33. Jiang WY. Therapeutic wisdom in traditional Chinese medicine: a perspective from modern science. *Trends Pharmacol Sci*. 2005;26(11):558–63.
34. Muluye RA, Bian Y, Alemu PN. Anti-inflammatory and antimicrobial effects of heat-clearing chinese herbs: a current review. *J Tradit Complement Med*. 2014;4(2):93–8.
35. Maione F, De Feo V, Caiazza E, De Martino L, Cicala C, Mascolo N. Tanshinone IIA, a major component of *Salvia miltiorrhiza* Bunge, inhibits platelet activation via Erk-2 signaling pathway. *J Ethnopharmacol*. 2014;155(2):1236–42.
36. Shi CS, Huang HC, Wu HL, Kuo CH, Chang BI, Shiao MS, Shi GY. Salvianolic acid B modulates hemostasis properties of human umbilical vein endothelial cells. *Thromb Res*. 2007;119(6):769–75.
37. Peng C, Xie X, Wang L, Guo L, Hu T. Pharmacodynamic action and mechanism of volatile oil from *Rhizoma Ligustici Chuanxiong* Hort. on treating headache. *Phytomedicine*. 2009;16(1):25–34.
38. Gao XM. Chinese materia medica. Beijing: China Press Tradit Chin Med; 2007.
39. Yan ZH. Chinese materia medica. Beijing: People's Medical Publishing House; 2014.
40. Kohavi R, Provost F. Glossary of terms. *Mach Learn*. 1998;30(2–3):271–4.
41. Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. *Nat Rev Genet*. 2015;16(6):321–2.
42. Giguère S, Lavolette F, Marchand M, Tremblay D, Moineau S, Liang X, Corbeil J. Machine learning assisted design of highly active peptides for drug discovery. *Plos Comput Biol*. 2015;11(4):e1004074.
43. Raccuglia P, Elbert KC, Adler PD, Falk C, Wenny MB, Mollo A, Norquist AJ. Machine-learning-assisted materials discovery using failed experiments. *Nature*. 2016;533(7601):73–6.
44. Zhang S. Application of machine learning in drug discovery and development. *Chemo-informatics and advanced machine learning perspectives: complex computational methods and collaborative techniques*. 2010. p. 235.
45. Jones DE, Ghandehari H, Facelli JC. A review of the applications of data mining and machine learning for the prediction of biomedical properties of nanoparticles. *Comput Methods Program Biomed*. 2016;132:93–103.
46. Nazarenko DV, Kharyuk PV, Oseledets IV, Rodin IA, Shpigun QA. Machine learning for LC–MS medicinal plants identification. *Chemomet Intell Lab Syst*. 2016;156:174–80.
47. Bornstein AT, McLoughlin MH, Aguilar J, Wong WS, Solomon BD. Tracking medical genetic literature through machine learning. *Mol Genet Metab*. 2016;118(4):255–8.
48. Asri H, Mousannif H, Moatassime HA, Noel T. Using machine learning algorithms for breast cancer risk prediction and diagnosis. *Procedia Comput Sci*. 2016;83:1064–9.
49. Heineremann J, Kramer O. Machine learning ensembles for wind power prediction. *Renew Energ*. 2016;89:671–9.
50. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2016;521(7553):436–44.
51. Mitchell JBO. Machine learning methods in chemoinformatics. *Wires Comput Mol Sci*. 2014;4(5):468–81.
52. Unterthiner T, Mayr A, Klambauer G, Steijaert M, Wegner JK, Ceulemans H, Hochreiter S. Deep learning as an opportunity in virtual screening. In: Proceedings of the deep learning workshop at NIPS. 2014. p. 27.
53. Mamoshina P, Vieira A, Putin E, Zhavoronkov A. Applications of deep learning in biomedicine. *Mol Pharm*. 2016;13(5):1445–54.
54. Markoff J. Scientists see promise in deep-learning programs. *New York Times*. 2012.
55. Bousquet O, Boucheron S, Lugosi G. Introduction to statistical learning theory. *Advanced lectures on machine learning*. Heidelberg: Springer; 2004. p. 169–207.
56. Pednault EPD. *Statistical learning theory*. New York: IBM Thomas J. Watson Research Division; 1997.
57. Vapnik VN, Vapnik V. *Statistical learning theory*. New York: Wiley; 1998.
58. Vapnik VN. An overview of statistical learning theory. *IEEE Trans Neural Netw*. 1999;10(5):988–99.

59. Utkin LV, Chekh AI, Zhuk YA. Binary classification SVM-based algorithms with interval-valued training data using triangular and Epanechnikov kernels. *Neur Netw*. 2016;80:53–66.
60. Santhanam T, Padmavathi MS. Application of K-means and genetic algorithms for dimension reduction by integrating SVM for diabetes diagnosis. *Procedia Comput Sci*. 2015;47:76–83.
61. Zhang N, Shetty D. An effective LS-SVM based approach for surface roughness prediction in machined surfaces. *Neurocomputing*. 2016;198:35–9.
62. Duolin W, Guihe Q, Jinnan D, Shuai Y. LS-SVM based human recognition on automotive active safety system. Heidelberg: In *Mechanical Engineering and Technology*. Springer; 2012. p. 75–80.
63. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. *Mach Learn*. 2002;46(1–3):389–422.
64. de Oliveira FSS, de Carvalho FAO, Silva AC, de Paiva AC, Gattass M. Classification of breast regions as mass and non-mass based on digital mammograms using taxonomic indexes and SVM. *Comput Biol Med*. 2015;57:42–53.
65. Ma J, Sheridan RP, Liaw A, Dahl GE, Svetnik V. Deep neural nets as a method for quantitative structure-activity relationships. *J Chem Inf Model*. 2015;55(2):263–74.
66. Remus JJ, Morton KD, Torriano PA, Tantum SL, Collins LM. Comparison of a distance-based likelihood ratio test and k-nearest neighbor classification methods. In: *IEEE workshop on machine learning for signal processing*. 2008. p. 362–7.
67. Fukunaga K, Hostetler LD. k-Nearest-neighbor Bayes-risk estimation. *IEEE Trans Inf Theory*. 1975;21(3):285–93.
68. Supardi NZ, Mashor MY, Harun NH, Bakri FA, Hassan R. Classification of blasts in acute leukemia blood samples using k-nearest neighbour. In: *IEEE international colloquium on signal processing and its applications*. 2012. p. 461–5.
69. Khoa NM, Viet DT, Hieu NH. Classification of power quality disturbances using wavelet transform and K-nearest neighbor classifier. In: *2013 IEEE international symposium on industrial electronics (ISIE)*. 2013. p. 1–4.
70. Li X, Xiang C. Correlation-based K-nearest neighbor algorithm. In: *IEEE international conference on software engineering and service science*. 2012. p. 185–7.
71. Peterson LE. K-nearest neighbor. *Scholarpedia*. 2009;4(2):1883.
72. Bengio Y, Lamblin P, Popovici D, Larochelle H. Greedy layer-wise training of deep networks. *Adv Neural Inf Process Syst*. 2007;19:153.
73. Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. *Neural Comput*. 2006;18(7):1527–54.
74. Hinton GE. Training products of experts by minimizing contrastive divergence. *Neural Comput*. 2002;14(8):1771–800.
75. Tao QQ, Zhan S, Li XH, Kurihara T. Robust face detection using local CNN and SVM based on kernel combination. *Neurocomputing*. 2016;211:98–105.
76. Lee H, Grosse R, Ranganath R, Ng AY. Unsupervised learning of hierarchical representations with convolutional deep belief networks. *Commun ACM*. 2011;54(10):95–103.
77. Zhang W, Qu C, Ma L, Guan J, Huang R. Learning structure of stereoscopic image for no-reference quality assessment with convolutional neural network. *Pattern Recogn*. 2016;59:176–87.
78. Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. *J Physiol*. 1968;195(1):215–43.
79. Zeiler MD, Fergus R. Stochastic pooling for regularization of deep convolutional neural networks. *arXiv preprint arXiv*. 2013. p. 1301.
80. Sainath TN, Mohamed AR, Kingsbury B, Ramabhadran B. Deep convolutional neural networks for LVCSR. In: *2013 IEEE international conference on acoustics, speech and signal processing*. 2013. p. 8614–18.

Submit your next manuscript to BioMed Central  
and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

